UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ENVIRONMENT AND INTERSPECIFIC INTERACTIONS AT THE MARGIN OF SPECIES RANGES: A SPATIAL ANALYSIS OF FOREST COMMUNITIES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

DARYN ROBERT HARDWICK
Norman, Oklahoma
2017

ENVIRONMENT AND INTERSPECIFIC INTERACTIONS AT THE MARGIN OF SPECIES RANGES: A SPATIAL ANALYSIS OF FOREST COMMUNITIES

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL SUSTAINABILITY

BY

Dr. Bruce Hoagland, Chair

Dr. Scott Greene

Dr. Jennifer Koch

Dr. Heather McCarthy

Dr. Lara Souza

To my friends and family who have kept me relatively sane these past four years. Particularly to my siblings, Daniel, Stephanie, and Dylan, my uncle, Robert, my father, Shane, and my mother, Angela, for their consistent support during my time at the University of Oklahoma.

Acknowledgements

Firstly, I would like to mention several of the courses and people that have directed me towards this degree and field of study. At the University of Wisconsin - Stevens Point, Dr. Lisa Theo led me to geography and demonstrated the value of geography as a discipline. Dr. Mikhail Blinnikov, my master's advisor at Saint Cloud State University, first introduced me to biogeography. I am also grateful to Dr. Gareth John for encouraging me to pursue a doctoral degree. In my attempts to become a better plant ecologist, two courses at the University of Oklahoma were especially informative: Plant Physiological Ecology taught by Dr. Heather McCarthy and Community Ecology seminar with Dr. Michael Kaspari.

Additionally, I would like to express my gratitude to my doctoral committee. Dr. Scott Greene, Dr. Jennifer Koch, Dr. Heather McCarthy, and Dr. Lara Souza provided valuable guidance and feedback throughout all stages of this dissertation including time before its formulation. The well-rounded foci of these committee members in statistics, geography, landscape ecology, physiology, and ecology translated to productive discussion during the proposal, general examination, writing, and defense of this dissertation. Finally, I am grateful to my doctoral advisor, Dr. Bruce Hoagland, for the direction, guidance, and critiques of ideas, theoretical ponderings, and of the dissertation these past four years. The doctoral degree and dissertation were long, but ultimately rewarding, and would not have been accomplished without my committee.

Table of Contents

List of Tables

List of Figures

Abstract

Problem: Species' geographic ranges are determined, in part, by suitable environmental conditions, the ability to reach sites possessing those environmental conditions, and the capability to survive interactions with other species until reproductive age. Species' geographic ranges and community composition have been related to environmental conditions frequently, to dispersal limitation infrequently, and rarely to biotic interactions. This dissertation utilizes spatially explicit analysis to further elucidate the effect of geography on community composition, beta diversity, interspecific interactions, and the intersections of them on species' geographic ranges. Contradictory evidence suggests that interspecific interactions become either more facilitative or competitive with increasing stress, however results appear to be affected by the scale of study. Therefore, I performed a regional scale study to analyze interspecific interactions across a regional stress gradient, the proximity to a species' range margin.

Methods: Bottomland and upland forests consisting of 91 species across 307 sites were analyzed across spatial and environmental gradients. Sites were separated on the basis of species presence using a Nonmetric Multidimensional Scaling ordination and grouped using an agglomerative hierarchical clustering algorithm. To explain variance in community composition, climate, spatial, and edaphic variables were related to species presence across sites with a Redundancy Analysis ordination which uses multiple Canonical Correspondence Analyses. Species' range margins were delineated using species occurrence data from the Forest Inventory and Analysis with MaxEnt modeling software. In order to test how interspecific interactions change across species' ranges, a spatially explicit model assessing co-occurrence in groups of nearest neighbor sites was created using Python programming language. Interspecific interactions were subsequently compared to randomly generating communities.

Results: Species within site groups respond primarily to average annual precipitation and secondarily to the standard deviation in monthly precipitation. Sites are more aggregated in environmental space than physical space. Variation in community composition is best explained by climatic variables (22%) followed by spatial (9.9%) and edaphic (9.8%) variables. Beta diversity is significantly positively correlated with climate distance, mammal and bird beta diversity, the variance in distances to species' range margins, and soil texture distance between sites. Net interspecific interactions monotonically shift from competitive to facilitative with proximity to species' range margins. Species exhibit monotonic, unimodal, and multimodal relationships between net interaction intensity and proximity to their respective range margins in approximately equal proportions. As conditions become more favorable, species within a genus interact more competitively than species of different genera. Both locally rare and locally dominant species experience net competitive interactions, while locally common species experience net facilitative interactions. Regionally dominant species experience a greater intensity of net competitive interactions and do not experience net facilitative interactions across the stress gradient. Regionally rare species experience relatively net neutral interactions; however, they experience net facilitative interactions beyond their range margins. Net interactions in null communities are of higher competitive intensity regardless of proximity to range margins.

Significance: By discerning the effects of environment, space, and interspecific interactions on community composition, this dissertation serves to improve our understanding about the structure of species' ranges. Projected climate change may force species to adapt to novel conditions or migrate to suitable habitat. Additionally, species will likely exist in communities consisting of different species assemblages. Modeling community composition is difficult since research on the change of biotic interactions through space and through species' ranges is deficient. Analyzing the proximity to range margins as a surrogate for stress provides a spatial framework for utilizing biotic interactions in the modeling of species' geographic ranges. For sessile organisms such as trees, dispersal of propagules is the primary method of shifting or expanding the species' range. Facilitative interactions at shifting range margins may provide refuge for propagules facing stressful environmental conditions. Finally, results on how interactions affect species performance across their geographic range inform broader discussion of the species niche. I have provided spatial evidence that the realized niche of a species is expanded through net facilitative interactions beyond species' range margins. Understanding the role of facilitative interactions at range margins may improve upon and direct conservation efforts for species at risk under projected climate change.

Chapter I

Overview

Introduction

The nature of the geographic range or extent of species has been a focus of inquiry for centuries by naturalists, ecologists, and biogeographers alike (Forster 1778; von Humboldt and Bonpland 1807; Darwin 1859; Gaston 2003). When describing species' geographic ranges two concepts are considered. First is the fundamental niche, which consists of all abiotic conditions in which a species can survive at broad spatial scales (Hutchinson 1957). The fundamental niche is defined as locations that individuals of a species can reach by migration via dispersal, where scale is dependent on the dispersal mechanism involved. Second is the realized niche, where biotic conditions (e.g., mutualism, competition, predation, herbivory, etc.) determine microsite conditions and whether the site is habitable (Hutchinson 1957). The realized niche is the product of biotic, abiotic, and migration factors acting in combination and results in the realized niche always being a subset of the fundamental niche.

Although the size of species' geographic ranges varies widely, they share several properties that are predictable. Most geographic ranges tend to be small across many different taxa and the number of species with large geographic range sizes is relatively uncommon (Gaston 2003). Geographic range size also tends to be positively related to latitude (Rapoport 1982; Stevens 1989). This is likely caused by increased species richness, the subsequent increase in competition, and decrease in available space and habitat in the tropics (Gaston et al. 1998). In plants, some species that excel at dispersing long distances have large geographic range sizes (Edwards and Westoby 1996) but this is not a ubiquitous pattern. Species with large geographic ranges tend to be those that are more abundant (Brown 1984). Species with a larger population size or larger populations in the metacommunity often have higher birth rates and lower mortality rates, and therefore a lower probability of local extinction. Moreover, larger population size tends to equate to higher genetic diversity within a species (Karron 1987). This allows for species with large geographic range sizes to have a higher phenotypic plasticity, an indicator of the ability to adapt to changing environmental conditions (Bradshaw 1965).

Abiotic variables within a species’ geographic range are distributed along a gradient, and therefore species exhibit different environmental optima and respond to environmental conditions based on preferences for those abiotic variables (Whittaker 1967). Whittaker (1967) proposed the species response curve to model how response to an environmental gradient is normally distributed around an optimized value where abundance is highest and decreases as that variable increases or decreases away from its optima (Figure 1.1). Near the tails of the species response curve are stressful conditions in which few individuals of a species can survive. Outside the species response curve represents species absence due to intolerance of the environmental conditions at that portion of the gradient.

A species usually does not exist monotypically, however, but rather within a group or community of other species. In this community, a species will interact with other species both positively and negatively. These interactions alter the shape of the species response curve. For example, if a species is out-competed by another species in environmental conditions under which it may otherwise exist, the species response curve could become skewed (Gauch and Whittaker 1972). A set of species responses curves for a given community are called coenoclines, the shape of which depend on the species richness of a region (Gauch and Whittaker 1972). Regions with higher species richness will have a greater density of coenoclines across a given gradient. In species diverse regions such as tropical forests, the modes of coenoclines tend to be shorter (individual species have lower relative abundance) and the dispersion is narrower (species turnover is high). In contrast, coenoclines tend to have higher modes and broader dispersion in species poor regions (Gauch and Whittaker 1972).

Like species, communities can be measured across environmental gradients. Measured characteristics of ecological communities are species richness (the number of species at a location), evenness (the variability in abundance of species at a location), and diversity (an index which accounts for species richness and evenness) (Bell 2003). Diversity can also be calculated based on the scale of observation as alpha (the number of species in a sample or at a site; i.e., species richness), beta (the amount of change in species composition between samples or communities), and gamma (the number of species in the study area) diversity (Whittaker 1960). Alpha and gamma diversity are similar in that they are descriptive; however, beta is not. A higher beta diversity score indicates low similarity in species composition between multiple sites.

If all species are ranked by a metric describing their commonness (i.e., abundance, density, importance, etc.) and plotted with their rank on the x-axis against their commonness value on the y-axis, the result is a dominance-diversity curve (Whittaker 1965). Since evenness describes the distribution of abundances in a community, it alters the shape of the dominance-diversity curve. Communities with low evenness (those with species of relatively unequal abundance) produce steeper dominance-diversity curves than even communities. In Oklahoma, upland tree communities are dominated by two *Quercus* L. species, while bottomland tree communities are more even (Figure 1.2). Non-monotypic communities are composed of relatively common and rare species; however, their proportions are not equal. Often there are a far greater number of rare species than common species in a community (Preston 1948). In order to persist, rare species must possess some ecological advantage over common species, such as higher growth rates, higher reproductive rates, or lower mortality rates (Connell et al. 1984). Rarity can manifest itself in several ways: a species can be locally rare but common in the region, regionally rare but locally abundant under optimal conditions, or both locally and regionally rare (Schoener 1987).

Species presence and the co-existence of similar species at a location is likely to occur based on tradeoffs between competitive ability for resources, dispersal ability, and survival capability (Tilman 1994). Climatic factors are often shown to be the dominant factor driving the presence of a species at a location (O'Brien 1993; Cottenie 2005; García-Valdés et al. 2015; Jones et al. 2016; Zellweger et al. 2016). Although species sorting appears to be the dominant mechanism in community assembly in many studies, mass effects also provide a large contribution to community composition (Cottenie 2005). Mass effects are the colonization of a site(s) by a species that cannot persist without continued dispersal inputs to those sites (Schmida and Wilson 1985). Additionally, edaphic variables, particularly phosphorous content and pH, have also been demonstrated to drive plant species turnover at small scales (Jones et al. 2016; Zellweger et al. 2016).

Habitat patches across a landscape vary in quality, and therefore species presence is affected by the ability of propagules to reach sites in which they are able to survive to maturity (Hanski 1998; Thomas et al. 2001). Individuals can reach a site in which abiotic conditions are suitable through dispersal. Plant species can be dispersed in a variety of ways including: gravity, by animals (zoochory), by water (hydrochory), and by wind (anemochory). In the forest data analyzed in this dissertation, there are two distinct communities (bottomland and upland). The species within these two communities disperse quite differently. In the upland community, most species are dispersed through some combination of gravity and zoochory – mostly rodents. Therefore, dispersal occurs within a relatively short distance from the parent tree (Nathan and Muller-Landau 2000). In the bottomland community, the majority of species are dispersed by some combination of wind, water, and zoochory – mostly via birds (Burns and Honkala 1990). Dispersal of these bottomland species occurs relatively farther from the parent tree (Clark et al. 1999; Nathan et al. 2008).

Given that the geographic range of species is regulated by biotic and abiotic factors operating at various scales, the biotic-abiotic-migration (BAM) model was developed to provide a framework for the study of these general factors (Peterson et al. 2011). The BAM model suggests that there is an abiotic space in which a species can survive that is constrained by a species' ability to reach a site (the fundamental niche) and persist there in a community of organisms (the realized niche). Modeling species' geographic ranges attempts to relate species presence in geographic space with species presence in environmental space. To do this, data for species presence (and sometimes absence) and environmental conditions must be obtained. Abiotic data are often relatively easy to find and more straightforward to model than migration or biotic data. Migration is more difficult to model since dispersal rates (the distribution of transport distances for a species) are often unknown and requires the modeling effort to incorporate time. Biotic interactions are not usually modeled spatially since data about how intra and interspecific interactions change throughout species' geographic ranges has typically been studied at small scales through the use of common garden experiments, and thus usually does not exist at the scale required for distribution modeling. Therefore, modeled niches are most often the fundamental niche, not the realized niche.

Scale is a crucial factor in any modeling exercise. Despite the lack of research incorporating biotic interactions into a distribution modeling framework, abiotic variables may limit species' geographic ranges at large scales, while biotic interactions are, to a greater extent, a local phenomenon (Pearson and Dawson 2003). However, recent research is finding biotic interactions to be more pervasive at larger scales (Meier et al. 2010). The resolution of a model determines which variables to use and, conversely, is usually constrained by which variables are available and at what resolutions. Scale should also vary depending on what is being modeled. A restricted species should be modeled at a fine spatial resolution, a widespread species at a medium resolution, and communities at a coarse resolution. In other words, geographic range size and resolution should be correlated (Peterson et al. 2011).

Research Questions

Species’ geographic ranges are determined, in part, by suitable environmental conditions, the ability to reach sites possessing those environmental conditions, and the capability to survive interactions with other species until reproductive age (Peterson et al. 2011). Species’ geographic ranges and community composition have been related to environmental conditions frequently, to dispersal limitation infrequently, and rarely to biotic interactions (Pearson and Dawson 2003). To further elucidate how environmental conditions, dispersal, and biotic interactions explain community composition at a location and thus, relates to species’ geographic ranges, this dissertation has three foci: 1) community composition, 2) beta diversity, and 3) species interactions.

Communities gradually turnover between sites since species respond individualistically to environmental gradients (Gauch and Whittaker 1972; Collins et al. 1993; Hoagland and Collins 1997). The intensity of turnover between communities in physical space is positively correlated to beta diversity over that space (Anderson et al. 2011). Community composition is also affected by species interactions (Bertness and Callaway 1994; Bruno et al. 2003; Peterson et al. 2011). Interspecific interactions are known to work at fine scales by providing water or shade for another plant (Bonanomi et al. 2011), but whether these interactions have consequences at larger scales affecting species’ geographic ranges is still unclear. The work I present in this dissertation utilizes spatially explicit analysis to further elucidate the effect of geography on community composition, beta diversity, interspecific interactions, and the intersections of them on species’ geographic ranges.

