

SHIFTING SCALES AND PASSÉ PARADIGMS:
alternative spatio-temporal perspectives on
environmental heterogeneity and plant community
assembly

Andrew D. Letten

Centre for Ecosystem Science
School of Biological, Earth and Environmental Science
UNSW Australia

Thesis Submitted for the Degree of

Doctor of Philosophy

June 2017

PLEASE TYPE**THE UNIVERSITY OF NEW SOUTH WALES**
Thesis/Dissertation Sheet

Surname or Family name: Letten

First name: Andrew

Other name/s: David

Abbreviation for degree as given in the University calendar: PhD

School: Biological, Earth, and Environmental Sciences Faculty: Science

Title: Shifting scales and passe paradigms: alternative spatio-temporal perspectives on environmental heterogeneity and plant community assembly

Abstract 350 words maximum: (PLEASE TYPE)

Environmental heterogeneity has been a focal point of ecological research for decades and yet we have tended to stick to comparatively formulaic interpretations of its dimensionality. In this thesis I explore alternative perspectives on the relationship between heterogeneity and plant community assembly in both spatial and temporal dimensions. Beginning with perhaps the most basic property of an ecological community, its taxonomic diversity, in Chapter 2 I investigate the explanatory power of fine-scale spatial heterogeneity in temporal climate variability. Consistent with coexistence theory, the results of this study indicate that climate variability can be a better predictor of richness than more commonly used climatic averages, and highlight the need for ecologists to expand their purview beyond absolutes and averages. Turning to multivariate distance-based descriptors of community structure, in Chapters 3 and 4, I consider the under-explored role of temporal scale and variability on inference in community phylogenetic and trait-dispersion studies. Given a classic Brownian model of trait evolution, in Chapter 3, I show that the expected functional displacement of any two taxa is most parsimoniously represented as a linear function of time's square root. On this basis I argue that existing methods overweight deep time relative to recent time. Taking into account this methodological adjustment, in Chapter 4, I use standard phylogenetic and functional-trait metrics to evaluate the temporal stability of community structure through succession in a fire-prone heathland. Contrary to widely-held assumptions, community structure did not become increasingly functionally and phylogenetically dispersed with time. This contributes to an emerging body of evidence indicating that limits to the similarity of coexisting species are rarely observed at fine-scales. In Chapter 5, I adopt a model-based approach to the analysis of species co-occurrence patterns in response to fine-scale spatial environmental heterogeneity. This study confirms the vital role of hydrological niches for the maintenance of within-community plant diversity. Two primary conclusions emerge from the thematically broad investigations within the thesis: subtle shifts in the scales at which spatial and temporal heterogeneity are examined can yield new insights into community assembly; and reliable mapping of processes from patterns requires the continuous scrutiny of paradigmatic assumptions.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

  21/07/15
Signature Witness Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY

Date of completion of requirements for Award:

THIS SHEET IS TO BE GLUED TO THE INSIDE FRONT COVER OF THE THESIS

Copyright Statement

I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation. I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International. I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.

Signed 
Date 21/07/15

Authenticity Statement

I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.

Signed 
Date 21/07/15

Originality Statement

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed



Date 21/07/15

To Ash *et al.*

Acknowledgements

For what has been an incredibly rewarding and wonderfully challenging 45 months I have several people to thank, several times over. David, for being a superb supervisor, whose unstinting zen and benevolent wisdom are a template for academic mentorship. Richard and Will, who both played the role of unofficial supervisors and always left their doors ajar. Sharon, whose gravitational pull kept me (and everyone else) in orbit. Dan, Mick and John for providing valuable advice along the way. Steve and Bill who as panel members took a much appreciated interest in my progression. Cam who provided valuable insights into my original project proposal. Various other fantastic UNSW folk that have contributed to the fun, including in no particular order: Sam, Sylvia, Eve, Mitch, Nick, Evan, Jo, Chris, Ben, Anna, Chantel, Francis, David and the Ecostats crew, Angela, Haba, Rhiannon and the rest of the Big Ecology lab. The brilliant academics of the UCT Botany Dept, and in particular Jeremy and Tony who as honours supervisors made me think differently, in a really good way. My patient parents who have supported me unconditionally in my every pursuit. My generous brother who amongst many other things has done his best to keep me in fashion. My wonderful wife Shan, a mention in the acknowledgements of my thesis almost makes a mockery of your contribution; you rock like Kilimanjaro. Finally, I thank my amazing little Ash Mae, who makes me stop and smell the roses (and the nappies).

"Space and Time! Two minor omissions that no one is likely to notice"

Isaac Newton in *The System of the World* by Neal Stephenson

Abstract

Although environmental heterogeneity has been a focal point of ecological research for decades, we have tended to stick to comparatively formulaic interpretations of its dimensionality. In this thesis I explore alternative perspectives on the relationship between heterogeneity and plant community assembly in both spatial and temporal dimensions. Beginning with perhaps the most basic property of an ecological community, its taxonomic diversity, in Chapter 2 I investigate the explanatory power of fine-scale spatial heterogeneity in temporal climate variability. Consistent with coexistence theory, the results of this study indicate that climate variability can be a better predictor of richness than more commonly used climatic averages, and highlight the need for ecologists to expand their purview beyond absolutes and averages. Turning to multivariate distance-based descriptors of community structure, in Chapters 3 and 4, I consider the under-explored role of temporal scale and variability on inference in community phylogenetic and trait-dispersion studies. Given a classic Brownian model of trait evolution, in Chapter 3, I show that the expected functional displacement of any two taxa is most parsimoniously represented as a linear function of time's square root. On this basis I argue that existing methods overweight deep time relative to recent time. Taking into account this methodological adjustment, in Chapter 4, I use standard phylogenetic and functional-trait metrics to evaluate the temporal stability of community structure through succession in a fire-prone heathland. Contrary to widely-held assumptions, community structure did not become increasingly functionally and phylogenetically dispersed with time. This contributes to an emerging body of evidence indicating that limits to the similarity of coexisting species are rarely observed at fine-scales. In Chapter 5, I adopt a model-based approach to the analysis of species co-occurrence patterns in response to fine-scale spatial environmental heterogeneity. This study confirms the vital role of hydrological niches for the maintenance of within-community plant diversity. Two primary conclusions emerge from the thematically broad investigations within the thesis: subtle shifts in the scales at which spatial and

temporal heterogeneity are examined can yield new insights into community assembly; and reliable mapping of processes from patterns requires the continuous scrutiny of paradigmatic assumptions.

Table of contents

Acknowledgements	vii
Abstract	ix
1 Introduction	1
1.1 Pattern, process and heterogeneity	1
1.2 Diversity under spatio-temporal heterogeneity	3
1.3 Community composition, heterogeneity and temporal scale	4
1.4 Modelling fine-scale responses to heterogeneity	6
1.5 Concluding remarks	8
2 The importance of temporal climate variability for spatial patterns in plant diversity	9
2.1 Abstract	10
2.2 Introduction	11
2.3 Materials and Methods	13
2.3.1 Study area	13
2.3.2 Floristic data	14
2.3.3 Environmental data	15
2.3.4 Analysis	16
2.4 Results	19
2.5 Discussion	22
2.5.1 Conclusion	26
3 Trees, branches and (square) roots: why evolutionary relatedness is not linearly related to functional distance	27
3.1 Summary	28
3.2 Introduction	29
3.3 Simulations	34
3.4 Other models of trait evolution	35

3.5 Conclusion	36
4 Phylogenetic and functional dissimilarity does not increase during temporal heathland succession	39
4.1 Summary	40
4.2 Introduction	41
4.3 Materials and methods	44
4.3.1 Study site, sampling and fire history	44
4.3.2 Species pool and community phylogeny	45
4.3.3 Functional traits	47
4.3.4 Temporal change in phylogenetic and functional community structure	47
4.3.5 Temporal trends in community-weighted mean trait values	49
4.3.6 Temporal phylogenetic and functional beta turnover	49
4.4 Results	50
4.4.1 Temporal change in phylogenetic and functional community structure	50
4.4.2 Temporal trends in community-weighted mean trait values	53
4.4.3 Temporal phylogenetic and functional beta turnover	55
4.5 Discussion	56
5 Fine-scale hydrological niche differentiation through the lens of multi-species co-occurrence models	61
5.1 Abstract	62
5.2 Introduction	63
5.3 Materials and methods	66
5.3.1 Study area and floristic sampling	66
5.3.2 Environmental sampling	67
5.3.3 Data preparation	67
5.3.4 Model design	69
5.4 Results	74
5.5 Discussion	80
6 Conclusions	87
References	93
Appendix A Supporting information: Chapter 2	109
A.1 Supplementary figures and tables	110

Appendix B Supporting information: Chapter 3	123
B.1 R code to reproduce simulations	123
Appendix C Supporting information: Chapter 4	129
C.1 Phylogeny construction details	129
C.2 Equations for calculating D_{nn} and D_{pw}	131
C.3 Phylogenetic signal in traits	131
C.4 Temporal change in phylogenetic and functional community structure: supplementary figures	133
C.5 Temporal phylogenetic and functional beta turnover: supp. figs. .	137
C.6 Species accumulation curve	141
Appendix D Supporting information: Chapter 5	143
D.1 Supplementary figures and tables	143
D.2 LVM Primer	147
D.2.1 fitLVM-auxiliaryfunctions.R	158

Chapter 1

Introduction

1.1 Pattern, process and heterogeneity

Ecologists like to think about heterogeneity. Since beginning my doctoral studies in 2011, 350 articles containing the word ‘heterogeneity’ in the title have been published in the peer-reviewed ecological literature¹. At a glance, such an impressive number could be taken as evidence of an emerging ‘hot-topic’ or perhaps even a bandwagon (Fox, 2011), and yet the concept has pervaded ecological discourse in some shape or form for decades (e.g. Hutchinson, 1961, Wiens, 1977, Connell, 1978, Kotler & Brown, 1988, Li & Reynolds, 1995). Given such longevity, you may expect ecologists to have a pretty good handle on the topic. The reality is that we do and we don’t. Whilst theoreticians have made significant progress, particularly with respect to understanding the role of spatial and temporal environmental heterogeneity on species coexistence (Shmida & Ellner, 1984, Chesson, 1985, 2000b, Chesson & Huntly, 1997, Amarasekare, 2003), empiricists arguably have had a harder time of it. Historical contingencies (Belyea & Lancaster, 1999, Chase, 2003, Fukami *et al.*, 2005), system-specific properties (Lawton, 1999), and inconsistencies across different spatial and temporal scales (Levin, 1992, Swenson *et al.*, 2006) are just some of the factors that can obfuscate clear lines of inference and generalisation. It is the last of these ‘issues’ that has been my preoccupation for the last four years, and that provides the overarching, albeit heterogenous, focus of this thesis.

The concept of environmental heterogeneity is multi-dimensional in both scope and scale. Probably the first distinction that needs to be made is between spatial and temporal heterogeneity. Both are considered important for diversity and are

¹According to Web of Science.

to some degree analogous in their effect on species coexistence. Too little of it and one or two species out-compete the rest; too much of it, and no species are favoured over a sufficient spatial extent, or length of time, to maintain a stable population (Levine & Rees, 2004, Adler & Drake, 2008, Shurin *et al.*, 2010). Dig a little deeper into the heterogeneity hierarchy and things get a little murkier. The perceived amount of heterogeneity in any given community, landscape or region is entirely scale dependent. It will vary depending on the extent of the area under study, the length of time it is under observation and the spatial and temporal grain of the analysis. It gets particularly interesting when *spatio-temporal* processes are added to the mix. For instance, to what extent does temporal variability exhibit spatial heterogeneity, or conversely to what extent is spatial heterogeneity temporally stable? Measuring even a small subset of these diverse properties at community and/or ecosystem scales, and relating them to assembly processes, has long been both a logistical and analytical challenge. Fortunately, with the recent availability of inexpensive instrumentation, the development of sophisticated statistical models, and the accumulation of spatially broad, long-term datasets, ecologists are currently in a position to make significant progress in our understanding of heterogeneity across multiple scales and dimensions.

A common theme running through each chapter of this thesis is how we perceive heterogeneity/variability/dissimilarity, and how this bears on how we interpret community assembly and species diversity. With the exception of Chapter 4, which is taxonomically neutral, a further unifying thread is the usage of observational data derived from a range of terrestrial plant communities in southeast Australia. There has long been an antagonism in community ecology research between the mechanistic insight afforded by manipulative experimentation versus the potential for broader-scale inference derived from less resource-intensive observational studies. The case against the latter typically invokes the inherent uncertainty in inferring processes from observed patterns (Weiher & Keddy, 2001, Gotelli & McGill, 2006, Vellend *et al.*, 2014). In recent years, several authors have put forward cogent arguments outlining the potential for misattributing different ecological processes to observed patterns in community ecology, particularly in relation to the increasingly popular trait-based and phylogenetic approaches described in greater detail below (Cavender-Bares *et al.*, 2009, Mayfield & Levine, 2010, Fox, 2012, Adler *et al.*, 2013a, Kraft *et al.*, 2014). At the same time, there has been a string of exemplary manipulative experiments and long-term demographic studies that have helped bridge the gap between (coexistence) theory and data (Adler *et al.*, 2006, Angert *et al.*, 2009, Levine & HilleRisLambers, 2009, Adler *et al.*, 2010, Narwani *et al.*, 2013, Fritschie *et al.*, 2014, Godoy *et al.*, 2014, Alexandrou

et al., 2015, Kraft *et al.*, 2015). Together these studies have undoubtedly fostered a more nuanced understanding of community assembly and coexistence, and yet they almost exclusively derive from a rather narrow spectrum of ecological systems (e.g. annual plant and phytoplankton communities) which are particularly amenable to manipulative experiments. In contrast, in many perennial systems dominated by comparatively long-lived species, experimental approaches are less tenable and so we must often rely on observational data. It is also not clear to what degree the experimental findings from annual plant and phytoplankton communities scale-up to more complex systems comprising a much greater variety of life-history strategies. To this end, observational studies remain enormously valuable, but necessitate careful evaluation of the potential processes generating emergent patterns in community structure.

1.2 Diversity under spatio-temporal heterogeneity

Perhaps the most basic property of a community is its richness i.e. the number of species (or other taxonomic unit) observed per unit area. When measured in concert with other attributes of the environment (e.g. climate, disturbance, topographical heterogeneity etc.), it can provide insight into community structure and the processes driving it. If local diversity is high, we may infer some (multi-variate) property of the environment acts to stabilize the coexistence of a large number of species. In contrast, if diversity is low we may infer either that the environment falls beyond the fundamental niche of most species, or that it favours a minority of species to such a degree that they competitively exclude other would-be inhabitants. Given the comparative ease with which richness data can be obtained, a vast literature has built up exploring the role of heterogeneity as a driver of species richness across a variety of scales (reviewed in Stein *et al.*, 2014). In particular, there has been considerable interest in how spatial variation in climatic conditions, such as temperature and rainfall, regulates species richness (e.g. O'Brien, 1998, Francis & Currie, 2003, Currie *et al.*, 2004, Kozak & Wiens, 2012). While these studies have tended to focus primarily on mean climatic conditions, an emerging body of theory and empirical evidence suggests that local scale temporal climate variability may also play an important role in stabilizing species coexistence and thus fostering diversity. Nevertheless, few studies have compared the predictive capacity of temporal climate variability with respect to spatial patterns in species richness. To this end, in Chapter 2, I explore the relationship between temporal climate variability and plant

diversity using a large dataset (2,400 standardized floristic plots) from temperate forests in Southeast Australia, together with fine-resolution climate grids derived from data collected over two years by near surface climate loggers. The main objective of the study was to compare the relative strength of local-scale temporal temperature variability and absolute temperature, as predictors of regional scale spatial patterns in plant diversity. As part of the analysis, I also consider the shape of the relationship between temperature variability and species diversity in light of various predictions drawing on coexistence theory and population viability analysis.

1.3 Community composition, heterogeneity and temporal scale

Whilst species richness provides an effective first approximation of community structure, the distillation of community structure into a single metric comes at the cost of a tremendous amount of potentially useful information. For instance, we may also want to know what types of species are in the community; to what extent do they share similar life-history strategies; or what are their relative abundances and patterns of dominance/rarity. This type of multivariate information can potentially be used to make more direct inferences on the kinds of processes structuring the community. In particular, over the last 15-20 years there has been a rapid proliferation in trait-dispersion analyses, where community-level information on functional traits is used as a means of differentiating the relative contribution of biotic (e.g. competition) vs. abiotic processes (e.g. environmental filtering) to community structure (e.g. Weiher & Keddy, 1995, Weiher *et al.*, 1998, Cavender-Bares *et al.*, 2004, Cornwell & Ackerly, 2009, Kraft & Ackerly, 2010). This in turn has given rise to the closely related, and hugely popular, phylogenetic approaches to community structure analysis, where evolutionary relatedness is used as a proxy for the functional distance between taxa (Webb, 2000, Tofts & Silvertown, 2000, Slingsby & Verboom, 2006, Cavender-Bares *et al.*, 2006, Purschke *et al.*, 2013).

The classical paradigm, dating back to Darwin, holds that competition should inhibit species with high niche overlap from coexisting (limiting similarity *sensu* MacArthur & Levins, 1967), while environmental filtering has the opposite effect of limiting the range of successful ecological strategies at any one location (Weiher & Keddy, 1995, Diaz *et al.*, 1998, Stubbs & Wilson, 2004). If niche overlap can

be approximated by the distance between functional traits, which in turn can be approximated by phylogenetic relatedness, then we might infer that communities comprising of functionally similar or closely-related species are a product of strong environmental filters, whilst highly functionally and/or phylogenetically dispersed communities are indicative of strong competitive interactions (Webb, 2000, Webb *et al.*, 2002). This intuitively appealing line of reasoning has been adopted in numerous studies conducted in recent years, but as increasingly recognised, such straight-forward inference is not always justified.

Even before functional and phylogenetic based approaches gained momentum in the early 2000's, several authors had noted that the concept of limiting similarity was over-simplistic. Specifically, it ignores the concurrent role of competition in 'filtering' out all but the best competitors in the community, thus also limiting the *dissimilarity* of coexisting species (e.g. Abrams, 1986, Leibold, 1998). More recently, Mayfield & Levine (2010) provided a clear articulation of how this duality pertains to community phylogenetic approaches. Drawing on the coexistence framework laid out by Chesson (2000b), Mayfield & Levine (2010) argued that because the outcome of competition depends on both niche and fitness differences (with the former favouring coexistence and the latter driving exclusion), competition may favour the coexistence of both distantly and closely related species, contingent on the relative strength of stabilizing mechanisms (those that enhance niche differences) or equalizing mechanisms (those that reduce fitness differences). A further complication was recently proffered by Godoy *et al.* (2014) who showed that phylogenetic signal in niche and fitness difference may not always be correlated.

Notwithstanding these shortcomings, the widespread adoption of phylogenetic and functional based approaches has undoubtedly fostered a timely revival of research at the interface of ecology and evolutionary biology (Emerson & Gillespie, 2008, Cavender-Bares *et al.*, 2009, Vamosi *et al.*, 2009). Despite their intrinsic links, as researchers have become increasingly specialised in their work, ecology and evolutionary biology have to some extent grown artificially apart. Indeed, that it took so long for some of the weaker assumptions in community phylogenetics to be widely acknowledged likely points to both the paucity of deep phylogenetic understanding amongst contemporary ecologists, and the lack of exposure to modern coexistence theory amongst evolutionary biologists. As such, rather than abandon these approaches altogether, there is considerable merit in working towards the development of more robust forms of analysis and inference. This is particularly true given that functional traits and phylogenetic

information may remain for some time the only logically accessible sources of data on community assembly in complex and highly diverse systems, e.g. the tropical forests on which Webb (2000) originally pioneered the community phylogenetic approach.

In the interest of working towards a more nuanced understanding of the interaction between ecological and evolutionary processes, Chapters 3 & 4 of this thesis consider the comparatively weakly explored role of temporal scale and variability on inference in community phylogenetic and trait-dispersion studies. More specifically, Chapter 3 questions the adequacy of using divergence time as a linear approximation of functional distance between pairs of taxa, and via simulations of community assembly offers a simple solution to the overweighting of early-diverged clades. In contrast, Chapter 4 evaluates the stability of phylogenetic and functional community structure over ecological time-scales. To date, most studies of phylogenetic and functional community structure have been limited to a single snapshot in time, where by necessity observed patterns are assumed to be indicative of dominant assembly processes through time. While the robustness of this assumption seems plausible in long-lived late successional systems (e.g. tropical forests), questions remain as to how stable measures of community structure are through time in more frequently disturbed communities, and what this means for inference. To explore this problem, I investigated within- and between-community measures of phylogenetic and functional community structure in a fire-prone heathland along a 21-year time-series. The results of this study are interpreted in light of the conflicting predictions forwarded by ‘classical’ and more recent perspectives on non-random compositional patterns.

1.4 Modelling fine-scale responses to heterogeneity

While the analysis of community-level phylogenetic and functional trait data provides a deeper level of inference than is possible from investigations of species richness alone, there is still a significant element of potential information loss that comes with collapsing multivariate data into summary metrics (e.g. the distance-based approaches used in Chapters 3 & 4). Indeed, different species, or even different individuals within species, may be expected to respond in complex ways to the same environmental gradient. In recent years, the importance of variability in inter- and intra-specific responses at fine spatial and temporal scales has come under greater scrutiny (e.g. Sears & Chesson, 2007a, Fridley

et al., 2011, Dwyer *et al.*, 2015, Lai *et al.*, 2015). One of the current challenges is the development of analytical methods that are able to delineate individualistic responses at fine-scales.

An emerging array of model-based approaches to community level analysis has the potential to provide more transparent and statistically robust lines of inference for multivariate data-sets (Ovaskainen *et al.*, 2010, Ives & Helmus, 2011, Warton *et al.*, 2012, Clark *et al.*, 2014, Pollock *et al.*, 2014, Hui *et al.*, 2014, Harris, 2015). In particular, there has been several recent studies advocating the use of models for disentangling patterns of co-occurrence and their abiotic and/or biotic drivers (Ovaskainen *et al.*, 2010, Pollock *et al.*, 2014, Harris, 2015). What distinguishes these multivariate co-occurrence models apart from conventional null-randomisation based approaches is that they provide a direct means of assessing shared environmental responses separately from other processes that may generate non-random patterns of co-occurrence. Moreover, they facilitate more transparent and accurate representations of the statistical properties of the data (e.g. overdispersion of counts, Warton *et al.*, 2012), as well as provide a means of accounting for complexity and uncertainty in a hierarchical framework (Cressie *et al.*, 2009).

In Chapter 5, I adopt a novel model-based approach to co-occurrence analysis in order to evaluate the comparative importance of fine-scale hydrological niche segregation for plant species co-occurrence. Although theory suggests fine-scale heterogeneity can promote species coexistence (Chesson, 2000b, Amarasekare, 2003, Snyder & Chesson, 2004), owing to a dearth of studies at sufficiently high spatial resolution, the empirical evidence remains comparatively sparse. One fine-scale axis of differentiation, for which there is growing evidence of potentially wide taxonomic and geographic generality, is the partitioning of species along fine-scale hydrological gradients (Silvertown *et al.*, 2015). This evidence, however, stems largely from studies evaluating plant species responses to soil moisture independently of other environmental factors. In this final chapter, I assess the comparative importance of fine-scale hydrological niche differentiation for species co-occurrence using a high resolution study of soil hydrology and other edaphic variables, coupled with the same long-term study of heathland community dynamics analysed in Chapter 4.

1.5 Concluding remarks

In sum, the individual chapters of this thesis are thematically broad but all address in a different way the role of spatio-temporal scale in investigating ecological and environmental heterogeneity and their interaction. Chapter 2 explores the explanatory power of fine-scale spatial heterogeneity in temporal climate variability as a predictor of species richness; Chapter 3 offers a more parsimonious definition for the scaling of functional trait distance with divergence time in the context of contemporary ecological processes; Chapter 4 evaluates the temporal stability of phylogenetic and functional community structure; and Chapter 5 assesses the importance of fine scale spatial environmental heterogeneity in driving within-community patterns of co-occurrence. With the possible exception of Chapters 3 & 4, there is no strict linear format to the ordering of the thesis. However, by presenting the chapters in chronological order², the thesis follows a natural progression reflecting my own search for greater mechanistic insight into assembly processes. In the final chapter (Ch. 6), I briefly summarize the contributions of each of the preceding chapters, draw attention to their limitations and where relevant outline opportunities for future enquiry.

A note on formatting

Each data chapter is a standalone body of work that has already been published in the peer-reviewed literature (Chapters 2–5; Letten *et al.*, 2013, Letten & Cornwell, 2014, Letten *et al.*, 2014, 2015), with some small modifications specifically for this thesis. As such the formatting of each chapter is tailored to the journal in which it was published. The only departure from journal specific formatting is in the use of a consistent citation style, with references and appendices/supporting information provided at the end of the thesis, rather than at the end of each chapter. The contribution of each author is stated at the start of each chapter.

²Refers to the order in which the studies were conducted and written-up (Chs. 3 & 4 were written concurrently).

Chapter 2

The importance of temporal climate variability for spatial patterns in plant diversity

Andrew D. Letten, Michael B. Ashcroft, David A. Keith, John R. Gollan and Daniel Ramp

Ecography (2013), 36: 1341–1349
<http://doi.wiley.com/10.1111/j.1600-0587.2013.00346.x>



This study was conceived by ADL with input from MAB, DAK, JRL & DR. MAB and JRL collected climate data and modelled surfaces. ADL compiled floristic data, conducted analyses and wrote the manuscript, with contributions from MAB, DAK, JRL & DR.

2.1 Abstract

Spatial variation in absolute climatic conditions (means, maxima or minima) is widely acknowledged to play a fundamental role in controlling species diversity patterns. In contrast, while evidence is accumulating that variability around mean climatic conditions may also influence species coexistence and persistence, the importance of spatial variation in temporal climatic variability for species diversity is still largely unknown. We used a unique dataset capturing fine-scale spatial heterogeneity in temperature variability across 2,490 plots in Southeast Australia to examine the comparative strength of absolute temperature and temperature variability in explaining spatial variation in plant diversity. Across all plots combined and in three of five forest types, temperature variability emerged as the better predictor of diversity. In all but one forest type, diversity also exhibited either a significant unimodal or positive linear correlation with temperature variability. This relationship is consistent with theory that predicts diversity will initially increase along a climate variability gradient due to temporal niche partitioning, but at an intermediary point, may decline as the risk of stochastic extinction exceeds competitive stabilization. These findings provide critical empirical evidence of a linkage between spatial variation in temporal climate variability and plant species diversity, and in light of changing climate variability regimes, highlight the need for ecologists to expand their purview beyond absolutes and averages.

2.2 Introduction

Ecologists have long been interested in the role of environmental gradients in driving patterns of biodiversity (Fox *et al.*, 2011). The relationship between species richness and spatial variation in climatic conditions, such as temperature and rainfall, has in particular received considerable attention (e.g. O'Brien, 1998, Currie *et al.*, 2004, Kozak & Wiens, 2012). While the overarching emphasis has been on richness responses to spatial variation in mean (including maxima or minima) climate variables, recent evidence indicates that deterministic and stochastic elements of variability around mean climatic conditions may play an underappreciated role in species coexistence and persistence, and therefore in the maintenance of species diversity (Chesson, 2000b, Levine & Rees, 2004, Adler *et al.*, 2006). At the same time, the spectre of altered climatic variability regimes under future climates, including more frequent extreme weather events (Easterling *et al.*, 2000, IPPC, 2007, Hansen *et al.*, 2012), is providing an urgent mandate for ecologists to expand their purview beyond mean climate measures (Adler & Drake, 2008, Shurin *et al.*, 2010, White *et al.*, 2010).

At broad biogeographic scales, gradients in climatic means undoubtedly have a fundamental role in controlling richness gradients (Francis & Currie, 2003). But at the local scale, where ecological processes such as competition take increasing precedence, contemporary coexistence theory makes a salient argument for giving greater consideration to fluctuations around mean climate. Specifically, coexistence theory stipulates that under certain conditions variability can stabilize competition, and thus promote diversity, by increasing the number of temporal niches available within a fixed space (temporal niche partitioning) (Chesson & Hulley, 1997, Chesson, 2000b). Competitive stabilization occurs when competing species prosper under different conditions, and have the ability to 'bank' fitness gains made during good times via what are termed 'storage effects', such as long-lived adults or seed banks (Chesson, 2000b). At the same time, it is widely recognised by practitioners of population viability analysis that too much environmental variability can be detrimental to population persistence because it reduces long-term population growth rates and increases vulnerability to stochastic extinction (Alvarez, 2001, Boyce *et al.*, 2006). These differing perspectives make opposing predictions on the long-run trajectory of species diversity under climatic variability (Levine & Rees, 2004), but in some instances may be expected to generate a richness peak at intermediate levels of climate variability (Adler & Drake, 2008).

Unlike the rich body of literature providing evidence for broad scale richness-responses along mean climatic gradients (as reviewed in Pausas & Austin, 2001), evidence relating temporal climate fluctuations to species coexistence and diversity has only recently begun to accumulate (White *et al.*, 2010). The favoured approach has been to analyse historical population dynamics amongst coexisting plant species in temporally variable environments (Levine & Rees, 2004, Adler *et al.*, 2006, 2009, Angert *et al.*, 2009), but a small number of studies have employed alternative approaches including lab-based microcosm experiments (Descamps-Julien & Gonzalez, 2005, Jiang & Morin, 2007, Tuck & Romanuk, 2012) and comparative analyses of species richness across sites characterized by different levels of climate variability (Shurin *et al.*, 2010). Notably, in their study of lake zooplankton communities, Shurin *et al.* (2010) demonstrated that richness increased with increasing temporal variability in water temperature and was more strongly correlated with variability than mean temperature. To our knowledge theirs is the only study to date to relate spatial variation in within-community species diversity to temporal climate variability.

In a recent review, White *et al.* (2010) highlighted the need for more comparative spatial studies on diversity-variability relationships to complement the traditional emphasis on population dynamics. A novel approach to testing diversity-variability relationships that reduces the confounding effects of large-scale climate patterns is to compare diversity at sites that share the same regional climate but exhibit heterogeneous temporal climate variability at the landscape-scale. For instance, organisms occupying a sheltered gully will typically experience less climatic variability, over various time-scales, than those inhabiting a nearby ridgeline. Capturing this fine-scale heterogeneity in climate variability across landscapes was until recently constrained by the limited availability of accurate climate data at fine spatial scales. Indeed, the coarse-scale grids routinely utilized by ecologists for modelling species distributions are inappropriate for this purpose because they are typically derived from data collected by standardized weather stations (i.e. sparsely distributed Stevenson screens positioned at ~1.5–2 m on flat, cleared terrain) which are specifically designed to reduce the effects of landscape features, such as topographic shelter, that generate fine-scale heterogeneity (Ashcroft *et al.*, 2012). Fortunately, the availability of small, low-cost microclimatic sensors has made it more viable for researchers to produce fine-resolution climate surfaces that consider the effects of a much wider variety of climate forcing factors, including cold air drainage, topographic exposure and canopy cover (Ashcroft & Gollan, 2012, Ashcroft *et al.*, 2012). With these sensors

and methods it is possible to produce climate surfaces that are more tightly coupled to the conditions experienced by most organisms.

Here we provide a unique analysis of the influence of temporal climate variability on species diversity in the terrestrial plant realm. To explore the relationship between temporal climate variability and plant diversity we assembled a dataset comprising more than 2,400 standardized floristic plots from temperate forests in Southeast Australia, together with fine-resolution climate grids derived from data collected over two years by near surface climate loggers (see Ashcroft *et al.*, 2012). Our two main objectives were: (1) to compare the relative strength of temperature variability and absolute temperature, alongside other climate variables, as predictors of plant species diversity; and (2) to investigate the shape of the relationship between temperature variability and species diversity; i.e. does it exhibit consistency with either the predictions of coexistence theory (positive monotonic relationship), population viability theory (negative monotonic relationship) or a unified unimodal model (*sensu* Adler *et al.*, 2006).

2.3 Materials and Methods

2.3.1 Study area

The study area comprises approximately 60,000 km² in the Hunter Valley region of New South Wales, Australia (Fig. 2.1). Although parts of the region have been cleared or modified for agriculture and mining, there remain large patches of native vegetation within protected areas, including parts of two world heritage sites. These relatively undisturbed areas are topographically complex with an altitudinal range of more than 1,400m. The vegetation is dominated by dry and wet sclerophyll eucalypt forest, with smaller areas of heath, upland swamp, and temperate rainforest (Keith, 2004).

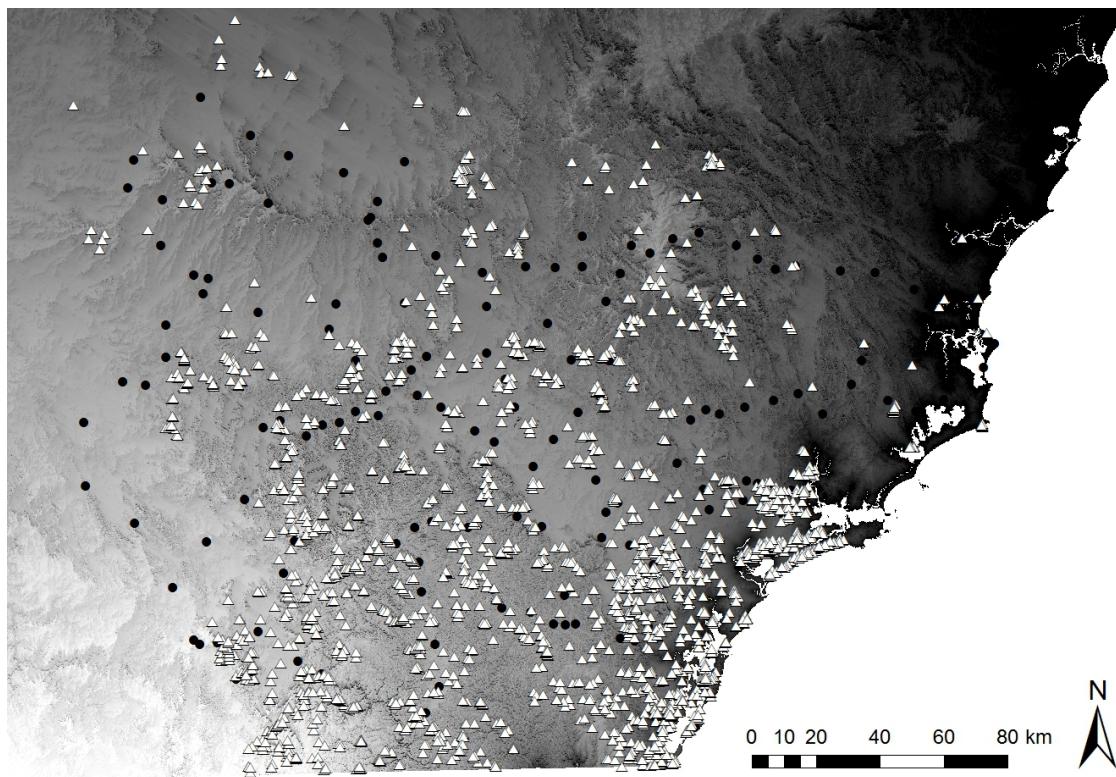


Fig. 2.1 Map of greater Hunter Valley study area ($31.4\text{--}33.4^{\circ}\text{S}$, $149.4\text{--}152.6^{\circ}\text{E}$) showing locations of survey plots (white triangles) and climate loggers (black circles) overlaid on topoclimatic grid ($25 \times 25 \text{ m}$) of near-surface average variability in maximum temperature. Light to dark shading indicates transition from variable to stable localities.

2.3.2 Floristic data

Floristic data was assembled from the New South Wales Office and Environment and Heritage's YETI 3.2 vegetation plot database (www.environment.nsw.gov.au/research/VISplot.htm). Before performing analyses, the metadata for all surveys conducted in the study area were intensively scrutinized to ensure included records were from plots of a standard size (0.04 ha) and sampling effort; comprised a complete list of vascular plants; were accurately geo-referenced; and coincided geographically with regions mapped as native vegetation types on digitized maps. After implementing the evaluation criteria, a total of 2490 standardized 0.04 ha full-floristic plot records conducted between 1998 and 2010 were included in the analysis.

Interspecific competition can occur between plants differing in size by several orders of magnitude. For instance, herbaceous species are known to limit the establishment of tree seedlings or even compete directly with adult trees (e.g. Riginos, 2009). All vascular plants were therefore included in the assessment of

species richness, which was quantified as the total number of species in each quadrat and ranged from 3 to 108 in any given plot, with a sum total of 2889 species representing 189 families. To evaluate whether observed response patterns varied across different growth-forms, supplementary analyses were performed on richness quantified for two separate groups: trees or arborescent shrubs; and all other growth-forms (hereafter trees and non-trees). Allotment of species into these groups was based on descriptions given in published floras (Harden, 1993), and may not reflect the actual growth form at a particular site. We performed all analyses on both the entire dataset and on subsets of the data grouped by forest type, which enabled us to explore whether the relationship between richness and climate variability varies between habitats, as has been shown for richness-disturbance relationships in wet vs. dry forests (Bongers *et al.*, 2009). The exclusion of all forest types (derived from digitized maps) with less than 50 plots allowed for the partitioning of the data into five distinct forest types: grassy woodlands (GW; $n = 225$), dry sclerophyll forest (DSF; $n = 1600$), wet sclerophyll forest (WSF; $n = 349$), rainforest (RF; $n = 101$), and forested wetlands (FW; $n = 215$).

2.3.3 Environmental data

For the purposes of the study we refer to absolute temperature as any variable that provides an indication of the extreme or mean temperature at a site, but does not provide any indication of variability around extreme/mean temperature. For instance, two sites might have a mean annual temperature of 20 °C, but while one might experience relatively constant temperatures throughout the year, the other might fluctuate seasonally from a low of 5 °C to a high of 35 °C. Although there are a number of approaches to quantifying absolute temperature, we specifically focused on extreme temperature, rather than mean temperature, as extreme temperature is more tightly coupled with fine-scale topographical heterogeneity (Suggitt *et al.*, 2011); is more likely to affect individual fitness than mean temperature (Stenseth *et al.*, 2002, Reyer *et al.*, 2012); and is predicted to fluctuate with greater frequency under climate change (Smith, 2011). As such, we defined absolute (extreme) temperature as the 95th percentile of daily maximum temperature and 5th percentile of daily minimum temperature over a one-year period.

For each quadrat location, temperature variables were extracted from a series of fine-resolution (25 m) topoclimatic grids interpolated from 113 climate loggers deployed within the study area for a total of 666 days from June 2009 to May 2011

(Fig. 2.1). To produce the grids of absolute temperature and temperature variability, the topoclimatic data was interpolated using a regional regression approach (Daly 2006), which involves fitting a multiple linear regression of temperature variables against climate-forcing factors including elevation, distance to coast, canopy cover, latitude, cold-air drainage, and topographic exposure (see Ashcroft & Gollan, 2012, for full details). Variability in maximum temperature was initially partitioned into three time-scales: (i) intra-seasonal variation in maximum temperatures, calculated as the 95th percentile of summer (December–February) maximums minus the 5th percentile of summer maximums; (ii) intra-annual variation in maximum temperatures, calculated as the 95th percentile of summer (December–February) maximum temperatures minus the 95th percentile of winter (June–August) maximum temperatures; and (iii) inter-annual variation in maximum temperatures, calculated as the difference in the 95th percentile of maximum temperatures between the two years. In order to obtain a measure of overall variability in maximum temperature, we then calculated the average of the three variability grids at each locality (Ashcroft & Gollan, 2012, Ashcroft *et al.*, 2012). To enable comparative analyses of temperature variability against other absolute climate variables, we also extracted the raw maximum and minimum temperature and humidity data (95th and 5th percentile of maximum and minimum temperatures) from the topoclimate surfaces, as well as five measures of rainfall derived from BioClim (Houlder *et al.*, 2003). Given that variability in minimum temperature was strongly correlated with absolute climate factors, we constrained our measure of variability to variability in maximum temperatures, while also controlling for potential direct effects of canopy cover on species richness by including remotely sensed canopy cover estimates (DECC, 2008) as covariates in the multiple predictor models. The complete list of variables is provided in (Table 2.1). All explanatory variables were Z-standardized.

2.3.4 Analysis

In order to provide a direct comparison between temperature variability and absolute temperature as predictors of plant species richness (Objective 1), we first fitted single-predictor generalized linear models (GLMs) for species richness as a quadratic and a linear function of each of average variability in maximum temperatures and the 95th percentile of maximum temperatures. This simultaneously enabled us to investigate the shape of the relationship between species

Variable	Abbreviation	Source
Average variability in 95th percentile of maximum temperature	VarMT	Ashcroft et al., 2012
95th percentile of maximum temperature	MaxT95	Ashcroft & Gollan, 2011
5th percentile of maximum temperature	MaxT5	Ashcroft & Gollan, 2011
95th percentile of minimum temperature	MinT95	Ashcroft & Gollan, 2011
5th percentile of minimum temperature	MinT5	Ashcroft & Gollan, 2011
95th percentile of maximum humidity	MaxH95	Ashcroft & Gollan, 2011
5th percentile of maximum humidity	MaxH5	Ashcroft & Gollan, 2011
95th percentile of minimum humidity	MinH95	Ashcroft & Gollan, 2011
5th percentile of minimum humidity	MinH5	Ashcroft & Gollan, 2011
Mean annual precipitation	AP	Bioclim
Precipitation of warmest quarter	PWaQ	Bioclim
Precipitation of coldest quarter	PCQ	Bioclim
Precipitation of driest quarter	PDQ	Bioclim
Precipitation of wettest quarter	PWeQ	Bioclim
Canopy cover	CC	DECC, 2008

Table 2.1 Factors available for selection as correlates of species richness in multi-variable models.

richness and climate variability in each forest type (Objective 2). The models were initially fitted with Poisson error-distributions to account for the strictly non-normal distribution of count data, but due to overdispersion (variance > mean) we subsequently corrected the standard errors using a quasi-Poisson GLM model. For all significant quadratic and linear terms, we calculated the quasi-AIC (QAIC) value, a modification of AIC based on quasi-likelihood appropriate for overdispersed response variables (Burnham & Anderson, 2002), and the percent deviance explained by the model as: $[1 - (\text{observed deviance} - \text{residual deviance})] \times 100$. If both linear and quadratic models showed statistical significance, we chose the model with the lowest QAIC value (using the dispersion parameter derived from the global quadratic model).

To determine how important temperature variability was, compared to other climate variables in predicting species richness patterns across the different forest types, we employed a multi-model information theoretic approach (Burnham & Anderson, 2002). The explanatory variables available for inclusion in the master model were first selected from the pool of climate variables (plus canopy cover) in Table 2.1. Prior to model selection, independent relationships (linear and quadratic) between species richness and each potential explanatory variable were evaluated for significance. To mitigate the problematic effects of collinearity in explanatory variables (Dormann *et al.*, 2013), we evaluated correlations between all significant variables. Wherever two variables were strongly correlated (Pearson's $|r| > 0.7$), we excluded the variable exhibiting the weakest (greater QAIC) independent relationship with species richness.

Model selection was performed using the dredge function in the R package 'MuMIN' (Barton, 2012), whereby GLMs with quasi-Poisson error distributions were run for all variable combinations in each forest type and were evaluated and ranked according to their QAIC. We intentionally employed such an indiscriminate method in order to make any inference on the relationship between temperature variability (1 factor), versus absolute climate variables (14 factors) (Table 2.1), and species richness as conservative as possible. That is to say our approach ensured we would only conclude there was a significant relationship between richness and temperature variability if no other factors could explain this trend.

To evaluate the predictive power of each explanatory variable, we first constructed a 95% confidence set of models for each forest type by summing the cumulative Akaike weights, w_i , of the highest ranked models until the sum exceeded 0.95. The relative importance of each variable was then calculated as the sum of the Akaike weights ' $w_+(j)$ ' for all the models in which the variable of interest occurred in the 95% confidence set. Model averaging was finally applied to determine model averaged parameter coefficients for the 95% confidence set of models in each forest type (Burnham & Anderson, 2002).

Given the spatially clustered nature of some of the quadrats, we tested for spatial autocorrelation using Moran's I tests. Moran's I tests revealed significant autocorrelation in the residuals of the global models for each habitat type. While there a number of methods for dealing with spatial auto-correlation in normally distributed data, the available methods for dealing with non-normal data are more limited, and either suffer from a lack of precision in their ability to accurately estimate model coefficients (Dormann *et al.*, 2007), or have only been tested on a limited number of datasets (Murphy *et al.*, 2010). However, because the coefficients obtained in our models are almost identical to those obtained assuming normally distributed data (Table S2.1), we were satisfied that running simultaneous autoregressive (SAR) models, an effective approach to removing spatial autocorrelation from residuals assuming normal error distributions (Kissling & Carl, 2008), would provide a valid test of the effect of spatial autocorrelation on the model results. SAR models were generated using the R package 'spdep' (Bivand 2012).

2.4 Results

The percentage variation in species richness explained by temperature variability in the single-predictor models followed an increasing trend from the driest to the wettest (based on annual precipitation) forest types (Table 2.2, Appendix A, Fig. S2.1). Variability in maximum temperature performed better (greater explanatory power) than absolute maximum temperature as a univariate predictor of species richness in all plots combined and in the three wetter forest types but performed poorer in the two driest forest types: grassy woodland and dry sclerophyll forest (Table 2.1). Species richness exhibited a significant ($P < 0.05$) unimodal relationship with temperature variability across all plots combined and across all individual forest types, with the exception of forested wetlands and grassy woodlands, which showed a significant positive linear relationship and no relationship respectively (Fig. 2.2). Relative performance of the two temperature metrics in predicting tree species richness were similar to those described for all vascular plants in each of the five different forest types, but in all plots combined absolute temperature performed better than temperature variability (12.8 vs. 5.1 percentage deviance explained) (Table S2.4). For all non-tree species, patterns were again similar to those observed for all vascular plants, with the exception of dry sclerophyll forest and rainforest, where there was very little difference in performance between absolute temperature and temperature variability (Table S2.5).

