

SPATIAL AND TEMPORAL SCALING OF SPECIES
COMPOSITION AT THE TALLGRASS PRAIRIE
PRESERVE, OKLAHOMA: IMPLICATIONS
FOR THEORY AND CONSERVATION

By

DANIEL J MCGLINN

Bachelor of Science in Biology
University of North Carolina
Chapel Hill, NC
2004

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2009

SPATIAL AND TEMPORAL SCALING OF SPECIES
COMPOSITION AT THE TALLGRASS PRAIRIE
PRESERVE, OKLAHOMA: IMPLICATIONS
FOR THEORY AND CONSERVATION

Dissertation Approved:

Dr. Michael W. Palmer

Dissertation Adviser

Dr. Samuel D. Fuhlendorf

Dr. Janette A. Steets

Dr. Ronald J. Tyrl

Dr. A. Gordon Emslie

Dean of the Graduate College

PREFACE

My dissertation is composed of five chapters. Chapter I is a general introduction to the topic of scaling in species composition and a summary of my main conclusions. Chapter II, entitled “A long-term study on the scaling of vascular plant composition in a tallgrass prairie”, was coauthored by my advisor, M.W. Palmer, and will be submitted as a Data Paper to the *Ecological Archive*. Chapter III, entitled “Modeling the sampling effect in the species-time-area relationship”, was coauthored by M.W. Palmer and published in the journal *Ecology* in March 2009. Chapter IV, entitled “Quantifying the influence of environmental texture on the rate of species turnover – evidence from two habitats” was coauthored by M.W. Palmer. Chapter V, entitled “The influence of management relative to inherent landscape heterogeneity on the vegetation of a tallgrass prairie” was coauthored by M.W. Palmer and P.G. Earls. Although much of the fieldwork work in this dissertation was carried on by others prior to my arrival in Oklahoma, I use the first person singular in this document for stylistic reasons

ACKNOWLEDGMENTS

First, I thank my dissertation committee for their encouragement and help in developing my research ideas. I gratefully acknowledge my primary advisor, Mike Palmer, for his guidance, insight, generosity with his time, and his careful attentiveness to my writing and development as a scientist. I am also thankful for Mike's enthusiasm with ecology and science in general, his breadth of ideas, and carefree attitude towards academia. It was extremely comforting to feel that new ideas were always up for debate. Ron Tyrl has always taken a strong interest in my development and provided me with a lot of personal encouragement and guidance identifying plants and uncovering interesting aspects of natural history and ecology. Janette Steets offered me many thorough critiques of my manuscripts, was an excellent and caring person to work for during my teaching assistantship, and provided me insight in developing into a successful young ecologist. Sam Fuhlendorf helped me to better understand the applied side of ecology and guided my thinking in new and stimulating directions. Other colleagues that have provided insightful critique and assistance with my work include E.P. White, J.D. Fridley, J.C. Nekola, B. Crauder, and A. Noel.

I am very thankful to my fellow graduate students who have stimulated my ideas. In particular, Zachary Roehrs who was always up for a brainstorming session. The members of the Laboratory for Innovative Biodiversity Research and Analysis for their

encouragement of my ideas and assistance in the field. Particularly I wish to mention Pete Earls, Matt Allen, Vaskar Thapa, Fumiko Shirakura, Kelly DeRennaux, Will Lowry, Andrew Potter, and Olga Blinkova.

My family was very supportive during my five years studying in Oklahoma. My father and mother, James and Jeanne McGlenn, aided in the development of my ideas, proofed many of my manuscripts, and in general are responsible for my curiosity of the natural world. I am thankful for the support of my sisters Meghan Manfra and Alison Turner.

I especially thank my partner Maurine Gilmore without whom this dissertation would not have been possible. Maurine brought joy and creativity into my life that sustained me throughout my studies. I must also thank her for being my dedicated editor and biggest supporter.

I acknowledge the Environmental Protection Agency for funding the majority of my research through the GRO fellowship, NSF Grant Number EPS-0447262, and the OSU Botany department (particularly Ann Yankunas and Paula Shryock) for additional funding support and their assistance with the logistics of my fellowship.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Literature Cited.....	6
II. A LONG-TERM STUDY ON THE SCALING OF VASCULAR PLANT COMPOSITION IN A TALLGRASS PRAIRIE	10
Introduction.....	10
Metadata.....	12
Class I. Data Set Descriptions.....	12
Class II. Research Origin Descriptors.....	13
Class III. Data Set Status and Accessibility.....	18
Class IV. Data Structural Descriptors.....	19
Class V. Supplemental Descriptors.....	31
Acknowledgements.....	32
Literature Cited.....	33
Tables.....	37
Figures.....	39
III. MODELING THE SAMPLING EFFECT IN THE SPECIES-TIME-AREA RELATIONSHIP	41
Abstract.....	41
Introduction.....	42
Sampling Effect Model.....	44
Model assumptions.....	44
Model derivation.....	45
Model parameters.....	47
Model results.....	49
Empirical Patterns.....	51
Empirical analysis.....	51
Empirical results.....	51

Chapter	Page
III. MODELING THE SAMPLING EFFECT IN THE SPECIES-TIME-AREA RELATIONSHIP	41
Discussion	52
Spatial and temporal scale dependence	52
Time-by-area interaction	53
Importance of the RAD of the species pool	55
Relaxing the zero-sum assumption	55
Relevance and future utility of the model	56
Conclusions	58
Acknowledgements	59
Literature Cited	60
Tables	66
Figures	67
IV. QUANTIFYING THE INFLUENCE OF ENVIRONMENTAL TEXTURE ON THE RATE OF SPECIES TURNOVER – EVIDENCE FROM TWO HABITATS	71
Abstract	71
Introduction	72
Methods	74
Study site	74
Sampling methods	75
Analytical methods	76
Results	82
Environmental control of species composition	82
Geometry of the environment and species turnover	82
Discussion	83
Implications for the ETH and future tests	83
Euclidean and chi-squared variograms as metrics of species turnover	85
Habitat differences	86
Conclusions	88
Acknowledgements	89
Literature Cited	90
Tables	94
Figures	95

Chapter	Page
V. THE INFLUENCE OF MANAGEMENT RELATIVE TO INHERENT LANDSCAPE HETEROGENEITY ON THE VEGETATION OF A TALLGRASS PRAIRIE.....	100
Abstract.....	100
Introduction.....	101
Methods.....	104
Study site.....	104
Management.....	105
Data collection	106
Data analysis	108
Results.....	112
Variance in species richness	112
Variance in species composition.....	114
Discussion.....	115
Role and drivers of site-to-site variation.....	115
Role and drivers of year-to-year variation	117
Independent management effects.....	118
Management implications	120
The importance of observational studies	121
Conclusions.....	122
Acknowledgements.....	122
Literature Cited	124
Tables	133
Figures.....	136
 APPENDICES	 141
Appendix A.....	142
Appendix B.....	144
Appendix C.....	146
Literature Cited	151
Appendix D.....	152
Appendix E.....	153
Appendix F.....	154
Appendix G.....	155

LIST OF TABLES

Table	Page
2.1. The linear dimension and area of the five spatial grains/levels	37
2.2. The visual cover classes estimated for each species at the 100 m ² grain	38
3.1. The evenness of each of the relative abundance distributions	66
4.1. RDA and permutation tests for the two habitat types	94
5.1. Variation partitioning of richness into its components for site, year and management variables	133
5.2. The conditional explanatory power of specific variables on species richness	134
5.3. The results of the randomization tests on the pCCAs	135
E1. The formulation of five isotropic variogram models used in the text	153
G1. The abbreviation and functional group for each species that appears in Fig. 5.5 – the pCCA biplot of management effects	155

LIST OF FIGURES

Figure	Page
2.1. A map of the Tallgrass Prairie Preserve	39
2.2. The sampling design for the permanent plots	40
3.1. The SAR and STR calculated from the sampling model for several values of R	67
3.2. The SAR and STR calculated from the sampling model for five different RADs.....	68
3.3. The time-by-area interaction as a function of log richness for three values of R and the five RADs.....	69
3.4. The empirical SAR and STR for data from the TGPP.....	70
4.1. Diagram of a transect (A.), a subtransect (B.), and a section (C.)	95
4.2. Diagram of the random rotation and random reflection permutation scheme.....	96
4.3. The nine environmental and community Euclidean variograms in the grassland habitat.....	97
4.4. The nine environmental and community Euclidean variograms in the woodland habitat.....	98
4.5. The relationship between the rates of compositional and environmental distance decay	99
5.1. Location of permanent plots, weather tower, and management units at the study site.....	136
5.2. The average species richness of five functional groups over the course of the study period	137

Figure	Page
5.3. DCA scatterplot displaying the 20 sites from 1998 to 2008.....	138
5.4. A Venn diagram displaying the variation in species composition explained by site, year, and management factors using pCCA.....	139
5.5. A pCCA biplot displaying the influence of the management variables with the sites and years as covariables.....	140
C1. The \log_{10} probability rank diagrams for all nine of the different RADs	147
C2. The species-area relationship for five values of R and the five RADs	148
C3. The species-time relationship for five values of R and the five RADs	149
C4. The distribution of time-by-area interaction as a function of \log_{10} richness for five values of R and the five RADs.....	150
D1. The distribution of each management variable and the degree of collinearity with the other variables	152
F1. The relationship between richness and several diversity indices	154

CHAPTER I

INTRODUCTION

Species composition is structured in space and time (Preston 1960, MacArthur 1965, Rosenzweig 1995, Hubbell et al. 1999). This fundamental aspect of communities reflects the influence of various biotic and abiotic drivers including dispersal limitation (MacArthur and Wilson 1963, Hubbell et al. 1999), competitive exclusion (Tilman 1994, Pacala and Levin 1997, Wilson 2007), pathogenic effects (Janzen 1970, Connell 1971, Petermann et al. 2008), environmental gradients (Whittaker 1956, Wilson and Mohler 1983, Palmer and Dixon 1990), clonal growth (Pecháčová et al. 1999), and disturbance events (Levin and Paine 1974, White 1979, Arévalo et al. 2000). The influence of these various drivers on species turnover is generally expected to be scale and ecosystem dependent (Reed et al. 1993, Palmer 2007a).

Despite the degree of complexity underlying the spatial and temporal structure of community patterns, ecologists have noted that some quantitative community metrics, such as community similarity and species richness, change systematically as a function of spatial and temporal scale in almost all communities. Distance decay is one such scaling relationship. Distance decay reflects what is often referred to as the first law of geography: the spatial and/or temporal distance between two samples is inversely proportional to their similarity (Tobler 1970, Nekola and White 1999). This general rule

applies to patterns of species composition, and it provides a basis for understanding the rate at which new species are accumulated as a function of area or time, which is reflected in turn by two other scaling relationships: the species-area relationship and the species-time relationship, respectively. All three of these scaling relationships provide community-wide quantitative expressions of species aggregation (He et al. 1996, Plotkin et al. 2000a, He and Legendre 2002, Morlon et al. 2008) and have the potential to implicate which drivers are shaping species turnover across scales (Conner and McCoy 1979, Drakare et al. 2006). Therefore, the distance decay relationship, the species-area relationship, and the species-time relationship are useful for testing multi-scale hypotheses related to species turnover (Williams 1943, Storch et al. 2003, Adler 2004, Fattorini 2007, Palmer 2007a).

The goal of this study was to advance our understanding of the spatial and temporal patterns of change in plant species composition or species turnover and to examine hypotheses related to the driving mechanisms of these patterns. I accomplish this goal with four complementary studies in the following chapters.

In Chapter II, I provide metadata for an 11 year, multi-scale dataset that M.W. Palmer, my advisor, initiated at the Tallgrass Prairie Preserve in Osage Co., OK. The function of this contribution is to make this valuable, multipurpose dataset available to the broader community of ecologists and conservationists. I helped to manage and update this dataset over my five years of study, and these data formed the basis of my analyses in Chapters III and V of this dissertation. A detailed description of the dataset will help to supplement these later chapters, as well as to provide advanced notice that these data will be submitted into the public domain in the near future.

In Chapter III, I examine whether empirical patterns of spatial and temporal scaling of species richness reflect an ecological signature and not simply the passive sampling of species. I developed an analytical model for the species-time-area relationship (STAR) based on the sampling effect and compared it with an empirical STAR generated by the data described in Chapter II. The model generated the expected STAR under the assumption that species were sampled at random from a species pool (i.e., species are neutral). The model demonstrated that the average rate of replacement and the relative abundance distribution of the species pool can exert a strong influence on the STAR. Given the degree of variation observed in the expected STAR, I was unable to reject the sampling effect as an explanation for an empirical STAR. This chapter represents the first attempt to analytically model the STAR, the first empirical example of a complete nested STAR, and the first study to derive patterns of the time-by-area interaction which is the fundamental basis of the STAR. The material in Chapter III has been published in *Ecology* (2009).

In Chapter IV, I continue my theoretical examination of species turnover with a test of a generalized version of the Environmental Texture Hypothesis (ETH) (Palmer 2007a). Specifically I tested whether more rapid environmental distance decay was positively correlated with more rapid community distance decay. For this test, I sampled the vascular plant communities and environmental variables along grassland and woodland transects at the Tallgrass Prairie Preserve (Osage Co., OK). I found that the rate of species turnover was positively correlated with the rate of environmental distance decay, but this relationship was only significant in the grassland habitat. I expected the relationship between environmental distance decay and community distance decay would

not be as strong in the woodland because species composition was not as strongly correlated with the measured environmental variables in this habitat. The primary implication of this study is that the geometry or texture of the environment can influence the rate of species turnover if the environment appears to influence species composition. This is the first direct empirical test of the assumptions of the ETH.

In Chapter V, I focus on conservation implications of environmental heterogeneity. Specifically, I quantified the ability of temporal variation in management variables, including prescribed fire and cattle/bison grazing, to explain variation in species composition in a tallgrass prairie relative to spatial and temporal sources of heterogeneity. I found that although management variables explained significant variation in species richness and species composition, the contribution of these variables was small relative to inherent spatial heterogeneity between samples. The analyses suggested that the spatial variation between samples was primarily related to belowground differences. These findings demonstrate the importance of understanding the influence of management in a broader context and suggest the exact details of management plans may not be of critical concern when attempting to meet flexible management goals.

The common thread between these four studies is a focus on the central drivers of species turnover. Overall my findings suggest that scaling relationships will continue to provide a fruitful avenue of research into the determinants of species turnover. However, some important obstacles, such as empirically quantifying the influence of the sampling effect and estimating the relative abundance distribution of the species pool, may continue to thwart predictive models of scaling relationships. In addition, my research

demonstrates that the geometry of the environment deserves further attention as a predictor of species turnover. Lastly, the influence of management heterogeneity on species composition should be examined both through controlled experimental designs as well as with observational studies that place the management effects into a broader context.

LITERATURE CITED

- Adler, P. B. 2004. Neutral models fail to reproduce observed species-area and species-time relationships in Kansas grasslands. *Ecology* 85:1265-1272.
- Arévalo, J. R., J. K. DeCoster, S. McAlister, and M. W. Palmer. 2000. Changes in two Minnesota forests during 14 years following catastrophic windthrow. *Journal of Vegetation Science* 11:833-840.
- Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in forest trees. Pages 298-312 *in* *Dynamics of Populations*, Eds. P. J. denBoer and G. R. Gradwell. Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands.
- Conner, E. F. and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. *American Naturalist* 113:791-833.
- Drakare, S., J. J. Lennon, and H. Hillebrand. 2006. The imprint of the geographical, evolutionary and ecological context on species-area relationships. *Ecology Letters* 9:215-227.
- Fattorini, S. 2007. Non-randomness in the species-area relationship: testing the underlying mechanisms. *Oikos* 116:678-689.
- He, F. L. and P. Legendre. 2002. Species diversity patterns derived from species-area models. *Ecology* 83:1185-1198.
- He, F. L., P. Legendre, and J. V. LaFrankie. 1996. Spatial pattern of diversity in a tropical rain forest in Malaysia. *Journal of Biogeography* 23:57-74.

- Hubbell, S. P., R. B. Foster, S. T. O'Brien, K. E. Harms, R. Condit, B. Wechsler, S. J. Wright, and S. L. de Lao. 1999. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science* 283:554-557.
- Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. *American Naturalist* 105:501-528.
- Levin, S. A. and R. T. Paine. 1974. Disturbance, patch formation, and community structure. *Proceedings of the National Academy of Sciences of the United States of America* 71:2744-2747.
- MacArthur, R. H. 1965. Patterns of species diversity. *Biological Reviews* 40:510-533.
- MacArthur, R. H. and E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. *Evolution* 17:373-387.
- McGlenn, D. J. and M. W. Palmer. 2009. Modeling the sampling effect in the species-time-area relationship. *Ecology* 90:836-846.
- Morlon, H., G. Chuyong, R. Condit, S. Hubbell, D. Kenfack, D. Thomas, R. Valencia, and J. L. Green. 2008. A general framework for the distance-decay of similarity in ecological communities. *Ecology Letters* 11:904-917.
- Nekola, J. C. and P. S. White. 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* 26:867-878.
- Pacala, S. W. and S. A. Levin. 1997. Biologically generated spatial pattern and the coexistence of competing species. Pages 204-232 *in* D. Tilman and P. Kareiva, editors. *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Princeton University Press, Princeton, NJ.

- Palmer, M. W. 2007. Species-area curves and the geometry of nature. Pages 15-31 *in* Scaling Biodiversity, Eds. D. Storch, P. L. Marquet, and J. H. Brown. Cambridge University Press, Cambridge, UK.
- Palmer, M. W. and P. M. Dixon. 1990. Small-scale environmental heterogeneity and the analysis of species distributions along gradients. *Journal of Vegetation Science* 1:57-65.
- Pecháčová, S., H. J. During, V. Rydlová, and T. Herben. 1999. Species-specific spatial pattern of below-ground plant parts in a montane grassland community. *Journal of Ecology* 87:569-582.
- Petermann, J. S., A. J. F. Fergus, L. A. Turbnull, and B. Schmid. 2008. Janzen-Connell effects are widespread and strong enough to maintain diversity in grasslands *Ecology* 89:2399-2406.
- Plotkin, J. B., M. D. Potts, N. Leslie, N. Manokaran, J. LaFrankie, and P. S. Ashton. 2000. Species-area curves, spatial aggregation, and habitat specialization in tropical forests. *Journal of Theoretical Biology* 207:81-99.
- Preston, F. W. 1960. Time and space and the variation of species. *Ecology* 41:611-627.
- Reed, R. A., R. K. Peet, M. W. Palmer, and P. S. White. 1993. Scale dependence of vegetation-environment correlations - a case-study of a North Carolina piedmont woodland. *Journal of Vegetation Science* 4:329-340.
- Rosenzweig, M. L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, Cambridge, UK.

- Storch, D., A. L. Šizling, and K. J. Gaston. 2003. Geometry of the species-area relationship in central European birds: testing the mechanism. *Journal of Animal Ecology* 72:509-519.
- Tilman, D. 1994. Competition and biodiversity in spatially structured habitats. *Ecology* 75:2-16.
- Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46:234-240.
- White, P. S. 1979. Pattern, process, and natural disturbance in vegetation. *Botanical Review* 45:229-299.
- Whittaker, R. H. 1956. Vegetation of the Great Smoky Mountains. *Ecological Monographs* 26:1-80.
- Williams, C. B. 1943. Area and number of species. *Nature* 152:264-267.
- Wilson, M. V. and C. L. Mohler. 1983. Measuring compositional change along gradients. *Vegetatio* 54:129-141.
- Wilson, S. D. 2007. Competition, resources, and vegetation during 10 years in native grassland. *Ecology* 88:2951-2958.

CHAPTER II

A LONG-TERM STUDY ON THE SCALING OF VASCULAR PLANT COMPOSITION IN A TALLGRASS PRAIRIE

INTRODUCTION

Identifying how species richness changes as a function of scale has important theoretical and applied implications. For example, the well known species-area relationship is an expression of how species richness changes as a function of spatial grain. This relationship has revealed how a diverse array of factors influence the scale dependence of species richness, including the rarefaction effect (McGlenn and Palmer 2009), environmental heterogeneity (Palmer 2007a), dispersal limitation (Rosenzweig 1995), and evolutionary isolation (Drakare et al. 2006). Indeed, it is now relatively common place for ecologists to consider patterns of richness at multiple spatial scales. However, time, another important facet of scale, has received considerably less attention in studies of biodiversity (White et al. 2006a, White 2007). This omission occurred despite early recognition that the temporal scale of a sample is an important determinate of richness (Fisher et al. 1943, Preston 1960). The importance of considering the temporal scaling of diversity is compounded by the growing body of evidence that demonstrates that the scaling of richness in space depends upon the temporal scale over

which it is examined (Adler et al. 2005, Fridley et al. 2005, McGlinn and Palmer 2009). The interdependence of the spatial and temporal scaling of diversity (i.e., the species-time-area relationship, STAR) has the potential to provide new theoretical insights by requiring that models simultaneously account for changes in diversity in space and time (e.g., Adler 2004).

Applied ecology also may benefit through the development of novel methods for carrying out space-for-time substitutions (Adler and Lauenroth 2003, Adler et al. 2005). One potential application of space-for-time substitutions is to predict future temporal patterns of diversity in light of climate change with the aid of current spatial patterns of diversity (Adler and Levine 2007).

Given the importance of temporal patterns of diversity, it appears that the current paucity of studies considering simply time or the joint influence of space and time is due in large part to a lack of suitable datasets in the public domain. Here, I describe a multi-scale dataset in space and time on vascular plants that I hope, if accepted, will become part of the publically accessible *Ecological Archives*. Portions of this dataset have already addressed a range of applied and theoretical questions. Palmer et al. (2002) used part of this dataset to compare strategies for efficiently conducting a thorough taxonomic inventory. Palmer et al. (2003) examined the relevance of the species pool hypothesis to explain the relationship between species richness and soil reaction. Brokaw (2004) compared the ability of modern measures of the soil environment (e.g., total C, residual P) with traditional measures of soil properties (e.g., soil cations) to explain plant composition using only samples in this dataset collected in 2002. Palmer et al. (2008a) examined how the relationships between native and exotic richness as well as the species

to genus ratio changed as a function of scale. McGlinn and Palmer (2009) constructed an empirical example of a STAR with the data. M.W. Palmer has also used the data to provide The Nature Conservancy progress reports related to changes in the vegetation of the Tallgrass Prairie Preserve (TGPP).

In addition to stimulating additional studies into the relationship of biodiversity and scale, this dataset will be valuable to practitioners interested in the functioning and conservation of the tallgrass prairie ecosystem.

METADATA

Class I. Data Set Descriptions

A. Data set identity:

Title: Multi-scale vascular plant composition from long-term monitoring at the Tallgrass Prairie Preserve, Oklahoma

B. Data set identification code:

Suggested Data Set Identity Code: TGPP_plants

C. Data set description

Principal Investigator: Michael W. Palmer, Department of Botany, Oklahoma State

University Oklahoma USA

Daniel J. McGlinn, Department of Botany, Oklahoma State University, Oklahoma USA

Abstract:

I describe a dataset that was collected as part of a monitoring project on vascular plant composition at the Tallgrass Prairie Preserve (TGPP) in Osage County, Oklahoma. The purpose of this description is to precede the submission of this information into the

public domain as a Data Paper in *Ecological Archives*. The dataset is valuable for both theoretical and applied questions given the relevance of scaling patterns to theoretical models of biodiversity which guide our conservation of diversity. Furthermore, these data will provide a reference point for tallgrass prairie restoration projects in the Flint Hills. Over the course of the 11 year period, I sampled 20 permanent plots annually. The permanent plots were selected semi-randomly from a UTM grid using the criteria that they contain less than 20 % of woody cover, standing water, or exposed rock. Plant species presence was recorded at five spatial scales: 0.01, 0.1, 1.0, 10, and 100 m² in each of the four corners of a 100 m² square quadrat. Plant species were assigned to a percent cover class at the 100 m² grain. In addition to information on plant composition, I provide data on topography, soil variables, monthly total rainfall, monthly average temperature, and management records related to fire and grazing history. I hope this dataset will stimulate further research into the scaling of biodiversity and insight into the functioning and conservation of tallgrass prairie plant communities.

D. Key words:

tallgrass prairie, restoration, species-time-area relationship, bison, vegetation monitoring, spatial scale, vascular plants, Flint Hills

Class II. Research Origin Descriptors

A. Overall project description

Identity: Multiscale vascular plant composition from long-term monitoring at the Tallgrass Prairie Preserve, Oklahoma

Originators: M.W. Palmer

Period of Study: Multiscale vascular plant data and environmental site data from the month of June, 1998-2008. Climate data from January 1993 to December 2008.

Objectives:

Abstract: same as above

Source(s) of funding: The Oklahoma State University College of Arts and Science, The Oklahoma Nature Conservancy, The Spatial and Environmental Information Clearinghouse, The Philecology Trust, The Swiss Federal Institute for Forest, Snow and Landscape Research, and the Oklahoma Water Resources Research Institute provided financial assistance at various stages of research at the Tallgrass Prairie Preserve.

B. Specific subproject description

Study Site:

The Tallgrass Prairie Preserve (TGPP, ca. 15,700 ha in size) is located between 36.73° and 36.90° N latitude, and 96.32° and 96.49° W longitude in Osage County, Oklahoma. The elevation on the preserve varies from 253 to 366 m, and over the course of the study period (1998 to 2008) the total annual rainfall averaged 942 mm and ranged from 593 to 1217 mm. The preserve is located in the southern terminus of the Flint Hills (see Hamilton 2007 Fig. 2.1) which is an ecoregion characterized by shallow soils derived from Permian sediment (Oviatt 1998). Due to long-term erosion, the surface layers of soil are thin and young; limestone and sandstone are frequently exposed at the surface. Because of these shallow rocky soils, the Flint Hills, including the TGPP, has remained unplowed and is utilized primary for cattle grazing (Kindscher and Scott 1997).

The TGPP is owned and operated by The Nature Conservancy (TNC) who purchased the bulk of the preserve (the 11,800 ha Barnard Ranch) in 1989. Since that time the TNC has made additional land acquisitions that increased the preserve's area to its current size. Prior to the acquisition of the preserve by TNC in 1989, the majority of the site was managed for cow-calf and yearling cattle production with a 4- to 5-year rotation of prescribed burning and aerial application of broadleaf herbicides (1950-1989) (Hamilton 2007). In 1993, 300 bison (*Bos bison*) were introduced onto a 1,960 ha portion of the preserve. Over time the bison unit has ground grown to a herd size of ca. 2,600 and occupies an area of ca. 8,500 ha (shaded region, Fig. 2.1). Approximately 1/3 of the burn units (watersheds) within the bison unit are randomly selected for prescribed burning annually. Some areas experience periods as long as 10 years without fire due to the random nature of burn unit selection. The remainder of the preserve is managed for seasonal cow-calf production with a more frequent application of fire. Lastly, bison are allowed to graze year round, but cattle grazing is only during the spring and summer months. Hamilton (2007) provides additional details on the management of the TGPP.

Approximately 90 % of the TGPP consists of grasslands. The majority of the grasslands are composed of tallgrass prairie habitats dominated by *Andropogon gerardii*, *Sorghastrum nutans*, *Sporobolus compositus*, *Panicum virgatum*, and *Schizachyrium scoparium*. Shortgrass prairie habitat occurs to a lesser extent on more xeric sites and is dominated by *Bouteloua* spp. Other notable vegetation types on the preserve are oak woodlands of the cross timbers which are composed primarily of *Quercus stellata* and *Quercus marilandica*, gallery forests along the main tributaries, and ephemeral wetland communities on shallow slopes and plateaus. Despite the application of herbicide earlier

in the 20th century, the flora of the TGPP appears relatively intact with a total of 763 species of vascular plants present (to date) of which 12.1% are exotic (Palmer 2007b). The referenced voucher specimens for my study are deposited in two locations: 1) in the Oklahoma State University Herbarium (herbarium code: OKLA), Stillwater, OK 74074 and 2) in the TGPP Herbarium, Pawhuska, OK (located at study site).

Research methods:

A suitable sampling design for understanding the scaling of diversity within and amongst samples requires objectively placed permanent plots. This is necessary to ensure that the results are not biased by the investigators subjective impression of homogeneity or representativeness of the site (Palmer 1993). Other important aspects of suitable long-term data include accurately relocating the samples and maintaining the consistency and accuracy of taxa identification (Milberg et al. 2008). Therefore I selected twenty permanent 100 m² plots randomly from a UTM NAD27 1 km grid of 151 plots. The only criteria I imposed on plot selection were that plots not contain artificial structures or more than 20 % of woody cover, standing water, or exposed rock.

I sampled the plots every June (when I could readily identify both early and last season plants) from 1998 to 2008. Depending on weather, sampling typically required 10 days in the field to complete. Each plot was 10×10m with iron reinforcement bars at the corners sunk to ground level and topped by Surv-Kap® aluminum caps stamped with the plot ID number. The plots were relocated with a GPS and a magnetic locator. Each corner has a series of square nested subplots with areas of 0.01, 0.1, 1, and 10 m² (Fig. 2.2). I recorded all vascular plant species rooted within each subplot, as well as the entire plot. Species not rooted in the quadrat but leaning into the quadrat were also recorded

but with a special code (see Table 2.1). I recorded a cover class for each species at the 100 m² grain (Table 2.2). M.W. Palmer estimated visual cover and made the final identification on all recorded taxa to maintain consistency and accuracy throughout the study. Additionally, I recorded height of the tallest grass, forb, and woody plant; estimated cover of woody plants, rock, bare soil, and water; and recorded slope and aspect. I took two 15 cm deep soil cores 50 cm outside each quadrat corner, for a total of 8 cores per quadrat; I varied the direction annually to minimize disturbance. Soils were analyzed by Brookside Labs (New Knoxville, OH) for total exchange capacity, pH, percent organic matter, bulk density, and, using a Mehlich 3 extractant, available sulfur, phosphorus, calcium, magnesium, potassium, sodium, boron, iron, manganese, copper, zinc, and aluminum (Mehlich 1984).

Climate data was downloaded from the Foraker Mesonet tower (36.841° N, - 96.428° W; elevation: 330 m) that is located on the preserve (Fig. 2.1). The Mesonet tower is 10 m tall and collects data every 5 minutes on a wide range of meteorologically relevant information (<http://www.mesonet.org/mcdguide.pdf>; McPherson et al. 2007). However, for the purposes of this dataset I only accessed monthly data on total precipitation and average temperature. Precipitation was measured with a Met One Tipping-Bucket Rain Gauge located just off the ground. Temperature was recorded with a Thermometrics Fast Air Temperature sensor 1.5 m above the ground. Although I provide the monthly precipitation and temperature data here, all other measured variables are freely accessible via the Mesonet webpage on a daily interval (<http://www.mesonet.org/>).

Data on fire events that occurred at the twenty quadrats were extracted from topographic burn maps created by R. G. Hamilton. The burn boundaries were visually digitized in ArcView v3.3 with the aid of a digital 3 m² resolution aerial photograph of the preserve and scanned USGS topo quads. The burn boundaries typically followed the edge of an unpaved road or tributary and are therefore accurate within a reasonable margin of error. Grazing history was reported by R. G. Hamilton and for this dataset consists simply of years of bison grazing. All other sites were within cattle units.