Bottomland Forest Structure

The second chapter analyzes how bottomland forests are structured geographically and across a precipitation gradient in the central United States. Forest communities are assembled by a variety of mechanisms that have been studied for decades (Whittaker 1962). There are four prevailing theories in ecology as to how communities are spatially structured: 1) patch dynamics, 2) species sorting, 3) mass effects, and 4) neutral theory (Leibold et al. 2004). Patch dynamics states that species within patches become extinct over time and are replaced through dispersal of conspecifics or heterospecifics (Hanski 1998). Species sorting is a reflection of environmental preferences between species (Whittaker 1967). Mass effects are the colonization of sites by a species in which they cannot persist without continued dispersal to those sites (Schmida and Wilson 1985). Finally, neutral theory states that environmental niches and dispersal ability is equivalent between species (Hubbell 2001).

Quantifying the contribution of these mechanisms (species sorting, mass effects, patch dynamics, and neutral theory) across gradients and scales is a relatively new endeavor due to the increase in computing power and development of multivariate methodology (Borcard et al. 1992). A recent meta-analysis suggests that local environmental factors may explain community composition to a greater extent than regional spatial factors concluding that approximately half of variation in community composition can be explained by spatial and environmental factors, with environmental (22%) factors explaining more than spatial (16%) factors, and from correlations between spatial and environmental factors (10%) (Cottenie 2005). However, spatial factors likely affect assemblage to a greater extent at larger scales than smaller scales (Cottenie 2005). Being that this is a regional scale study, results from this chapter will add to current knowledge about how spatial mechanisms affect species assemblage and to what extent. I address two research questions in Chapter II:

R1. Are bottomland forest species in the Johnson dataset aggregated to a greater extent in physical or environmental space?

Currently there is debate about the amount of control spatial factors have on community assemblage, which is scale dependent. This question aims to determine the extent to which spatial mechanisms such as dispersal limitation and mass effects drive community composition across the region by aggregating bottomland forest sites based on their species composition. If groups of sites are strictly clustered in geographic space without overlap, this would indicate that spatial mechanisms such as dispersal limitation and mass effects are primarily driving community composition.

R2. Is variance in community composition better explained by environmental (species sorting) or spatial (mass effects) mechanisms?

This question analyzes the varying contributions of environmental and spatial mechanisms to determine their relative effects on community composition. Theoretically, this involves analyzing how sites are grouped based on their species composition and determining how these groups are distributed in space. If groups are more clustered in environmental space than in physical space it can be considered that environmental conditions, and therefore species sorting, are the dominant driver of species assemblage. Current research suggests this to true (O’Brien 1993; Cottenie 2005; García-Valdés et al. 2015; Jones et al. 2016; Zellweger et al. 2016). Climate has been found to be the dominant driver, space the secondary driver, and soils the tertiary driver of community composition in several studies (Cottenie 2005; Soininen 2016; Arellano et al. 2016).

Beta Diversity and Stress

The third chapter analyzes the effect of heterogeneity of stress (measured as the variation of distances to species' range margins) and net interspecific interactions on beta diversity between sites. Heterogeneity drives beta diversity between sites in a variety of ways. For example, beta diversity exhibits spatial autocorrelation and has been demonstrated to decrease with distance between sites (Condit et al. 2002; Soininen et al. 2007) and elevation (Bryant et al. 2008). In addition, beta diversity and species richness increase with increasing habitat heterogeneity (Nguyen and Gómez-Zurita 2016).

Therefore, it may be reasonable to posit that as variance of stress increases (e.g., some species in the community are near their optima while others are near their range margins) similarity of species between sites decreases and therefore beta diversity increases. I address two research questions in Chapter III:

R1. Does variation in stress affect beta diversity between sites?

This question attempts to discover whether sites with similar environmental conditions have lower beta diversity than sites with different environmental conditions. Here I use proximity to a species' range margin as a surrogate for stress based on the assumption that individuals within a species are more stressed at the range margin (Gaston 2003). Beta diversity increases with increasing distance between sites (Condit et al. 2002) which could be caused by dispersal limitation and environmental (and therefore habitat) change to varying degrees. Deciphering the relative contributions of mass effects and species sorting on beta diversity would add to the body of literature concerning beta diversity and its drivers.

R2. Does beta diversity change with net interspecific interaction?

This question addresses whether or not interspecific interactions effect beta diversity. In theory, facilitation (where one species benefits from the presence of another species) should allow a species to exist in locations it would not otherwise (Bruno et al. 2003). Facilitation should, therefore, increase species richness and beta diversity relative to what they would have been without facilitative interactions if the species being facilitated does not already exist at the group of sites being compared. Likewise, if a common species is out-competed and replaced by a species not already present at the group of sites, then beta diversity increases. However, if a rare species is out-competed to the point of extirpation from the site and replaced by a common species already present at the group of sites, beta diversity will decrease. Determining the effect of interspecific interactions on beta diversity will elucidate whether competitive and facilitative interactions primarily affect common or rare species. Analyzing these relationships will provide novel insights on whether beta diversity is affected by interspecific interactions at the regional scale.

Stress and Species Interactions

The fourth chapter analyzes how interspecific interactions change with increasing stress, utilizing the novel stress gradient of proximity to species' range margins and whether this change differs from random expectations. The Stress Gradient Hypothesis states that interactions between organisms shift from competitive to facilitative with increasing stress (Bertness and Callaway 1994). Bertness and Callaway (1994) proposed two mechanisms for these positive interactions. First, plants can gain associational benefits from other species of plants that can ameliorate consumer pressure. For instance, an edible plant may benefit from reduced consumer pressure by growing in association with inedible plants. Second, plants can ameliorate environmental stress. Presence of neighbors can reduce light stress, increase soil moisture, and provide shelter from wind disturbance (Bonanomi et al. 2011).

Chapter IV focuses on the relatively understudied effect of biotic interactions on species' geographic ranges and community composition utilizing spatial analysis as a proxy for analysis from common garden experiments. I, therefore, address five questions in Chapter IV:

R1. Do interspecific interactions change with proximity to a species' range margin?

This question focuses on the directionality of interspecific interactions with increasing stress, using distance from a species' range margin as a surrogate for stress. It analyzes whether the importance of a focal species increases or decreases in the presence of another species with decreasing distance to the focal species' range margin. Current theory suggests that as stress increases, facilitative interactions become more prevalent (Bertness and Callaway 1994). Since stress is different for different species, the goals of this research question are to determine 1) whether distance to a species' range margin can be used as a more reliable measure of species stress and if so, 2) whether interactions become more facilitative near range margins. Both of which would provide novel insights into how stress affects interspecific interactions.

R2. What is the magnitude of interaction change with decreasing distance to a species' range margin compared to the random expectation?

This research question serves as a null hypothesis from which to analyze the results of R1 of this chapter. Null models have been used to elucidate broad ecological patterns (Bell 2001) and to test divergence of community composition from random (Gotelli 2001). Compositional deviation from random may suggest that species assemblages are individualistic and not driven by ecological processes (e.g., niche differentiation) (Hubbell 2001). The methods employed to ascertain interaction intensity and direction in R1 use importance values from the underlying community data. In order to compare to the real community, a null community matrix is generated by randomly selecting species and their importance values while maintaining the species richness at each site to determine 1) if the trend of interaction change is similar between the real and null communities (i.e., do interactions in both communities become more facilitative with proximity to range margins), and 2) the difference in interaction change between the two (i.e., does the real community exhibit more competitive or facilitative interactions than the null community).

R3. What proportion of species exhibit monotonic or unimodal relationships between net interspecific interactions and proximity to range margin?

Recently, there has been some debate about net interspecific interactions in extreme stress conditions and some have suggested a reshaping of the curve proposed by Bertness and Callaway (1994) since some studies have found that the net interactions shift back to competitive in extreme stress conditions (Castanho et al. 2015; McIntire and Fajardo 2011), while others still find support for the original hypothesis (Dohn et al. 2013; Richardson et al. 2012). Holmgren and Scheffer (2010) posit three explanations for the unimodal curve along stress gradients. First, there may be a threshold at which the facilitation between two individuals is overruled by competition for a resource when that resource becomes scarce. Second, facilitation alone may not be enough to allow growth under extreme conditions. Third, because organisms in a community are adapted for conditions in that community, facilitative interactions may be more prevalent under conditions that appear less stressful than those that are not.

In a meta-analysis, Soliveres and Maestre (2014) discovered fewer facilitative interactions under higher relative stress measured as increasing elevation in alpine plant communities and larger range in annual temperature in dryland plant communities. However, several authors suggest that the disparate results when testing the SGH could be caused by only considering a few species and not entire communities as formulated in the original hypothesis (Soliveres and Maestre 2014), or by not analyzing the entire stress gradient (López et al. 2016). Therefore, I will quantify the number of species exhibiting monotonic or unimodal relationships between net interactions and proximity to range margins.

R4. Do interactions between species within a family or genus differ from interactions between species of different families or genera?

In the columnar cacti forest of the Tehuacán-Cuicatlán valley in southeastern Mexico, Valiente-Banuet and Verdú (2008) analyzed interspecific interactions along one km^2 transects between 102 woody species (761 species pairs). The mean phylogenetic distance between facilitative species pairs was significantly higher than random expectation, while mean phylogenetic distance between competing species was significantly lower. Within a given community, taxonomic differences between species differ based on the taxonomic scale studied. For instance, if focusing on the small taxonomic scale (i.e., a genus) species are often less related than expected while at a larger taxonomic scale (i.e., flowering plants) species are usually more phylogenetically similar than expected (Cavender-Bares et al. 2006). This is likely due to competition between closely related species which often co-exist at smaller spatial scales and have more functional similarity (Swenson et al. 2007). I therefore expect to find a greater intensity of competitive interactions between species within families and genera than between species in different families and genera with greater distance to range margins.

R5. Do interactions differ based on the commonness or rarity of a species?

I also posit that the commonness or rarity of a species being analyzed should affect the types of interactions they experience since species presence in stressful areas is thought to be determined by a tradeoff between stress tolerance and competitive ability where more broadly tolerant species are affected to greater extent by competition (Grime 1979; Liancourt et al. 2005). Due to this tradeoff, I expect to find that dominant species are subjected to a greater intensity of competitive interactions than rare species. Conversely, rare species should be facilitated more often than dominant species since rare species tolerate stressful conditions where facilitative interactions are more common.

Research Goals

The work I present in this dissertation utilizes spatially explicit analysis to further elucidate the effect of geography on community composition, beta diversity, interspecific interactions, and the intersections of them on species' geographic ranges. Therefore, the main purpose of this dissertation is to discern the importance of drivers on community assembly and how these drivers are altered across spatial, environmental, and stress gradients. Stress gradients are species specific and species' geographic ranges are partly the manifestation of these stress gradients. Viewing 'stress' through this framework, stress for a species is likely higher beyond, at, and just within its range margin. The use of spatial analysis provides a novel perspective to tests of the SGH, the assemblage of communities, and modeling assemblages spatially.

Modeling species assemblages spatially has proven a difficult endeavor; however, individual species can be modeled for likelihood of occurrence in multidimensional environmental space, and subsequently these models can be overlaid to generate a list of species that are likely to occur at a site, termed stacked species distribution models (S-SDMs, Dubuis et al. 2011). Modeling the geographic range of a single species minimally requires environmental data and occurrence data (Peterson et al. 2011). However, methodology for incorporating dispersal limitations through use of global and regional species pools and biotic interactions through co-occurrence matrices has been recently developed (Guisan and Rahbek 2011). Although methodology for including these factors represents a step towards modeling geographic ranges more accurately, there is room for improvement. For example, modeling biotic interactions through co-occurrence (i.e., assuming two species competitively exclude each other based on lack of co-occurrence) is imprecise since it assumes accurate and sufficient sampling across a study area. Incorporating the change in interspecific interactions based on the geographic position of an individual relative to the species' geographic range may provide a more reliable measure of biotic interactions over space.

Finally, results on the effect of interspecific interactions on species performance across their geographic ranges will inform broader discussion of the species niche. Facilitation produces a scenario in which a species' fundamental niche may extend beyond that of the realized niche if the beneficial interactions allows an individual to exist in environmental conditions outside of the realized niche (Bruno et al. 2003). Since the realized niche is projected into geographic space in order to predict species' geographic ranges (Peterson et al. 2011), species' range margins represent boundaries of the realized niche. However, measuring biotic interactions at range margins has not previously been attempted. In order to accurately predict species' geographic ranges, and subsequently species assemblages, explaining the effect of interspecific interactions on species occurrence at the margins of geographic ranges is a necessary start.