	Single-predictor models		Relative explanatory power (VarMT/MaxT95)	Multiple-predictor models 95% confidence set *
	VarMT	MaxT95		
ALL	4.85 (0.64)	1.00 6.56	4.9 0.1	16.19 - 16.49 11.11 - 12.03
DSF	1.94	3.32	0.6	17.35 - 17.61
WSF	10.32	3.63	2.8	13.50 - 16.82
RF	13.03	(4.65)	2.9	29.35 - 38.15
FW	19.28	10.90	1.8	26.27 - 31.10

* see Table S2.3 for models comprising the 95% confidence set.

Table 2.2 Percentage deviance explained by variability in maximum temperature (VarMT) and absolute maximum temperature (MaxT95) as independent predictors of species richness, and the range of percentage deviance explained in the 95% confidence set of multiple predictor models for each forest type and all plots combined (ALL = all plots combined, DSF = dry sclerophyll forest, WSF = wet sclerophyll forest, GW = grassy woodland, RF = rainforest and FW = forested wetlands). Non-significant results are shown in brackets. The ratio VarMT/MaxT95 provides a measure of the relative explanatory power of temperature variability and absolute temperature.

As is to be expected, a number of the climate variables available for model selection were highly correlated ($|r| > 0.7$), necessitating the exclusion of the weaker ($>\text{QAIC}$) of each collinear pair (Table S2.2). Variability in maximum temperature generally exhibited weak to moderate collinearity with the other variables, but in dry sclerophyll forest it was strongly negatively correlated with the 95th percentile of minimum temperature ($r = -0.704$), while in forested wetlands it was strongly negatively correlated with both precipitation of the driest quarter ($r = -0.746$) and precipitation of the coldest quarter ($r = -0.734$). In each instance, variability in maximum temperature exhibited the lowest QAIC and thus was retained.

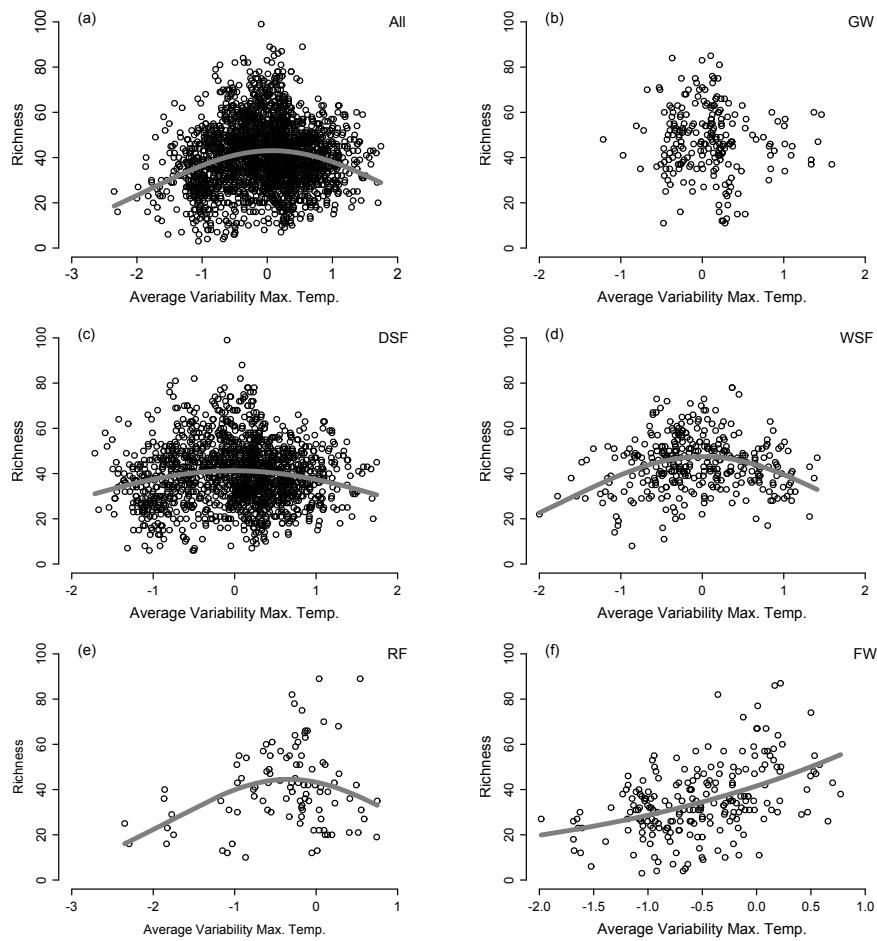


Fig. 2.2 Relationship between species richness and average variability in maximum temperature (VarMT) in: all plots combined (ALL, $n = 2490$) (a), grassy woodland (GW, $n = 255$) (b), dry sclerophyll forest (DSF, $n = 1600$) (c), wet sclerophyll forest (WSF, $n = 349$) (d), rainforest (RF, $n = 101$) (e), and forested wetlands (FW, $n = 215$) (f). Solid grey lines are the lines of best fit for significant single-predictor GLMs (linear model in FW exhibits slight upward curvature due to log-link function of Poisson regression). Percentage deviance explained by each model is given in Table 2.2.

ALL				GW				DSF						
Variable	Estimate	SE	P	w+(j)	Variable	Estimate	SE	P	w+(j)	Variable	Estimate	SE	P	w+(j)
VarMT	1.45	0.33	***		PWaQ	-0.79	0.22	***		VarMT	0.02	0.32		
VarMT²	-2.61	0.22	***	1.00	PWaQ ²	-1.61	0.20	***	1.00	VarMT²	-1.51	0.20	***	1.00
MinT ₅	-2.84	0.45	***		MinT ₅	0.45	0.18	*		MinT ₅	0.16	0.33		
MinT ₅ ²	2.21	0.18	***	1.00	MinT ₅ ²	-0.14	0.18		0.20	MinT ₅ ²	1.55	0.17	***	1.00
MinT ₉₅	3.61	0.51	***		MaxH ₅	-0.29	0.19			MaxH ₅	-0.40	0.30		
MinT ₉₅ ²	-0.01	0.37		1.00	MaxH ₅ ²	0.09	0.18		0.16	MaxH ₅ ²	-1.74	0.20	***	1.00
MaxH ₅	-3.47	0.47	***		MaxH ₉₅	-0.41	0.22	.		MaxH ₉₅	-0.72	0.29	*	
MaxH ₅ ²	-2.77	0.21	***	1.00	MaxH ₉₅ ²	-0.17	0.23		0.11	MaxH ₉₅ ²	-1.59	0.19	***	1.00
MaxH ₉₅	0.76	0.39	.							MinH ₅	1.88	0.44	***	
MaxH ₉₅ ²	-0.66	0.26	*	0.43						MinH ₅ ²	-1.19	0.22	***	1.00
MinH ₅	2.91	0.43	***							PCQ	-0.91	0.36	*	
MinH ₅ ²	-1.80	0.24	***	1.00						PCQ ²	-2.73	0.24	***	1.00
PCQ	-0.52	0.38								CC	1.02	0.23	***	0.83
PCQ ²	-4.00	0.23	***	1.00										
CC	1.07	0.23	***	0.86										
WSF				RF				FW						
Variable	Estimate	SE	P	w+(j)	Variable	Estimate	SE	P	w+(j)	Variable	Estimate	SE	P	w+(j)
VarMT	0.25	0.24			MaxH ₉₅	0.84	0.46	.		VarMT	2.01	0.31	***	1.00
VarMT²	-1.60	0.20	***	1.00	MaxH ₉₅ ²	-1.51	0.65	*	0.82	MaxH ₅	-1.98	0.27	***	
PDQ	-0.64	0.23	**		MinT ₉₅	1.64	0.39	***		MaxH ₅ ²	-1.44	0.25	***	1.00
PDQ ₂	-0.74	0.19	***	0.98	MinT ₉₅ ²	-1.48	0.36	***	0.76	MinT ₉₅	1.32	0.35	***	
MaxT ₅	0.14	0.28			VarMT	1.04	0.52	*		MinT ₉₅ ²	-0.28	0.19		0.87
MaxT ₅ ²	-0.67	0.20	**	0.56	VarMT²	-0.55	0.23	*	0.45	MaxH ₉₅	1.34	0.33	***	0.81
MaxT ₉₅	0.08	0.29			PDQ	-0.01	0.42			AP	0.21	0.38		0.28
MaxT ₉₅ ²	-0.56	0.21	**	0.40	PDQ ²	-1.01	0.31	**	0.45	MinH ₅	0.79	0.55		
MinH ₉₅	0.17	0.18			MinH ₅	0.70	0.41	.		MinH ₅ ²	-0.15	0.22		0.23
MinH ₉₅ ²	-0.40	0.18	*	0.31	MinH ₅ ²	-0.01	0.23		0.20	MaxT ₉₅	0.83	0.67		
MinT ₅	0.10	0.22								MaxT ₉₅ ²	0.14	0.21		0.22
MinT ₅ ²	-0.07	0.18		0.11										

Table 2.3 Coefficient estimates, standard errors and associated P-values (*P < 0.05, **P < 0.01, ***P < 0.001) of model averaged parameters in the 95% confidence set for each forest type. Summed Akaike weights [w₊(j)] provide a measure of the importance of each covariate. (ALL = all plots combined, DSF = dry sclerophyll forest, WSF = wet sclerophyll forest, GW = grassy woodland, RF = rainforest and FW = forested wetlands.

In the multiple-predictor models, variability in maximum temperature emerged as an important predictor of species richness relative to the absolute climate variables across all plots combined and in all forest types with the exception of grassy woodlands (Table 2.3, Table S2.3). In particular, variability in maximum temperatures was the single most important climate variable explaining species richness in wet sclerophyll forests, and the joint most important variable (together with the 5th percentile of maximum humidity) in forested wetlands. In rainforest, variability in maximum temperature was the joint third most important variable, while in dry sclerophyll forest and all plots combined, almost all the variables, including variability in maximum temperature, occurred in the 95% confidence set of best models.

In support of the single-predictor models, coefficient estimates derived from model averaging confirmed that species richness was consistently best modelled as a significant negative quadratic function of variability in maximum temperature in three of five forest types, as well as the combined data. Exceptions were forested wetlands, for which variability in maximum temperature was consistently positive and linear, and grassy woodlands where non-significant univariate models precluded the inclusion of variability in maximum temperature in the model selection process.

The global models for all forest types and all plots combined exhibited significant spatial correlation in the residuals. However, parameter estimates obtained with spatial regression models (SAR error models) were almost identical in sign and significance to those obtained with non-spatial GLMs (Table S2.1). In particular, the sign and significance of the linear and quadratic terms for variability in maximum temperature were consistent across almost all global models, regardless of whether the models were regressed using a non-spatial quasi-Poisson model, a non-spatial Gaussian model or a spatially sensitive SAR error model. The only exception was for dry sclerophyll forest where the quadratic term for variability in maximum temperature was no longer significant when modelled using SAR models. We therefore concluded that the results were largely insensitive to spatial autocorrelation in the data.

2.5 Discussion

The results of this study provide novel evidence that at local spatial scales temperature variability may be a better predictor of plant species diversity than absolute measures of temperature such as means, maxima, or minima. Specifically, our most important finding was that across all plots combined and in three of five individual forest types, the explanatory power of temperature variability was in the order of around two to five times that of absolute temperature (Table 2.2). This relationship supports the contention that spatial variation in temporal temperature variability plays a more important role in regulating species coexistence and persistence than comparative spatial variation in absolute temperature. In a previous study that compared diversity patterns across localities differing in their temporal climate variability profile, Shurin *et al.* (2010) similarly found that variation in freshwater zooplankton richness was better explained by temperature variability than mean temperature. To our knowledge these results provide the

first empirical evidence of a linkage between spatial variation in temporal climate variability and plant species diversity at local scales.

It is instructive that in both the single- and multiple-predictor models the performance of temperature variability as a predictor of plant diversity increased from the driest to the wettest forest types, which would suggest that temperature variability has a more crucial role in structuring plant communities where water is non-limiting. However, given that absolute temperature also performed poorly as a predictor of species richness in the drier forest types, we cannot rule out the possibility of higher level water-energy dynamics precluding richness responses to temperature in dry habitats (Francis & Currie, 2003). It is also likely that the greater frequency of fires in dry forest types (Clarke *et al.*, 2005) supersedes the influence of climatic factors, such as temperature, on species richness patterns. Indeed, the best multiple-predictor models for grassy woodlands, the driest vegetation type in our study, were poorer than all other forest types in explaining variation in species richness, suggesting that climate factors in general may have a less important role in structuring plant communities in drier habitats. Notably, Bongers *et al.* (2009) found that a disturbance index integrating primarily non-climatic variables such as logging and fire was better at explaining variation in tree species diversity in dry forest types than in wet forest types. This is consistent with our findings that climatic factors appear to be less important drivers of richness patterns in dry habitats than mesic habitats.

The statistically strong independent relationship between temperature variability and species richness for all forest types combined is counterbalanced by weak explanatory power (4.85%), albeit within the range typical of many ecological studies (Møller & Jennions, 2002) and almost five times that explained by absolute temperature. One plausible explanation is that factors other than temperature variability exert a stronger influence on species richness in our study area. A number of variables that were not included in the models may be expected to exert a substantial influence on fine-scale richness patterns, including soil and nutrient profiles, water availability, fire frequency, logging history and other forms of anthropogenic disturbance (Pausas & Austin, 2001). It is also possible that inconsistencies in the shape of the richness response across the individual forest types will serve to cancel each other out, precipitating a weak relationship across all plots combined. For example, percentage variability in species richness explained univariately by temperature variability was comparatively high in the wetter forest types (10.32% in wet sclerophyll forest, 13.03% in rainforest and 19.28% in forested wetlands), yet the shape differed between them (two

unimodal and one positive linear response). As such, these results suggest climate variability may play a more significant role in determining species richness at the within-community level rather than across communities.

One potential source of uncertainty in our analyses arises from the necessary interpolation of the climate data to obtain climate measures at each of the quadrat locations. However, because the identical interpolation method was used for both temperature variability and absolute measures of temperature, this should not have introduced any bias into the analyses. As such, while uncertainty derived from the interpolation may have weakened the overall explanatory power of all variables, the relative performance of temperature variability versus absolute temperature should have been unaffected. It is also notable that our estimate of temperature variability is likely subject to some measurement error on account of the climate model being derived from only two years of data, which may dilute its explanatory power. In particular the estimates of inter-annual variability would presumably benefit most significantly from a longer time-series as more climatically extreme years are sampled. Nevertheless, we are confident that the locations we identify as variable/stable are relatively static in their positioning along the variability gradients given their distinct topographic, geographic and environmental features (Ashcroft *et al.*, 2012). Furthermore, it is instructive that even with just two years of data we were able to show that temperature variability is, in many cases, a better predictor of plant diversity.

The separate univariate analyses conducted for each of the two sub-groups (trees and non-trees) provides some evidence that the relative performance of temperature variability and absolute temperature as predictors of diversity may vary depending on the life-history of the taxa in question. Although temperature variability exhibited better (or equal) performance relative to absolute temperature in the three wettest forest types for both trees and non-trees, the performance of absolute temperature exceeded that of temperature variability for trees when all plots were combined (Table S2.4). While it is tempting to speculate that variability at the temporal scales considered in this study may be less important for species with longer life histories, such an inference is weakened by the unavailability of data on the site-specific growth-habit of each species within each quadrat. Further work would benefit from targeted investigations of richness-variability relationships across different functional groups.

The finding that the relationship between plant diversity and variability in maximum temperature exhibited significant unimodality across multiple forest types (Fig. 2.2, Table 2.3) corroborates recent theory (Adler & Drake, 2008) reconciling

the opposing effects of climate variability on coexistence in ecological communities. Whilst strictly correlational, the inference is that along a gradient of temporal climate variability, richness will at first increase in response to temporal niche partitioning, but at some intermediary point, will decline as the risk of stochastic extinction exceeds competitive stabilization. Aside from grassy woodlands, where temperature variability was not significantly correlated with species richness, forested wetlands was the only forest type to exhibit a significant non-unimodal (positive linear response) relationship between temperature variability and species richness. It is noteworthy however, that maximum temperature variability tended to be lower in forested wetlands compared to the other forest types (Fig. S2.2). As such, it may be that temperature fluctuations rarely reach sufficient amplitude in the forested wetlands in our study area to elicit negative effects on species persistence.

The notion that temperature variability can be an important regulator of plant diversity may seem intuitive, but in the past biogeographers and ecologists have tended to give greater precedent to spatial heterogeneity in mean temperature (Currie *et al.*, 2004). Where temporal variability is acknowledged as a potential driver of richness patterns, conventional thinking holds that climate stability supports high diversity by facilitating adaptation to narrower niches, thus increasing the number of species per unit area (Stevens, 1989), or by fostering persistence and greater length of time for speciation (McGlone, 1996, Hopper, 2009). To this end, relative climate stability in the tropics compared to temperate regions has been invoked to explain the latitudinal gradients in species richness (Stevens, 1989; but see Gaston *et al.*, 1998). However, together with a number of previous studies, we provide further evidence that climate stability may, in some environments and over shorter time-scales, actually limit the maintenance of species diversity. This contention that the relative influence of climate variability on diversity may vary between ecological and evolutionary time scales, and across spatial scales, is particularly pertinent to the identification of refugia under future climates. Both short and long period environmental stochasticity is thought to be detrimental to the persistence of relict populations under climate change (Hampe & Jump, 2011), yet our results suggest sites that maintain high rates of occupancy under future climate regimes may not necessarily be those that are the most climatically stable at fine spatial scales.

2.5.1 Conclusion

In light of changes to climate variability regimes under climate change, empirical studies of this kind highlight the need for ecologists to broaden their focus beyond mean climate shifts. As evidenced in our system, it is likely that the comparative influence of variability and absolute climate on community structuring is conditional on multiple limiting factors. Our ability to predict where, when and at what spatio-temporal scales climatic variability or absolutes take precedence will benefit from further empirical research in different taxa and communities. To complement the array of experimental and population demographic approaches, we hope that the observational evidence presented here will provide motivation for others to investigate the strength and shape of richness-variability relationships across space in other systems.

Acknowledgements

This research was part of Australian Research Council Linkage Project LP100200080 in collaboration with the NSW Office of Water, Australian Museum, Central West Catchment Management Authority, the Australian Wetlands, Rivers and Landscapes Centre at the University of New South Wales and the University of Technology Sydney. We are grateful to Renee Woodward for assistance with acquisition of the YETI vegetation dataset, Cam Webb for valuable feedback on the original research proposal, and Habacuc Flores-Moreno for valuable comments on an earlier draft of the manuscript.

Chapter 3

Trees, branches and (square) roots: why evolutionary relatedness is not linearly related to functional distance

Andrew D. Letten and William K. Cornwell

Methods in Ecology and Evolution (2014)

<http://doi.wiley.com/10.1111/2041-210X.12237>



This study was conceived by ADL and WCK. ADL and WCK contributed equally to scripting code for simulations and writing the manuscript.

3.1 Summary

1. An increasingly popular practice in community ecology is to use the evolutionary distance amongst interacting species as a proxy for their overall functional similarity.
2. At the core of this approach is the implicit, yet poorly recognized, assumption that trait dissimilarity increases linearly with divergence time, i.e. all evolutionary time is considered equal. However, given a classic Brownian model of trait evolution, we show that the expected functional displacement of any two taxa is more appropriately represented as a linear function of time's square root.
3. In light of this mismatch between theory and methodology, we argue that current methods at the interface of ecology and evolutionary biology often greatly overweight deep time relative to recent time.
4. An easy solution to this weighting problem is a square-root transformation of the phylogenetic distance matrix. Using simulated models of trait evolution and community assembly, we show that this transformation yields considerably higher statistical power, with improvements in 92% of trials. This methodological update is likely to improve our understanding of the connection between evolutionary relatedness and contemporary ecological processes.

3.2 Introduction

With increasingly precise estimates of the most common ancestor amongst interacting species, modern phylogenetics offers the promise of a synthesis of contemporary ecology and evolutionary history (Webb *et al.*, 2002, Johnson & Stinchcombe, 2007, Cavender-Bares *et al.*, 2009, Cadotte *et al.*, 2013, Swenson, 2013). Following on this, the last 10–15 years have seen a precipitous rise in the number of studies examining ecological patterns and processes through the lens of evolutionary relatedness. Whilst this integrative approach has undoubtedly yielded new insights, much of the foregoing research has proved inconclusive; contemporary ecological interactions often appear, using conventional methods, to be unrelated to evolutionary history (Cahill *et al.*, 2008, Bennett *et al.*, 2013, Narwani *et al.*, 2013, Fritschie *et al.*, 2014). In this comment we offer one explanation for the poor performance of the phylogenetic metrics used in contemporary ecology, as well as a partial solution.

From the beginning of the phylogeny–ecology synthesis, evolutionary relatedness has often been used a proxy for the traits mediating species’ interactions with each other and their environment. It is impossible to measure all the relevant traits for complex ecological interactions, but because evolution is a conservative branching process and traits are on average more conserved than random, phylogenies have the potential to provide an integrative measure across all traits (Webb, 2000). It follows that the strength of this inference is contingent on an accurate model of how phylogenetic and functional distance covary. There is a widespread assumption in the literature (see Figure 1b of Cadotte *et al.*, 2013) that phylogenetic and functional distance scale linearly. This assumption is implicit in many of the conventional metrics for assessing phylogenetic community structure and phylogenetic diversity (Vellend *et al.*, 2010). Below we consider this assumption critically, using current theory on trait evolution.

The classic, “default” model for trait evolution is Brownian motion (Figure 3.1b and Felsenstein, 1985). The diffusion equation for Brownian motion (Einstein, 1905) has the form:

$$\frac{\partial \rho}{\partial t} = \sigma^2 \frac{\partial^2 \rho}{\partial x^2} \quad (3.1)$$

where σ^2 is the diffusion constant, t is time, ρ is density and x is position in space. That equation has the solution:

$$\rho(x, t) = \frac{\rho_0}{\sqrt{4\pi\sigma^2 t}} e^{-\frac{x^2}{4\sigma^2 t}} \quad (3.2)$$

which implies that the second moment of the distribution is:

$$\overline{x^2} = 2\sigma^2 t. \quad (3.3)$$

In other words, the variance goes up linearly with time and the standard deviation rises with time's square root, or as Einstein put it: "the mean displacement is therefore proportional to the square root of the time". Applied in a phylogenetic context this means that while among species variance in trait values goes up linearly with time (Felsenstein, 1985), the expected displacement of any two taxa in trait space does not increase linearly with time, but rather with time's square root (Figure 3.1). This non-linearity is true both for one trait and for Euclidean distance in n -dimensional trait space. Indeed, there is no plausible model for expected trait dissimilarity to be linearly related to evolutionary time. For there to be a linear relationship, after a speciation event, the functional distance between two lineages would increase constantly and continuously as their traits evolve away from each other. It is difficult to imagine a scenario where that would be the case.

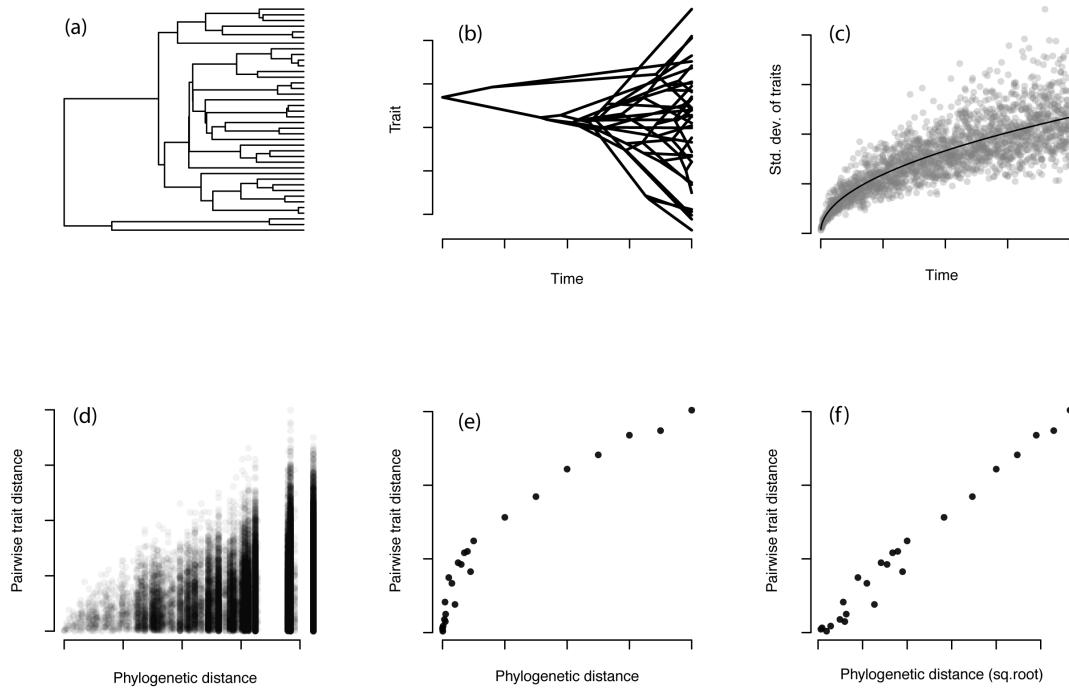


Fig. 3.1 Panel (a) shows a Yule phylogeny. Panel (b) shows a “traitgram” with one Brownian motion simulation with the trait value on the y-axis. Panel (c) shows the effect of time on the standard deviation of trait values at the tips for 2000 simulations with the same Brownian motion rate parameter; each point represents the standard deviation of the trait values of extant species within a separate simulation. Panel (d) shows the pairwise trait differences for 50000 simulations plotted against phylogenetic distance. Panel (e) simply takes the data from panel (d) and places them in bins, to show the statistical expectation at a given relatedness. Panel (f) shows the effect of taking the square root of the phylogenetic distance matrix on the relationship between the phylogenetic distance and the expected trait difference. Note that the expectation for the trait standard deviation and the pairwise difference is linear with respect to the square root of time as shown analytically in the text. All simulations use code from FitzJohn (2012) and Revell (2012).

While this technical point is well understood in parts of the comparative methods literature (Hardy & Pavoine, 2012), the implications for many ecological applications has gone largely unnoticed. If evolutionary relatedness is used as a proxy for functional distance, the non-linearity of their scaling relationship means that more recent evolution should have a disproportional influence on contemporary ecology. Current methods used in community phylogenetics (see review of methods in Vellend *et al.*, 2010) typically treat evolutionary relatedness linearly; 1 and 6 million year relatedness difference is treated as having the same expected effect as a 101 and 106 million year relatedness difference. All evolutionary time is considered equal, whether that time occurred over the last 5 million years or more than 100 million years ago. Combined with imbalanced trees, this creates a problem that is well known in empirical investigations: the sta-

tistical over-weighting of early-diverged, low diversity clades (Kembel & Hubbell, 2006). When those clades are included in the sample (or in a randomization), they have a disproportionate weight on the test statistic, a weight that is highly disproportionate to the expected trait difference under a Brownian model.

The root of this problem is in the distribution of pairwise phylogenetic distances. In the basic simulated birth-death trees, there is an equal probability at every time step for every branch of either a speciation or an extinction event. This creates what are known as balanced trees (Heard, 1997, also Figure 3.2). Balanced trees are rare in empirical studies, as heterogeneity in net diversification (Alfaro *et al.*, 2009) creates trees that are imbalanced (Mooers, 1995) and have peculiar distributions of pairwise distances (e.g. the vascular plant tree from Zanne *et al.* (2014) in Figure 3.3).

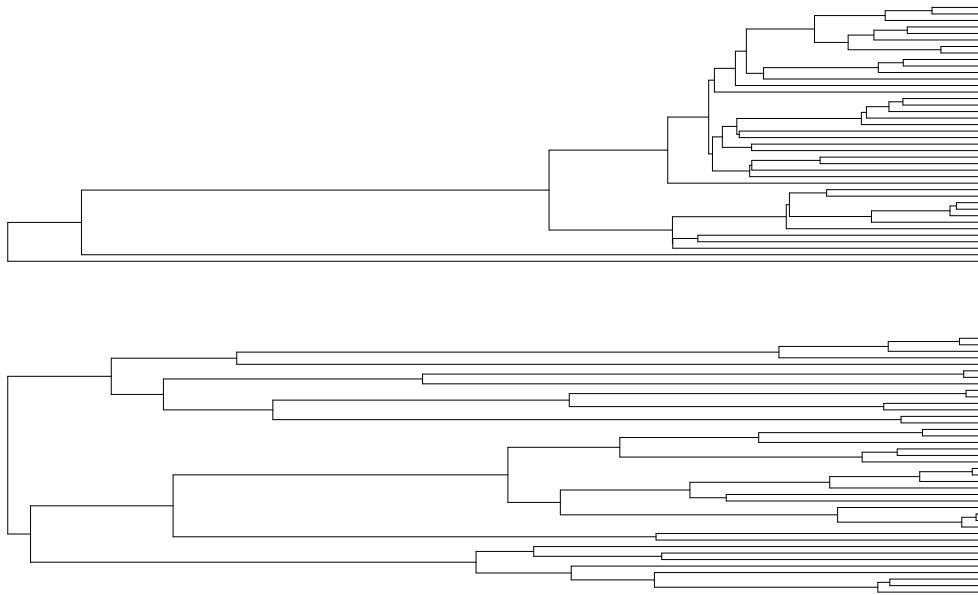


Fig. 3.2 Example of a real tree (top), obtained by randomly selecting taxa from the Zanne *et al.* (2014) tree, compared with a homogeneous birth-death simulation tree (bottom).

Drawing from theory on Brownian motion reveals a simple solution to the weighting problem: a square-root transform of the phylogenetic distance matrix. For example, the mean pairwise distance (sensu Webb, 2000):

$$MPD = \frac{2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n d_{i,j}}{(n)(n - 1)} \quad (3.4)$$

can be redefined as the mean of the square-root transformed pairwise distances:

$$MPD^* = \frac{2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sqrt{d_{i,j}}}{(n)(n-1)} \quad (3.5)$$

where n is the number of taxa and $d_{i,j}$ is the pairwise phylogenetic distance between species i and species j . This quantity MPD^* is proportional to the mean of the expected pairwise differences for traits evolving under a Brownian model, for both one trait in one dimension and for m traits in m dimensional space. This equation is provided as an example: very similar adjustments are possible to most of the common community phylogenetics statistics (see definitions within Vellend *et al.*, 2010), simply via a square-root transform of the distance matrix.

With this transformation, long-ago time is down weighted compared to recent time. As such, the effect of the presence or absence of a species on MPD is weighted in proportion to the mean pairwise expected trait difference to all other species under a Brownian model. In practical terms this re-scaling can be accomplished exceedingly easily with one additional line of code in combination with the tools in widely available statistical packages (Kembel *et al.*, 2010).

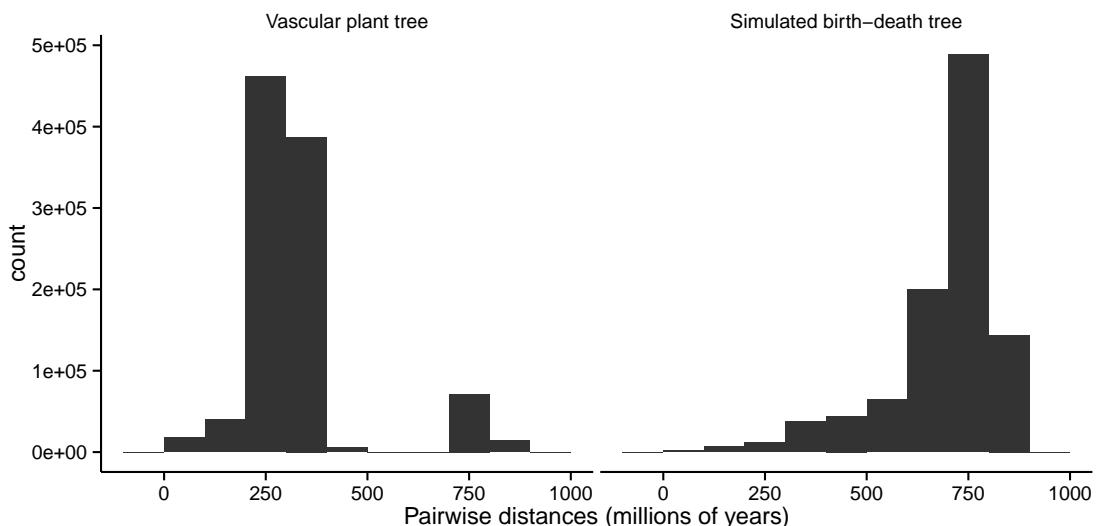


Fig. 3.3 The pairwise phylogenetic distances from a recent phylogeny of vascular plants by Zanne *et al.* (2014) and a simulated phylogeny of the same age and number of extant species.

3.3 Simulations

To explore the empirical implications of this idea, we conducted simulations applying a similar framework to that described by Kraft *et al.* (2007, in this case repeating the ‘filtering-derived’ and ‘neutral assembly’ algorithms). We simulated trait evolution by Brownian motion on both a real tree and a homogenous birth-death simulated tree (Figure 3.2). To keep the pool size (number of tree tips) consistent across the real and simulated trees, the real tree was randomly pruned down to 200 taxa. In each run, we ‘evolved’ a trait across the phylogeny and then applied one of two community assembly filters to obtain a final community of 40 taxa. Under the ‘filtering derived’ assembly filter, the most derived (extreme) trait-value was treated as the optimum, with the remaining 39 places in the community selected from taxa having the nearest trait values to that optimum. This process simulated community assembly via habitat filtering (Diaz *et al.*, 1998), whereby the abiotic environment sets some threshold on the range of strategies (and thereby trait values) that are able to sustain a positive population growth rate (e.g. tolerance of inundation along a hydrological gradient). In contrast with the deterministic nature of the ‘filtering-derived’ algorithm, under the ‘neutral assembly’ algorithm the community was obtained by randomly selecting 40 species independent of their trait values. One-thousand runs were conducted for each community assembly algorithm on each of the real and simulated phylogenies. Finally, we quantified the effect of the filter using conventional community phylogenetics methods (MPD and MNTD - mean nearest taxon distance), and compared the standard approach with that of a square-root transform of the phylogenetic distance matrix (all code to perform replicate simulations is provided in Appendix B).

Simulation results indicate the transformed test has considerably higher statistical power (Figure 3.4) for detecting the signal of community assembly. Using the square-root transform improved 92% of trials for both MPD and MNTD. Using the standardized effect size metric developed by Webb (2000), whereby:

$$SES_{METRIC} = \frac{METRIC_{observed} - mean(METRIC_{null})}{sd(METRIC_{null})}$$

the median improvement in standardized effect size was 0.64 for MPD and 0.43 for MNTD (simulated and real trees combined). This is a comparatively large increase in effect size from a simple statistical adjustment. The improvement was

similar for real and simulated trees, but may be more crucial in the real case because the general power of community phylogenetics is lower for the real tree (Kembel & Hubbell, 2006). While a comprehensive exploration of the effect of the transformation on type 1 error rates is beyond the scope of this paper (see Kraft *et al.*, 2007), simulations indicated that the expected reduction in type 2 error rate is in the order of 5–25% (depending on the community to pool size ratio and other factors).

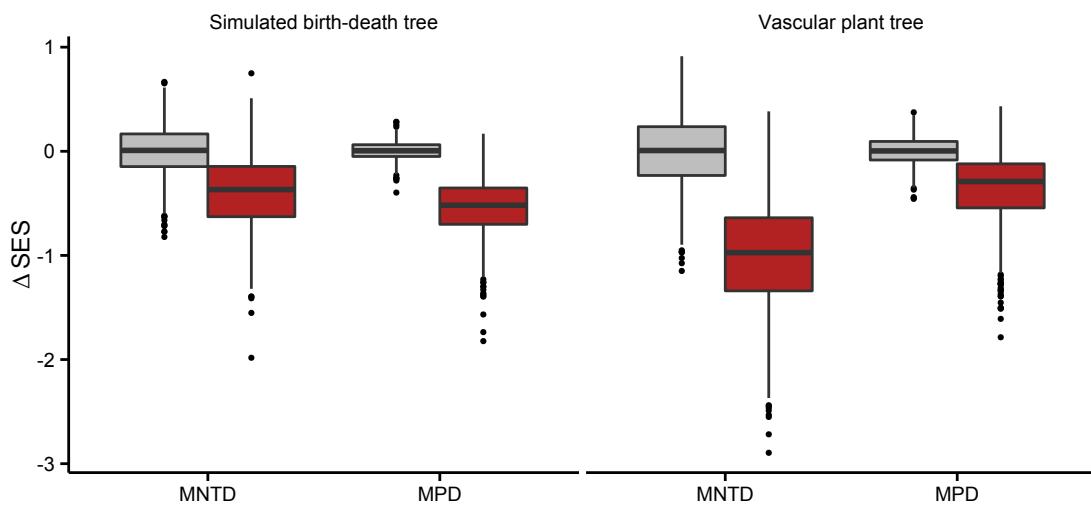


Fig. 3.4 Improved statistical power from down-weighting long-ago evolution: change in standardized effect size for mean pairwise distance (MPD) and mean nearest taxa distance (MNTD) using a square-root transformed phylogenetic distance matrix versus the conventional approach. Box plots with grey-fill group communities assembled under a neutral (random sample) model; box plots with red-fill group communities assembled under a 'habitat filter' model.

3.4 Other models of trait evolution

While a highly useful “default” model, we do not expect that a Brownian motion model will prove to be a fully adequate model for trait evolution at large scales. In general, the current evidence suggests that actually the square-root transform does not go far enough toward down-weighting long-ago evolution compared to recent evolution in many cases (Butler & King, 2004, Harmon *et al.*, 2010, Smith *et al.*, 2010). In the event that there are bounding or mean-reverting processes (e.g. Ornstein–Uhlenbeck (OU) Butler & King, 2004), phylogenetic signal will be less strong than under a Brownian model. Under these alternative models the effect of evolutionary relatedness decays more rapidly, a phenomena defined as “phylogenetic half-life” by Hansen *et al.* (2008). If trait evolution typically includes

this type of process, the problem we describe here will be even more extreme. In this case (see Kelly *et al.*, 2014), the square-root transformation will not go far enough. For the alternatives to Brownian motion, such as OU and heterogeneous models where rates of evolution vary among clades (Beaulieu *et al.*, 2012), there are analogous tree scaling approaches (Pearse *et al.*, 2013), but these, unlike the square-root transformation, require *a priori* information about trait evolution in the relevant clade.

3.5 Conclusion

We recommend a square-root transform of the phylogenetic distance matrix for all uses where phylogenetic relatedness is used as a proxy for current-day functional disparity. There are some cases where the number of years of evolutionary history in a place may be an interesting quantity in and of itself (Purvis & Hector, 2000). In those cases linear relatedness may still be of interest; however, many ecological studies use evolutionary relatedness as a proxy for trait dissimilarity and in these cases using relatedness linearly will decrease the power of the investigation.

While we have made a statistical argument above, in conclusion we stress that this is actually a conceptual point. We argue that conventional approaches over-weight long-ago evolutionary time and under-weight recent evolution both conceptually and statistically, and in doing so inadvertently limit the statistical power and success of efforts to leverage phylogenetic information in ecological contexts. There are many other reasons why the mapping of ecological process to phylogenetic community patterns may be inconclusive (Mayfield & Levine, 2010, Godoy *et al.*, 2014); many of these issues are hard to address. Here, we have identified one problem—the weighting of evolutionary history—where a simple adjustment may help.

Of course, by using the square-root transformation we make the assumption that trait differences scale linearly with ecological fitness. While disentangling the ecological/evolutionary processes is somewhat intractable in this instance, this is at least a parsimonious assumption. Instead, justification for using more complex measures of functional distance (e.g. squared distance) in the context of ecological selection/assembly should be contingent on supporting theory. To our knowledge none exists. The magnitude of fitness-trait relationships has received some attention (Kimball *et al.*, 2011, Adler *et al.*, 2013b) but explicitly exploring their scaling properties could well prove an invaluable area of future research.

In general, there needs to be a more nuanced use of evolutionary relatedness within community ecology. By improving the connection between metrics within community phylogenetics and trait evolution, we can increase the power and utility of using evolutionary relatedness to ask ecological questions. This methodological update will not be a one-off: as our understanding of the processes and patterns in trait macro-evolution at large scales grows (O'Meara, 2012, Pennell & Harmon, 2013), the phylogenetic metrics used in contemporary ecology will need to be continually updated.

Acknowledgements

Thanks to Nathan Kraft, Rich FitzJohn, Rob Kooyman, Sandrine Pavoine and an anonymous reviewer for insightful comments.

Data accessibility

R code to reproduce simulations provided in Appendix B.

Chapter 4

Phylogenetic and functional dissimilarity does not increase during temporal heathland succession

Andrew D. Letten, David A. Keith and Mark G. Tozer

Proceedings of the Royal Society B (2014), 281, 20142102

<http://dx.doi.org/10.1098/rspb.2014.2102>



This study was conceived by ADL with input from DAK. DAK and MGT provided floristic plot data. ADL compiled plant trait data, assembled phylogeny, conducted analyses and wrote the manuscript, with contributions from DAK & MGT.

4.1 Summary

Succession has been a focal point of ecological research for over a century, but thus far has been poorly explored through the lens of modern phylogenetic and trait-based approaches to community assembly. The vast majority of studies conducted to date have comprised static analyses where communities are observed at a single snapshot in time. Long-term datasets present a vantage point to compare established and emerging theoretical predictions on the phylogenetic and functional trajectory of communities through succession. We investigated within, and between, community measures of phylogenetic and functional diversity in a fire-prone heathland along a 21-year time-series. Contrary to widely-held expectations that increased competition through succession should inhibit the coexistence of species with high niche overlap, plots became more phylogenetically and functionally clustered with time since fire. There were significant directional shifts in individual traits through time indicating deterministic successional processes associated with changing abiotic and/or biotic conditions. However, relative to the observed temporal rate of taxonomic turnover, both phylogenetic and functional turnover were comparatively low, suggesting a degree of functional redundancy amongst close relatives. These results contribute to an emerging body of evidence indicating that limits to the similarity of coexisting species are rarely observed at fine spatial scales.

4.2 Introduction

Given limited scope for experimental manipulation in natural systems, a common approach in community ecology is to infer the mechanisms structuring communities from the distribution of their component species and traits. Inferring processes from patterns is of course non-trivial, relying as it necessarily does on a raft of assumptions about how the components of communities (i.e. species) respond to each other and their environment. This *modus operandi* is nowhere more apparent than in the phylogenetic and trait-based analyses of community assembly that have proliferated over the last 10–15 years (e.g. Webb, 2000, Cornwell & Ackerly, 2009, Kraft & Ackerly, 2010). To date, the vast majority of phylogenetic and trait-based studies of community assembly have comprised ‘static’ analyses where assembly processes are inferred from patterns observed at a single snapshot in time (as reviewed in Swenson, 2013, Götzenberger *et al.*, 2012). By necessity, static studies of this kind either ignore the dynamic properties of communities, treat community assembly as a one-off event, or at best assume that observed patterns are representative of prevailing processes. While this assumption may hold in some late successional systems, in dynamic or frequently disturbed systems, the processes that govern community structure may fluctuate considerably over time.

Disentangling sequential assembly processes from observed temporal patterns is complicated by competing and/or unresolved theoretical predictions. One oft-repeated axiom of community ecology holds that competition inhibits species with high niche overlap from coexisting, while environmental filtering has the opposite effect of limiting the range of successful ecological strategies at any one location (Weiher & Keddy, 1995, Stubbs & Wilson, 2004, Purschke *et al.*, 2013). It follows logically that if ecological niches are phylogenetically conserved, these two apparently opposing processes will leave different signatures on the phylogenetic structure of communities; competition will drive phylogenetic divergence, whilst strong environmental filters will lead to communities consisting largely of close relatives (Webb, 2000, Webb *et al.*, 2002). Coupling this framework with classical successional theory (Clements, 1916, Connell & Slatyer, 1977, Walker & Chapin, 1987, Wilson, 1999), we might anticipate communities will transition from exhibiting functional and phylogenetic convergence early in succession to becoming increasingly functionally and phylogenetically dispersed as competition increases in relative importance. However, even when niches are phylogenetically conserved, it has recently been argued that this dichotomous

framework makes untenable assumptions about the relative importance of niche differences and fitness differences in determining the outcome of community assembly (Chesson, 2000b, Mayfield & Levine, 2010). As recognised by Mayfield & Levine (2010), when differences in competitive ability exceed niche differences for a large proportion of the species pool, competition may exclude all but the most effective resource competitors. From this alternative perspective, we might predict phylogenetic and functional convergence, rather than divergence, if competition increases through succession. Evidence that competition may indeed drive phylogenetic convergence has recently begun to emerge from a variety of systems and taxa (Kunstler *et al.*, 2012, Bennett *et al.*, 2013, Price & Pärtel, 2013, Narwani *et al.*, 2013).

Given the paucity of suitable long-term datasets, most phylogenetic and trait-based research on the effects of disturbance and/or succession on community structure has been limited to static comparisons of relatedness and functional similarity in disturbed versus non-disturbed communities (Verdu & Pausas, 2007, Knapp *et al.*, 2008, Dinnage, 2009, Helmus *et al.*, 2010), or across different successional states in a chronosequence (i.e. a space-for-time substitution) (Verdu *et al.*, 2009, Letcher, 2010, Kunstler *et al.*, 2012, Purschke *et al.*, 2013). With a few notable exceptions (Verdu *et al.*, 2009, Kunstler *et al.*, 2012), most studies have reported greater functional and/or phylogenetic dispersion in undisturbed or late successional communities, including along a rare time series (Norden *et al.*, 2012). Nevertheless, given the known dangers of space-for-time substitutions in ecological research (Johnson & Miyanishi, 2008), additional temporal successional studies are needed to more robustly explore the generality of this pattern.