Nomenclature follows the PLANTS database (USDA NRCS 2008).

Project personnel:

M.W. Palmer was responsible for establishing the plots, gathering of all species data, data input and error checking. D.J. McGlenn assisted in vegetation sampling, data management, digitizing of burn layers, data input and error checking, and maintenance of the species and environmental components of the dataset. P.G. Earls developed the GIS database of management information, assisted in sampling, data input and checking. Many others assisted in the process of sampling the vegetation and soils (see *Acknowledgements*).

Class III. Data Set Status and Accessibility

A. Status

Latest Update: May 2008 for the final format of all files.

Latest Archive date: June 2009

Metadata status: Metadata are complete for this period and are stored with the data (see B. below).

Data verification: M.W. Palmer verified all species data. The soil data was checked for consistent values between years by D.J. McGlenn. The management data was extracted by P.G. Earls and D.J. McGlenn.

B. Accessibility

Storage location and medium: All digital data exist on M.W. Palmer's personal computer in ASCII format.

Contact person: Michael W. Palmer, Department of Botany, Oklahoma State University, Stillwater Oklahoma 74078 USA; tel 405-744-7717; fax 405-744-7074; mike.palmer@okstate.edu.

Copyright restrictions: None

Proprietary restrictions: None

Costs: None

Class IV. Data Structural Descriptors

A. Data Set Files

Identity:

TGPP_spe.csv for the species composition from 1998 to 2008

TGPP_specodes.csv for the species names

TGPP_env.csv contains all environmental variables including management and climate information for the study period.

TGPP_rich.csv contains species richness for each corner and each level (spatial scale) of each sample

TGPP_clim.csv contains monthly total rainfall and average temperature for all years of available Mesonet data, including years prior to the origination of sampling (1994-1997).

Size:

TGPP_spe.csv -- 16819 lines, not including header row.

TGPP_specodes.csv -- 321 lines, not including header row.

TGPP_env.csv -- 220 lines, not including header row.

TGPP_clim.csv -- 180 lines, not including header row.

TGPP_rich.csv -- 220 lines, not including header row.

Comments:

TGPP_spe.csv

Species occurrence is recorded at the highest level (smallest grain, Table 2.1) it was observed for each corner. Because the subplots are nested within one another (Fig. 2.2), species that occur at a given level are assumed to occur in all levels below. For example if a species is recorded at level 3 then it is also considered present at levels 2 and 1. If a species was only observed at level 1 (i.e. it was not present in a subplot but was rooted within the plot) then a 1 was recorded only in the column corresponding to presences in corner 4.

Evidence of spot applications of herbicide to the invasive species *Lespedeza cuneata* (sericea lespedeza) was observed occasionally in my plots. Therefore trends in this species should be interpreted with respect to this fact.

TGPP_env.csv and TGPP_clim.csv

The total monthly precipitation and monthly average temperature information in TGPP_env.csv reflects the monthly conditions from June of the previous calendar year to May of the current sampling year. Additionally I provide the datafile TGPP_clim.csv, which contains monthly precipitation and temperature records beginning 4 years prior to the initiation of my study (1994 to 2008). Therefore the data in the two files are redundant in part. I included this redundant information primarily because I felt that others would find it convenient that the climate variables were already included with the other site variables and because I wanted to provide others the option to calculate climatic lag effects for years prior to the beginning of sampling.

Lastly, in the month of February 1998 the rain gauge at the Mesonet tower did not record any data and therefore I provide no estimate of total rainfall for this month.

Format and Storage mode:

ASCII text, comma delimited. No compression schemes used.

B. Variable definitions

TGPP_spe.csv

Variable name	Variable definition	Units	Storage type	Range numeric values	Missing value codes
plot	plot number	number	integer	1–20	N/a
year	calendar year	number	integer	1998–2008	N/a
corner1	the finest grain (i.e., highest level) of species occurrence in corner 1 (see Table 2.1 and Fig.2)	number	integer	0; 2–5	No missing data, but a blank indicates a true absence of the species
corner2	the finest grain of species occurrence in corner 2	number	integer	0; 2–5	As above
corner3	the finest grain of species occurrence in corner 3	number	integer	0; 2–5	As above
corner4	the finest grain of species occurrence in corner 4 (levels 2–5) or if the species has not occurred in any of the subplots but is in the plot (level 1); level 0 applies to both corner 4 and the remainder of the plot not sampled in the subplots	number	integer	0–5	As above
cover	cover class (see Table 2.2)	number	integer	1–9	N/a
idnum	numeric ID for each species; as in TGPP_specodes.csv	number	integer	1–321	N/a
code	eight-letter code uniquely identifying species; typically first four letters of genus and species; as in TGP_specodes.csv	text	string	N/a	N/a
species	species name according to PLANTS	text	string	N/a	N/a

TGPP_specodes.csv

Variable name	Variable definition	Units	Storage type	Range numeric values	Missing value codes
idnum	numeric ID for each species; as in TGPP_specodes.csv	number	integer	1–321	N/a
code	eight-letter code uniquely identifying species; typically first four letters of genus and species; as in TGP_specodes.csv	text	string	0–4	N/a
species	species name according to PLANTS (2008)	text	string	N/a	N/a

TGPP_env.csv

Variable name	Variable definition	Units	Storage type	Range numeric values	Missing value codes
plot	plot number; as in TGPP_spe.csv	number	integer	1–20	N/a
yr	calendar year	number	integer	1998–2008	N/a
plot_yr	plot number and calendar year concatenated; as in TGPP_spe.csv	numeric code	string	n/a	N/a
date_samp	calendar date of sampling MM/DD/YY	numeric code	string	06/04/98– 06/21/08	N/a
jul_samp	Julian day of sample relative to Jan. 1 of the calendar year of sampling	number	integer	150–181	N/a
easting	UTM coordinate; NAD27 Conus zone 14	m	integer	727000–738000	N/a
northing	UTM coordinate; NAD27 Conus zone 14	m	integer	4069000– 4086000	N/a
grass ht	distance from the ground to the highest blade of grass in the plot	m	floating point	0.3–1.8	N/a
forb ht	distance from the ground to the highest forb leaf in the plot	m	floating point	0.3–1.6	N/a
woody ht	distance from the ground to the highest shrub leaf in the plot	m	floating point	0–0.9	N/a
woody %	percent cover of woody plants in the plot	%	floating point	0–15	N/a
water %	percent cover of water in the plot	%	floating point	0–15	N/a
rock %	percent cover of rock in the plot	%	floating point	0–40	N/a
bare %	percent cover of bare soil in the plot	%	floating point	0–55	N/a

slope	slope	%	integer	1–8	N/a
aspect	aspect	°	integer	28–310	N/a
TEC	Total Exchange Capacity	MEQ/100g	floating point	4.86–32.67	N/a
PH	pH	pH units	floating point	5.5–7.6	N/a
ORG	Organic Matter (humus)	%	floating point	1.91–8.65	N/a
S	Soluble Sulfur	ppm	floating point	8–87	N/a
P	Easily extractable Phosphorus	ppm	floating point	3–23	N/a
CA	Calcium	ppm	floating point	769–5001	N/a
MG	Magnesium	%	floating point	78–673	N/a
K	Potassium	%	floating point	61–658	N/a
NA	Sodium	%	floating point	14–322	N/a
BCA	Saturation of Calcium	%	floating point	44.31–86.61	N/a
BMG	Saturation of Magnesium	%	floating point	9.15–24.68	N/a
BK	Saturation of Potassium	%	floating point	1.28–6.11	N/a
BNA	Saturation of Sodium	%	floating point	0.24–5.11	N/a
BH	Saturation of Hydrogen	%	floating point	0–39.73	N/a

B	Boron	ppm	floating point	0.23–1.87	N/a
FE	Iron	ppm	floating point	68–330	N/a
MN	Manganese	ppm	floating point	8–99	N/a
CU	Copper	ppm	floating point	0.67–4.92	N/a
ZN	Zinc	ppm	floating point	1.48–8.03	N/a
AL	Aluminum	ppm	floating point	344–919	N/a
rain6	total monthly rainfall in June of the previous calendar year	mm	integer	24–269	N/a
rain7	total monthly rainfall in July of the previous calendar year	mm	integer	14–176	N/a
rain8	total monthly rainfall in Aug. of the previous calendar year	mm	integer	0–240	N/a
rain9	total monthly rainfall in Sept. of the previous calendar year	mm	integer	13–152	N/a
rain10	total monthly rainfall in Oct. of the previous calendar year	mm	integer	25–210	N/a
rain11	total monthly rainfall in Nov. of the previous calendar year	mm	integer	1–116	N/a
rain12	total monthly rainfall in Dec. of the previous calendar year	mm	integer	6–138	N/a
rain1	total monthly rainfall in Jan. of the calendar year of sampling	mm	integer	1–94	N/a
rain2	total monthly rainfall in Feb. of the calendar year of sampling	mm	integer	0–101	NA

rain3	total monthly rainfall in Mar. of the calendar year of sampling	mm	integer	18–217	N/a
rain4	total monthly rainfall in Apr. of the calendar year of sampling	mm	integer	32–159	N/a
rain5	total monthly rainfall in May of the calendar year of sampling	mm	integer	32–170	N/a
temp6	average monthly temperature in June of the previous calendar year	°C	floating point	21.9–24.8	N/a
temp7	average monthly temperature in July of the previous calendar year	°C	floating point	24.6–29	N/a
temp8	average monthly temperature in Aug. of the previous calendar year	°C	floating point	24–28.6	N/a
temp9	average monthly temperature in Sept. of the previous calendar year	°C	floating point	18.9–25.1	N/a
temp10	average monthly temperature in Oct. of the previous calendar year	°C	floating point	11.9–16.9	N/a
temp11	average monthly temperature in Nov. of the previous calendar year	°C	floating point	4.7–12.4	N/a
temp12	average monthly temperature in Dec. of the previous calendar year	°C	floating point	-3.9–4.7	N/a
temp1	average monthly temperature in Jan. of the calendar year of sampling	°C	floating point	0.2–6.8	N/a
temp2	average monthly temperature in Feb. of the calendar year of sampling	°C	floating point	1.5–8.2	N/a
temp3	average monthly temperature in Mar. of the calendar year of sampling	°C	floating point	6–13.6	N/a
temp4	average monthly temperature in Apr. of the calendar year of sampling	°C	floating point	12.8–17.9	N/a
temp5	average monthly temperature in May of the calendar year of sampling	°C	floating point	17.8–21.8	N/a

Obison	binary variable indicating plots that were grazed by bison for at least half of a year prior to June sampling in 1998 (=1) or were grazed by cattle (=0)	numeric code	integer	0–1	N/a
bison	binary variable indicating plots that were grazed by bison for at least half of a year prior to the date of sampling (=1) or were grazed by cattle (=0)	number	integer	0–1	N/a
YrsOB	the years at plot was considered in the bison unit relative to sampling date	number	floating point	0–14.66	N/a
BP5Yrs	the number of burns in the past five years relative to the date of sampling	number	integer	0–5	N/a
YrsSLB	the years since the last burn relative to the date of sampling	number	floating point	0.15–10.25	N/a
burn	a binary variable indicating a plot was reported as burned less than one year prior to sampling (=1) or was not burned (=0)	number	integer	0–1	N/a
date_burn	calendar date of burn for burns that less than one year prior to sampling date	MM/DD/Y Y	character string	01/28/98– 04/07/08	N/a
jul_burn	Julian day of burns that occurred less than one year prior to sampling, calculated relative to January 1 of the calendar year of the burn	number	integer	28–349	N/a

TGPP_clim.csv

Variable name	Variable definition	Units	Storage type	Range numeric values	Missing value code
year	calendar year	number	integer	1994–2008	N/a
mo	calendar month	number	integer	1–12	N/a
rain	total monthly rainfall	mm	integer	0–269	N/a
temp	average monthly temperature	°C	floating point	-3.9–29	N/a

TGPP_rich.csv

Variable name	Variable definition	Units	Storage type	Range numeric values	Missing value codes
plot	plot number; as in TGPP_spe.csv	number	integer	1–20	N/a
year	calendar year	number	integer	1998–2008	N/a
L1	number of species at level 1 (see Table 2.1)	number	integer	48–104	N/a
L2C1	number of species at level 2 in corner 1 (see Table 2.1 and Fig. 2.2)	number	integer	24–68	N/a
L2C2	number of species at level 2 in corner 2	number	integer	25–70	N/a
L2C3	number of species at level 2 in corner 3	number	integer	21–67	N/a
L2C4	number of species at level 2 in corner 4	number	integer	24–74	N/a
L3C1	number of species at level 3 in corner 1	number	integer	10–40	N/a
L3C2	number of species at level 3 in corner 2	number	integer	10–44	N/a
L3C3	number of species at level 3 in corner 3	number	integer	3–40	N/a
L3C4	number of species at level 3 in corner 4	number	integer	7–39	N/a
L4C1	number of species at level 4 in corner 1	number	integer	2–21	N/a
L4C2	number of species at level 4 in corner 2	number	integer	4–23	N/a
L4C3	number of species at level 4 in corner 3	number	integer	0–24	N/a
L4C4	number of species at level 4 in corner 4	number	integer	1–25	N/a
L5C1	number of species at level 5 in corner 1	number	integer	1–10	N/a
L5C2	number of species at level 5 in corner 2	number	integer	1–11	N/a
L5C3	number of species at level 5 in corner 3	number	integer	0–9	N/a
L5C4	number of species at level 5 in corner 4	number	integer	0–9	N/a

Class V. Supplemental Descriptors

A. Data acquisition

Data forms: data forms

Location of completed data forms: The completed species data forms are stored at Oklahoma State University Department of Botany (M.W. Palmer's Office).

B. Quality assurance/quality control procedures Field sheets were proofed for concerns after every day in the field as well as during digitization.

C. Related material: n/a

D. Computer programs and data processing algorithms: n/a

E. Archiving: n/a

F. Publications and results:

These data have been used in the following publications:

Brokaw, J. M. 2004. Comparing explanatory variables in the analysis of species composition of a tallgrass prairie. *Proceedings of the Oklahoma Academy of Science* 84:33-40.

McGlenn, D. J. and M. W. Palmer. 2009. Modeling the sampling effect in the species-time-area relationship. *Ecology* 90:836-846.

Palmer, M. W., J. R. Arévalo, M. C. Cobo, and P. G. Earls. 2003. Species richness and soil reaction in a northeastern Oklahoma landscape. *Folia Geobotanica* 38:381-389.

Palmer, M. W., P. G. Earls, B. W. Hoagland, P. S. White, and T. Wohlgemuth. 2002. Quantitative tools for perfecting species lists. *Environmetrics* 13:121-137.

Palmer, M. W., D. J. McGlinn, and J. F. Fridley. 2008. Artifacts and artificions in biodiversity research. *Folia Geobotanica* 43:245-257.

G. History of data set usage: see F. above for references that use the data

H. Data set update history: All of the data were last updated in June 2008.

Review history: n/a

Questions and comments from secondary users: n/a

ACKNOWLEDGMENTS

DJM received funding from the U.S. Environmental Protection Agency (EPA) under the Greater Research Opportunities (GRO) Graduate Program. The U.S. EPA has not officially endorsed this publication, and the views expressed herein may not reflect the views of the Agency. Additionally I thank Bob Hamilton, members of the Laboratory for Innovative Biodiversity Research and Analysis, and the members of the Osage Nation for assistance with field work.

LITERATURE CITED

- Adler, P. B. 2004. Neutral models fail to reproduce observed species-area and species-time relationships in Kansas grasslands. *Ecology* 85:1265-1272.
- Adler, P. B. and W. K. Lauenroth. 2003. The power of time: spatiotemporal scaling of species diversity. *Ecology Letters* 6:749-756.
- Adler, P. B. and J. M. Levine. 2007. Contrasting relationships between precipitation and species richness in space and time. *Oikos* 116:221-232.
- Adler, P. B., E. P. White, W. K. Lauenroth, D. M. Kaufman, A. Rassweiler, and J. A. Rusak. 2005. Evidence for a general species-time-area relationship. *Ecology* 86:2032-2039.
- Brokaw, J. M. 2004. Comparing explanatory variables in the analysis of species composition of a tallgrass prairie. *Proceedings of the Oklahoma Academy of Science* 84:33-40.
- Drakare, S., J. J. Lennon, and H. Hillebrand. 2006. The imprint of the geographical, evolutionary and ecological context on species-area relationships. *Ecology Letters* 9:215-227.
- Fisher, R. A., A. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12:42-58.
- Fridley, J. D., R. K. Peet, T. R. Wentworth, and P. S. White. 2005. Connecting fine- and broad-scale species-area relationships of Southeastern US Flora. *Ecology* 86:1172-1177.

- Hamilton, R. G. 2007. Restoring heterogeneity on the Tallgrass Prairie Preserve: applying the fire-grazing interaction model. Pages 163-169 *in* Proceedings of the 23rd Tall Timbers Fire Ecology Conference: Fire in Grassland and Shrubland Ecosystems, Eds. R. E. Masters and K. E. M. Galley. Allen Press, Tall Timbers Research Station, Tallahassee, Florida, USA.
- Kindscher, K. and N. Scott. 1997. Land ownership and tenure of the largest land parcels in the Flint Hills of Kansas, USA. *Natural Areas Journal* 17:131-135.
- McGlenn, D. J. and M. W. Palmer. 2009. Modeling the sampling effect in the species-time-area relationship. *Ecology* 90:836-846.
- McPherson, R. A., C. A. Fiebrich, K. C. Crawford, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A. Kloesel, A. D. Melvin, H. Shrivastava, J. M. Wolfenbarger, J. P. Bostic, D. B. Demko, R. L. Elliott, S. J. Stadler, J. D. Carlson, and A. J. Sutherland. 2007. Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology* 24:301-321.
- Mehlich, A. 1984. Mehlich 3 soil test extraction modification of Mehlich 2 extractant. *Communications in Soil Science and Plant Analysis* 15:1409-1416.
- Milberg, P., J. Bergstedt, J. Fridman, G. Odell, and L. Westerberg. 2008. Observer bias and random variation in vegetation monitoring data. *Journal of Vegetation Science* 19:633-644.
- Oviatt, C. G. 1998. Geomorphology of Konza Prairie. Pages 35-47 *in* A. Knapp, J. Briggs, D. Hartnett and S. Collins (eds.), *Grassland Dynamics: Long-term ecological research in tallgrass prairie*, Eds. Oxford University Press, NY.

- Palmer, M. W. 1993. Potential biases in site and species selection for ecological monitoring. *Environmental Monitoring and Assessment* 26:277-282.
- Palmer, M. W. 2007a. Species-area curves and the geometry of nature. Pages 15-31 *in* *Scaling Biodiversity*, Eds. D. Storch, P. L. Marquet, and J. H. Brown. Cambridge University Press, Cambridge, UK.
- Palmer, M. W. 2007b. The vascular flora of the Tallgrass Prairie Preserve, Osage county, Oklahoma. *Castanea* 72:235-246.
- Palmer, M. W., J. R. Arévalo, M. C. Cobo, and P. G. Earls. 2003. Species richness and soil reaction in a northeastern Oklahoma landscape. *Folia Geobotanica* 38:381-389.
- Palmer, M. W., P. G. Earls, B. W. Hoagland, P. S. White, and T. Wohlgemuth. 2002. Quantitative tools for perfecting species lists. *Environmetrics* 13:121-137.
- Palmer, M. W., D. J. McGlenn, and J. F. Fridley. 2008. Artifacts and artificions in biodiversity research. *Folia Geobotanica* 43:245-257.
- Preston, F. W. 1960. Time and space and the variation of species. *Ecology* 41:611-627.
- Rosenzweig, M. L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, Cambridge, UK.
- USDA NRCS. 2008. The PLANTS Database. National Plant Data Center, Baton Rouge, LA 70874-4490 USA.
- White, E. P. 2007. Spatiotemporal scaling of species richness: patterns, processes, and implications. Pages 325-346 *in* *Scaling Biodiversity*, Eds. D. Storch, P. A. Marquet, and J. H. Brown. Cambridge University Press, Cambridge, UK.

White, E. P., P. B. Adler, W. K. Lauenroth, R. A. Gill, D. Greenberg, D. M. Kaufman, A. Rassweiler, J. A. Rusak, M. D. Smith, J. R. Steinbeck, R. B. Waide, and J. Yao. 2006. A comparison of the species-time relationship across ecosystems and taxonomic groups. *Oikos* 112:185-195.

TABLES

Table 2.1. The linear dimension and area of the five spatial grains/levels are noted below.

The grain at which a species was first encountered (the level) is denoted in

TGPP_spe.csv.

Level	Linear Dimension (m)	Area (m ²)
1	10.0×10.0	100
2	3.16×3.16	10
3	1.0×1.0	1
4	0.36×0.36	0.1
5	0.10×0.10	0.01

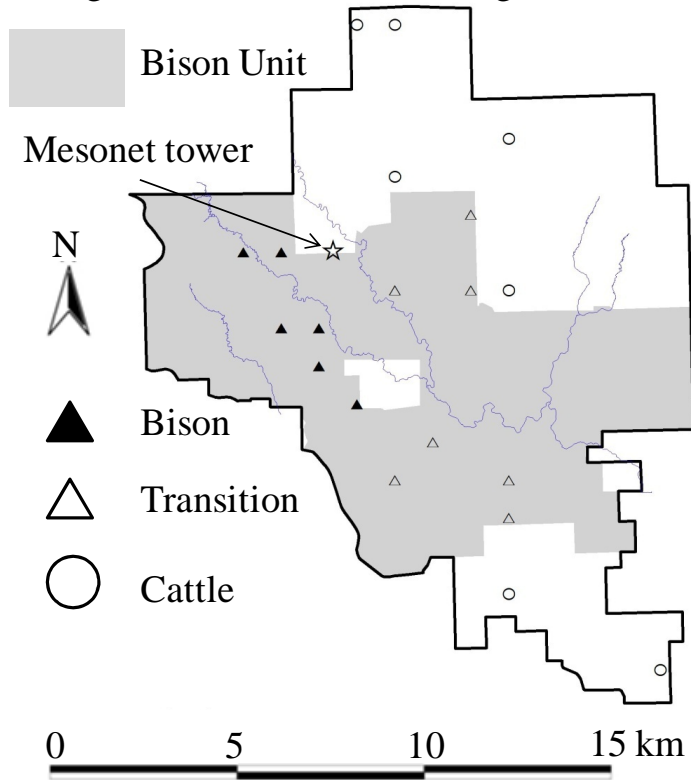
* a level of zero was recorded for unique species leaning over but not rooted in the quadrat

Table 2.2. Each species was placed in a visual cover class at the 100 m² grain (level 1).

Cover class	% range
1	trace
2	< 1
3	1-2
4	2-5
5	5-10
6	10-25
7	25-50
8	50-75
9	75-100

FIGURES

The Nature Conservancy's
Tallgrass Prairie Preserve, Osage Co., OK, USA



Preserve Area \approx 15,700 ha
Bison Unit Area \approx 8,500 ha (54%)*

Fig. 2.1 A map of the Tallgrass Prairie Preserve. The shaded area denotes the bison unit, which increased in area during the duration of the study. The Mesonet tower where the climate data was recorded is marked on the map as a star (☆). The twenty quadrats sampled each year of this study are displayed on the map (as triangles and circles). The sites that were grazed by bison at the beginning of the study (1998) are displayed with filled triangles (▲), those that transitioned during the study from cattle to bison are denoted by unfilled triangles (△), and the other cattle grazed samples are denoted by unfilled circles (○). *Area of bison unit as of May 1, 2008.

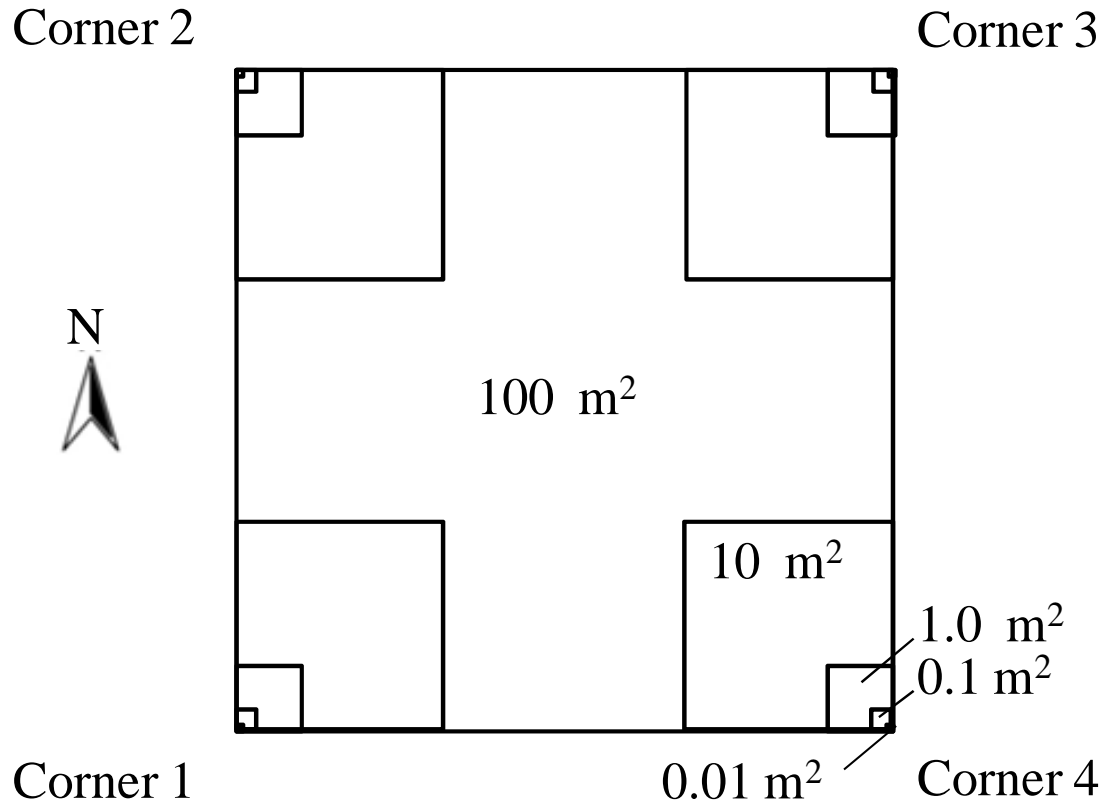


Fig. 2.2 Sampling design for the permanent plots. The presence of each species was recorded in each corner at each spatial grain and percent cover was visually estimated at the 100 m² grain.

CHAPTER III

MODELING THE SAMPLING EFFECT IN THE SPECIES-TIME-AREA RELATIONSHIP

Abstract. Recent empirical work in numerous systems has demonstrated the interdependence of spatial and temporal accumulation of species in the species-time-area relationship (STAR). I developed a process-based stochastic model for the STAR that assumes species neutrality, and compared the model's expectations to data collected on plant species in a tallgrass prairie. I varied two important aspects of the neutral species assemblage: evenness in the species pool and individual replacement rate (R). When R is larger than approximately 0.5 and evenness is intermediate to high, the neutral STAR generates patterns qualitatively similar to the empirical STAR. My model also indicates that space and time were not symmetrical in their effects on species accumulation, except in the special case of $R=1.0$. I observed both positive and negative time-by-area interactions in the sampling model, which indicates that nonzero interactions are not necessarily evidence of ecological processes. Furthermore, as accumulated richness approaches the size of the species pool, the time-by-area interaction becomes increasingly negative in my model. This suggests that negative time-by-area interactions should be expected *a priori* in empirical systems if rates of species accumulation decrease due to

increasing rarity of unique species. Given the wide range of STARs that the sampling model generated, the difficulty in estimating key parameters, and the complexity of assessing the relative abundance distribution and scale of the species pool, I cannot refute the sampling effect and I suggest caution in accepting ecologically-oriented explanations of empirical STARs.

Keywords: biodiversity, grassland, sampling effect, rarefaction effect, scaling, species richness, species turnover, STAR, scale dependence

INTRODUCTION

The species-area relationship (SAR) and the species-times relationship (STR) were recently unified conceptually and empirically as the species-time-area relationship (STAR) (Adler and Lauenroth 2003, Adler et al. 2005, White 2007). The basis for this unification was Preston's (1960) conceptual model that both the STR and SAR are driven by analogous sampling, ecological, and evolutionary processes and the empirical finding that the influences of space and time on accumulated richness are not mutually exclusive (Adler et al. 2005). Specifically, Adler et al. (2005) found that for all datasets they investigated, the slope of the log-log STR, w , decreased as the spatial scale of the sample increased. Simultaneously, the slope of the log-log SAR, z , decreased with increasing temporal scale. The rate at which w changes as a function of log area, and z changes as a function of log time, are identical to each other and are, by definition, an interaction effect (Ai and Norton 2003). Adler et al. (2005) referred to the decrease in w and z as area and time increased respectively as a negative time-by-area interaction. This finding

has renewed interest in Preston's model, and many have found that the SAR and STR are qualitatively similar (Rosenzweig 1995, 1998, McKinney and Frederick 1999, Hadly and Maurer 2001, Adler and Lauenroth 2003, White 2004, Adler et al. 2005, Fridley et al. 2006, Ulrich 2006, White et al. 2006b, Carey et al. 2007, White 2007) and likely driven by similar processes, but there is little empirical documentation of the drivers of the STR (White 2007).

Although Preston's argument is conceptually appealing, I still do not have an *a priori* reason to expect accumulation of species to be similar in space and time and to depend on one another. Such an argument is difficult to formulate due to the complexity and stochasticity of ecological and evolutionary drivers of species assemblages in space and time (Brown 1971, Pickett and White 1985, Chesson and Huntly 1989, Russell et al. 1995). However, it should be possible to formulate a model for the sampling component of the STAR from first principles (Rosenzweig 1998).

The *sampling effect* (also known as the *rarefaction effect*; Palmer et al. 2000a) describes a process by which species are gained by collecting more individuals or samples from a fixed "universe" (Palmer 1991, Goldberg and Estabrook 1998, Palmer et al. 2000, Gotelli and Colwell 2001). Although the sampling universe has not changed, the samples will not contain identical species due to chance. This process has also been referred to as random placement (Turner and Tjørve 2005) and passive sampling (Conner and McCoy 1979, McGuinness 1984a, 1984b). The influence of the sampling effect extends easily to samples taken in both space and time (Fisher et al. 1943, Williams 1943, Preston 1960, Brewer and Williamson 1994, White 2004). The sampling effect is generally thought to dominate species accumulation at fine temporal and spatial grains

although its effect never goes to zero (Preston 1960, Palmer and van der Maarel 1995, White 2004). Ecological explanations for SARs and STRs are not warranted unless the sampling effect can be rejected; therefore, it has played an important role as a null model (Conner and McCoy 1979, McGuinness 1984a, 1984b, Rosenzweig 1995, Gotelli and Graves 1996).

Many authors have modeled the sampling effect for the SAR (e.g. Arrhenius 1921, Williams 1943, Coleman 1981), and recently White (2004) has extended Coleman's (1981) approach to the STR, but none have yet modeled the sampling effect in both space and time on first principles (see Adler et al. 2005 for a randomization approach). If the STAR is to be used to test ecological hypotheses and provide evidence for ecological mechanisms (Adler et al. 2005), I must develop a more sophisticated understanding of its basic underpinnings, and the sampling effect is a solid starting point. Therefore, the purpose of this study is to derive a model to investigate the behavior of the sampling phase of the STAR and to qualitatively compare the results to an empirical STAR from a tallgrass prairie ecosystem.