Tables and Figures

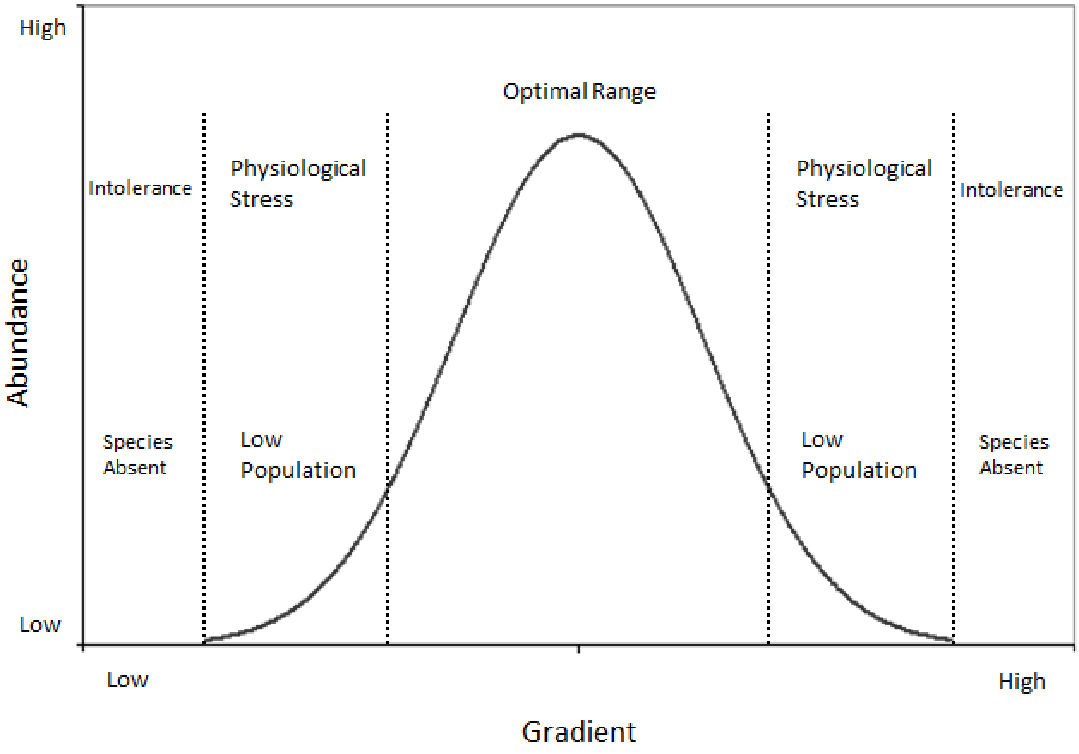


Figure 1.1. A theoretical species response curve to an environmental gradient

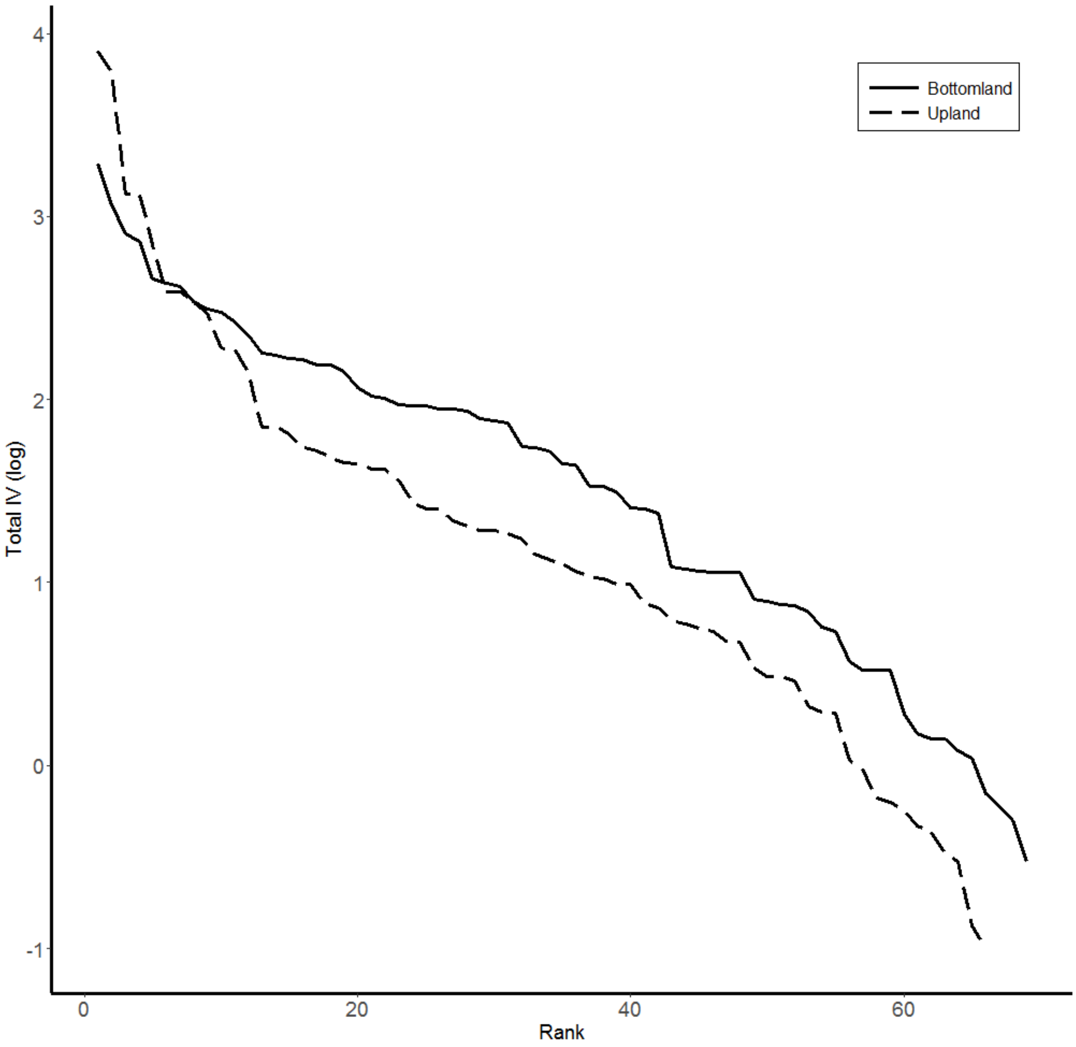


Figure 1.2. Dominance-diversity curves for bottomland (solid) and upland (dashed) communities

Chapter II

Bottomland Forest Structure

Introduction

Drivers of species composition and the subsequent assembly of a community at a site have been long studied in ecology. Whether communities respond to these drivers as a distinct unit (Clements 1916) or species are individualistic in their responses across sites (Gleason 1926) has been a subject of debate. The concept of a community as a unit is scale dependent, in that community assembly is thought to be governed by different processes at the landscape versus patch level. In addition to the effect of scale, spatial mechanisms driving community assembly have received a great deal of attention. As a result, there are four prevailing theories in ecology as to how communities are spatially structured: 1) patch dynamics, 2) species sorting, 3) mass effects, and 4) neutral theory (Leibold et al. 2004). In this chapter I focus on only species sorting and mass effects since both can be measured from a non-temporal field survey across sites.

Species sorting is a reflection of environmental preferences between species. Abiotic variables within a species' geographic range are distributed along a gradient, and therefore species respond to environmental conditions based on preferences for these variables (Whittaker 1967). The composition of a community changes due to differing environmental optima between species and the subsequent inter and intraspecific competition following dispersal and establishment at a site. Therefore, the response of an individual species to an environmental gradient becomes skewed when it is present in a community (Gauch and Whittaker 1972). Climatic factors are often shown to be the dominant factor driving species sorting and subsequently community assembly (O’Brien 1993; Cottenie 2005; García-Valdés et al. 2015; Jones et al. 2016; Zellweger et al. 2016). However, edaphic variables, particularly phosphorous content and pH, have also been demonstrated to drive species at small scales (Jones et al. 2016; Zellweger et al. 2016). O’Brien (1993) found that climate explained 78% of the variation in species richness of woody plants in southern Africa. Approximately half of the variation in the structure of a multitude of communities can be explained by spatial and environmental factors, with environmental (22%) factors explaining more than spatial (16%) factors (Cottenie 2005). Additionally, regardless of spatial scale, environmental and spatial variables explained 25% of community composition of plant life along an elevational gradient in Bolivia, with environmental variables explaining more than spatial variables on average (Arellano et al. 2016).

Although many studies have demonstrated species sorting to be the dominant mechanism in community assembly, others have demonstrated that mass effects are a primary contributor (Cottenie 2005). Mass effects are the colonization of sites by a species that cannot persist without continued dispersal to those sites (Schmida and Wilson 1985). For instance, Schmida and Wilson (1985) observed that 36% of desert plants that occurred along a Judean Ridge transect were present due solely to mass effects. Soininen (2016) conducted a meta-analysis of 158 datasets across many taxa (plants, herbivores, omnivores, carnivores, and decomposers) and found that spatial variables explained a mean of 11% of community variation. In some cases, dispersal explains the most variation in colonization success and extinction rates in patches (García-Valdés et al. 2015). Dispersal also affects alpha, beta, and gamma diversity.

Mass effects usually increase local species richness except for cases where highly competitive species (i.e., invasive species) are benefitting from high dispersal rates (Mouquet and Loreau 2003). Species richness is often highest if the species pool consists of intermediate dispersers, while beta and gamma diversity decrease with long distance dispersers which saturate the species pool and extirpate short dispersing species (Mouquet and Loreau 2003; Hubbell 2001).

Species sorting and mass effects likely work in tandem through a tradeoff between dispersal ability, and the ability of an individual to survive the environmental conditions upon arrival (Tilman 1994; Hassell et al. 1994). Linking species sorting and mass effects into a single framework describing change in community composition is difficult since they explain processes at differing spatial scales. At the landscape scale suitable environmental conditions, along with the differing quality of habitat patches and the proximity of resources, drive community assembly and beta diversity between sites (Dunning et al. 1992). Mass effects resulting from dispersal limitation is a regional process, where species sorting across environmental gradients is local (Ricklefs 1987; Mouquet and Loreau 2002). However, several studies have succeeded at linking species sorting and mass effects. For example, dispersal ability, environmental heterogeneity between sites, and differences between species preferences at the regional scale all affect co-occurrence between species and, therefore, species diversity (Mouquet and Loreau 2002). Additionally, species exist along environmental gradients based on their niche and competitive outcomes between similar species, and this can be affected by high dispersal rates through a heterogeneous landscape (Chase and Leibold 2003).

To analyze the effects of species sorting and mass effects on community composition, I juxtaposed environmental and spatial variables. Environmental variables are further classified into climatic and edaphic categories, while spatial variables are the two-dimensional distance between sites. I used environmental variables as a surrogate for habitat filtering, and thus species sorting, across sites. In the bottomland forests of Oklahoma most tree species are dispersed by some combination of wind, water, and zoochory – mostly via birds (Burns and Honkala 1990). Therefore, I assume that increasing distance between sites is akin to higher dispersal limitation and likewise an appropriate estimator of mass effects. Mass effects in bottomland forest communities would allow individuals to grow under otherwise unsuitable conditions based solely on proximity to conspecifics and continued dispersal to those unsuitable conditions.

Species sorting and mass effects rely on two assumptions: 1) there are fundamental environmental differences between sites, and 2) species do not respond similarly to those differences (Leibold et al. 2004). Therefore, the approach of this study is to determine the primary driver (species sorting or mass effects) of these fundamental differences between bottomland forest sites, species response to these differences along climatic gradients, and the effect of spatial distance on co-occurrence of species between sites. Bottomland forests occur along relatively linear water networks and therefore successful dispersal predominately occurs within these networks. Due to this, bottomland forests should be relatively more dispersal limited, and thus mass effects should be a larger driver of community composition than in previous studies elaborated on above. To analyze whether species sorting or mass effects are the primary driver of species assemblage, I addressed two research questions. 1) Are bottomland forest species in the Johnson dataset aggregated to a greater extent in physical or environmental space? If species are clustered into groups of sites which are more spatially aggregated than environmentally aggregated in ordination space, spatial drivers (mass effects) are more influential than environmental drivers (species sorting). 2) Is variance in community composition better explained by environmental (species sorting) or spatial (mass effects) mechanisms? Likewise, if the variance in community composition is better explained by distance between sites than by difference in climate, mass effects are more influential in driving composition than species sorting.

Study Area

Bottomland forests along various climatic and spatial gradients were studied within the state of Oklahoma, USA. The minimum bounding box around sites encompasses an area just over 180,000 km^2. The topography of the state consists of rolling hills and plains with the exception of three mountain ranges: Ouachitas in the southeast (~750 m relief), Arbuckles in the south-central (~130 m relief), and Wichitas in the southwest (~300 m relief). These mountain ranges notwithstanding, surface geology becomes younger as you transition east to west across the study area. Pennsylvanian aged rocks are characteristic of the eastern half, Permian aged rocks are predominant in the western half, and Tertiary aged rocks are found at the surface in the far northwest portion of the study area (Johnson 2008). Eighty-four out of the 102 bottomland sites reside on Quaternary, two on Cretaceous, seven each on Permian and Pennsylvanian, and two on Devonian aged strata. Seventy-four sites are situated on alluvium, ten each on sand and shale, six on sandstone, and one each on conglomerate and limestone.

Floristically, the study area is situated at the conjunction of the temperate forests in the east and the Great Plains in the west (Figure 2.1a) and consists of ten ecoregions (Woods et al. 2005). In the southeast of the study area is the Ouachita Mountain ecoregion. Bottomland forests in the Ouachita Mountain ecoregion are dominated by *Ulmus americana* L. and *Carya cordiformis* (Wangenh.) K. Koch. The Cross Timbers ecoregion lies north-south across the center of the study area of rolling hills where forest land is interspersed with rangeland and grassland. *Ulmus americana* and *Celtis laevigata* Willd. are the dominant bottomland tree species in Cross Timbers forest land. The Central Great Plains ecoregion encompasses much of the western part of the study area and is characterized by grasslands with scattered forest stands dominated by *Ulmus americana*, *Populus deltoides* W. Bartram ex Marshall, and *Celtis laevigata*. Other ecoregions not described here are used for reference in this chapter. Mean annual precipitation is the dominant climatic gradient in this region and effects vegetation types longitudinally (Figure 2.1b). Mean annual precipitation in this region decreases from east (1480 mm) to west (520 mm) (Figure 2.1b) and mean annual temperature increases from north (14 deg. C) to south (18 deg. C) (Figure 2.1c).

Methods

Data

Johnson (1982) surveyed bottomland forest sites in Oklahoma between 1977-1982. Average distance between nearest sites is approximately twenty km, and are spatially randomly distributed according to a nearest neighbor analysis in ArcGIS. Field data were collected at forty randomly selected points per site using an augmented variable radius technique (Rice and Penfound 1955). A tree was recorded if the distance from a sample point was less than 33 times its diameter. Calculation of basal area is the number of trees selected multiplied by ten (i.e., four trees equals a basal area of forty ft^2/acre). Frequency and density were calculated by forty arm-length transects of approximately forty m^2 each totaling 1600 m^2 at each site. Density was converted to a unit of individuals per acre. This method was standard practice in Germany for expedited forest surveys and introduced to American foresters by Grosenbaugh (1952). Relative frequency, density, and basal area per species were averaged to generate importance values (IV) for each species at each site. The bottomland dataset consists of 69 species across 102 sites, which I complied into a site by species matrix. Finally, the location of each site was determined based on township, range, and section information recorded by Johnson (1982).

Soil series data was retrieved from the United States Department of Agriculture's (USDA) STATSGO product database (Soil Survey Staff 2017). Each bottomland site was attributed with pH, available water capacity, and percent of silt, sand, and clay. The proportion of silt, sand, and clay at a site were used as an estimator of soil texture. Monthly climate data (PRISM Climate Group 2017) for maximum, minimum, and mean temperature, as well as total precipitation were collected for the thirty years prior to the culmination of the bottomland survey (1953-1982) and utilized to create monthly climate normals, which were calculated with monthly data over the time period over a one km^2 grid. Precipitation data from PRISM contains approximately 5% error annually (between 3 and 6% monthly) in Oklahoma on average, with slightly larger error in western than in eastern Oklahoma (Daly et al. 2008). However, values are not systematically under or over-predicted and therefore, error is random.

Ordination and Clustering

The sites by species matrix was analyzed using Nonmetric Multidimensional Scaling (NMS) in PC-ORD 6 (McCune and Mefford 2011) to determine the dissimilarity between species assemblages. NMS utilizes rank-order data in an attempt to avoid the assumption that species response to environmental gradients is monotonic (Gauch 1982). Although constrained canonical correspondence analysis (CCA) was used for variance partitioning (see “Variance Partitioning” section) and provides an objective method for relating composition to environmental data, it assumes that species responses to environmental gradients are linear and that these environmental gradients are not correlated and is therefore relatively less effective at separating sites based on their composition (Gauch and Wentworth 1976). Due to assumptions CCA makes, I chose to use NMS to separate sites in ordination space and CCA to partition the variation in composition between environmental axes (Lemly and Cooper 2011). I removed all species that occurred at five or fewer sites (Table B1), as recommended by

Poos and Jackson (2012) and used the Sorenson (Bray-Curtis) distance method since it is more sensitive to heterogeneity in a species matrix, is not influenced as greatly by outliers as Euclidian methods, and is more robust to variation between sites than Jaccard methods (Beals 1984).