An important advantage of phylogenetic and functional analyses of temporal datasets is that it enables the compilation of species pools that are a truer representation of potential colonisers. In the past, static studies have been criticised for deriving species pools from regional species-lists which may include numerous taxa that may never colonise a site even in the absence of competitors (Grime, 2006, de Bello *et al.*, 2012). This coarse approach may potentially bias analyses towards finding phylogenetic or functional convergence resulting from broad-scale environmental filtering (Götzenberger *et al.*, 2012). An alternative approach is to try to eliminate the effects of large scale filters *a priori* by constraining the species pool to known colonisers of a site (de Bello *et al.*, 2012). This of course is limited by the availability of data collected over a sufficiently long period of time to record not only those species present at any given time but also the aptly-named dark diversity (Pärtel *et al.*, 2011) i.e. species that may only be

competitive (and therefore more likely to be detectable) for a brief period during community assembly/succession. Long-term studies, where the presence of species is monitored at intervals at permanent sites, make this achievable.

In this study we investigated temporal trends in the phylogenetic and functional community structure of understorey plants in fire-prone heathland in southeast Australia. Previous work has provided evidence of strong competitive hierarchies related to vertical stature in this system, with overstorey shrubs typically eliminating understorey species through succession post-fire (Keith & Bradstock, 1994, Tozer & Bradstock, 2003, Keith *et al.*, 2007). However, unlike much of the existing literature on phylogenetic and functional community structure through succession in plant communities (e.g. Letcher, 2010, Norden *et al.*, 2012, Kunstler *et al.*, 2012, Bhaskar *et al.*, 2014), here we explicitly focus on understorey communities. With access to compositional data collected over more than 20 years through multiple fire events, this study represents one of the most comprehensive assessments of temporal dynamics in both phylogenetic and functional community structure to date.

Whilst the phylogenetic and functional structure of plant communities may arise through a complex interplay of various evolutionary (e.g. trait evolution and niche conservatism) and ecological processes (e.g. competition, environmental filtering, herbivory etc.), we concentrated on a subset of hypotheses that reflect the competing theories which have received the most attention in the recent literature. Firstly, assuming fire acts as a filter on the species pool we hypothesised that plots would exhibit functional clustering immediately following fire, and correspondingly also exhibit phylogenetic clustering if the traits mediating early dominance are conserved. Alternatively, if key assembly traits are not conserved, or if measured functional traits have little bearing on community assembly, then we would expect functional and phylogenetic patterns to be uncorrelated. In addition, we made alternative predictions about the trajectories of communities in the period following fire. If increased competition through succession enforces a limit on the similarity of coexisting species, communities should become increasingly functionally dispersed though time, and therefore also phylogenetically dispersed if species function is conserved. Alternatively, if increased competition results in the exclusion of all but the most dominant resource competitors (*sensu* Mayfield & Levine, 2010), or if assembly is primarily driven by fluctuating environmental processes, functional and phylogenetic clustering should remain static or increase through time.

Our secondary aim was to infer what processes are likely to be driving any observed trends. To this end, we not only considered within community structure but also trends in community-weighted mean trait values and rates of temporal phylogenetic and functional turnover relative to taxonomic turnover. This provided a means to evaluate the extent to which community assembly through succession is structured by deterministic or stochastic processes (e.g. Swenson *et al.*, 2012). For instance, a deterministic model of community assembly assumes that species turnover through time is non-random with respect to species function, and by inference phylogeny. If biotic and abiotic conditions are relatively constant through time, we would predict less phylogenetic and/or functional turnover relative to the observed rate of taxonomic turnover. Conversely, if conditions fluctuate though time, we would predict greater than expected phylogenetic/functional turnover. Finally, if function and phylogeny have no bearing on community assembly (i.e. a stochastic or neutral model *sensu* Hubbell, 2001), we should expect taxonomic turnover to be random with respect to phylogeny and function.

4.3 Materials and methods

4.3.1 Study site, sampling and fire history

The study was conducted in an area of fire-prone coastal heathland in Royal National Park, New South Wales, Australia (centred on $34^{\circ}05'46.00''$ S $151^{\circ}09'02.73''$ E). The vegetation in the area is characterised by sclerophyllous plants, with a herbaceous ground layer dominated by species within the Restionaceae, Cyperaceae and Poaceae families, and woody overstorey layers dominated by shrubs within the Proteaceae, Myrtaceae, Ericaceae and Fabaceae. The vegetation is fire-prone, with the herbaceous component regenerating rapidly and gradually becoming overtapped by shrubs within 5-6 years post fire (Keith *et al.*, 2007). The soils, which derive from sandstone, tend to be highly infertile, acidic and siliceous. The topography is relatively flat with elevations ranging from 68 to 72 m above sea level.

In 1990, fifty-six permanent 0.25 m^2 plots were established along eight transects arranged in pairs, with each transect comprising seven plots. Plots in each transect pair are spaced an average of 18 metres apart (min = 5 m, max = 45 m), while plots in different transect pairs are separated by an average distance of 211

metres (min = 104 m; max = 323 m). Henceforth we refer to these two discrete spatial scales as the ‘plot’ scale ($n = 56$) and the ‘site’ scale ($n = 4$). Since 1990, the total abundance (number of stems) of all herbaceous species within each plot has been censused on nine separate occasions (1990–1994, 1999, 2002, 2007 and 2011) (Keith & Tozer, 2012). A fire in 1988 prior to the first census burnt the entire site, with subsequent fires in 1994 (entire site burnt) and 2001 (14 plots along one transect-pair burnt). The 1994 fire occurred prior to the census of that year.

4.3.2 Species pool and community phylogeny

A species pool was defined comprising all herbaceous angiosperm species (49 unique taxa) recorded across all 56 plots since monitoring began. A species accumulation curves generated from random permutation of samples indicated adequate sampling of the species pool (Appendix C). Non-angiosperms were excluded because they were at low abundance and are known, due to their early divergence from all other species in the pool, to have a disproportionately large effect on metrics of phylogenetic community structure (Kembel & Hubbell, 2006). A community phylogeny (figure 4.1), derived from DNA sequence data for two commonly used plastid gene regions (*rbcL* and *matK*), was generated using the programs phyloGenerator (Pearse & Purvis, 2013) and BEAST (Drummond & Rambaut, 2007). Full phylogeny construction details are provided in Appendix C.

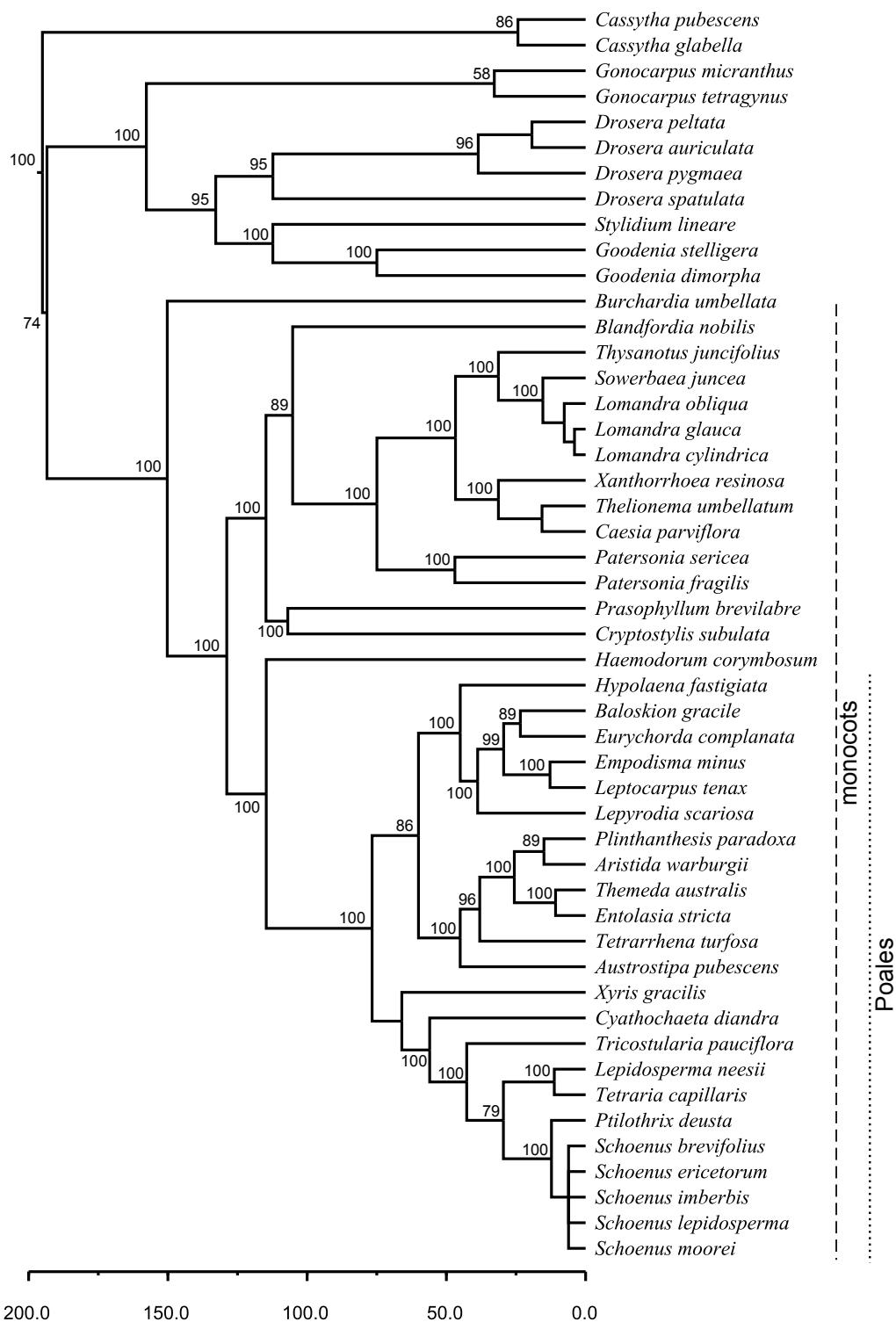


Fig. 4.1 Community phylogeny for all species recorded in plots across the full monitoring period. Vertical dashed/dotted lines denote species belonging to two nested taxonomic groups (monocots and Poales) on which separate analyses were performed. Node labels denote posterior support values, with unlabelled nodes indicating points where taxa were manually added to the phylogeny post processing. Timescale is in millions of years before present.

4.3.3 Functional traits

All recorded species were scored for seven functional traits related to competitive ability and/or tolerance of disturbance (seed mass, plant height, Raunkiær life-form, fire response, fecundity, longevity and seedbank persistence) (Westoby *et al.*, 2002, Adler *et al.*, 2013b). All trait data were obtained from existing databases, the primary literature and expert knowledge, with the exception of seed mass which was supplemented with measurements made from herbarium specimens. Seed mass (mg) and plant height (cm) were scored on a continuous scale, while the five remaining traits were coded on an ordinal scale. This included two strictly categorical traits (fire response and life-form) and three implicitly continuous traits (fecundity, longevity and seedbank persistence) that owing to the absence of sufficiently high resolution quantitative data were classified in bins. Fire response was coded with three levels (killed = 1, facultative resprouter = 2, obligate resprouter = 3); life-form with three levels (geophyte = 1, hemicryptophyte = 2, epiphyte = 3); fecundity with four levels (low = 1, low-moderate = 2, moderate = 3, high = 4); longevity with five levels (<5 years = 1, 5–10 years = 2, 10–25 years = 3, 25–50 years = 4, >50 years = 5); and seedbank persistence with three levels (transient = 1, moderate persistence = 2, persistent = 3). Several commonly measured leaf traits recognised to represent important axes of niche differentiation (e.g. specific leaf area and leaf dry matter content) were not included in the study due to a large proportion of the dominant species lacking true leaves.

In order to assess the correlation between phylogenetic relatedness and functional similarity we tested for significant phylogenetic signal in continuous traits using Blomberg's K statistic (Blomberg *et al.*, 2003) and in ordinally coded traits using the 'fixed tree, character randomly reshuffled' model of Maddison & Slatkin (1991) with ordered costs for character state transitions.

4.3.4 Temporal change in phylogenetic and functional community structure

Phylogenetic community structure within individual plots at each census was evaluated using two commonly used metrics: mean nearest taxon distance (MNTD, mean distance separating each species in each community from its closest relative), and mean pairwise phylogenetic distance (MPD, mean pairwise distance between all species in each community) (Webb, 2000, Webb *et al.*, 2002).

While these two metrics are typically correlated they provide complementary information, with MNTD being more sensitive to clustering or dispersion near the tips of the phylogeny, and MPD being more sensitive to tree-wide patterns of clustering or dispersion (Kembel *et al.*, 2010). For both MNTD and MPD we used a square-root transformation of phylogenetic distance in order to account for non-linear scaling of phylogenetic relatedness and functional distance (Lettent & Cornwell, 2014). Standardized effects sizes (SES) for MNTD and MPD were obtained by comparing observed values to those expected under a null model of community assembly. Both observed and null values were quantified using abundance-weighted data. A null model was used that randomly shuffled the names of individuals across the tips of the phylogeny 999 times. This is considered to be the most appropriate null model for temporal analyses (Swenson *et al.*, 2012, Norden *et al.*, 2012).

An identical framework was used to evaluate functional community structure at each census, where the functional analogue of MNTD (F-MNTD) represents the mean distance to each species' nearest neighbour in multi-trait space, and the functional analogue of MPD (F-MPD) represents the mean pairwise distance in multi-trait space between all species in the community. A Gower distance (which allows for range-standardised quantitative and qualitative data) was used to generate the functional distance matrix representing the functional similarity of species in multivariate trait space.

In order to account for potential sensitivity of community-wide patterns to spatial scale (Swenson *et al.*, 2006), all analyses of individual plots were replicated at a larger spatial scale by summing plot composition within each site (transect-pair, $n = 4$). In addition, whilst we were primarily interested in the trajectory of the full understorey community through time, we replicated all analyses at three nested phylogenetic scales: i) all species; ii) monocots; and iii) Poales (figure. 4.1).

To examine trends in phylogenetic and functional community structure, linear models and linear mixed-models were fit for phylogenetic and functional community structure as a function of time. To account for spatial and temporal non-independence, random effects were fit for transect-pair (random intercept; transect-pairs that didn't burn in 2001 only) and plots (random slope and intercept) at the plot scale, and for transect-pair (random intercept; transect-pairs that didn't burn in 2001 only) at the site scale. Models were fit for either side of the 1994 fire i.e. 1990-1993 and 1994-2011. For the period from 1994-2011, separate models were fit for those plots/transect-pairs that haven't burnt since 1994 ($n = 42/3$) and those that burnt in the 2001 fire ($n = 14/1$). Given that coefficients may

be biased for random effects with fewer than five levels, we checked our results against those obtained when using transects ($n = 8$) as a random effect at the plot scale, and when treating transect as the site grouping at the site scale. All analyses of phylogenetic and functional community structure were performed with the R-package 'picante' (Kembel *et al.*, 2010). An R^2 summarising the variance in phylogenetic and functional community structure explained by time was calculated using the approach for mixed-effects models provided by Nakagawa & Schielzeth (2013) and extended by Johnson (2014). Finally Welch's t-test was used to determine whether observed differences in the temporal trajectory between sites with different burn frequencies may be attributable to different initial values.

4.3.5 Temporal trends in community-weighted mean trait values

To complement the core analyses, we also investigated temporal trends in community-weighted mean trait values at the plot scale. Community-weighted mean trait values were calculated for each trait in each plot at each time-step by weighting species' trait values by their proportional abundance. For the purposes of obtaining a single value for each trait in each plot, ordinal traits were treated quantitatively. As for measures of phylogenetic and functional community structure, to examine temporal trends, linear models and linear mixed effect models were fit for each community-weighted mean trait value as a function of time.

4.3.6 Temporal phylogenetic and functional beta turnover

To quantify temporal phylogenetic and functional beta turnover we used between community equivalents of MNTD and MPD which provide a measure of the phylogenetic and functional dissimilarity between pairs of plots sampled over different years (Ricotta & Burrascano, 2009, Swenson *et al.*, 2012). The formulas for calculating phylogenetic and functional nearest neighbour dissimilarity (D_{nn} , the beta diversity analogue of MNTD) and pairwise dissimilarity (D_{pw} , the beta diversity analogue of MPD) are provided in Appendix C. To determine whether temporal phylogenetic and functional beta diversity was different from that expected given the rate of taxonomic turnover, we compared observed values to those expected under null models. As for MNTD and MPD, null models for D_{nn} and D_{pw} were generated by randomly shuffling individuals across the tips of the phylogeny or the columns of the functional distance matrix 999 times. Given the large number of possible temporal pairwise comparisons, we only considered

the rate of phylogenetic and functional beta turnover of all census points relative to the first census in 1990 and relative to the immediately preceding census point. As for within-community measures, the above analyses were replicated at the two additional phylogenetic scales (monocots and Poales).

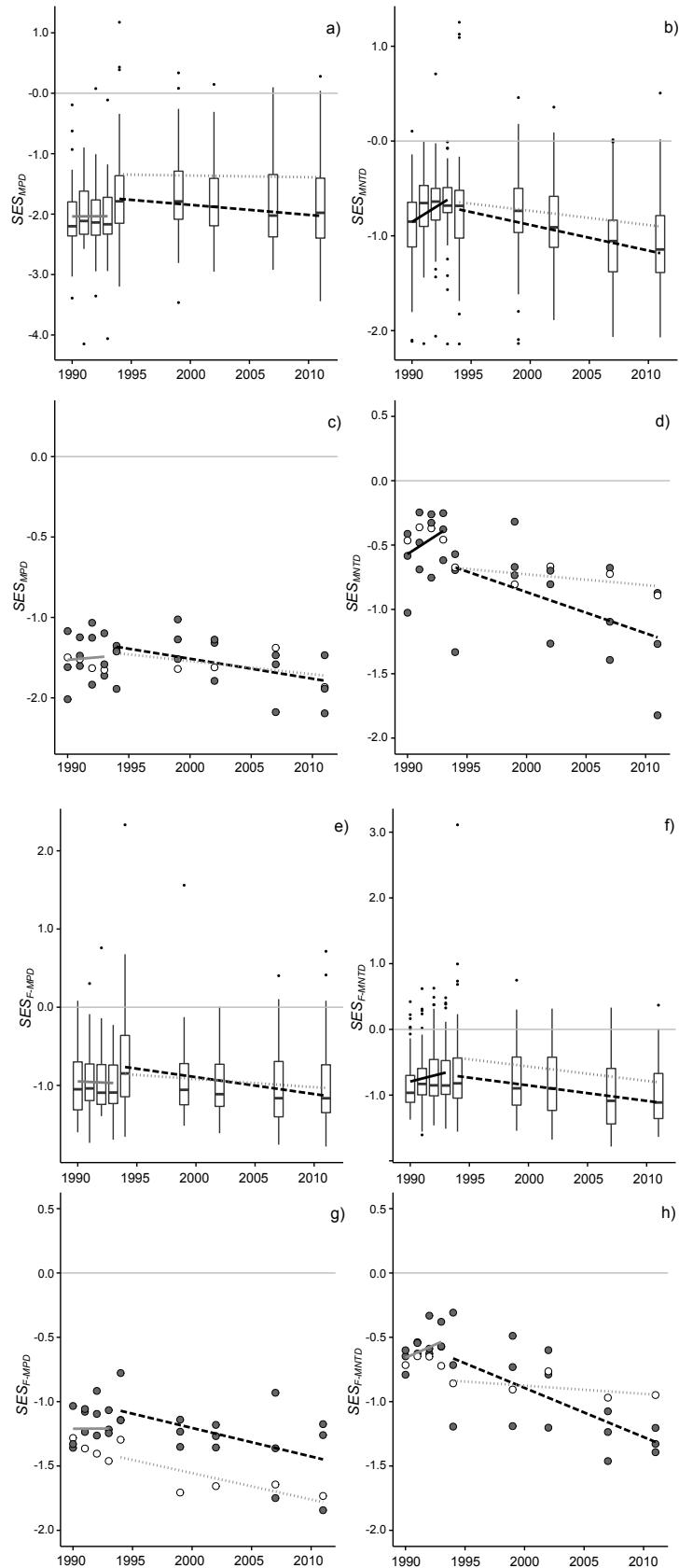
4.4 Results

4.4.1 Temporal change in phylogenetic and functional community structure

Throughout the 20 year survey period, species composition at both spatial scales was consistently phylogenetically clustered relative to the species pool (figure 4.2). Clustering was particularly pronounced for MPD with 53% of plots and 100% of sites exhibiting significant clustering ($SES_{MPD} < -1.96$) over the full sampling period. Between 1990 and 1993 there was a weak increase (towards zero from negative) in MNTD at both the plot scale ($R^2 = 0.05$, $p < 0.001$) and the site scale ($R^2 = 0.10$, $p < 0.05$) (figure 4.2 b & d), but no significant trend in MPD (figure 4.2 a & c). In contrast, following the 1994 fire, both MNTD and MPD exhibited a consistent decreasing trend in those plots (MNTD: $R^2 = 0.10$, $p < 0.001$; MPD: $R^2 = 0.02$, $p < 0.05$) and sites (MNTD: $R^2 = 0.29$, $p < 0.005$; MPD: $R^2 = 0.16$, $p = 0.06$) that did not burn again in 2001. The plots and single site that burnt again in 2001 had flatter, non-significant slopes for both MNTD and MPD over the same time interval.

When the species pool was constrained to just monocots or Poales, results were similar to those described for the full species pool with all spatial-phylogenetic

Fig. 4.2 (facing page) Phylogenetic and functional community structure through time based on the standardized effect size of: mean pairwise phylogenetic distance (MPD) between all species in each a) plot and c) site; mean phylogenetic distance between nearest taxonomic neighbours (MNTD) in each b) plot and d) site; mean pairwise functional distance (F-MPD) between all species in each e) plot and g) site; mean functional distance between nearest taxonomic neighbours (F-MNTD) in each f) plot and h) site. Trendlines correspond to models of MNTD/MPD vs. time for all plots/sites through the first four years of sampling (solid line); plots/sites that only burnt in 1994 (dashed line) and plots that burnt in 1994 and 2001 (dotted lines). Trend lines shaded black indicate significant slope coefficients at $p < 0.05$; grey lines indicate insignificant slopes. In a, b, e & f, boxplot midline corresponds to median value for all 56 plots at each census point; upper and lower hinges give the first and third quartiles; whiskers extend from the hinge to the highest value that is within $1.5 \times$ interquartile range; and data points beyond whiskers are shown as outliers. In c, d, g & h, open circles correspond to sites that burnt in the 2001 fire; grey-filled circles correspond with sites that did not burn in 2001.



scale combinations exhibiting a decreasing trend in MPD and MNTD following the 1994 fire (Appendix C, figure S4.1-S4.2). The only notable difference between the main results and those obtained with reduced species pools was less negative effect sizes in the latter case. This was particularly true for species pools limited to Poales, for which MPD and MNTD was mostly positive, but still low with only MNTD at the plot scale having any (< 1%) significantly phylogenetically over-dispersed plots over the entire monitoring period.

All seven measured traits exhibited significant ($p < 0.05$) phylogenetic signal (Appendix C, table S4.1). In addition, patterns in functional community structure broadly mirrored those observed for phylogenetic community structure. Almost all plots and sites exhibited functional clustering throughout the monitoring period (figure 4.2 e-h), although notably within the -1.96 threshold for statistical significance. In addition, temporal trends were similar to the phylogenetic analyses, with plots/sites that last burnt in 1994 becoming increasingly functionally clustered through time (plots [F-MNTD: $R^2 = 0.07$, $p < 0.001$; F-MPD: $R^2 = 0.06$, $p < 0.001$]; sites [F-MNTD: $R^2 = 0.45$, $p < 0.001$; F-MPD: $R^2 = 0.27$, $p < 0.05$]). As with the phylogenetic analyses, the plots and single site that burnt again in 2001 had flatter, non-significant slopes, although it is important to note that at the site scale statistical power was low given the small number ($n = 5$) of data points. Prior to the 1994 fire, the only significant trend was a very weak increase (towards zero from negative) in F-MNTD at the plot scale ($R^2 = 0.01$, $p < 0.01$). When the species pool was constrained to just monocots or Poales results were qualitatively identical to the main results (Appendix C, figure S4.3-S4.4), but as for the phylogenetic analyses, effect sizes were reduced.

The observed differences in temporal trajectory cannot be attributed to different initial values given that there was no significant differences at the start of the sampling period between sites with different subsequent burn frequencies (MNTD: $t = -0.8406$, $p = 0.4312$; MPD: $t = 0.7769$, $p = 0.4431$; F-MNTD: $t = 0.7769$, $p = 0.4431$; F-MPD: $t = 1.3654$, $p = 0.1812$). In addition, coefficients and standard errors were nearly identical when transects was treated as a random effect at the plot scale, or when the site scale was formed by summing abundance across transects rather than transect pairs. Finally, it is worth noting that effect sizes were weaker but remained predominately negative when null randomisations were restricted to those species observed within any given year.

4.4.2 Temporal trends in community-weighted mean trait values

In common with observed patterns of phylogenetic and functional community structure, temporal trends in mean-trait values tended to differ between plots with different burn frequencies (figure. 4.3). Over the four years prior to the 1994 fire, mean-trait values in all plots were comparatively static, with time explaining no more than 3% of the variation in community weighted mean trait values for any given trait. In contrast, following the 1994 fire, plots within those transects that went unburnt for the remainder of the sampling period exhibited a significant increase in mean seed-weight ($R^2 = 0.07$, $p < 0.005$), longevity ($R^2 = 0.09$, $p < 0.001$) and the proportion of obligate resprouters ($R^2 = 0.16$, $p < 0.001$), and a decrease in fecundity ($R^2 = 0.18$, $p < 0.001$) and the proportion of hemicryptophytes ($R^2 = 0.10$, $p < 0.001$). In contrast, plots that burnt again in 2001 exhibited mostly static mean-trait values, with only a small decrease in fecundity ($R^2 = 0.02$, $p < 0.001$).

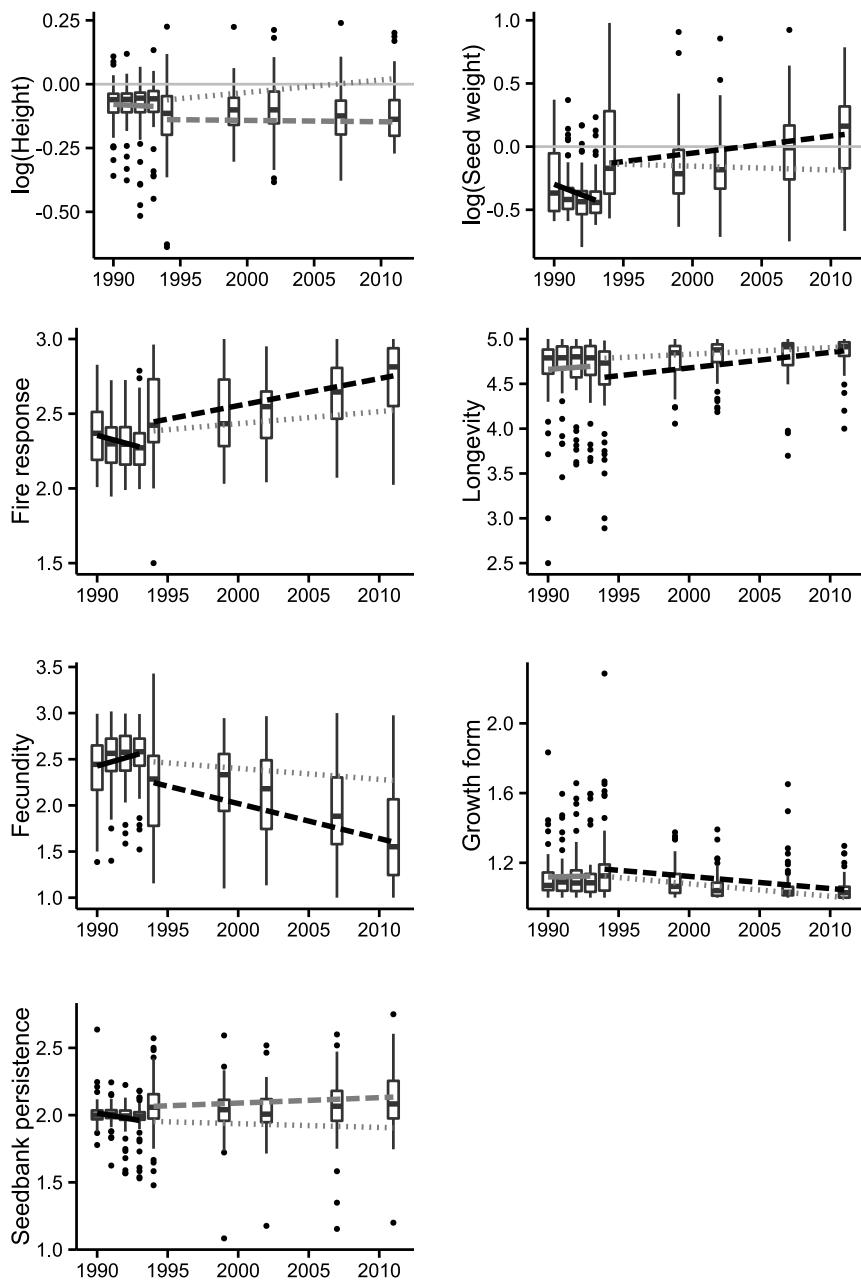


Fig. 4.3 Community-weighted mean trait-values through time. Trendlines correspond to models of individual traits vs. time for all plots/sites through the first four years of sampling (solid line); plots/sites that only burnt in 1994 (dashed line) and plots that burnt in 1994 and 2001 (dotted lines). Trend lines shaded black indicate significant slope coefficients at $p < 0.05$; grey lines indicate insignificant slopes. Boxplot midline corresponds to median value for all 56 plots at each census point; upper and lower hinges give the first and third quartiles; whiskers extend from the hinge to the highest value that is within $1.5 \times$ interquartile range; and data points beyond whiskers are shown as outliers.

4.4.3 Temporal phylogenetic and functional beta turnover

Patterns of phylogenetic and functional beta turnover relative to the first census in 1990 were similar to those relative to the immediately preceding census point (for the latter see Appendix C, figure S4.5). Observed temporal phylogenetic turnover was strongly dependent on the metric used and the phylogenetic resolution of the analysis (figure 4.4). When evaluated under a nearest-neighbour dissimilarity metric for all species (D_{nn}), phylogenetic turnover tended to be relatively random with respect to taxonomic turnover, except during early census-point comparisons when some plots showed greater than expected phylogenetic turnover. In contrast, under a pairwise dissimilarity metric (D_{pw}), temporal phylogenetic turnover was mostly less than that expected given observed taxonomic turnover for all species, but tended to exhibit more random patterns of turnover when the species pool was limited to monocots and in particular to Poales (Appendix C, figure S4.6). Taken together, these results indicate that plot composition through time tended to be constrained from a phylogeny-wide perspective, i.e limited to a small subset of clades (mainly Poales) relative to the community phylogeny. However, within those well represented clades, turnover tended to be either largely random as indicated by (D_{pw}) for Poales, or occasionally directional as indicated by high values for (D_{nn}) for the whole phylogeny. This latter inference is based on (D_{nn}) being more sensitive to turnover near the tips of the phylogeny.

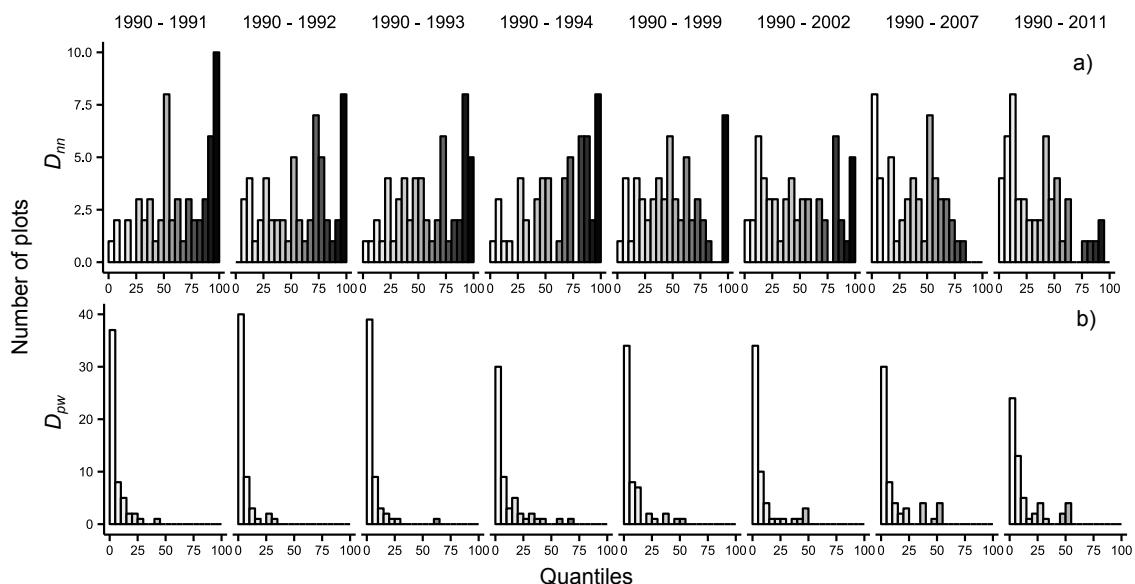


Fig. 4.4 Temporal phylogenetic beta turnover for all species quantified by a) D_{nn} , and b) D_{pw} . Low quantile scores (white) indicate low turnover in phylogenetic composition relative to the observed rate of taxonomic turnover; high quantile scores (black) indicate high turnover in phylogenetic composition relative to the observed rate of taxonomic turnover. Values <2.5 or >97.5 are significant at the 0.05 level.

For both metrics and all three higher phylogenetic scales, observed temporal functional turnover was less than expected given taxonomic turnover (Appendix C, figure S4.7-S4.9) suggesting that the overall functional composition, as derived from the seven measured traits, was relatively fixed through time. However, as indicated by trends in community-weighted mean trait values and observations of temporal turnover in individual traits, this appeared to be driven by low turnover in several traits, in particular plant height and seedbank persistence.

4.5 Discussion

A classical axiom of community ecology holds that abiotic processes (e.g. environmental filtering) dominate early in succession, while the relative importance of biotic processes (e.g competition) increases as communities mature (Clements, 1916, Connell & Slatyer, 1977, Walker & Chapin, 1987, Wilson, 1999). Assuming that closely related and functionally similar species compete most intensely, it follows that during succession communities may be expected to transition from those that are dominated by closely related and/or functionally similar taxa to those that comprise more phylogenetically and functionally dispersed assemblages (Puschke *et al.*, 2013, Bhaskar *et al.*, 2014). In the present study we found no evidence for an increase in phylogenetic and functional dispersion over 20 years of succession post-fire in a coastal heathland community. Instead, species composition at both the plot scale and the site scale tended on average to become more phylogenetically and functionally clustered over time, except notably in those plots where succession was interrupted by fire.

These findings contrast with a number of previous studies, where an opposite pattern of increased phylogenetic and/or functional dispersion has been reported in non-disturbed versus disturbed communities (Dinnage, 2009, Helmus *et al.*, 2010), in late successional stages along chronosequences (Letcher, 2010, Puschke *et al.*, 2013), and along a rare time series of succession (Norden *et al.*, 2012). However, in several rare exceptions to the rule that echo our own findings, Verdu *et al.* (Verdu *et al.*, 2009) found that the competitive exclusion of pioneer species appeared to reduce phylogenetic diversity in the very latter stages of a chronosequence of post-fire succession, while Kunstler *et al.* (Kunstler *et al.*, 2012) attributed greater functional and phylogenetic convergence with increasing forest plot age to competition sorting species along a competitive hierarchy in plant height. Most recently, Bhaskar *et al.* (Bhaskar *et al.*, 2014) found little evidence

for an increase in functional dispersion in secondary tropical forests at various stages post agricultural abandonment. Together with these earlier studies, our results reinforce growing awareness of the complex, and often unpredictable, interaction between community assembly processes and patterns of phylogenetic and functional community structure (Mayfield & Levine, 2010, Kunstler *et al.*, 2012, Price & Pärtel, 2013, Bennett *et al.*, 2013).

While it remains difficult to infer underlying processes from static snapshots of community structure, the observed directional shifts in several community-weighted traits through time are indicative of deterministic turnover in this system. These trends tended to be consistent with predictions based on known relationships between functional traits and life-history strategies (Westoby *et al.*, 2002, Adler *et al.*, 2013b) i.e. the transition from species with relatively fast life-histories (small seeds, high fecundity and short life-spans) in early succession, to those with slower life histories (large seeds, low fecundity and long life-spans) in late succession (figure. 4.3). In particular, the observed increase in community weighted mean seed size, a trait often correlated with shade tolerance (Coomes & Grubb, 2003), is consistent with the replacement of good dispersers and fast germinators/resprouters in the immediate post-fire environment by understorey species with greater shade tolerance as the shrub layer becomes more established. It is also telling that the same traits that exhibited directional shifts in the unburnt plots post 1994, were mostly temporally static in those plots that burnt again in 2001, suggesting that fire prevented the competitive exclusion of early successional species by late successional species.

Mediated by the shrub overstorey, light deprivation to the ground layer may act more like an abiotic stressor external to the system. As such, we might attribute phylogenetic and functional clustering later in succession to a filtering process that is biotic but largely external to the ground layer component of the community. However, even if we treat the effect of shading as an external process, competition for resources in the light deprived understorey is still likely to play a significant role in the assembly process. Indeed, at the local scale of interacting species, relative shade tolerance will translate into fitness differences, whereby the more shade tolerant species are at a competitive advantage. If there is a hierarchy of shade tolerance within the ground layer, competition amongst ground layer species will be expected to drive functional clustering. The notion that competition may drive functional and phylogenetic clustering via hierarchical differences in species' competitive ability was highlighted by Mayfield and Levine (Mayfield & Levine, 2010), and has received empirical backing by several recent

studies (Kunstler *et al.*, 2012, Bennett *et al.*, 2013, Narwani *et al.*, 2013). However, to our knowledge this is the first study to present patterns of phylogenetic and functional community structure consistent with this process emerging through a successional time-series.

In spite of the observed directional shifts in community weighted mean-trait values, overall functional and pairwise phylogenetic beta-diversity was comparatively constrained through time. This however was counter-balanced by a relatively high rate of temporal phylogenetic beta-diversity amongst closely related taxa early in succession (D_{nn} , figure 4.4). Together with largely stochastic patterns of turnover amongst the highly represented Poales, a picture emerges of a system in which the dominant species are consistently drawn from a small number of clades within the overall phylogeny. However, within those well represented clades, turnover appears largely stochastic or neutral, with functionally similar close relatives replacing each other through time. This disconnect between results observed at the individual trait level and summary statistics of overall phylogenetic and functional turnover highlights the importance of combined approaches. One explanation is that despite directional shifts in individual traits indicating that the relative fitness of different traits is changing through time, there is still substantial functional redundancy near the tips of the phylogeny, with the fate of individual taxa being relatively stochastic (Rosenfeld, 2002).

Our approach focused specifically on the ground layer within a vertically stratified community, whereas most previous studies of phylogenetic and functional community structure through succession in plant communities have focused on the canopy layer (e.g. Letcher, 2010, Norden *et al.*, 2012) or have been conducted in less vertically stratified communities such as grasslands (e.g. Purschke *et al.*, 2013). Where data are available for both canopy and ground layer components of a community, separate analyses for each strata may be preferable to avoid conflating the relative effects of different assembly processes in each strata. Phylogenetic and functional convergence through succession may be a more general feature of understorey communities than it is of less stratified or canopy ‘communities’ (but see Kunstler *et al.*, 2012, for an example of convergence amongst canopy species in temperate forests). Ultimately, more phylogenetic and functional based studies of community assembly through succession in the understorey of vertically stratified communities are needed to verify this hypothesis.

Disentangling the relative contribution of abiotic and biotic processes in driving community assembly through succession remains a major challenge for community ecologists. Contrary to expectations, here we have shown that phylogenetic

and functional dispersion is not the only, or necessarily the most likely, outcome of succession. This finding contributes to an emerging body of research re-evaluating the role of limiting functional similarity in determining the outcome of community assembly (Mayfield & Levine, 2010, Kunstler *et al.*, 2012, Price & Pärtel, 2013, Bennett *et al.*, 2013). Efforts to partition out the relative importance of the underlying processes driving functional and phylogenetic convergence through succession in this, and other systems, will be a valuable focus of future work.

Acknowledgements

Thanks to Will Pearse for assistance with phyloGenerator, Nate Swenson for providing code for the analysis of phylogenetic and functional beta diversity, and Jodi Price, Will Cornwell and two anonymous reviewers for valuable comments on an earlier draft of this paper.

Chapter 5

Fine-scale hydrological niche differentiation through the lens of multi-species co-occurrence models

Andrew D. Letten, David A. Keith, Mark G. Tozer and Francis Hui

Journal of Ecology (2015)
<http://dx.doi.org/10.1111/1365-2745.12428>



This study was conceived by ADL with input from DAK. DAK and MGT provided floristic plot data and edaphic data. ADL conducted soil moisture sampling and collected supplementary floristic and edaphic data. FH and ADL developed latent variable co-occurrence models. ADL conducted analyses and wrote the manuscript, with contributions from DAK, MGT & FH.

5.1 Abstract

1. Theory suggests spatial heterogeneity can facilitate species co-occurrence at fine-scales, but environmental data is rarely collected at sufficiently high resolution to test this empirically. While there is emerging evidence subtle variation in soil hydrology represents a fundamental fine-scale niche axis within plant communities, this is largely derived from studies of soil hydrology in isolation from other environmental factors.
2. We assessed the comparative importance of fine-scale hydrological niche differentiation for species co-occurrence using a high resolution study of soil hydrology and other edaphic variables, coupled with a long-term (24 years) dataset of herbaceous plant plots in a heathland community in southeast Australia.
3. For the analysis, we employed novel latent variable models (LVMs), which offer an explicit, model-based approach to partitioning out the different drivers of species co-occurrence patterns. While the regression component of an LVM models the species-specific environmental responses, the latent variable component can be used to identify residual patterns of co-occurrence, which may be attributable to unmeasured factors and/or biotic interactions.
4. Relative to a host of plant resources, non-resource factors and ‘unmeasured’ latent variables, soil hydrology emerged as the best predictor of negative co-occurrences within the community, with the dominant species exhibiting strongly differentiated responses across a comparatively narrow moisture gradient. Nevertheless, strong species-specific responses to environmental variability only emerged at scales greater than those at which plants may be expected to compete for resources, throwing doubt on the direct role of spatial heterogeneity as a mechanism for local-scale coexistence.
5. *Synthesis.* This study confirms the vital role of hydrological niches for the maintenance of within-community plant diversity, but also highlights the need for more rigorous analysis of scale dependencies to better understand the underlying coexistence mechanisms at play. In addition, it illustrates the inferential gains made possible with model-based approaches to the analysis of species co-occurrence.

5.2 Introduction

Spatial environmental heterogeneity is widely recognised to play a fundamental role in driving broad-scale patterns of diversity and compositional turnover. The underlying mechanisms are diverse, but include wider niche space, more refuges from adverse conditions, and stronger selection for evolutionary diversification (reviewed in Stein *et al.*, 2014). At finer, within-community scales, however, the extent to which heterogeneity explains observed patterns of species co-occurrence is less clear. In spite of a wealth of theory suggesting that ‘variation dependent’ coexistence mechanisms may be pervasive (Chesson, 2000b, Amarasekare, 2003, Snyder & Chesson, 2004), the empirical evidence is surprisingly sparse (but see Sears & Chesson, 2007a, Fridley *et al.*, 2011). One explanation for this disconnect between theory and data is that environmental heterogeneity is rarely examined at sufficiently high resolution to evaluate its influence on within-community scale patterns of diversity (Adler *et al.*, 2013a, Kraft *et al.*, 2014). More specifically, in observational studies, the local environment is typically treated as spatially homogenous, with coexistence attributed to variation independent processes such as resource partitioning, the temporal storage effect or neutral dynamics (Chesson, 2000b, Hubbell, 2001). Matching this missing fine-scale environmental data with floristic observations will aid identification of the proximal causes of plant diversity maintenance at within-community scales.

Species coexistence arising through spatial heterogeneity is contingent on multiple criteria, but the most fundamental prerequisite of these is the presence of different species-specific responses to the environment (Chesson, 2000b, Silvertown, 2004). More specifically, for a given spatially heterogeneous environmental variable to operate as a plant niche axis, and thus reduce niche overlap, species need to favour different regions of that variable. It follows that the spatial structure of the variable needs to exhibit sufficient temporal stability such that individuals may germinate, reach reproductive maturity, and preferably build-up a local population before the environment changes (Grubb, 1977, Chesson, 2000b, Muko & Iwasa, 2000, Amarasekare, 2003). Of the vital plant resources, water in the form of soil moisture has repeatedly been shown to exhibit a high degree of temporal stability, whereby across fine-scales, localities may often be consistently ranked on the basis of their soil water content (Vachaud *et al.*, 1985, Cassel *et al.*, 2000, Pachepsky *et al.*, 2005). As such, given the essential physiological role of water in plant function, and the vast array of plant adaptations to water stress and uptake, there is a strong *a priori* expectation that in some systems fine-scale variability in

soil moisture may provide one of the more important axes of niche differentiation (Silvertown *et al.*, 2015).