SAMPLING EFFECT MODEL

Model assumptions

If the number of species added to a community (by increasing area or time) is driven only by a sampling effect, then the relationship between species and individuals will be solely determined by the distribution of relative abundances in the species pool and will not be related to the environment or intrinsic differences between species. Thus, species interactions are here assumed to be neutral (*sensu* Hubbell 2001). For ease of

presentation, I start by assuming the neutral community experiences zero-sum dynamics (later, I relax this assumption without changing the model outcomes); therefore, the number of individuals in the local community (J) is held constant over time, and birth and death rates are simplified into one term for all species: the individual replacement rate, R . The replacement rate is the probability that individuals in the local community are replaced by individuals chosen randomly from the species pool during one time unit. Following Williams (1943) and consistent with zero-sum dynamics, one unit of area is defined as the area occupied by an individual. Therefore, the number of individuals at any given time is equal to the sampled area (A). At each new time unit (T), the number of individuals added to the sample is the product of the number of individuals in the sampled area, A , and the replacement rate (R), and if the first sampled individual is defined to occur at $A = 1$, $T = 1$, then the cumulative number of individuals sampled through time after the first sampling is equal to $AR(T - 1)$. Therefore, the total number of individuals sampled for a given area and temporal duration is given by:

$$J = A + AR(T - 1) \tag{1}$$

From these assumptions I derive the sampling model for the expected number of species.

Model derivation

The expected number of species for J randomly sampled individuals is equal to one minus the probability that no individuals of species i are present in the sample summed over all the species in the species pool (S_P).

$$E[S(J | R, \mathbf{p})] = \sum_{i=1}^{S_P} [1 - (1 - p_i)^J] = S_P - \sum_{i=1}^{S_P} (1 - p_i)^J \tag{2}$$

Where \mathbf{p} is a vector of length S_p that represents the relative abundance distribution (RAD) of the species pool. After substitution of Eq. 1 into Eq. 2 the expected number of species is:

$$E[S(A, T | R, \mathbf{p})] = S_p - \sum_{i=1}^{S_p} (1 - p_i)^{A[1+R(T-1)]} \quad (3)$$

To characterize the relative increase in richness as a function of area and time for the sampling based STAR (Eq. 3), I also derived formulas for the partial derivatives of the natural logarithm (ln) of the expected richness (referred to as S_E for brevity) as a function of ln area and ln time respectively, z_E and w_E (see Appendix A). I used the subscript E to indicate these are the *expected* values derived from the sampling effect model (Eq. 3).

$$z_E = \frac{\partial \ln S_E}{\partial \ln A} = \frac{A}{S} \frac{\partial S_E}{\partial A} = \frac{-A[1+R(T-1)] \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} \ln(1-p_i)]}{S_p - \sum_{i=1}^{S_p} (1-p_i)^{A[1+R(T-1)]}} \quad (4)$$

$$w_E = \frac{\partial \ln S_E}{\partial \ln T} = \frac{T}{S} \frac{\partial S_E}{\partial T} = \frac{-TAR \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} \ln(1-p_i)]}{S_p - \sum_{i=1}^{S_p} (1-p_i)^{A[1+R(T-1)]}} \quad (5)$$

The relationship between z_E and w_E is a linear function

$$z_E = w_E \left(\frac{1-R}{TR} + 1 \right) \quad (6)$$

and thus when $R=1.0$, $z_E = w_E$. Equations 4 and 5 describe the independent influence of area and time on the accumulation of richness, but the equations also demonstrate that area and time cannot be decoupled. To quantify the interdependence between area and time on richness I calculated the second-order partial derivative of log richness with

respect to \ln area and \ln time together (see Appendix B), resulting in the following equation.

$$u_E = \frac{\partial^2 \ln S_E}{\partial \ln A \partial \ln T} = \frac{\partial w_E}{\partial \ln A} = A \frac{\partial w_E}{\partial A} = T \frac{\partial z_E}{\partial T} \quad (7)$$

The mathematical basis of Eq. 7 is identical to the fitted interaction parameter u in the “full model” of Adler et al. (2005) (i.e., both u and u_E are equal to $\frac{\partial^2 \ln S_E}{\partial \ln A \partial \ln T}$).

Therefore, I will refer to u_E using the term applied by Adler et al. (2005), the time-by-area interaction. There are two important differences between u and u_E : 1) u is a single fitted parameter and u_E is a continuous function, and 2) u is a statistically estimated interaction and u_E is a mathematically exact numerical interaction.

Equation 3 is conceptually a neutral model, not to be confused with the Unified Neutral Theory (UNT, Hubbell 2001). Unlike the UNT, here I am sampling from an infinite metacommunity with no dynamics in abundance, speciation, or extinction. However, like the UNT, biological interactions between individuals are assumed to be neutral and of no consequence except for the imposition of zero-sum dynamics. I do not expect many real communities to follow this relationship, but I seek to determine whether its qualitative behavior can reveal whether a sampling effect is a reasonable explanation for fine-scale STARS. The computer code to run my models in R v2.6.2 (R Developmental Core Team 2008) is provided as a supplement (S1).

Model parameters

I do not attempt here to estimate or constrain the parameters of the sampling model based on my empirical data set from a tallgrass prairie (described later), and therefore my model should not be considered a null model in this analysis. The clonal nature of many

plants in my study system cause estimates of abundance such as visual cover and stem counts to become decoupled from numbers of individuals. Furthermore, extrapolating from a series of observed plots to the relative abundance distribution of the entire species pool (which has an unknown spatial extent) is problematic (Bunge and Fitzpatrick 1993, Palmer 1995, Gotelli and Colwell 2001, Brose et al. 2003). Here, I assume an *a priori* RAD of the species pool, random sampling of individuals from the species pool. I then compare the qualitative behavior of my neutral model to my empirical STARS.

I calculated expectations for the neutral sampling model using a wide range of parameters. Preliminary results (not shown) were qualitatively robust to the size of the species pool, so I simply set the species pool to 800 which is slightly larger than the number of recorded vascular plants at my study site (Palmer 2007b). However, I recognize that the species pool may greatly exceed 800. I varied area and time from 1 to 16384 by successive doublings of scale and set the replacement rate (R) at five values: 0.00, 0.01, 0.10, 0.50, and 1.00. I generated nine RADs for the species pool which differed dramatically in evenness (following Siegel and German 1982) to investigate the influence that the relative abundance distribution (RAD) of the species pool had on the neutral STARS. One distribution was uniform (even), three distributions were lognormally distributed [LOGN(μ , σ)] with three levels of standard deviation ($\sigma = 1, 2,$ and 4), the other distributions included the geometric ($k = 0.9$), broken stick, Zipf ($\gamma = 1.3$), Zipf-Mandelbrot ($\gamma = 1.3, \beta = 100$) RADs (see Wilson 1991 for details on these RADs), and the uneven. The uneven distribution had almost complete dominance where one species had probability of 0.99 and the remaining 0.01 was distributed evenly among the other $S_p - 1$ species (Table 3.1, Fig. C1). I considered many different RADs for the

species pool because I have no *a priori* reason to prefer one model over the other for the entire species pool. My interest is primarily in documenting the variability in the expectations of the model. I varied the standard deviation of the lognormal because this provided a straightforward method for decreasing the evenness of an assemblage.

Model results

The neutral STAR displayed a large range of variation in response to the replacement rate (R) and to different RADs; however, there were some important generalizations that emerged. The SAR and STR generated with the broken stick, Zipf, and Zipf-Mandelbrot RADs were qualitatively similar to the results of the LOGN(0,1) RAD under the chosen parameters. To avoid redundancy, I will not discuss or display the results of these four RADs in this section. The SAR in log-log space was predominately linear except when it approached the size of the species pool in which case it became more concave-down (i.e., $\frac{\partial^2 \log S}{\partial^2 \log A} < 0$) in shape (Fig. 3.1, Fig. 3.2) or when evenness was low in which case it displayed concave-up (i.e., $\frac{\partial^2 \log S}{\partial^2 \log A} > 0$) curvature (Fig. 3.2). In contrast, the STR in log-log space was concave-up except when R was larger than approximately 0.50 and evenness was intermediate to high in which case the STR was linear to concave-down (Fig. 3.1, Fig. 3.2). It is worth noting that although the geometric RAD had an intermediate level of evenness (Table 3.1), it did not necessarily produce a SAR and STR intermediate to the other RADs (Fig. 3.2). This is because the majority of the species were very rare (median $p_i = 5.254 \times 10^{-20}$, Fig. C1) which caused the SAR and STR to be constrained below the asymptote of the species pool. As expected from Eq. 6, the SAR and the STR were identical when $R = 1.0$ because area

and time were varied at equal intervals and with this parameterization the scaling of species richness is equal in space and time (i.e., $z_E = w_E$); however, I did not expect the SAR and STR to display such similar patterns when R was as low as 0.50 (Fig. 3.1). As expected, when $R = 0$ the SAR and the STR were completely independent because there was no accumulation of species through time.

The time-by-area interaction, u_E , displayed a range of values $[-0.41, 0.25]$ and was only zero in special cases (Fig. 3.3). In Figure 3.3 I display only a subset of the parameterizations for clarity, but in appendix C there is a more thorough comparison (Fig. C4). Positive values of u_E resulted when evenness was low at relatively small scales. A positive time-by-area interaction means the SARs diverge from one another as the temporal scale increases, and equivalently the STRs diverge as the spatial scale increases (e.g., at small scales for uneven RAD, see Fig. 3.2). As the SARs and STRs converged towards the species pool (or asymptote of the model), z_E and w_E decreased resulting (as one may expect) in predominately negative time-by-area interactions. Negative time-by-area interactions were also observed in the geometric RAD when the expected number of individuals was well below the size of the species pool because this RAD had a high degree of rarity. Negative values were more commonly observed across the range of parameterizations I chose (Fig. 3.3). When the expected richness finally saturated at the size of the species pool, z_E and w_E became fixed at zero and therefore u_E was also forced to go to zero (see Eq. 7 and Fig. 3.3). The time-by-area interaction was also zero when the replacement rate was zero because there was no accumulation of species through time; the SARs and the STRs were therefore parallel.

EMPIRICAL PATTERNS

Empirical analysis

A temporal moving-window approach (Adler et al. 2005) was used to construct the empirical STAR from the case study quadrats described in Chapter II (Fig 2.2). Species richness at each spatial scale was calculated for every possible one year time span, then calculated for every two year time span, and so on to a time span of ten years. The richness values were then averaged across the four corners of each quadrat and across all twenty quadrats for each time span – area combination. My method of data collection and analysis are equivalent to what Carey et al. (2007) referred to as nested and complete nested designs for the SAR and STR, respectively.

Empirical results

The empirical SAR and the STR displayed strong patterns of scale dependence (Fig 3.4). The rate of change of log richness decreased as log area increased resulting in a concave-down SAR. The STR appeared more linear in log-log space; however, closer examination of the rates of change in log richness as a function of time (not shown) indicated that the STR was also concave-down although not as extremely as the SAR. More relevant to the characterization of the empirical STAR was the observation that the slope of the SAR decreased as the temporal duration increased, and equivalently the slope of the STR decreased as the spatial scale was increased. In other words, there was a strong negative time-by-area interaction.

Difficulties in estimating the RAD of the species pool, plus other key parameters, preclude us from making rigorous quantitative comparisons between my model and my empirical system. It is important to note that the units of time and space in my sampling

model are not arbitrary because they are defined on the basis of R and individual size respectively (i.e., individual size equals one unit of area). It may be possible in some ecological systems to choose a spatial unit that is defined on the basis of average individual size and a temporal scale that corresponds to individual temporal turnover; however, for most grassland systems (with large variation in body size, clonality, unknown longevities, and unknown spatial scale of the species pool) this remains problematic.

DISCUSSION

My objective was to investigate the sampling phase of the STAR with a neutral sampling model and to qualitatively compare my model with empirical data. The sampling model indicated that a neutral, zero-sum sampling process is a simple mechanism by which the SAR and the STR can be linked. Furthermore, area and time do not exert symmetrical sampling effects on the STAR except in the special case in which all individuals turnover each unit of time. Different parameters of the sampling model can lead to a wide diversity of STARs, and negative values for the time-by-area interactions are expected when richness is limited by a finite species pool.

Spatial and temporal scale dependence

The empirical and neutral SAR and STR both displayed systematic patterns of spatial and temporal scale dependence, respectively. Although scale dependence is not considered in the most common models of the SAR and STR (the power and exponential models), the scale-dependent behavior I observed was not unexpected. It is well documented that the slope of the SAR exhibits marked spatial scale dependence (Shmida

and Wilson 1985, Palmer and White 1994, Rosenzweig 1995, Plotkin et al. 2000b, Crawley and Harral 2001, Fridley et al. 2005, Turner and Tjørve 2005, Fridley et al. 2006, Palmer 2007a), and it appears the STR exhibits temporal scale dependence although this has not been documented as extensively (Rosenzweig 1998, White 2007). The general pattern of spatial scale dependence of the global log-log SAR (and likely of temporal scale dependence in the STR) is characterized by a steep slope over fine and broad scales and a shallow slope over intermediate scales. My study spanned five orders of magnitude in space but only one in time; therefore, I expected the STAR to potentially show more negative spatial scale dependence than temporal scale dependence. Negative spatial and temporal scale dependence (concave-down curvature) was also observed for the neutral STAR when the species pool or a high degree of rarity exerted a negative influence on the rate of accumulated richness.

The different patterns of scale dependence between the neutral SAR and the neutral STR illustrate that area and time exert non-equivalent sampling effects on accumulated richness. In contrast to previous empirical studies I suggest that the STR is not simply a temporal analogue of the SAR with respect to the sampling effect. This is especially true when the replacement rate in a community falls below approximately 0.5.

Time-by-area interaction

The third type of scale dependence I observed for both the neutral and empirical STARs was the time-by-area interaction (Figs. 3.1-3.3, Fig 3.4). The strength and pattern of the time-by-area interaction is the most important quantitative description of the STAR. If the interaction is positive then the curves of the STAR are diverging from one another, and if they are negative then they are converging (to an asymptote for example).

If the interaction is zero then the curves of the STAR will be parallel which indicates that the SAR and STR are independent of one another. My empirical results as well as my sampling model demonstrated that the interaction can be quite strong and displays systematic patterns of scale dependence.

I stress that the ‘interaction’ in the neutral model is a purely *numerical* interaction, because no interaction between space and time is incorporated in the model. As neutrality can result in both negative and positive interactions depending on the characteristics of the species pool and the replacement rate, deviations of u_E from zero cannot be used to infer processes of ecological significance within communities.

To date all published empirical STARs (including my empirical results) have displayed negative time-by-area interactions (Adler et al. 2005), and my neutral model indicates that this behavior should be expected *a priori* if the species pool is finite. Specifically as the SAR or STR approached the species pool in my model, the interaction became increasingly negative until the asymptote was achieved in which case the interaction was forced to zero. Although in nature asymptotes do not exist for empirical SARs or STRs (Williamson et al. 2001), in small scale patterns the rarity of unsampled species will cause accumulation rates to decrease (Palmer and White 1994, Rosenzweig 1995, Plotkin et al. 2000b). Thus even though empirical STARs are not strictly asymptotic, one should expect a negative time-by-area interaction if the probability of encountering a new species is quite low and/or decreasing. Regardless of which processes are driving species accumulation (sampling or ecological), the nature of this expectation does not change. Thus, I have established a general *a priori* expectation that

the SAR and STR should not be independent if rarity decreases the relative rate of species accumulation in space or time.

Importance of the RAD of the species pool

My comparison of the nine different RADs for the species pool indicated the choice of model and parameterization of the RAD can lead to dramatically different STARs. This is most clearly illustrated for the geometric RAD, in which the degree of rarity was so severe that the SAR and STR converged well below the size of the species pool (Fig. 3.2, Fig. 1C). It was not entirely unexpected that the broken stick, Zipf, and Zipf-Mandelbrot would generate similar STARs as the lognormal RAD because these distributions had similar slopes and curvature on the rank-log probability plot with the parameterizations I chose (Fig. 1C). However, it is worth noting that the Zipf and Zipf-Mandelbrot RADs can also give very similar results to the geometric RAD under certain parameterizations (not shown).

Relaxing the zero-sum assumption

In my model of the sampling effect, I assume a fixed number of individuals occupying a fixed area. This assumption can be relaxed if I recognize that the expected number of species is a function of the expected number of individuals. If individuals are distributed in space and time by a stationary process (i.e., the mean and variance of the process do not vary) then the expected cumulative number of individuals will not be altered by variance in this number. Therefore, the density (number of individuals per unit area or ρ , after Hubbell 2001) is a random variable with expectation equal to the average density or $\bar{\rho}$. Each time unit, a constant fraction (R) of the local community is recruited and distributed in space such that $\bar{\rho}$ is not changed. In this sense, R is the long-term

average proportion of the community that 'turns over' and is replaced by new individuals, even if the actual number of new individuals varies by time unit. Therefore the new equation for expected number of individuals is:

$$J = \bar{\rho}A + \bar{\rho}AR(T - 1) \tag{8}$$

Equation 8 is very similar to Eq. 1 and could be inserted into Eq. 3 in a similar manner; however, now I am considering that density is a random variable. The replacement rate (R) is no longer the probability that an individual is replaced, but rather it is the fraction of the local community that recruits each generation. The assumption of stationarity does not prohibit the possibility of aggregation or clumping of individuals in space or time, it simply requires that the mean and variance of the *process* of clumping does not change (Wagner and Fortin 2005). These changes to my sampling model imply that my results should generally hold for communities with variable (but stationary) population sizes and for communities in which individuals display variable (but stationary) patterns of spatial aggregation as long as the individuals are randomly drawn from the species pool (i.e., species neutrality still holds).

Relevance and future utility of the model

My primary motivation in developing a model of the sampling effect was to generate expectations of the STAR under a set of constraining assumptions. If difficulties in estimating key parameters can be overcome, my model may also provide a null model for the STAR which can be compared to ecological models incorporating non-neutral dynamics and spatial and temporal trends in abundance. Systems in which unique individuals can be identified and monitored over time, such as small mammal

communities or annual plant communities may provide ideal systems for studying the sampling phase of the STAR in the future.

My model can be viewed as an analytical extension of two previous approaches to modeling aspects of the STAR. A neutral simulation model was used to predict values of average richness, z , and w , and to compare these with estimates from an empirical grassland community (Adler 2004). Adler's (2004) simulation model produced realistic looking SARs and STRs, but it was not able to simultaneously generate reasonable estimates for all the empirical community's terms and it did not consider the time-by-area interaction. My model is also an extension of a randomization-based null model in which individuals were randomly drawn without replacement from the empirically observed pool until the observed number of individuals for a particular site in a particular year was achieved (Adler et al. 2005). This null model implicitly assumed that the replacement rate (R) was equal to one. Adler et al. (2005) found that their sampling model did a poor job of predicting the empirical STAR when compared with multiple regression models. Both of these previous models (like my own model) rely on the assumption of random spatial and temporal structure in individuals (i.e., individuals are independent of each other). My model extends the Adler (2004) simulation model by explicitly considering the time-by-area interaction in an analytical framework, and it extends the Adler et al. (2005) null model by analytically allowing R to vary from one. Although these are improvements, the assumptions of spatial and temporal homogeneity as well as species neutrality (that all the models require) could be relaxed in future models of the STAR. There have been many successful attempts to incorporate information on spatial aggregation into models of the SAR (e.g. Plotkin et al. 2000b, He and Legendre 2002,

Picard et al. 2004) and recently White and Gilchrist (2007) also demonstrated that temporal population structure can be incorporated into models of the STR to yield more accurate predictions.

Although the assumptions of my model are necessarily false, it was still able to reproduce patterns that were similar to my data when evenness and the replacement rate were intermediate to high. Furthermore, the particular shapes and attributes of the sampling STAR strongly depend upon the replacement rate and the particular RAD (e.g. lognormal vs. geometric), such that almost any conceivable monotonic SAR or STR can result. Therefore, I cannot rule out the possibility that my empirical patterns are shaped by sampling effects. These findings imply that a sampling effect should be considered a null hypothesis for observed STARs (Adler et al. 2005, White 2007), and that it may be difficult to differentiate a sampling effect from ecological patterns without detailed information on turnover rates, the nature of the RAD, and the number of individuals per unit area. Given these data, analytical and randomization-based null models which incorporate fewer constraining assumptions should aid in differentiating the ecological and sampling effects on the STAR in the future.

Conclusions

I formulated a sampling-based model that yielded a diversity of STARs based on the nature of the species pool and individual replacement rate. Unless the individual replacement rate is equal to one, time and space are not symmetrical. Due to the diversity of outcomes, and the difficulty in estimating key parameters, I cannot discount the possibility that the sampling effect structured my empirical STAR. Strong time-by-area interactions are not evidence of ecological drivers of the STAR. Furthermore, the

existence of a finite species pool leads to an expectation of negative time-by-area interactions regardless of the processes shaping the SAR and STR. Lastly, both my simulated and empirical STARS indicated that spatial and temporal scale dependence are fundamental characteristics of fine-scale species accumulation relationships that should be incorporated into future statistical models of the STAR.

ACKNOWLEDGEMENTS

DJM received funding by the United States Environmental Protection Agency (EPA) under the Greater Research Opportunities (GRO) Graduate Program. EPA has not officially endorsed this publication and the views expressed herein may not reflect the views of the EPA. DJM acknowledges B.C. Crauder for mathematical assistance, J.A. Steets and other members of LIBRA for helpful advice, and M.J. Gilmore, J.E., and J.M. McGlenn for encouraging words. MWP acknowledges NSF Grant Number EPS-0447262 for partial financial support. N.J. Gotelli, E.P. White, and an anonymous reviewer gave constructive comments that improved the manuscript.

LITERATURE CITED

- Adler, P. B. 2004. Neutral models fail to reproduce observed species-area and species-time relationships in Kansas grasslands. *Ecology* **85**:1265-1272.
- Adler, P. B. and W. K. Lauenroth. 2003. The power of time: spatiotemporal scaling of species diversity. *Ecology Letters* **6**:749-756.
- Adler, P. B., E. P. White, W. K. Lauenroth, D. M. Kaufman, A. Rassweiler, and J. A. Rusak. 2005. Evidence for a general species-time-area relationship. *Ecology* **86**:2032-2039.
- Ai, C. and E. C. Norton. 2003. Interaction terms in logit and probit models. *Economics Letters* **80**:123-129.
- Arrhenius, O. 1921. Species and area. *Journal of Ecology* **9**:95-99.
- Brewer, A. and M. Williamson. 1994. A new relationship for rarefaction. *Biodiversity and Conservation* **3**:373-379.
- Brose, U., N. D. Martinez, and R. J. Williams. 2003. Estimating species richness: Sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology* **84**:2364-2377.
- Brown, J. H. 1971. Mammals on mountaintops - nonequilibrium insular biogeography. *American Naturalist* **105**:467-478.
- Bunge, J. and M. Fitzpatrick. 1993. Estimating the Number of Species - a review. *Journal of the American Statistical Association* **88**:364-373.
- Carey, S., A. Ostling, J. Harte, and R. del Moral. 2007. Impact of curve construction and community dynamics on the species-time relationship. *Ecology* **88**:2145-2153.

- Chesson, P. and N. Huntly. 1989. Short-term Instabilities and long-term community dynamics. *Trends in Ecology & Evolution* **4**:293-298.
- Coleman, B. D. 1981. On random placement and species-area relations. *Mathematical Biosciences* **54**:191-215.
- Conner, E. F. and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. *American Naturalist* **113**:791-833.
- Crawley, M. J. and J. E. Hurrell. 2001. Scale dependence in plant biodiversity. *Science* **291**:864-868.
- Fisher, R. A., A. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* **12**:42-58.
- Fridley, J. D., R. K. Peet, E. van der Maarel, and J. H. Willems. 2006. Integration of local and regional species-area relationships from space-time species accumulation. *American Naturalist* **168**:133-143.
- Fridley, J. D., R. K. Peet, T. R. Wentworth, and P. S. White. 2005. Connecting fine- and broad-scale species-area relationships of Southeastern US Flora. *Ecology* **86**:1172-1177.
- Goldberg, D. E. and G. F. Estabrook. 1998. Separating the effects of number of individuals sampled and competition on species diversity: an experimental and analytic approach. *Journal of Ecology* **86**:983-988.
- Gotelli, N. J. and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* **4**:379-391.

- Gotelli, N. J. and G. R. Graves. 1996. Null models in ecology. Smithsonian Institution, Washington D.C.
- Hadly, E. A. and B. A. Maurer. 2001. Spatial and temporal patterns of species diversity in montane mammal communities of western North America. *Evolutionary Ecology Research* **3**:477-486.
- Hamilton, R. G. 1996. Using fire and bison to restore a functional tallgrass prairie landscape. Pages 208-214 *in* Transactions of the 61st North American Wildlife and Natural Resources Conference. Wildlife Management Institute, Washington D.C.
- He, F. L. and P. Legendre. 2002. Species diversity patterns derived from species-area models. *Ecology* **83**:1185-1198.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.
- McGuinness, K. A. 1984a. Equations and explanations in the study of species-area curves. *Biological Reviews* **59**:423-440.
- McGuinness, K. A. 1984b. Species area relations of communities on intertidal boulders - testing the null hypothesis. *Journal of Biogeography* **11**:439-456.
- McKinney, M. L. and D. L. Frederick. 1999. Species-time curves and population extremes: ecological patterns in the fossil record. *Evolutionary Ecology Research* **1**:641-650.
- Palmer, M. W. 1991. Patterns of species richness among North Carolina hardwood forests: tests of two hypotheses. *Journal of Vegetation Science* **2**:361-366.
- Palmer, M. W. 1995. How should one count species? *Natural Areas Journal* **15**:124-136.

- Palmer, M. W. 2007a. Species-area curves and the geometry of nature. Pages 15-31 *in* D. Storch, P. L. Marquet, and J. H. Brown, editors. *Scaling Biodiversity*. Cambridge University Press.
- Palmer, M. W. 2007b. The vascular flora of the Tallgrass Prairie Preserve, Osage county, Oklahoma. *Castanea* **72**:235-246.
- Palmer, M. W., D. A. Clark, and D. B. Clark. 2000a. Is the number of tree species in small tropical forest plots nonrandom? *Community Ecology* **1**:95-101.
- Palmer, M. W. and E. van der Maarel. 1995. Variance in species richness, species association, and niche limitation. *Oikos* **73**:203-213.
- Palmer, M. W. and P. S. White. 1994. Scale dependence and the species-area relationship. *American Naturalist* **144**:717-740.
- Palmer, M. W., T. Wohlgemuth, P. Earls, J. R. Arévalo, and S. D. Thompson. 2000b. Opportunities for long-term ecological research at the Tallgrass Prairie Preserve, Oklahoma. Pages 123-128 *in* Cooperation in long term ecological research in central and eastern Europe: proceedings of the ILTER regional workshop, Budapest, Hungary. Oregon State University, Corvallis, OR.
- Picard, N., M. Karembe, and P. Birnbaum. 2004. Species-area curve and spatial pattern. *Ecoscience* **11**:45-54.
- Pickett, S. T. A. and P. S. White. 1985. *The Ecology of Natural Disturbance and Patch Dynamics*. Academic Press, San Diego, CA.
- Pielou, E. C. 1975. *Ecological Diversity*. Wiley, New York, NY.
- Plotkin, J. B., M. D. Potts, D. W. Yu, S. Bunyavejchewin, R. Condit, R. Foster, S. Hubbell, J. LaFrankie, N. Manokaran, L. H. Seng, R. Sukumar, M. A. Nowak,

- and P. S. Ashton. 2000. Predicting species diversity in tropical forests. *Proceedings of the National Academy of Sciences of the United States of America* **97**:10850-10854.
- Preston, F. W. 1960. Time and space and the variation of species. *Ecology* **41**:611-627.
- R Developmental Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge, UK.
- Rosenzweig, M. L. 1998. Preston's ergodic conjecture: the accumulation of species in space and time. Pages 311-348 *in* M. L. McKinney and J. A. Drake, editors. Biodiversity dynamics: turnover of populations, taxa, and communities. Columbia University Press, New York, NY.
- Russell, G. J., J. M. Diamond, S. L. Pimm, and T. M. Reed. 1995. A century of turnover - community dynamics at three timescales. *Journal of Animal Ecology* **64**:628-641.
- Shmida, A. and M. V. Wilson. 1985. Biological determinants of species diversity. *Journal of Biogeography* **12**:1-20.
- Siegel, A. F. and R. Z. German. 1982. Rarefaction and taxonomic diversity. *Biometrics* **38**:235-241.
- Turner, W. R. and E. Tjørve. 2005. Scale-dependence in species-area relationships. *Ecography* **28**:721-730.
- Ulrich, W. 2006. Decomposing the process of species accumulation into area dependent and time dependent parts. *Ecological Research* **21**:578-585.

- Wagner, H. H. and M. J. Fortin. 2005. Spatial analysis of landscapes: Concepts and statistics. *Ecology* **86**:1975-1987.
- White, E. P. 2004. Two-phase species-time relationships in North American land birds. *Ecology Letters* **7**:329-336.
- White, E. P. 2007. Spatiotemporal scaling of species richness: patterns, processes, and implications. Pages 325-346 *in* D. Storch, P. A. Marquet, and J. H. Brown, editors. *Scaling Biodiversity*. Cambridge University Press.
- White, E. P., P. B. Adler, R. A. Laurenroth, R. A. Gill, D. Greenberg, D. M. Kaufman, A. Rassweiler, J. A. Rusak, M. D. Smith, J. R. Steinbeck, R. B. Waide, and J. Yao. 2006. A comparison of species-time relationship across ecosystems and taxonomic groups. *Oikos* **112**:185-195.
- White, E. P. and M. A. Gilchrist. 2007. Effects of population-level aggregation, autocorrelation, and interspecific association on the species-time relationship in two desert communities. *Evolutionary Ecology Research* **9**:1329-1347.
- Williams, C. B. 1943. Area and number of species. *Nature* **152**:264-267.
- Williamson, M., K. J. Gaston, and W. M. Lonsdale. 2001. The species-area relationship does not have an asymptote! *Journal of Biogeography* **28**:827-830.
- Wilson, J. B. 1991. Methods for fitting dominance/diversity curves. *Journal of Vegetation Science* **2**:35-46.

TABLES

Table 3.1. The evenness of each of the relative abundance distributions (RADs) calculated using Pielou's (1975) evenness index (E). This evenness index is the Shannon information scaled by maximum information and ranges between 0 and 1.