The data were then analyzed using Cluster Analysis in PC-ORD to group sites together based on their species assemblages (McCune and Mefford 2011). I used an agglomerative hierarchical clustering algorithm which creates a hierarchical group structure in which a threshold must be set to determine the final groups (Pielou 1984). I selected a threshold of fifty percent information retained after creating a scree plot from distance measures between successive grouping stages. I used the Euclidian distance method and Ward’s Method for linking groups which minimizes the total within-cluster variance when clustering groups. The NMS ordination separated sites according to their assemblages and the cluster analysis grouped sites together based on this separation in ordination space. Sites within the cluster groups were subsequently related to environmental variables and their spatial locations to test the first research question.

Variance Partitioning

In order to test the second research question, I analyzed the contributions of environmental and spatial variables to variance in the community composition using a redundancy analysis ordination (Borcard et al. 1992) in the R package ‘vegan’ (Oksanen et al. 2017). This method partitions the variance in a community matrix described by a set of variables using multiple CCA ordinations (Borcard et al. 1992). Canonical eigenvalues derived from models using a subset of variables are compared relative to the total eigenvalue score for all variables. This fraction is interpreted of the percentage of variation that can be explained by that subset of variables.

I categorized the explanatory variables into three subsets: climatic (C), edaphic (E), and spatial (S). From this categorization I tested five models for their contributions to variance in community composition: 1) climatic variables only [C|E+S] (median and standard deviation of maximum, minimum and mean temperature, and precipitation derived from the climate normal data), 2) edaphic variables only [E|C+S] (average water capacity, pH, and percent sand, silt and clay), 3) spatial variables only [S|C+E] (latitude and longitude), 4) all variables [C+E+S], and 5) unexplained variance [1 – (C+E+S)]. Models 1, 2, and 4 were reduced using a forward selection iteration process (‘ordiR2step’ function) using the adjusted r^2 as a goodness of fit measure. The forward selection process was initiated with the variable explaining the most variance in the full model. I permutated the community matrix 999 times to test whether the linear relationship between the community matrix and each model is significantly stronger than random community matrices using ANOVA (‘anova.cca’ function) in R. If the C or E model explains the most variance in community composition, species sorting is the primary driver of composition. In contrast, the S model explaining the most variance would be indicative of mass effects being the primary driver.

There are two constraints for the variance partitioning method. First, disentangling pure spatial effects from environmental effects cannot be fully accomplished at the regional scale particularly where a well-defined spatial-environmental gradient exists such as the east to west precipitation gradient in the study area (Borcard et al. 1992). Since these variables are not independent, the actual variance explained may be more than or less than what has been modeled. For instance, it is possible that dispersal limited species (such as species dispersed by gravity or species outside of the regional species pool) may be able to tolerate climatic conditions over a portion of the environmental gradient it has not yet dispersed to. If so, these species contribute to variance explained by environmental variables due to the change of environment over space rather than contributing to the variance explained by the spatial model as it should. For example, median precipitation is highly correlated with longitude in the study area and therefore can be used to explain the same amount of variance (compare Figure C1 and Figure C2). Second, variance explained increases with the number of variables modeled (Borcard et al. 1992). The forward selection iteration process I utilized will ameliorate this issue. However, more climate variables were used after the forward selection iteration process than either spatial or edaphic variables (three climate, two spatial, one edaphic).

Results

Ulmus americana is the most dominant bottomland species in the study area (Table B2). It occurs at the most sites (95%) and has the highest mean importance value (18.9) across sites. *Morus rubra* L. is the second most frequently occurring species (90% of sites), however is not as important as several species that do not occur as frequently (*Celtis laevigata* and *Populus deltoides*). *Celtis laevigata* is relatively more important (higher relative frequency, density, and basal area) than *Morus rubra* at sites where it is present. In contrast, *Morus rubra* and *Populus deltoides* have a similar relative number of stems (*Morus rubra* – 6.4%, *Populus deltoides* – 5.1%), but individuals of *Populus deltoides* have thicker stems, and thus higher relative basal area than *Morus rubra* (*Morus rubra* – 2.8%, *Populus deltoides* – 10.7%). Nineteen species occur at more than 25% of sites, and 46 species occur at more than five sites.

Local tree species richness of bottomland sites is greater in the eastern (22 species) than the western portion (six species) of the region (Figure 2.2a), correlating with increased mean annual precipitation. Bottomland sites have a median richness of thirteen species. Sites are generally more even in the south and east (0.8, where 1.0 is completely even), only becoming relatively uneven in the north-northwest part (0.66) of the region (Figure 2.2b).

The NMS ordination of the matrix (102 sites x 46 species) computed a three axis solution. Final stress was 15.5 after 142 iterations. The third NMS axis was not highly correlated with any of the explanatory variables, and is therefore excluded from further discussion here. Cluster groups were separated in ordination space and assigned group letters based on average annual precipitation where Group A is wettest and Group H is driest (Figure 2.3). Axis one was highly correlated with precipitation variables, particularly mean/median annual ($r^2 = 0.77$) and Sept.-April monthly precipitation, while axis two was most correlated with a mixture of precipitation (most notably the standard deviation of monthly precipitation, $r^2 = 0.22$) and summer to late fall monthly minimum temperatures (Table 2.1).

Preserving 50% of the information contained within the underlying community data, eight groups were delineated. Group A contained the highest species richness, Group B was least species diverse and had the lowest Shannon index value, Group D had the lowest Simpson index value, Group E was comprised of sites with the most even communities, and Group D consisted of the least even sites (Table 2.2). Median richness was highest in Group A (seventeen species) and lowest in Group G (ten species). Finally, the range of species richness at sites within groups was largest for Group A (sixteen; maximum richness = 23, minimum = seven) and smallest for Group G (five; maximum = thirteen, minimum = eight).

Within each group, the dominance of species was determined by calculating mean importance (including sites in which the species is absent) and relative frequency of site presence was calculated. Species with a mean importance value >5% within each group are shown in Table 2.3. Sites dominated by *Populus deltoides* and *Ulmus americana* (Group H, mean annual precipitation = 578 mm) and by *Sapindus saponaria* var. *drummondii* and *Celtis laevigata* (Group G, 608 mm) form spatially clustered groups of sites within the Central Great Plains ecoregion at the dry end of the precipitation gradient in western Oklahoma (Figure 2.4). Group A is dominated by *Quercus nigra* L. and is confined to the South Central Plains and Ouachita Mountains ecoregions at the wet end of the precipitation gradient (average annual precipitation = 955 mm) in southeast Oklahoma. In contrast, the other five groups are less spatially clustered and inhabit sites in the middle portion of the precipitation gradient (average annual precipitation between 733 and 846 mm). The overlap in directional standard distance (one standard deviation) within groups was higher (50% overlap, Figure 2.4) in physical space than in environmental space (17% overlap, Figure 2.3) indicating that groups are better clustered in environmental space.

There were 46 species present at five or more sites. Out of these, 28 species were present within more than half of the cluster groups while 18 species were present at less than half. Seven species were present in all eight cluster groups and therefore likely tolerate the range of the precipitation gradient: *Acer negundo* L., *Carya cordiformis*, *Carya illinoinensis* (Wangenh.) K. Koch, *Fraxinus pennsylvanica* Marsh., *Morus rubra*, *Quercus macrocarpa* Michx., and *Ulmus americana*. Two species only occurred within one cluster group: *Cephalanthus occidentalis* L. and *Celtis reticulata* Torr. Both of these species exist in Group H at the dry end of the precipitation gradient. The eastern range margin of *Celtis reticulata* occurs within the study area which likely explains it being confined to the dry, western cluster. In contrast, *Cephalanthus occidentalis* is an eastern United States species that has been found in all but two counties within the study area (Hoagland et al. 2012). It is only present at seven sites in the dataset however, and is locally uncommon at all of them (importance never greater than 3.6). The confinement of *Cephalanthus occidentalis* to Group H is peculiar since it is a common wetland species across the eastern United States (Snyder 1991). Therefore, the absence of *Cephalanthus occidentalis* at sites under wetter conditions in the bottomland dataset is likely not a genuine pattern.

The relationship of cluster groups to the precipitation gradient is illustrated in Figure 2.5. Mean annual precipitation (Figure 2.5a) primarily decreases from left to right across NMS axis one, while the standard deviation of monthly precipitation (Figure 2.5b) primarily decreases from bottom to top across NMS axis 2. Groups H and A are at the dry and wet end of the precipitation gradient, respectively. The remaining six groups occupy a portion of ordination space where precipitation changes steeply. These six groups are better separated in two-dimensional climate space. For example, Group D exists at sites that receive consistently wetter monthly conditions, while sites in Group G are drier and receive more variation in monthly rainfall (Figure 2.5c-d).

Approximately 71% of the variance in the community matrix was unexplained (29% is explained by the variables tested) (Table 2.4). Out of the three sets of variables tested, climatic variables explained the most variance (22%), followed by the spatial and edaphic variables, 9.9% and 9.8% respectively. ANOVA of all models returned significant results ($p < 0.001$) indicating that the linear relationships between each model and the community matrix are stronger than random expectation. The standard deviation of minimum temperature and precipitation, along with longitude, explained the most variance in the model with all variables (Figure C1). The standard deviation of minimum temperature and precipitation, along with median annual precipitation, explained the most variance in the model with only climatic variables (Figure C2). pH explained the most variance in the model with only edaphic variables (Figure C3). Latitude and longitude explained roughly equal variance in the spatial model (Figure C4).

Discussion

Here I addressed two research questions: 1) are bottomland forest species aggregated to a greater extent in physical or environmental space, and 2) is variance in community composition better explained by environmental (species sorting) or spatial (mass effects) mechanisms? I analyzed whether species sorting (environmental explanations for composition change) or mass effects (using the distance between sites as a measure of dispersal limitation) are the primary driver of species assemblage.

Bottomland forest communities are generally spatially clustered and transition longitudinally, responding to the east-west precipitation gradient. Mean annual precipitation, as well as monthly precipitation between September and April are most highly correlated with this transition. The spatial clustering of communities at the wettest and driest ends of the precipitation gradient are well defined suggesting that species that possess some competitive advantage at either end of the precipitation gradient tend to cluster together spatially. This indicates that species sorting at large scales and mass effects at regional scales drive species occurrence in these communities.

In contrast, communities in the middle of the precipitation gradient are more spatially heterogeneous. Species in those communities may therefore not be driven by large scale climatic gradients, but rather by microsite environmental conditions and other biotic factors such as interspecific competition. Communities in the middle of the precipitation gradient appear to be more clustered in environmental space than physical space, where the gradient in the standard deviation in monthly precipitation, in addition to mean annual precipitation, partly drives compositional change suggesting that species sorting is the primary driver of composition at the center of the precipitation gradient.

However, explaining this central portion of the gradient is difficult. Some of the unexplained variance is likely explained by a combination of species sorting, mass effects, patch dynamics, biotic interactions, and randomness. Species sorting may occur at the microsite scale, responding to small changes in the environment and/or topography. I was unable to test this due to the inadequate spatial resolution (one km^2) of PRISM climate data for the period between 1953-1982. Species also likely occur at sites in which their population is not sustainable (i.e., outside of the area in which reproduction > death). Without rigorous testing of what these tolerances are for each species from greenhouse or transplant studies, it is difficult to determine how prevalent this occurs for any individual species or within communities as a whole.

Additionally, there is likely to be more overlap in niche space at the center of the precipitation gradient than at either the dry or wet ends due to confluence of species that are drought tolerant, drought intolerant, and those which are broadly tolerant (Gauch and Whittaker 1972) and therefore, species occurrence at a given site is more contextually and historically dependent (i.e., dependent on the probability of a propagule reaching a site and surviving the environmental conditions/competitive pressures of the pre-existing community) at the center of the precipitation gradient.

Difference in climate, and thus species sorting, explains the most variation in community composition (22%), followed by spatial distance between sites (mass effects), and then difference in soils (both approximately 10%). These results corroborate the findings of the clustering analysis since cluster groups are generally better aggregated in environmental space than physical space. Moreover, environmental variables explaining more community variation than spatial variables agrees with previous research (Cottenie 2005; Soininen 2016; Arellano et al. 2016). Climate explained 22% of the variation in composition across 158 datasets of various taxa (Cottenie 2005), 11% in tropical forest woody species (Arellano et al. 2016), and 22% in the bottomland forest tree communities here. Spatial variables appear to consistently explain less variation than climatic variables at 10% in a meta-analysis of many different taxa (Cottenie 2005), 5% in tropical forests (Arellano et al. 2016), 11% in a meta-analysis of 322 studies using variance partitioning across many taxa (insects, birds, butterflies, fish, plankton, corals, bats, ants, vascular plants, bryophytes, bees, spiders, bacteria, and algae) (Soininen 2016), and 10% in the bottomland tree communities of Oklahoma.

Species sorting may, therefore, be the dominant driver of compositional change regardless of the focal community studied. Mass effects may drive compositional change at larger scales; however, it has been demonstrated not to deviate significantly from 11% explained variance when different dispersal types are considered (Soininen 2016). It is clear, however, that both mass effects and species sorting across the precipitation gradient explains community compositional change longitudinally when the entire gradient is considered. At the central portion of the gradient, the relationship between climate and space with compositional change is less clear which likely contributes to the relatively high amount of unexplained variance (71%).

These results provide additional evidence toward climate gradients, and thus species sorting, being the main driver of compositional change over regional scales (O’Brien 1993; Cottenie 2005; García-Valdés et al. 2015; Jones et al. 2016; Zellweger et al. 2016). In addition, distance between sites (a proxy for dispersal limitation, mass effects) and soil are also significant drivers. The unexplained variance in community composition is likely not all due to random variation in community assembly. Past land use, management (burning or cutting), stand age, or fragmentation of habitat may account for some of the unexplained variance. Additionally, biotic factors should explain part of it; however, research analyzing how interspecific interactions affect species’ geographic ranges, and how those interactions change within a species’ geographic range, is in its infancy (Jones and Gilbert 2016). Compositional change over space is the product of biotic and abiotic factors, which are not independent of each other. For instance, the Stress Gradient Hypothesis posits that interspecific interactions become more facilitative under stressful abiotic conditions and more competitive under relatively benign conditions (Bertness and Callaway 1994).

The effects of these interactions on occurrence at the site level are likely species-specific depending on the stress gradient that effects the plant. Other biotic factors (i.e., herbivory, soil mycorrhizal relationships, predation) are spatially heterogeneous across species' geographic ranges. Understanding how these factors alter species occurrence within its geographic range is essential for interpreting compositional change over space, particularly in the portion of gradients where many co-occurring species' niches overlap. Research concerning how biotic factors change through species' geographic ranges is only just beginning. Future work should attempt to incorporate this research into studies analyzing drivers of compositional change over space.