Whilst ecologists have long been aware of the tendency for plant species to segregate along strong soil moisture gradients (Pickett & Bazzaz, 1978), evidence that subtle variation in soil moisture can also moderate co-occurrence patterns is only just emerging. Silvertown *et al.* (1999) first demonstrated fine-scale hydrological niche differentiation across topographically invariant landscapes in European wet meadows. More recently, Araya *et al.* (2011) underscored the potential generality of the phenomenon in *Restio* dominated fynbos communities in the Cape Floristic region of South Africa. Given their combined taxonomic and geographic breadth, together the evidence from these two studies is compelling, and yet there remains a strong mandate for wider investigation. In particular, we know very little about the comparative strength of soil moisture relative to other important plant resources (e.g. nitrogen, phosphorous, organic carbon) as drivers of fine-scale plant co-occurrence patterns.

Whilst manipulative experimental approaches arguably have the greatest potential to disentangle underlying mechanisms in annual plant communities (e.g. Sears & Chesson, 2007b, Godoy *et al.*, 2014), they are less amenable to perennial systems where individuals can take several years to reach reproductive maturity. Fortunately, recent advances in statistical methods and computing power have greatly enhanced the capacity of researchers to draw robust inferences from the kinds of purely observational datasets that are typical in studies of perennial communities. In particular, an emerging family of community-level model-based approaches have the potential to yield significant new insights (Ferrier & Guisan, 2006, Warton *et al.*, 2014). These include a subset of models which are specifically orientated towards disentangling patterns of co-occurrence and their abiotic and/or biotic drivers (Ovaskainen *et al.*, 2010, Pollock *et al.*, 2014, Harris, 2015). In contrast with conventional null-model based approaches, multivariate co-occurrence models (also known as Joint Species Distribution Models (JSDM), after Clark *et al.*, 2014) provide a direct means of assessing shared environmental responses separately from other (abiotic and biotic) processes that may generate non-random patterns of co-occurrence. Furthermore, as they are built upon the standard regression framework of generalized linear models (McCullough & Nelder, 1989), they facilitate more transparent and accurate representations of the statistical properties of the data (e.g. overdispersion of counts, Warton *et al.*, 2012), as well as provide a means of accounting for structural complexity and uncertainty in a hierarchical framework (Cressie *et al.*, 2009).

In this study, we assessed the comparative importance of hydrological niche segregation for species co-occurrence, using a high resolution study of soil hydrology and other edaphic variables, coupled with a long-term (24 years) dataset of herbaceous plant plots in a coastal heathland community in southeast Australia. To this end, we modelled pairwise co-occurrence using an extension of the latent variable model (LVM, Skrondal & Rabe-Hesketh, 2004) approach for model-based unconstrained ordination recently proposed by Walker & Jackson (2011) and Hui *et al.* (2014). LVMs offer an explicit, model-based approach to partitioning out the different drivers of species co-occurrence patterns. In particular, the regression component of an LVM models the species-specific environmental responses, based on which we can evaluate the prevalence of species-specific responses to soil hydrology, in comparison with a range of other important plant resources and non-resource factors. The latent variable component of the LVM is used to identify the residual patterns of species co-occurrence, which may be attributable to unmeasured factors and/or biotic interactions.

If heterogeneity in fine-scale hydrology is an important niche parameter, we predicted species would exhibit strongly differentiated responses to the local soil moisture gradient, and that relative to other environmental variables, this would translate into large numbers of negative correlations between species. To evaluate the overall influence of the environment on patterns of species co-occurrence, we further examined the residual patterns of co-occurrence in an LVM including multiple environmental predictors in a single model. If the environment is the dominant force structuring patterns of negative species co-occurrence, we expected weak residual correlations induced by the latent variables. Conversely, if any unmeasured environmental factors and/or biotic processes are equally, or more, important than measured environmental factors, we expected strong correlations based on the latent variables.

Given that observations of species distributions at fine spatial grains are known to be sensitive to stochastic processes (Chase, 2014), we performed our primary analysis on all plots from all censuses combined, whilst accounting for spatial and temporal non-independence in a hierarchical framework. This aggregation of census data collected over a long period of time not only served to minimize the masking affect of stochastic phenomena on underlying species-environment associations, but also to reduce measurement error that might arise from missed detections in any given single survey. Nevertheless, to aid interpretation and disentangle spatial and temporal processes, we additionally examined species

species-specific responses at each individual census, as well as at a finer spatial scale to the main analysis.

5.3 Materials and methods

5.3.1 Study area and floristic sampling

Sampling was conducted in an area (approx. 4 ha) of fire-prone coastal heathland in Royal National Park, New South Wales, Australia. The herbaceous ground layer is dominated by species within the Restionaceae, Cyperaceae and Poaceae families, while the overstorey consists mainly of sclerophyllous shrubs within the Proteaceae, Myrtaceae, Ericaceae and Fabaceae families (Keith *et al.*, 2007). The soils derive from a sandstone substrate, and tend to be low in nutrients, sandy and acidic. The topography is relatively flat, with minimum and maximum elevations of 67 and 71 metres respectively. A previous study found that terrain-based hydrological models provided poor estimates of true soil moisture at the site (Holman, 1999), as is typical of low relief areas (Anderson & Kneale, 1980).

Fifty-six 0.25 m² plots within the study area were sampled ten times over 24 years (1990–1994, 1999, 2002, 2007, 2011 and 2014). At each census, the abundance (number of stems) of all herbaceous species in each plot was recorded. The plots are spatially clustered into four groups of 14, with a mean distance of 211 metres between each cluster (min = 104 m; max = 323 m). Within each cluster, plots are spaced a mean distance of 18 metres apart (min = 5 m, max = 45 m). A fire in October 1988 burnt the entire site prior to the commencement of the study, with subsequent fires in January 1994 (all plots burnt) and January 2001 (14 plots in one cluster burnt).

We restricted our analysis to the most dominant species, where dominance was defined as those species present in at least 20% of all 560 observations (56 plots × 10 censuses). Out of a total of 49 species, 12 met this criterion. The rationale for only considering the dominant species was both ecological and methodological. Firstly, we asserted *a priori* that for any given environmental variable to be considered an important niche axis in the community, at least two of the dominant members should exhibit distinct responses to that variable. Secondly, the model-based framework we employed precluded the analysis of rare species for which too few data points were available.

5.3.2 Environmental sampling

In March-May 2013, 56 100 cm access tubes for a PR2/6 soil moisture profile probe (Delta-T Devices: <http://www.delta-t.co.uk>) were installed within 10 cm of each plot. The probe is fitted with six sensors allowing for instantaneous soil moisture measurements at six depths (10/20/30/40/60/100 cm). After a two month settling-in period, soil moisture sampling commenced in early August 2013 and was repeated at monthly intervals until July 2014. At each monthly interval, three readings (x six depths) were taken at each plot, with the probe rotated through 120 degrees between each reading to account for any small scale variability in moisture at each depth.

In addition to present day measurement of soil moisture, a range of edaphic variables known to be important for plant growth and survival have been sampled in the immediate vicinity of each plot on seven occasions spread across the sampling period (1991, 1993, 1994, 1995, 1999, 2002 and 2014). At each plot, 5-10 soil cores (diameter: 27 mm; length = 100 mm) were sampled randomly within 10 cm of the plot boundary. In each given sampling year, the cores from each plot were combined and analysed for a range of variables, including organic carbon (all years), ammonium nitrogen (NH_4) (all years excluding 1991, 1999 and 2014), nitrate nitrogen (NO_3) (all years excluding 1991), total iron (all years excluding 1991, 1995 and 2002), phosphorous (all years excluding 1991, 1995 and 2002), conductivity (all years), pH (all years), cation-exchange capacity (CEC) (all years) and various exchangeable cations (Na, K, Ca, Mg & Al) (all years).

5.3.3 Data preparation

Prior to inclusion in models, soil variables were evaluated for several criteria including sufficient spatial heterogeneity, relative temporal stability (a prerequisite for a spatial niche axis), and the absence of outliers. Both nitrate and exchangeable K were at such low levels that very little variability in their concentration could be detected, and were consequently excluded from the analysis. Of the remaining variables, nitrogen (ammonium), conductivity, exchangeable Na and exchangeable Ca were found to exhibit substantial temporal variability with rapid temporal decay in their spatial correlation (mean Spearman correlation coefficient between sample years = 0.23, 0.17, 0.20 & 0.32 respectively) and so were also excluded. The mean value was then calculated for each of the remaining variables at each plot, following which outliers (never more than two) were excluded for

several variables to avoid them having a disproportional influence on parameter estimates. The final set of variables included phosphorous, organic carbon, total iron, pH, cation exchange capacity, exchangeable Mg and exchangeable Al, as well as two measures of soil moisture detailed further below.

In the absence of between year samples of soil moisture at each plot, the temporal stability of soil moisture at each depth was evaluated on the basis of the variability observed over the 12 month sampling period. The vertical profile (100, 200, 300, 400, 600 and 1,000 mm) for soil moisture exhibited a gradual decay in correlation such that values at 100 mm were near independent of those at 1,000 mm ($r = 0.09$), while those at intermediate depth fell along a spectrum between these two extremes. As a result, we subsequently restricted further evaluation of soil moisture to these two depths, one representing 'shallow' near surface soil moisture (100 mm) and one representing 'deep' soil moisture (1,000 mm). To evaluate the temporal stability of soil moisture, both in terms of relative spatial structure and absolute levels, we considered the mean Spearman rank correlation coefficient of all pairwise monthly comparisons (relative stability), and the overall standard deviation in mean monthly soil moisture (absolute stability). Deep soil moisture was highly stable in both relative and absolute terms, with mean pairwise correlation between months, $r = 0.96$, and standard deviation in average monthly soil moisture = 1.017. Shallow soil moisture was only slightly less stable with mean pairwise correlation between months = 0.86, and standard deviation in average monthly soil moisture = 4.7. The high Spearman rank coefficients observed at both depths indicate high temporal stability in the relative spatial structure of soil moisture at the study site, whilst the increase in stability with depth is consistent with previous work (Cassel *et al.*, 2000). As such there is strong evidence that the spatial structure of soil moisture observed over the 12 month monitoring period is typical of that which would have been observed throughout the floristic sampling period. Furthermore, variability in monthly rainfall over the 12 month monitoring period ($sd = 57.3$) was near that observed for the entire period since floristic surveys began in 1990 ($sd = 78.4$). The final per plot measurements of deep and shallow soil moisture were obtained by averaging across all 12 monthly samples. Notably, average shallow soil moisture was unrelated to elevation ($p = 0.832$), while deep soil moisture was weakly, and positively related ($R^2 = 0.1031 p = 0.0179$), indicating a very slight increase in deep soil moisture with increasing elevation.

Finally, to obtain an estimate of light availability in the final year of sampling, bottom-up canopy photographs were taken over two days in March 2014. At

each plot, two replicate photographs were taken using an SLR camera with a fish-eye lens positioned facing directly upwards on a level surface in an east-west orientation. The images were then digitally converted to black and white based on an automatic threshold using the GIMP graphics editor (www.gimp.org), and the ratio of black to white pixels (representing vegetation cover) was quantified. The two counts for each plot were then averaged.

5.3.4 Model design

We used a novel approach to analyzing co-occurrence patterns based on latent variable models (LVM). LVMs can be regarded as an extension of factor analysis (Knott & Bartholomew, 1999) to non-normally distributed responses. For each of the nine environmental predictors (excl. light availability), we ran three different LVMs: one leveraging all plots and censuses combined (full model); one for each of the four clusters of plots (individual cluster model) to infer to what extent patterns observed for the full dataset were driven by variation within or between clusters; and one for each of the ten individual censuses (individual census model) to evaluate the temporal stability of pairwise co-occurrence patterns. Additionally, we fitted an LVM for the full dataset which included multiple environmental predictors (excl. light availability) in order to identify remaining patterns of co-occurrence after accounting for the environment. Due to collinearity ($r > 0.7$) amongst several variables, three variables (pH, CEC and Exchangeable Mg) were left out of the multi-predictor model. We also fitted a single LVM for species co-occurrence as a function of light availability for 2014 only. For the individual cluster models, we restricted our analysis to those species that occurred in at least 20% of all plots in each cluster over the 10 censuses (as was done for the full model). For the individual census point models, the criterion was relaxed to 10% of all plots at each time-step. A schematic summarizing the relationship between the different models and associated analyses (detailed further below) is provided in Fig. 5.1.

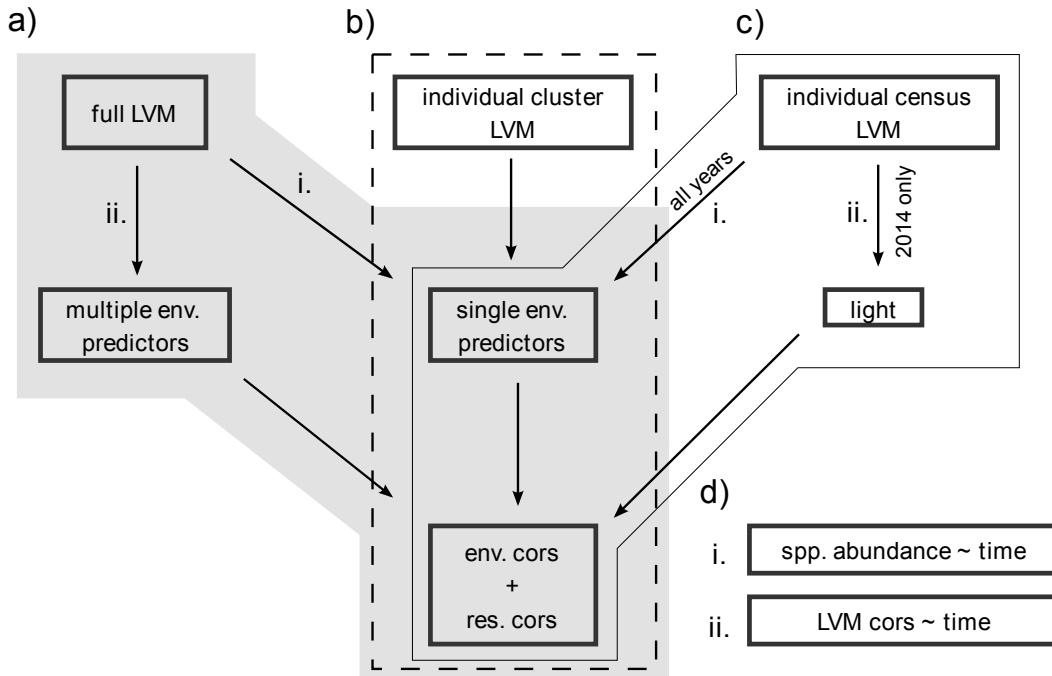


Fig. 5.1 Schematic summary of core analysis. Environmental and residual correlations calculated for: a) LVMs fitted to the full dataset for i) each of the nine environmental predictors independently, and ii) for multiple predictors in a single model; b) LVMs fitted to each individual cluster for each predictor independently; and c) i) LVMs fitted to each census for each predictor independently (all census years), and ii) additionally for light in 2014. Supplementary GAMMS were fitted for: d) i) species abundance as a function of time for comparison with the residual correlations from the multiple predictor model, and ii) for environmental and residual correlations from the individual census LVMs as a function of time.

Similar to other co-occurrence models such as JSDMs (e.g., Ovaskainen *et al.*, 2010, Clark *et al.*, 2014, Pollock *et al.*, 2014), the extent to which species exhibit distinct environmental responses was inferred through a fitted regression model, specifically using the framework of generalized linear models. However, instead of employing an unstructured covariance matrix to account for the missing environmental covariates and/or biotic interactions (as was done in Ovaskainen *et al.*, 2010, Pollock *et al.*, 2014, for instance), LVMs utilize latent variables as a parsimonious means of modelling residual species correlation (see also Harris, 2015). As such, we use the term ‘residual’ in reference to remaining patterns in the data after accounting for one or more predictors, rather than in terms of its definition in the context of residual analysis.

Our model can be regarded as an extension of the LVM proposed for model-based unconstrained ordination in Hui *et al.* (2014), and can be written in the following hierarchical form:

$$\begin{aligned}
 \text{Responses: } & [y_{ij} | \mathbf{u}_i, \mathbf{x}_i] \sim \text{Neg-Bin}(y_{ij}; \mu_{ij}, \phi_j) \\
 & \log(\mu_{ij}) = \eta_{ij} = \mathbf{x}'_i \boldsymbol{\beta}_j + \mathbf{u}'_i \boldsymbol{\lambda}_j \\
 \text{Latent Variables: } & [\mathbf{u}_i] \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\
 \text{Priors: } & [\boldsymbol{\beta}_j] \sim \mathcal{N}(\mathbf{0}, c_0 \mathbf{I}), [\boldsymbol{\lambda}_j] \sim \mathcal{N}(\mathbf{0}, c_0 \mathbf{I}), [\phi_j] \sim \text{Unif}(0, c_1),
 \end{aligned} \tag{5.1}$$

where ‘~’ denotes “is distributed as”, $\mathcal{N}(\cdot, \cdot)$ denotes a multivariate normal distribution with mean and covariance matrix given by the first and second arguments respectively, $\text{Unif}(0, c_1)$ denotes a uniform distribution with minimum zero and maximum c_1 , and \mathbf{I} denotes an identity matrix. To elaborate, y_{ij} denotes the observed count for species j at site i , and is assumed to come from a negative binomial (Neg-Bin) distribution with mean μ_{ij} and species-specific overdispersion parameter ϕ_j . We used a negative binomial distribution as it exhibits a quadratic mean variance relationship, $\text{Var}(y_{ij}) = \mu_{ij} + \phi_j \mu_{ij}^2$, which allowed us to explicitly account for overdispersion present in the species counts.

Using a log-link function, the mean μ_{ij} was regressed against two sets of variables. Firstly, a vector of explanatory variables for each site, \mathbf{x}_i , which included an intercept term, environmental covariates (e.g. soil moisture), a fixed effect for the four clusters of plots to account for spatial non-independence within clusters (full and individual year models), and a random effect (intercept) for plot to account for temporal non-independence (full and individual cluster models). In order to account for unimodal species responses as predicted by niche theory (Austin, 2002), quadratic terms were included for the environmental covariates in each model. The second set of variables comprised a vector of two latent variables for each site, $\mathbf{u}_i = (u_{i1}, u_{i2})$, which were assumed to be drawn from independent, standard normal distributions. The species-specific regression coefficients $\boldsymbol{\beta}_j$ and $\boldsymbol{\lambda}_j$ describe how the mean changes as a function of the explanatory and latent variables respectively.

The key element in the LVM in equation (5.1) are the latent variables \mathbf{u}_i , which can account for any residual correlation between species not attributable to spatial heterogeneity in the measured environmental covariates \mathbf{x}_i . This correlation may be driven by biotic interactions such as competition (negative) or facilitation

(positive), or alternatively to missing predictors, where λ_j are the coefficients corresponding to these missing predictors. In addition, in our full model which combines multiple censuses, residual correlation may arise due to negatively or positively correlated fluctuations in species abundance through time. If unaccounted for, missing covariates, ecological interactions or temporal correlation will induce residual correlation between species which can potentially lead to erroneous inference. Latent variables offer an attractive approach to dealing with this issue. In particular, they require significantly fewer parameters to model species residual correlation compared to the unstructured (full rank) correlation matrices used by Ovaskainen *et al.* (2010) and Pollock *et al.* (2014). Finally, note that more than two latent variables could have been used, although our preliminary testing with this dataset suggested that two was sufficient to characterize the main patterns of species residual correlations.

We used Bayesian Markov Chain Monte Carlo (MCMC) methods to estimate the LVMs, with sampling performed through JAGS v3.4.0 (Plummer, 2003) using the package R2jags v0.03-08 (Su & Yajima, 2012) in R v3.1.1. We assigned uninformative priors for all parameters, β_j , λ_j and ϕ_j , by setting $c_0 = c_1 = 100$ in equation (5.1). For the full and individual cluster models, we also assigned an uninformative uniform prior for the variances of the random effect for plot. For each LVM fitted, we ran three chains with a burnin period of 10,000 iterations followed by 100,000 iterations with a thinning lag of 50 for each chain. This produced a final combined sample of 6000 MCMC samples for each LVM. We assessed parameters to have converged when traceplots were well mixed and the Gelman-Rubin statistic was below 1.1.

R code demonstrating the fitting and analysis of latent variable models of co-occurrence is provided in Appendix D.

Model inference and supplementary analysis

After fitting the LVMs, in order to visualize patterns of co-occurrence arising from the different environmental factors, we calculated two types of correlation matrices. The first was constructed by calculating, for any two species, the correlation between their fitted values $x_i' \beta_j$ (across all plots). This is the same as equation (4) in Pollock *et al.* (2014), with the resulting matrix representing the correlation between species that can be attributed to a shared/diverging environmental response. The second type of correlation matrix was calculated using the latent variable coefficients, λ_j , also known as factor loadings. Specifically, if

we let Λ be the two-column matrix formed by stacking the factor loadings on top of each other, then a covariance matrix is obtained as $\Lambda\Lambda'$ from which the residual correlation matrix can be calculated. This second residual correlation matrix represents the correlation between species that may be attributable to biotic interactions or missing environmental covariates. Since Bayesian MCMC estimation was used, the correlation between fitted responses was calculated for each MCMC sample, which made it possible to obtain a posterior distribution for each cell of the environmental and residual correlation matrix. As such, correlation ‘significance’ was evaluated on the basis of the 95% credible intervals for the posterior mean excluding zero.

Having extracted environmental and residual correlation matrices for each environmental factor in each of the three model types (full, individual cluster, individual census point), we assessed the importance of each factor as a niche axis on the basis of two criteria: i) the number of significant negative pairwise environmental responses, where significance is defined as 95% credible intervals that don’t cross zero; and ii) the minimum number of species that need to be excluded to remove all negative pairwise responses, referred to in network theory as the minimum vertex cover. In this application, the minimum vertex cover quantifies the extent to which a given number n of negative pairwise environmental responses is driven by one species exhibiting a different response to all other species, or at the other extreme consists of completely unique species pairs. In the former case, only one species need be excluded to remove all negative correlations whereas in the latter case n species must be removed. As such, with 12 species, if all 66 unique pairs exhibit significant negative correlations (an implausible scenario necessitating extremely narrow niche breadths), 11 would need to be excluded. In addition to quantifying negative pairwise environmental responses, we also counted the number of positive (shared) environmental responses to evaluate the extent to which a given environmental factor influences the fine-scale distribution of species even in the absence of any clear niche differentiation (e.g. if all species exhibit a similar response to the gradient).

After considering species-specific environmental responses, we then examined the residual correlation matrix for each model, with particular emphasis on the full model including all of the environmental predictors, as this represents patterns of co-occurrence unexplained by all our measured predictors combined. To aid interpretation of the residual correlation matrix from the full multi-predictor model, we also fitted a generalized additive mixed model (GAMMs) for each species as a function of time, and compared the correlation in the fitted values

from the GAMMs with the residual correlation matrix based on the full LVM model including multiple environmental predictors. The temporal GAMMs for each species were fitted using the MGCV package (Wood, 2012), with a thin-plate spline smoother for time, a fixed effect for plot cluster, a random effect for plot, an AR(1) correlation structure nested within plot, and assuming a negative binomial error distribution with log link. If patterns in the residual correlation matrix of the full model are largely a product of temporal correlation (negative or positive) in species responses, we expected a strong correlation with the fitted value correlation matrix derived from the individual GAMMs. Correlation between the two matrices was assessed via a Mantel test. Note that the GAMM for one species, *Empodisma minus*, failed to converge (most likely due to overdispersion), and as a result had to be excluded from the temporal analyses.

To investigate temporal trends in environmental and residual correlations, we also fitted GAMMs for mean pairwise environmental and residual correlation as a function of time for all variables combined. Each model included a thin-plate spline smoother for time, a random effect for environmental variable and an AR(1) correlation structure nested within environmental variable. Owing to the mean correlations values being consistently positive and right skewed, we used a gamma error distribution with a log link function. To correct over-smoothing observed in the residual correlation trends model, the maximum number of degrees of freedom was reduced from 9 (package default) to 4.

Finally, we performed *post-hoc* exploratory analysis of the differences in species response to soil moisture with their pairwise differences in two readily available aboveground traits, plant height and seed mass. In each case, we used a Mantel test to compare pairwise difference in soil moisture ($-1 \times$ environmental correlation) with pairwise trait differences, and visualised the relationship with scatterplots.

5.4 Results

Of the nine environmental variables considered independently in each of the single-predictor models for the full dataset, all but pH produced significant negative correlations due to diverging species-specific responses (Fig. 5.2). Deep soil-moisture was associated with both the largest number of negative correlations (18) and the largest number of node removals required to eliminate all negative correlations (4). As is to be expected given such strong correlations, the

12 species displayed a range of differentiated responses to the deep soil-moisture gradient, with some increasing monotonically (or near monotonically) towards the extremes of the gradient (e.g. *Xanthorrhoea resinosa*, *Haemodorum corymbosum* and *Entolasia stricta*), whilst others exhibited unimodal peaks closer to the middle of the gradient (e.g. *Ptilothrix deusta*, *Empodium minus* and *Xyris gracilis*) (Fig. 5.3). Exchangeable aluminium was associated with the second largest number of negative correlations (17), followed by organic carbon (11), shallow soil-moisture and total iron (10), exchangeable magnesium (6), cation exchange capacity (3), and phosphorous (1) (Fig. 5.2). Whilst pH was not associated with any negative pairwise correlations, 27 species pairs exhibited significant positive correlations to pH. Deep soil-moisture was associated with the second highest number of significant positive correlations, followed by exchangeable aluminum and exchangeable magnesium (21), organic carbon (18), shallow soil-moisture (17), total iron (16), cation exchange capacity (12), and phosphorous (8) (see Fig. S5.1 in Appendix D).

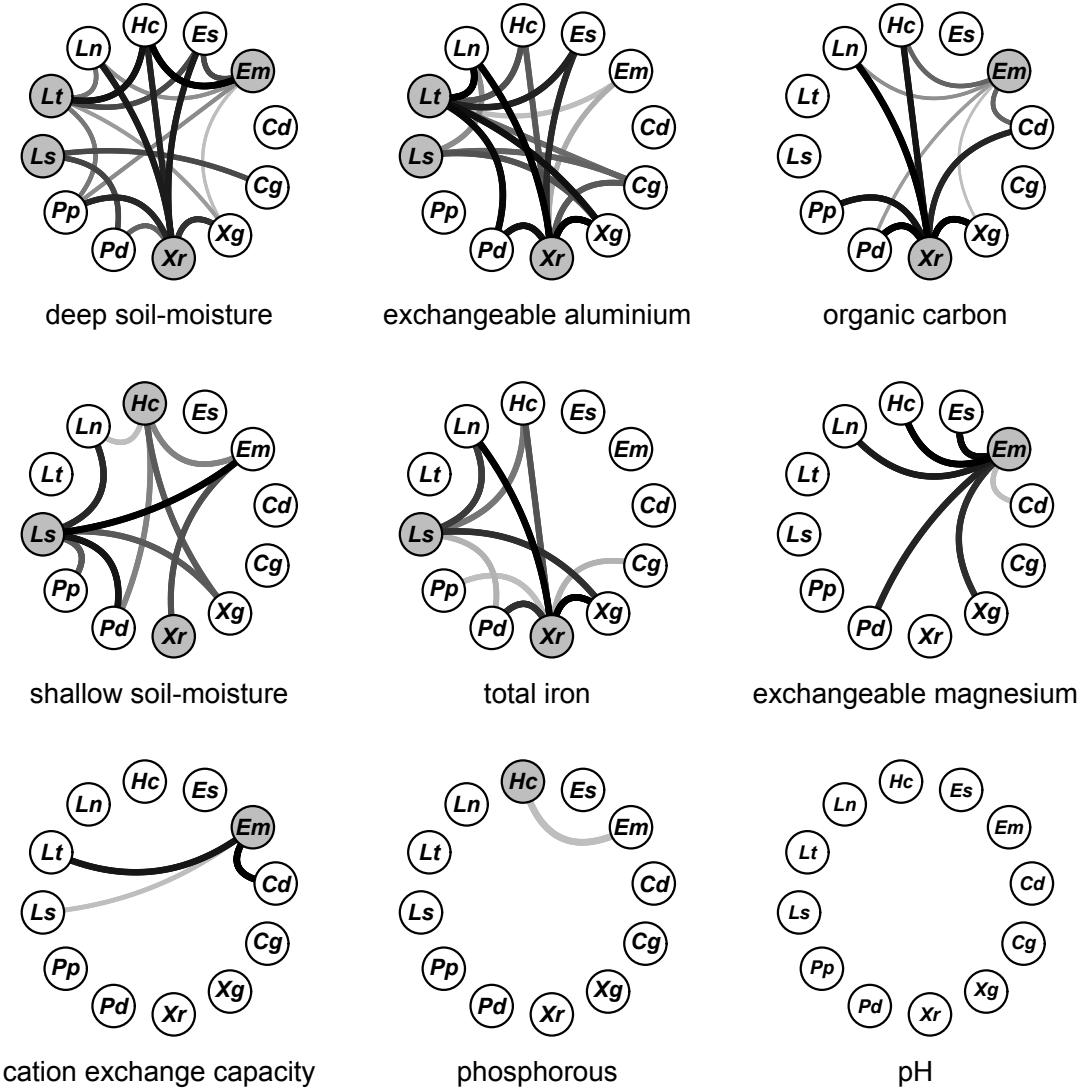


Fig. 5.2 Negative pairwise species correlations to the environment derived from single-predictor LVMs for the full dataset. Connecting lines between species nodes denote negative mean posterior correlations with credible intervals excluding zero. Line colour and thickness indicates the strength of the negative correlation where darker and thicker lines are closer to -1. Nodes shaded grey indicate the minimum vertex cover, i.e. the smallest combination of species that need to be removed to break all negative associations (note that while the minimum number has a single solution, the species composition making up the minimum vertex cover set can vary). Species node labels combine the first letters of the genus and specific epithet given in full in Fig. 5.3.

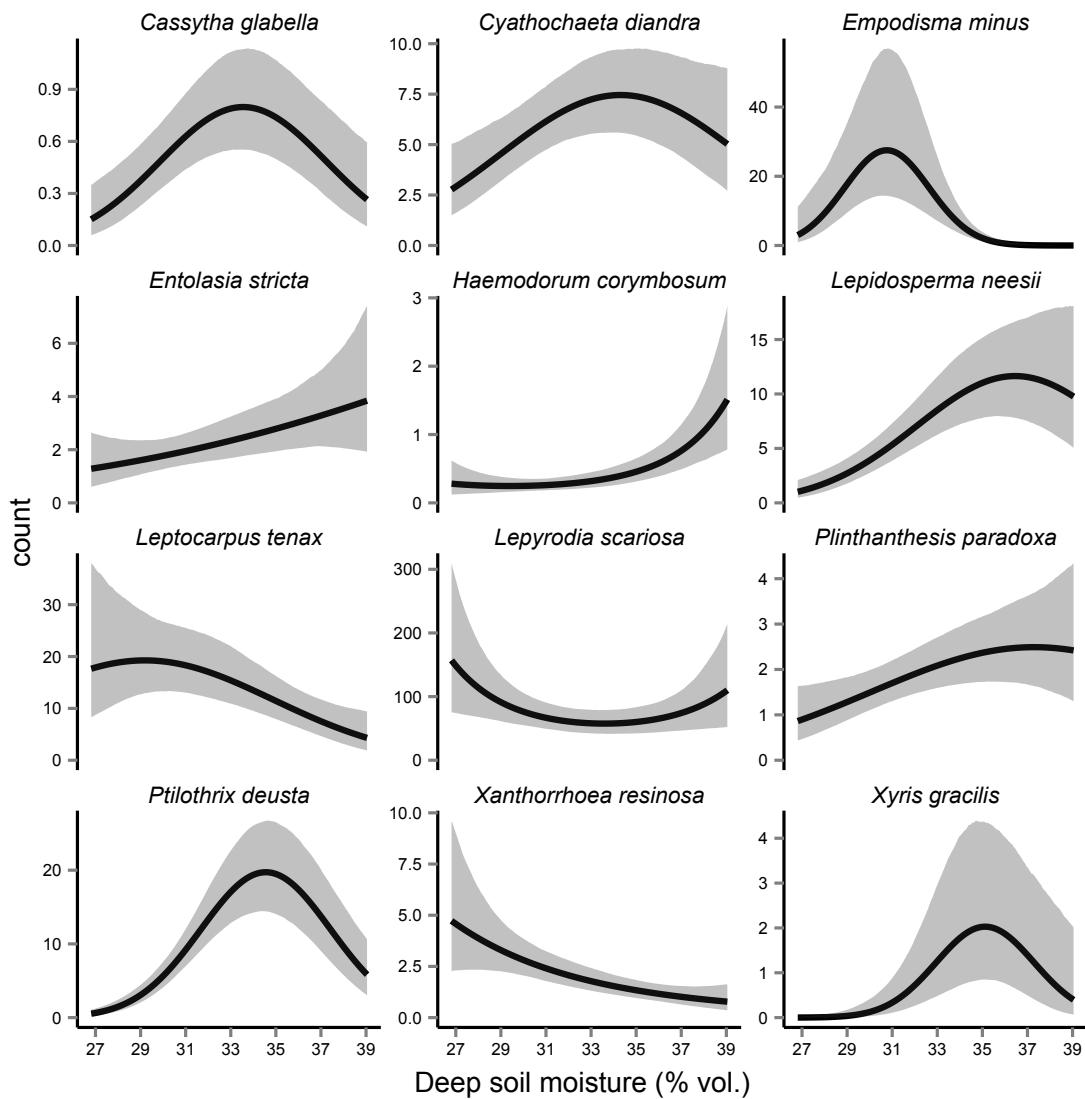


Fig. 5.3 Species specific responses to the deep soil-moisture gradient. Fitted values derived from regression parameter estimates from the deep soil-moisture LVM (full dataset). Shaded region denotes 95% credible intervals.

Residual correlations based on single-predictor LVMs fitted to the full dataset were strongly associated with each other, and with residual correlations from the multi-predictor model ($r = 0.68\text{--}0.97$). As such, the consistency of the observed residual correlations suggests a potentially important variable was not accounted for in any of the LVMs. Nevertheless, most of the significant correlations in the residual correlation matrix of the multi-predictor model were positive (23) rather than negative (2), thus negating the probability that an important fine-scale niche axis was omitted from our analysis (Fig. 5.4). Furthermore, a substantial portion of the residual correlation appears to be attributable to shared and differentiated temporal responses between species, with the residual correlation matrix from the

multi-predictor model being moderately associated with the temporal response correlation matrix derived from the temporal models (Mantel test $r = 0.43$) (see Fig. S5.2 for smoothed species responses through time).

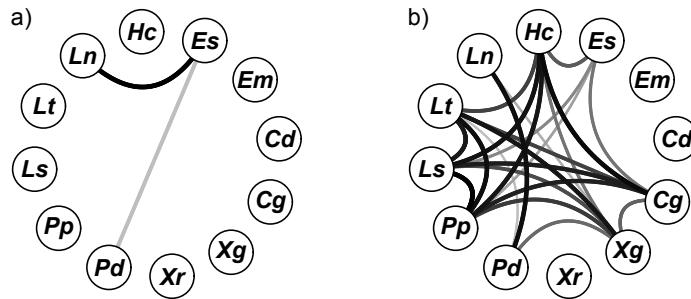


Fig. 5.4 Negative (a) and positive (b) pairwise residual correlations derived from the multi-predictor LVM for the full dataset. Connecting lines between species nodes denote posterior correlations with credible intervals excluding zero. Line colour and thickness indicates the strength of the correlation where darker and thicker lines are closer to $|1|$. Species node labels combine the first letters of the genus and specific epiphets.

Relative to the models run on the full dataset, the number of significant negative correlations due to diverging environmental responses between species was considerably smaller for models run on each individual cluster (plots separated by 5–45 metres). As such, most of the observed patterns of co-occurrence in the full models appears to be driven by environmental variation over spatial scales of 10s of metres, rather than metres (i.e. between clusters rather than within clusters). The one exception to this pattern was for phosphorous, which exhibited a relatively high number of significant negative co-occurrence associations in three of the four clusters. Notably, a substantial number of species did exhibit negative pairwise associations in the residual correlation matrices for at least two of the clusters, suggesting that biotic interactions may still influence co-occurrence at within-cluster scales (see Table S1).

The number of significant negative pairwise associations in species environmental response was also comparatively small for the individual census models. It is notable however that even if we ignore ‘significance’ criteria, the sign of each pairwise species association is largely consistent with those observed for the single predictor models for the full dataset. In addition, the models for each individual census provide insight into the relative dominance of positive and negative interactions through time. Whilst there were no obvious trends in

mean environmental correlation through time, residual correlations underwent a conspicuous upwards positive trend in the second half of the sampling period (Fig. 5.5). Pairwise species associations derived from the single LVM fitted for light availability in 2014 were almost universally positive, with all but one species (*Xanthorrhoea resinosa*) being more abundant in less shaded locations.

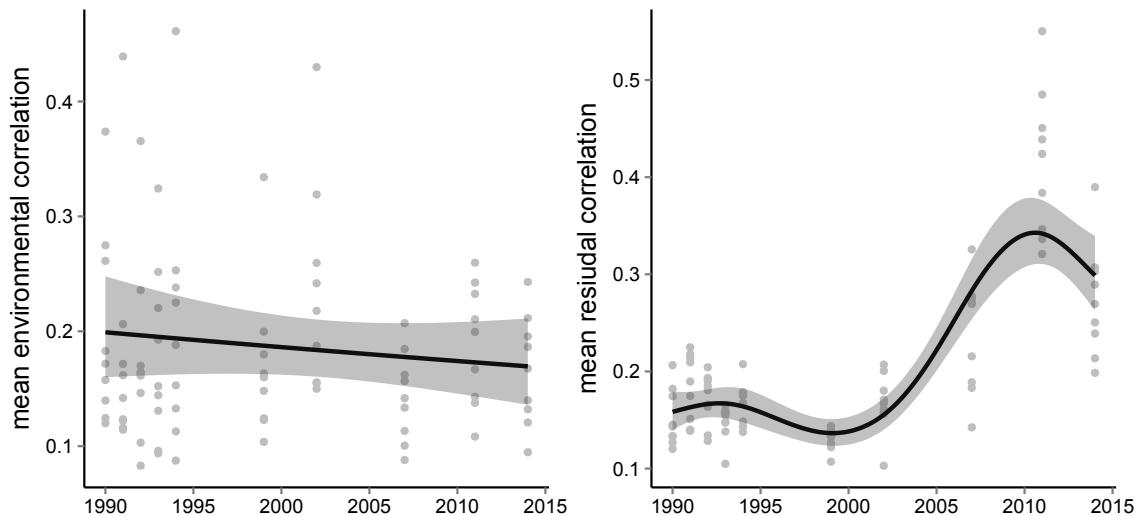


Fig. 5.5 Trends in mean environmental (left) and residual (right) correlation through time. Trend lines obtained with thin-plate splines in a GAMM framework (see main text for model fitting). Shaded region represents 95% confidence intervals.

Differences in species pairwise responses to deep soil moisture were moderately positively related to pairwise difference in plant height ($r = 0.6259, P = 0.002$) but unrelated to differences in seed mass ($r = -0.1176, P = 0.809$) (Fig. 5.6).

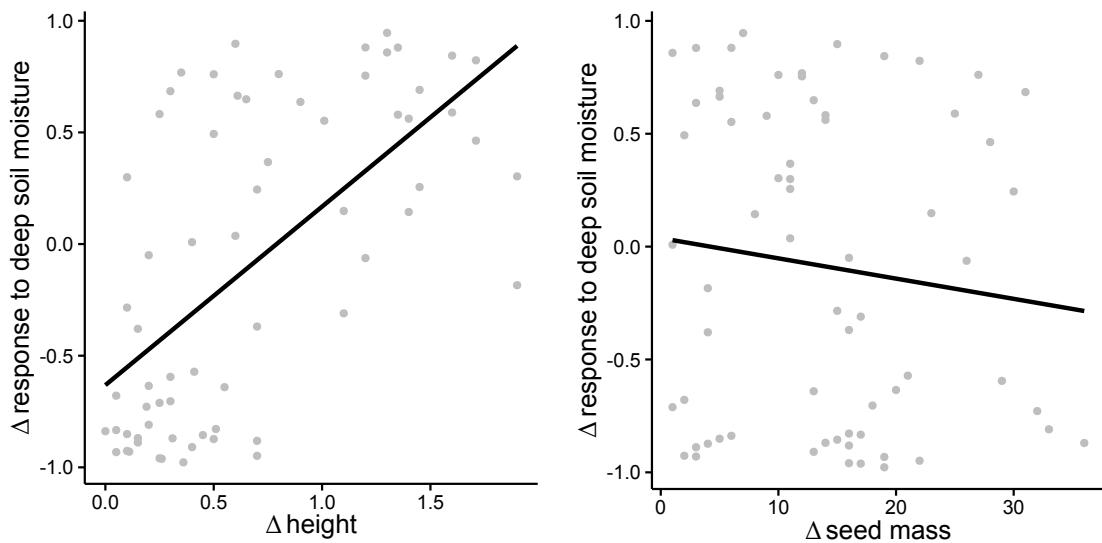


Fig. 5.6 Scatterplots of pairwise differences in species responses to deep soil moisture ($-1 \times$ environmental correlation) in relation to pairwise difference in plant height (left) and seed mass (right). Lines of best fit obtained from linear models.

5.5 Discussion

The notion that fine-scale spatial environmental variability may foster plant species coexistence is founded upon a rich body of theory (Chesson, 2000b, Amarasekare, 2003, Adler *et al.*, 2013a), and yet empirical efforts to establish which, if any, environmental factors are associated with strong patterns of fine-scale niche differentiation are surprisingly sparse. We found strong empirical support for the hypothesis that differentiation along hydrological gradients represents one of the most important fine-scale plant niche axes at within-community scales (*sensu* Silvertown *et al.*, 2015). On the basis of multiple criteria, deep soil moisture emerged as the single best predictor of negative co-occurrence patterns (Fig. 5.2), with the dominant community members exhibiting strongly differentiated responses across a comparatively narrow moisture gradient (Fig. 5.3). Together with several earlier studies (e.g. Silvertown *et al.*, 1999, Araya *et al.*, 2011), these results are amongst the first to demonstrate hydrological niche differentiation in the absence of significant topographical complexity. Furthermore, they highlight the robustness of the hypothesis when confronted with a comprehensive set of alternative prospective niche axes.

The adoption of a latent variable modelling framework enabled us to not only evaluate the comparative importance of each environmental factor, but also to determine the likelihood that an important factor was unaccounted for. Given that

the residual correlation matrix from the multi-predictor model was dominated by positive pairwise associations, we can be confident that we accounted for the most important factors driving niche segregation amongst the dominant community members (Fig. 5.4). As such, our approach allows robust conclusions about the relative importance of soil hydrology not only with respect to measured variables but also unmeasured variables. Notably, a considerable portion of the residual correlation from the multi-predictor model appeared to be attributable to both positive, and to a lesser extent negative, temporal correlations in species peak abundance. These in turn appear to largely reflect differential interspecific temporal responses to fire events in 1988, 1994 and 2001 (Fig. S5.2). As such, with time-series data, co-occurrence models may also be used to draw inferences on successional processes and temporal niche partitioning.

While the results provide strong evidence for fine-scale hydrological niche differentiation, the extent to which hydrological niches provide opportunities for local-scale species coexistence is to a large extent contingent on the scale at which the community is defined (Amarasekare, 2003, Adler *et al.*, 2013a). For instance, in the current study, negative co-occurrence patterns associated with the deep soil moisture gradient were only detectable when comparing plots across clusters separated by 100–320 metres. Similarly, Silvertown *et al.* (1999) made what is probably the strongest case for hydrological niches to date on the basis of plots dispersed across 10's of hectares, whilst even the ca. 50 × 50 metre plots of Araya *et al.* (2011) presumably allowed for individuals to be spaced as much as 70 metres apart. In contrast, Fridley *et al.* (2011) reported species-specific responses to microsite variation in soil depth in 3 × 3 metre plots in a limestone grassland, but this appeared to be unrelated to soil water potential. As such, the evidence for hydrological niches at the local scale within which neighbouring plants directly compete for resources arguably remains relatively weak. This conclusion would be in line with the dominant paradigm in community ecology that negative patterns of co-occurrence at the local scale are predominately indicative of competitive exclusion (Gotelli & McCabe, 2002, but see Fridley *et al.* 2011). Indeed, only phosphorous exhibited more than six negative pairwise associations (compared with just one for the full dataset) in any one plot cluster (Table S5.1). In contrast, the residual correlation matrices for two plot clusters exhibited a relatively large number of negative associations, which do not appear to be driven by any of the measured environmental factors. Nevertheless, even in the absence of localised spatial partitioning of resources, hydrological niches may still promote local coexistence via source-sink dynamics and spatial storage effects (Shmida & Ellner, 1984, Amarasekare, 2003, Snyder & Chesson, 2004). This

will arise when species are able to disperse from favourable to less favourable locations at a sufficiently high rate to offset locally negative population growth rates and thus prevent local extinction (Snyder & Chesson, 2004). Since monitoring began in the current study, all but one species has been recorded at all four clusters, suggesting that most of the dominant community members do indeed regularly disperse to and establish in locations where they are less competitive. It is also notable, that within short distances (<100 m) of our study plots, the topography is more varied, and as such may provide nearby opportunities for even finer scale hydrological niche differentiation.

In contrast to the individual plot cluster models, the lack of strong negative environmental correlations in the individual census models appears to be an artefact of low power, particularly given that the raw correlation values corresponded closely with those from the full model. Nevertheless, it is interesting to note that whereas mean environmental correlations were relatively static through time, residuals correlations exhibited a strong upwards trend towards more positive values (Fig. 5.5). This trend most likely reflects a shared environmental response to increased overstorey shading, as indicated by the observed positive response of the majority of species to light availability in the most recent 2014 census. In a previous study from the same site, Letten *et al.* (2014) described an increase in functional and phylogenetic similarity of community members through temporal succession consistent with this observation.