RAD	E
<i>High evenness</i>	
Even	1.00
Zipf-Mandelbrot ($\gamma = 1.3, \beta = 100$)	0.95
broken stick	0.94
LOGN(0,1)	0.93
<i>Intermediate evenness</i>	
LOGN(0,2)	0.71
Zipf ($\gamma = 1.3$)	0.55
Geometric ($k = 0.9$)	0.49
<i>Low evenness</i>	
LOGN(0,4)	0.20
uneven	0.02

FIGURES

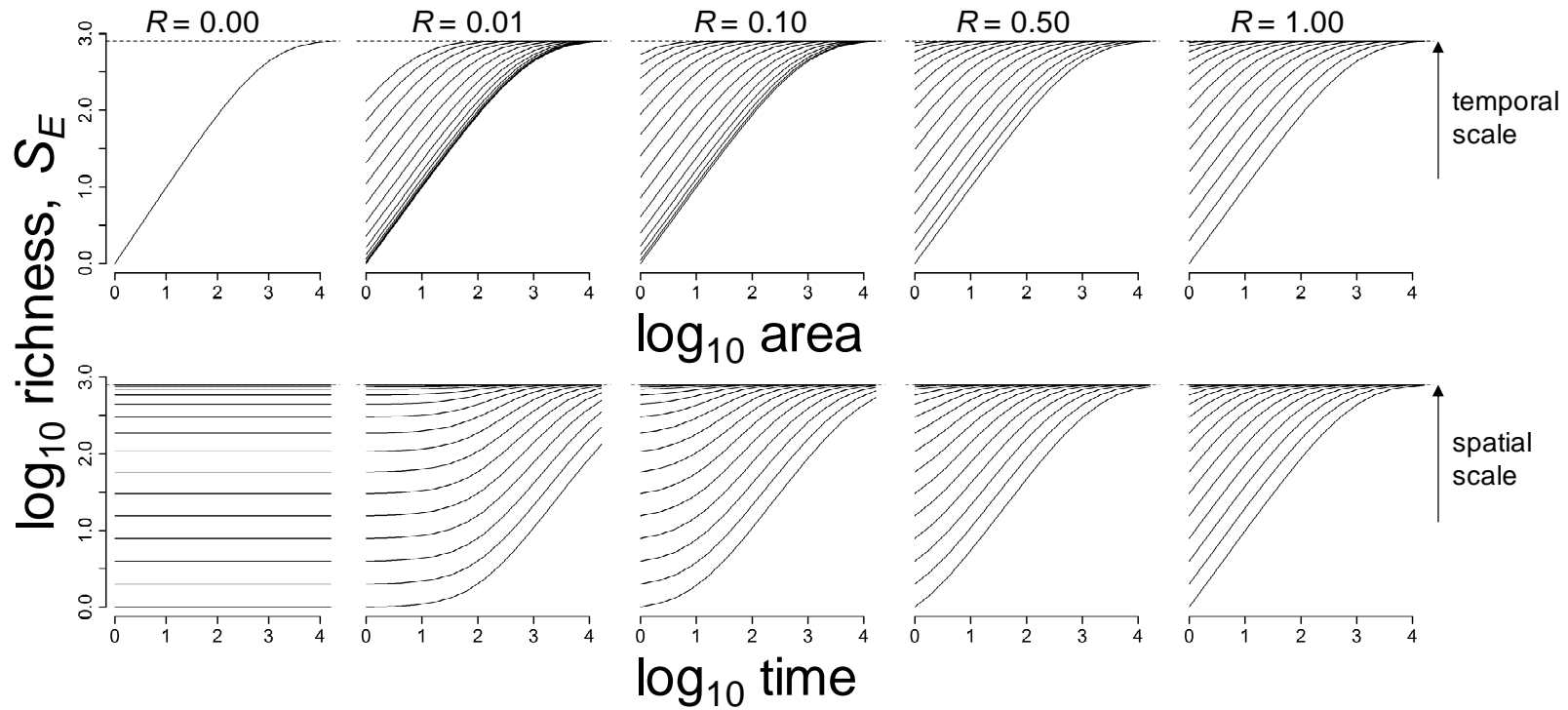


Fig. 3.1. The SAR and STR calculated with the sampling model (Eq. 3). The evenness in the assemblage was high (LOGN(0,1), $E = 0.93$) and the replacement rate (R) was at one of five levels: 0.00, 0.01, 0.10, 0.50, or 1.00. For each curve the species pool (S_p) was set to 800 and points were calculated from 1 to 16384 by successive doublings of scale. The size of the species pool is indicated as the horizontal dashed grey line, and the arrows on the right margin indicate the influence of increasing the scale on the relationships.

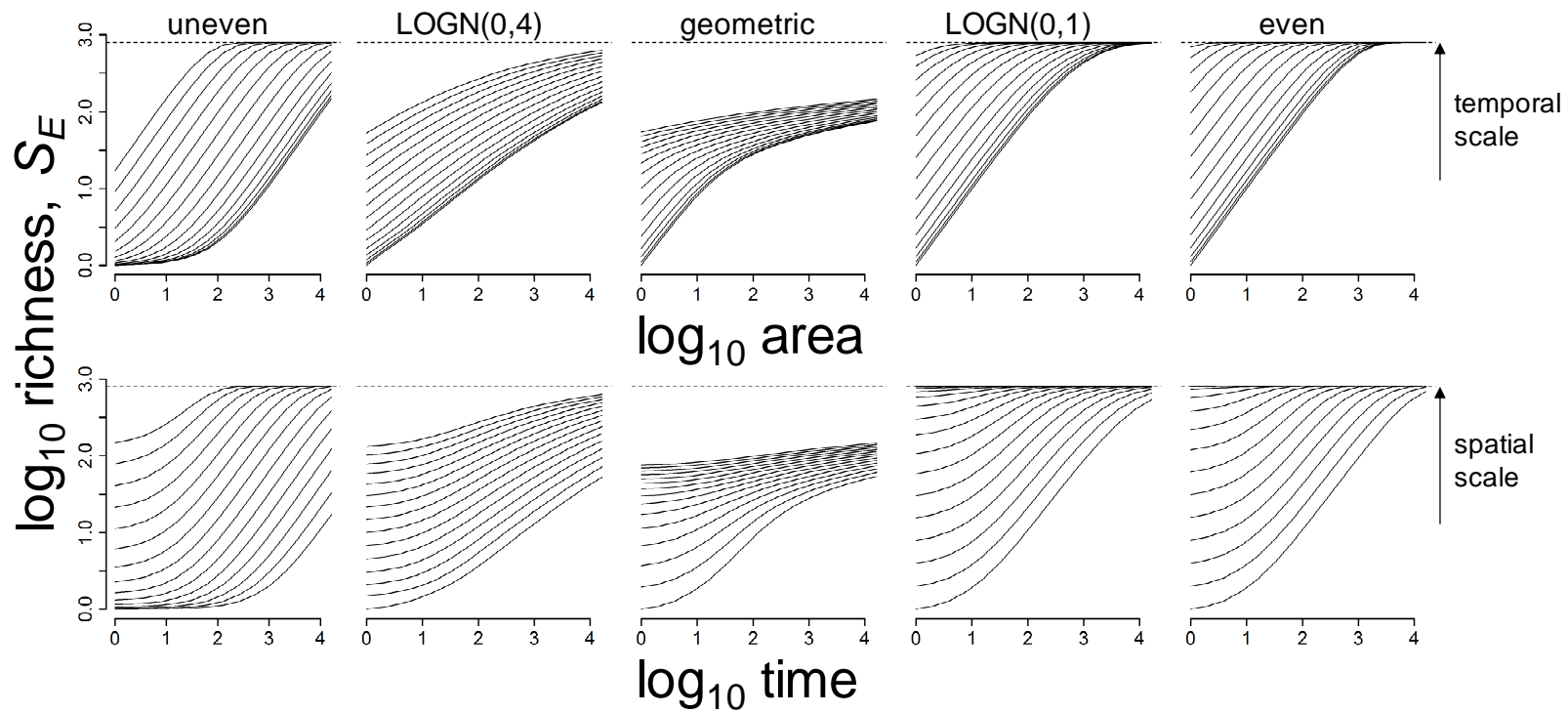


Fig. 3.2. The SAR and STR calculated with the sampling model (Eq. 3). The replacement rate (R) was set to 0.10 and the RAD of the species pool was either uneven, LOGN(0,4), geometric ($k = 0.9$), LOGN(0,1), or even. The other parameters of the model were the same as in Fig. 1.

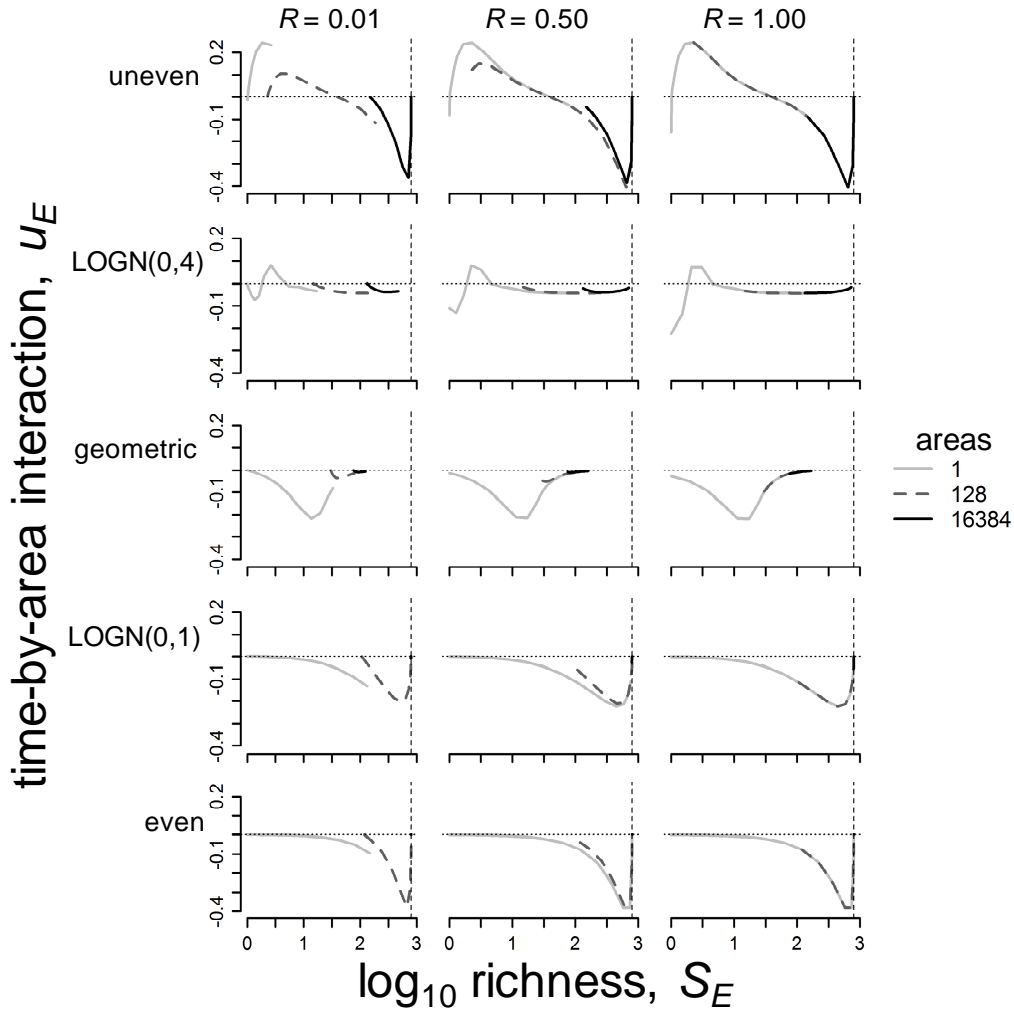


Fig. 3.3. The time-by-area interaction (u_E) as a function of log richness (S_E) for three values of R (columns) and the five RADs (rows). Each curve was generated by holding area constant at three different values (see legend) and varying time. The dotted horizontal line indicates zero, and the dashed vertical line indicates the size of the species pool (S_p). Positive values of u_E were only observed at small scales under low evenness. The most negative values of u_E were observed near the species pool, and when richness finally reached the species pool (or in the case of the geometric RAD the probability of a new species became extremely small) u_E was forced to zero. Similar results occur when time instead of area is fixed.

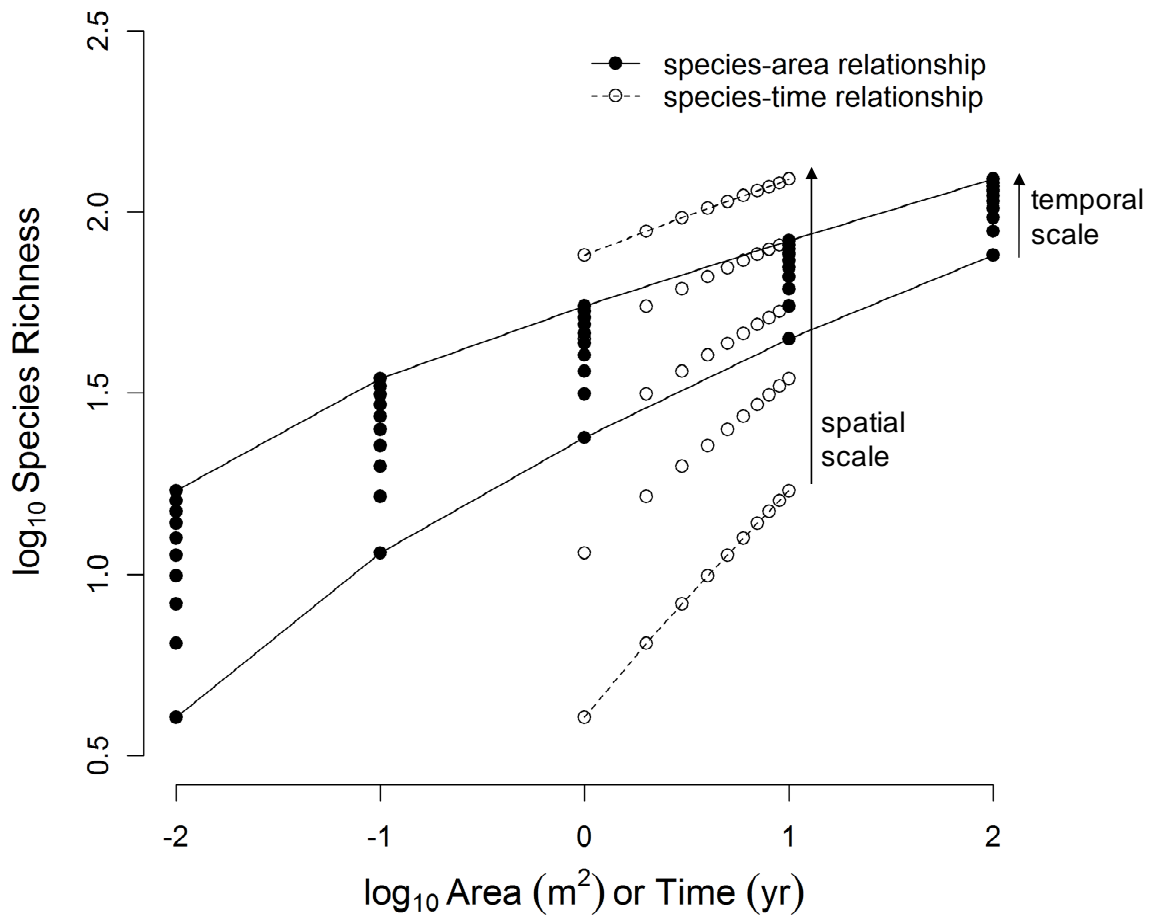


Fig. 3.4. The empirical SAR (filled circles) and STR (open circles) for data from the TGPP. The SAR is not as steep as the STR and is distinctly concave-up. Note how the slope of the SAR decreased as the temporal scale increased, and the slope of the STR decreased as the spatial scale increased. Only the lowermost and uppermost curves of the SAR and STR are connected with lines to minimize clutter and to draw attention to the change in slope as a function of scale. Comparisons between the SAR and STR should be only made along the ordinate as the exact position along the abscissa depends on an arbitrary choice of measurement units.

CHAPTER IV

QUANTIFYING THE INFLUENCE OF ENVIRONMENTAL TEXTURE ON THE RATE OF SPECIES TURNOVER – EVIDENCE FROM TWO HABITATS

Abstract. The Environmental Texture Hypothesis proposes that the spatial geometry or texture of the environment influences the rate at which new species are accumulated in space or time. This suggests that regions that exhibit a larger rate of environmental distance decay should exhibit more rapid rates of species turnover. This hypothesis should apply over any range of scales where the quantified environment is correlated with species composition. I tested for a positive relationship between the rate of change in soil cations and vascular plant species composition of a grassland and woodland habitat. I recorded presence-absence data along a 1,883 m transect in each habitat and estimated the rate of turnover and environmental distance decay for spatial lags of 1 to 41 m. I found that the soil cation environment explained spatial patterns of species composition more accurately in the grassland habitat compared to the woodland habitat. Consequently, the rate of change in soil cations as a function of spatial distance was significantly positively correlated with the rate of species turnover in the grassland but not the woodland. My study suggests that one of the central premises of the Environmental Texture Hypothesis is relevant for local patterns of species turnover if the

environment influences species composition.

Keywords: distance decay, scaling of biodiversity, spatial dependence, spatial autocorrelation, Euclidean variogram, soil cations, crosstimbers, tallgrass prairie

INTRODUCTION

Conservation of species diversity depends on my ability to explain and predict variation in patterns of species turnover or distance decay (Wilson and Mohler 1983, Nekola and White 2002, Buckley and Jetz 2008). Historically patterns of turnover were related almost entirely to the degree of environmental change between samples (e.g., Whittaker 1960). However, ecologists later recognized that spatial or temporal distance between samples, irrespective of environmental change, was positively correlated with the degree of species turnover (Nekola and White 1999). Legendre (1993) referred to these endogenously driven patterns of species turnover as false gradients of species composition. From a modeling standpoint, the importance of false gradients shifted the focus away from the environment and towards biological processes such as dispersal limitation (e.g., Hubbell 2001). The recent increase in models that address the inherent spatial autocorrelation in ecological communities provides ecologists with powerful new tools for dissecting spatial patterns; however, in the process the role of the environment which can induce spatial dependence in the community (i.e. true gradients), has been neglected.

The Environmental Texture Hypothesis (ETH) is one recent attempt to shift the debate back towards the relevance of environmental heterogeneity on patterns of species turnover (Palmer 2007). Specifically the ETH suggests that the rate of species

accumulation is determined by the rate new environments are sampled (i.e. the texture or geometry of the environment influences the rate of species accumulation). This hypothesis may offer an explanation for the triphasic pattern of the species-area relationship (SAR). It predicts that at local and global scales the environment varies in a smooth manner (i.e., low fractal dimension), which causes rapid accumulation of new species. At intermediate scales the environment is expected to be rougher (i.e., high fractal dimension) in which case new species accumulate more slowly (because many of the species in the region are encountered in the first few samples).

Although the ETH makes several specific predictions related to the triphasic SAR, testing these predictions is difficult due to a lack of multi-scale richness data that span many orders of magnitude. However, testing whether the rate of species accumulation is determined by the rate of change in the environment is logistically feasible because this prediction should apply at any range of scales over which the environment influences species composition. Therefore, the purpose of my study was to test the ETH by examining if areas of more rapid environmental distance decay (DD) possess more rapid species accumulation.

Examining this hypothesis does not necessarily require the usage of the SAR which is, ideally, constructed from spatially nested or contiguous quadrats. A powerful but more flexible alternative approach, in terms of sampling design, is to quantify the rate of community DD. However, care must be taken when selecting the choice of metrics for quantifying community DD because different metrics of community dis/similarity or species turnover carry different interpretations and sometimes important subtleties (Koleff et al. 2003). The semivariance of the Euclidean community variogram (aka, the

variogram of complementarity) is the ideal metric to test the this hypothesis because it is equivalent to the expected number of unique species encountered at a given spatial distance from the focal quadrat (Wagner 2003, 2004).

A test of my hypothesis relies on the assumption that the environmental variables we include in the calculation of environmental DD are relevant predictors of species composition. Therefore, the objectives of my study were to: 1) identify relevant environmental variables explaining the spatial structure of species composition, 2) quantitatively describe the spatial geometry of the environment and species composition, and 3) test for a positive correlation between the rate of environmental and community DD. I accomplished these objectives with data from a grassland and woodland habitat along two 1,883 m transects.

METHODS

Study Site

I conducted my study on The Nature Conservancy's Tallgrass Prairie Preserve between 36.73° and 36.90° N latitude, and 96.32° and 96.49° W longitude, in Osage County, Oklahoma. The vegetation of the preserve is composed of approximately 90% grasslands and 10% forests or woodlands. The grasslands are dominated by tallgrass prairie plant species such as *Andropogon gerardii*, *Sorghastrum nutans*, *Sporobolus compositus*, *Panicum virgatum*, and *Schizachyrium scoparium*. Shortgrass prairie habitat occurs to a lesser extent on more xeric sites and is dominated by *Bouteloua* spp. The woodland habitats can be classified primarily as Cross Timbers, and they vary in structure from open and savanna-like to rather dense closed canopy forests. The two most common tree species are *Quercus stellata* and *Q. marilandica*. The herbaceous

layer of the woodlands commonly consists of *Parietaria pensylvanica*, *Oxalis violacea*, *Solidago ulmifolia*, *Andropogon gerardii*, and *Carex gravida*.

Sampling methods

An investigation of DD requires objectively sampled quadrats, otherwise the rate of turnover in species composition may simply reflect the investigators' sampling biases (Palmer 1993). However, in this study I was interested in examining whether my hypothesis was robust as to habitat type, and therefore several criteria were developed for a constrained objective placement of quadrats within the grassland and woodland habitats: transects were to be oriented in a cardinal direction, not run parallel to roads (crossing was considered acceptable), not intersect large bodies of water, be located in the bison management unit, and be located almost entirely within their respective habitats (grassland or woodland) according to aerial photography. With these criteria in mind two sets of UTM coordinates for the origin of each transect were selected from a USGS topographic map. Once the initial coordinates of each transect were set all other points were determined by the sampling grid. The grassland transect was oriented north-south, the woodland transect was oriented east-west, and the transects were separated by approximately nine kilometers. The transects were sampled in different months due to differences in flowering phenology between species of the two habitats. The grassland transect was sampled over a 10 day period in early July 2006. The woodland transect was sampled over a 16 day period from May to early June of 2007. Although the transects were oriented in different directions and sampled in different years, we do not expect these differences to bias our results.

The two transects were both composed of 200, 1 m² square quadrats arranged identically in ten subtransects (Fig. 4.1). Each subtransect was separated by a distance of 117 m and covered an extent of 83 m. Within each subtransect, four quadrats were arranged continuously into five sections. Each section was separated by 17 meters. Within each 1 m² quadrat I recorded the presence of every vascular plant species rooted within the quadrat, and collected three 10 cm soil cores at 0, 0.5, and 1 m perpendicular to the transect. The three cores were aggregated into a single sample and sent to Brookside Labs (New Knoxville, Ohio) to be analyzed for the following cations: Ca, Mg, K, Na, B, Fe, Mn, Cu, Zn, Al, and P. I collected all non-woody above-ground biomass below 1.3 m (i.e. the herbaceous layer) from the corner of each quadrat in an area of 0.01 m² (Fig. 4.1). The wet and dry mass of the biomass were recorded in the lab. In the woodland, a spherical densiometer (Model A, Forest Densiometers, Bartlesville, Oklahoma) reading was recorded in each cardinal direction for each section. I converted the four readings to percent canopy cover and averaged them. I recorded a single field measurement of slope and aspect for each section at a distance of 10 m from the transect.

Analytical methods

Studies of compositional distance decay typically use an analytical framework in which two or more distance matrices are related to one another with the Mantel statistic (see Legendre et al. 2005 for review); however, for my purposes I used the Euclidean community variogram also known as the “variogram of complementarity” (equation 5 of Wagner 2003):

$$\hat{\gamma}(h) = \frac{1}{2n_h} \sum_{a,b|h_{ab} \approx h} \|\mathbf{X}_a - \mathbf{X}_b\|^2 = \sum_i^S \frac{1}{2n_h} \sum_{a,b|h_{ab} \approx h} (x_{ia} - x_{ib})^2 \quad (1)$$

where, $\hat{\gamma}(h)$ is the estimated semivariance at a spatial lag of h , n_h is the number of pairs of samples separated by a distance of h , and \mathbf{X}_a and \mathbf{X}_b are vectors of species abundance or presence/absence in samples a and b respectively that are separated by distance h . Note that decomposing equation 1 into its species specific components reveals that the Euclidean community variogram does not consider pairwise covariances between species (i.e. interspecific associations) but simply the sum of the squared differences in abundance or presence of species i in samples a and b across all S species.

Equation 1 provides both an intuitive metric of species turnover and a clear link to environmental models of species composition that are necessary for addressing my hypothesis. When based upon presence-absence data, the semivariance [$\hat{\gamma}(h)$] of the Euclidean community variogram is the expected number of unique species that will be encountered at a given distance from the focal quadrat (Wagner 2003, 2004). This aspect of the Euclidean community variogram makes it a good metric of turnover for my study because my hypothesis is tied to the relationship between the rate of encountering new environments and the rate of encountering new species as a function of distance.

Additionally, the semivariance can be thought of as a scale-dependent decomposition of the variance-covariance matrix associated with the “linear” ordination techniques: principal components analysis (PCA) and redundancy analysis (RDA). This is beneficial because it provides a direct link between the variance explained by a non-spatial model of the species environment relationship and the patterns of semivariance in the environment and species composition. The scale specific examination of ordination techniques is generally referred to as multi-scale ordination (Noy-Meir and Anderson 1971, Wagner 2004).

I used RDA to test whether variance in species composition was related to the environment (ter Braak and Šmilauer 2002, Palmer et al. 2008b). RDA was the natural choice of explanatory model in my study because the total variance of this analysis can be decomposed by the Euclidean variogram into its variance profile. The only environmental variables in my analysis were standardized, log transformed soil cations (ppm). This decision was made in part because other studies have demonstrated that soil cations were strongly correlated with vascular plant composition at our study site (Chapter V). Additionally soil cations provide a relatively precise measurement of spatial environmental heterogeneity. I standardized the cations because two cations in the grassland displayed a few strong outliers, and standardization was able to decrease the influence of these samples on the PCA. I quantified eleven different soil cations: P, Ca, Mg, K, Na, B, Fe, Mn, Cu, Zn, and Al. Because many of these cations are highly correlated with one another, I summarized the variance in the soil cations with the first four axes of a PCA on the soil cations from each habitat. Four axes were necessary to capture at least 80% of the variance in the cations within both habitats. The four orthogonal PCA axes for each habitat type were then used as explanatory variables in the RDA analyses. Additionally, I performed a backward stepwise selection of soil cations that minimized the AIC value of the RDA analysis (not shown). These two methods resulted in the same qualitative results and therefore only the environment defined by the PCA axes will be discussed further. The amount of variance explained by the PCA soil axes was quantified as the ratio of the sum of constrained variance over the total variance (R_{RDA}^2). The unbiased or adjusted version of this statistic (R_{RDAadj}^2) was also calculated using Ezekiel's formulation (Peres-Neto et al. 2006). It should be noted that because I

am not factoring out the purely spatial component of compositional variance these estimates of variance explained are inflated (Legendre 1993, Legendre et al. 2005).

I tested whether the RDAs explained significantly more variance in the species composition than expected due to chance with constrained permutation tests. The permutation algorithm simultaneously carried out random reflections and rotations of the rows (or sites) of the raw site-by-species matrix (Palmer and van der Maarel 1995, ter Braak and Šmilauer 2002). These permutations were designed to maintain the spatial autocorrelation and interspecific associations between the samples while nullifying their spatial dependence (i.e. relationship between a sample and the environment). The permutation scheme was applied separately at three hierarchical scales that corresponded with the three levels of my sampling design: quadrats, sections, and subtransects (see Fig. 4.2 for an example of the permutation technique applied to subtransects).

The permutation algorithm can be thought of conceptually as first breaking the overall site-by-species matrix along its rows into the subunits of interest (e.g., sections). Next there was a 50% chance that the order of the rows (sites) in a particular submatrix were reversed (the random reflection). Following White and Gilchrist's (2007) analogy, each submatrix was then rolled into a cylinder with the last quadrat in the series adjacent to the first quadrat in the series and rotated a random number of positions. The rotations between submatrices were independent of one another during each permutation. Lastly, each cylinder was broken and the overall matrix was reconstructed for usage with the observed site-by-environment matrix in an RDA. This permutation scheme only preserves a fraction of the original spatial autocorrelation because the cylinder is broken at different positions (White and Gilchrist 2007). The test statistic for this analysis was

the F -value (Legendre and Legendre 1998, p608). When testing a model at the subtransect scale there were only 20 possible spatial orientations (including the empirical ordering), and therefore it was possible to perform a complete permutation test in which every possible permutation was enumerated. In this case the smallest possible p -value was 0.05 ($=1/20$). Complete permutation tests were impractical for all other tests and 999 permutations were conducted instead. To judge significance, I used a Bonferonni corrected alpha value of $\alpha = 0.05/n$ where n is the number of tests performed on a single dataset.

I quantified the rate of change in the environment as a function of spatial lag by first calculating multivariate Euclidean variograms for the variables of interest. A weighted least squares regression model was used to estimate the slope of the relationship between the log of the semi-variance in the environment on the log of spatial lag. The weights of the model were determined by the number of pair-wise comparisons that were available for a given spatial lag. The number of pair-wise comparisons (for both species composition and the environment) were 15, 15, 40, 20, 22 and 30 at spatial lags of 1, 2.33, 19, 21.4, 35.55, and 41 m respectively. In general, ordinary least squares regression models agreed closely with the weighted regression model (not shown).

The same general approach was used to calculate the rate of species turnover as a function of spatial lag. I quantified spatial community turnover with Euclidean variograms (Wagner 2003, 2004). As with the environmental variograms, I log transformed the semivariance of species composition and spatial lag before estimating the slope of the relationship with a weighted linear regression model. Only spatial lags up to

half of the distance of the replicate were considered to reduce the influence of spatial lags with poor replication. With respect to the subtransects this was a distance of 41 m.

I chose to examine the variograms in log-log space because I was interested in estimating the relative rate of DD rather than absolute rate of change. Furthermore, if patterns of DD are relatively linear in log-log space then it suggests that they may be scale invariant which would aid in my ability to predict patterns outside the domain of my sampling scales. Lastly, I expected that much of the change in composition and the environment may take place rapidly at relatively small scales given the small grain of my sampling unit (1 m^2) and in order to better qualitatively and quantitatively characterize this pattern, a log transformation of spatial lag was beneficial.

I tested if a positive relationship between the rate of change in the environment and the rate of change in species turnover existed between the subtransects with constrained but spatially random permutation tests. The spatial positions of each subtransect were randomly swapped, but the order of quadrats within each subtransect were kept fixed in the observed spatial arrangement. The test statistic for this randomization test was the t -value for the slope of the OLS regression model of the rates of species turnover on the rates of environmental DD. Only data on 9 of the 10 subtransects were included in the test of the ETH for each habitat due either to compromised soil samples or to exposed bedrock which prevented soil coring. This missing environmental data precluded the calculation of spatial lags at intermediate scales for the subtransects in which they occurred and therefore these specific subtransects were not used to estimate the rates of DD. The vegetation information from these subtransects (which was complete) was still used in the randomization procedures.