Tables and Figures

Axis 1			Axis 2		
Variable	ρ	r^2	Variable	ρ	r^2
Dec. Precip.	-0.88	0.78	STD of Precip.	-0.46	0.22
Jan. Precip.	-0.88	0.78	Range in Precip.	-0.46	0.21
Apr. Precip.	-0.88	0.77	July Min. Temp.	-0.24	0.06
Mean Precip.	-0.88	0.77	July Mean Temp.	-0.22	0.05
Feb. Precip.	-0.87	0.76	Jul. Precip.	0.20	0.04
Mar. Precip.	-0.87	0.75	Aug Precip.	0.19	0.04
Nov. Precip.	-0.86	0.74	Aug. Min. Temp.	-0.19	0.03
Sep. Precip.	-0.84	0.70	Oct. Min. Temp.	-0.18	0.03
Median Precip.	-0.82	0.68	Aug. Mean Temp.	-0.18	0.03
Oct. Precip.	-0.82	0.67	Nov. Precip.	0.16	0.02

Table 2.1. Highest correlates (Spearman's ρ) between environmental gradients and NMS ordination axes 1 and 2. Only correlates >0.8 for ordination axis 1 are shown

Group	# of Sites	Richness	Evenness	Shannon Index	Simpson Index	Mean Annual Precip	Range of Precip	STD of Precip
A	11	50	0.57	0.83	0.43	955	896	241
B	3	19	0.63	0.59	0.37	846	997	288
C	20	44	0.65	1.23	0.54	802	1036	284
D	7	29	0.50	0.69	0.36	748	1020	281
E	12	28	0.67	1.26	0.58	734	1021	297
F	18	38	0.56	1.11	0.49	733	1060	294
G	7	22	0.54	0.78	0.40	608	1108	303
H	24	36	0.60	1.19	0.51	578	1034	289

Table 2.2. Species richness, diversity, and precipitation measures for sites within each cluster group. Precipitation is measured in mm

Group	Species	Avg IV	RF (%)
A	*Quercus nigra*	12.3	81.8
	Fraxinus pennsylvanica	9.3	100.0
	Quercus lyrata	8.2	54.5
	Ulmus alata	7.9	54.5
	Quercus phellos	6.7	63.6
	Liquidambar styraciflua	6.5	45.5
	Celtis laevigata	5.4	81.8
	Ulmus americana	5.4	63.6
B	*Quercus palustris*	38.7	100.0
	Fraxinus pennsylvanica	16.7	100.0
	Ulmus americana	13.5	100.0
	Carya illinoinensis	8.8	100.0
C	*Ulmus americana*	18.7	100.0
	Carya illinoinensis	13.7	70.0
	Celtis laevigata	11.7	100.0
	Quercus shumardii	9.3	85.0
	Fraxinus pennsylvanica	5.0	90.0
D	*Fraxinus pennsylvanica*	32.8	100.0
	Ulmus americana	17.0	100.0
	Celtis laevigata	13.1	100.0
	Populus deltoides	8.3	100.0
	Acer negundo	6.5	85.7
E	*Celtis occidentalis*	25.6	100.0
	Ulmus americana	17.9	100.0
	Fraxinus pennsylvanica	15.4	100.0
F	*Ulmus americana*	30.3	100.0
	Celtis laevigata	24.4	100.0
	Acer negundo	7.8	83.3
	Fraxinus pennsylvanica	5.9	72.2
	Morus rubra	5.1	94.4
G	*Sapindus saponaria var. drummondii*	28.3	100.0
	Celtis laevigata	27.1	100.0
	Ulmus americana	14.7	100.0
	Quercus macrocarpa	10.0	85.7
H	*Populus deltoides*	24.6	95.8
	Ulmus americana	19.9	95.8
	Juniperus virginiana	6.5	58.3
	Morus rubra	6.5	87.5
	Celtis laevigata	5.7	58.5

Table 2.3. Common species in each of the clustering groups, their average important value, and relative frequency among all sites within that cluster group

Model	Var. Explained	Pr(>F)	Model Var.	Resid. Var.
C|E+S	22.06%	0.001	10.146	35.854
E|C+S	9.76%	0.001	4.489	41.511
S|C+E	9.93%	0.001	4.565	41.435
C+E+S	29.15%	0.001	13.407	32.593
1 – (C+E+S)	70.85%			

Table 2.4. Results from the variance partitioning modeling exercise showing the percent of variance explained, the p-value, model variance, and residual variance

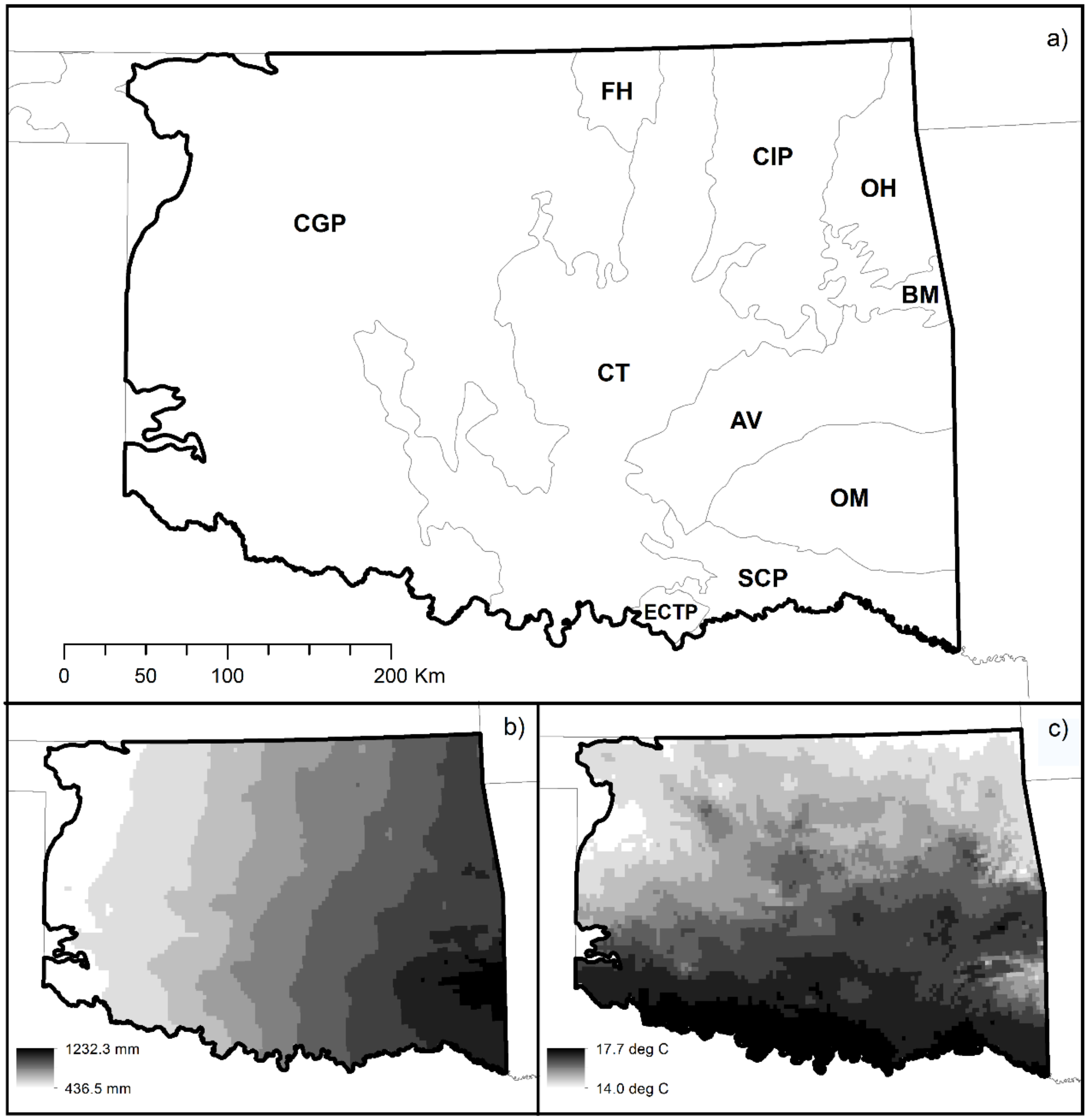


Figure 2.1. a) Ecoregions across the study area (Woods et al. 2005). CGP – Central Great Plains, FH – Flint Hills, CT – Cross Timbers, CIP – Central Irregular Plains, OH – Ozark Highlands, BM – Boston Mountains, AV – Arkansas Valley, OM – Ouachita Mountains, SCP – Southern Coastal Plain, ECTP – East Central Texas Plains. b) Annual precipitation gradient, dark colors indicate more precipitation and light colors indicate less precipitation. c) Mean annual temperature gradient, dark colors indicate warmer temperatures and light colors indicate cooler temperatures

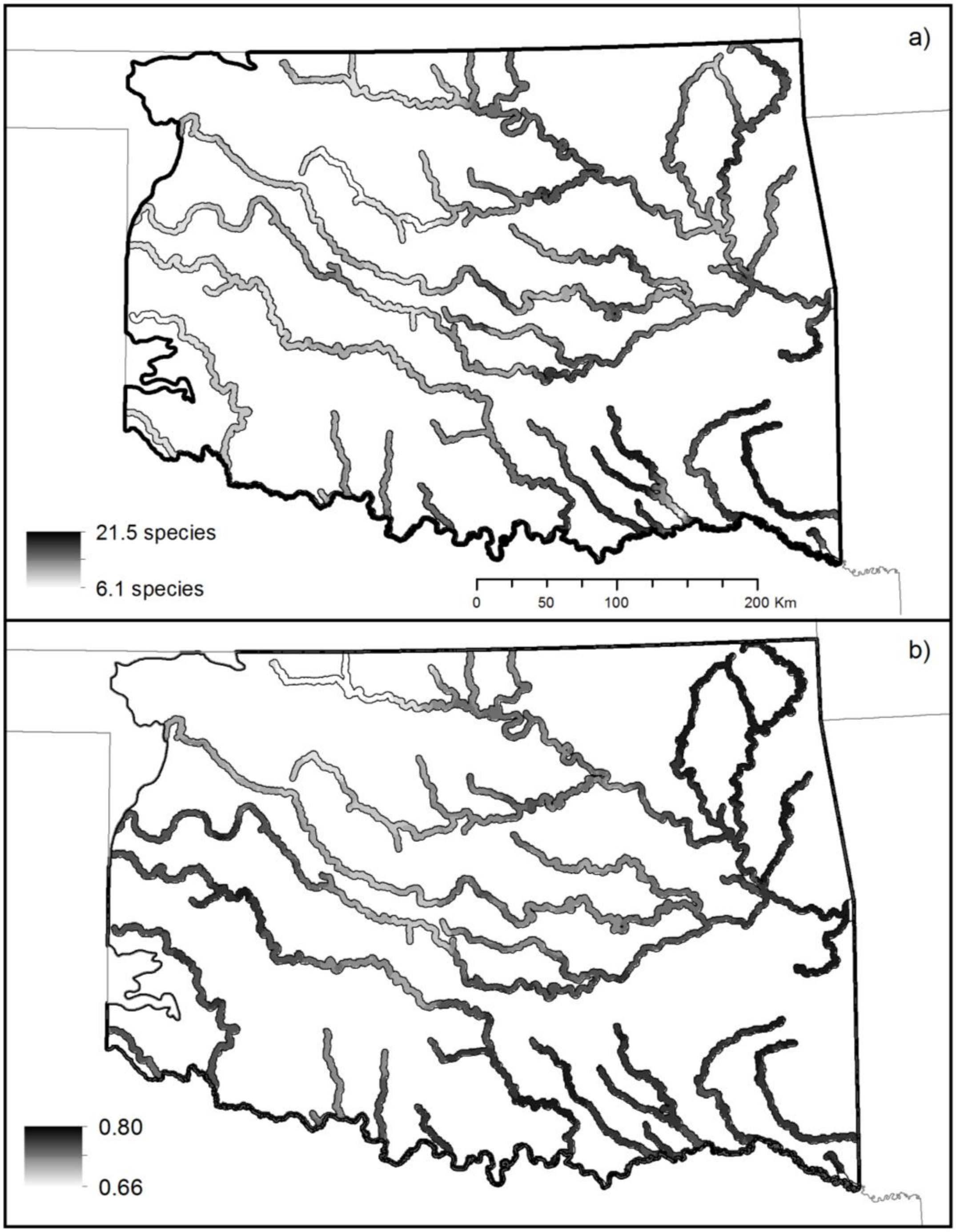


Figure 2.2. a) Species richness for bottomlands in the study area. Darker colors indicate greater richness. b) Evenness for bottomlands in the study area. Darker colors indicate greater evenness

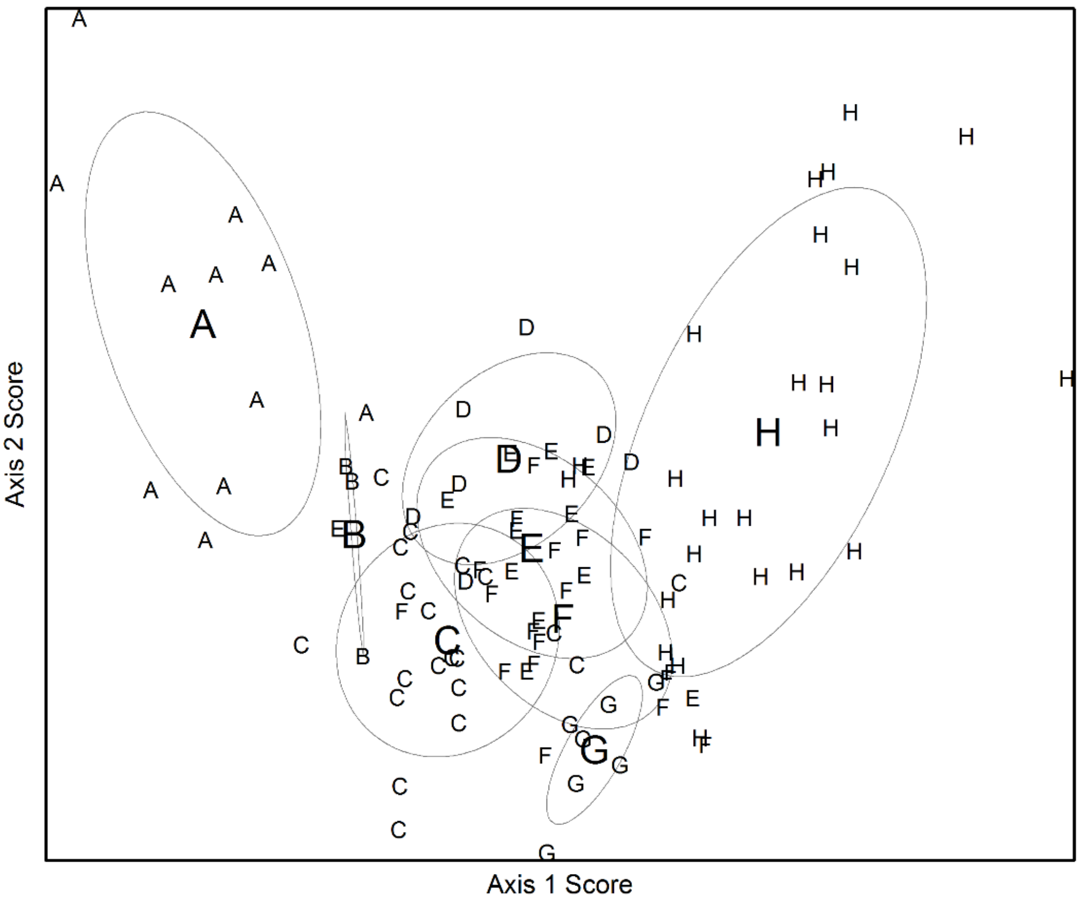


Figure 2.3. NMS ordination of bottomland sites. Letters refer to the cluster groups identified in Table 3. Axis 1 is negatively correlated with mean annual precipitation (r^2 = 0.77), while Axis 2 is negatively correlated with the standard deviation in monthly precipitation (r^2 = 0.22). Ellipses illustrate the directional standard distance (one standard deviation) between points within a group. Group centroids are indicated by larger letters