Our focus was on negative environmental correlations, but positive correlations also provide insight into other aspects of community assembly. For instance, positive correlations in species responses were highest for soil pH, suggesting pH may be a useful predictor of local richness patterns, whilst also acting as niche axis over larger spatial scales. It is also notable that although this system is typically associated with low-pH adapted species, many of the dominant herbaceous community members exhibited a positive response to the pH gradient (data not shown). As such, it would be interesting to explore the extent to which rarer species exploit this apparently ‘unoccupied’ niche space within the landscape, particularly given the known importance of soil pH for plant community structure in other systems (Laliberté *et al.*, 2014).

Although a comprehensive examination of trait-environment relationships was beyond the scope of this study, comparisons of species pairwise differences in seed mass and plant height relative to hydrological niches differences still proved informative. Whereas differences in seed mass, which may be associated with survival from drought stress (Pérez-Harguindeguy *et al.*, 2013), were unrelated to

hydrological niches, differences in plant height were strongly positively related with differences in species response to soil moisture (Fig. 5.6; cf. Fig 2c in Adler *et al.* 2013). This appeared to be driven predominately by taller species (e.g. *Leptocarpus tenax* and *Xanthorrhoea resinosa*) favouring drier areas of the deep soil moisture gradient, while shorter species (e.g. *Plinthanthesis paradoxa* and *Xyris gracilis*) were more abundant in wetter areas. This we expect reflects the greater tolerance of shorter species to anaerobic conditions in periodically saturated sites. Given that soil moisture measurements at all depths were typically around, or above, the field capacity of sandy-loam soil ($\sim 23\%$ vol., Atwell *et al.*, 1999), this possibly reflects the greater tolerance of shorter species to low oxygen in periodically saturated sites. In addition it suggests fine scale variability in soil moisture may have important implications for the structural complexity of standing vegetation.

Beyond differences in plant height and seed mass, it is notable that the dominant herbaceous members of the community actually exhibit remarkably similar phenotypes, with all but one of the twelve study species being a geophytic or hemicryptophytic monocot, and nine of the twelve belonging to just three families within the poales (*Restionaceae*, *Poaceae* and *Cyperaceae*). This apparent constraint on functional strategies is likely attributable to the fire-prone and nutrient poor nature of the system, which makes it all the more remarkable that these same species exhibit such differentiated responses to the soil moisture gradient. One likely explanation is that these species exhibit a range of different below-ground adaptations to water acquisition and flooding/drought tolerance (Silvertown *et al.*, 2015).

An open question arising from our results is why species co-occurrence patterns should exhibit a stronger relationship with deep, rather than shallow, soil moisture. This is particularly true given that most of the study species are likely to have the bulk of their rooting volume above 1,000 mm where the deep soil moisture measurements were made. The most parsimonious explanation is that shallow soil moisture is unevenly modified by plant water usage and episodic rainfall recharge, and thus provides a poorer approximation of baseline soil moisture across the study area. More specifically, shallow soil moisture is likely to experience greater seasonal and interannual variability contingent on the density of standing vegetation. Observed trends in monthly soil moisture support this notion, with shallow soil moisture exhibiting a relatively strong seasonal trend over the 12 months of sampling, which appeared to be mostly attributable to increased plant water usage over spring/summer rather than decreased rainfall.

In contrast, deep soil moisture was comparatively stable. This may explain the apparent lack of a strong relationship between species specific responses and near surface soil moisture observed by (Fridley *et al.*, 2012). Ultimately, the depth of soil moisture that best predicts hydrological niches is likely to depend on numerous system-specific factors, but our findings highlight the potentially confounding role of plant water usage on inferences drawn from near surface measurements. Furthermore, it is notable that elevation was a poor predictor of soil moisture at any depth, confirming that the use of terrain-based hydrological models in plant niche studies may lead to misleading conclusions.

One challenge to interpreting plant compositional responses to spatial variability in soil moisture is the dual role of water as both a depletable resource factor (cf. nutrients or light) subject to density-dependent feedbacks and as a non-resource factor (cf. temperature or soil type) that mediates density-independent responses (e.g. via reduced oxygen diffusion). It follows that depending on supply rates, spatial heterogeneity in soil moisture may produce complex interactions between niche and fitness differences amongst interacting species (*sensu* Chesson, 2000b). It might therefore be expected that the breadth of plant responses to soil moisture will allow for higher-dimensional trade-offs relative to more narrowly defined environmental factors (Tilman, 1982, Chase & Leibold, 2003). In the current study, soil moisture appeared to be consistently around or above the likely field capacity of the soil. This suggests that in this system soil moisture may exert a stronger influence on compositional patterns by acting as a stressor rather than a spatially heterogeneous limiting resource. In the absence of data on spatial variation in air-filled porosity such inferences are unfortunately somewhat speculative, but this represents an important consideration for future studies.

It is important to recognise that demonstrating species-specific environmental responses along a given gradient, is not on its own sufficient evidence that it facilitates species coexistence. To this end, rigorous tests of species coexistence necessitate experimental manipulations of intra- and inter-specific competition (e.g. Levine & Rees, 2002, Sears & Chesson, 2007b, Godoy *et al.*, 2014), or alternatively the mathematical parametrization of models from long-term demographic studies (e.g. Adler *et al.*, 2006, Angert *et al.*, 2009). Unfortunately these approaches are not always practical, particularly in perennial systems such as heathlands where individuals can take several years to reach reproductive maturity. Nevertheless, as advocated by (Silvertown *et al.*, 2015), providing evidence of fine-scale hydrological niche differentiation is an important first step, that will hopefully provide stimulus for further research.

Here we have provided robust evidence for species-specific responses to fine-scale variation in soil hydrology, whilst also illustrating the valuable inferential gains made possible with model-based approaches to co-occurrence analysis. Our results are consistent with emerging empirical evidence, but go a step further than previous studies by demonstrating the apparent strength of hydrological niche differentiation when compared to a range of other potential niche axes. Future research efforts should focus on identifying the different physiological processes and traits associated with hydrological niches (Silvertown *et al.*, 2015), and on exploring scale dependencies to better understand the underlying coexistence mechanisms at play Adler *et al.* (2013a). In light of perturbations to hydrological regimes under climate change, a better understanding of hydrological niches could prove critical to conservation efforts.

Acknowledgements

For their assistance with fieldwork we are grateful to Chantel Benbow, Sam Dawson, Chris Gordon, Anna Feit, Ben Feit, Sylvia Hay, Mitch Lyons and Max Mallen-Cooper. Thanks also to Martin Westgate whose ‘circleplot’ package was used to produce the network plots in Fig. 1.

Chapter 6

Conclusions

In common with many modern doctoral theses/dissertations, the studies presented in each chapter of this thesis were conducted, and taken through to completion (i.e. publication), in a chronological order, rather than the traditional model where manuscript writing is left till last. An interesting by-product of this approach is that it provides a vantage point to critically examine one's own work from an appreciable distance. For instance, a final draft of Chapter 2 was written-up by late 2012 and accepted for publication in April 2013 (Letten *et al.*, 2013). In the two years that have since passed, my knowledge of the field and general skill-set have, I like to think, expanded. This raises the question, if I had the time again, would I do it differently? In this final chapter, I briefly summarise the contributions of each of the preceding chapters, whilst also considering their limitations, and where relevant the scope for future enquiry.

In Chapter 2, I provided evidence that spatial heterogeneity in temporal climate variability may in some systems be a better predictor of local-scale plant species richness patterns than more commonly used absolute climate measures. Modern coexistence theory (*sensu* Chesson, 2000b) provides an intuitively appealing explanation for this result, and yet perhaps more than any other chapter in this thesis, this study illustrates the tension between the mechanistic insight afforded by experimental manipulation, and the broad scale, but strictly correlative, insight derived from observational studies. Is climate variability a better predictor than absolute climate because it stabilizes coexistence, or is climate variability merely a better proxy for some other factor regulating richness patterns? Unfortunately, no amount of statistical duress will yield a conclusive answer to this question from the dataset at hand. Species richness patterns are a product of a myriad

factors, some deterministic and some stochastic, and thus the extent to which causation may be inferred from correlation will always be less than satisfactory.

In juxtaposition with the low-resolution but broad-scale potential of richness data, data-intensive single-site studies afford opportunities for deeper interrogation of underlying assembly processes. Long-term demographic datasets have proven particularly effective in demonstrating the role of temporal climate variability in annual plant coexistence (Adler *et al.*, 2006, 2009, Angert *et al.*, 2009). An exciting middle ground between these approaches, would be to investigate inter- and intra-specific demographic variability across sites characterised by different levels of temporal climate variability. The analysis of demographic data along environmental gradients has recently been shown to be a powerful method for exploring spatially varying niche relationships (Diez *et al.*, 2014). The logistical challenge in this instance will be locating study sites that exhibit sufficient spatial variation in temporal climate variability but that are close enough together to facilitate regular monitoring. Given these constraints, the topographically heterogeneous terrain typical of alpine grasslands may represent one potential system for such a study. However, it also remains unclear to what extent relatively low intra-seasonal and intra-annual climate variability (as exhibited in sheltered gorges compared to exposed ridges) translates into more buffered inter-annual variability. Given that inter-annual climate variability is likely to play a considerably more important role in species coexistence than intra-annual variability, this assumption warrants further investigation.

The exploration of several largely unexplored assumptions specific to phylogenetic approaches in community ecology formed the basis of both Chapters 3 and 4 of this thesis. More specifically, Chapter 3 highlighted the absence of any theoretical basis to the "*widespread assumption in the literature that phylogenetic and functional distance scale linearly*" (Letten & Cornwell, 2014). In fairness, this is probably less an explicit assumption and more an oversight. Nonetheless, it has consequences for assessing phylogenetic community structure. The simple, but theoretically robust, solution to avoid the over-weighting of early diverged clades is a square root transformation of the phylogenetic distance matrix. As briefly mentioned in the discussion to that chapter, while there are almost certainly more sophisticated approaches available, a square-root transformation is for now the most parsimonious approach to the scaling of evolutionary relatedness and functional distance.

Notwithstanding the importance of appropriate phylogenetic and functional scaling, it is important to keep in mind that functional traits are themselves

proxies for the underlying processes structuring communities. For instance, plant height is commonly used to predict competitive dominance in competition for light (Westoby *et al.*, 2002, McGill *et al.*, 2006). However, it is less clear how these two factors scale. Does a doubling in height confer a doubling in competitive fitness? There has been some compelling recent work examining the relationship between functional traits and demographic vital rates (Adler *et al.*, 2013a) as well as fitness/niche differences (Kraft *et al.*, 2014), but whether any generalisations can be made as to the precise scaling of these relationships remains an unknown. Kraft *et al.* (2015) only tested for linear relationships owing to the apparent absence of strong non-linear relationships in their data but acknowledged that more complex relationships may occur. This deserves to be a prominent focus of future research in trait-based approaches in community ecology.

In Chapter 4, I used standard phylogenetic and functional-trait metrics to evaluate the temporal stability of community structure through succession. Contrary to paradigmatic assumptions relating to both the changing importance of competition through succession and the community-level signature of competitive interactions, community structure did not become increasingly functionally and phylogenetically over-dispersed with time. Again, from the results alone it is unclear to what extent this reflects a diminished role for competition in this system, or rather simply a misconception in how competition structures communities (*sensu* Mayfield & Levine, 2010). Certainly, the circumstantial evidence for the latter has grown rapidly in recent years, with numerous studies re-emphasizing the role of equalizing mechanisms (and not just stabilizing mechanisms) in mediating coexistence. Nevertheless, the observed temporal increase in positive patterns of co-occurrence presented in Chapter 5 (Fig. 5.6) does lend credibility to the notion that the importance of competition may indeed also tail-off late in succession.

Despite being central to two chapters (or perhaps as a consequence of it), my overall enthusiasm for inferring community assembly processes from phylogenetic patterns has waned over the timeline of this thesis. With the possible exception of the most diverse, hard to study, systems, it is difficult not to conclude that the fragility and contingency of the underlying assumptions outweighs the attractive simplicity of the approach. As a brief scan of the recent literature will attest, this is by no means an original perspective. Nevertheless, I would argue that the widespread adoption of community phylogenetic approaches over the last 10–15 years has still been a net gain for the field. Generally, it is only once the uptake of a given method reaches a critical mass that the assumptions come into sharp focus. In the case of community phylogenetics, it seems this critical process

has fostered much broader exposure to fundamental theory amongst empirically minded ecologists, whilst also motivating a timely mini-synthesis of current eco-evolutionary theory. At the same time, it is also important not to lose sight of other motivations for combining phylogenetic and community-level data, e.g. to identify how coexistence and community assembly control macro-evolutionary processes (Gerhold *et al.*, 2015)

While still strictly an observational study, of all the chapters in this thesis, Chapter 5 arguably provided the tightest link between pattern and process. Whereas Chapter 4 considered temporal patterns in community structure independent of environmental factors other than time since fire, the aim of Chapter 5 was to explicitly model the environmental basis for observed patterns of co-occurrence. To this end, a strength of Chapter 5 was the adoption of a latent variable model approach. As illustrated with R code in Appendix D, latent variable models provide a robust means of decomposing non-random co-occurrence patterns into that which is attributable to measured and unmeasured factors. Nevertheless, even the most sophisticated models can only go so far. Niche dimensions modelled using observed abundances reflect species' realised niches, but say little about their potential breadth in the absence of competition. Ultimately it is not clear from these results to what extent spatial heterogeneity in soil moisture (or other factors) stabilizes coexistence or merely 'filters' species into their respective fundamental niches.

Documenting significant patterns of co-occurrence is undoubtedly an important first step, but as of yet I am unaware of any strict tests of species coexistence under soil moisture heterogeneity. To this end, it is necessary to not only demonstrate species-specific environmental responses, but also that the effect of the environment and competition positively covary such that intraspecific competition is concentrated relative to interspecific competition (Chesson, 2000a,b, Sears & Chesson, 2007b, Adler *et al.*, 2007). In an effort to draw an explicit link between differences in functional traits, relevant physiological processes and the underlying coexistence mechanism, an ideal future study would investigate the extent to which environment-competition covariance is positively related to differences in traits related to plant water usage and stress tolerance (Adler *et al.*, 2013a).

A feature of soil moisture that makes it such a worthy focal point of plant coexistence research is its dual role as both a depletable resource factor (cf. nutrients or light) subject to density-dependent feedbacks and as a non-resource factor (cf. temperature, soil type or topography) mediating density-independent responses. As such, depending on supply rates, spatial heterogeneity in soil

moisture will likely produce complex interactions between niche and fitness differences amongst interacting species. While there is some evidence that opportunities for coexistence may be greater for heterogeneity in non-resource factors (Schoolmaster, 2013), it might be expected that the breadth of plant responses to soil moisture will allow for higher-dimensional trade-offs relative to more narrowly defined environmental factors. Isolating the underlying physiological processes and trade-offs should be a high priority. As recently argued by Silvertown *et al.* (2015), the time is ripe for an integrated research programme into hydrological niches drawing on the combined expertise of community ecologists, plant-physiologists, and hydrologists.

Shifting scales and passé paradigms: concluding remarks

My overall goal in this thesis was to explore alternative spatio-temporal perspectives on the relationship between heterogeneity and plant community assembly. While the individual chapters provide unique insights, the thesis as a whole admittedly only scratched the surface of this broad overarching theme. Nevertheless, I hope the reader will indulge me in extrapolating two concluding points that emerge from the chapters examined in unison.

1. Subtle shifts in the scales at which we examine spatial and temporal heterogeneity can yield new insights into patterns of diversity and their organizing processes.
2. Reliable mapping of processes from patterns requires that paradigmatic assumptions are subject to continuous theoretical and empirical scrutiny, and are updated accordingly.

References

- Abrams, P.A. (1986) The competitive exclusion principle: Other views and a reply. *Trends in Ecology & Evolution* **1**, 131–2.
- Adler, P.B. & Drake, J.M. (2008) Environmental variation, stochastic extinction, and competitive coexistence. *The American Naturalist* **172**, 186–95.
- Adler, P.B., Ellner, S.P. & Levine, J.M. (2010) Coexistence of perennial plants: an embarrassment of niches. *Ecology Letters* **13**, 1019–29.
- Adler, P.B., Fajardo, A., Kleinhesselink, A.R. & Kraft, N.J.B. (2013a) Trait-based tests of coexistence mechanisms. *Ecology Letters* **16**, 1294–1306.
- Adler, P.B., HilleRisLambers, J., Kyriakidis, P.C., Guan, Q. & Levine, J.M. (2006) Climate variability has a stabilizing effect on the coexistence of prairie grasses. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 12793–8.
- Adler, P.B., Hillerislambers, J. & Levine, J.M. (2007) A niche for neutrality. *Ecology Letters* **10**, 95–104.
- Adler, P.B., HilleRisLambers, J. & Levine, J.M. (2009) Weak effect of climate variability on coexistence in a sagebrush steppe community. *Ecology* **90**, 3303–3312.
- Adler, P.B., Salguero-Gómez, R., Compagnoni, A., Hsu, J.S., Ray-Mukherjee, J., Mbeau-Ache, C. & Franco, M. (2013b) Functional traits explain variation in plant life history strategies. *Proceedings of the National Academy of Sciences of the United States of America* .
- Alexandrou, M.A., Cardinale, B.J., Hall, J.D., Delwiche, C.F., Fritschie, K.J., Narwani, A., Venail, P.A., Bentlage, B., Pankey, M.S. & Oakley, T.H. (2015) Evolutionary relatedness does not predict competition and co-occurrence in natural or experimental communities of green algae. *Proceedings. Biological sciences / The Royal Society* **282**, 20141745.
- Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G. & Harmon, L.J. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**, 13410–13414.
- Alvarez, L.H.R. (2001) Does increased stochasticity speed up extinction? *Journal of Mathematical Biology* **43**, 534–544.

- Amarasekare, P. (2003) Competitive coexistence in spatially structured environments: a synthesis. *Ecology Letters* **6**, 1109–1122.
- Anderson, M. & Kneale, P. (1980) Topography and hillslope soil water relationships in a catchment of low relief. *Journal of Hydrology* **47**, 115–128.
- Angert, A.L., Huxman, T.E., Chesson, P. & Venable, D.L. (2009) Functional tradeoffs determine species coexistence via the storage effect. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 11641–5.
- Araya, Y.N., Silvertown, J., Gowing, D.J., McConway, K.J., Linder, H.P. & Midgley, G. (2011) A fundamental, eco-hydrological basis for niche segregation in plant communities. *New Phytologist* **189**, 253–8.
- Ashcroft, M.B. & Gollan, J.R. (2012) Fine-resolution (25 m) topoclimatic grids of near-surface (5 cm) extreme temperatures and humidities across various habitats in a large (200 × 300 km) and diverse region. *International Journal of Climatology* **32**, 2134–2148.
- Ashcroft, M.B., Gollan, J.R., Warton, D.I. & Ramp, D. (2012) A novel approach to quantify and locate potential microrefugia using topoclimate, climate stability, and isolation from the matrix. *Global Change Biology* **18**, 1866–1879.
- Atwell, B.J., Kriedemann, P.E. & Turnbull, C.G. (1999) *Plants in action: adaptation in nature, performance in cultivation*. Macmillan Education AU.
- Austin, M. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* **157**, 101–118.
- Barton, K. (2012) MuMIn: multi-model inference. *R package version 1.8.0*.
- Beaulieu, J.M., Jhwueng, D.C., Boettiger, C. & O'Meara, B.C. (2012) Modeling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution* **66**, 2369–2383.
- Belyea, L.R. & Lancaster, J. (1999) Assembly rules within a contingent ecology. *Oikos* pp. 402–416.
- Bennett, J.A., Lamb, E.G., Hall, J.C., Cardinal-McTeague, W.M. & Cahill, J.F. (2013) Increased competition does not lead to increased phylogenetic overdispersion in a native grassland. *Ecology Letters* **16**, 1168–76.
- Bhaskar, R., Dawson, T.E. & Balvanera, P. (2014) Community assembly and functional diversity along succession post-management. *Functional Ecology* **28**, 1256–1265.
- Blomberg, S.P., Garland, T. & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**, 717–45.
- Bongers, F., Poorter, L., Hawthorne, W.D. & Sheil, D. (2009) The intermediate disturbance hypothesis applies to tropical forests, but disturbance contributes little to tree diversity. *Ecology Letters* **12**, 798–805.
- Boyce, M.S., Haridas, C.V., Lee, C.T. & The Nceas Stochastic Demography Working Group (2006) Demography in an increasingly variable world. *Trends in Ecology & Evolution* **21**, 141–8.

- Burnham, K. & Anderson, D. (2002) *Model Selection and Multi-Model Inference*. Springer, 2nd edn.
- Butler, M.A. & King, A.A. (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist* **164**, 683–695.
- Cadotte, M., Albert, C.H. & Walker, S.C. (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters* **16**, 1234–1244.
- Cahill, J.F., Kembel, S.W., Lamb, E.G. & Keddy, P.a. (2008) Does phylogenetic relatedness influence the strength of competition among vascular plants? *Perspectives in Plant Ecology, Evolution and Systematics* **10**, 41–50.
- Cassel, D., Wendoroth, O. & Nielsen, D. (2000) Assessing Spatial Variability in an Agricultural Experiment Station Field: Opportunities Arising from Spatial Dependence. *Agron. J.*.
- Cavender-Bares, J., Keen, A. & Miles, B. (2006) Phylogenetic Structure of Floridian Plant Communities Depends on Taxonomic and Spatial Scale. *Ecology* **87**, 109–122.
- Cavender-Bares, J., Kitajima, K. & Bazzaz, F.A. (2004) Multiple trait associations in relation to habitat differentiation among 17 Floridian oak species. *Ecological Monographs* **74**, 635–662.
- Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters* **12**, 693–715.
- Chase, J. & Leibold, M. (2003) *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago, IL.
- Chase, J.M. (2003) Community assembly: when should history matter? *Oecologia* **136**, 489–98.
- Chase, J.M. (2014) Spatial scale resolves the niche versus neutral theory debate. *Journal of Vegetation Science* **25**, 319–322.
- Chesson, P. (2000a) General theory of competitive coexistence in spatially-varying environments. *Theoretical population biology* **58**, 211–37.
- Chesson, P. (2000b) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* **31**, 343–366.
- Chesson, P. & Huntly, N. (1997) The roles of harsh and fluctuating conditions in the dynamics of ecological communities. *The American Naturalist* **150**, 519–53.
- Chesson, P.L. (1985) Coexistence of competitors in spatially and temporally varying environments: A look at the combined effects of different sorts of variability. *Theoretical Population Biology* **28**, 263–287.
- Clark, J.S., Gelfand, A.E., Woodall, C.W. & Zhu, K. (2014) More than the sum of the parts: forest climate response from joint species distribution models. *Ecological Applications* **24**, 990–999.

- Clarke, P.J., Knox, K.J.E., Wills, K.E. & Campbell, M. (2005) Landscape patterns of woody plant response to crown fire: disturbance and productivity influence sprouting ability. *Journal of Ecology* **93**, 544–555.
- Clements, F.E. (1916) *Plant succession; an analysis of the development of vegetation*. Carnegie Institution of Washington, Washington.
- Connell, J.H. (1978) Diversity in tropical rain forests and coral reefs. *Science* **199**, 1302–1310.
- Connell, J.H. & Slatyer, R.O. (1977) Mechanisms of Succession in Natural Communities and Their Role in Community Stability and Organization. *The American Naturalist* **111**, 1119.
- Coomes, D.A. & Grubb, P.J. (2003) Colonization, tolerance, competition and seed-size variation within functional groups. *Trends in Ecology & Evolution* **18**, 283–291.
- Cornwell, W.K. & Ackerly, D.D. (2009) Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs* **79**, 109–126.
- Cressie, N., Calder, C.A., Clark, J.S., Hoef, J.M.V. & Wikle, C.K. (2009) Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications* **19**, 553–570.
- Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guegan, J.F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. & Turner, J.R.G. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters* **7**, 1121–1134.
- de Bello, F., Price, J.N., Münkemüller, T., Liira, J., Zobel, M., Thuiller, W., Gerhold, P., Götzenberger, L., Lavergne, S., Lepš, J., Zobel, K. & Pärtel, M. (2012) Functional species pool framework to test for biotic effects on community assembly. *Ecology* **93**, 2263–2273.
- DECC (2008) NSW Interim Native Vegetation Extent (2008-Version 1). Report and data prepared by NSW Department of Environment and Climate Change for the National Land and Water Resources Audit Sydney, Australia .
- Descamps-Julien, B. & Gonzalez, A. (2005) Stable Coexistence in a Fluctuating Environment: an Experimental Demonstration. *Ecology* **86**, 2815–2824.
- Diaz, S., Cabido, M. & Casanoves, F. (1998) Plant functional traits and environmental filters at a regional scale. *Journal of Vegetation Science* **9**, 113–122.
- Diez, J.M., Giladi, I., Warren, R. & Pulliam, H.R. (2014) Probabilistic and spatially variable niches inferred from demography. *Journal of Ecology* **102**, 544–554.
- Dinnage, R. (2009) Disturbance alters the phylogenetic composition and structure of plant communities in an old field system. *PloS one* **4**, e7071.
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46.

- Dormann, C.F., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609–628.
- Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology* **7**, 214.
- Dwyer, J.M., Hobbs, R.J., Wainwright, C.E. & Mayfield, M.M. (2015) Climate moderates release from nutrient limitation in natural annual plant communities. *Global Ecology and Biogeography* **24**, 549–561.
- Easterling, D.R., Meehl, G.A., Parmesan, C., Changnon, S.A., Karl, T.R. & Mearns, L.O. (2000) Climate extremes: observations, modeling, and impacts. *Science New York NY* **289**, 2068–2074.
- Einstein, A. (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* **322**, 549–560.
- Emerson, B.C. & Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution* **23**, 619–30.
- Felsenstein, J. (1985) Phylogenies and the comparative method. *American Naturalist* pp. 1–15.
- Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology* **43**, 393–404.
- FitzJohn, R.G. (2012) Diversitree : comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* **3**, 1084–1092.
- Fox, G.A., Scheiner, S.M. & Willig, M.R. (2011) Ecological Gradient Theory: A Framework for Aligning Data and Models. *The Theory of Ecology* (eds. S.M. Scheiner & M. Willig), pp. 283–307, University of Chicago Press.
- Fox, J.W. (2011) Bandwagons in ecology. *Oikos Blog* <https://oikosjournal.wordpress.com/2011/09/14/band/>.
- Fox, J.W. (2012) When should we expect microbial phenotypic traits to predict microbial abundances? *Frontiers in microbiology* **3**, 268.
- Francis, A.P. & Currie, D.J. (2003) A globally consistent richness-climate relationship for angiosperms. *The American Naturalist* **161**, 523–36.
- Fridley, J.D., Grime, J.P., Askew, A.P., Moser, B. & Stevens, C.J. (2011) Soil heterogeneity buffers community response to climate change in species-rich grassland. *Global Change Biology* **17**, 2002–2011.
- Fridley, J.D., Grime, J.P., Huston, M.A., Pierce, S., Smart, S.M., Thompson, K., Börger, L., Brooker, R.W., Cerabolini, B.E.L., Gross, N., Liancourt, P., Michalet, R. & Le Bagousse-Pinguet, Y. (2012) Comment on "Productivity is a poor predictor of plant species richness". *Science (New York, N.Y.)* **335**, 1441; author reply 1441.

- Fritschie, K.J., Cardinale, B.J., Alexandrou, M.A. & Oakley, T.H. (2014) Evolutionary history and the strength of species interactions: testing the phylogenetic limiting similarity hypothesis. *Ecology* **95**, 1407–17.
- Fukami, T., Martijn Bezemer, T., Mortimer, S.R. & Putten, W.H. (2005) Species divergence and trait convergence in experimental plant community assembly. *Ecology Letters* **8**, 1283–1290.
- Gaston, K.J., Blackburn, T.M. & Spicer, J.I. (1998) Rapoport's rule: time for an epitaph? *Trends in Ecology & Evolution* **13**, 70–74.
- Gerhold, P., Cahill, J.F., Winter, M., Bartish, I.V. & Prinzing, A. (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology* .
- Godoy, O., Kraft, N.J.B. & Levine, J.M. (2014) Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology Letters* **17**, 836–44.
- Gotelli, N.J. & McCabe, D.J. (2002) Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. *Ecology* **83**, 2091–2096.
- Gotelli, N.J. & McGill, B.J. (2006) Null Versus Neutral Models: What's The Difference? *Ecography* **29**, 793–800.
- Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A., Lepš, J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel, K. & Zobel, M. (2012) Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biological Reviews* **87**, 111–27.
- Grime, J.P. (2006) Trait convergence and trait divergence in herbaceous plant communities: Mechanisms and consequences. *Journal of Vegetation Science* **17**, 255–260.
- Grubb, P.J. (1977) The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews* **52**, 107–145.
- Hampe, A. & Jump, A.S. (2011) Climate Relicts: Past, Present, Future. *Annual Review of Ecology Evolution and Systematics* **42**, 1–51.
- Hansen, J., Sato, M. & Ruedy, R. (2012) Perception of climate change. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2415–23.
- Hansen, T.F., Pienaar, J. & Orzack, S.H. (2008) A comparative method for studying adaptation to a randomly evolving environment. *Evolution* **62**, 1965–1977.
- Harden, G.J. (1993) *Flora of New South Wales*. UNSW Press.
- Hardy, O.J. & Pavoine, S. (2012) Assessing phylogenetic signal with measurement error: a comparison of Mantel tests, Blomberg et al.'s K, and phylogenetic distograms. *Evolution* **66**, 2614–21.
- Harmon, L.J., Losos, J.B., Jonathan Davies, T., Gillespie, R.G., Gittleman, J.L., Bryan Jennings, W., Kozak, K.H., McPeek, M.A., Moreno-Roark, F., Near, T.J. & Others (2010) Early bursts of body size and shape evolution are rare in comparative data. *Evolution* **64**, 2385–2396.

- Harris, D.J. (2015) Generating realistic assemblages with a Joint Species Distribution Model. *Methods in Ecology and Evolution* pp. n/a–n/a.
- Heard, S.B. (1997) Inferring evolutionary process from phylogenetic tree shape. *The quarterly review of Biology* **72**, 31–54.
- Helmus, M.R., Keller, W.B., Paterson, M.J., Yan, N.D., Cannon, C.H. & Rusak, J.A. (2010) Communities contain closely related species during ecosystem disturbance. *Ecology Letters* **13**, 162–174.
- Holman, L. (1999) *Spatial analysis of soil moisture, nutrient status and vegetation within a wet heathland*. Honours, University of Sydney.
- Hopper, S.D. (2009) OCBIL theory: towards an integrated understanding of the evolution, ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes. *Plant and Soil* **322**, 49–86.
- Houlder, D., Hutchinson, M., Nix, H. & McMahon, J. (2003) *ANUCLIM 5.1 user's guide*. Australian National University: Canberra.
- Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*, vol. 17 of *Monographs in Population Biology*. Princeton University Press.
- Hui, F.K., Taskinen, S., Pledger, S., Foster, S.D. & Warton, D.I. (2014) Model-Based Approaches to Unconstrained Ordination. *Methods in Ecology and Evolution* .
- Hutchinson, G.E. (1961) The paradox of the plankton. *American Naturalist* pp. 137–145.
- IPPC (2007) *Fourth Assessment Report: Climate Change 2007: The AR4 Synthesis Report*. Geneva: IPCC.
- Ives, A.R. & Helmus, M.R. (2011) Generalized linear mixed models for phylogenetic analyses of community structure. *Ecological Monographs* **81**, 511–525.
- Jiang, L. & Morin, P.J. (2007) Temperature fluctuation facilitates coexistence of competing species in experimental microbial communities. *The Journal of animal ecology* **76**, 660–8.
- Johnson, E.A. & Miyanishi, K. (2008) Testing the assumptions of chronosequences in succession. *Ecology Letters* **11**, 419–31.
- Johnson, M.T.J. & Stinchcombe, J.R. (2007) An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology & Evolution* **22**, 250–7.
- Johnson, P.C. (2014) Extension of Nakagawa & Schielzeth's R 2 GLMM to random slopes models. *Methods in Ecology and Evolution* **5**, 944–946.
- Keith, D.A. (2004) *Ocean Shores to Desert Dunes: The Native Vegetation of New South Wales and the ACT*. Department of Environment and Conservation, Sydney.
- Keith, D.A. & Bradstock, R.A. (1994) Fire and competition in Australian heath: a conceptual model and field investigations. *Journal of Vegetation Science* **5**, 347–354.

- Keith, D.A., Holman, L., Rodoreda, S., Lemmon, J. & Bedward, M. (2007) Plant functional types can predict decade-scale changes in fire-prone vegetation. *Journal of Ecology* **95**, 1324–1337.
- Keith, D.A. & Tozer, M.G. (2012) Vegetation dynamics in coastal heathlands of the Sydney basin. *Proceedings of the Linnean Society of New South Wales* **134**, B181.
- Kelly, S., Grenyer, R. & Scotland, R.W. (2014) Phylogenetic trees do not reliably predict feature diversity. *Diversity and Distributions* **20**, 600–612.
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–4.
- Kembel, S.W. & Hubbell, S.P. (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology* **87**, S86–S99.
- Kimball, S., Gremer, J.R., Angert, A.L., Huxman, T.E. & Venable, D.L. (2011) Fitness and physiology in a variable environment. *Oecologia* **169**, 319–29.
- Kissling, W.D. & Carl, G. (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography* **17**, 59–71.
- Knapp, S., Kühn, I., Schweiger, O. & Klotz, S. (2008) Challenging urban species diversity: contrasting phylogenetic patterns across plant functional groups in Germany. *Ecology Letters* **11**, 1054–64.
- Knott, M. & Bartholomew, D.J. (1999) *Latent variable models and factor analysis*. 7, Edward Arnold.
- Kotler, B.P. & Brown, J.S. (1988) Environmental Heterogeneity and the Coexistence of Desert Rodents. *Annual Review of Ecology and Systematics* **19**, 281–307.
- Kozak, K.H. & Wiens, J.J. (2012) Phylogeny, ecology, and the origins of climate-richness relationships. *Ecology* **93**, S167–S181.
- Kraft, N.J.B. & Ackerly, D.D. (2010) Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecological Monographs* **80**, 401–422.
- Kraft, N.J.B., Adler, P.B., Godoy, O., James, E., Fuller, S. & Levine, J.M. (2014) Community assembly, coexistence, and the environmental filtering metaphor. *Functional Ecology* .
- Kraft, N.J.B., Cornwell, W.K., Webb, C.O. & Ackerly, D.D. (2007) Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist* **170**, 271–283.
- Kraft, N.J.B., Godoy, O. & Levine, J.M. (2015) Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences* p. 201413650.
- Kunstler, G., Lavergne, S., Courbaud, B., Thuiller, W., Vieilledent, G., Zimmermann, N.E., Kattge, J. & Coomes, D.A. (2012) Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. *Ecology Letters* **15**, 831–40.

- Lai, H.R., Mayfield, M.M., Gay-des combes, J.M., Spiegelberger, T. & Dwyer, J.M. (2015) Distinct invasion strategies operating within a natural annual plant system. *Ecology Letters* **18**, 336–346.
- Laliberté, E., Zemunik, G. & Turner, B.L. (2014) Environmental filtering explains variation in plant diversity along resource gradients. *Science* **345**, 1602–1605.
- Lawton, J.H. (1999) Are there general laws in ecology? *Oikos* pp. 177–192.
- Leibold, M.A. (1998) Similarity and local co-existence of species in regional biotas. *Evolutionary Ecology* **12**, 95–110.
- Letcher, S.G. (2010) Phylogenetic structure of angiosperm communities during tropical forest succession. *Proceedings. Biological sciences / The Royal Society* **277**, 97–104.
- Letten, A.D., Ashcroft, M.B., Keith, D.A., Gollan, J.R. & Ramp, D. (2013) The importance of temporal climate variability for spatial patterns in plant diversity. *Ecography* **36**, 1341–1349.
- Letten, A.D. & Cornwell, W.K. (2014) Trees, branches and (square) roots: why evolutionary relatedness is not linearly related to functional distance. *Methods in Ecology and Evolution* .
- Letten, A.D., Keith, D.A. & Tozer, M.G. (2014) Phylogenetic and functional dissimilarity does not increase during temporal heathland succession. *Proceedings. Biological sciences / The Royal Society* **281**, 20142102–.
- Letten, A.D., Keith, D.A., Tozer, M.G. & Hui, F.K. (2015) Fine-scale hydrological niche differentiation through the lens of multi-species co-occurrence models. *Journal of Ecology* .
- Levin, S.A. (1992) The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology* **73**, 1943.
- Levine, J.M. & HilleRisLambers, J. (2009) The importance of niches for the maintenance of species diversity. *Nature* **461**, 254–7.
- Levine, J.M. & Rees, M. (2002) Coexistence and relative abundance in annual plant assemblages: the roles of competition and colonization. *The American Naturalist* **160**, 452–67.
- Levine, J.M. & Rees, M. (2004) Effects of temporal variability on rare plant persistence in annual systems. *The American Naturalist* **164**, 350–63.
- Li, H. & Reynolds, J.F. (1995) On definition and quantification of heterogeneity. *Oikos* **73**, 280–284.
- MacArthur, R.H. & Levins, R. (1967) The limiting similarity, convergence and divergence of coexisting species. *The American Naturalist* **101**, 377–385.
- Maddison, W.P. & Slatkin, M. (1991) Null Models for the Number of Evolutionary Steps in a Character on a Phylogenetic Tree. *Evolution* **45**, 1184–1197.
- Mayfield, M. & Levine, J.M. (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters* **13**, 1085–1093.

- McCullough, P. & Nelder, J.A. (1989) *Generalized linear models*. Chapman & Hall, London.
- McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution* **21**, 178–185.
- McGlone, M. (1996) When history matters: scale, time, climate and tree diversity. *Global Ecology and Biogeography Letters* **5**, 309–314.
- Møller, A. & Jennions, M. (2002) How much variance can be explained by ecologists and evolutionary biologists? *Oecologia* **132**, 492–500.
- Mooers, A.O. (1995) Tree balance and tree completeness. *Evolution* **49**, 379–384.
- Muko, S. & Iwasa, Y. (2000) Species coexistence by permanent spatial heterogeneity in a lottery model. *Theoretical population biology* **57**, 273–84.
- Murphy, B.P., Paron, P., Prior, L.D., Boggs, G.S., Franklin, D.C. & Bowman, D.M.J.S. (2010) Using generalized autoregressive error models to understand fire-vegetation-soil feedbacks in a mulga-spinifex landscape mosaic. *Journal of Biogeography* **37**, 2169–2182.
- Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution* **4**, 133–142.
- Narwani, A., Alexandrou, M.A., Oakley, T.H., Carroll, I.T. & Cardinale, B.J. (2013) Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters* **16**, 1373–81.
- Norden, N., Letcher, S.G., Boukili, V., Swenson, N.G. & Chazdon, R. (2012) Demographic drivers of successional changes in phylogenetic structure across life-history stages in plant communities. *Ecology* **93**, S70–S82.
- O'Brien, E. (1998) Water-energy dynamics, climate, and prediction of woody plant species richness: an interim general model. *Journal of Biogeography* **25**, 379–398.
- O'Meara, B.C. (2012) Evolutionary inferences from phylogenies: a review of methods. *Annual Review of Ecology, Evolution, and Systematics* **43**, 267–285.
- Ovaskainen, O., Hottola, J. & Siitonen, J. (2010) Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology* **91**, 2514–2521.
- Pachepsky, Y.A., Guber, A.K. & Jacques, D. (2005) Temporal persistence in vertical distributions of soil moisture contents. *Soil Sci. Soc. Am. J.* .
- Pärtel, M., Szava-Kovats, R. & Zobel, M. (2011) Dark diversity: shedding light on absent species. *Trends in Ecology & Evolution* **26**, 124–8.
- Pausas, J.G. & Austin, M.P. (2001) Patterns of plant species richness in relation to different environments: An appraisal. *Journal of Vegetation Science* **12**, 153–166.
- Pearse, W.D., Jones, A. & Purvis, A. (2013) Barro Colorado Island's phylogenetic assemblage structure across fine spatial scales and among clades of different ages. *Ecology* **94**, 2861–2872.

- Pearse, W.D. & Purvis, A. (2013) phyloGenerator: An automated phylogeny generation tool for ecologists. *Methods in Ecology and Evolution* **4**, 692–698.
- Pennell, M.W. & Harmon, L.J. (2013) An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. *Annals of the New York Academy of Sciences* **1289**, 90–105.
- Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jau-reguiberry, P., Bret-Harte, M.S., Cornwell, W.K., Craine, J.M., Gurvich, D.E., Urcelay, C., Veneklaas, E.J., Reich, P.B., Poorter, L., Wright, I.J., Ray, P., Enrico, L., Pausas, J.G., de Vos, A.C., Buchmann, N., Funes, G., Quétier, F., Hodgson, J.G., Thompson, K., Morgan, H.D., ter Steege, H., Sack, L., Blonder, B., Poschlod, P., Vaieretti, M.V., Conti, G., Staver, A.C., Aquino, S. & Cornelissen, J.H.C. (2013) New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany* **61**, 167.
- Pickett, S.T.A. & Bazzaz, F.A. (1978) Organization of an Assemblage of Early Successional Species on a Soil Moisture Gradient. *Ecology* **59**, 1248.
- Plummer, M. (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. March, pp. 20–22.
- Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A. & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution* **5**, 397–406.
- Price, J.N. & Pärtel, M. (2013) Can limiting similarity increase invasion resistance? A meta-analysis of experimental studies. *Oikos* **122**, 649–656.
- Purschke, O., Schmid, B.C., Sykes, M.T., Poschlod, P., Michalski, S.G., Durka, W., Kühn, I., Winter, M. & Prentice, H.C. (2013) Contrasting changes in taxonomic, phylogenetic and functional diversity during a long-term succession: insights into assembly processes. *Journal of Ecology* **101**, 857–866.
- Purvis, A. & Hector, A. (2000) Getting the measure of biodiversity. *Nature* **405**, 212–219.
- Revell, L.J. (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217–223.
- Reyer, C., Leuzinger, S., Rammig, A., Wolf, A., Bartholomeus, R.P., Bonfante, A., Lorenzi, F., Dury, M., Gloning, P., Abou Jaoudé, R., Klein, T., Kuster, T.M., Martins, M., Niedrist, G., Riccardi, M., Wohlfahrt, G., Angelis, P., Dato, G., François, L., Menzel, A. & Pereira, M. (2012) A plant's perspective of extremes: Terrestrial plant responses to changing climatic variability. *Global Change Biology* **19**, 75–89.
- Ricotta, C. & Burrascano, S. (2009) Testing for differences in beta diversity with asymmetric dissimilarities. *Ecological Indicators* **9**, 719–724.
- Riginos, C. (2009) Grass competition suppresses savanna tree growth across multiple demographic stages. *Ecology* **90**, 335–340.

- Rosenfeld, J.S. (2002) Functional redundancy in ecology and conservation. *Oikos* **98**, 156–162.
- Schoolmaster, D.R. (2013) Resource competition and coexistence in heterogeneous metacommunities: many-species coexistence is unlikely to be facilitated by spatial variation in resources. *PeerJ* **1**, e136.
- Sears, A.L.W. & Chesson, P. (2007a) NEW METHODS FOR QUANTIFYING THE SPATIAL STORAGE EFFECT: AN ILLUSTRATION WITH DESERT ANNUALS. *Ecology* **88**, 2240–2247.
- Sears, A.L.W. & Chesson, P. (2007b) New methods for quantifying the spatial storage effect: an illustration with desert annuals. *Ecology* **88**, 2240–2247.
- Shmida, A. & Ellner, S. (1984) Coexistence of plant species with similar niches. *Vegetatio* **58**, 29–55.
- Shurin, J.B., Winder, M., Adrian, R., Keller, W.B., Matthews, B., Paterson, A.M., Paterson, M.J., Pinel-Alloul, B., Rusak, J.A. & Yan, N.D. (2010) Environmental stability and lake zooplankton diversity - contrasting effects of chemical and thermal variability. *Ecology Letters* **13**, 453–63.
- Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology & Evolution* **19**, 605–611.
- Silvertown, J., Araya, Y. & Gowing, D. (2015) Hydrological niches in terrestrial plant communities: A review. *Journal of Ecology* **103**, 93–108.
- Silvertown, J., Dodd, M.E., Gowing, D.J.G. & Mountford, J.O. (1999) Hydrologically defined niches reveal a basis for species richness in plant communities. *Nature* **400**, 61–63.
- Skrondal, A. & Rabe-Hesketh, S. (2004) *Generalized latent variable modeling: Multi-level, longitudinal, and structural equation models*. CRC Press.
- Slingsby, J.A. & Verboom, G.A. (2006) Phylogenetic relatedness limits co-occurrence at fine spatial scales: evidence from the schoenoid sedges (Cyperaceae: Schoeneae) of the Cape Floristic Region, South Africa. *The American Naturalist* **168**, 14–27.
- Smith, F.A., Boyer, A.G., Brown, J.H., Costa, D.P., Dayan, T., Ernest, S.K.M., Evans, A.R., Fortelius, M., Gittleman, J.L., Hamilton, M.J. & Others (2010) The evolution of maximum body size of terrestrial mammals. *Science* **330**, 1216–1219.
- Smith, M.D. (2011) The ecological role of climate extremes: current understanding and future prospects. *Journal of Ecology* **99**, 651–655.
- Snyder, R.E. & Chesson, P. (2004) How the spatial scales of dispersal, competition, and environmental heterogeneity interact to affect coexistence. *The American Naturalist* **164**, 633–50.
- Stein, A., Gerstner, K. & Kreft, H. (2014) Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* **17**, 866–80.