RESULTS

In the grassland average species richness was 18.02 with a range of 7-33 and in the woodland it was 13.62 with a range of 0-29. The woodland habitat had higher gamma diversity with 40 more species than the grassland habitat (211 species in the woodland compared to 171 in the grassland). Total compositional variance was similar in the two habitats (Table 4.1). Total variance in the standardized soil cations was also similar in the two habitats: 9.31 and 9.01 in the woodland and grassland, respectively. Average aboveground dry biomass in the grassland was 25.88 g compared to 5.86 g in the woodland. Average canopy cover in the woodland was 60%, but ranged from 0-90%.

Environmental control of species composition

The species composition in the grassland responded more strongly to variance in soil cations (Table 4.1). Specifically, the RDA explained 12% of the variance in grassland species composition and was either significant or marginally significant at each of the three scales of randomization. In contrast, the RDA in the woodland explained 5% of the variance in species composition and was only significant or marginally significant at the section and subtransect scales, respectively.

Geometry of the environment and species turnover

The log-log variograms for both the environment and species composition were generally well described by linear models in log-log space (Figs. 4.3-4.4). The estimated semi-variance for both the environment and species composition did not consistently deviate from the regression line in either the positive or negative direction for a certain spatial lag which would be indicative of spatial dependence. However, I refrain from describing them as self-similar or scale free because considerable variation exists around

some of the regression lines (given that this is a log-log scale). The average rate of DD (i.e., the weighted OLS regression slope of the log-log variogram) for the grassland was 0.38, which was higher than the average rate of environmental decay observed in the woodland (0.30); however, in both habitats the subtransects displayed a range of environmental decay rates (Figs. 4.3-4.5). The rate of community turnover was positively correlated with the rate of environmental DD in only the grassland habitat, and there was no relationship in the woodland (Fig. 4.5). In both habitats the relationship was rather noisy.

DISCUSSION

I hypothesized that if the environment structures community composition, then the rate of DD in the environment should be positively correlated with the rate of DD in species composition. My results suggest that in both habitats the rate of community DD increased with the rate of environmental DD, however, I observed a significant positive correlation only in the grassland habitat. The lack of significant correlation in the woodlands did not strongly support or strongly contradict my hypothesis. My results suggest that one of the central premises of the environmental texture hypothesis (ETH) may be relevant at local spatial scales.

Implications for the ETH and future tests

My study was stimulated in large part by the ETH, which hypothesizes that the triphasic pattern of the species-area relationship (SAR) is due to changes in the geometry of the environment as a function of spatial grain (Palmer 2007, Qian et al. 2007). An underlying premise of the ETH is that the rate at which unique species are accumulated (z

of the species-area relationship) is determined by the rate at which new environments are sampled. Many conceptual models and empirical studies include a metric of environmental or habitat diversity when attempting to predict species richness (Conner and McCoy 1979, Storch et al. 2003, Triantis et al. 2003); however, the hypothesis I addressed suggests that it is not simply the total variability in the environment but the spatial (or temporal) structure of that variability which is relevant to understanding the rate of species accumulation or turnover. If the environmental variability is strongly spatially structured it will have a steep Euclidean log-log variogram, which is characteristic of an environmental gradient (Palmer 1988), and new species should accumulate rapidly.

Although the ETH was first suggested as an explanation of the triphasic SAR that is only observed across many orders of magnitude in area, my study suggests that the underlying assumption that the geometry of the environment influences the rate of species turnover may be also relevant for vascular plants at local scales (1-50 m²). If my sampling design captured a larger range of spatial scales, the ETH would predict that the log-log variograms for both the environment and species composition would display a scale dependent deceleration in the relative rate of accumulation of new species and new environments as a function of spatial scale. A potential test of the ETH is to examine if this is indeed the case and if the scale at which the environment changes geometry corresponds with a change in the rate of species turnover. However, a strong obstacle in performing such a test would be that the same environmental variable that is relevant at local scales is not necessarily the most important at regional scales.

Euclidean and chi-squared variograms as metrics of species turnover

I used Euclidean variograms to estimate the rate of compositional turnover and environmental change between two sampling units. I chose this metric because when based upon presence/absence data the semivariance is the expected number of unique species between two sampling units. Euclidean variograms also provide a spatial decomposition of the variance explained by RDA. Additionally, the mathematical properties of variograms are generally well understood due to their rich history in the field of geostatistics, unlike newer methods of studying species turnover such as dissimograms and correlograms (Journel and Huijbregts 1978, Wagner 2003). Over the scales I examined both the environmental and compositional variograms generally appeared linear in log-log space. This justifies in part my usage of linear regression to estimate the slopes of these relationships, but I hesitate to refer to them as self-similar. Palmer (1988) demonstrated that considering variation in the slope of the log-log variogram (via fractograms) can yield additional insight into the geometry of species composition even when the relationship appears approximately linear.

Wagner (2004) noted that although the Euclidean community variogram has many useful properties (some of which I have mentioned here), it may be best suited for describing species turnover over relatively short environmental gradients in which species are expected to display linear responses to the environment. If species display unimodal responses to an environmental gradient, then it is typically argued that weighted averaging ordination techniques such as correspondence analysis (CA) or canonical correspondence analysis (CCA) offer a superior representation of community variation (Gauch 1982, ter Braak 1986, ter Braak and Prentice 2004). Therefore Wagner

(2004) developed the chi-square community variogram, which is a spatial decomposition of the chi-squared variance-covariance matrix used by CA and CCA. I refrained from the interpretation of the chi-squared variogram in my study because the semivariance of this method does not have an intuitive link to the expected number of unique species between two samples. In a study of several different datasets, Schlup and Wagner (2008) found that the Euclidean and chi-squared variograms generally agreed with one another. However, in their study increasing quadrat grain consistently increased the semivariance of the Euclidean variogram but not the chi-squared variogram. More case studies are required to better understand the differences and appropriateness these two community variograms.

Habitat differences

The geometry of the soil environment appeared to exert a stronger influence on the rate of species turnover in the grassland than in the woodland. This was expected, in part, because the RDAs indicated that in the woodland species composition was less related to the soil environment than in the grassland. Given how little of the total variation the woodland RDA explained (5%), it seems likely that other environmental variables may be shaping community spatial patterns that were not included in the analysis. In the woodland, heterogeneity in degree of light limitation may have influenced understory plant composition given the range of overstory canopy cover I observed in my study (0-94%, standard deviation c. 25%). However, additional RDA analyses (not shown) indicated that average canopy cover explained only a small additional portion of the total variability (less than 0.1%). The woodlands were also more topographically complex than the grasslands and crossed several small gullies or

arroyos, which may have decreased the relative importance of measures of the soil environment. But here again additional analyses utilizing slope and aspect (not shown) indicated that these variables also only contributed a small fraction of explained variance. Moisture and nutrient availability are also known to shape herbaceous understory community composition (Ludwig et al. 2004, Galhidy et al. 2006, Graves et al. 2006), but I did not collect information on either of these variables.

An alternative explanation for the weaker environmental signature in the woodland, is that historical drivers are playing a stronger role in this habitat. Both habitats were in the bison management unit but they differed in frequency of prescribed burning. The woodland sites were located in a management unit that experienced a more frequent burning regime (seven spring burns in the past seven years) than the management unit in which grassland site was located (two burns in the past seven years). Because of the greater opportunity for fire exclusion in the forest (due to rockiness and bare patches without fuel) it is unknown if the actual frequency of fire at the woodland sites was lower than what was reported by The Nature Conservancy. It seems reasonable that a high fire frequency in the woodlands may have contributed to decreased response in the understory vegetation to soil cations. However, a study conducted on the understory of a mixed-oak forest found that frequency of fire explained relatively little total variation in species composition (Hutchinson et al. 2005). Reilly et al. (2006) found that a single wildfire increased (not decreased) the correlation between species composition and the environment. However, this finding was primarily due to a strong gradient in elevation which influenced the severity of the fire.

A third possible explanation is that because the vegetation was sparser in the woodlands, the rarefaction or sampling effect may be exerting a stronger influence in this habitat. The rarefaction effect is due to the necessary link between the number of individuals sampled and the number of species observed (Palmer et al. 2008a, McGlinn and Palmer 2009). The rarefaction effect will exert the strongest influence on the probability of observing a particular species at fine grains (e.g., 1 m² quadrats). Even if species are responding strongly to the environment our ability to detect this will be inhibited if the density of individuals is relatively low. If the number of individuals of each species is recorded in each sample, then the expectation of species turnover due only to rarefaction effects can be developed. A simple randomization approach would be to place all individuals into a pool of potential colonists and randomly select the observed number of individuals without replacement for each quadrat. However, this may not be an option for many vegetation studies (including the present study) because many plants exhibit clonal growth which makes estimating the number of unique individuals (or genets) logistically unfeasible. An alternative solution is to sample at multiple grains such that the ecological pattern may be observed at presumably different levels of the rarefaction effect.

Conclusions

The rate of species turnover varies across landscapes. If species composition is related to measured environmental variables (as in the grassland habitat), then the rate of change in the environment will be positively correlated with the rate of species turnover. Therefore, the texture of the environment can help to explain variation in the rate of species turnover across spatial scales in objectively placed samples. Future work on

linking the environment to patterns of turnover should consider the strengths of the Euclidean variogram, which is appropriate for non-contiguous samples, provides an intuitive definition of turnover, and is linked to environmental models of species composition. The chi-squared variogram may be a useful metric of turnover as well, but more work is necessary to understand its properties. At local scales, the rarefaction effect may decrease the ability to detect environmental signatures and should be considered as a potential confounding factor in comparative studies.

ACKNOWLEDGEMENTS

DJM received funding from the U.S. Environmental Protection Agency (EPA) under the Greater Research Opportunities (GRO) Graduate Program. The U.S. EPA has not officially endorsed this publication, and the views expressed herein may not reflect the views of the agency. Additionally, I thank The Nature Conservancy for logistical support and members of LIBRA for assistance in the field particularly P. Earls, W. Lowry, and C. Huang.

LITERATURE CITED

- Buckley, L. B. and W. Jetz. 2008. Linking global turnover of species and environments. *Proceedings of the National Academy of Sciences* 105:17836-17841.
- Conner, E. F. and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. *American Naturalist* 113:791-833.
- Galhidy, L., B. Mihok, A. Hagyo, K. Rajkai, and T. Standovar. 2006. Effects of gap size and associated changes in light and soil moisture on the understorey vegetation of a Hungarian beech forest. *Plant Ecology* 183:133-145.
- Gauch, H. G., Jr. 1982. *Multivariate Analysis in Community Ecology*. Cambridge University Press, New York.
- Graves, J. H., R. K. Peet, and P. S. White. 2006. The influence of carbon-nutrient balance on herb and woody plant abundance in temperate forest understories. *Journal of Vegetation Science* 17:217-226.
- Hubbell, S. P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, New Jersey, USA.
- Hutchinson, T. F., R. E. J. Boerner, S. Sutherland, E. K. Sutherland, M. Ortt, and L. R. Iverson. 2005. Prescribed fire effects on the herbaceous layer of mixed-oak forests. *Canadian Journal of Forest Research* 35:877-890.
- Journel, A. G. and C. Huijbregts. 1978. *Mining geostatistics*. Academic Press, London, UK.
- Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology* 72:367-382.

- Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74:1659-1673.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecological Monographs* 75:435-450.
- Legendre, P. and L. Legendre. 1998. *Numerical Ecology*. Elsevier, Boston.
- Ludwig, F., H. de Kroon, F. Berendse, and H. H. T. Prins. 2004. The influence of savanna trees on nutrient, water and light availability and the understorey vegetation. *Plant Ecology* 170:93-105.
- McGlenn, D. J. and M. W. Palmer. 2009. Modeling the sampling effect in the species-time-area relationship. *Ecology* 90:836-846.
- Nekola, J. C. and P. S. White. 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* 26:867-878.
- Nekola, J. C. and P. S. White. 2002. Conservation, the two pillars of ecological explanation, and the paradigm of distance. *Natural Areas Journal* 22:305-310.
- Noy-Meir, I. and D. Anderson. 1971. Multiple pattern analysis or multiscale ordination: towards a vegetation hologram. Pages 207–232 *in* *Statistical ecology: populations, ecosystems, and systems analysis*, Eds. E. C. P. G. P. Patil, and E. W. Water. Pennsylvania State University Press, University Park, Pennsylvania, USA.
- Palmer, M. W. 1988. Fractal geometry: a tool for describing spatial patterns of plant communities. *Vegetatio* 74:91-102.
- Palmer, M. W. 1993. Potential biases in site and species selection for ecological monitoring. *Environmental Monitoring and Assessment* 26:277-282.

- Palmer, M. W. 2007. Species-area curves and the geometry of nature. Pages 15-31 *in* Scaling Biodiversity, Eds. D. Storch, P. L. Marquet, and J. H. Brown. Cambridge University Press, Cambridge, UK.
- Palmer, M. W., D. J. McGlinn, and J. F. Fridley. 2008a. Artifacts and artificions in biodiversity research. *Folia Geobotanica* 43:245-257.
- Palmer, M. W., D. J. McGlinn, L. Westerberg, and P. Milberg. 2008b. Indices for detecting differences in species composition: some simplifications of RDA and CCA. *Ecology* 89:1769-1771.
- Palmer, M. W. and E. van der Maarel. 1995. Variance in species richness, species association, and niche limitation. *Oikos* 73:203-213.
- Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology* 87:2614-2625.
- Qian, H., J. D. Fridley, and M. W. Palmer. 2007. The latitudinal gradient of species-area relationships for vascular plants of North America. *American Naturalist* 170:690-701.
- Reilly, M. J., M. C. Wimberly, and C. L. Newell. 2006. Wildfire effects on beta-diversity and species turnover in a forested landscape. *Journal of Vegetation Science* 17:447-454.
- Schlup, B. M. and H. H. Wagner. 2008. Effects of study design and analysis on the spatial community structure detected by multiscale ordination. *Journal of Vegetation Science* 19:621-632.

- Storch, D., A. L. Šizling, and K. J. Gaston. 2003. Geometry of the species-area relationship in central European birds: testing the mechanism. *Journal of Animal Ecology* 72:509-519.
- ter Braak, C. J. F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67:1167-1179.
- ter Braak, C. J. F. and I. C. Prentice. 2004. A theory of gradient analysis. *Advances in Ecological Research* 34:235-282.
- ter Braak, C. J. F. and P. Šmilauer. 2002. *Canoco for Windows Version 4.5*. Biometris - Plant Research International, Wageningen, The Netherlands.
- Triantis, K. A., M. Mylonas, K. Lika, and K. Vardinoyannis. 2003. A model for the species-area-habitat relationship. *Journal of Biogeography* 30:19-27.
- Wagner, H. H. 2003. Spatial covariance in plant communities: Integrating ordination, geostatistics, and variance testing. *Ecology* 84:1045-1057.
- Wagner, H. H. 2004. Direct multi-scale ordination with canonical correspondence analysis. *Ecology* 85:342-351.
- White, E. P. and M. A. Gilchrist. 2007. Effects of population-level aggregation, autocorrelation, and interspecific association on the species-time relationship in two desert communities. *Evolutionary Ecology Research* 9:1329-1347.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279-338.
- Wilson, M. V. and C. L. Mohler. 1983. Measuring compositional change along gradients. *Vegetatio* 54:129-141.

TABLES

Table 4.1. The results of the RDA and permutation tests on the sum of all canonical eigenvalues for the two habitat types. The first four axes of a PCA on the soil cations from a particular habitat were used as explanatory variables in each model respectively. The permutations were spatially constrained at three scales: quadrats, sections, and subtransects (see Methods – *Analytical methods* for details). All tests were conducted with 999 permutations except for those at the subtransect scale in which a complete permutation test on all 20 possible constrained orderings of the data was considered. The Bonferroni corrected alpha value for each habitat is 0.017 (= 0.05/3).

Habitat	Total variance	R_{RDA}^2	R_{RDAadj}^2	F	p -values		
					Quadrat	Section	Subtransect*
Grassland	11.08	0.14	0.12	7.59	0.001	0.001	0.050
Woodland	10.25	0.07	0.05	3.65	0.794	0.005	0.050

*the smallest possible p -value is 0.05 (=1/20)

FIGURES

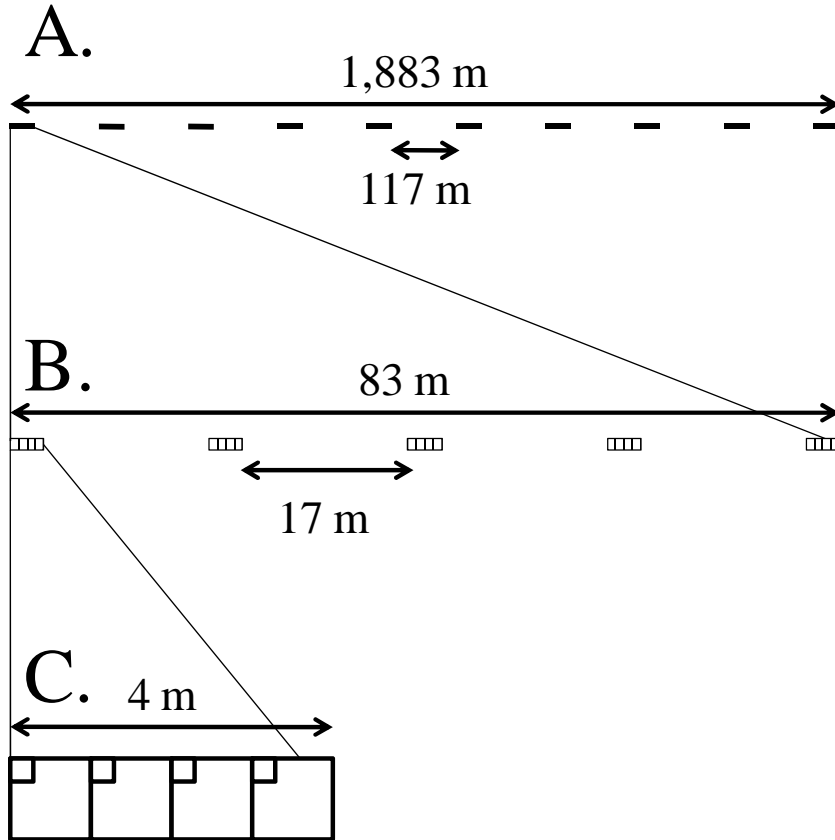


Fig. 4.1. Diagram of a transect (A.), a subtransect (B.), and a section (C.). Each transect covered a total extent of 1,883 m and was composed of ten subtransects which each covered an extent of 83 m. The gaps in the transects indicate areas that were not sampled, and the unfilled squares indicate where the vegetation and the soil environment was sampled. Each quadrat occupied an area of 1 m^2 , and above ground biomass was sampled in the 0.01 m^2 corner subquadrats.

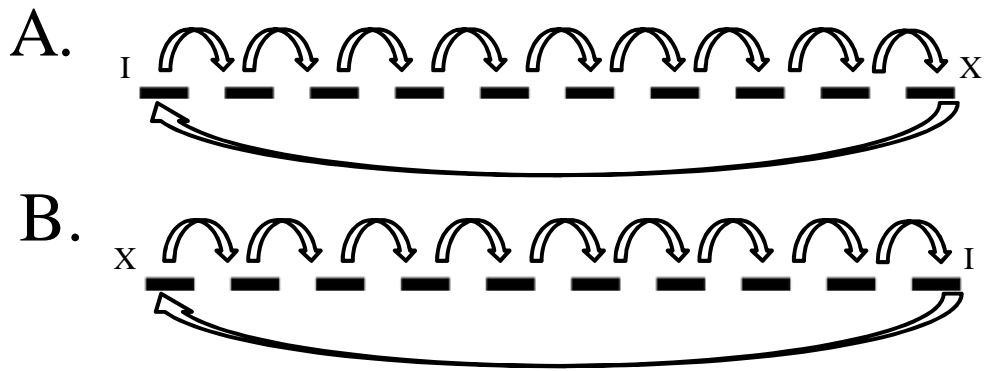


Fig. 4.2. Diagram of the random rotation and random reflection permutation scheme. A single random shift without a reflection (A.) and a single random shift with a reflection (B.) applied to the ten subtransects. The Roman numerals indicate the orientation of the transect. Although the position of each subtransect was shifted, the spatial order of the samples within each subtransect was as observed in the empirical sample with the exception of the subtransect that must be returned to the beginning of the series.

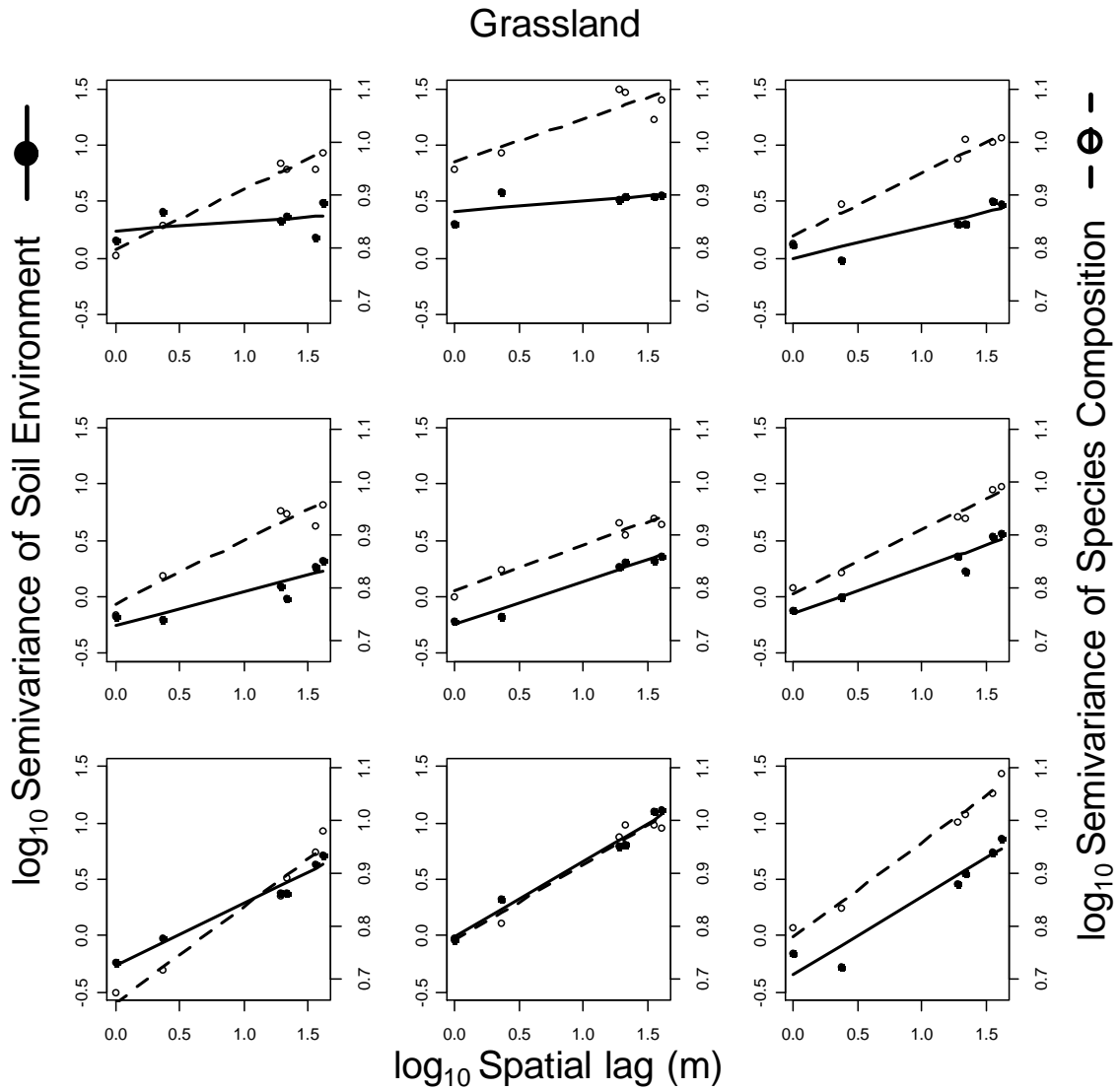


Fig. 4.3. The environmental and community Euclidean variograms for the nine subtransects in the grassland habitat. The figures are ordered left to right and top to bottom in increasing rate of environmental distance decay. The axis on the left applies to the degree of difference in the selected environmental variables and the axis on the right applies to the degree of species turnover in species composition. All axes are \log_{10} transformed and the fitted lines are weighted linear regression models.

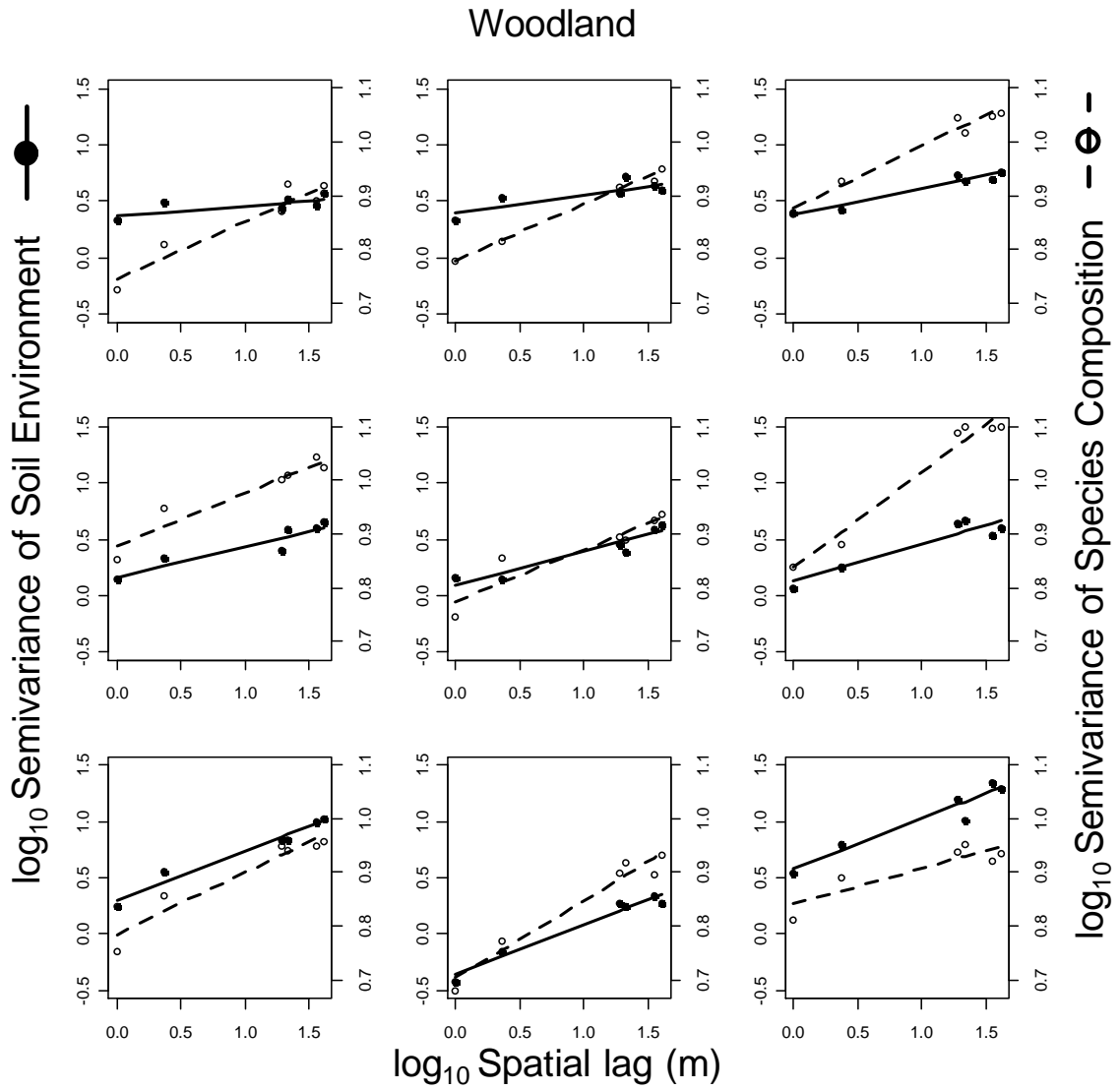


Fig. 4.4. The environmental and community Euclidean variograms for the nine subtransects in the woodland habitat. The figures are ordered left to right and top to bottom in increasing rate of environmental distance decay. See Fig. 4.3 for additional explanation of the graph.

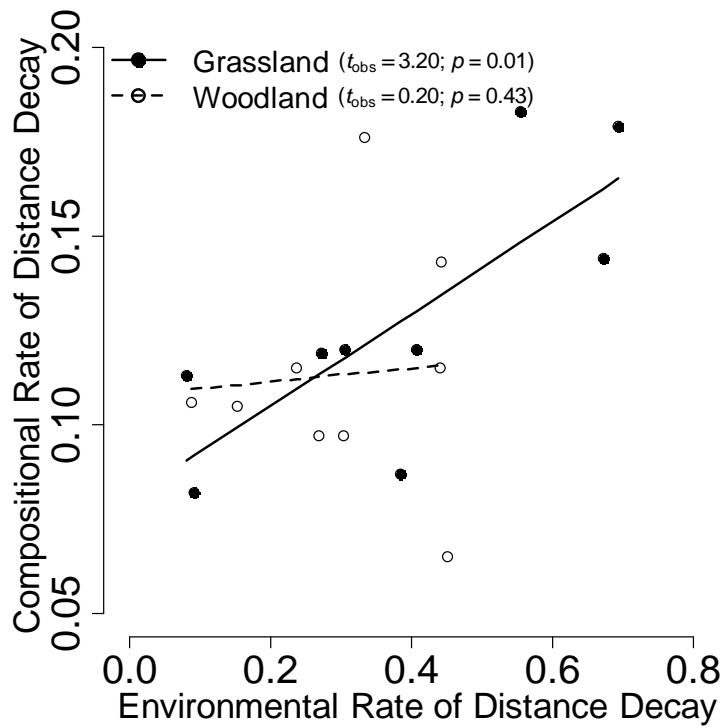


Fig. 4.5. Relationship between the rates of compositional and environmental distance decay. The solid regression line is for the grassland subtransects (solid circles), and the dotted regression line is for the woodland subtransects (open circles). The p -values are based on 999 constrained permutations (see Methods - *Analytical methods* for more details).

CHAPTER V

THE INFLUENCE OF MANAGEMENT RELATIVE TO INHERENT LANDSCAPE HETEROGENEITY ON THE VEGETATION OF A TALLGRASS PRAIRIE

Abstract. It is widely recognized that management using prescribed fire and grazing influences plant species richness and composition in many ecological communities. However, the contextual relevance of management is often unclear because the majority of our knowledge is based upon small-scale manipulative studies that are not ideally suited to quantify the importance of management relative to inherent landscape heterogeneity. The purpose of my study was to quantify the importance of fire and grazing by bison and/or cattle relative to inherent site and year effects on the vascular plant community in a tallgrass prairie. I accomplished this objective with an 11 year observational study on a preserve where management decisions are geared towards increasing community heterogeneity through the application of randomized burning and freely ranging grazers. I used variation partitioning and explanatory modeling within multiple regression and canonical ordination frameworks on species richness and composition respectively. My results indicated that site effects, due to belowground differences, explained the majority of variation in richness and composition. Year effects, related to seasonal precipitation, were more strongly correlated with variation in

richness than composition. Management practices were relatively unimportant relative to inherent site and year drivers with respect to both richness and management; however, management effects were significant and interpretable, and bison management was positively correlated with plant richness. The strength of inherent landscape heterogeneity on the plant community suggests that fine-tuning management is not critical for maintaining this community as long as woody plant encroachment is kept in check. Furthermore, our study demonstrates how observational studies can be used to place management effects into a broader ecological context.