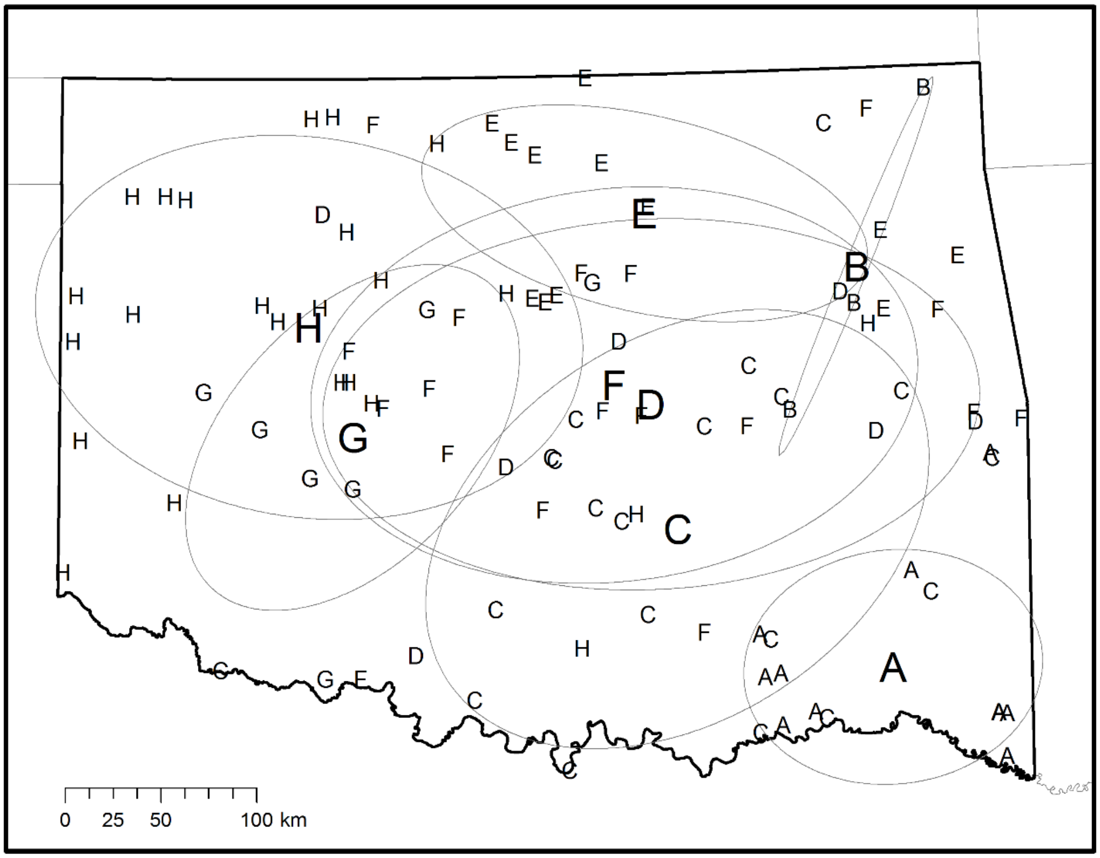


Figure 2.4. Geographic distribution of bottomland forest types. Letters refer to the cluster groups identified in Table 2.3. Ellipses illustrate the directional standard distance (one standard deviation) between points within a group. Group centroids are indicated by larger letters

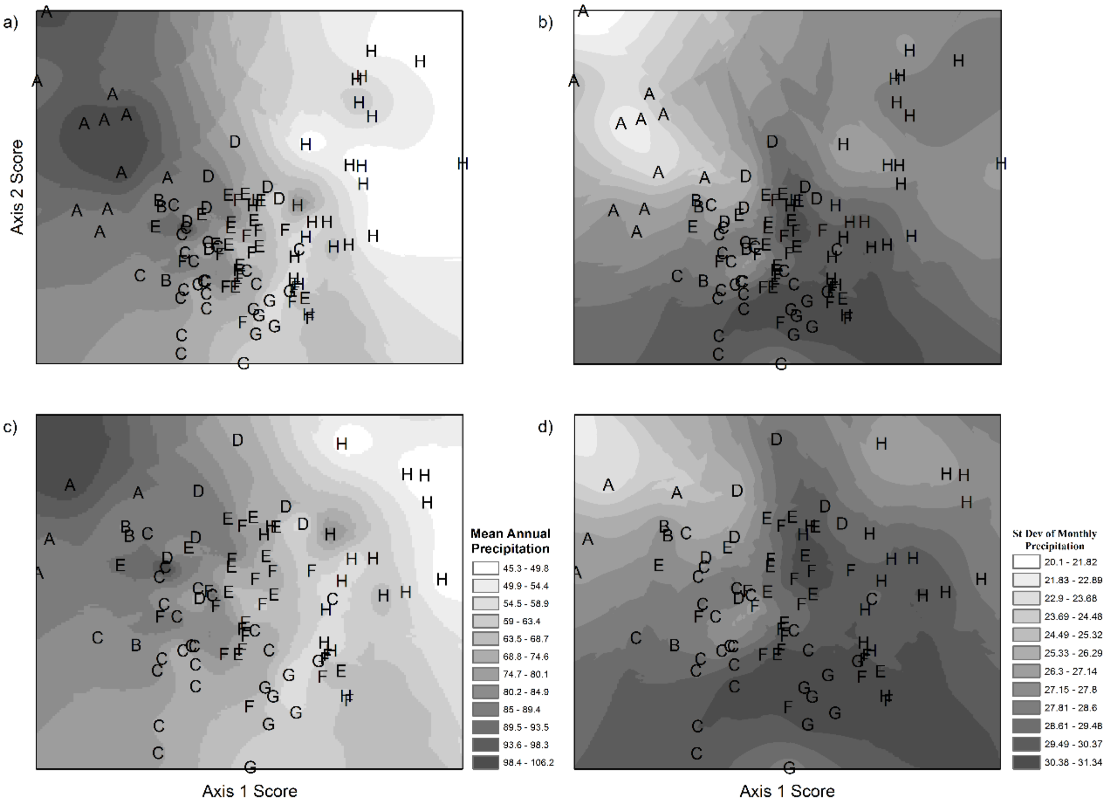


Figure 2.5. Plots showing mean annual precipitation (a), the standard deviation in monthly precipitation (b), mean annual precipitation of the six cluster groups at the middle of the precipitation gradient (c), and the standard deviation in monthly precipitation of the six cluster groups at the middle of the precipitation gradient (d) across NMS ordination space. Letters refer to the cluster groups identified in Table 3. In all panes darker colors indicate greater precipitation/range in precipitation

Chapter III

Beta Diversity and Stress

Introduction

Stress affects plant life in a variety of ways scaling from the physiological limits to growth and survival to the resilience of communities against disturbances (Buchanan 2000). Inadequate environmental conditions stress individuals, while lack of dispersal avenues to recolonize disturbed sites stress populations. Thus, stress has a prevalent effect on individuals and populations of a species. Due to this pervasiveness, ‘stress’ has been used to describe many different effects on organisms and hence, its usefulness has been debated (Grime 1989; Körner 2003). Körner (2003) argues that from the position of the plant, departure from optimal conditions is anything but stressful, it is normal, and therefore suggests that stress is an extreme form of resource limitation. Limitation or overabundance of resources used for photosynthesis or respiration relative to the temporal deviations from the normal conditions experienced by an individual stress individual plants (Körner 2003). Furthermore, species are individualistic in their response to resource gradients, and therefore stress (Gauch and Whittaker 1972). In light of this, researchers should consider stress in two facets: resource (e.g., water, nutrients, light) and non-resource (e.g., temperature, altitude, disturbance, etc.) based (Maestre et al. 2009).

Within a community, individuals compete or facilitate other individuals across gradients of stress that impact photosynthesis or respiration, such as water or sunlight. The Stress Gradient Hypothesis (SGH) states that interactions between organisms monotonically shift from competitive to facilitative with increasing stress, with net neutral (equal proportions of competitive and facilitative interactions) interactions at intermediate stress (Bertness and Callaway 1994). Bertness and Callaway (1994) defined stress as “potentially limiting harsh physical conditions” (such as low nutrient levels, soil moisture, heat, and disturbance). Bertness and Callaway (1994) proposed two mechanisms for positive interactions. First, plants can gain associational benefits from other species of plants that can ameliorate consumer pressure. For instance, an edible plant may benefit from reduced consumer pressure by growing in association with inedible plants. The second mechanism is the amelioration of environmental stress. Presence of neighbors can reduce light stress, increase soil moisture, and provide shelter from wind disturbance (Bonanomi et al. 2011).

Recently, there has been some debate about redefining the SGH proposed by Bertness and Callaway (1994) since many studies have found that the net interspecific interactions shift back to competitive in extreme stress conditions (Chu et al. 2008; Bowker et al. 2010; Graff and Aguiar 2011; McIntire and Fajardo 2011; Cavieres and Sierra-Almeida 2012; Castanho et al. 2015), while others still find support for the original hypothesis (Dohn et al. 2013; Richardson et al. 2012). Therefore, the relationship between interactions and stress may be unimodal instead of the monotonic relationship posited by Bertness and Callaway (1994). Holmgren and Scheffer (2010) give three possible explanations for the unimodal curve along stress gradients. First, there may be a threshold at which competition overcomes facilitation between two individuals when a resource becomes too scarce. Second, facilitation alone may not be enough to allow growth under extreme conditions. Third, because organisms in a community are adapted for conditions in that community, facilitation may dominate conditions that appear less stressful than conditions that are not. For instance, *Betula papyrifera* facilitates seedling growth of conifer species through mycorrhizal diversity in the temperate deciduous forest of southern British Columbia (Simard and Vyse 2006). Facilitation can also explain the hump-backed relationship between productivity and species diversity. Michalet et al. (2006) postulated that as conditions transition from benign to relatively stressful, species ameliorate conditions that stress-intolerant species would otherwise be unable to persist under and therefore increases species richness. Once environmental conditions become too severe, however, increasing stress on stress-intolerant species (and the species acting as the facilitator itself) outweighs any amelioration afforded by a facilitative species, and therefore species richness declines.

There are two types of interspecific interactions: direct and indirect. Direct interactions occur when one organism affects conditions for another, such as by providing shade. Indirect interactions occur when an organism alters the effects of another individual on a third organism, for instance through associational benefits (Tálamo et al. 2015). Associational benefits are scenarios in which an unpalatable species decreases consumer pressure on neighboring, palatable, species (Tahvanainen and Root 1972). Direct interactions are more prevalent in communities with a single well-defined gradient, while indirect interactions predominate in communities with multiple relatively equivalent gradients (Brooker et al. 2008). Thus, common garden experiments measure direct interactions by comparing performance of an individual between monocultures and in mixture with another individual while manipulating the stress gradient (see Maestre et al. 2005; Cavieres and Sierra-Almeida 2012). Co-occurrence or spatial aggregation studies analyze indirect interactions (see Graff and Aguiar 2011; Meier et al. 2011; López et al. 2016). This chapter assesses species interactions between sites using nearest neighbor methods at farther distances than direct interactions occur. Therefore, any reference to interactions are referring to indirect interspecific interactions.

Interspecific interactions are likely to affect individuals of a species differently at different parts of their geographic ranges since conditions for a given species are more suitable near the center of its geographic range and increasingly stressful with proximity to its range margin (Gaston 2003). For instance, abundance decreases nearer to a species' range margin (Brown 1995). Birth rates decrease and mortality rates increase near a species' range margin (Gaston 2003). Additionally, genetic variation of individuals decreases at range margins indicating that individuals of a species are less likely to adapt to changing conditions at or outside of their margins (Karron 1987). These factors emphasize that individuals at range margins experience more stress than their counterparts experience near the center of a species' geographic range.

In classical ecological theory, species have a fundamental niche that consists of all places that a species could occupy if that species could migrate there and if there were no biotic interactions (Hutchinson 1957). The realized niche is therefore always a subset of the fundamental niche since competitive interactions and lack of dispersal ability always constrict the fundamental niche. Facilitation provides a mechanism by which the realized niche can extend beyond the fundamental niche (Bruno et al. 2003) by allowing individuals to grow in areas otherwise uninhabitable. It also appears that facilitation is a ubiquitous driver of species richness, and potentially beta diversity, despite being less conspicuous than competition (McIntire and Fajardo 2014).

Beta diversity, the change in species composition across space, and the drivers of compositional change between sites is a focal theme in ecology (Bell 2001). The earliest formulation of beta diversity was a function of species richness at a single location and the regional species pool (Whittaker 1960). Since then, beta diversity has generally been defined as a measure of either turnover or variation in community data (Anderson et al. 2011). Gradient analysis uses beta diversity expressed as turnover, while beta diversity as a measure of variance is calculated when groups of sites are involved. Measures of community variance assess three components to derive beta diversity: species only at the focal site, species not at the focal site but at other sites being compared, and shared species between sites. In this chapter, beta diversity is measured as the variation of species occurrence in a community matrix between sites.

Several mechanisms drive beta diversity. For instance, beta diversity exhibits spatial autocorrelation and has been demonstrated to increase with distance between sites (Condit et al. 2002; Soininen et al. 2007) and elevational change (Bryant et al. 2008). Dispersal also effects beta diversity. Beta diversity is inversely related to the number of long distance dispersers which displace short dispersing species for available habitat (Mouquet and Loreau 2003; Hubbell 2001). From this, I assume a positive relationship between the diversity of dispersers (birds, mammals, wind, water, and gravity) and variation of tree species. Moreover, turnover, and therefore beta diversity, may have an optimum along productivity gradients much like species richness (Chalcraft et al. 2004) or may increase with increased productivity due to increased stochastic community assembly under productive conditions (Chase 2010). However, stochastic assembly is the dominant mechanism in plant communities, even when productivity does not drive beta diversity (Jonsson et al. 2016).

Due to different environmental optima of species in an assemblage, a group of sites with heterogeneous conditions typically have higher beta diversity than a group of sites with homogeneous conditions. Beta diversity and species richness increases with increasing habitat heterogeneity, for instance (Nguyen and Gómez-Zurita 2016). Likewise, if the length of an environmental gradient is long between a group of sites, a higher number of species are likely to be stressed by the variation in conditions. The higher variance of stressful conditions (e.g., some species in the community are near their optima while others are near their range margins) should therefore increase beta diversity due to increasing dissimilarity between sites. For example, deforestation of rainforest decreases beta diversity until a threshold where highly deforested sites increase beta diversity due to disturbance and light-insensitive species taking advantage of new habitat (Arroyo-Rodríguez et al. 2013). The variance in stress is high at this threshold since light-insensitive plants are stressed by the remaining undisturbed habitat within the patch and light-sensitive plants are stressed by the increase in patch edge.