- Stenseth, N.C., Mysterud, A., Ottersen, G., Hurrell, J.W., Chan, K.S. & Lima, M. (2002) Ecological effects of climate fluctuations. *Science (New York, N.Y.)* **297**, 1292–6.
- Stevens, G. (1989) The latitudinal gradient in geographical range: how so many species coexist in the tropics. *American Naturalist* **133**, 240–256.
- Stubbs, W.J. & Wilson, J.B. (2004) Evidence for limiting similarity in a sand dune community. *Journal of Ecology* **92**, 557–567.
- Su, Y.S. & Yajima, M. (2012) R2jags: A Package for Running jags from R. *R package version 0.03-08, URL http://CRAN.R-project.org/package=R2jags*.
- Suggitt, A.J., Gillingham, P.K., Hill, J.K., Huntley, B., Kunin, W.E., Roy, D.B. & Thomas, C.D. (2011) Habitat microclimates drive fine-scale variation in extreme temperatures. *Oikos* **120**, 1–8.
- Swenson, N.G. (2013) The assembly of tropical tree communities - the advances and shortcomings of phylogenetic and functional trait analyses. *Ecography* **36**, no–no.
- Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. & Zimmerman, J.K. (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology* **87**, 2418–2424.
- Swenson, N.G., Stegen, J.C., Davies, S.J., Erickson, D.L., Forero-Montaña, J., Hurlbert, A.H., Kress, W.J., Thompson, J., Uriarte, M., Wright, S.J. & Zimmerman, J.K. (2012) Temporal turnover in the composition of tropical tree communities: functional determinism and phylogenetic stochasticity. *Ecology* **93**, 490–499.
- Tilman, D. (1982) *Resource Competition and Community Structure. (Mpb-17)*. Princeton University Press.
- Tofts, R. & Silvertown, J. (2000) A phylogenetic approach to community assembly from a local species pool. *Proceedings. Biological sciences / The Royal Society* **267**, 363–369.
- Tozer, M. & Bradstock, R. (2003) Fire-mediated effects of overstorey on plant species diversity and abundance in an eastern Australian heath. *Plant Ecology* **164**, 213–223.
- Tuck, C. & Romanuk, T.N. (2012) Robustness to thermal variability differs along a latitudinal gradient in zooplankton communities. *Global Change Biology* **18**, 1597–1608.
- Vachaud, G., Passerat De Silans, A., Balabanis, P. & Vauclin, M. (1985) Temporal Stability of Spatially Measured Soil Water Probability Density Function. *Soil Sci. Soc. Am. J.* .
- Vamosi, S.M., Heard, S.B., Vamosi, J.C. & Webb, C.O. (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology* **18**, 572–592.
- Vellend, M., Cornwell, W.K., Magnuson-Ford, K. & Mooers, A.O. (2010) Measuring phylogenetic biodiversity. *Biological diversity: frontiers in measurement and assessment. Oxford University Press, Oxford, UK* .

- Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleynhans, E.J., Kraft, N.J.B., Letaw, A.D., Macdonald, A.A.M., Maclean, J.E., Myers-Smith, I.H., Norris, A.R. & Xue, X. (2014) Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* **123**, 1420–1430.
- Verdu, M. & Pausas, J.G. (2007) Fire drives phylogenetic clustering in Mediterranean Basin woody plant communities. *Journal of Ecology* **95**, 1316–1323.
- Verdu, M., Rey, P.J., Alcantara, J.M., Siles, G. & Valiente-Banuet, A. (2009) Phylogenetic signatures of facilitation and competition in successional communities. *Journal of Ecology* **97**, 1171–1180.
- Walker, L. & Chapin, S. (1987) Interactions among Processes Controlling Successional Change. *Oikos* **50**, 131 – 135.
- Walker, S.C. & Jackson, D.A. (2011) Random-effects ordination: describing and predicting multivariate correlations and co-occurrences. *Ecological Monographs* **81**, 635–663.
- Warton, D.I., Foster, S.D., De'ath, G., Stoklosa, J. & Dunstan, P.K. (2014) Model-based thinking for community ecology. *Plant Ecology* pp. 1–14.
- Warton, D.I., Wright, S.T. & Wang, Y. (2012) Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution* **3**, 89–101.
- Webb, C.O. (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist* **156**, 145–155.
- Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics* **33**, 475–505.
- Weiher, E., Clarke, G.D.P. & Keddy, P.A. (1998) Community assembly rules, morphological dispersion, and the coexistence of plant species. *Oikos* **81**, 309–322.
- Weiher, E. & Keddy, P. (2001) *Ecological assembly rules: perspectives, advances, retreats*. Cambridge University Press.
- Weiher, E. & Keddy, P.A. (1995) Assembly Rules, Null Models, and Trait Dispersion: New Questions from Old Patterns. *Oikos* **74**, 159–164.
- Westoby, M., Falster, D.S., Moles, A.T., Vesk, P.A. & Wright, I.J. (2002) Plant Ecological Strategies: Some Leading Dimensions of Variation Between Species. *Annual Review of Ecology and Systematics* **33**, 125–159.
- White, E.P., Ernest, S.K.M., Adler, P.B., Hurlbert, A.H. & Lyons, S.K. (2010) Integrating spatial and temporal approaches to understanding species richness. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **365**, 3633–43.
- Wiens, J.A. (1977) On Competition and Variable Environments: Populations may experience "ecological crunches" in variable climates, nullifying the assumptions of competition theory and limiting the usefulness of short-term studies of population patterns. *American Scientist* pp. 590–597.

- Wilson, S. (1999) Plant interactions during secondary succession. *Ecosystems of disturbed ground* (ed. L. Walker), pp. 629–650, Elsevier, Amsterdam.
- Wood, S. (2012) mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation .
- Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G., McGlinn, D.J., O'Meara, B.C., Moles, A.T., Reich, P.B., Royer, D.L., Soltis, D.E., Stevens, P.F., Westoby, M., Wright, I.J., Aarssen, L., Bertin, R.I., Calamitus, A., Govaerts, R., Hemmings, F., Leishman, M.R., Oleksyn, J., Soltis, P.S., Swenson, N.G., Warman, L., Beaulieu, J.M. & Ordonez, A. (2014) Data from: Three keys to the radiation of angiosperms into freezing environments. \url{http://dx.doi.org/10.5061/dryad.63q27.2}.

Appendix A

Supporting information: Chapter 2

Figure S2.1 – Spread of average annual precipitation amongst plots grouped by habitat type. Boxes show median (bold middle line) and 25% and 75% quartiles; whiskers show 1.5 times interquartile range or maximum value (whichever is smaller); and dots show outliers.

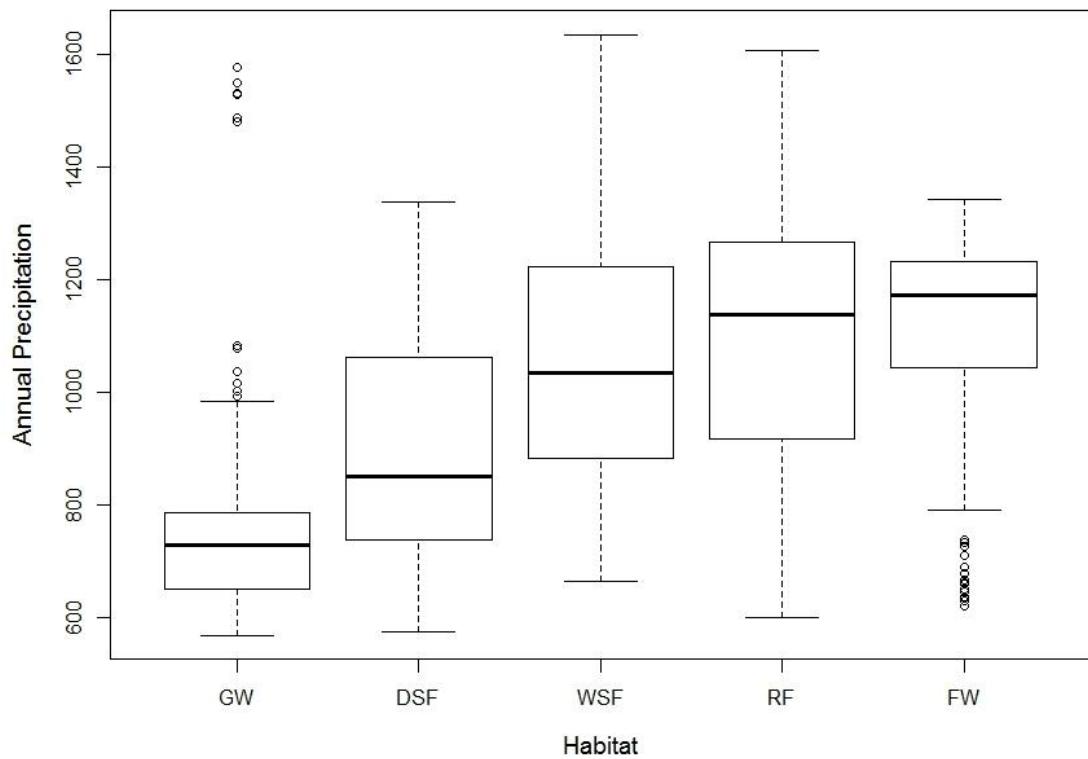


Figure S2.2 – Spread of average variability in maximum temperature amongst plots grouped by habitat type. Boxes show median (bold middle line) and 25% and 75% quartiles; whiskers show 1.5 times interquartile range or maximum value (whichever is smaller); and dots show outliers.

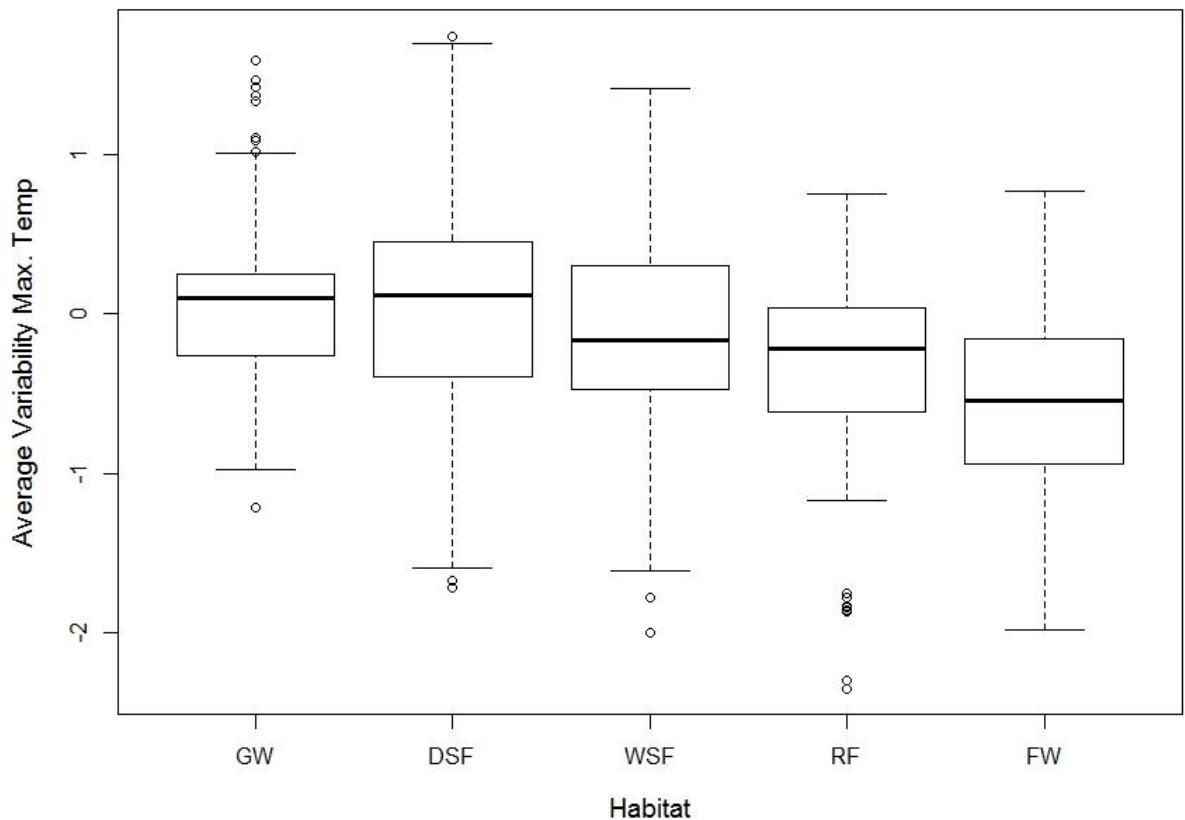


Table S2.1 - Parameter coefficient estimates for global models for each forest type and all plots combined using non-spatial Gaussian models or spatially sensitive SAR error models.

Quasipoisson regression										SAR error model									
Quasipoisson regression					Gaussian regression					SAR error model									
All		Estimate	Std.Error	t value	Pr(> t)	Estimate		Std.Error	t value	Estimate		Std.Error	t value	z value	Pr(> z)				
(Intercept)	3.696218	0.006422	575.565	5.7E-16	***	(Intercept)	40.7253	0.2581	157.771	<2e-16	***	(Intercept)	40.82572	0.88545	46.1071	<2e-16	***		
VarMT	1.359183	0.6222	2.184	0.029021	*	VarMT	25.1395	2.198	11.4347	2.8162	0.00486	VarMT	84.53426	30.0167	0.00850	<2e-16	**		
VarMT ²	-2.61569	0.446667	-5.856	0.57E-09	***	VarMT ²	16.4538	5.861	5.22E-09	*	***	VarMT ²	56.40293	21.43447	2.6314	0.00850	<2e-16	**	
CC	1.007941	0.455195	2.214	0.026898	*	CC	34.1994	18.1306	1.886	5.94E-02	.	CC	19.79844	18.29735	1.082	0.27924			
MaxH5	-3.726337	0.845806	-4.406	1.10E-05	***	MaxH5	-165.3472	34.5353	-4.788	1.79E-06	***	MaxH5	52.86746	46.04523	1.1482	0.25990			
MaxH5 ²	-2.780683	0.41191	-6.751	1.83E-11	***	MaxH5 ²	-113.2059	16.3644	-6.918	5.82E-12	***	MaxH5 ²	-68.15748	16.91243	-4.03	0.00006			
MaxH5	0.735995	0.718797	1.024	0.30597		MaxH5	33.5855	29.2633	1.148	2.51E-01		MaxH5	21.51259	28.22374	-0.7622	0.44593			
MaxH5 ²	-0.64232	0.519398	-1.237	0.216311		MaxH5 ²	-24.6956	20.5184	-1.204	0.228866		MaxH5 ²	-23.93515	20.43361	-1.1714	0.24145			
MinH5	2.657025	0.716274	3.71	0.000212	***	MinH5	29.0608	29.0608	4.267	2.06E-05	***	MinH5	133.6478	31.166339	4.2209	0.00002			
MinH5 ²	-1.70111	0.446819	-3.807	0.000144	***	MinH5 ²	-73.8534	17.9574	-4.113	4.04E-05	***	MinH5 ²	-52.9319	17.62095	-3.0039	0.00267	**		
MinT5	-2.99024	0.805254	-3.713	0.000269	***	MinT5	-118.6318	32.5768	-3.642	0.000277	***	MinT5	34.76203	40.45071	0.8594	0.39014			
MinT5 ²	2.173175	0.361598	6.01	2.13E-09	***	MinT5 ²	88.615	14.5705	6.082	1.37E-09	***	MinT5 ²	65.48044	14.3674	4.5576	0.00001	***		
MinT95	3.876975	1.045259	3.709	0.000213	***	MinT95	159.7826	42.1256	3.793	1.52E-04	***	MinT95	52.31951	51.81628	1.0097	0.31263			
MinT95 ²	0.315785	0.556954	0.567	0.570775		MinT95 ²	18.0257	21.8781	0.824	0.410066		MinT95 ²	-23.30304	26.7452	-0.8713	0.38359			
PCQ	-0.50873	0.729185	-0.698	0.485448		PCQ	-12.1073	29.7105	-0.408	0.68367		PCQ	-130.7871	47.55012	-2.7505	0.00535	**		
PCQ ²	-3.98798	0.4533585	-8.792	<2e-16	***	PCQ ²	-136.8466	16.6488	-8.22	3.25E-16	***	PCQ ²	-51.17131	25.24666	-2.0269	0.04268	*		
DSF										DSF									
(Intercept)	3.67149	0.00775	473.741	<2e-16	***	(Intercept)	30.7212	0.3047	130.344	<2e-16	***	(Intercept)	40.14681	1.16348	34.5058	<2e-16	***		
VarMT	-0.029683	0.57952	-0.051	0.95916		VarMT	22.9591	0.162	8.7169	7.7862	7.78E-05	VarMT	17.9572	18.08225	2.5916	0.00955	*		
VarMT ²	-1.51345	0.39905	-3.871	0.000113	***	VarMT ²	15.2088	3.962	8.7169	7.7862	7.78E-05	VarMT ²	17.9572	18.08225	2.5916	0.00955	*		
CC	1.01935	0.44876	2.272	0.02325	*	CC	37.5759	17.382	2.162	0.03079	*	CC	7.66776	16.67918	0.4597	0.64572			
MaxH5	-0.40636	0.586867	-0.692	0.488774		MaxH5	-22.0054	23.6389	-0.931	3.52E-01		MaxH5	106.3558	36.0087	2.9536	0.00314	**		
MaxH5 ²	-1.74976	0.38743	-4.516	6.76E-06	***	MaxH5 ²	-66.4928	14.8973	-4.463	8.64E-06	***	MaxH5 ²	-39.32521	14.32823	-2.7446	0.00606	**		
MaxH5	-0.76983	0.53228	-1.446	0.148291		MaxH5	-33.4624	20.6923	-1.617	1.06E-01		MaxH5	-0.91862	21.65697	-0.0424	0.96617			
MaxH5 ²	-1.58216	0.36755	-4.305	1.78E-05	***	MaxH5 ²	-56.0575	13.8313	-4.053	5.30E-05	***	MaxH5 ²	-47.16642	14.08297	-3.3492	0.00081	***		
MinH5	1.76947	0.7127	2.483	0.013139	*	MinH5	77.0276	28.3832	2.714	0.00672	*	MinH5	97.8897	28.21926	3.4689	0.00052	***		
MinH5 ²	-1.2034	0.41224	-2.919	0.003559	**	MinH5 ²	-46.5905	16.3783	-2.845	0.0045	**	MinH5 ²	-24.46748	15.31293	-1.5978	0.11008			
MinT5	0.10163	0.58556	0.174	0.862229		MinT5	10.5619	23.1463	0.456	6.48E-01		MinT5	71.91678	28.51156	2.5224	0.01166	*		
MinT5 ²	1.53865	0.32393	4.75	2.22E-06	***	MinT5 ²	64.3623	13.0544	4.93	9.07E-07	***	MinT5 ²	47.67709	12.59039	3.7968	0.00015	***		
PCQ	-0.86009	0.66982	-1.284	0.199307		PCQ	-32.6372	26.5229	-1.231	2.19E-01		PCQ	-143.0363	43.53363	-3.2852	0.00012	**		
PCQ ²	-2.7765	0.42068	-6.6	5.98E-11	***	PCQ ²	-102.2591	16.0176	-6.384	2.26E-10	***	PCQ ²	0.64838	27.64998	0.0234	0.98129			

WSF	Pr(> t)										Pr(> z)											
	Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> z)			
(Intercept)	3.78221	0.01348	281.065	<2e-16	***	(Intercept)	44.4785	0.5951	74.74	<2e-16	***	(Intercept)	44.58861	0.81464	54.734	<2e-16	***	(Intercept)	26.67831	22.91098	1.164	0.2425
VarMT	0.31937	0.50763	0.629	0.5297	***	VarMT	13.1729	22.2028	0.593	0.5534	9.65E-06	VarMT	-57.3296	14.69991	-3.9	9.62E-05	***	VarMT	14.795	33.33632	-0.982	0.3261
VarMT ²	-1.53583	0.33262	-4.617	5.54E-06	***	VarMT ²	-60.3494	13.4305	-4.493	9.65E-06	***	VarMT ²	-13.04588	14.19708	-0.9189	0.5814	***	VarMT ²	18.8652	35.33949	0.5371	0.59118
MaxT95	0.04725	0.77171	0.061	0.9512	0.9512	MaxT95	-0.9637	33.8897	-0.028	0.9773	0.3127	MaxT95	23.37391	16.09595	-1.4522	0.14645	0.14645	MaxT95	-1.427	0.1546	0.09535	-0.0165
MaxT95 ²	-0.3193	0.32711	-0.976	0.3297	0.3297	MaxT95 ²	-14.2515	14.0951	-1.011	0.3127	0.8652	MaxT95 ²	23.37391	16.09595	-1.4522	0.14645	0.14645	MaxT95 ²	-1.427	0.1546	0.09535	-0.0165
MaxT5	0.07799	0.75258	0.9175	0.104	0.104	MaxT5	-5.6712	33.3702	0.17	0.8652	0.6596	MaxT5	-6.40688	12.56355	-0.51	0.61008	0.61008	MaxT5	-1.427	0.1546	0.09535	-0.0165
MaxT5 ²	-0.55822	0.38408	-1.453	0.1471	0.1471	MaxT5 ²	-23.816	16.694	-1.427	0.1546	0.1919	MaxT5 ²	-32.05906	19.83676	-1.6161	0.10606	0.10606	MaxT5 ²	-1.427	0.1546	0.09535	-0.0165
MinH95	0.13797	0.33471	0.412	0.6805	0.6805	MinH95	4.9674	14.8011	0.336	0.7374	0.9512	MinH95	-16.50522	12.64776	-1.305	0.1919	0.1919	MinH95	-1.427	0.1546	0.09535	-0.0165
MinH95 ²	-0.35066	0.29793	-1.195	0.2329	0.2329	MinH95 ²	-14.6355	12.3861	-1.182	0.2382	0.9651	MinH95 ²	-17.4242	17.4242	-0.044	0.9651	0.9651	MinH95 ²	-1.427	0.1546	0.09535	-0.0165
MinT5	-0.01932	0.39786	-0.049	0.9613	0.9613	MinT5	-0.7629	17.4242	-0.044	0.9651	0.6596	MinT5	-0.44101	17.20265	0.0256	0.97955	0.97955	MinT5	-1.427	0.1546	0.09535	-0.0165
MinT5 ²	-0.10326	0.2854	-0.362	0.7177	0.7177	MinT5 ²	-5.4678	12.4032	-0.441	0.6596	0.6596	MinT5 ²	-22.1218	16.9174	-1.308	0.1919	0.1919	MinT5 ²	-1.427	0.1546	0.09535	-0.0165
PDQ	-0.52566	0.38344	-1.366	0.1729	0.1729	PDQ	-22.1218	16.9174	-1.308	0.1919	0.1919	PDQ	-26.0399	12.3407	-2.11	0.0956	*	PDQ	-28.23253	14.8213	-1.9043	0.05888
PDQ ²	-0.6996	0.29722	-2.354	0.0192	*	PDQ ²	-26.0399	12.3407	-2.11	0.0956	*	PDQ ²	-28.23253	14.8213	-1.9043	0.05888	*	PDQ ²	-28.23253	14.8213	-1.9043	0.05888
RF	Pr(> t)										Pr(> z)											
Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> z)				
(Intercept)	3.666	0.03698	99.141	<2e-16	***	(Intercept)	40.851	1.507	27.109	<2e-16	***	(Intercept)	1.7049	1.7049	24.32	<2e-16	***	(Intercept)	1.7049	1.7049	24.32	<2e-16
VarMT	1.35961	0.43346	1.829	0.0707	.	VarMT	53.834	30.807	1.747	0.0884	.	VarMT	54.0033	31.7114	1.703	0.08838	.	VarMT	19.1096	29.0014	19.1096	0.12911
VarMT ²	-0.41787	0.46482	-0.899	0.3711	*	VarMT ²	-18.807	17.141	-1.097	0.2755	.	VarMT ²	10.0624	29.0624	0.3462	0.72923	.	VarMT ²	0.5045	0.5045	0.3462	0.72923
MaxH95	0.29358	0.71821	0.409	0.6837	0.6837	MaxH95	21.022	31.37	0.67	0.4568	0.4568	MaxH95	-29.9884	28.9751	-1.035	0.30688	0.30688	MaxH95	-1.427	0.1546	0.09535	-0.0165
MaxH95 ²	-0.85301	0.79949	-1.067	0.2889	0.2889	MaxH95 ²	-23.745	31.77	-0.247	0.4568	0.1898	MaxH95 ²	29.5606	20.5896	1.4357	0.15109	0.15109	MaxH95 ²	-1.427	0.1546	0.09535	-0.0165
MinH5	0.68639	0.53811	1.276	0.2054	0.2054	MinH5	31.023	23.483	1.321	0.269	0.7886	MinH5	-12.4381	15.3498	0.8103	0.41776	0.41776	MinH5	-1.427	0.1546	0.09535	-0.0165
MinH5 ²	0.11337	0.43935	0.258	0.797	0.797	MinH5 ²	4.815	17.903	0.269	0.7886	0.7886	MinH5 ²	-33.02	34.1543	0.0701	0.10687	0.10687	MinH5 ²	-1.427	0.1546	0.09535	-0.0165
MinT95	1.75532	0.84702	2.072	0.0411	*	MinT95	63.319	34.1543	1.833	0.2272	0.2272	MinT95	-33.02	27.154	-0.216	0.2272	0.2272	MinT95	-1.427	0.1546	0.09535	-0.0165
MinT95 ²	-1.45053	0.73293	-1.979	0.0509	.	MinT95 ²	-33.02	27.154	-0.216	0.2272	0.2272	MinT95 ²	-13.9662	24.92327	-0.57523	0.57523	0.57523	MinT95 ²	-1.427	0.1546	0.09535	-0.0165
PDQ	0.27733	0.65232	0.425	0.6717	0.6717	PDQ	11.148	26.301	0.424	0.6727	0.6727	PDQ	-10.8864	29.0015	0.3754	0.70738	0.70738	PDQ	-24.6172	22.267	-1.1055	0.26892
PDQ ²	-0.71889	0.58286	-1.233	0.2206	0.2206	PDQ ²	-15.687	21.28	-0.737	0.4629	0.4629	PDQ ²	-26.0399	25.2647	0.823	0.41414	0.41414	PDQ ²	-26.0399	25.2647	0.823	0.41414
FW	Pr(> t)										Pr(> z)											
Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> t)		Estimate	Std.Error	t value	Pr(> z)				
(Intercept)	3.52061	0.02581	136.392	<2e-16	***	(Intercept)	34.8419	0.8808	39.558	<2e-16	***	(Intercept)	34.8253	1.1964	29.1077	<2e-16	***	(Intercept)	84.7265	26.1117	3.2448	0.001175
VarMT	1.91409	0.62345	3.07	0.00243	**	VarMT	62.74	22.4623	2.793	0.00572	**	VarMT	31.58	28.0595	1.1255	0.260392	0.260392	VarMT	1.272	0.204931	0.47343	0.140962
MaxT95	1.06369	0.77202	1.378	0.16978	0.16978	MaxT95	35.113	27.6108	1.718	0.1429	0.1429	MaxT95	34.2919	25.1842	1.3616	0.17331	0.17331	MaxT95	1.429	0.154659	0.382266	0.152638
MaxT95 ²	0.18889	0.48017	0.387	0.69906	0.69906	MaxT95 ²	12.8563	17.8995	0.718	0.47343	0.47343	MaxT95 ²	27.7726	18.8645	1.4722	0.140962	0.140962	MaxT95 ²	1.429	0.154659	0.382266	0.152638
MinT95	0.82781	0.65511	1.264	0.20781	0.20781	MinT95	34.5573	24.1898	1.429	0.154659	0.154659	MinT95	16.1313	-8.6305	-0.535	0.152638	0.152638	MinT95	1.429	0.154659	0.382266	0.152638
MinT95 ²	-0.32106	0.40464	-0.793	0.42844	0.42844	MinT95 ²	-13.1064	14.968	-0.876	0.47343	0.47343	MinT95 ²	-21.029	-2.7612	0.005759	**	MinT95 ²	-1.427	0.154659	0.382266	0.152638	
MaxH5	2.14298	0.45289	-4.732	4.17E-06	***	MaxH5	-84.6479	17.88	-4.734	4.12E-06	***	MaxH5	-40.4713	14.7794	-2.7384	0.006175	0.006175	MaxH5	-1.427	0.154659	0.382266	0.152638
MaxH5 ²	-1.48432	0.42801	-3.468	0.00064	***	MaxH5 ²	-57.8518	15.8617	-3.647	0.000337	***	MaxH5 ²	-45.3044	24.2803	1.866	0.063499	0.063499	MaxH5 ²	-1.427	0.154659	0.382266	0.152638
MaxH5	1.03856	0.65476	1.586	0.11425	0.11425	MaxH5	11.148	22.8442	1.03	0.304437	0.304437	MaxH5	16.1941	22.4557	0.7212	0.470812	0.470812	MaxH5	-1.427	0.154659	0.382266	0.152638
MinH5	0.84765	0.64492	1.314	0.19021	0.19021	MinH5	-10.0318	16.6336	-0.603	0.54716	0.54716	MinH5	-26.6027	16.2337	-1.6384	0.101337	0.101337	MinH5	-1.427	0.154659	0.382266	0.152638
MinH5 ²	-0.23079	0.45823	-0.504	0.61505	0.61505	MinH5 ²	-20.7957	25.2647	0.823	0.41414	0.41414	MinH5 ²	-14.3864	29.8841	0.4814	0.630227	0.630227	MinH5 ²	-1.427	0.154659	0.382266	0.152638
AP	0.61744	0.69676	0.886	0.37659	*	AP	-1.427	0.154659	0.382266	0.152638	*	AP	-1.427	0.154659	0.382266	0.152638	*	AP	-1.427	0.154659	0.382266	0.152638

Table S2.2 – Pearson correlations (r) of all independent variables available for model selection for each forest type and all plots combined.
 Values greater than $|0.7|$ are in bold. See Table 1 in main text for abbreviations.

All	VarMT	AP	CC	MaxHS	MaxH95	MaxTS	MaxT95	MinH5	MinH95	MinT5	MinT95	PWeQ	PDQ	PWaQ	PCQ
VarMT	1.000	-0.571	-0.024	-0.383	0.324	-0.416	0.314	-0.430	-0.150	-0.433	-0.628	-0.529	-0.531	-0.440	-0.617
AP	-0.571	1.000	0.235	0.556	0.157	-0.031	-0.665	0.674	0.619	0.531	0.236	0.969	0.922	0.934	0.952
CC	-0.024	0.235	1.000	0.091	0.390	-0.531	-0.519	0.579	0.573	0.318	-0.099	0.248	0.194	0.267	0.185
MaxH5	0.556	0.091	1.000	0.210	0.192	-0.275	0.458	0.582	-0.020	0.273	0.517	0.466	0.460	0.558	
MaxH95	0.324	0.157	0.390	1.000	0.761	-0.496	0.429	0.668	-0.238	-0.681	0.156	0.203	0.221	0.106	
MaxT5	-0.416	-0.031	-0.531	0.192	-0.761	1.000	0.626	-0.382	-0.534	0.043	0.757	-0.038	-0.108	-0.116	0.023
MaxT95	0.314	-0.665	-0.519	-0.275	-0.496	0.626	1.000	-0.851	-0.767	-0.409	0.230	-0.626	-0.681	-0.638	-0.636
MinH5	-0.430	0.674	0.579	0.458	0.429	-0.382	-0.851	1.000	0.766	0.420	0.041	0.670	0.597	0.679	0.614
MinH95	-0.150	0.619	0.573	0.582	0.668	-0.534	-0.767	0.766	1.000	0.322	-0.142	0.604	0.549	0.618	0.553
MinT5	-0.433	0.531	0.318	-0.020	-0.238	0.043	-0.409	0.420	0.322	1.000	0.507	0.551	0.394	0.537	0.464
MinT95	-0.628	0.236	-0.099	0.273	-0.681	0.757	0.230	0.041	-0.142	0.507	1.000	0.272	0.042	0.197	0.220
PWeQ	-0.529	0.969	0.248	0.517	0.156	-0.038	-0.626	0.670	0.604	0.551	0.272	1.000	0.811	0.985	0.853
PDQ	-0.531	0.922	0.194	0.466	0.203	-0.108	-0.681	0.597	0.549	0.394	0.042	0.811	1.000	0.755	0.976
PWaQ	-0.440	0.934	0.267	0.460	0.221	-0.116	-0.638	0.679	0.618	0.537	0.197	0.985	0.755	1.000	0.785
PCQ	-0.617	0.952	0.185	0.558	0.106	0.023	-0.636	0.614	0.553	0.464	0.220	0.853	0.976	0.785	1.000

DSF	VarMT	AP	CC	MaxHS	MaxH95	MaxTS	MaxT95	MinH5	MinH95	MinT5	MinT95	PWeQ	PDQ	PWaQ	PCQ
VarMT	1.000	-0.597	-0.026	-0.457	0.358	-0.476	0.360	-0.509	-0.179	-0.454	-0.704	-0.534	-0.573	-0.433	-0.662
AP	-0.597	1.000	0.201	0.640	0.064	0.065	-0.672	0.690	0.626	0.560	0.350	0.968	0.922	0.921	0.957
CC	-0.026	0.201	1.000	0.098	0.385	-0.528	-0.517	0.547	0.536	0.263	-0.092	0.226	0.159	0.245	0.152
MaxH5	-0.457	0.640	0.098	1.000	0.177	0.186	-0.406	0.572	0.616	0.007	0.297	0.603	0.548	0.553	0.629
MaxH95	0.358	0.064	0.385	0.177	1.000	-0.777	-0.434	0.308	0.652	-0.252	-0.692	0.077	0.099	0.136	0.003
MaxT5	-0.476	0.065	-0.528	0.186	-0.777	1.000	0.531	-0.253	-0.511	0.093	0.769	0.043	0.001	-0.030	0.125
MaxT95	0.360	-0.672	-0.517	-0.406	-0.531	1.000	-0.850	-0.433	-0.807	-0.433	0.093	-0.635	-0.670	-0.628	-0.642
MinH5	-0.509	0.690	0.547	0.572	0.308	-0.253	-0.850	1.000	0.777	0.452	0.212	0.682	0.596	0.675	0.636
MinH95	-0.179	0.626	0.536	0.616	0.652	-0.511	-0.807	0.777	1.000	0.296	-0.125	0.623	0.547	0.641	0.547
MinT5	-0.454	0.560	0.263	0.007	-0.252	0.093	-0.433	0.452	0.296	1.000	0.521	0.581	0.434	0.576	0.495
MinT95	-0.704	0.350	-0.092	0.297	-0.692	0.769	0.093	0.212	-0.125	0.521	1.000	0.365	0.185	0.301	0.354
PWeQ	-0.534	0.968	0.226	0.603	0.077	0.043	-0.635	0.682	0.623	0.581	0.365	1.000	0.807	0.982	0.859
PDQ	-0.573	0.922	0.159	0.548	0.099	0.001	-0.670	0.596	0.547	0.434	0.185	0.807	1.000	0.725	0.976
PWaQ	-0.433	0.921	0.245	0.553	0.136	-0.030	-0.628	0.641	0.576	0.301	0.982	0.725	1.000	0.772	
PCQ	-0.662	0.957	0.152	0.629	0.003	0.125	-0.642	0.636	0.547	0.495	0.354	0.859	0.976	0.772	1.000

GW	AP	MaxH95	MaxT5	MaxT95	MinH95	MinT5	MinT95	PWeQ	PDQ	PWaQ	PCQ	
AP	1.000	0.748	-0.642	-0.805	0.841	0.639	-0.065	-0.684	0.974	0.974	0.978	
MaxH95	0.748	1.000	-0.782	-0.748	0.773	0.861	-0.268	-0.782	0.682	0.752	0.689	
MaxT5	-0.642	-0.782	1.000	0.920	-0.680	-0.764	0.053	0.809	-0.546	-0.733	-0.553	
MaxT95	-0.805	-0.748	0.920	1.000	-0.831	-0.729	-0.017	0.752	-0.741	-0.841	-0.744	
MinH95	0.841	0.773	-0.680	-0.831	1.000	0.817	0.096	-0.515	0.858	0.729	0.862	
MinT5	0.639	0.861	-0.764	-0.729	0.817	1.000	0.124	-0.516	0.623	0.558	0.623	
MinT95	-0.065	-0.268	0.053	-0.017	0.096	0.124	1.000	0.505	0.040	-0.233	0.027	
PWeQ	0.974	0.682	-0.782	0.809	0.752	-0.515	0.505	1.000	-0.542	-0.846	-0.556	
PDQ	0.936	0.752	-0.733	-0.841	0.729	0.558	-0.233	-0.846	1.000	0.839	0.998	
PWaQ	0.974	0.689	-0.553	-0.744	0.862	0.623	0.027	-0.556	0.998	0.844	1.000	
PCQ	0.978	0.748	-0.679	-0.824	0.784	0.588	-0.161	-0.765	0.910	0.984	0.912	
<hr/>												
WSF	VarMT	AP	MaxH95	MaxT5	MaxT95	MinH95	MinT5	MinT95	PWeQ	PDQ	PWaQ	PCQ
VarMT	1.000	-0.513	0.560	-0.551	0.094	0.180	-0.536	-0.666	-0.528	-0.341	-0.483	-0.456
AP	-0.513	1.000	-0.230	0.212	-0.412	0.168	0.603	0.296	0.953	0.861	0.934	0.910
MaxH95	0.560	-0.230	1.000	-0.743	-0.335	0.564	-0.622	-0.815	-0.259	-0.030	-0.194	-0.199
MaxT5	-0.551	0.212	-0.743	1.000	0.634	-0.483	0.331	0.814	0.242	-0.007	0.169	0.185
MaxT95	0.094	-0.412	-0.335	0.634	1.000	-0.552	-0.225	0.393	-0.324	-0.549	-0.354	-0.443
MinH95	0.180	0.168	0.564	-0.483	-0.552	1.000	-0.012	-0.274	0.134	0.175	0.146	0.162
MinT5	-0.536	0.603	-0.622	0.331	-0.225	-0.012	1.000	0.632	0.593	0.416	0.558	0.543
MinT95	-0.666	0.296	-0.815	0.844	0.393	-0.274	0.632	1.000	0.380	-0.041	0.316	0.188
PWeQ	-0.528	0.953	-0.259	0.242	-0.324	0.134	0.593	0.380	1.000	0.681	0.993	0.744
PDQ	-0.341	0.861	-0.030	-0.007	-0.549	0.175	0.416	-0.041	0.681	1.000	0.662	0.962
PWaQ	-0.483	0.934	-0.194	0.169	-0.354	0.146	0.558	0.316	0.993	0.662	1.000	0.709
PCQ	-0.456	0.910	-0.199	0.185	-0.443	0.162	0.543	0.188	0.744	0.962	0.709	1.000
<hr/>												
RF	VarMT	AP	MaxH95	MaxT5	MaxT95	MinH95	MinT5	MinT95	PDQ	PCQ		
VarMT	1.000	-0.531	0.496	-0.368	-0.333	-0.474	0.044	-0.494	-0.554			
AP	-0.531	1.000	0.189	0.027	0.535	0.044	0.920	0.917				
MaxH95	0.496	0.189	1.000	-0.591	0.398	-0.667	0.173	0.086				
MaxT5	-0.368	0.027	-0.591	1.000	-0.494	0.768	-0.004	0.134				
MinH95	-0.333	0.535	0.398	-0.494	1.000	-0.290	0.484	0.438				
MinT5	-0.474	0.044	-0.667	0.768	-0.290	1.000	-0.087	0.069				
PDQ	-0.494	0.920	0.173	-0.004	0.484	-0.087	1.000	0.974	1.000			
PCQ	-0.554	0.917	0.086	0.134	0.438	0.069	0.974	1.000				

FW	VarMT						MinT95						MinT95						PWaQ						PCQ					
	AP	MaxH5	MaxH95	MaxT95	MinH5	MinH95	MinT5	MinT95	PWeQ	PDQ	PWaQ	PCQ	AP	MaxH5	MaxH95	MaxT95	MinH5	MinH95	MinT5	MinT95	PWeQ	PDQ	PWaQ	PCQ						
VarMT	1.000	-0.669	-0.364	0.507	0.611	-0.393	-0.620	-0.585	-0.570	-0.746	-0.442	-0.734	-0.669	1.000	-0.449	-0.605	0.303	0.750	0.580	0.945	0.895	0.944	0.944	0.944	0.944					
MaxH5	-0.669	1.000	0.450	-0.449	-0.605	0.303	0.750	0.580	0.961	0.945	0.895	0.944	0.450	-0.246	0.267	0.051	0.332	0.390	0.460	0.313	0.477	0.477	0.477	0.477	0.477					
MaxH95	0.450	-0.364	1.000	0.069	-0.246	-0.246	0.267	0.051	-0.504	-0.558	-0.382	-0.480	-0.321	-0.449	1.000	0.193	0.132	-0.632	-0.183	-0.545	-0.629	-0.482	-0.611	-0.611	-0.611	-0.611				
MaxT95	0.507	-0.449	0.069	1.000	-0.246	0.193	1.000	-0.677	-0.677	-0.632	-0.183	-0.227	-0.227	-0.227	0.132	-0.677	1.000	0.340	0.221	0.279	0.292	0.251	0.313	0.313	0.313	0.313				
MinH5	0.611	-0.605	-0.246	0.267	0.132	-0.677	1.000	0.340	0.221	-0.227	-0.227	-0.227	-0.227	-0.227	0.193	-0.677	1.000	0.340	0.221	0.279	0.292	0.251	0.313	0.313	0.313	0.313				
MinH95	-0.393	0.303	0.267	0.267	0.132	-0.677	1.000	0.340	0.221	-0.227	-0.227	-0.227	-0.227	-0.227	0.193	-0.677	1.000	0.340	0.221	0.279	0.292	0.251	0.313	0.313	0.313	0.313				
MinT5	-0.620	0.750	0.051	-0.504	-0.632	-0.340	1.000	0.640	0.649	0.711	0.655	0.720	0.720	0.720	0.332	-0.558	-0.183	0.221	0.640	1.000	0.517	0.551	0.454	0.609	0.609	0.609				
MinT95	-0.585	0.580	0.580	0.580	-0.558	-0.558	-0.183	0.221	0.640	1.000	0.517	0.551	0.454	0.454	0.454	0.390	-0.382	-0.545	0.279	0.699	0.517	1.000	0.826	0.974	0.825	0.825	0.825			
PWeQ	-0.570	0.961	0.390	-0.382	-0.545	-0.545	-0.545	0.279	0.699	0.517	1.000	0.500	0.500	0.500	0.500	0.500	0.460	-0.480	-0.629	0.292	0.711	0.551	0.826	1.000	0.708	0.991	0.991			
PDQ	-0.746	0.945	0.460	-0.480	-0.629	-0.629	-0.629	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292	-0.482	-0.321	0.251	0.655	0.544	0.974	0.708	1.000	0.703	0.703	0.703			
PWaQ	-0.442	0.895	0.313	-0.321	-0.482	-0.482	-0.482	0.251	0.251	0.251	0.251	0.251	0.251	0.251	0.251	0.251	-0.484	-0.611	0.313	0.720	0.609	0.825	0.991	1.000	0.703	0.703	0.703			
PCQ	-0.734	0.944	0.477	-0.484	-0.611	-0.611	-0.611	0.313	0.313	0.313	0.313	0.313	0.313	0.313	0.313	0.313	-0.484	-0.611	0.313	0.720	0.609	0.825	0.991	1.000	0.703	0.703	0.703			

Table S2.3 – Best fitting models comprising the 95% confidence set for each forest type and all plots combined. Models are ranked according to

QAIC, which measures the relative goodness of fit of each model. Unshaded cells indicate variables excluded from the model in that row. The percentage deviance explained by each model is given by Dev.(%) and model Akaike weights are given by w_i .

ALL											
VarMT	CC	MinT5	MinT95	MaxH5	MaxH95	MinH5	PCQ	Dev. (%)	QAIC	ΔQ_{AIC}	w_i
								16.39	5915.10	0.00	0.507
								16.49	5915.80	0.77	0.345
								16.33	5918.80	3.70	0.080
								16.19	5919.20	4.13	0.064

DSF											
VarMT	CC	MinT5	MaxH5	MaxH95	MinH5	PCQ	Dev. (%)	QAIC	ΔQ_{AIC}	w_i	
								17.61	3985.20	0.00	0.801
								17.35	3988.30	3.19	0.162

WSF									
VarMT	MaxT95	MaxT5	MinT5	MinH95	PDQ	Dev. (%)	QAIC	ΔQ_{AIC}	w_i
						15.92	1071.78	0.00	0.228
						15.76	1072.44	0.67	0.164
						16.57	1073.05	1.27	0.121
						14.48	1073.81	2.03	0.083
						16.37	1073.88	2.11	0.080
						15.34	1074.22	2.44	0.067
						16.21	1074.53	2.75	0.058
						16.00	1075.47	3.70	0.036
						16.82	1075.99	4.22	0.028
						15.78	1076.36	4.58	0.023
						16.62	1076.82	5.05	0.018
						14.62	1077.22	5.45	0.015
						14.53	1077.62	5.84	0.012
						16.40	1077.77	5.99	0.011
						13.50	1077.93	6.15	0.011

RF	VarMT	MinT95	MaxH95	MinH5	PDQ	Dev. (%)	QAIC	ΔQ_{AIC}	w_i
						33.60	219.69	0.00	0.215
						35.42	220.98	1.30	0.113
						35.04	221.54	1.86	0.085
						34.78	221.93	2.24	0.070
						32.08	221.94	2.26	0.070
						29.35	221.99	2.30	0.068
						37.03	222.59	2.90	0.051
						37.02	222.61	2.93	0.050
						31.53	222.75	3.06	0.047
						36.51	223.36	3.67	0.034
						33.8	223.39	3.70	0.034
						33.75	223.46	3.77	0.033
						33.74	223.47	3.78	0.033
						35.50	224.86	5.17	0.016
						38.15	224.93	5.24	0.016
						30.00	225.03	5.35	0.015
						32.53	225.26	5.58	0.013

FW	VarMT	MaxT95	MinT95	MaxH5	MinH5	AP	Dev. (%)	QAIC	ΔQAIC	w _i
						30.39	480.46	0.00	0.392	
						30.40	482.40	1.95	0.148	
						30.48	484.17	3.71	0.061	
						30.44	484.29	3.83	0.058	
						26.27	485.27	4.82	0.035	
						27.22	485.71	5.25	0.028	
						30.54	485.97	5.51	0.025	
						30.48	486.15	5.70	0.023	
						26.64	486.51	6.05	0.019	
						29.72	486.61	6.15	0.018	
						29.05	486.75	6.29	0.017	
						27.17	486.81	6.36	0.016	
						30.85	486.97	6.51	0.015	
						28.32	487.10	6.64	0.014	
						31.10	487.23	6.77	0.013	
						29.51	487.28	6.82	0.013	
						30.12	487.31	6.85	0.013	

G_W	MinT5	MaxH5	MaxH95	PWQ	Dev. (%)	QAIC	Δ QAIC	w_i
					11.11	538.90	0.00	0.526
					11.74	541.20	2.36	0.162
					11.53	541.80	2.90	0.123
					11.42	542.00	3.19	0.107
					12.03	544.40	5.59	0.032

Table S2.4 - Percentage deviance explained by variability in maximum temperature (VarMT) and absolute maximum temperature (MaxT95) as independent predictors of non-tree species richness for each forest type and all plots combined (ALL = all plots combined, DSF = dry sclerophyll forest, WSF = wet sclerophyll forest, GW = grassy woodland, RF = rainforest and FW = forested wetlands). Non-significant results are shown in brackets. Both linear and quadratic parameter estimates and standard errors (χ/χ^2) are provided for quadratic models.