Keywords: bison, Flint Hills, grassland, natural variability concept, Oklahoma, restoration, vascular plants, and vegetation monitoring

INTRODUCTION

Natural variability concepts of land management, which promote spatial and temporal variability, are increasingly used in restoration ecology (Palmer et al. 1997). Underlying these concepts are two premises: 1) historical conditions and processes can provide guidance for management, and 2) spatial and temporal variability generated by disturbance are vital components of nearly all ecosystems (Landres et al. 1999). Managing for historical conditions is thought to benefit species that have evolved in that system and to minimize human alterations (Swanson et al. 1994). Spatial and temporal variability in management is thought to maintain biological diversity (MacArthur 1965, Petraitis et al. 1989). Although these concepts are grounded in ecological theory, as

Palmer et al. (1997) note, the importance of managing for natural variability is rarely examined experimentally.

The application of natural variability concepts is relevant to the conservation and restoration of the North American tallgrass prairie ecosystem. In this ecosystem, fire and grazing were important components of the pre-Columbian North America disturbance regime (Anderson 1990) and still are today (Daubenmire 1968, Abrams et al. 1986, Collins 1992). Although historically the region was grazed by bison (*Bos bison* L.) and was burned in a variety of seasons, presently much of the remaining tallgrass prairie ecosystem is managed for cattle with annual spring burns (Fuhlendorf and Engle 2001). The purpose of annual spring burning is to maximize yields of palatable C4-grasses (Towne and Owensby 1984). Howe (1994) also noted that many restoration efforts on prairie remnants promote production of C4-grasses with spring season burning and a lack of grazing. These management practices may be a threat to the ecosystem's biodiversity by only benefiting one competitively superior plant-functional group and by homogenizing an entire region's disturbance regime (Howe 1994, Fuhlendorf and Engle 2001, Nekola and White 2002).

In an effort to restore natural variability to grazing systems, Fuhlendorf and Engle (2001, 2004) suggested the interaction between fire and grazing could be used in a more variable manner in space and time to create a *shifting mosaic* in contrast to the traditional homogenous application of these management tools. They argued that a mosaic of burned and unburned patches more closely approximates the historical variability that would have existed on the landscape and will result in higher biodiversity than traditional homogenous management practices (annual spring burning). This hypothesis is guiding

the management of the Tallgrass Prairie Preserve (TGPP), one of the largest tallgrass prairie preserves (Hamilton 1996, 2007).

The scale of the TGPP (15,700 ha) coupled with the spatially and temporally varying application of fire provides both important opportunities and challenges for experimentally evaluating aspects of the natural variability hypothesis. One of the opportunities that the TGPP offers is a chance to carry out an observational study that investigates the relative importance of management in a tallgrass prairie ecosystem that is presumably in somewhat of a more pre-Columbian ('natural') state given its intact native vegetation, the presence of free ranging bison, and the variable burning regime. This opportunity is valuable because the majority of our knowledge on tallgrass prairie ecology either originates from relatively small scale experiments (e.g., Hulbert 1988) or from strictly controlled watershed treatments (e.g., Konza prairie LTER studies). Controlled studies are extremely valuable in elucidating ecological mechanisms that underlie a system, but they are less useful in providing the ability to examine multiple driving factors of community change within a broader ecological context. This can be better achieved by an observational study if properly designed (Hobbs et al. 2007, Weiher 2007) Examining more than two or three treatments at several levels requires more replication than is typically feasible in long-term ecological research, however, the effect of several factors can be statistically separated in an observational study with relatively few replicates.

The purpose of this study was to investigate the variation in plant species richness and composition in space and time at the TGPP, and to quantify the relative importance of management, which included prescribed burning and grazing by bison and/or cattle.

My goal was not to directly test whether or not variable management regimes are effective tools for meeting conservation targets, but rather to examine the relative influence that management has on the plant community when management is guided by the natural variability hypothesis. Differences due to management within sites are likely to be strongly confounded by year-to-year variation in climate (Anderson 1982, Gibson and Hulbert 1987, Adler and Levine 2007) and between sites by variation in soil (Critchley et al. 2002) and topography (Abrams and Hulbert 1987, Briggs and Knapp 1995). However, confounding differences between sites and years can be statistically controlled for, and the independent and shared effects of temporal variation in management relative to site and year effects can be estimated simultaneously.

Using a mixture of explanatory modeling and ordination techniques, I addressed three related questions. Is there directional change in species richness and species composition through time? What are the most important environmental variables for explaining plant richness and composition? If site and year effects are controlled for, does temporal variation in management significantly correlate with richness or composition and, if so, in what way?

METHODS

Study Site

The TGPP is a 15,700 ha nature preserve located between 36.73° and 36.90° N latitude, and 96.32° and 96.49° W longitude, in Osage County, Oklahoma and owned by The Nature Conservancy (TNC). Over the course of the 11 year study period (1998-2008), total annual rainfall varied from 490 to 1250 mm. The preserve is situated at the

southern extent of the Flint Hills region. The elevation of the preserve ranges from 253 to 366 m, and the underlying bedrock of the region is characterized by soils deriving from Permian sediment (Oviatt 1998). Due to long-term erosion, the surface layers of soil are thin and young; limestone and sandstone are frequently exposed at the surface, sometimes within close proximity of each other. Because of this rockiness the Flint Hills region, including the Tallgrass Prairie Preserve, has remained unplowed and has been instead utilized primarily as rangeland for cattle. Prior to the acquisition of the preserve by TNC in 1989, the majority of the site was managed for cow-calf and yearling cattle production with a 4- to 5-year rotation of prescribed burning and aerial application of broadleaf herbicides (1950-1989) (Hamilton 2007).

Approximately 90 % of the TGPP consists of grasslands. The majority of the grasslands are composed of tallgrass prairie habitats dominated by *Andropogon gerardii*, *Sorghastrum nutans*, *Sporobolus compositus*, *Panicum virgatum*, and *Schizachyrium scoparium*. Shortgrass prairie habitat occurs to a lesser extent on more xeric sites and is dominated by *Bouteloua* spp. Despite the application of herbicide earlier in the 20th century, the flora of the preserve appears relatively intact with a total of 763 species of vascular plants (to date) of which 12.1% are exotic (Palmer 2007).

Management

The management at the TGPP was variable in space and time. In 1993, 300 bison were introduced year-round onto a 1,960 ha portion of the preserve (Hamilton 1996, 2007). As the bison herd increased in size, the area allotted to the herd was increased eight times to an area of 8,517 ha by 2007 (Fig. 5.1, 54% of preserve area). Initial bison stocking rates were increased in 1999 to 2.1 animal-unit months ha⁻¹ (see Hamilton 2007

for additional details). Within the bison unit, animals were allowed to range freely and their movement was not obstructed by internal fences. Watersheds within the bison unit were considered randomly for burning only if they met the minimum fuel criteria of 900 kg ha⁻¹ of fine fuels. Within a given year, the season of burn of the bison unit was split as follows: 40 % dormant spring (March - April), 20 % late growing season (August - September), and 40 % dormant winter (October - December). The remainder of the preserve was seasonally grazed by cattle and typically burned more frequently in the dormant spring season, but some of the cattle pastures were utilized for smaller scale (2,350 ha) patch-burn experiments in which only one-third of a given management unit was burned annually (Hamilton 2007). Stocking within the cattle pastures included both intensive-early stocking and season-long stocking, which contrasted with the year-round stocking in the bison unit.

Data collection

Because of the temporally variable and spatially aggregated nature of the management, I opted to annually re-sample a semi-random, spatially stratified set of twenty square 100 m² plots located at the intersections of the 1-km UTM grid (Fig. 5.1.). The only criteria that I imposed on the selection of my plots were that they had not standing water, and less than 20 % cover of woody plants or exposed rock. A semi-unbiased spatial stratification of samples is an ideal sampling method for ensuring that samples are representative of the broader variation in grassland communities across the preserve (Palmer 1995). I opted to annually resample my sites to observe the vegetation at a given site in multiple management states (e.g., one year post burn, bison vs. cattle grazed), and to provide information on inherent year-to-year variation. I resampled my

plots every June from 1998 to 2008 and recorded the percent cover of all vascular plant species at the 100 m² scale.

Each year I combined four 15 cm soil cores collected at each corner of the quadrat and sent these to Brookside Labs (New Knoxville, Ohio) to be analyzed for soil cations, pH, and other variables. I recorded topographic data on slope and aspect in the field. Climatic variation was quantified with total monthly precipitation data that was downloaded from the Oklahoma Mesonet Foraker site (36.841° N, -96.428° W; elevation: 330 m; Fig. 5.1), which is located on the preserve (McPherson et al. 2007). I calculated total precipitation for three arbitrarily defined season variables (four months each): June through September (summer), October through January (winter), and February through May (spring).

I derived management variables, including years of bison grazing, years since last burn, and number of burns in the past five years, from a GIS database that I developed based on TNC's recorded burn and grazing history. I recognize that the number of burns in the past five years is negatively correlated with years since last burn (see Appendix D) and reflects somewhat of an arbitrary cutoff point (five years), but I chose to include this variable in my models despite these shortcomings because I wished to quantify the short term burn history of a site. If a site is burned frequently, then years since burn does not convey any information on the history of the site beyond the last burn. I chose not to include season of burn as an explanatory variable because 83% (67 out of 80) of the prescribed fire events recorded on my study sites took place during the dormant season.

The management variables were dynamic in time. Specifically, the fire return interval was not fixed in any of the management units we sampled; additionally, seven of our samples changed from cattle to bison managed over the course of the study.

I grouped species into five functional groups to aid in the ecological interpretation of patterns of species richness and composition we observed. The functional groups were forbs, legumes, C3 grasses, C4 grasses, and shrubs.

Data analysis

My goal when modeling richness and species composition was to develop explanatory models and not predictive models, *sensu* Mac Nally (2000). Therefore, I was primarily concerned with comparing the explanatory strength of variables that were chosen *a priori* rather than developing a single most accurate or necessarily most parsimonious model. I coded site and years as dummy variables to quantify site and year effects relative to management variables. I then performed a post-hoc examination of the specific environmental variables (described in *Data collection*) that we believed may explain the site and year effects.

Given the observational nature of this study, many of my environmental variables (e.g., soil cations, total spring rain) were strongly collinear and likely act as proxy variables. Therefore, to increase the clarity of my results, I only examined a small number of environmental variables. I selected calcium as the soil variable to examine because previous published analyses that used portions of my dataset indicated that this variable was strongly correlated with richness and acted as a proxy limestone (rather than sandstone derived soils) (Palmer et al. 2003, Brokaw 2004). Aspect was converted to an index of northness [northness = $\cos(\text{aspect})$] (Roberts 1986). I did not consider a

corresponding index of eastness as exploratory results indicated it was unimportant (not shown). We examined seasonal totals of rainfall because exploratory analyses indicated it greatly out performed total rainfall (not shown).

I used ordinary least squares (OLS) and generalized least squares (GLS) to dissect relationships between species richness and my various explanatory variables. GLS models were considered because they allow us to incorporate correlational models for the residuals of the model. When examining specific explanatory models, I compared 10 isotropic correlational models (5 one parameter and 5 two parameter) for the residuals of each explanatory model (see Appendix E for description of models). The model with the lowest AIC value was then chosen to estimate effect sizes and carry out conditional F -tests for each explanatory variable of interest. If a model with one less parameter had an AIC within 3 of the minimum, then visual examination of model fit was used to judge which was more the more appropriate model. The GLS models were fit with a restricted maximum likelihood algorithm, which is the preferred method for generating unbiased estimates of variance in models that have a relatively large number of parameters (Diggle et al. 1994). The R package *nlme* version 3.1-90 was used to carry out all GLS model fitting and diagnostics (Pineiro et al. 2008).

I used variation partitioning (or commonality analysis) to estimate the unique and shared fractions of explained variation in three factors (or classes of variables), sites, years, and management, on richness (Legendre and Legendre 1998, Peres-Neto et al. 2006). Sites and years were coded as dummy variables and the management factor was composed of the three management variables described above. The independent component of variation attributable to management should be interpreted as variation

within-sites which was independent of specific years and was associated with temporal changes in the management variables that occurred over the study period. For example, over the course of the study, seven of the sites changed from cattle grazed to bison grazed. Therefore, variation in these sites before and after the change in grazer that was independent of specific year effects was attributed to differences between bison and cattle. In contrast, if a site remained in the bison or cattle unit for the duration of the study then the grazer variable would attribute no explained variance to temporal changes at that particular site.

I carried out partitioning with OLS as well as with GLS multiple regression. For each fraction of the OLS analysis, I report both the coefficient of determination and its adjustment for number of variables, R^2 and R_{adj}^2 respectively. For the GLS fractions I calculated a generalized formulation of the coefficient of determination, R_{GLS}^2 , that is appropriate for GLS models (Nagelkerke 1991). For the variation partitioning the correlation structure of all models was assumed to be first-order autoregressive so that comparisons between fractions could be more easily interpreted.

In all analyses I treated species richness as a continuous variable with normally distributed error. I recognized that in some modeling contexts it is more appropriate to consider richness as a Poisson distributed variable (Candy 1997, Palmer and Hussain 1997). However, because the normal distribution provides a good approximation of a Poisson random variable when the mean is larger than approximately 20; I feel that my usage of normal errors (instead of Poisson) is justified given that richness was never below 48 in my samples. I also recognize that some authors consider richness to only be one facet of grassland biodiversity and that other diversity indices can yield more insight

into the ecology of the system (Wilsey et al. 2005, Chalcraft et al. 2009). However, in this study richness was so strongly correlated with all of the indices examined in Wilsey et al. (2005) that additional analyses on various diversity indices would have yielded little additional insight (see Appendix F).

I used both indirect and direct gradient analyses to examine species composition. Prior to all ordinations, I square-root transformed all species cover data to decrease the influence of the most abundant species on the analysis, and I down-weighted rare species. Detrended Correspondence Analysis (DCA; Hill and Gauch 1980) was used to visually inspect directional change in the plant composition through time and to estimate the amount of variation composition displayed across a hypothetical environmental gradient. Non-metric multidimensional scaling (NMS), another indirect ordination technique, was used to confirm the general conclusions of the DCA.

For the analysis of species composition I also carried out variation partitioning but with partial canonical correspondence analysis (pCCA) (Borcard et al. 1992, Økland 1999). The factors used in the variation partitioning were as defined above for richness. I calculated Peres-Neto et al.'s (2006) formulation of the adjusted fraction of variation explained in CCA using a permutation approach (R_{CCAadj}^2). To my knowledge, correlational models for the residuals have not been theoretically developed within the direct ordination context and therefore these methods were not applied in my analysis of species composition. However, when testing the importance of dependent variables that accounted for the within-site variability in species composition, I conducted a toroidal shift Monte Carlo test (Legendre 1993, ter Braak and Šmilauer 2002b). This method of permutation maintains the observed order of the samples within a site while nullifying the

temporal relationship of samples between sites. When the constraining variable was site specific, I carried out permutations that randomly shuffled samples only within their year of occurrence which agrees with my assumption of spatial independence between sites but not years. All randomization tests were conducted with 999 iterations to determine significance of all canonical axes for all partial CCA analyses under the reduced model (Legendre and Legendre 1998, p308). As in the analysis on richness, I constructed models to examine the importance of specific explanatory variables with estimates of variation explained and conditional F -tests. I performed multivariate analyses with CANOCO version 4.5 (ter Braak and Šmilauer 2002a) and the R package *vegan* version 1.15-1 (Oksanen et al. 2008, R Development Core Team 2008).

RESULTS

Variance in species richness

The average species richness over the 11 year study was 76.24 (± 0.86). The between-year variance ($\sigma = 9.13$) in richness was approximately equal to the between-site variance ($\sigma = 8.93$). There were not strong directional changes in total richness (not shown) or in the five functional groups through time (Fig. 5.2). Forbs were more species-rich than all other functional groups. Variation partitioning of the unbiased OLS estimate of variance explained (R_{adj}^2 ; Table 5.1), indicated that between-site differences accounted for the majority of variability in richness ($R_{\text{adj}}^2 = 0.48$), followed by year-to-year differences ($R_{\text{adj}}^2 = 0.13$). However, after the within-site temporal autocorrelation was considered (using a first-order auto-regressive term), the strength of site identity relative to year identity was diminished quite a bit according to Nagelkerke's (1991) generalized

definition of the coefficient of determination. In both OLS and GLS based variation partitioning, the management class of variables accounted for a much smaller percentage of explained variance in richness (2-4 %), but still contributed a significant amount of explained variation as judged by conditional (type III) *F*-tests at an alpha level of 0.05 (Table 5.2). The shared component of variation between management and site was larger than the influence of management alone, but the same was not true when comparing year and management after correcting for temporal autocorrelation in which case the shared component was effectively zero.

The most important variable for site-to-site differences was log Ca, which was negatively correlated with richness (standardized coefficient, $\beta = -0.28$) (Table 5.2). Topography did not seem to have a strong influence on richness as neither slope nor northness explained much variation in richness.

Temporal variation in richness within the sites was attributable to both climate and management variables. The most important seasonal rainfall variable was summer rain, which was negatively correlated with richness ($\beta = -0.19$); winter rain and spring rain appeared equally important and both were positively associated with richness ($\beta = 0.15$ and 0.14 , respectively). The single most important management variable was years of bison, which was positively associated with richness ($\beta = 0.43$). Both years since burn and number of burns in the past five years were negatively associated with richness but neither variable was found to explain significantly more variation in richness than due to chance.

Variance in species composition

The DCA scatterplot indicated that change in species composition through time was not directional (Fig. 5.3). Furthermore, sites remained relatively separate from one another in ordination space, which indicates that variation in species composition was greater between sites than within sites. There was not a strong difference in sample scores between sites that were bison- versus cattle-grazed along the first or second DCA axes. Furthermore, when samples switched from cattle to bison, their trajectory through time did not change substantially. The results of the NMDS (not shown) corroborated these qualitative results.

Variation partitioning based upon CCA indicated that site effects explained the majority of the variation in species composition ($R_{CCAadj}^2 = 0.46$, Fig. 5.4). Year and management effects on species composition were negligible after adjustment, but both were still found to be significant in randomization tests at the 0.05 level (Table 5.3). There was a large shared site and management fraction ($R_{CCAadj}^2 = 0.08$) of explained variation. The pCCAs that considered site specific explanatory variables after factoring out year and management effects indicated that log Ca was the most important site specific variable ($R_{CCAadj}^2 = 0.07$; Table 5.3).

Although management explained only a very small proportion of the variation in species composition, I feel that for applied purposes it is still worth examining the management effects on particular species using a pCCA biplot (Fig. 5.5). The pCCA biplot displays only the 90 most abundant species (although all 307 species were included in the analysis) with respect to the three management variables after factoring out year and site effects as dummy variables. It appeared that both C3 and C4 grasses (open and

filled black circles on biplot respectively) appeared to decrease in cover with increased years of bison grazing. *Lespedeza cuneata* (sericea lespedeza), an invasive species, was positively correlated with years of bison grazing. However, this species is a target for herbicide application by the TNC so care must be taken in interpretation of this result. Annual and ruderal species, such as *Chamaesyce nutans* (eye bane), *Plantago virginica* (Virginia plantain), *Ambrosia artemisiifolia* (annual ragweed), and *Andropogon virginicus* (broomsedge bluestem), all were positively associated with the number of burns in the past five years.

DISCUSSION

The purpose of my study was to examine the influence of a variable management regime on plant richness and composition relative to site-to-site and year-to-year variability. My results suggest that the grasslands of the TGPP are not undergoing strong directional changes in richness or composition through time (i.e. they do not appear to be on a clear trajectory). Rather, samples appear to maintain their differences over time and respond somewhat idiosyncratically to management effects. This is not to say that management effects on richness and composition independent of site and year effects were irrelevant and uninterpretable, but simply that the management effects were relatively unimportant relative to inherent sources of landscape heterogeneity.

Role and drivers of site-to-site variation

The bulk of the variation in richness and composition was due to differences between sites, which I believe primarily reflects belowground differences between my samples. I found that calcium explained the predominant amount of site-to-site

variability in both richness and composition (Table 5.2). I interpret calcium's importance to be due to its role as a proxy variable for many other soil properties which are generally indicative of limestone derived soils (Palmer et al. 2003, Brokaw 2004). Using the same 20 sites analyzed in this study (only for the year 2002), Brokaw (2004) found that out of a set of 12 different soil nutrients and total carbon that the two most important variables for explaining plant composition (using pCCA) were total carbon and residual phosphorus. These variables were strongly positively correlated with the majority of soil nutrients and calcium and interpreted as a general indicator of limestone parent materials. These findings suggest that the important role that calcium played in my study does not necessarily contradict the body of work that suggests that below ground nutrients are the predominant drivers (not cations) of plant richness and composition in tallgrass prairie ecosystems (Turner et al. 1997, Burke et al. 1998, Baer et al. 2003, 2004).

Why was the relationship between calcium and richness negative? The positive relationship between calcium and soil nutrients (that Brokaw [2004] found in my samples) suggests one possible explanation. If soil high in calcium are the productive, nutrient rich sites, then plant species at these sites may experience stronger competitive exclusion and therefore have lower richness (Grime 1973, Tilman 1982). Pärtel's (2002) species pool hypothesis offers an alternative hypothesis for the negative correlation between richness and calcium (Palmer et al. 2003). The species pool hypothesis postulates that negative correlations between richness and pH (or calcium given their tight correlation) will occur in regions in which habitats high in pH were evolutionarily scarce. Palmer et al. (2003) found that the data from the Tallgrass Prairie Preserve

(including this study's sites as well as many others) appeared to support Pärtel's hypothesis in the grasslands but not in the woodlands.

Role and drivers of year-to-year variation

Year-to-year variation was a better explanatory variable of richness than composition. This indicates that changes in year-to-year composition were primarily idiosyncratic. Relative to site and management factors, the year factor explained almost equal amounts of variation in richness as site effects (21% compared to 25%) with 9 fewer parameters. With respect to species composition, year effects were more comparable with the small influence of management (both were < 1%).

Climatic variability is the most obvious driver of year-to-year variation (that is independent of management effects). In the tallgrass prairie plant community water is typically viewed as an important limiting resource, and it is generally positively associated with richness in grasslands (Cornwell and Grubb 2003, Adler and Levine 2007, Wilson 2007). In my study, richness was positively correlated with the winter and spring seasonal precipitation variables, but negatively correlated with the summer precipitation variable. Although all the season variables explained a significant amount of variance in richness, it appeared that the independent negative effect of the summer variable was largest with respect to richness (Table 5.2). Following Adler and Levine (2007), increased precipitation may act on richness directly by making the environment more benign and thus allowing more species to coexist. However, if this was generally true, then it is difficult to explain the strong negative correlation observed for summer rainfall. An alternative season specific hypothesis that may explain the negative relationship with summer rainfall is that increased precipitation in the summer previous

to sampling yields higher aboveground cover of C4 grasses. High yields of aboveground biomass could act to both inhibit germination of annuals and biennials later at the beginning of the next growing season as well as increase asymmetric light competition due to increased litter (Grace 2001). The positive influence of spring and winter rains may also be explained by the critical role that rainfall plays in stimulating germination in the annuals and biennials during the dormant seasons. In this sense precipitation is viewed as a non-resource variable that indirectly influences richness by increasing aboveground primary productivity the summer prior to sampling and by influencing germination of species rich plant groups.

Independent management effects

Despite the occurrence of 80 recorded fire events across my twenty sites and the usage of different grazers, my results indicated that management effects were relatively minor in comparison to inherent site and year heterogeneity. This result is in conflict with the dominant paradigm of tallgrass prairie ecology, which asserts that management tools such as fire and grazing are the dominant drivers of diversity and composition in the plant community (Collins 1987, Gibson and Hulbert 1987, Anderson 1990). Although the importance of site and year effects is not a novel insight in this system, they are rarely observed to dominate the influence of management so strongly. There are several reasons that likely led to this result. My samples cover a larger spatial extent than the majority of other tallgrass prairie studies, given the large size of the TGPP. Additionally, due to the observational nature of my study, I did not compare my samples to control sites that received no burning and/or grazing, although some of my sites went as long as 10 years without fire. And lastly, by utilizing site and year dummy variables, I provided

a conservative (but I believe more honest) estimate of the independent influence of management.

Although management explained little variation overall, years of bison grazing had a strong positive relationship with richness (Table 5.2). This suggests that “bison management” is having a positive influence on richness at the TGPP. I stress bison *management* rather than the influence of bison directly because different burning regimes are used for these two species: the average fire return interval was generally longer and the season of burn was more variable in the bison units when compared with the cattle units (Hamilton 2007). Also the bison units were grazed year round, while the cattle units were only seasonally grazed. Therefore, I am fairly confident bison management is having a positive effect on richness, but I am less confident that this is due actually to the presence of bison in comparison to cattle.

One explanation for the positive correlation between years of bison and richness may be related to the dietary differences between cattle and bison. Bison are thought to selectively forage on graminoids rather than forbs (in contrast to cattle which behave more as generalists) (Coppedge et al. 1998b, Knapp et al. 1999). This dietary behavior may release forbs from competitive exclusion due to the dominant C4 grasses. Years of bison grazing was negatively correlated with the cover of the majority of C3 and C4 species (Fig. 5.5). The increase in richness and decrease in graminoid cover in the bison managed sites lend modest support to the hypothesis that bison may increase richness of tallgrass prairie by decreasing the cover of graminoids relative to sites that were grazed seasonally by cattle.

Towne et al. (2005) undertook a more controlled approach to investigating the effects of bison and cattle grazing on plant richness and composition at the Konza tallgrass prairie over a 10 year period in annually burned pastures. They also found that overall the differences between bison and cattle were slight but that bison grazed patches had a higher cover of some forb species and gained forb species at a more rapid rate through time. They concluded the differential responses of vegetation to bison or cattle grazing may be predominately due to differences in management of these grazers rather than inherent differences in their biology. A similar effect may be taking place at my study site as well.

Years since last burn was most strongly correlated with changes in species composition (Fig. 5.5), but the other two management variables explained comparable amounts of variation (Table 5.3). Other studies have noted the strong positive relationship between cover of legumes and C4 grasses with frequency of dormant season burning (Hulbert 1988, Towne and Knapp 1996, Coppedge et al. 1998a, Peterson et al. 2007), although 84% of the burns in my samples occurred during the dormant season, there was not a clear relationship between fire frequency and either of these functional groups (Fig. 5.5). I found anecdotal evidence that ruderal species were positively associated with the number of burns in the past five years (Fig. 5.5), which is to be expected given that these samples were likely to have higher grazing pressure due to their more nutritious regrowth (Coppedge and Shaw 1998, Fuhlendorf and Engle 2001).

Management implications

Our results suggest that management decisions can cause significant changes in plant species richness and composition; however, the exact details of the management

plan for tallgrass prairie plant communities may be of less importance given the large sources of inherent landscape heterogeneity we observed. This may be welcome news for land managers because it suggests that intact tallgrass prairie plant communities may be relatively insensitive to the exact management prescription, as long as some combination of fire and grazing is present on the landscape. Although our study did not examine long-term unburned and/or ungrazed sites, it is worth noting that tallgrass prairie landscapes managed in this way are likely to experience increased woody encroachment (e.g., Briggs et al. 2002).

At our study site, TNC is utilizing a variable application of prescribed fire to meet several conservation goals. One goal is to maintain or increase the biological diversity of the plant community (Hamilton 2007). The non-directional changes in species composition and relatively weak increase in richness I observed over the 11-year period suggest that the management decisions are at the very least not detrimental to the plant community. TNC is also attempting to manage for wildlife habitat and diversity. Structural heterogeneity in the vegetation, attributed to the variable application of fire, resulted in the development of suitable habitat for a wider breadth of grassland bird species at the TGPP (Fuhlendorf et al. 2006, Coppedge et al. 2008). These results in conjunction with our findings suggest that the management decisions at the preserve contribute to important conservation goals, even if their effects on plant richness and composition are slight.

The importance of observational studies

Even without control treatments, observational studies can provide vital information about the ecology of a system (Underwood et al. 2000, Legg and Nagy 2006,

MacNeil 2008). To examine long-term trends at my study site, it was more appropriate to conduct an observational study (rather than a manipulative one) given my interest in the role of inherent heterogeneity as well as the temporally dynamic nature of the management at the preserve. My study demonstrates that long-term plots may be a potentially valuable source of information despite a lack of clearly defined treatments, if they are placed objectively (Palmer 1993). This is not to say that experimental studies are not necessary for uncovering driving mechanisms, but rather that it is often valuable to complement them with observational studies such that the context of their results can be better interpreted.

Conclusions

Management effects on plant richness and composition were relatively minor relative to inherent variation between sites and years. Management effects on species composition were nevertheless significant and interpretable, and I detected a significant positive correlation between bison management and plant richness. Given the overriding influence of inherent landscape heterogeneity on the plant community, the exact details of a management plan that incorporates the variable application of fire and grazing may not be of critical importance to maintaining tallgrass prairie plant communities.

ACKNOWLEDGMENTS

DJM received funding from the U.S. Environmental Protection Agency (EPA) under the Greater Research Opportunities (GRO) Graduate Program. The U.S. EPA has not officially endorsed this publication, and the views expressed herein may not reflect the views of the Agency. Additionally I thank Bob Hamilton, members of Laboratory for

Innovative Biodiversity and Analysis, and the Osage Nation for assistance with field work. NSF Grant Number EPS-0447262, The Oklahoma State University College of Arts and Science, The Oklahoma Nature Conservancy, The Spatial and Environmental Information Clearinghouse, The Phileology Trust, The Swiss Federal Institute for Forest, Snow and Landscape Research, and the Oklahoma Water Resources Research Institute for assistance at various stages of research at the Tallgrass Prairie Preserve.

LITERATURE CITED

- Abrams, M. D. and L. C. Hulbert. 1987. Effect of topographic position and fire on species composition in tallgrass prairie in northeast Kansas. *American Midland Naturalist* 117:442-445.
- Abrams, M. D., A. K. Knapp, and L. C. Hulbert. 1986. A ten-year record of aboveground biomass in a Kansas tallgrass prairie: effects of fire and topographic position. *American Journal of Botany* 73:1509-1515.
- Adler, P. B. and J. M. Levine. 2007. Contrasting relationships between precipitation and species richness in space and time. *Oikos* 116:221-232.
- Anderson, E. 1990. The historic role of fire in North American grassland. *in* Fire in North American tallgrass prairies, Eds. S. L. Collins and L. L. Wallace. University of Oklahoma Press, Norman, Oklahoma.
- Anderson, R. C. 1982. An evolutionary model summarizing the roles of fire, climate, and grazing animals in the origin and maintenance of grasslands: an end paper. Pages 297-308 *in* Grasses and grasslands: systematics and ecology, Eds. J. R. Estes, R. J. Tyrl, and J. N. Brunken. University of Oklahoma Press, Norman, Oklahoma, USA.
- Baer, S. G., J. M. Blair, S. L. Collins, and A. K. Knapp. 2003. Soil resources regulate productivity and diversity in newly established tallgrass prairie. *Ecology* 84:724-735.
- Baer, S. G., J. M. Blair, S. L. Collins, and A. K. Knapp. 2004. Plant community responses to resource availability and heterogeneity during restoration. *Oecologia* 139:617-629.

- Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73:1045-1055.
- Briggs, J. M., G. A. Hoch, and L. C. Johnson. 2002. Assessing the rate, mechanisms, and consequences of the conversion of tallgrass prairie to *Juniperus virginiana* forest. *Ecosystems* 5:578-586.
- Briggs, J. M. and A. K. Knapp. 1995. Interannual variability in primary production in tallgrass prairie - climate, soil-moisture, topographic position, and fire as determinants of aboveground biomass. *American Journal of Botany* 82:1024.
- Brokaw, J. M. 2004. Comparing explanatory variables in the analysis of species composition of a tallgrass prairie. *Proceedings of the Oklahoma Academy of Science* 84:33-40.
- Burke, I. C., W. K. Lauenroth, M. A. Vinton, P. B. Hook, R. H. Kelly, H. E. Epstein, M. R. Aguiar, M. D. Robles, M. O. Aguilera, K. L. Murphy, and R. A. Gill. 1998. Plant-soil interactions in temperate grasslands. *Biogeochemistry* 42:121-143.
- Candy, S. G. 1997. Poisson vs normal-errors regression in Mac Nally (1996). *Austral Ecology* 22:233-235.
- Chalcraft, D. R., B. J. Wilsey, C. Bowles, and M. R. Willig. 2009. The relationship between productivity and multiple aspects of biodiversity in six grassland communities. *Biodiversity and Conservation* 18:91-104.
- Collins, S. L. 1987. Interaction of disturbances in tallgrass prairie: a field experiment. *Ecology* 68:1243-1250.
- Collins, S. L. 1992. Fire frequency and community heterogeneity in tallgrass prairie vegetation. *Ecology* 73:2001-2006.

- Coppedge, B. R., D. M. Engle, C. S. Toepfer, and J. H. Shaw. 1998a. Effects of seasonal fire, bison grazing and climatic variation on tallgrass prairie vegetation. *Plant Ecology* 139:235-246.
- Coppedge, B. R., S. D. Fuhlendorf, W. C. Harrell, and D. M. Engle. 2008. Avian community response to vegetation and structural features in grasslands managed with fire and grazing. *Biological Conservation* 141:1196-1203.
- Coppedge, B. R., D. M. Leslie, and J. H. Shaw. 1998b. Botanical composition of bison diets on tallgrass prairie in Oklahoma. *Journal of Range Management* 51:379-382.
- Coppedge, B. R. and J. H. Shaw. 1998. Bison grazing patterns on seasonally burned tallgrass prairie. *Journal of Range Management* 51:258-265.
- Cornwell, W. K. and P. J. Grubb. 2003. Regional and local patterns in plant species richness with respect to resource availability. *Oikos* 100:417-428.
- Critchley, C. N. R., B. J. Chambers, J. A. Fowbert, R. A. Sanderson, A. Bhogal, and S. Rose. 2002. Association between lowland grassland plant communities and soil properties. *Biological Conservation* 105:199-215.
- Daubenmire, R. 1968. Ecology of fire in grasslands. *Advances in Ecological Research* 5:209-266.
- Diggle, P. J., K.-Y. Liang, and S. L. Zeger. 1994. *Analysis of Longitudinal Data*. Clarendon Press, Oxford.
- Fuhlendorf, S. D. and D. M. Engle. 2001. Restoring heterogeneity on rangelands: ecosystem management based on evolutionary grazing patterns. *Bioscience* 51:625-632.

- Fuhlendorf, S. D. and D. M. Engle. 2004. Application of the fire-grazing interaction to restore a shifting mosaic on tallgrass prairie. *Journal of Applied Ecology* 41:604-614.
- Fuhlendorf, S. D., W. C. Harrell, D. M. Engle, R. G. Hamilton, C. A. Davis, and D. M. Leslie. 2006. Should heterogeneity be the basis for conservation? Grassland bird response to fire and grazing. *Ecological Applications* 16:1706-1716.
- Gibson, D. J. and L. C. Hulbert. 1987. Effects of fire, topography and year-to-year climatic variation on species composition in tallgrass prairie. *Plant Ecology* 72:175-185.
- Grace, J. B. 2001. The roles of community biomass and species pools in the regulation of plant diversity. *Oikos* 92:193-207.
- Grime, J. P. 1973. Competitive exclusion in herbaceous vegetation. *Nature* 242:344-347.
- Hamilton, R. G. 1996. Using fire and bison to restore a functional tallgrass prairie landscape. Pages 208-214 *in* Transactions of the 61st North American Wildlife and Natural Resources Conference, Eds. Wildlife Management Institute, Washington D.C.
- Hamilton, R. G. 2007. Restoring heterogeneity on the Tallgrass Prairie Preserve: applying the fire-grazing interaction model. Pages 163-169 *in* Proceedings of the 23rd Tall Timbers Fire Ecology Conference: Fire in Grassland and Shrubland Ecosystems, Eds. R. E. Masters and K. E. M. Galley. Allen Press, Tall Timbers Research Station, Tallahassee, Florida, USA.
- Hill, M. O. and H. G. Gauch. 1980. Detrended Correspondence analysis: an improved ordination technique. *Vegetatio* 42:47-58.

- Hobbs, R. J., S. Yates, and H. A. Mooney. 2007. Long-term data reveal complex dynamics in grassland in relation to climate and disturbance. *Ecological Monographs* 77:545-568.
- Howe, H. F. 1994. Managing species diversity in tallgrass prairie: Assumptions and implications. *Conservation Biology* 8:691-704.
- Hulbert, L. C. 1988. Causes of fire effects in tallgrass prairie. *Ecology* 69:46-58.
- Knapp, A. K., J. M. Blair, J. M. Briggs, S. L. Collins, D. C. Hartnett, L. C. Johnson, and E. G. Towne. 1999. The keystone role of bison in North American Tallgrass Prairie. *Bioscience* 49:39-50.
- Landres, P. B., P. Morgan, and F. J. Swanson. 1999. Overview of the use of natural variability concepts in managing ecological systems. *Ecological Applications* 9:1179-1188.
- Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74:1659-1673.
- Legendre, P. and L. Legendre. 1998. *Numerical Ecology*. Elsevier, Boston.
- Legg, C. J. and L. Nagy. 2006. Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management* 78:194-199.
- Mac Nally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between - and reconciliation of - 'predictive' and 'explanatory' models. *Biodiversity and Conservation* 9:655-671.
- MacArthur, R. H. 1965. Patterns of species diversity. *Biological Reviews* 40:510-533.
- MacNeil, M. A. 2008. Making empirical progress in observational ecology. *Environmental Conservation* 35:193-196.

- McPherson, R. A., C. A. Fiebrich, K. C. Crawford, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A. Kloesel, A. D. Melvin, H. Shrivastava, J. M. Wolfenbarger, J. P. Bostic, D. B. Demko, R. L. Elliott, S. J. Stadler, J. D. Carlson, and A. J. Sutherland. 2007. Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology* 24:301-321.
- Nagelkerke, N. J. D. 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78:691-692.
- Nekola, J. C. and P. S. White. 2002. Conservation, the two pillars of ecological explanation, and the paradigm of distance. *Natural Areas Journal* 22:305-310.
- Økland, R. H. 1999. On the variation explained by ordination and constrained ordination axes. *Journal of Vegetation Science* 10:131-136.
- Oksanen, J., R. Kindt, P. Legendre, B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2008. *Vegan: community ecology package*.
- Oviatt, C. G. 1998. Geomorphology of Konza Prairie. Pages 35-47 in A. Knapp, J. Briggs, D. Hartnett and S. Collins (eds.), *Grassland Dynamics: Long-term ecological research in tallgrass prairie*, Eds. Oxford University Press, NY.
- Palmer, M. A., R. F. Ambrose, and N. L. Poff. 1997. Ecological theory and community restoration. *Restoration Ecology* 5:291-300.
- Palmer, M. W. 1993. Potential biases in site and species selection for ecological monitoring. *Environmental Monitoring and Assessment* 26:277-282.
- Palmer, M. W. 1995. How should one count species? *Natural Areas Journal* 15:124-136.

- Palmer, M. W. 2007. The vascular flora of the Tallgrass Prairie Preserve, Osage county, Oklahoma. *Castanea* 72:235-246.
- Palmer, M. W., J. R. Arévalo, M. C. Cobo, and P. G. Earls. 2003. Species richness and soil reaction in a northeastern Oklahoma landscape. *Folia Geobotanica* 38:381-389.
- Palmer, M. W. and M. Hussain. 1997. The unimodal (species richness-biomass) relationship in microcommunities emerging from soil seed banks. *Proceedings of the Oklahoma Academy of Sciences* 77:17-26.
- Pärtel, M. 2002. Local plant diversity patterns and evolutionary history at the regional scale. *Ecology* 83:2361-2366.
- Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology* 87:2614-2625.
- Peterson, D. W., P. B. Reich, and K. J. Wrage. 2007. Plant functional group responses to fire frequency and tree canopy cover gradients in oak savannas and woodlands. *Journal of Vegetation Science* 18:3-12.
- Petraitis, P. S., R. E. Latham, and R. A. Niesenbaum. 1989. The maintenance of species diversity by disturbance. *Quarterly Review of Biology* 65:393-418.
- Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R. C. Team. 2008. *nlme: Linear and Nonlinear Mixed Effects Models*.
- R Development Core Team. 2008. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Roberts, D. W. 1986. Ordination on the basis of fuzzy set theory. *Vegetatio* 66:123-131.

- Swanson, F. J., J. A. Jones, D. A. Wallin, and J. H. Cissel. 1994. Natural variability -- implications for ecosystem management Pages 85-99 *in* Eastside forest ecosystem health assessment. Volume II. Ecosystem management: principles and applications., Eds. M. E. Jense and P. S. Bourgeron. USDA Forest Service, Missoula, Montana, USA.
- ter Braak, C. J. F. and P. Šmilauer. 2002a. Canoco for Windows Version 4.5. Biometris - Plant Research International, Wageningen, The Netherlands.
- ter Braak, C. J. F. and P. Šmilauer. 2002b. CANOCO Reference manual and CanoDraw for Windows User's guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power, Ithaca, NY, USA.
- Tilman, D. 1982. Resource Competition and Community Structure. Princeton University Press, Princeton, NJ.
- Towne, E. G., D. C. Hartnett, and R. T. Cochran. 2005. Vegetation trends in tallgrass prairie from bison and cattle grazing. *Ecological Applications* 15:1550-1559.
- Towne, E. G. and A. K. Knapp. 1996. Biomass and density responses in tallgrass prairie legumes to annual fire and topographic position. *American Journal of Botany* 83:175-179.
- Towne, G. and C. Owensby. 1984. Long-term effects of annual burning at different dates in ungrazed Kansas tallgrass prairie. *Journal of Range Management* 37:392-397.
- Turner, C. L., J. M. Blair, R. J. Scharz, and J. C. Neel. 1997. Soil N and plant responses to fire, topography, and supplemental N in tallgrass prairie. *Ecology* 78:1832-1843.

- Underwood, A. J., M. G. Chapman, and S. D. Connell. 2000. Observations in ecology: you can't make progress on processes without understanding the patterns. *Journal of Experimental Marine Biology and Ecology* 250:97-115.
- Weiher, E. 2007. On the status of restoration science: obstacles and opportunities. *Restoration Ecology* 15:340-343.
- Wilsey, B. J., D. R. Chalcraft, C. M. Bowles, and M. R. Willig. 2005. Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity. *Ecology* 86:1178-1184.
- Wilson, S. D. 2007. Competition, resources, and vegetation during 10 years in native grassland. *Ecology* 88:2951-2958.

TABLES

Table 5.1. Variation partitioning of richness into its components for site, year, and management (abbreviated as manage) variables. Note that the factors: site, year, and management consist of q explanatory variables. The site and year factors were coded as dummy variables (see *Methods* for explanation). The partitioning was conducted with respect to the unadjusted and adjusted OLS coefficients of determination (R^2 and R_{adj}^2 respectively) as well as coefficients of determination for GLS models incorporating a single autoregressive term (R_{GLS}^2).

Factors (q)	Cofactors	OLS		GLS (AR1)
		R^2	R_{adj}^2	R_{GLS}^2
site (19)	year + management	0.48	0.48	0.27
year (10)	site + management	0.13	0.13	0.18
management (3)	site + year	0.04	0.04	0.02
site + year	management	0.01	-0.02	0.25
site + management	year	0.05	0.05	0.04
year + management	site	0.08	0.07	< 0.01
site + year + management	NA	-0.01	-0.02	< 0.01
Total (32)		0.77	0.74	0.77

Table 5.2. The conditional explanatory power of specific variables on species richness.

The variables site, year, and management are as defined in Table 5.1. The model used to account for within site temporal autocorrelation in the residuals is displayed with the estimated range and nugget (see *Methods* for details on the selection of the correlation model). Cells marked as "--" indicate that their values are identical to cells above. The standardized regression coefficient (β) is also given to indicate the strength and direction of the response of richness as well as an estimate of variation explained and the results of conditional *F*-tests (approximate tests).

Explanatory Variables	Covariables	Temporal Autocorrelation Model (range, nugget)	β	R^2_{GLS}	<i>F</i> -ratio	<i>p</i> -value
Factors (<i>q</i>)						
site (19)	year + management	Gaussian (3.92, 0.55)	NA	0.25	5.80	< 0.001
year (10)	site + management	--	NA	0.21	12.87	< 0.001
management (3)	site + year	--	NA	0.01	4.07	0.008
Site Specific Variables						
log Ca	slope + northness + year + management	Exponential (10.91, 0.32)	-0.28	0.05	6.51	0.012
slope	log Ca + northness + year + management	--	-0.10	< 0.01	0.71	0.400
northness	log Ca + slope + year + management	--	0.24	0.01	3.71	0.056
Year Specific Variables						
summer rain	winter rain + spring rain + site + management	Exponential (0.85)*	-0.19	0.03	26.65	< 0.001
winter rain	summer rain + spring rain + site + management	--	0.15	0.02	15.47	< 0.001
spring rain	summer rain + winter rain + site + management	--	0.14	0.01	11.43	0.001
Management Variables						
years of bison	years since burn + # of burns in 5 years + site + year	Gaussian (3.92, 0.55)	0.43	0.01	7.61	0.006
years since burn	years of bison + # of burns in 5 years + site + year	--	-0.11	< 0.01	3.52	0.062
# of burns in 5 years	years of bison + years since burn + site + year	--	-0.11	< 0.01	1.34	0.249

*a one parameter model (i.e., the nugget was assumed to be zero)

Table 5.3. The results of Monte Carlo randomization tests using the results of pCCA under the reduced model. Each randomization test was conducted with 999 iterations. The variables site, year, and management are as defined in Table 5.1. Cells marked as “--” indicate that their values are identical to cells above.

Explanatory Variable	Covariables	Type of Randomization *	R^2_{CCA}	R^2_{CCAadj}	F-ratio	p-value
Factors (q)						
site (19)	year + management	random shuffle	0.50	0.46	13.471	0.001
year (10)	site + management	toroidal shift	0.04	< 0.01	2.248	0.001
management (3)	site + year	random shuffle	0.01	< 0.01	2.145	0.001
--	--	toroidal shift	--	--	--	0.001
Site Specific Variables						
log Ca	slope + northness + year + management	random shuffle	0.13	0.12	19.735	0.001
slope	log Ca+ northness + year + management	random shuffle	0.03	0.03	8.152	0.001
northness	log Ca + slope + year + management	random shuffle	0.02	0.01	5.219	0.001
Management Variables						
years of bison	years since burn + # of burns in 5 years + site + year	random shuffle	< 0.01	< 0.01	2.314	0.001
--	--	toroidal shift	--	--	--	0.001
years since burn	years of bison + # of burns in 5 years + site + year	random shuffle	< 0.01	< 0.01	2.102	0.001
--	--	toroidal shift	--	--	--	0.001
# of burns in 5 years	years of bison + years since burn + site + year	random shuffle	< 0.01	< 0.01	1.616	0.001
--	--	toroidal shift	--	--	--	0.021

* random shuffle permutations were constrained to occur within a year, toroidal shifts were constrained within a given a site and preserved the temporal order of samples (see *Methods* for more explanation)

FIGURES

The Nature Conservancy's
Tallgrass Prairie Preserve, Osage Co., OK, USA

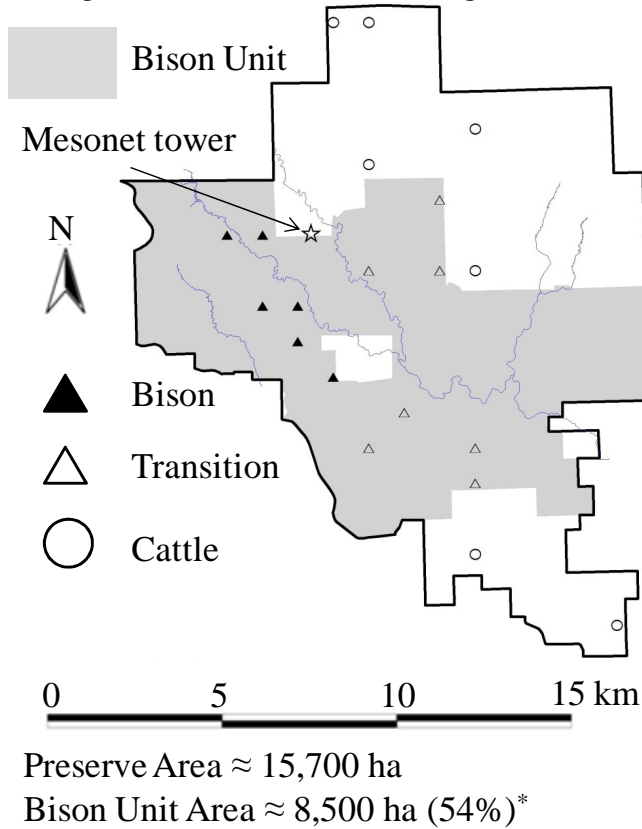


Fig. 5.1. A map of the Tallgrass Prairie Preserve. The shaded area denotes the bison unit which increased in area during the duration of the study. The Mesonet tower where the climate data was recorded is marked on the map as a star (\star). The twenty quadrats sampled each year of this study are displayed on the map. The sites that were bison at the beginning of the study (1998) are displayed with filled triangles (\blacktriangle), those that transitioned during the study from cattle to bison are denoted by unfilled triangles (\triangle), and the other cattle grazed samples are denoted by unfilled circles (\circ). *Area of bison unit is as of May 1, 2008.

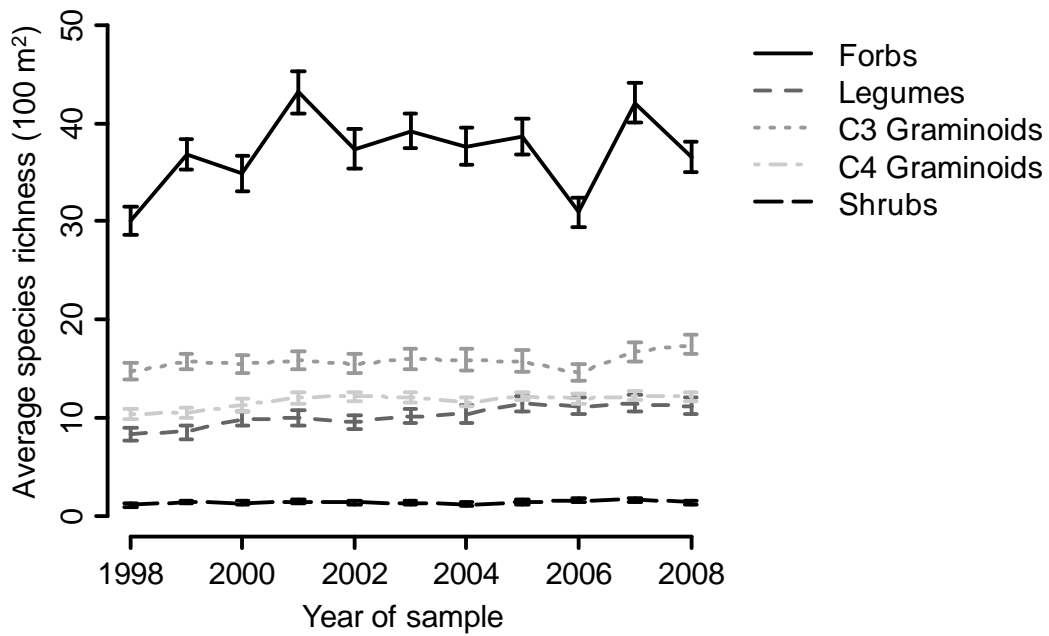


Fig. 5.2. The average species richness of five functional groups: forbs (that are not legumes), legumes, C3 graminoids (grasses, sedges, and rushes), C4 graminoids, and shrubs (woody plants) over the course of the study. The error bars display $\bar{x} \pm 1$ standard error.

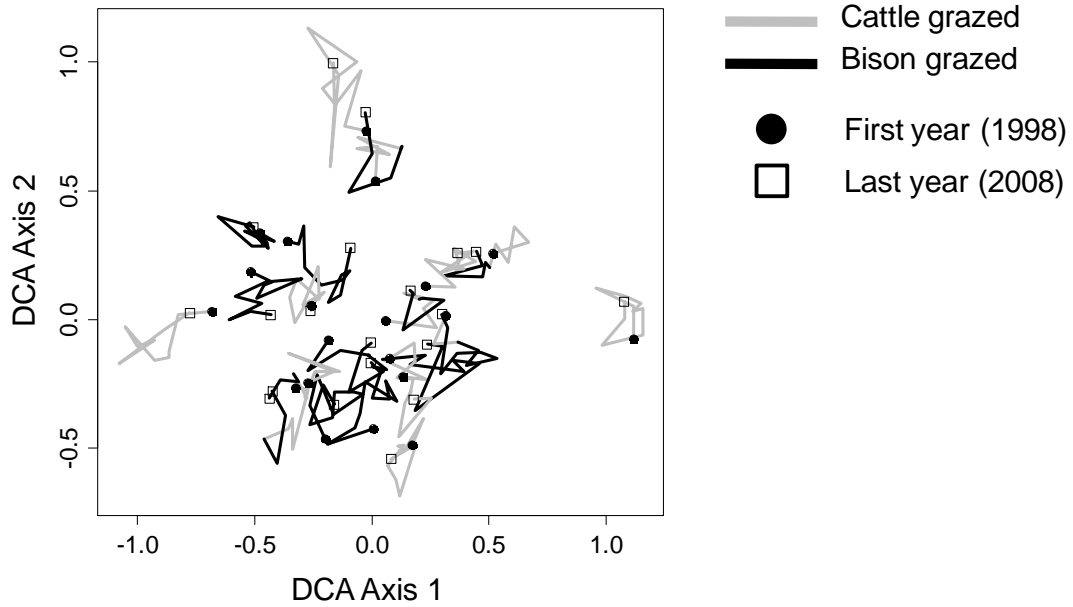
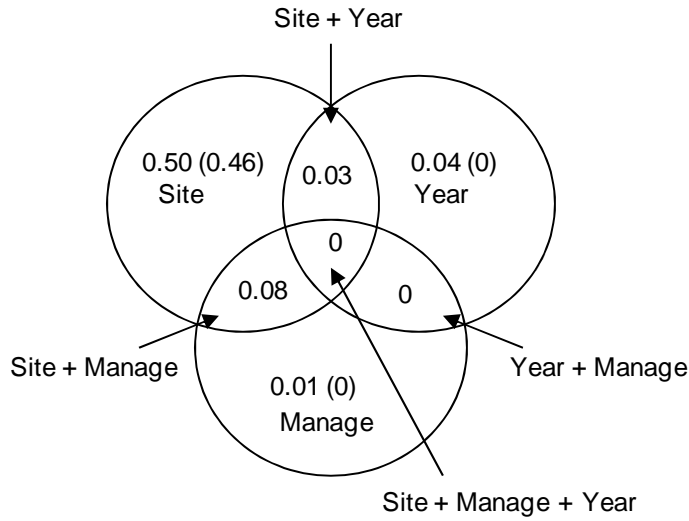


Fig. 5.3. DCA scatterplot displaying the 20 sites from 1998 to 2008. The eigenvalues were 0.152 and 0.108 for the first and second axis respectively. The time series is indicated by a line segment (—) with joints at each year, a dark black line joins years in which the plot experienced bison grazing and a grey line demarcates years in which the plot experienced cattle grazing. The 1998 samples are indicated by circles (●) and the 2008 samples are indicated by squares (□).



Total inertia = 1.48
 $R^2 = 0.63 (0.57)$

Fig. 5.4. A Venn diagram displaying the R^2_{CCA} (R^2_{CCAadj}) of each fraction resulting from the variation partitioning of species composition using pCCA. The three factors (classes of explanatory variables) were sites, years, and management. Note that R^2_{CCAadj} in this context should not be interpreted as fraction of explained variance but rather fraction of explained inertia. For the shared fractions I only report the unbiased R^2_{CCAadj} values.

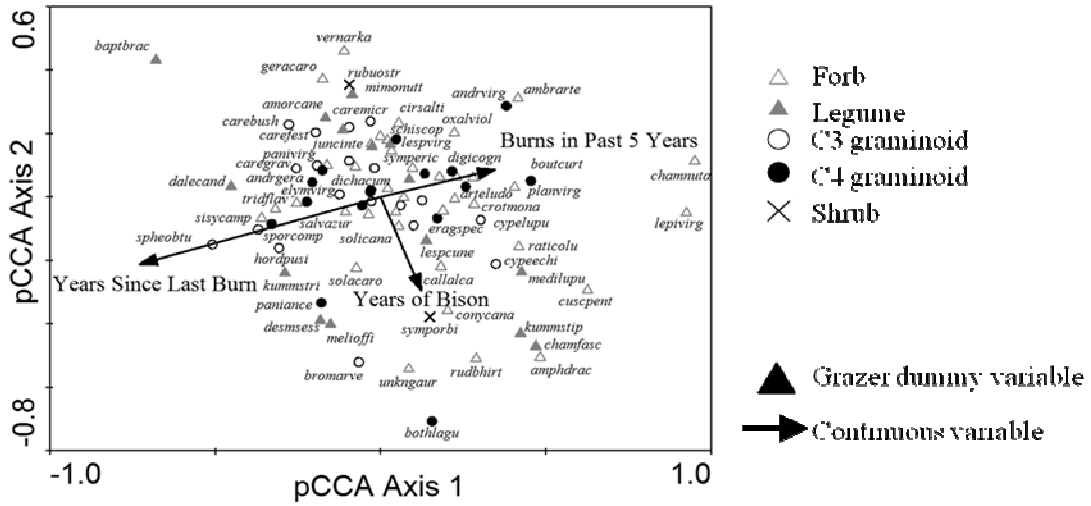


Fig. 5.5. A pCCA biplot displaying the influence of the management variables with the sites and years as covariables. The management variables in the model are indicated by black arrows. Only the 90 most abundant species of the 307 total are displayed for clarity. Abbreviations represent the first four letters of the genus and the first four of the letters of the species (see Appendix G). The symbol of each species depends on which functional type it belongs to (see legend).

APPENDICES

APPENDIX A

DERIVATION OF z_E AND w_E

Proof

From Eq. 3 in the main text the expected number of species is equal to:

$$E[S(A, T | R, \mathbf{p})] = S_p - \sum_{i=1}^{S_p} (1 - p_i)^{A[1+R(T-1)]} \quad (\text{A.1})$$

The Arrhenius definition of z is the slope of the SAR in log-log space; therefore, I can define z_E for the sampling model as the partial derivative of the natural logarithm (\ln) of the expected richness as a function of the \ln of area. For notational simplicity I will define this as:

$$z_E = \frac{\partial \ln S_E}{\partial \ln A} \quad (\text{A.2})$$

Because Eq. A.1 is given with respect to area and not $\ln(\text{area})$ and defined for S_E and not $\ln(S_E)$ I must use the chain rule to see that Eq. A.2 is actually:

$$z_E = \frac{A}{S_E} \cdot \frac{\partial S_E}{\partial A} = \frac{A}{S_p - \sum_{i=1}^{S_p} (1 - p_i)^{A[1+R(T-1)]}} \cdot \frac{\partial S_E}{\partial A} \quad (\text{A.3})$$

And using the rules of differentiation for exponential functions and the chain rule once more I find that:

$$\frac{\partial S_E}{\partial A} = -[1 + R(T - 1)] \sum_{i=1}^{S_p} \left[(1 - p_i)^{A[1+R(T-1)]} \ln(1 - p_i) \right] \quad (\text{A.4})$$

Combining this equation with Eq. A.3 I can see that the equation for z_E is:

$$z_E = \frac{-A[1 + R(T-1)] \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} \ln(1-p_i)]}{S_p - \sum_{i=1}^{S_p} (1-p_i)^{A[1+R(T-1)]}} \quad (\text{A.5})$$

A similar process can be used to find w_E .

$$w_E = \frac{\partial \ln S_E}{\partial \ln T} = \frac{T}{S_E} \cdot \frac{\partial S_E}{\partial A} = \frac{T}{S_p - \sum_{i=1}^{S_p} (1-p_i)^{A[1+R(T-1)]}} \cdot \frac{\partial S_E}{\partial T} \quad (\text{A.6})$$

And using the rules of differentiation for exponential functions and the chain rule I find that:

$$\frac{\partial S_E}{\partial T} = -AR \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} \ln(1-p_i)] \quad (\text{A.7})$$

Combining this equation with Eq. A.6 I can see that the equation for w_E is:

$$w_E = \frac{-TAR \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} \ln(1-p_i)]}{S_p - \sum_{i=1}^{S_p} (1-p_i)^{A[1+R(T-1)]}} \quad (\text{A.8})$$

APPENDIX B

DERIVATION OF u_E – THE PARTIAL DERIVATIVE OF z_E AND w_E WITH RESPECT TO THE
NATURAL LOGARITHM OF TIME AND AREA RESPECTIVELY.