Given that species respond to environmental gradients individualistically and are stressed by dissimilar gradients, I attempted to generalize ‘stress’ to all species by using proximity to a species’ range margin as a proxy for the stress gradient. I therefore addressed two research questions: 1) does variation in stress affect beta diversity between sites, and 2) does beta diversity change with net interspecific interaction? By using proximity to range margins to generalize stress across species, I am aiming to describe a mechanism that affects beta diversity independent of the focal system and therefore broadly applicable to ecological research. To my knowledge, this is the first study that has examined the quantitative relationship between the SGH and beta diversity.

Study Area

Bottomland and upland forests along various climatic and spatial gradients were studied within the state of Oklahoma, USA. The minimum bounding box around sites encompasses an area just over 180,000 km^2. The topography of the state consists of rolling hills and plains with the exception of three mountain ranges: Ouachitas in the southeast (~750 m relief), Arbuckles in the south-central (~130 m relief), and Wichitas in the southwest (~300 m relief). These mountain ranges notwithstanding, surface geology becomes younger as you transition east to west across the study area. Pennsylvanian aged rocks are characteristic of the eastern half, Permian aged rocks predominate the western half, and Tertiary aged rocks are found at the surface in the far northwest portion of the study area (Johnson 2008).

The study area is located in central United States within the state of Oklahoma at the conjunction of the temperate forests in the east and the Great Plains in the west (Figure 3.1a) and consists of ten ecoregions (Woods et al. 2005). The Ouachita Mountain ecoregion is in the southeast of the study area. The topography of the Ouachita Mountains consists of folded east-west ridges. *Quercus stellata* Wangenh. and *Pinus echinata* Mill. dominate the upland forests of the Ouachita Mountain ecoregion, while bottomland forests are dominated by *Ulmus americana* and *Carya cordiformis*. The Cross Timbers ecoregion lies north-south across the center of the study area and consists of rolling hills where forest land is interspersed with rangeland and grassland. *Quercus stellata* and *Quercus marilandica* Münchh. are the dominant tree species in Cross Timbers upland, as well as *Ulmus americana* and *Celtis laevigata* in bottomland, forests respectively. The Central Great Plains ecoregion encompasses much of the western part of the study area and is characterized by grasslands with scattered forestland. This scattered forestland consists of stands dominated by *Quercus marilandica*, *Quercus stellata* in the uplands, and *Ulmus americana* in the bottomlands. Other ecoregions not described here are used for reference in this chapter. Precipitation is the dominant climatic gradient in this region and effects vegetation types longitudinally. Mean annual precipitation in this region decreases from east (1480 mm) to west (520 mm) and mean annual temperature increases from north (14 deg. C) to south (18 deg. C).

Methods

Data

Two datasets for forest communities across a precipitation gradient were combined and utilized here; an upland forest dataset consisting of 70 species over 205 sites and collected between 1953-1957 by Rice and Penfound (1959) and a bottomland forest dataset consisting of 69 species across 102 sites between 1977-1982 (Johnson 1982), each organized in a site by species matrix. Mean distance between nearest sites is approximately twelve km, and are spatially randomly distributed according to a nearest neighbor analysis in ArcGIS. Both studies utilized the same data collection protocols introduced to American foresters by Grosenbaugh (1952) in which data were collected at forty randomly selected points per site using an augmented variable radius technique. A tree was recorded if the distance to the tree from a sample point was less than 33 times its diameter. Calculation of basal area is the number of trees selected multiplied by ten (i.e., four trees equals a basal area of forty ft^2/acre). Frequency and density were calculated by forty arm-length transects of approximately forty m^2 each totaling 1600 m^2 at each site. Density was converted to a unit of individuals per acre. Relative frequency, density, and basal area variables were averaged to generate importance values (IV) for each species at each site. The combined matrix consists of 91 species by 307 sites (see Table B3 for species list). Finally, each site had township, range, and section information that were used to georeference site locations.

I retrieved soil data from the United States Department of Agriculture's (USDA) STATSGO database (Soil Survey Staff 2017). I attributed each site with percent of silt, sand, and clay and used those proportions as an estimator for soil texture. Monthly climate data (PRISM Climate Group 2017) for maximum (tmax), minimum (tmin), and mean (tmean) temperature, as well as total precipitation (ppt) were collected for the thirty years prior to the culmination of the bottomland survey (1953-1982) and utilized to create monthly climate normals. Climate normals were calculated by averaging monthly data over the time period per grid cell. Precipitation data from PRISM contains approximately 5% error annually (between 3 and 6% monthly) in Oklahoma on average, with slightly larger error in western than in eastern Oklahoma (Daly et al. 2008). However, values are not systematically under or over-predicted and therefore, error is random.

In order to test whether increased variation of common dispersers increases the beta diversity of tree species, information on bird and mammal species occurrence was collected from the Breeding Bird Survey (Pardieck et al. 2016) and NatureServe (Patterson et al. 2007) respectively. Both bird and mammal datasets are qualitative (presence/absence only); however, I generated mammal beta diversity from geographic range shapefiles and bird beta diversity from 20 km^2 grid files. Finally, I utilized a 30-meter resolution digital elevation model (DEM) from the USGS National Elevation Dataset to generate slope.

Range margins were delineated by generating species' geographic ranges using MaxEnt (Elith et al. 2011). I downloaded species presence information from the Forest Inventory and Analysis (FIA) dataset (Bechtold and Patterson 2005) for species within the upland and bottomland datasets. FIA data were used to generate range margins to avoid analyzing the relationship between proximity to range margins and beta diversity from a single source, the community dataset. I then transformed the climate variables described above using Principal Components Analysis (PCA). I chose six climate axes as explanatory variables in the distribution models based on diminishing returns of variance explained by successive PCA axes using a scree plot. Species' geographic ranges for the entirety of the contiguous United States were modeled using ten subsample replicates withholding 10% of sample points for testing. Range margins were defined using the average likelihood value of the entire median MaxEnt result (following Cramer 2003). Distance to the range margin of a species was calculated as positive if the site falls within the geographic range and negative if the site is outside of the geographic range.

Distance Calculation

Pairwise spatial distance between sites was calculated as Euclidian distance between sets of site coordinates using the ‘pdist’ function in the SciPy Python library (Jones et al. 2014). The ‘pdist’ function calculates the multi-dimensional pairwise distance between all sites. In the case of spatial distance, this is two-dimensional (latitude and longitude). However, I computed distance three-dimensionally (percent silt, sand, and clay) to calculate distance in soil texture, and 48-dimensionally (monthly climate data for tmax, tmin, tmean, and ppt) to calculate climate distance between sites. As calculated distance in these factors increases, the dissimilarity between sites also increases.

Beta Diversity Calculation

Three metrics of beta diversity were calculated between sites: β_{rlb}, β_{sim}, and β_{w} (following Koleff et al. 2003), which describe different facets of beta diversity between sites. β_{rlb} is a measurement of continuity and loss in which a value of zero is returned if no species are shared and a value of one is achieved when all species are shared (i.e., species gained between sites is not accounted for). β_{w} is a measure of continuity that is directly influenced by the number of shared species between sites. β_{w} decreases with an increase in the proportion of shared species. β_{sim} is a measure of gain and loss since it considers shared species, species lost, and species gained between sites. β_{sim} and β_{w} are symmetric (remain unchanged if species occurrence at the focal site and compared site are transposed), and all measures are independent from changes in species richness. I calculated pairwise beta diversity between sites. For spatial representations, I calculated beta diversity for each site using the five nearest neighboring sites to the focal site.

Species Interactions

Knowing the relative intensity of the significant interactions between two species is useful to compare across studies and to determine changes in interactions across gradients. Markham and Chanway (1996) developed the "relative neighbor effect" (RNE) index to compare the performance of individuals grown in mixture and in monoculture (Eq. 1). The RNE index is symmetric around zero and produces intuitive scores (if RNE is between -1 and 0 the net interaction is considered facilitative, if RNE is between 0 and 1 the net interaction is considered competitive). The RNE is calculated as follows:

$$RNE = (P_{-N} - P_{+N})/\max(P_{-N}, P_{+N}) \qquad \text{Eq. 1}$$

where P_{-N} is the performance without a neighbor, and P_{+N} is performance with a neighbor.

Here I analyzed the difference in interaction intensity over space using a nearest neighbor analysis iteratively for each site. The Euclidean distance method selects the k nearest points to the focal point where distance is measured as a straight line between points. Calculating RNE requires that a set of nearest neighbors contain sites where both species are present and where the focal species is present but the other is not. Since species composition is spatially autocorrelated (Condit et al. 2002; Bryant et al. 2008), simple Euclidean distance will likely not select sites in both categories. Therefore, the algorithm selected the three closest sites where the two species co-occur and the three closest sites where only the focal species occurs. To address autocorrelation of species importance, the importance at the focal site was calculated using an inverse distance weighted technique assuming that as neighbors become further separated in space they become less influential with each other. To my knowledge, no model currently exists to access spatial change in the direction and intensity of interspecific interactions. I created this spatial model in Python programming language (Van Rossum 2007). The code is available in Code A1 in Appendix A.

Predictive Capability of Continuous Variables

I utilized several continuous explanatory variables (bird and mammal beta diversity, soil texture dissimilarity, distance, climate dissimilarity, difference in slope, net interaction difference, and variance in the proximity to range margins) organized in dissimilarity matrices to model contributions towards changes in tree beta diversity between sites. I performed the mantel tests using the ‘mantel’ function within the ‘vegan’ package (Oksanen et al. 2017) in R statistical software (R Core Team 2017). I chose mantel tests due to their ability to elucidate whether or not two dissimilarity matrices are related (Legendre et al. 2015). I calculated Spearman’s ρ over 10000 permutations for all eight explanatory variables. Spearman’s ρ assesses whether the relationship is monotonic, where positive ρ values indicate positive relationships. A general additive model (GAM) was utilized to fit relationships between beta diversity and the variance in proximity to range margins or net interaction. The GAM fits the model by smoothing partial residuals using a locally weighted linear least squares algorithm which estimates weights for data points based on proximity to the fitted mean (Hastie and Tibshirani 1990). No additional smoothing was performed on regression lines derived by the GAM.

Mantel tests can consolidate multiple variables into a single measure, allowing multivariate distance to be analyzed rather than simple difference, and it can be used with non-parametric data (Guillot and Rousset 2013). Despite these positives, there are several caveats. First, Euclidian geographic distance may not be suitable for a mantel test (Legendre et al. 2015). Second, due to how the mantel test permutes the distance matrices, it generates correlation coefficients that are under-dispersed relative to what would be expected using a null analysis and therefore produces significant results far more often than expected (Guillot and Rousset 2013). Third, mantel tests make two assumptions: 1) relationships between the two matrices is linear or monotonic, and 2) as distances in the first matrix increase distances in the second matrix increase. Therefore, the null hypothesis assessed by the mantel test is that there are no relationships between distances in the two matrices being compared (Legendre et al. 2015).

Results

The Cross Timbers is situated at the margin of many eastern and western tree species, and the density of tree species' range margins is greater in the Cross Timbers than anywhere else in the study area (Figure 3.1b). In the southern Cross Timbers, range margin density is highest at 5.1 km of margin per km^2. Range margin density is lowest in the southeast part of the Southern Coastal Plains at 1.8 km of margin per km^2.

Local tree species richness is greatest in the southeastern portion of the region, correlating with increased mean annual precipitation, reaching its maximum (22 species) in far southeast Oklahoma and its minimum (one species) across a large swath of the Central Great Plains in the west (Figure 3.1c). Evenness is greater along the entire eastern part (0.73-0.79, where a value of one is completely even) of the region and in three additional places such as the northern Cross Timbers (0.79), and in the southern (0.74) and western Central Great Plains (0.74). Sites only become uneven (0.29) in the northwest (Figure 3.1d). These uneven sites correspond to sites with low species richness dominated by a single species such as *Quercus marilandica* in the uplands.

Tree beta diversity is spatially heterogeneous (Figure 3.2). Areas of low diversity exist in several areas: in the southwest along the boundary between the Cross Timbers and Central Great Plains, in the southeast between the northern Ouachita Mountains and Arkansas Valley, and in the northeast straddling the Cross Timbers-Central Irregular Plains ecotone. Minimum tree beta diversity is 0.36 (where a value of zero and one represent all shared and no shared species between sites, respectively). In contrast, several areas of high diversity (up to 0.77) are present: in the Central Great Plains (0.71), at the northern boundary of the Cross Timbers and Central Great Plains (0.75), and at the confluences of the Cross Timbers-Southern Coastal Plain-East Central Texas Plains (0.74) and the Central Irregular Plains-Ozark Highlands-Boston Mountains (0.73) ecoregions.

Bird (Figure 3.3a) and mammal (Figure 3.3b) beta diversity share some spatial similarity with tree beta diversity; however, there are differences. Both birds and mammals have low beta diversities in the southwest Cross Timbers and in the Ouachita Mountains in the southeast. Local minima of beta for birds (0.03) and mammals (0.06) occur in the Ouachita Mountains. Both bird and mammal beta diversity is lower than tree beta diversity in the center of the study area. In addition, bird and mammal beta diversity is higher than tree beta diversity in the northwest part of the study area, where bird beta diversity reaches 0.16 and mammal beta diversity is 0.23. Bird and mammal beta diversity is also high in the northeast where the Cross Timbers and Central Irregular Plains ecoregions meet.

Climate distance (Figure 3.3c) and spatial distance between sites are positively correlated ($p < 0.001$, Figure 3.3d), and therefore share similar spatial patterns, except for in the central part of the study area. Climate distance in 48-dimensional space (48 axes consisting of monthly climate data for tmax, tmean, tmin, and ppt) between sites is three times higher in far southeast Oklahoma than at the ecotone between the Cross Timbers and Central Great Plains. Average distance between groups of five nearest neighbor sites is shortest in the Ouachita Mountains (~15 km) and farthest in northwest Oklahoma (~40 km). These differences also exist for difference in slope between sites (Figure 3.3e), where slope between sites is steepest in the Ozark Highlands (22 degrees) and most gentle in the southwest at the ecotone between the Cross Timbers and Central Great Plains (two degrees). Difference in soil texture between sites, however, shares most of its spatial pattern with tree beta diversity with the exception of the central part of the study area (Figure 3.3f). Soil texture distance in the center of the Central Great Plains and along the southern border of Oklahoma are four times higher than in the Ouachita Mountains.