		VarMT			MaxT95		
	%Dev. Exp.	Estimate	SE	%Dev. Exp.	Estimate	SE	
ALL	5.58 (0.57)	3.1/-3.9	0.4/0.4	1.38	1.6/-1.8	0.4/0.4	
GW		-0.4	0.3	5.55	0.2/-1.2	0.3/0.4	
DSF	2.29	0.9/-2.1	0.4/0.4	2.24	-0.7/-2.2	0.4/0.4	
WSF	8.54	1.5/-1.4	0.3/0.3	2.39	1.0	0.3	
RF	6.00	1.2	0.5	6.70	1.2	0.5	
FW	19.70	3.2	0.4	16.90	2.9	0.4	

Table S2.5 - Percentage deviance explained by variability in maximum temperature (VarMT) and absolute maximum temperature (MaxT95) as independent predictors of tree species richness for each forest type and all plots combined (ALL = all plots combined, DSF = dry sclerophyll forest, WSF = wet sclerophyll forest, GW = grassy woodland, RF = rainforest and FW = forested wetlands). Non-significant results are shown in brackets. Both linear and quadratic parameter estimates and standard errors (χ/χ^2) are provided for quadratic models.

		VarMT			MaxT95		
	%Dev. Exp.	Estimate	SE	%Dev. Exp.	Estimate	SE	
ALL	5.06 (0.04)	-5.2/-4.3	0.6/0.6	12.81 (0.10)	-10.2/-2.5	0.6/0.5	
GW		-0.2	0.6		-0.3	0.6	
DSF	3.36	-3.2/-1.4	0.5/0.5	13.10	-6.9/-0.9	0.5/0.5	
WSF	23.19	-3.9/-3.8	0.5/0.5	2.23	-0.1/-1.5	0.5/0.6	
RF	14.78	0.5/-2.9	0.7/0.8	(0.10)	-0.2	0.6	
FW	2.80	1.5	0.6	(1.80)	-1.3	0.6	

Appendix B

Supporting information: Chapter 3

B.1 R code to reproduce simulations

Note, the auxillary file `power_funcs.r` is available at:
www.onlinelibrary.wiley.com/doi/10.1111/2041-210X.12237/suppinfo

Code to reproduce power analysis in ‘Letten and Cornwell (2014) The relative importance of recent and ancient evolution for con- temporary ecology’

Load required packages into library.

```
library(picante)
library(ggplot2)
library(OUwie)
library(phytools)
```

Source functions provided in `power_funcs.r`

```
source('power_funcs.r')
```

Simulate a Yule phylogeny comprising 1000 tips/taxa.

```
tree.sim <- tree.evolve(tree.choice = "sim", pool.size = 1000)
```

Download real tree from <http://datadryad.org/resource/doi:10.5061/dryad.63q27.2/3.1> (Zanne et al. 2013)

```
temp <- tempfile(fileext=".zip")
download.file(
  "http://datadryad.org//bitstream/handle/10255/dryad.55548/PhylogeneticResources.zip",
  temp)
tanktree <- read.tree(unz(temp, 'PhylogeneticResources/Vascular_Plants_rooted.dated.tre'))
unlink(temp)
```

Randomly sample 1000 tips/taxa from real tree.

```
tree.real <- tree.evolve(tree.choice = "real", pool.size = 1000)
```

Simulate trait evolution (1000 reps) on simulated and real tree under Brownian motion.

```
out.mc.bm.sim <- trait.evolve(tree = tree.sim, num.reps = 1000, mode.of.evolution = "BM")
out.mc.bm.real <- trait.evolve(tree = tree.real, num.reps = 1000, mode.of.evolution = "BM")
```

Calculate pairwise phylogenetic distances on both trees.

```
phydist.sim <- cophenetic(tree.sim)
phydist.real <- cophenetic(tree.real)
```

Calculate standardised effect sizes for MNTD and MPD under ‘filtering-derived’ and ‘neutral’ models of community assembly (filtering algorithms follow Kraft et al. 2007). Analyses conducted on both real and simulated trees using raw and square root transformed phylogenetic distances.

Neutral filter (40 taxa sampled randomly with respect to trait value).

```

num.runs <- 999 # number of null model simulations (slow!)

comm.matrix.sim <- make.picante.format.bm(out.mc.bm.sim, ca="neutral")

phy.clust.mntd <- ses.mntd(comm.matrix.sim, phydist.sim, null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)
phy.clust.mntd.sq <- ses.mntd(comm.matrix.sim, sqrt(phydist.sim), null.model = "phylogeny.pool",
                                 abundance.weighted = F, runs = num.runs)

out.df.neut.mntd <- data.frame(phy.clust = phy.clust.mntd$mntd.obs.z,
                                 phy.clust.sq = phy.clust.mntd.sq$mntd.obs.z, metric="mntd",
                                 comm.assem.algo="neutral", tree="sim")

phy.clust.mpd <- ses.mpd(comm.matrix.sim, phydist.sim, null.model = "phylogeny.pool",
                           abundance.weighted = F, runs = num.runs)
phy.clust.mpd.sq <- ses.mpd(comm.matrix.sim, sqrt(phydist.sim), null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)

out.df.neut.mpd <- cbind(phy.clust=phy.clust.mpd$mpd.obs.z,
                          phy.clust.sq = phy.clust.mpd.sq$mpd.obs.z, metric="mpd",
                          comm.assem.algo = "neutral", tree= "sim")

out.df <- rbind(out.df.neut.mntd, out.df.neut.mpd)

comm.matrix.real <- make.picante.format.bm(out.mc.bm.real, ca="neutral")

phy.clust.mntd <- ses.mntd(comm.matrix.real, phydist.real, null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)
phy.clust.mntd.sq <- ses.mntd(comm.matrix.real, sqrt(phydist.real), null.model = "phylogeny.pool",
                               abundance.weighted = F, runs = num.runs)

out.df.neut.mntd <- cbind(phy.clust = phy.clust.mntd$mntd.obs.z,
                           phy.clust.sq = phy.clust.mntd.sq$mntd.obs.z, metric = "mntd",
                           comm.assem.algo = "neutral", tree = "real")

phy.clust.mpd <- ses.mpd(comm.matrix.real, phydist.real, null.model = "phylogeny.pool",
                           abundance.weighted = F, runs = num.runs)
phy.clust.mpd.sq <- ses.mpd(comm.matrix.real, sqrt(phydist.real), null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)

out.df.neut.mpd <- cbind(phy.clust = phy.clust.mpd$mpd.obs.z,
                          phy.clust.sq = phy.clust.mpd.sq$mpd.obs.z, metric="mpd",
                          comm.assem.algo="neutral",tree="real")

out.df <- rbind(out.df, out.df.neut.mntd, out.df.neut.mpd)

```

‘Derived filter’ (takes the maximum simulated trait value and selects 40 taxa with nearest trait values).

```

comm.matrix.sim <- make.picante.format.bm(out.mc.bm.sim, ca="filter")

phy.clust.mntd <- ses.mntd(comm.matrix.sim, phydist.sim, null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)
phy.clust.mntd.sq <- ses.mntd(comm.matrix.sim, sqrt(phydist.sim), null.model = "phylogeny.pool",

```

```

abundance.weighted = F, runs = num.runs)

out.df.filt.mntd <- data.frame(phy.clust = phy.clust.mntd$mntd.obs.z,
                                phy.clust.sq = phy.clust.mntd.sq$mntd.obs.z,
                                metric = "mntd", comm.assem.algo = "filter", tree="sim")

phy.clust.mpd <- ses.mpd(comm.matrix.sim, phydist.sim, null.model = "phylogeny.pool",
                           abundance.weighted = F, runs = num.runs)
phy.clust.mpd.sq <- ses.mpd(comm.matrix.sim, sqrt(phydist.sim), null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)

out.df.filt.mpd <- data.frame(phy.clust = phy.clust.mpd$mpd.obs.z,
                                phy.clust.sq = phy.clust.mpd.sq$mpd.obs.z,
                                metric = "mpd", comm.assem.algo = "filter", tree = "sim")

out.df <- rbind(out.df,out.df.filt.mntd,out.df.filt.mpd)

comm.matrix.real <- make.picante.format.bm(out.mc.bm.real,ca="filter")

phy.clust.mntd <- ses.mntd(comm.matrix.real, phydist.real, null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)
phy.clust.mntd.sq <- ses.mntd(comm.matrix.real, sqrt(phydist.real), null.model = "phylogeny.pool",
                               abundance.weighted = F, runs = num.runs)

out.df.filt.mntd <- data.frame(phy.clust = phy.clust.mntd$mntd.obs.z,
                                 phy.clust.sq = phy.clust.mntd.sq$mntd.obs.z,metric="mntd",
                                 comm.assem.algo="filter",tree="real")

phy.clust.mpd <- ses.mpd(comm.matrix.real, phydist.real, null.model = "phylogeny.pool",
                           abundance.weighted = F, runs = num.runs)
phy.clust.mpd.sq <- ses.mpd(comm.matrix.real, sqrt(phydist.real), null.model = "phylogeny.pool",
                             abundance.weighted = F, runs = num.runs)

out.df.filt.mpd <- data.frame(phy.clust = phy.clust.mpd$mpd.obs.z,
                                phy.clust.sq = phy.clust.mpd.sq$mpd.obs.z,
                                metric="mpd", comm.assem.algo = "filter", tree = "real")

out.df<-rbind(out.df,out.df.filt.mntd,out.df.filt.mpd)

out.df$phy.clust<-as.numeric(out.df$phy.clust)
out.df$phy.clust.sq<-as.numeric(out.df$phy.clust.sq)

```

Quantify difference in standadized effect sizes (MNTD/MPD) when using raw vs square-root transformed phylogenetic distance.

```
out.df$diff <- out.df$phy.clust.sq-out.df$phy.clust
```

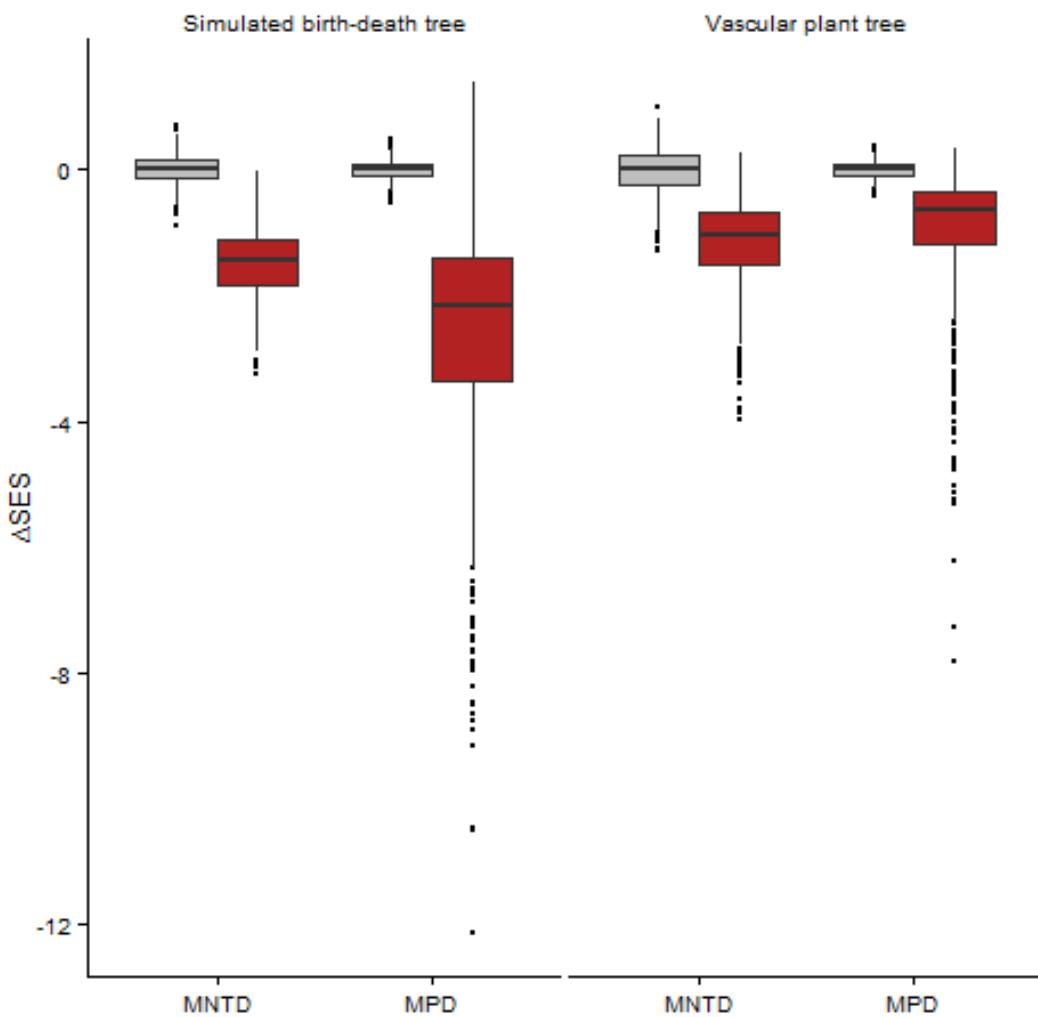
Generate Figure 4 from manuscript.

```
out.df.fig <- out.df
levels(out.df.fig$tree) <- c("Simulated birth-death tree", "Vascular plant tree")
levels(out.df.fig$metric) <- c("MNTD", "MPD")
```

```

p <- ggplot(out.df.fig,aes(x = metric, y = diff, fill = comm.assem.algo))
p + geom_boxplot(outlier.size = 1) + theme_bw() + facet_grid(.~tree) +
  theme(legend.key = element_blank(), legend.position= "none",
        axis.line = element_line(colour = "black"),
        legend.key = element_blank(),
        axis.line = element_line(colour = "black"),
        legend.key = element_blank(),
        panel.border = element_blank(),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        strip.background = element_rect(colour = 'NA', fill = 'NA')) +
  scale_fill_manual(values = c("grey", "firebrick")) +
  xlab("") + ylab(expression(Delta, "SES", sep = ""))

```



Appendix C

Supporting information: Chapter 4

C.1 Phylogeny construction details

The community phylogeny was constructed using the program phyloGenerator (Pearse and Purvis 2013) and its dependency programs as follows. DNA sequence data for two commonly used plastid gene regions (*rbcL* and *matK*) was searched for on GenBank (Benson et al. 2011). Of the 49 taxa in the pool, 30 species were represented, with a further 16 represented by congeneric taxa. Sequences were aligned using MAFFT (Katoh et al. 2005), and the community phylogeny (with divergence times) was estimated under a Bayesian framework using BEAST (Drummond and Rambaut 2007). A constraint tree, generated in Phylomatic (Webb and Donoghue 2005), and dated using the BLADJ algorithm of Phylocom (Webb, Ackerly, and Kembel 2008), was used to place strong priors on the ages and topology of existing highly supported clades. Four independent runs were performed using a GTR model assuming a lognormal relaxed clock with four rate categories, with each run comprising an MCMC chain run for 50,000,000 generations and sampled every 1,000 generations. After checking parameter statistics in TRACER (<http://tree.bio.ed.ac.uk/software/tracer/>) and removing burnins of 10-20%, the four independent runs were combined to generate a maximum clade credibility tree. Those taxa represented by congeners were then manually added to the tree, as were the three unrepresented taxa on the basis of their position on the Phylomatic derived constraint tree.

Phylogeny for all 49 species in Newick tree format

((((Gonocarpus micranthus: 32.78964725, Gonocarpus tetragynus: 32.78964725): 124.9441371, (((Drosera peltata: 19.27769113, Drosera auriculata: 19.27769113): 19.27769113, Drosera pygmaea: 38.55538226): 73.73868839, Drosera spatulata: 112.2940706): 20.52500615, (Stylium lineare: 112.3064351, (Goodenia stelligera: 74.95393085, Goodenia dimorpha: 74.95393085): 37.35250429): 20.51264166): 24.91470752): 35.6884448, (Burchardia umbellata: 150.2255016, (((Hypolaena fastigiata: 45.02808262, ((Baloskion gracile: 23.43160533, Eurychorda complanata: 23.43160533): 6.005965644, (Empodium minus: 12.79282614, Leptocarpus tenax: 12.79282614): 16.64474484): 9.335195904, Lepyrodia scariosa: 38.77276688): 6.25531574): 14.99415044, (((Plinthanthesis paradoxa: 14.96742327, Aristida warburgii: 14.96742327): 10.6327861, (Themeda australis: 10.795227, Entolasia stricta: 10.795227): 14.80498236): 12.42401627, Tetrarrhena turfosa: 38.02422563): 7.018168824, Austrostipa pubescens: 45.04239445): 14.9798386): 16.64744303, (((Tricostularia pauciflora: 42.6973992, (Lepidosperma neesii: 11.2529451, Tetraria capillaris: 11.2529451): 18.2730358, (Ptilothrix deusta: 12.26655391, (Schoenus brevifolius: 6.133276955, Schoenus ericerorum: 6.133276955, Schoenus imberbis: 6.133276955, Schoenus lepidosperma: 6.133276955, Schoenus moorei: 6.133276955): 6.133276955): 17.25942698): 13.17141831): 13.29991705, Cyathochaeta diandra: 55.99731626): 10, Xyris gracilis: 65.99731626): 10.67235984): 38.0120084, Haemodorum corymbosum: 114.6816845): 14.17530244, ((Blandfordia nobilis: 105.2018214, ((Thysanotus juncifolius: 31.32190603, (Sowerbaea juncea: 15.35052634, (Lomandra obliqua: 7.67526317, (Lomandra glauca: 4, Lomandra cylindrica: 4): 3.67526317): 7.67526317): 15.97137969): 15.40653782, (Xanthorrhoea resinosa: 31.31985316, (Thelionema umbellatum: 15.65992658, Caesia parviflora: 15.65992658): 15.65992658): 15.40859068): 28.21501178, (Patersonia sericea: 46.94615343, Patersonia fragilis: 46.94615343): 27.9973022): 30.25836574): 9.576595384, (Prasophyllum brevilabre: 106.9427746, Cryptostylis subulata: 106.9427746): 7.835642119): 14.07857018): 21.36851467): 43.1967275): 1.625295448, (Cassytha glabella: 24.29827518, Cassytha pubescens: 24.29827518): 170.7492493);

References

- Pearse, W. D. & Purvis, A. 2013 phyloGenerator: an automated phylogeny generation tool for ecologists. *Methods in Ecology and Evolution*, 4(7), 692-698.
- Drummond, A. J. & Rambaut, A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*, 7(1), 214.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. GenBank. *Nucleic acids research*, 39(Database issue), D32-7.
- Katoh, K., Kuma, K.-i., Toh, H. & Miyata, T. 2005 MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic acids research*, 33(2), 511-8.
- Webb, C. O. & Donoghue, M. 2005 Phylomatic: tree assembly for applied phylogenetics. *Molecular Ecology Notes*, 5(1), 181-183.
- Webb, C. O., Ackerly, D. D. & Kembel, S. W. 2008 Phylocom: soft654 ware for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* (Oxford, England), 24(18), 2098-100.

C.2 Equations for calculating D_{nn} and D_{pw}

Phylogenetic and functional nearest-neighbour dissimilarity (D_{nn} , the beta diversity analogue of MNTD) is given by:

$$D_{nn} = f_A \sum_{i=1}^{S_A} f_i \text{minff}_{ib} + f_B \sum_{j=1}^{S_B} f_j \text{minff}_{ja}$$

where S_A is the number of species in the community at time A, S_B is the number of species in the community at time B, minff_{ib} is the phylogenetic or functional distance of species i at time B to its nearest neighbour at time A, minff_{ja} is the phylogenetic or functional distance of species j at time B to its nearest neighbour at time A, f_i is the relative abundance of species i in the community at time A, and finally f_j is the relative abundance of species j in the community at time B.

Similarly, pairwise dissimilarity (D_{pw} , the beta diversity analogue of MPD) is given by:

$$D_{pw} = f_A \sum_{i=1}^{S_A} f_i \overline{\delta_{ib}} + f_B \sum_{j=1}^{S_B} f_j \overline{\delta_{ja}}$$

where terms shared with D_{nn} are equivalent, and $\overline{\delta_{ib}}$ is the mean pairwise phylogenetic or functional distance between species i in the community at time A and all species in the community at time B, and $\overline{\delta_{ja}}$ is the mean pairwise phylogenetic or functional distance between species j in the community at time B and all species in the community at time A. Note that as for MNTD and MPD, phylogenetic distance was first square-root transformed before input into both D_{nn} and D_{pw} .

References

- Ricotta, C. & Burrascano, S. 2009 Testing for differences in beta diversity with asymmetric dissimilarities. *Ecological Indicators*, 9(4), 719-724.
- Swenson, N. G., Stegen, J. C., Davies, S. J., Erickson, D. L., Forero-Montaña, J., Hurlbert, A. H., Kress, W. J., Thompson, J., Uriarte, M. et al. 2012 Temporal turnover in the composition of tropical tree communities: functional determinism and phylogenetic stochasticity. *Ecology*, 93(3), 490-499.

C.3 Phylogenetic signal in traits

Table S4.1. Phylogenetic signal in functional traits quantified using either Blomberg's K statistic for continuously defined traits or the 'Fixed Tree, Character Randomly Reshuffled' model of Maddison & Slatkin (1991) for traits coded ordinally.

Trait	K statistic	p-value
$\log_{10}(\text{seed weight})$	0.385	0.001
$\log_{10}(\text{maximum height})$	0.647	0.001
Fire response	-	< 0.001
Raunkiær life-form	-	0.005
Fecundity	-	0.003
Longevity	-	< 0.001
Seedbank persistence	-	< 0.001

C.4 Temporal change in phylogenetic and functional community structure: supplementary figures

The following figures, which correspond with Figure 2 in the main article, show phylogenetic and functional community structure through time when the species pool is constrained to only include monocots or only Poales. Trendlines correspond to models of MNTD/MPD/F-MNTD/F-MPD vs. time for all plots/sites through the first four years of sampling (solid line); plots/sites that only burnt in 1994 (dashed line) and plots that burnt in 1994 and 2001 (dotted lines). Trend lines shaded black indicate significant slope coefficients at $p < 0.05$; grey lines indicate insignificant slopes.

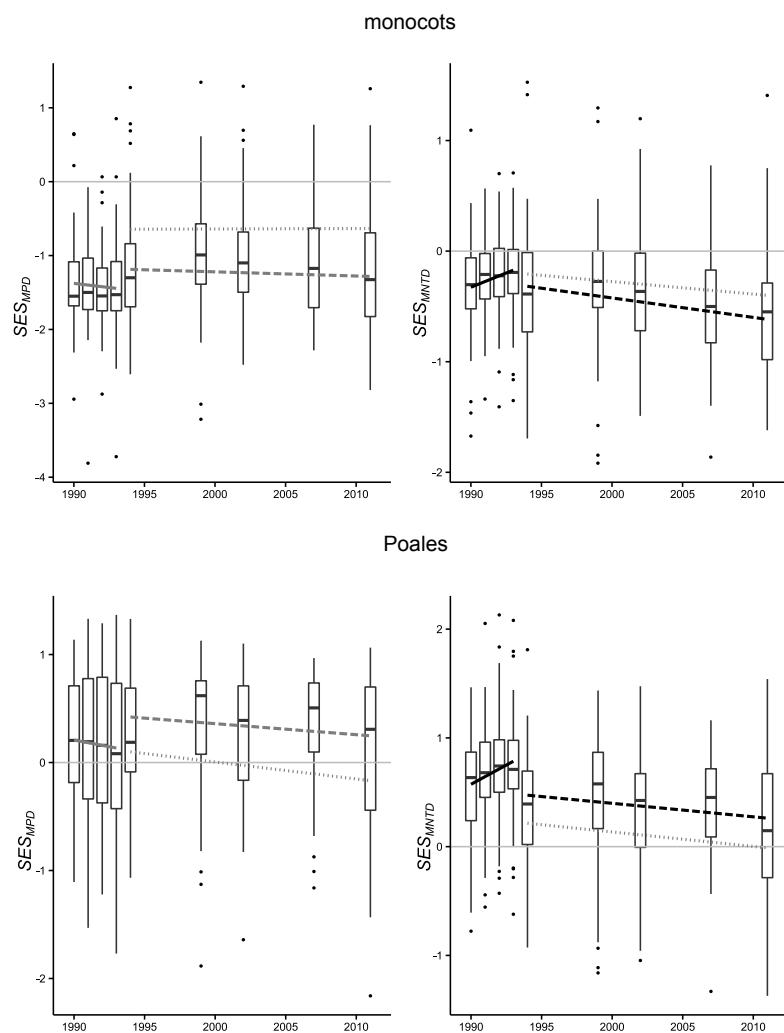


Figure S4.1. Phylogenetic community structure of monocots and Poales through time at the plot scale.

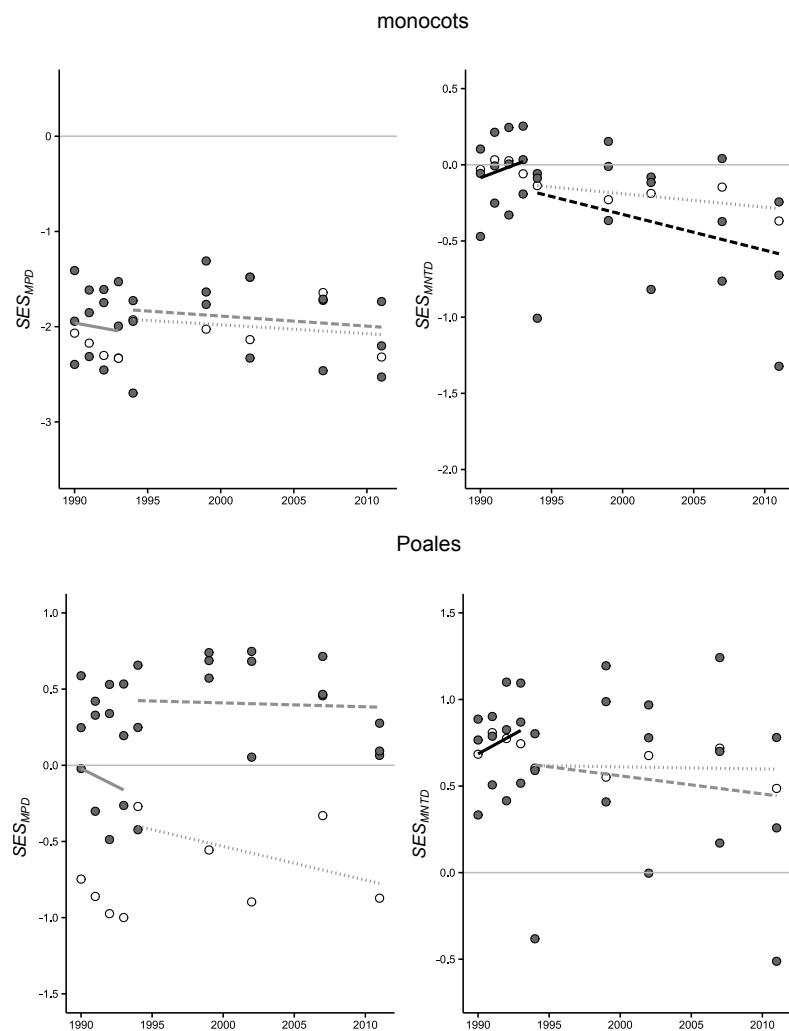


Figure S4.2. Phylogenetic community structure of monocots and Poales through time at the site scale.

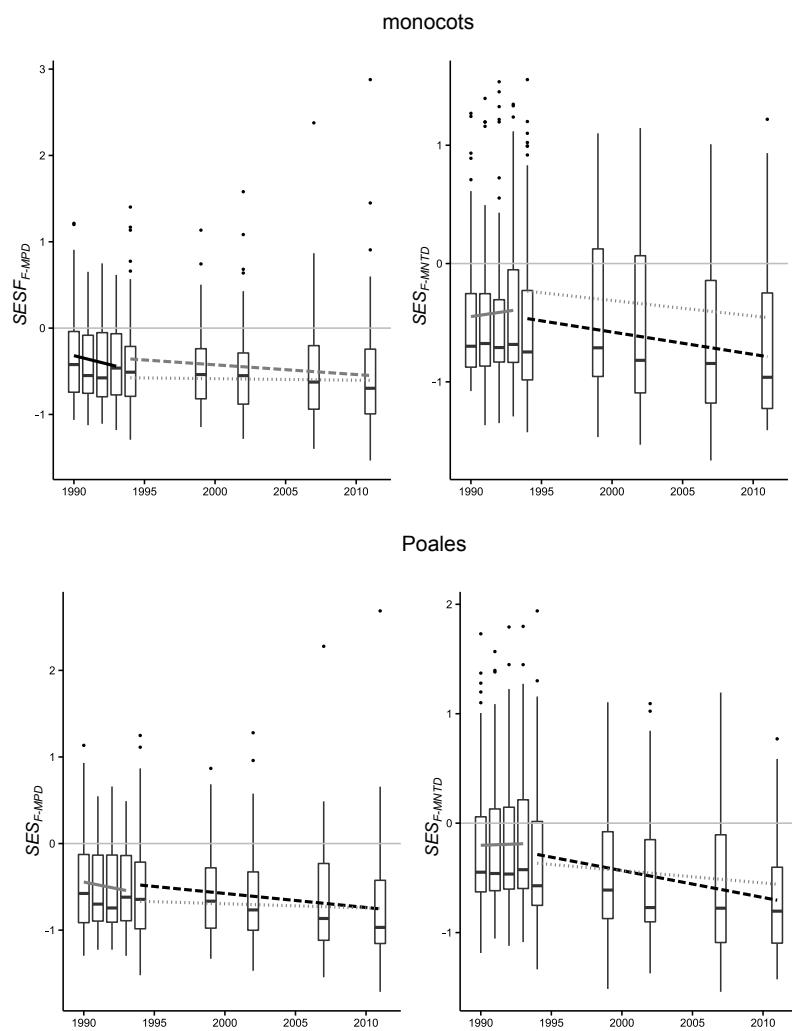


Figure S4.3. Functional community structure of monocots and Poales through time at the plot scale.

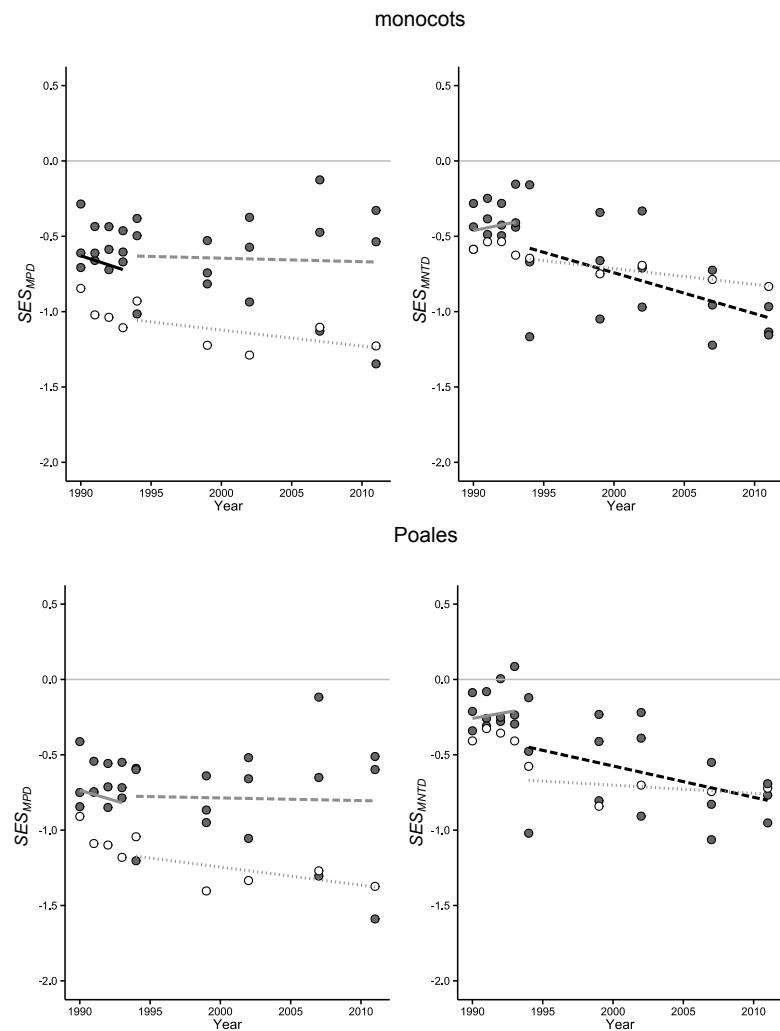


Figure S4.4. Functional community structure of monocots and Poales through time at the site scale.

C.5 Temporal phylogenetic and functional beta turnover: supp. figs.

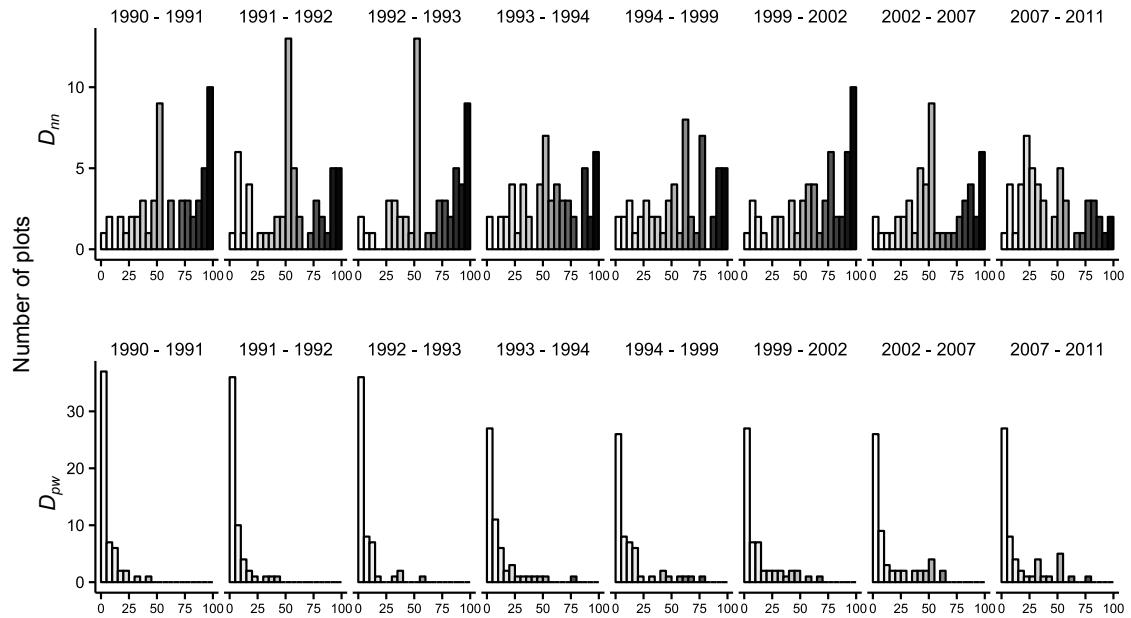


Figure S4.5. Temporal phylogenetic beta turnover quantified by D_{nn} and D_{pw} relative to the immediately preceding census point. Low quantile scores (white) indicate low turnover in phylogenetic composition relative to the observed rate of taxonomic turnover; high quantile scores (black) indicate high turnover in phylogenetic composition relative to the observed rate of taxonomic turnover. Values <2.5 or >97.5 are significant at the 0.05 level.

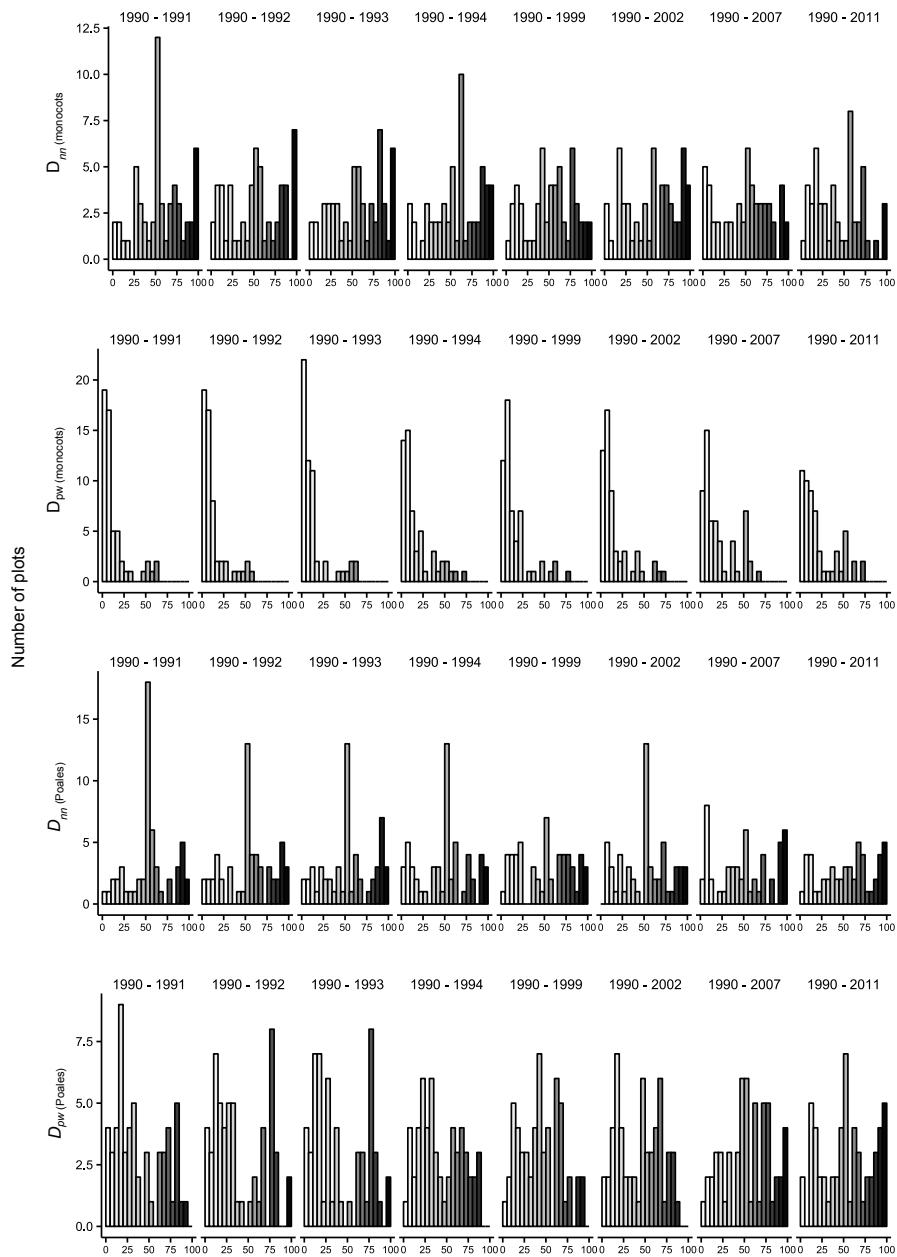


Figure S4.6. Temporal phylogenetic beta turnover (monocots and Poales) quantified by D_{nn} and D_{pw} relative to the first census point. Low quantile scores (white) indicate low turnover in phylogenetic composition relative to the observed rate of taxonomic turnover; high quantile scores (black) indicate high turnover in phylogenetic composition relative to the observed rate of taxonomic turnover. Values <2.5 or >97.5 are significant at the 0.05 level.

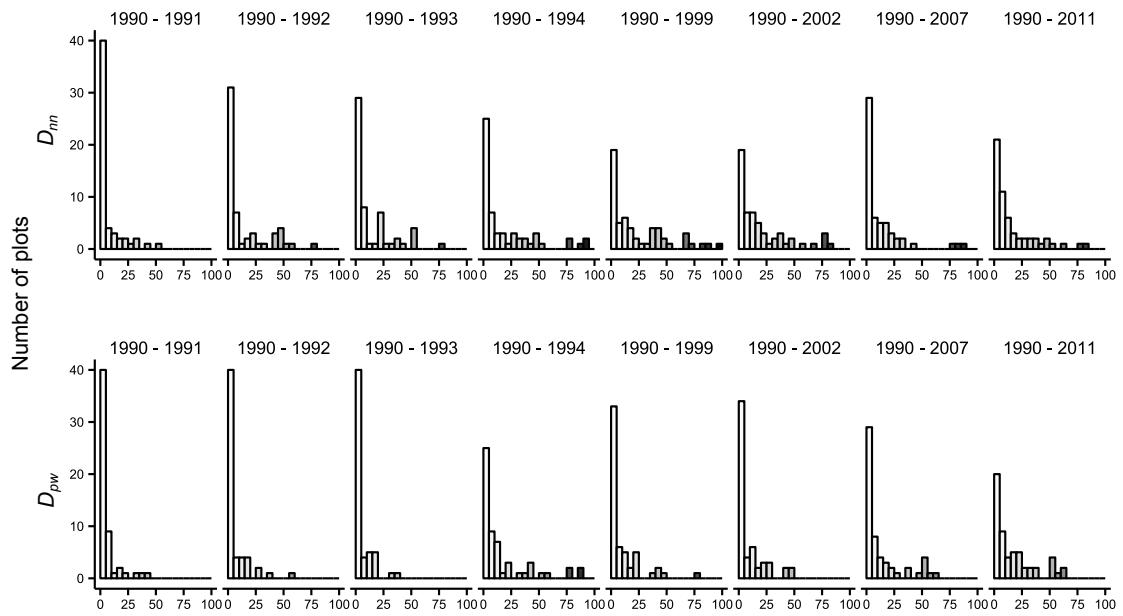


Figure S4.7. Temporal functional beta turnover quantified by D_{nn} and D_{pw} relative to the first census point. Low quantile scores (white) indicate low turnover in phylogenetic composition relative to the observed rate of taxonomic turnover; high quantile scores (black) indicate high turnover in phylogenetic composition relative to the observed rate of taxonomic turnover. Values <2.5 or >97.5 are significant at the 0.05 level.

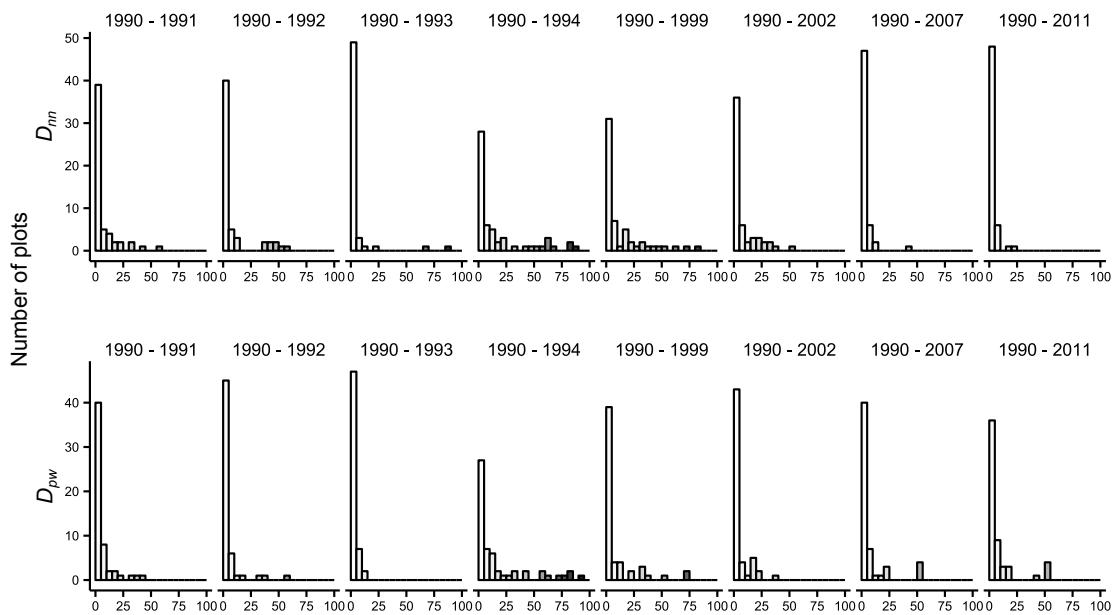


Figure S4.8. Temporal functional beta turnover quantified by D_{nn} and D_{pw} relative to the immediately preceding census point. Low quantile scores (white) indicate low turnover in phylogenetic composition relative to the observed rate of taxonomic turnover; high quantile scores (black) indicate high turnover in phylogenetic composition relative to the observed rate of taxonomic turnover. Values <2.5 or >97.5 are significant at the 0.05 level.