Proof

In appendix A, I derived Eq. A.8 which stated that:

$$w_E = \frac{\partial \ln S_E}{\partial \ln T} = \frac{T}{S} \frac{\partial S_E}{\partial T} = \frac{-TAR \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} \ln(1-p_i)]}{S_p - \sum_{i=1}^{S_p} (1-p_i)^{A[1+R(T-1)]}} \quad (\text{B.1})$$

From this equation I will calculate the partial of w_E with respect to $\ln A$. The formulation of this partial I will refer to as u_E :

$$u_E = \frac{\partial^2 \ln S_E}{\partial \ln A \partial \ln T} = \frac{\partial w_E}{\partial \ln A} = A \frac{\partial w_E}{\partial A} \quad (\text{B.2})$$

Before solving for this it is helpful to break Eq. B.1 into three separate functions and derive the partial derivative with respect to A for each function:

$$w_E = \frac{fg}{h} \text{ where} \quad (\text{B.3})$$

$$f = -TAR, \quad (\text{B.4})$$

$$f' = \frac{\partial f}{\partial A} = -TR, \quad (\text{B.5})$$

$$g = \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} \ln(1-p_i)], \quad (\text{B.6})$$

$$g' = \frac{\partial g}{\partial A} = [1 + R(T-1)] \sum_{i=1}^{S_p} [(1-p_i)^{A[1+R(T-1)]} (\ln(1-p_i))^2], \quad (\text{B.7})$$

$$h = S_p - \sum_{i=1}^{S_p} (1 - p_i)^{A[1+R(T-1)]}, \text{ and} \quad (\text{B.8})$$

$$h' = \frac{\partial h}{\partial A} = -[1 + R(T-1)] \sum_{i=1}^{S_p} [(1 - p_i)^{A[1+R(T-1)]} \ln(1 - p_i)]. \quad (\text{B.9})$$

After applying the quotient and product rules to Eq. B.3 I find that:

$$\frac{\partial w_E}{\partial A} = \frac{(f'g + g'f)h - h'(fg)}{h^2} \quad (\text{B.10})$$

Finally, u_E is calculated by substituting Eq. B.10 into Eq. B.2:

$$u_E = A \left[\frac{(f'g + g'f)h - h'(fg)}{h^2} \right] \quad (\text{B.11})$$

Unfortunately when equations B.4-B.9 are substituted into Eq. B.11, the resulting formula does not readily simplify, and therefore the final solution is too large to show here in its entirety. By Clairaut's theorem I know that Eq. B.11 will also result from taking the partial of z_E with respect to $\ln T$.

$$u_E = A \frac{\partial w_E}{\partial A} = T \frac{\partial z_E}{\partial T} = \frac{\partial^2 \ln S_E}{\partial \ln T \partial \ln A} \quad (\text{B.12})$$

APPENDIX C

GRAPHICS DISPLAYING THE PREDICTIONS OF THE SAMPLING MODEL FOR FIVE RELATIVE ABUNDANCE DISTRIBUTIONS (RADs) WITH DIFFERENT LEVELS OF EVENNESS AND FIVE VALUES OF THE REPLACEMENT RATE (R).

All of these figures were created with model parameters equal to those in the main text. Both area and time were varied from 1 to 16384 by successive doublings of scale and the size of the species pool (S_p) was 800. Figures C2 through C4 were generated with the aid of the R package *dichromat* v1.2-2 (Lumley 2008).

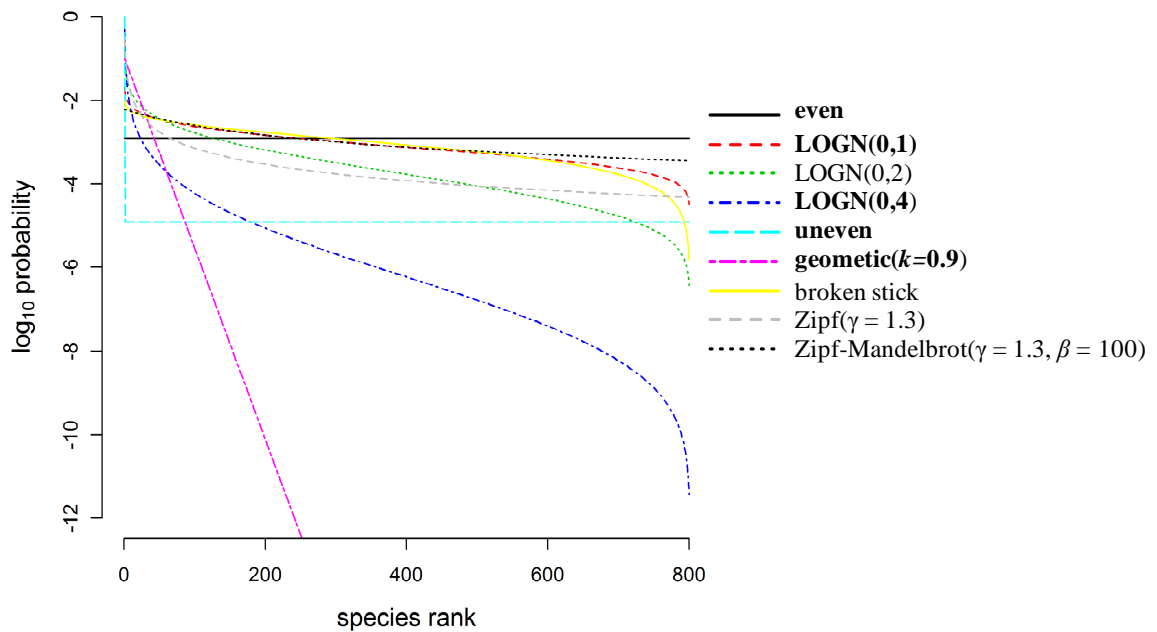


Fig. C1. The \log_{10} probability rank diagrams for all nine of the different relative abundance distributions (RADs): even, three lognormal, uneven, geometric, broken stick, Zipf, and Zipf-Mandelbrot. Note that the RADs that are bold in the figure legend are the five RADs which were chosen to represent the diversity of possible species-time-area relationships in the manuscript. Also note that the geometric RAD is linear in semi-log space over its entire range and is not shown in its entirety. The least common species in the geometric RAD had a \log_{10} probability of -37.56.

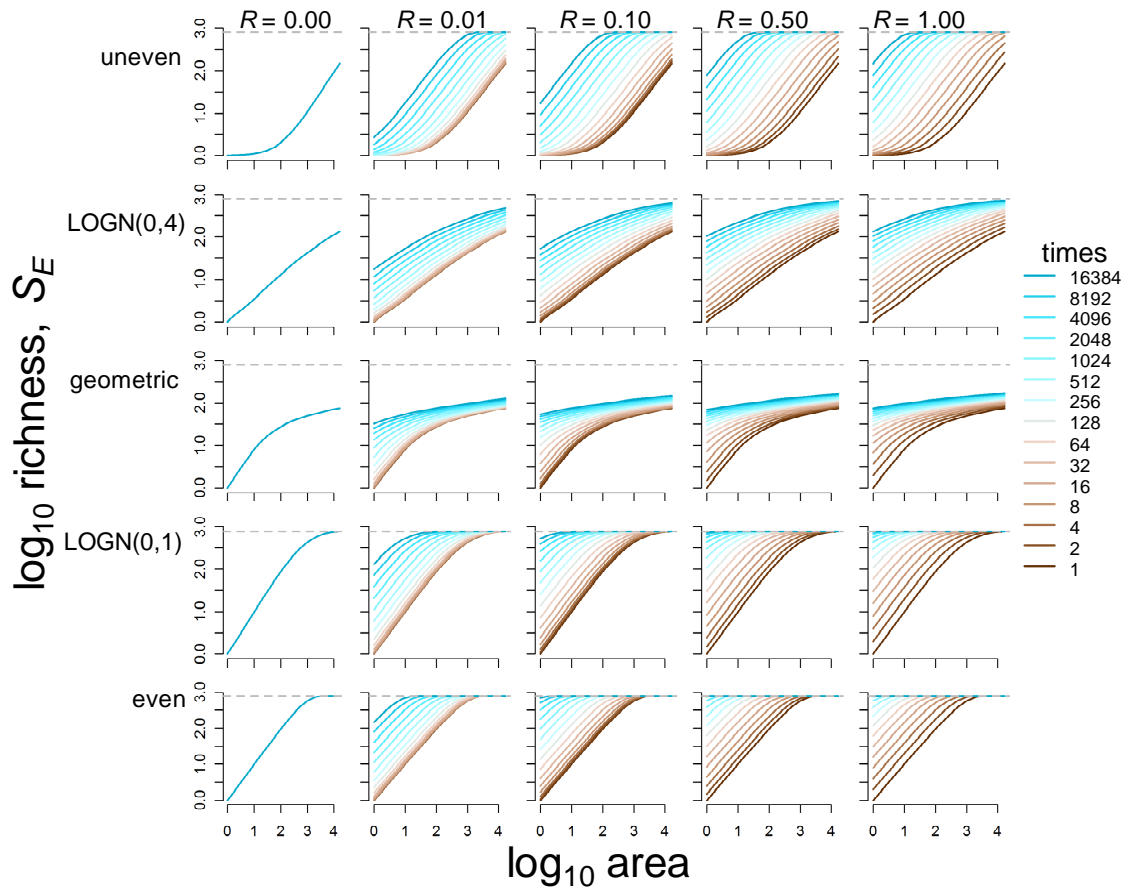


Fig. C2. The species-area relationship (SAR) for five values of R (columns) and the five RADs (rows). The color of the curves indicate the temporal scale of the SAR (see legend): brown indicates T was small and blue indicates T was large. The dashed grey line indicated the \log_{10} of the size of the species pool, S_P .

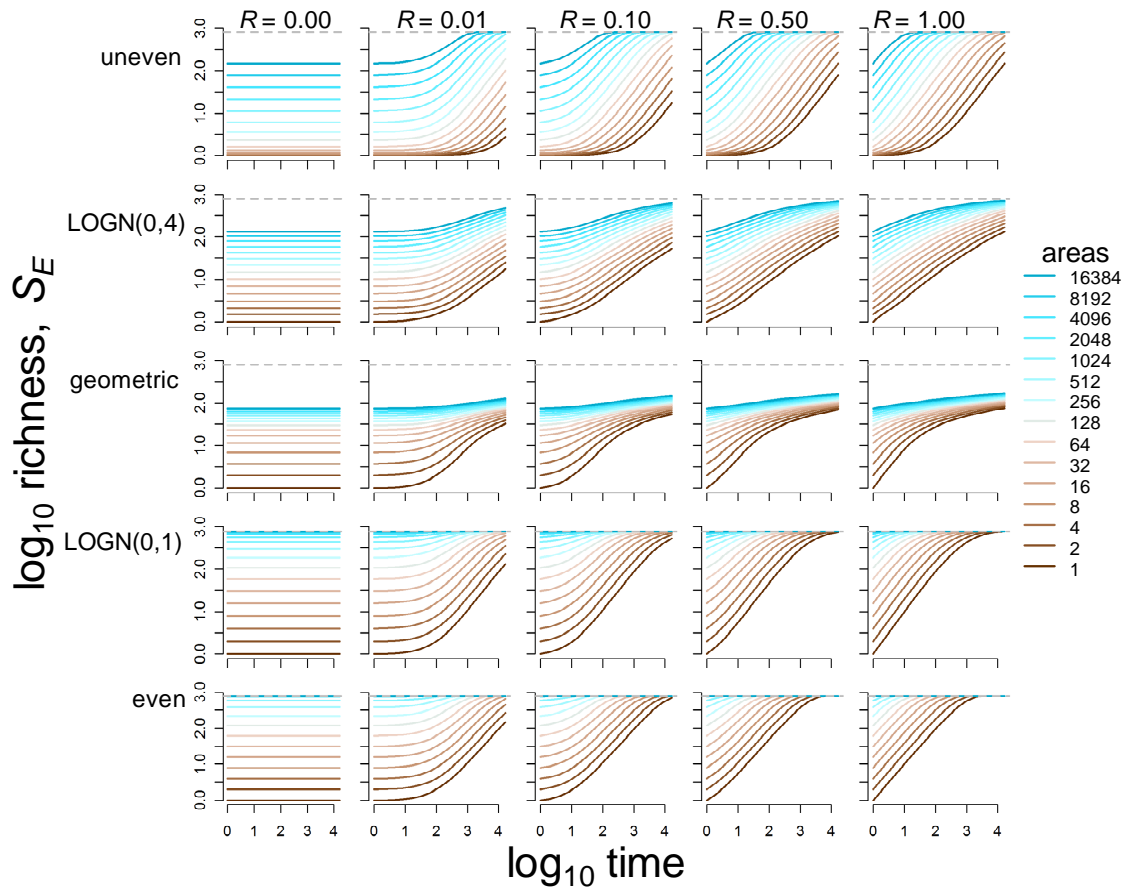


Fig. C3. The species-time relationship (STR) for five values of R (columns) and the five RADs (rows). The color of the curves indicate the spatial scale of the STR (see legend): brown indicates A was small and blue indicates A was large. The dashed grey vertical line indicated the \log_{10} of the size of the species pool, S_P .

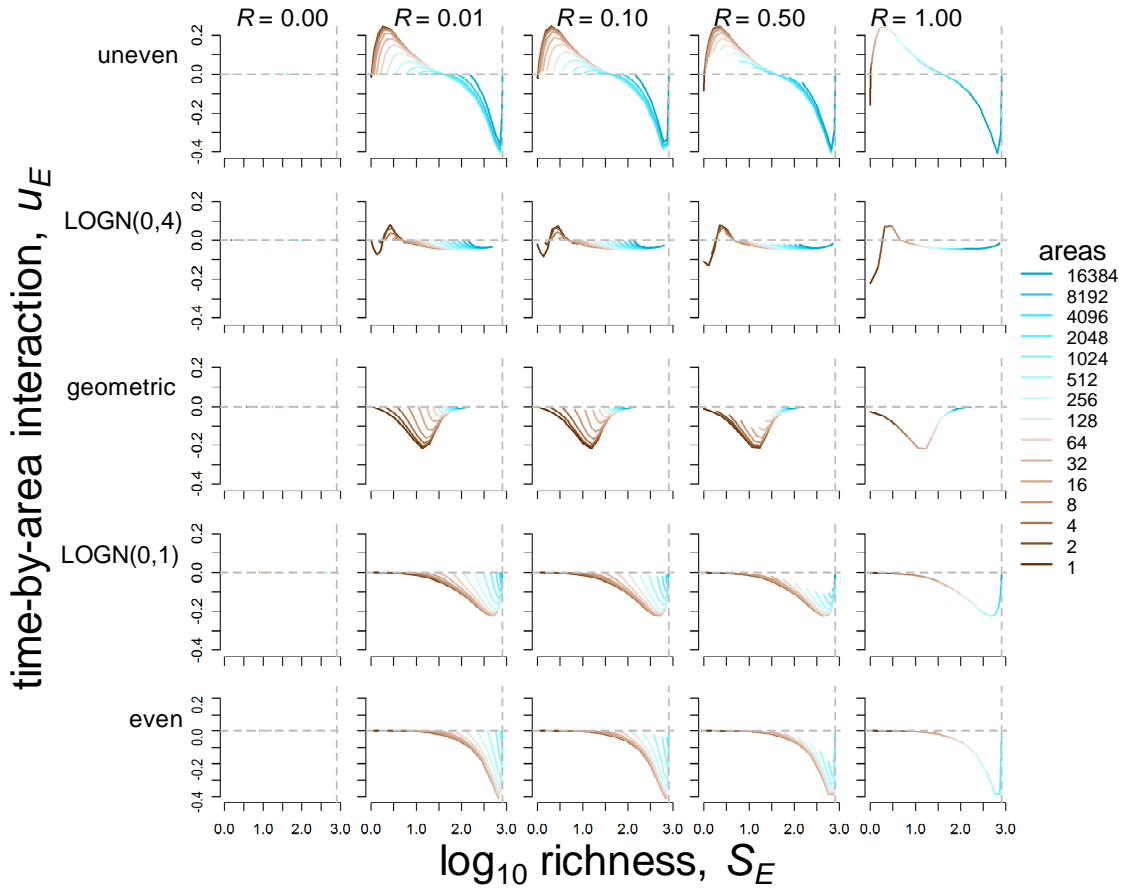


Fig. C4. The distribution of time-by-area interaction (u_E) as a function of \log_{10} richness (S_E) for five values of R (columns) and the five RADs (rows). Each curve was generated by holding area constant and varying time. There are no visible curves when $R = 0$ because u_E was equal to zero. Positive values were only observed at fine scales under low evenness. The color of the curves indicate at what scale u_E was calculated at (see legend): brown indicates A was small and blue indicates A was large. The dashed horizontal line indicates zero, and the dashed vertical line indicates the \log_{10} size of the species pool, S_P .

LITERATURE CITED

Lumley, T. 2008. Dichromat: color schemes for dichromats. R package version 1.2-2.

<http://cran.r-project.org/>.

APPENDIX D

THE DISTRIBUTION AND COLLINEARITY OF EACH THE MANAGEMENT VARIABLES.

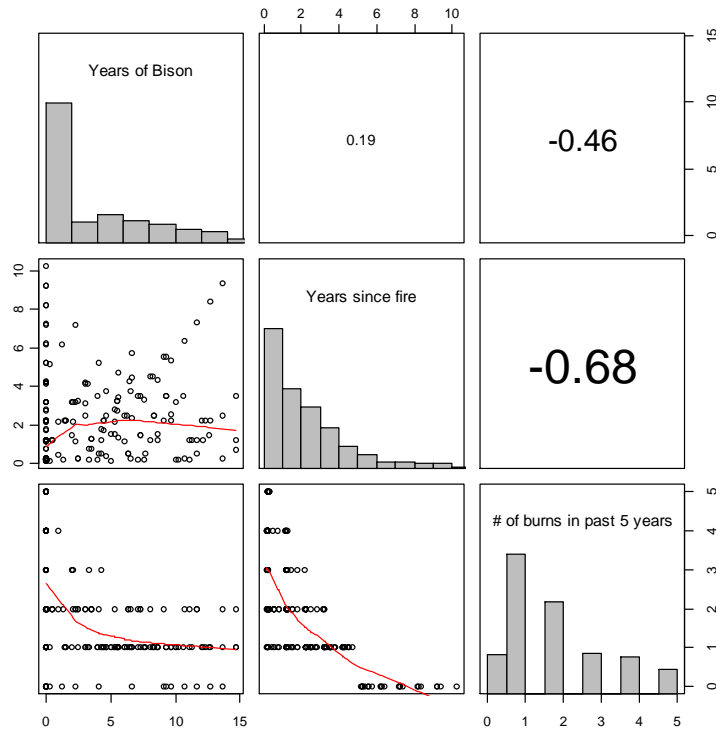


Fig. D1. The diagonal of the matrix displays a frequency histogram for each discrete management variable. The magnitude and direction of the correlation of two variables is indicated by the cells in the upper triangle. The lower triangle indicates the scatterplot of the two variables. A lowess smoothing function was applied to the scatterplots to aid visual interpretation.

APPENDIX E

ISOTROPIC VARIOGRAM MODELS FOR WITHIN SITE TEMPORAL AUTOCORRELATION OF THE RESIDUALS OF THE GLS MODELS OF SPECIES RICHNESS.

For each GLS model the single and double parameter formulations of five different models were compared using AIC which penalizes for the additional parameter.

Table E1. The single parameter formulation of five isotropic variogram models. This table is recreated from Table 2 in Pinherio et al. (2008).

Model name	Single parameter formulation
Exponential	$\gamma(s, \rho) = 1 - \exp(-s / \rho)$
Gaussian	$\gamma(s, \rho) = 1 - \exp[-(s / \rho)^2]$
Linear	$\gamma(s, \rho) = 1 - (1 - s / \rho)I(s < \rho)$
Rational quadratic	$\gamma(s, \rho) = (s / \rho)^2 / [1 + (s / \rho)^2]$
Spherical	$\gamma(s, \rho) = 1 - [1 - 1.5(s / \rho) + 0.5(s / \rho)^3]I(s < \rho)$

In the table above, s is the distance between two observations (in my case number of years) and ρ is the range of the model or the distance at which observations are no longer correlated. The two parameter version of each model can be formulated simply by the addition of a third parameter, the nugget (c_o):

$$\gamma_{nugg}(s, c_o, \rho) = \begin{cases} c_o + (1 - c_o)\gamma(s, \rho), & s > 0, \\ 0, & s = 0. \end{cases}$$

APPENDIX F

THE RELATIONSHIP BETWEEN RICHNESS AND SEVERAL OTHER DIVERSITY INDICES

EXAMINED IN WILSEY ET AL. (2005).

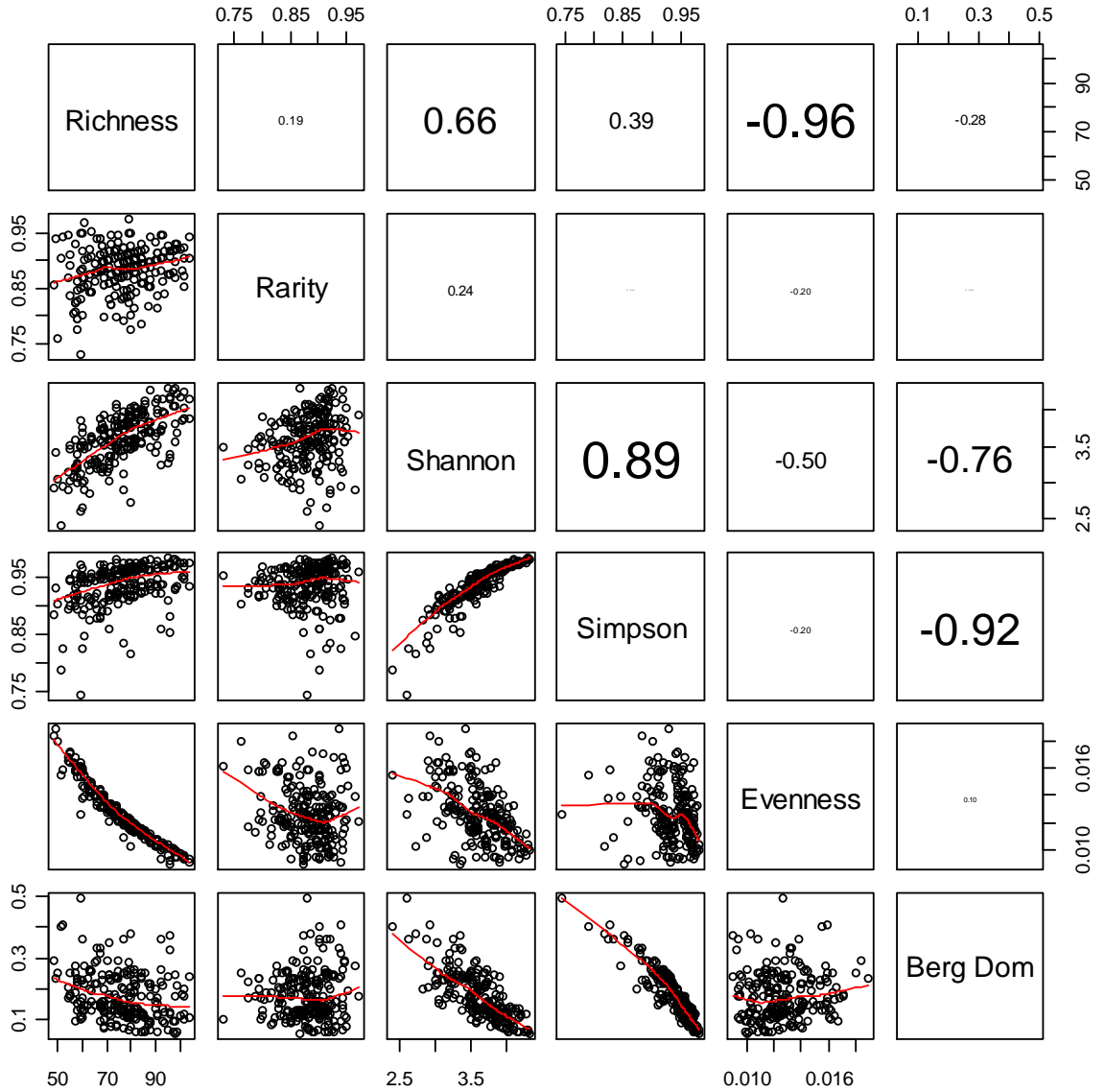


Fig. F1. The relationship between richness and each diversity index. The linear correlation coefficients are indicated in the upper triangular matrix. A lowess smoothing function was applied to the scatterplots to aid visual interpretation.

APPENDIX G

LEGEND OF SPECIES NAMES AND FUNCTIONAL GROUPS FOR FIG. 5.5.

Table G1. The abbreviation and functional group for each species displayed in Fig. 5.5 – the pCCA biplot of management effects. Nomenclature follows the PLANTS database (USDA NRCS 2008)

Abbreviated name	Binomial	Functional Group
ambrarte	<i>Ambrosia artemisiifolia</i>	Forb
amphdrac	<i>Amphiachyris dracunculoides</i>	Forb
arteludo	<i>Artemisia ludoviciana</i>	Forb
callalca	<i>Callirhoe alcaeoides</i>	Forb
chamnuta	<i>Chamaesyce nutans</i>	Forb
cirsalti	<i>Cirsium altissimum</i>	Forb
conycana	<i>Conyza canadensis</i>	Forb
crotmona	<i>Croton monanthogynus</i>	Forb
cuscpent	<i>Cuscuta pentagona</i>	Forb
gaurunko	<i>Gaura sp.</i>	Forb
geracaro	<i>Geranium carolinianum</i>	Forb
lepivirg	<i>Lepidium virginicum</i>	Forb
oxalviol	<i>Oxalis violacea</i>	Forb
planvirg	<i>Plantago virginica</i>	Forb
raticolu	<i>Ratibida columnifera</i>	Forb
rudbhirt	<i>Rudbeckia hirta</i>	Forb
salvazur	<i>Salvia azurea var. grandiflora</i>	Forb
sisycamp	<i>Sisyrinchium campestre</i>	Forb
solacaro	<i>Solanum carolinense</i>	Forb
solicana	<i>Solidago canadensis</i>	Forb
symperic	<i>Symphyotrichum ericoides</i>	Forb
vernarka	<i>Vernonia arkansana</i>	Forb
amorcane	<i>Amorpha canescens</i>	Legume
baptbrac	<i>Baptisia bracteata</i>	Legume
chamfasc	<i>Chamaecrista fasciculata</i>	Legume

dalecand	<i>Dalea candida</i>	Legume
desmsess	<i>Desmodium sessilifolium</i>	Legume
kummstip	<i>Kummerowia stipulacea</i>	Legume
kummstri	<i>Kummerowia striata</i>	Legume
lespcune	<i>Lespedeza cuneata</i>	Legume
lespvirg	<i>Lespedeza virginica</i>	Legume
medilup	<i>Medicago lupulina</i>	Legume
melioffi	<i>Melilotus officinalis</i>	Legume
mimonutt	<i>Mimosa nuttallii</i>	Legume
bromarve	<i>Bromus arvensis</i>	C3 grass
carebush	<i>Carex bushii</i>	C3 grass
carefest	<i>Carex festucacea</i>	C3 grass
caregrav	<i>Carex gravida</i>	C3 grass
caremicr	<i>Carex microdonta</i>	C3 grass
cypeechi	<i>Cyperus echinatus</i>	C3 grass
cypelupu	<i>Cyperus lupulinus</i>	C3 grass
dichacum	<i>Dichanthelium acuminatum</i>	C3 grass
elymvirg	<i>Elymus virginicus</i>	C3 grass
hordpusi	<i>Hordeum pusillum</i>	C3 grass
juncinte	<i>Juncus interior</i>	C3 grass
spheobtu	<i>Sphenopholis obtusata</i>	C3 grass
andrgera	<i>Andropogon gerardii</i>	C4 grass
andrvirg	<i>Andropogon virginicus</i>	C4 grass
bothlagu	<i>Bothriochloa laguroides ssp. torreyana</i>	C4 grass
boutcurt	<i>Bouteloua curtipendula</i>	C4 grass
digicogn	<i>Digitaria cognata</i>	C4 grass
eragspec	<i>Eragrostis spectabilis</i>	C4 grass
paniance	<i>Panicum anceps</i>	C4 grass
panivirg	<i>Panicum virgatum</i>	C4 grass
schiscop	<i>Schizachyrium scoparium</i>	C4 grass
sporcomp	<i>Sporobolus compositus</i>	C4 grass
tridflav	<i>Tridens flavus</i>	C4 grass
rubuostr	<i>Rubus ostryifolius</i>	Shrub
symporbi	<i>Symphoricarpos orbiculatus</i>	Shrub

VITA

Daniel J. McGlinn

Candidate for the Degree of

Doctor of Philosophy

Dissertation: SPATIAL AND TEMPORAL SCALING OF SPECIES COMPOSITION
AT THE TALLGRASS PRAIRIE PRESERVE, OKLAHOMA:
IMPLICATIONS FOR THEORY AND CONSERVATION

Major Field: Plant Ecology

Biographical:

Personal Data: Born in Lawrence, KS on August 27th, 1982 to James and
Jeanne McGlinn.

Education: Graduated from T.C. Roberson High School in Asheville, NC in
2000. In 2004, received a Bachelor of Science in Biology with Honors
from the University of North Carolina at Chapel Hill, Chapel Hill, NC.
Completed the requirements for the Doctor of Philosophy in Plant
Sciences at Oklahoma State University, Stillwater, OK in June of 2009.

Professional Memberships: Ecological Society of America and Southwestern
Association of Naturalists

Awards and Honors: Distinguished Graduate Fellowship, Greater Research
Opportunities Fellowship, Enhancing Linkages Between Mathematics &
Ecology Scholarship, and Audubon Award

Name: Daniel J. McGlinn

Date of Degree: July, 2009

Institution: Oklahoma State University

Location: Stillwater, Oklahoma

Title of Study: SPATIAL AND TEMPORAL SCALING OF SPECIES COMPOSITION AT THE TALLGRASS PRAIRIE PRESERVE, OKLAHOMA: IMPLICATIONS FOR THEORY AND CONSERVATION

Pages in Study: 156

Candidate for the Degree of Doctor of Philosophy

Major Field: Plant Ecology

Scope and Method of Study: The objective of this study was to advance our understanding of the drivers of species turnover using scaling relationships and patterns of species composition. To accomplish this objective I performed four studies in which: 1) I provided the meta-data for a long-term, multi-scale study on the vascular plants of a tallgrass prairie, 2) I developed a model of the species-time-area relationship (STAR) based on the sampling effect and compared it to an empirical STAR from a tallgrass prairie, 3) I tested if the rate of environmental and community distance decay were positively correlated in two habitat types, and 4) I quantified the importance of management-based heterogeneity relative to inherent sources of spatiotemporal heterogeneity on species richness and composition of a tallgrass prairie plant community over a period of 11 years. All four of these studies were based upon vegetation samples collected at the Tallgrass Prairie Preserve in Osage County, Oklahoma. Additional data were collected on soil cations and climate at the study site. Management records were provided by The Nature Conservancy who owned and managed the preserve.

Findings and Conclusions: I found that the sampling effect generated a STAR that was similar to the empirical relationship under certain conditions. Our model demonstrated that non-zero time-by-area interactions, which are the most important quantitative aspect of the STAR, are not necessarily attributable to ecological drivers and may result instead from a purely neutral sampling process. The geometry of the environment, as quantified by the rate of environmental distance decay, was positively correlated with the rate of species turnover in the grassland but not the woodland habitat. This suggests that one of the central tenets of the Environmental Texture Hypothesis is relevant at local spatial scales under certain conditions. Management had a significant but relatively unimportant influence of both species richness and composition of the tallgrass prairie plant community. Site effects were the most important source of heterogeneity, but year effects were comparable with respect to richness. The exact details of management may not be as critical for maintaining tallgrass prairie plant communities as long as woody plant encroachment is kept in check.

ADVISER'S APPROVAL: Dr. Michael W. Palmer