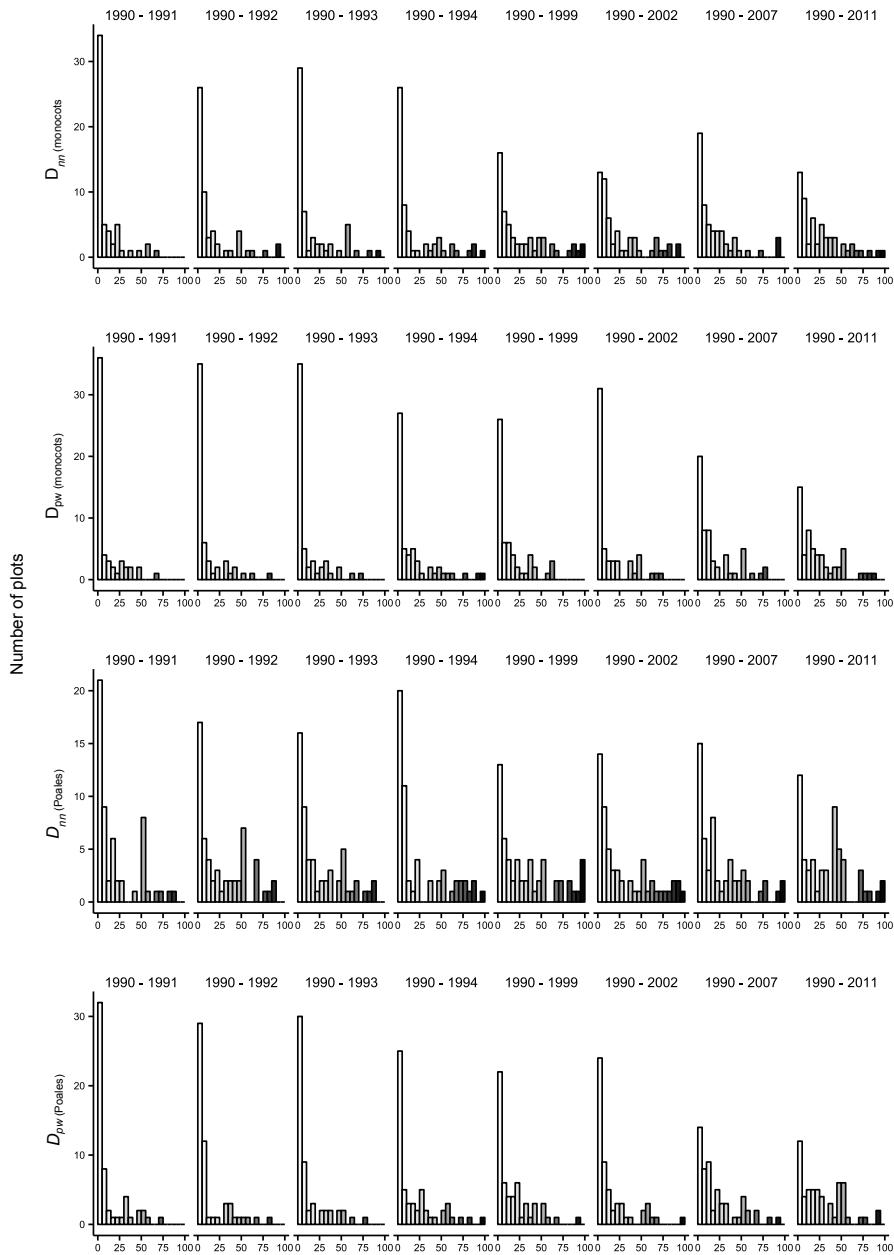


Figure S4.9. Temporal functional beta turnover (monocots and Poales) quantified by D_{nn} and D_{pw} relative to the first census point. Low quantile scores (white) indicate low turnover in phylogenetic composition relative to the observed rate of taxonomic turnover; high quantile scores (black) indicate high turnover in phylogenetic composition relative to the observed rate of taxonomic turnover. Values <2.5 or >97.5 are significant at the 0.05 level.

C.6 Species accumulation curve

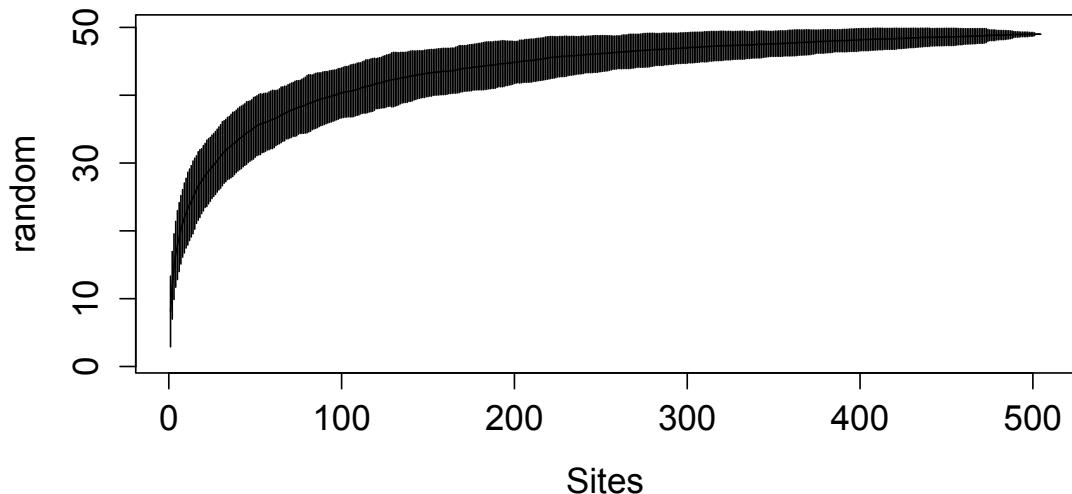


Figure S4.10. Species accumulation curve obtained from 100 permutations of site in random order. Width denotes standard error.

Appendix D

Supporting information: Chapter 5

D.1 Supplementary figures and tables

Figure S5.1. Positive pairwise species correlations to the environment derived from single-predictor LVMs for the full dataset. Connecting lines between species nodes denote positive mean posterior correlations with credible intervals excluding zero. Line colour and thickness indicates the strength of the positive correlation where darker and thicker lines are closer to 1. Species node labels combine the first letters of the genus and specific epithet.

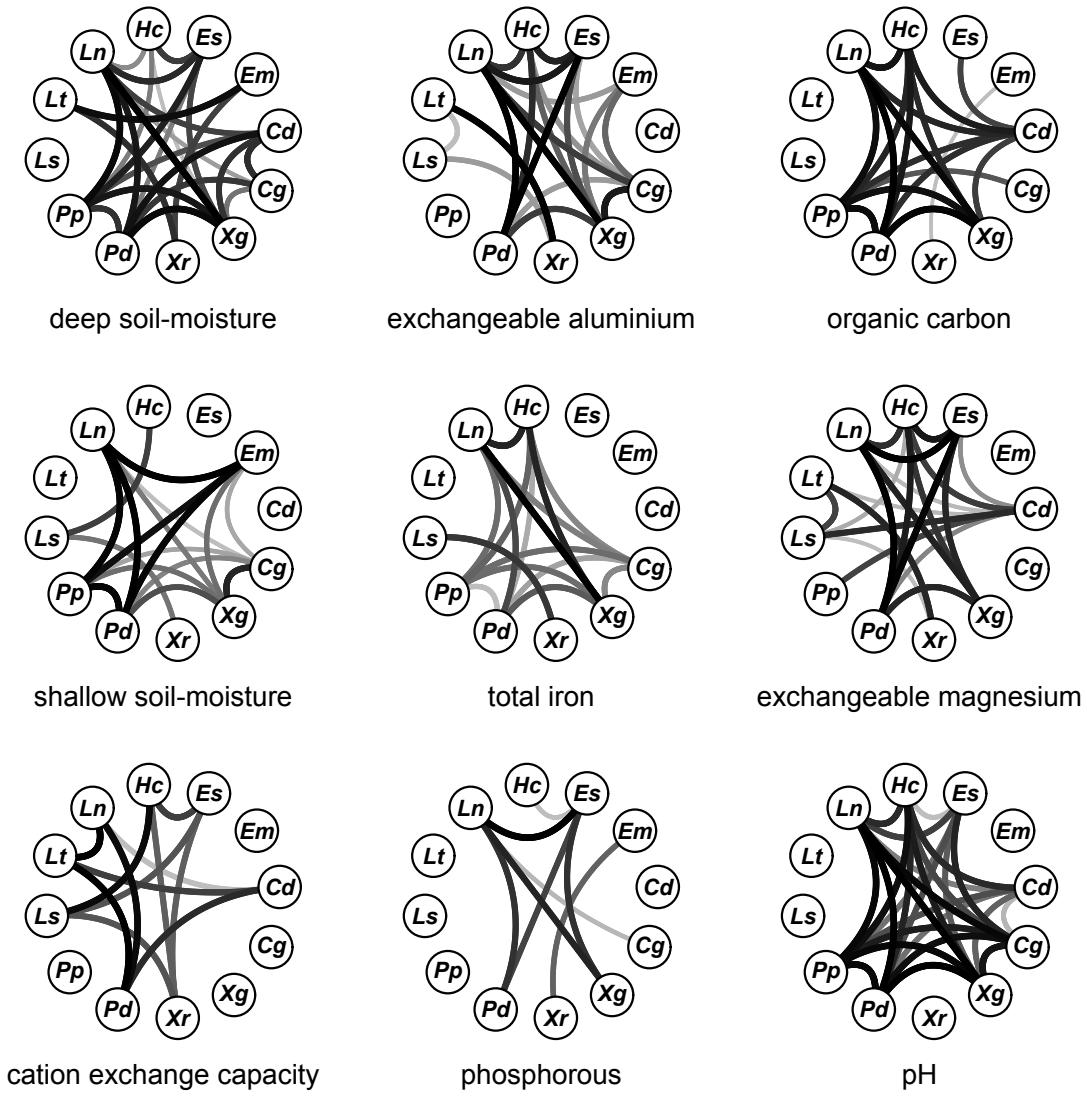


Figure S5.2. Abundance of each species (excl. *Empodiuma minus*) through time. Trend lines obtained with thin-plate splines in a GAMM framework (see main text for model fitting). Shaded region represents 95% confidence intervals.

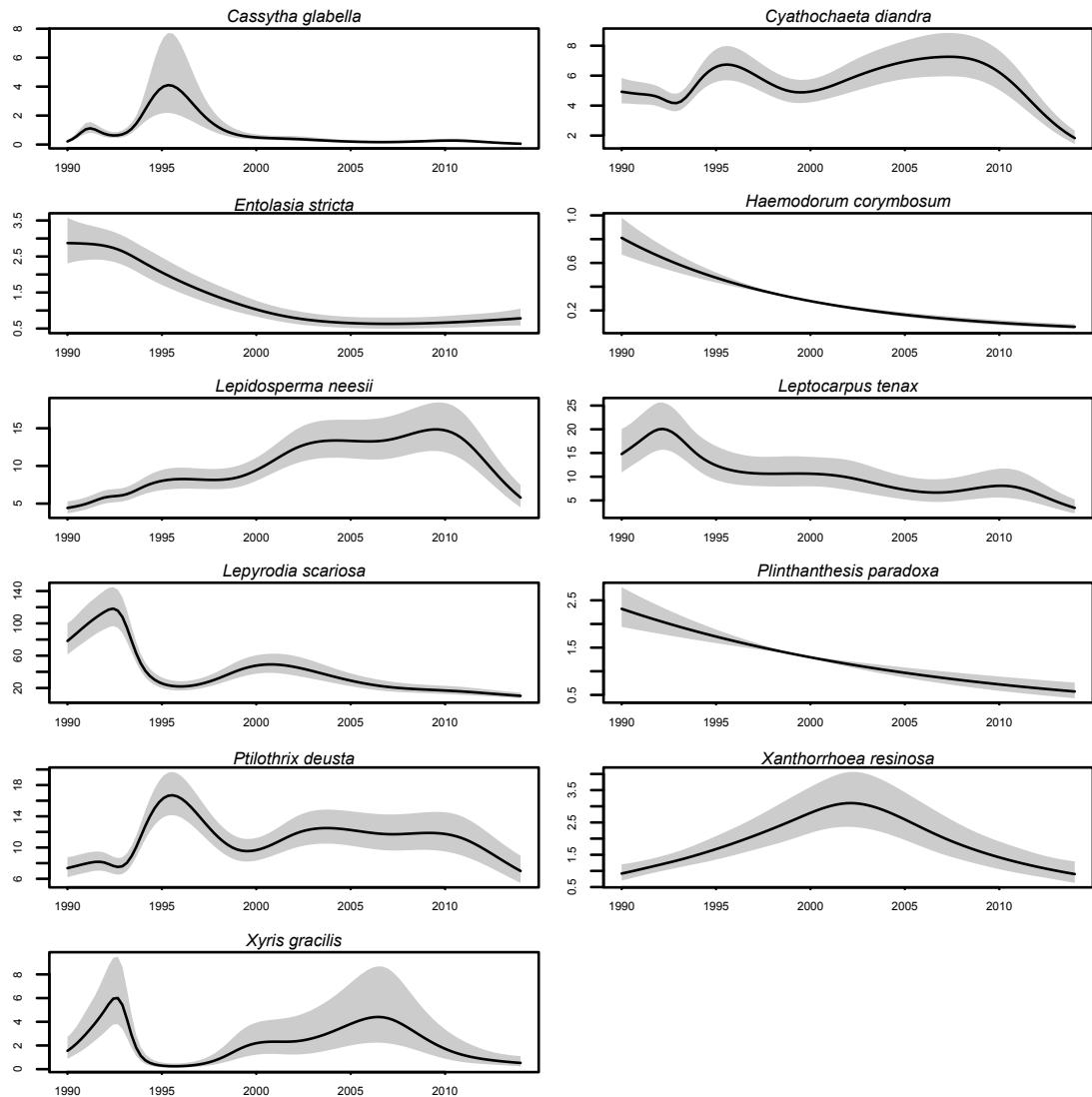


Table S5.1. Negative and positive pairwise environmental and residual correlations in each of the four plot clusters.

	No. sig. env. correlations		No. sig. res. correlations	
	-ve	+ve	-ve	+ve
deep soil-moisture	6,0,1,0	12,2,4,13	8,11,7,0	10,14,30,41
shallow soil-moisture	2,0,4,0	3,3,2,0	10,10,5,0	17,9,16,42
pH	0,0,3,0	1,1,11,0	7,10,6,0	11,13,23,41
total iron	2,0,2,0	2,1,3,1	10,11,7,0	6,8,17,41
phosphorous	4,0,9,3	3,5,7,3	12,11,4,0	7,10,26,34
organic carbon	1,0,6,0	0,0,10,16	8,10,6,0	6,12,27,39
exchangeable aluminium	2,0,2,6	4,8,2,12	10,16,5,0	15,20,25,41
exchangeable magnesium	3,0,5,1	3,1,9,8	10,11,6,0	11,13,28,41
cation exchange capacity	1,0,5,1	2,3,6,0	8,13,5,0	14,10,21,43

Primer on modelling species co-occurrence using latent variable models (LVM)

Supplement to ‘Letten, Keith, Tozer and Hui (2015) Fine-scale hydrological niche differentiation through the lens of multi-species co-occurrence models’

In this brief primer we demonstrate the efficacy of latent variable models (LVM) for partitioning out the different drivers of species co-occurrence patterns into that which is attributable to shared/diverging environmental responses and that which may be attributable to unmeasured covariates and/or biotic interactions.

To this end, we first simulate multivariate abundance data along two hypothetical environmental gradients (pH and soil moisture), and then fit two LVMs using each of the gradients independently as model predictors. We then derive environmental and ‘residual’ correlation matrices from the fitted regression and latent variable values respectively to show that the residual correlations induced by the latent variables closely correspond to the correlations due to the environmental response of the missing covariate. This reflects the ability of LVMs to perform inference which is robust to missing covariates/biotic interactions.

Before proceeding, make sure JAGS is installed (<http://mcmc-jags.sourceforge.net/>), as well as the following packages (note that R2jags is the only critical package for fitting LVMs, with the remainder required for data simulation, data wrangling, aesthetics, and inference). You will also need to load `fitLVM-auxiliaryfunctions.r` which includes a number of functions required for post-processing / analysis.

```
packages <- c("R2jags", "coenocliner",
             "ggplot2", "tidyverse",
             "vegan", "corrplot",
             "MASS", "gridExtra")
if (length(setdiff(packages, rownames(installed.packages()))) > 0) {
  install.packages(setdiff(packages, rownames(installed.packages())))
}

sapply(packages, require, character.only = TRUE)
source("fitLVM-auxiliaryfunctions.r")
```

Data simulation

Simulate data using the `coenocliner` (<http://cran.r-project.org/web/packages/coenocliner/vignettes/coenocliner.pdf>) package, which, like it says on the tin, “can be used to generate random count or occurrence data from parametrised species response curve”.

The very first step is to define the simulation parameters.

```
sp <- 6 # number of species

# Gradient 1: pH
grad1.locs <- seq(1, 14, length = 14) # pH gradient locations
grad1.opt <- runif(sp, min = min(grad1.locs), max = max(grad1.locs)) # species optima along gradient
grad1.tol <- ceiling(runif(sp, min = 0.5, max = 3)) # niche widths
h <- ceiling(rlnorm(sp, meanlog = 3)) # max abundances (only defined once)
grad1.par <- cbind(opt = grad1.opt, tol = grad1.tol, h = h) # combine into a matrix

# Gradient 2: soil moisture (% volume)
grad2.locs <- seq(1, 100, length = 20) # soil moisture gradient locations
```

```

grad2.opt <- runif(sp, min = min(grad2.locs), max = max(grad2.locs)) # species optima
grad2.tol <- ceiling(runif(sp, min = 5, max = 30)) # niche widths
grad2.par <- cbind(opt = grad2.opt, tol = grad2.tol) # combine into a matrix

pars <- list(px = grad1.par, py = grad2.par) # combine parameters for both gradients

```

We now generate a matrix of simulated species abundances using `coenocline` and visualise their bivariate distribution along the two gradients with `persp`. Note that each simulation run will generate different bivariate relationships.

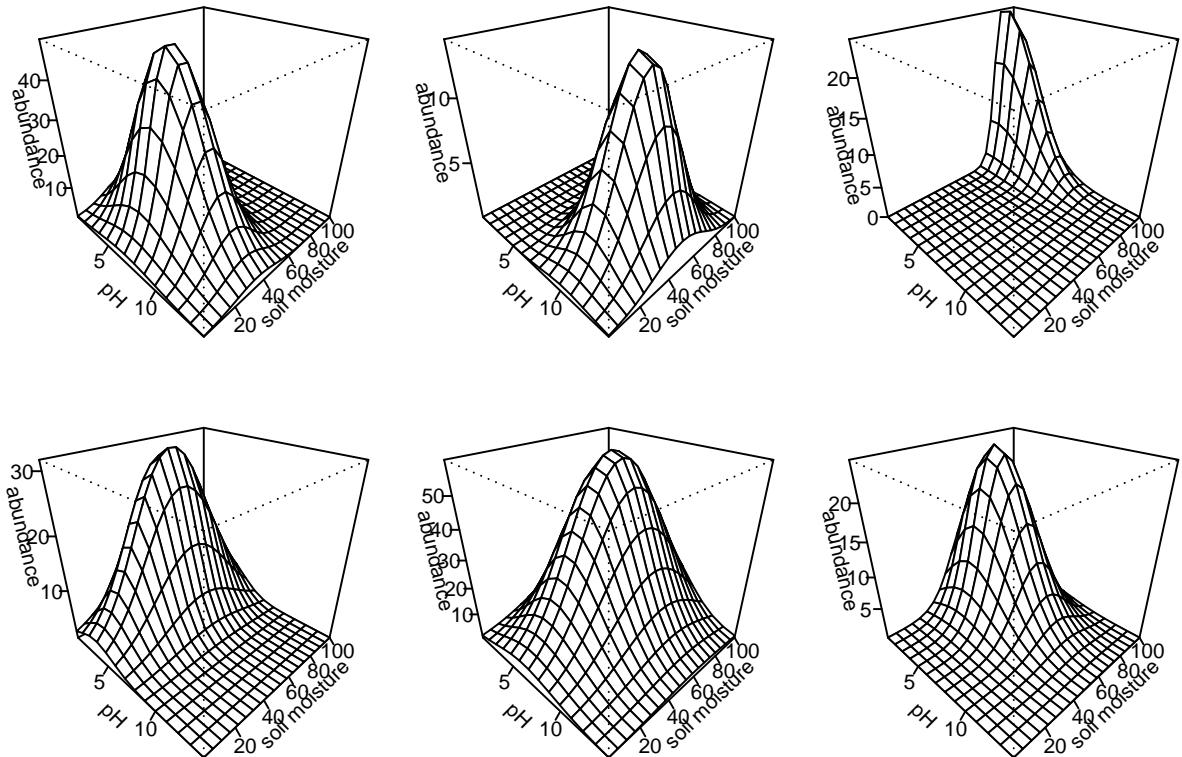
```

# List parameters
locs <- expand.grid(x = grad1.locs, y = grad2.locs) # put gradient locations together

multi.sim <- coenocline(locs, responseModel = "gaussian",
                        params = pars, extraParams = list(corr = 0),
                        expectation = TRUE)

layout(matrix(1:6, ncol = 3))
op <- par(mar = rep(1, 4))
for (i in c(1:6)) {
  persp(grad1.locs, grad2.locs, matrix(multi.sim[, i], ncol = length(grad2.locs)),
        ticktype = "detailed", zlab = "abundance", xlab = "pH", ylab = "soil moisture",
        theta = 45, phi = 30)
}

```



Simulate count (species abundance) data with negative-binomial errors.

```
multi.sim.count <- coenocline(locs, responseModel = "gaussian",
                               params = pars, extraParams = list(corr = 0),
                               expectation = FALSE, countModel = "negbin",
                               countParams = list(alpha = 1))
```

Model fitting

Bundle data into a JAGS friendly format.

```
grad1.X = model.matrix(~ poly(locs[,1],2) -1)

sim.data1 <- list(y    = as.matrix(multi.sim.count),
                  X    = grad1.X,
                  n.cov = ncol(grad1.X),
                  n.sites = nrow(multi.sim.count),
                  n.species = sp)

grad2.X = model.matrix(~ poly(locs[,2],2) -1)

sim.data2 <- list(y    = as.matrix(multi.sim.count),
                  X    = grad2.X,
                  n.cov = ncol(grad2.X),
                  n.sites = nrow(multi.sim.count),
                  n.species = sp)
```

Write JAGS model.

```
cat(
'
model {

## Observation level ##
for(i in 1:n.sites) { for(j in 1:n.species) {
  eta[i,j] <- spp.int[j] + inprod(X[i, ],spp.coef[j,]) +
    inprod(LV[i, ],lv.coef[j,])

## Parameterization of negbin as Poisson with multiplicative gamma random effect
## Var = mu + phi*mu^2
u[i,j] ~ dgamma(1/spp.phi[j],1/spp.phi[j])
y[i,j] ~ dpois(exp(eta[i,j])*u[i,j])
}

## Latent variables ##
for(i in 1:n.sites) { LV[i,1] ~ dnorm(0,1); LV[i,2] ~ dnorm(0,1) }
lv.coef[1,2] <- 0 ## Constraints
lv.coef[1,1] ~ dunif(0,30) ## Constraints
lv.coef[2,2] ~ dunif(0,30) ## Constraints
lv.coef[2,1] ~ dnorm(0,0.1)
for(j in 3:n.species) { lv.coef[j,1] ~ dnorm(0,0.01); lv.coef[j,2] ~ dnorm(0,0.01) }

## Process level ##
for(j in 1:n.species) {
  ## Draw intercepts and overdispersion parameter independently
  spp.int[j] ~ dnorm(0,0.01); spp.phi[j] ~ dunif(0,30)
  ## Draw spp slopes from common distribution
  for(l in 1:n.cov) {spp.coef[j,l] ~ dnorm(mu.beta[l], tau.beta[l])}
}

## Prior level ##
for(l in 1:n.cov) {mu.beta[l] <- 0; tau.beta[l] <- 0.01}
## Species slopes drawn from independent distributions. To draw from
## a common distribution (see Ovaskainen & Soininen 2011) change
## to vague priors for hyperparameters for spp slopes e.g.
## for(l in 1:n.cov) { mu.beta[l] ~ dnorm(0, 0.01); tau.beta[l] ~ dunif(0,50) }
}

',
  , file="LVM-simulate.txt")
```

We now run an individual LVM for each predictor (pH and soil moisture) by calling JAGS. Note this may take a fair bit of time (but hopefully under an hour!).

```
fit.LVM.1 <- jags(data=sim.data1,
                     inits = NULL,
                     parameters.to.save=c('spp.int', 'spp.coef', 'spp.phi', 'LV',
                                          'lv.coef', 'mu.beta', 'tau.beta'),
                     model.file="LVM-simulate.txt",
                     DIC=TRUE,
                     n.iter=10000,
                     n.burnin=1000,
                     n.chains=3,
                     n.thin=20)

fit.LVM.2 <- jags(data=sim.data2,
                     inits = NULL,
                     parameters.to.save=c('spp.int', 'spp.coef', 'spp.phi', 'LV',
                                          'lv.coef', 'mu.beta', 'tau.beta'),
                     model.file="LVM-simulate.txt",
                     DIC=TRUE,
                     n.iter=10000,
                     n.burnin=1000,
                     n.chains=3,
                     n.thin=20)

# Save data to model fits
fit.LVM.1$data <- sim.data1
fit.LVM.2$data <- sim.data2
```

Model inference

Extract environmental and residual correlations (both significant and non-significant mean and median values).

```
enviro.cor.1 <- extract.env.cor(fit.mod = fit.LVM.1,
                                  X = fit.LVM.1$data$X,
                                  y = fit.LVM.1$data$y)
residual.cor.1 <- extract.residual.cor(fit.mod = fit.LVM.1,
                                         X = fit.LVM.1$data$X,
                                         y = fit.LVM.1$data$y)
enviro.cor.2 <- extract.env.cor(fit.mod = fit.LVM.2,
                                  X = fit.LVM.2$data$X,
                                  y = fit.LVM.2$data$y)
residual.cor.2 <- extract.residual.cor(fit.mod = fit.LVM.2,
                                         X = fit.LVM.2$data$X,
                                         y = fit.LVM.2$data$y)
```

Compare environmental and residual correlation plots for the two gradients. The residual correlation from the LVM for pH should provide a good approximation of the environmental correlation attributable to soil moisture, and vice versa.

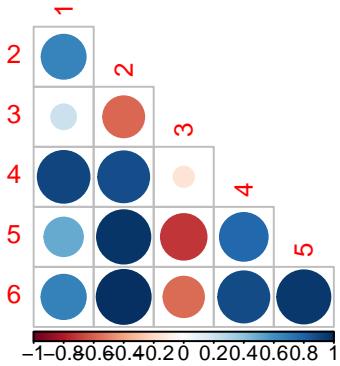
```

par(mfrow = c(2,2))
corrplot(enviro.cor.1$envcor.mean, diag = F, type = "lower",
         title = "Environmental correlation (pH)", mar=c(0,0,1,0))
corrplot(residual.cor.1$rescor.mean, diag = F, type = "lower",
         title = "Residual correlation (pH)", mar=c(0,0,1,0))

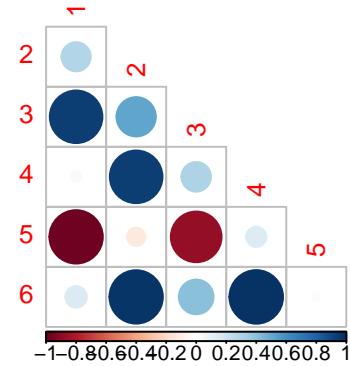
corrplot(enviro.cor.2$envcor.mean, diag = F, type = "lower",
         title = "Environmental correlation (soil moisture)", mar=c(0,0,1,0))
corrplot(residual.cor.2$rescor.mean, diag = F, type = "lower",
         title = "Residual correlation (soil moisture)", mar=c(0,0,1,0))

```

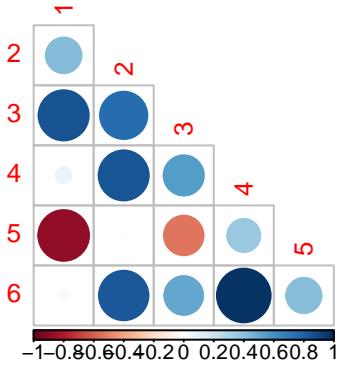
Environmental correlation (pH)



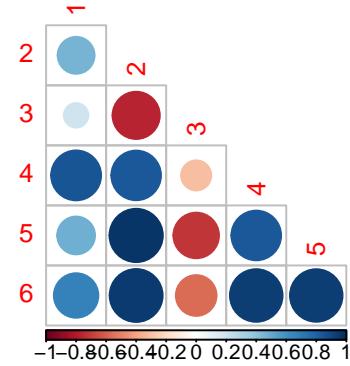
Residual correlation (pH)



Environmental correlation (soil moisture)



Residual correlation (soil moisture)



Large red circles indicate strong negative patterns of co-occurrence, while large blue circles indicate strong positive patterns of co-occurrence.

It is instructive to contrast the environmental and residual correlation plots with the species-specific responses to each gradient (note for simplicity these are fitted using the simulated data and `ggplot`'s internal model fitting functions, but could also be obtained via the predicted environmental responses derived from the LVMs).

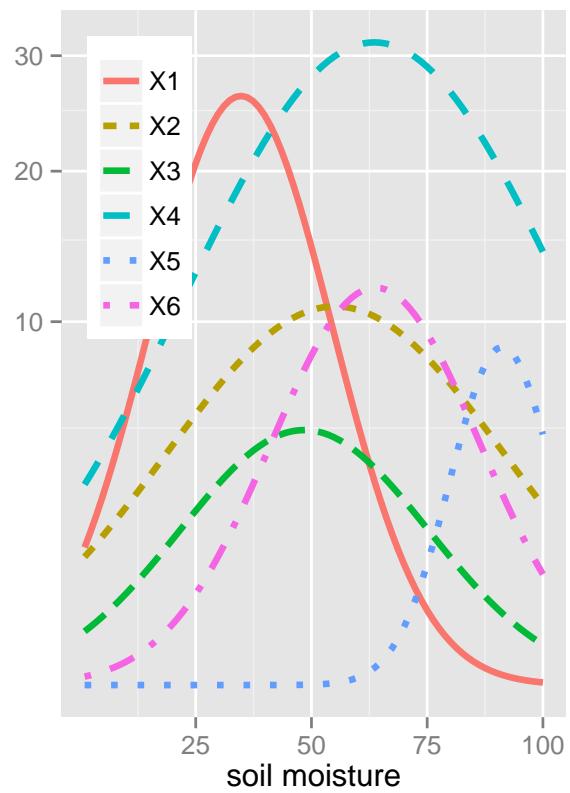
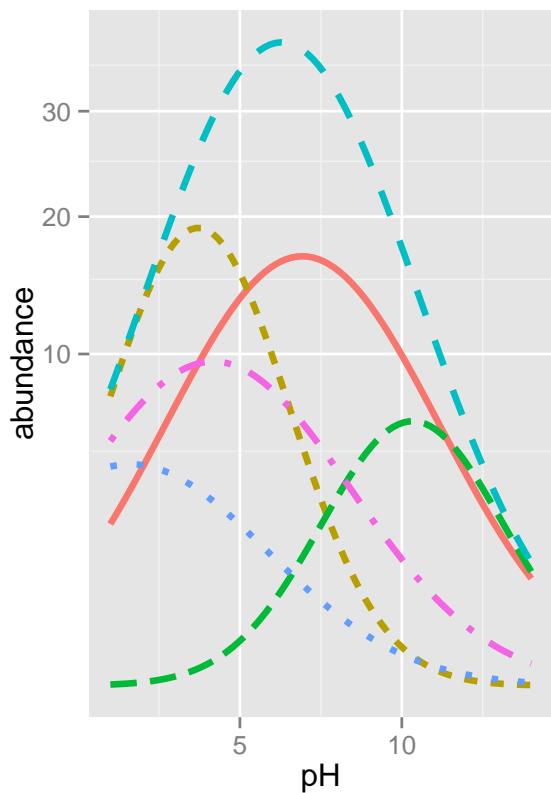
```
sim.frame1 <- data.frame(multi.sim, grad1 = locs[,1])
sim.gg1 <- gather(sim.frame1, species, value, -grad1)

p1 <- ggplot(sim.gg1, aes(y = value, x = grad1)) +
  geom_smooth(method = "glm", formula = y ~ poly(x,2),
              aes(group = species, col = species, linetype = species),
              size = 1.2, family = negative.binomial(theta = 1), se = FALSE) +
  theme(legend.position="none") +
  coord_trans(y = "sqrt") +
  xlab("pH") + ylab("abundance")

sim.frame2 <- data.frame(multi.sim, grad2 = locs[,2])
sim.gg2 <- gather(sim.frame2, species, value, -grad2)

p2 <- ggplot(sim.gg2, aes(y = value, x = grad2)) +
  geom_smooth(method = "glm", formula = y ~ poly(x,2),
              aes(group = species, col = species, linetype = species),
              size = 1.2, family = negative.binomial(theta = 1), se = FALSE) +
  theme(legend.title=element_blank(), legend.justification=c(0,1), legend.position=c(0,1)) +
  coord_trans(y = "sqrt") +
  xlab("soil moisture") + ylab(NULL)

grid.arrange(p1, p2, ncol = 2)
```



Mantel tests of the association between the environmental correlation for pH and residual correlation for soil moisture and vice versa.

```
mantel(enviro.cor.1$envcor.mean, residual.cor.2$rescor.mean)

## 'nperm' > set of all permutations; Resetting 'nperm'.

##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = enviro.cor.1$envcor.mean, ydis = residual.cor.2$rescor.mean)
##
## Mantel statistic r: 0.9931
##      Significance: 0.0041667
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.809 0.851 0.928 0.961
## Permutation: free
## Number of permutations: 720

mantel(enviro.cor.2$envcor.mean, residual.cor.1$rescor.mean)

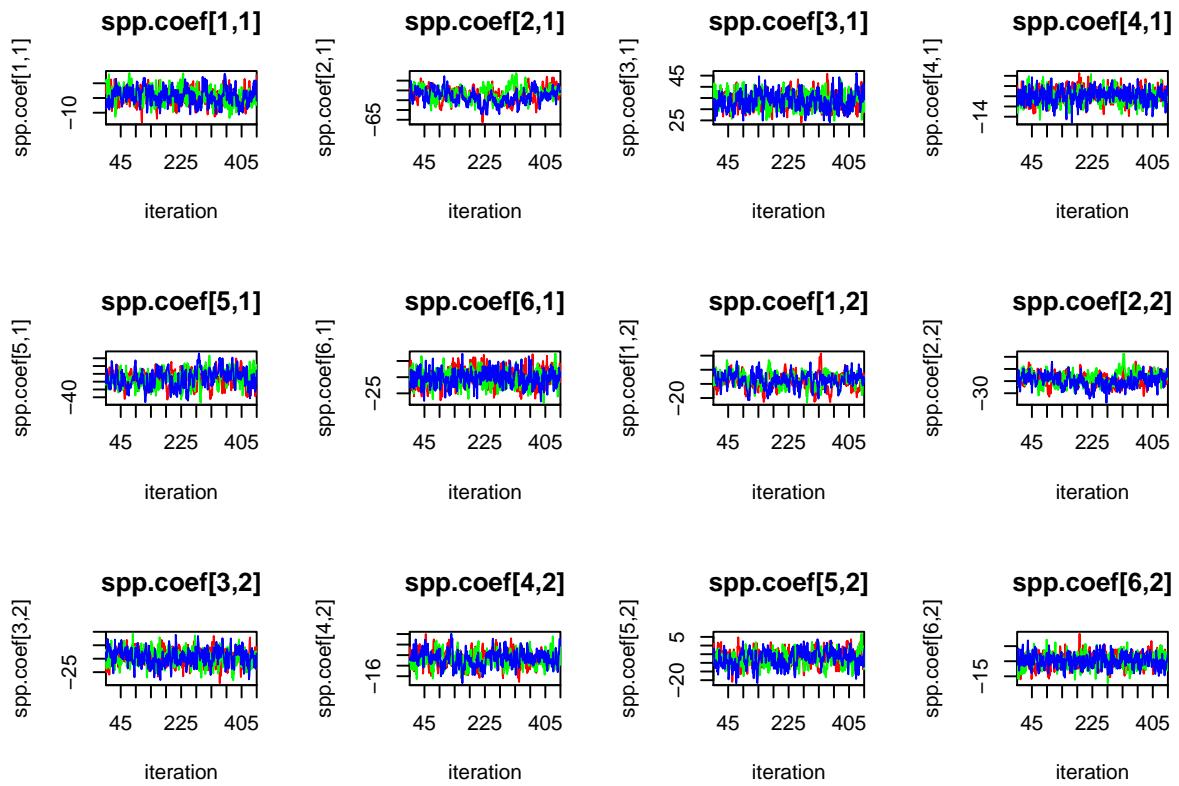
## 'nperm' > set of all permutations; Resetting 'nperm'.

##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = enviro.cor.2$envcor.mean, ydis = residual.cor.1$rescor.mean)
##
## Mantel statistic r: 0.9636
##      Significance: 0.0027778
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.453 0.567 0.676 0.859
## Permutation: free
## Number of permutations: 720
```

Model diagnostics

Remember to to assess mixing of chains, Rhats etc..

```
# e.g.  
out <- fit.LVM.1$BUGSoutput # extract BUGS output  
  
# view species coefficients, credible intervals and Rhats  
out$summary[grep("spp.coef", rownames(out$summary)),]  
  
##  
##               mean        sd      2.5%      25%      50%  
## spp.coef[1,1] -4.376308 2.658915 -9.623917 -6.211199 -4.368937  
## spp.coef[2,1] -52.244569 3.696968 -59.665952 -54.642950 -52.134381  
## spp.coef[3,1] 33.828181 3.625238 27.151757 31.306688 33.740443  
## spp.coef[4,1] -9.938732 1.412374 -12.817821 -10.836772 -9.902413  
## spp.coef[5,1] -28.113718 5.286011 -38.512741 -31.773215 -27.867749  
## spp.coef[6,1] -20.152278 2.512646 -25.208912 -21.722603 -20.155770  
## spp.coef[1,2] -13.850570 2.661883 -19.201696 -15.574995 -13.783175  
## spp.coef[2,2] -24.309061 2.676103 -29.866672 -26.093269 -24.186097  
## spp.coef[3,2] -19.474867 2.839791 -25.224109 -21.336707 -19.436900  
## spp.coef[4,2] -12.688023 1.474339 -15.530660 -13.673480 -12.719679  
## spp.coef[5,2] -7.328328 4.444105 -16.405253 -10.183379 -7.234160  
## spp.coef[6,2] -9.799089 2.251483 -14.159866 -11.275689 -9.820069  
##                 75%     97.5%    Rhat n.eff  
## spp.coef[1,1] -2.624422 1.022885 1.001172 1400  
## spp.coef[2,1] -49.668080 -45.426535 1.056914 40  
## spp.coef[3,1] 36.280006 41.139177 1.000929 1400  
## spp.coef[4,1] -9.014537 -7.126312 1.004041 480  
## spp.coef[5,1] -24.367833 -18.238094 1.005865 880  
## spp.coef[6,1] -18.569281 -15.106468 1.006487 430  
## spp.coef[1,2] -12.050343 -8.729608 1.017428 130  
## spp.coef[2,2] -22.469606 -19.265690 1.041300 53  
## spp.coef[3,2] -17.496958 -14.163351 1.002806 690  
## spp.coef[4,2] -11.716850 -9.720445 1.002209 860  
## spp.coef[5,2] -4.249257 1.338408 1.016139 140  
## spp.coef[6,2] -8.249307 -5.495947 1.000578 1400  
  
# view trace plots  
traceplot(fit.LVM.1, varname = c("spp.coef"), mfrow = c(3,4))
```



References

Ovaskainen, O. & Soininen, J. (2011). Making more out of sparse data: hierarchical modeling of species communities. *Ecology*, 92, 289-295.

D.2.1 fitLVM-auxiliaryfunctions.R

```
#####
## Auxilary functions for fitting LVMs ##
#####

## Calculate the correlation between spp due to shared/diverging env responses
extract.env.cor <- function(fit.mod, X, y) {
  n.species <- ncol(y)
  fit.mcmcBase <- fit.mod$BUGSoutput
  mcmc.runs <- mcmc(fit.mcmcBase$sims.matrix, start = 1,
                     thin = fit.mcmcBase$n.thin)
  rm(fit.mcmcBase)

  all.shared.env.mat <- matrix(0,nrow(mcmc.runs),n.species^2)

  ## For each MCMC sample, find the correlation between spp responses
  for(t in 1:nrow(mcmc.runs)) {
    cw.spp.coef <- matrix(mcmc.runs[t,grep("spp.coef\\\[",colnames(mcmc.runs))],
                           n.species,byrow=F)
    eta.mat <- X%*%t(cw.spp.coef) ## On linear predictor scale
    all.shared.env.mat[t,] <- as.vector(cor(eta.mat))
  }

  ## Average/Median over the MCMC samples
  env.mat.mean <- sig.env.mat.mean <- matrix(apply(all.shared.env.mat,2,mean),
                                                n.species,byrow=F)
  env.mat.median <- sig.env.mat.median <- matrix(apply(all.shared.env.mat,2,median),
                                                n.species,byrow=F)

  get.cor.intervals <- HPDinterval(as.mcmc(all.shared.env.mat), prob = 0.95)
  id.sign.cors <- which(get.cor.intervals[,1] > 0 | get.cor.intervals[,2] < 0)
  sig.env.mat.mean[-id.sign.cors] <- 0
  sig.env.mat.median[-id.sign.cors] <- 0

  return(list(envcor.mean = env.mat.mean,
              envcor.median = env.mat.median,
              sig.envcor.mean = sig.env.mat.mean,
              sig.envcor.median = sig.env.mat.median))
}

## Produce the residual correlation
extract.residual.cor <- function(fit.mod, X, y) {
  fit.mcmcBase <- fit.mod$BUGSoutput
  mcmc.runs <- mcmc(fit.mcmcBase$sims.matrix,
                     start = 1,
                     thin = fit.mcmcBase$n.thin)
  rm(fit.mcmcBase)

  n.species <- ncol(y)
  n.sites <- nrow(y)
  Tau.arr <- matrix(NA,nrow(mcmc.runs),n.species^2)
  for(t in 1:nrow(mcmc.runs)) {
    lvs <- matrix(mcmc.runs[t,grep("LV",colnames(mcmc.runs))],
                  n.sites,byrow=F)
    lv.coefs <- matrix(mcmc.runs[t,grep("lv.coef",colnames(mcmc.runs))],
```

```

n.species,byrow=F)
Tau.arr[t,] <- as.vector(cor(lvs%*%t(lv.coefs))) }

## Average/Median over the MCMC samples
Tau.mat.mean <- sig.Tau.mat.mean <- matrix(apply(Tau.arr,2,mean),
                                              n.species,byrow=F)
Tau.mat.median <- sig.Tau.mat.median <- matrix(apply(Tau.arr,2,median),
                                                 n.species,byrow=F)

get.cor.intervals <- HPDinterval(as.mcmc(Tau.arr), prob = 0.95)
id.sign.cors <- which(get.cor.intervals[,1] > 0 | get.cor.intervals[,2] < 0)
sig.Tau.mat.mean[-id.sign.cors] <- 0
sig.Tau.mat.median[-id.sign.cors] <- 0

return(list(rescor.mean = Tau.mat.mean,
            rescor.median = Tau.mat.median,
            sig.rescor.mean = sig.Tau.mat.mean,
            sig.rescor.median = sig.Tau.mat.median))
}

## Produce mean and median point estimates from fit
extract.params <- function(fit.mod) {
  fit.mcmcBase <- fit.mod$BUGSoutput
  mcmc.runs <- mcmc(fit.mcmcBase$sims.matrix, start = 1,
                      thin = fit.mcmcBase$n.thin)
  rm(fit.mcmcBase)
  n.species <- length(grep("spp.int", colnames(mcmc.runs)))

  all.spp.coef <- mcmc.runs[,grep("spp.coef", colnames(mcmc.runs))]
  spp.coef.mean <- matrix(apply(all.spp.coef,2,mean),n.species,byrow=F)
  spp.coef.median <- matrix(apply(all.spp.coef,2,median),n.species,byrow=F)

  all.spp.int <- mcmc.runs[,grep("spp.int", colnames(mcmc.runs))]
  spp.int.mean <- apply(all.spp.int,2,mean)
  spp.int.median <- apply(all.spp.int,2,median)

  all.spp.phi <- mcmc.runs[,grep("spp.phi", colnames(mcmc.runs))]
  spp.phi.mean <- apply(all.spp.phi,2,mean)
  spp.phi.median <- apply(all.spp.phi,2,median)

  all.mu.beta <- mcmc.runs[,grep("mu.beta", colnames(mcmc.runs))]
  mu.beta.mean <- apply(all.mu.beta,2,mean)
  mu.beta.median <- apply(all.mu.beta,2,median)

  all.tau.beta <- mcmc.runs[,grep("tau.beta", colnames(mcmc.runs))]
  tau.beta.mean <- apply(all.tau.beta,2,mean)
  tau.beta.median <- apply(all.tau.beta,2,median)

  return(list(spp.coef.mean = spp.coef.mean, spp.coef.median = spp.coef.median,
              spp.int.mean = spp.int.mean, spp.int.median = spp.int.median,
              spp.phi.mean = spp.phi.mean, spp.phi.median = spp.phi.median,
              mu.beta.mean = mu.beta.mean, mu.beta.median = mu.beta.median,
              tau.beta.mean = tau.beta.mean, tau.beta.median = tau.beta.median))
}

```