Net interspecific interactions are, for the most part, slightly competitive or facilitative across the study area. (Figure 3.3g – the lightest color represents net facilitative interactions). The strongest net competitive interactions (RNE of 0.09) between groups of nearest neighbor sites exists in the west-central part of the study area where the Cross Timbers-Central Great Plains ecotone is. This area is occupied by sites with low species richness (< 4 species) dominated by *Quercus marilandica* in which several species (such as *Morus rubra*, *Juniperus virginiana Celtis reticulata*, *Celtis laevigata*, and *Sideroxylon lanuginosum* ssp. *lanuginosum*) are present at a single site with low importance. Difference in the distances to range margins (Figure 3.3h) shares spatial patterns with tree beta diversity with the exception of the central portion of the study area. Species at sites in the center of the study area have similar distances to their respective range margins (variance in proximity to range margins are ~3 km); however, tree beta diversity is high. Maximum distance between the variance in distance to species' range margins (~40 km) occur in the far southwest part of the study area. Variance in proximity to range margins is positively correlated with beta diversity, especially when variance in proximity to range margins is high (Figure 3.4).

I performed mantel tests using these predictor variables to explain tree beta diversity. The highest correlated variable with tree beta diversity was climate distance (Spearman's $\rho = 0.26$) (Table 3.1). Six other variables were significantly positively correlated with tree beta diversity: mammal beta diversity ($\rho = 0.25$), spatial distance ($\rho = 0.24$), bird beta diversity ($\rho = 0.23$), variance in proximity to range margins ($\rho = 0.18$), soil texture distance ($\rho = 0.14$), and slope difference ($\rho = 0.07$). Net interaction difference was the only variable not significantly correlated to tree beta diversity.

Discussion

I addressed two research questions: 1) does variation in stress affect beta diversity between sites, and 2) does beta diversity change with net interspecific interaction? Stress can manifest itself in a variety of ways. Inadequate environmental conditions stress individuals, while lack of dispersal avenues to recolonize disturbed sites stress populations (Buchanan 2000). Here I have shown that variation in environmental stress (climate distance and soil texture distance) in addition to variation of dispersal avenues (bird and mammal beta diversity, as well as dispersal via gravity, wind, and water estimated by spatial distance between sites) correlate positively with tree beta diversity at the regional scale. Moreover, tree beta diversity is positively correlated with variation of stress when measured as the variation in distances to species' respective range margins.

The spatial and taxonomic scale of the field survey partly causes the lack of correlation between the spatial variation in tree beta diversity and ecotone boundaries (Figure 3.2). For example, the growing area of winter wheat, not the change in tree species composition, delineates the boundary between the Cross Timbers and Central Great Plains (Omernik and Griffith 2014). Additionally, sites were sampled at a fine spatial scale and are disparate in space. Therefore, composition changes at small scales characterized by the gain and loss of rare species in species rich areas will remain undiscovered.

The areas of low beta diversity in the southeast (Ouachita Mountain-Arkansas Valley boundary, 0.42) and northeast (Cross Timbers-Central Irregular Plains boundary, 0.45) are caused by groups of sites with moderate species richness (~10 species) in which turnover is minimal where the only differences in species assemblages are from gained species (Figure D1). In contrast, a group of sites in which species assemblages are nearly identical, and therefore, no species are lost or gained (Figure D2) explains the southwest area (Cross Timbers-Central Great Plains boundary) of low beta diversity (0.41). All three areas of low beta diversity, however, are located where no bottomland sites were sampled. Therefore, the interpolation algorithm is only calculating beta diversity between upland sites in these areas. Although the spatial boundary between bottomland and upland forests is ill-defined, species richness and evenness is generally greater in the bottomland community (Figure 1.2). Since water is a limiting resource over the entirety of the study area, proximity to water near the surface is a primary driver of species assemblage change. For this reason, areas of high beta diversity are located in areas where there is a mixture of bottomland and upland sites with distinct composition and ecoregion boundaries.

In theory, variance in stress should increase beta diversity between sites (Arroyo-Rodríguez et al. 2013; Nguyen and Gómez-Zurita 2016). For example, imagine an extremely dry area (i.e., an area where water availability is the main stressor) with a river running through it. There would be a steep species compositional gradient as proximity to the river increases, and therefore sites that are not spatially far apart may have substantially different assemblages. The beta diversity and the variance in stress between these two theoretical sites is very high. For instance, species turnover among spatially proximal sites was high along a transect beginning near the coast of southeastern Australia and extending into an arid environment (Gibson et al. 2017). Here I attempted to expand this concept to regional scales by using variance in the distance to species' range margins as a surrogate for stress since no single gradient is a suitable stressor for all species. If two sites have a mixture of species that are near their range margins and species that are not, beta diversity is likely to be higher due to differences between species assemblages at those two sites.

Generally, the variance in the distances to species’ range margins between sites increases from east to west across the study area. The vast majority of species in these datasets are eastern deciduous species (88%), many of which have range margins in the center of the study area or further west. Therefore, in the east part of the study area, individuals of these eastern deciduous species are well away from their species’ range margin. In the center of the study area, many individuals are near or beyond their species’ range margins lowering the variance in proximity to range margins between sites. In the west some eastern species cannot survive, other eastern species are approaching their range margins, and many of the western species are well away from their eastern range margins. These factors increase the variance in proximity to range margins between sites. The notable exception to this gradient is in the southeast, where the northwestern range margins of gulf coastal tree species such as *Taxodium distichum* (L.) Rich. causes high variance in the proximity to range margins between sites.

I predicted that net interspecific interactions would affect beta diversity across regional scales by positing that net facilitative interactions would allow species to survive at sites they otherwise would not be able to (Bruno et al. 2003). The effect of net competitive interactions on beta diversity may be more difficult to predict since it should vary depending on the commonness of the species being out-competed or facilitated. If a common species is out-competed and replaced by a species not already present at the site, then beta diversity increases. While if a rare species is out-competed to the point of extirpation from the site and replaced by an individual of a common species, beta diversity will decrease.

The complicated nature of which species interspecific interactions affect to a greater extent explains why I found no trend in beta diversity with change in net interaction (Figure 3.5). However, at highly competitive sites, the variance in beta diversity decreases and is generally higher than beta diversities at facilitative or neutral sites (RNE near zero) suggesting that competitive interactions are more likely to act on common species, allowing rarer species to occupy this empty niche space. In addition, net facilitative interactions do not increase beta diversity and thus, benefit common tree species by increasing their importance rather than providing rare species new sites. Therefore, interspecific interactions, whether facilitative or competitive, have a greater effect on whichever species are common between sites. This may indicate that interactions are density dependent, where more individuals of a common species at a site increases the probability of interactions between other species.

Some have criticized the utilization of mantel tests for elucidating patterns between datasets (see Predictive Capability of Continuous Variables section) (Guillot and Rousset 2013; Legendre et al. 2015). However, it has been widely used in ecological studies. Despite these criticisms, the mantel test is useful under certain circumstances that apply in this chapter. For example, the mantel test is acceptable in cases where both matrices are dissimilarity matrices (i.e., they measure distances between sites in multivariate space) (Legendre et al. 2015). Three of the explanatory variables (spatial distance, slope difference, and net interaction difference) are not dissimilarity matrices. Therefore, results are dubious for those three variables and, in particular, significant results for spatial distance and slope difference are likely to be false positives.

To my knowledge, this is the first study that has examined the quantitative relationship between the SGH and beta diversity. Analyzing a spatial gradient as a proxy for stress has been previously attempted (Choler et al. 2001); however, not in two-dimensional space. A commonly cited problem with species distribution modeling (SDMs) is the lack of incorporating biotic variables into the modeling effort (Meier et al. 2010; Urban et al. 2012; Svenning et al. 2014). Theory describing how to model species' geographic range dynamics affected by biotic factors is beginning to be developed (Jones and Gilbert 2016). For example, competition between species affects lower (both latitudinal and elevational) range margins to a greater extent than upper range margins (Hargreaves et al. 2014).

Species composition is affected by interspecific interactions and variance in stress based on proximity to range margins at regional scales. I have demonstrated that interspecific interactions drive beta diversity through competitive and facilitative forces on common species. However, many SDMs project geographic ranges of a single species at one time and I did not differentiate the identities of individual species in this analysis. Therefore, the next step in this research should focus on how individual species are affected by interspecific interactions with proximity to range margins in order to incorporate these results into a distribution modeling framework. In the bottomland forest dataset, 71% of the variance in the community data was determined to be unexplained (Chapter II). However, biotic interactions were not analyzed and their effects on beta diversity likely contribute to species composition. It is clear from this analysis that biotic interactions have some effect at the regional scale. Studying these effects further will improve our understanding about how biotic interactions drive species composition locally and alter species’ geographic ranges regionally.

Tables and Figures

Variable	ρ	P
Climate distance	0.26	0.001
Mammal β_w	0.25	0.001
Spatial distance	0.24	0.001
Bird β_w	0.23	0.001
Range Margin variance	0.18	0.001
Soil Texture distance	0.14	0.001
Slope difference	0.07	0.002
Net Interaction difference	0.02	0.174

Table 3.1. Results of Mantel tests on individual variables

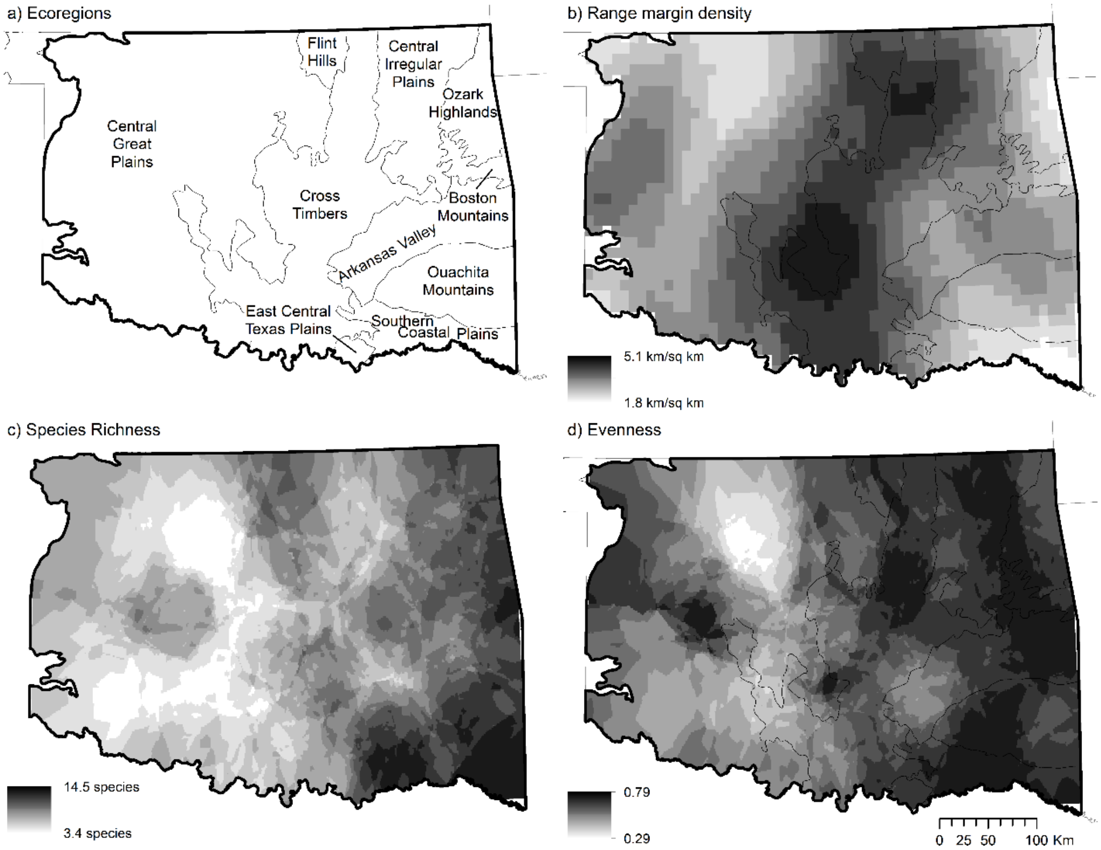


Figure 3.1. a) Ecoregions across the study area. b) Mean density of range margins. c) Species richness for sites in the study area. d) Evenness for sites in the study area. In b, c, and d darker colors indicate larger values

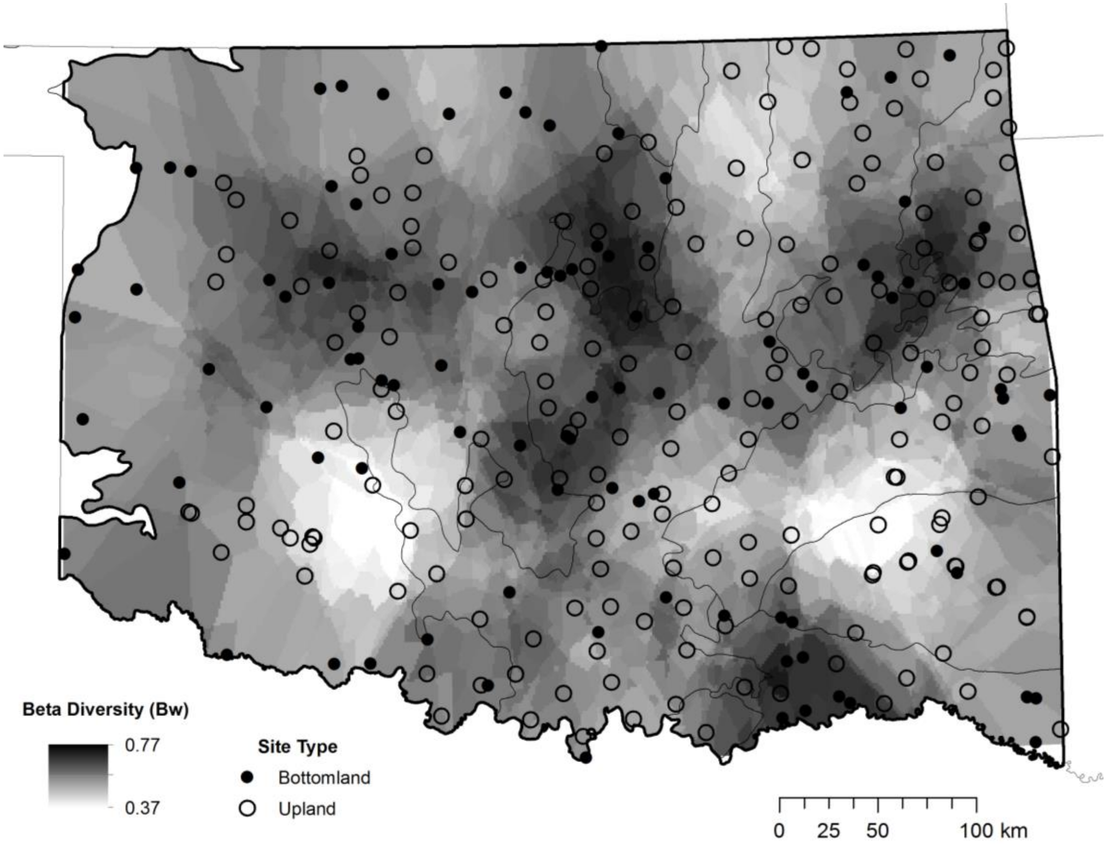


Figure 3.2. Spatial patterns in tree beta diversity using Whittaker's beta diversity measuring the shared species between sites from Koleff et al. (2003). Bottomland (Johnson 1982) and upland (Rice and Penfound 1959) sites are shown to provide spatial representations of the two community matrices

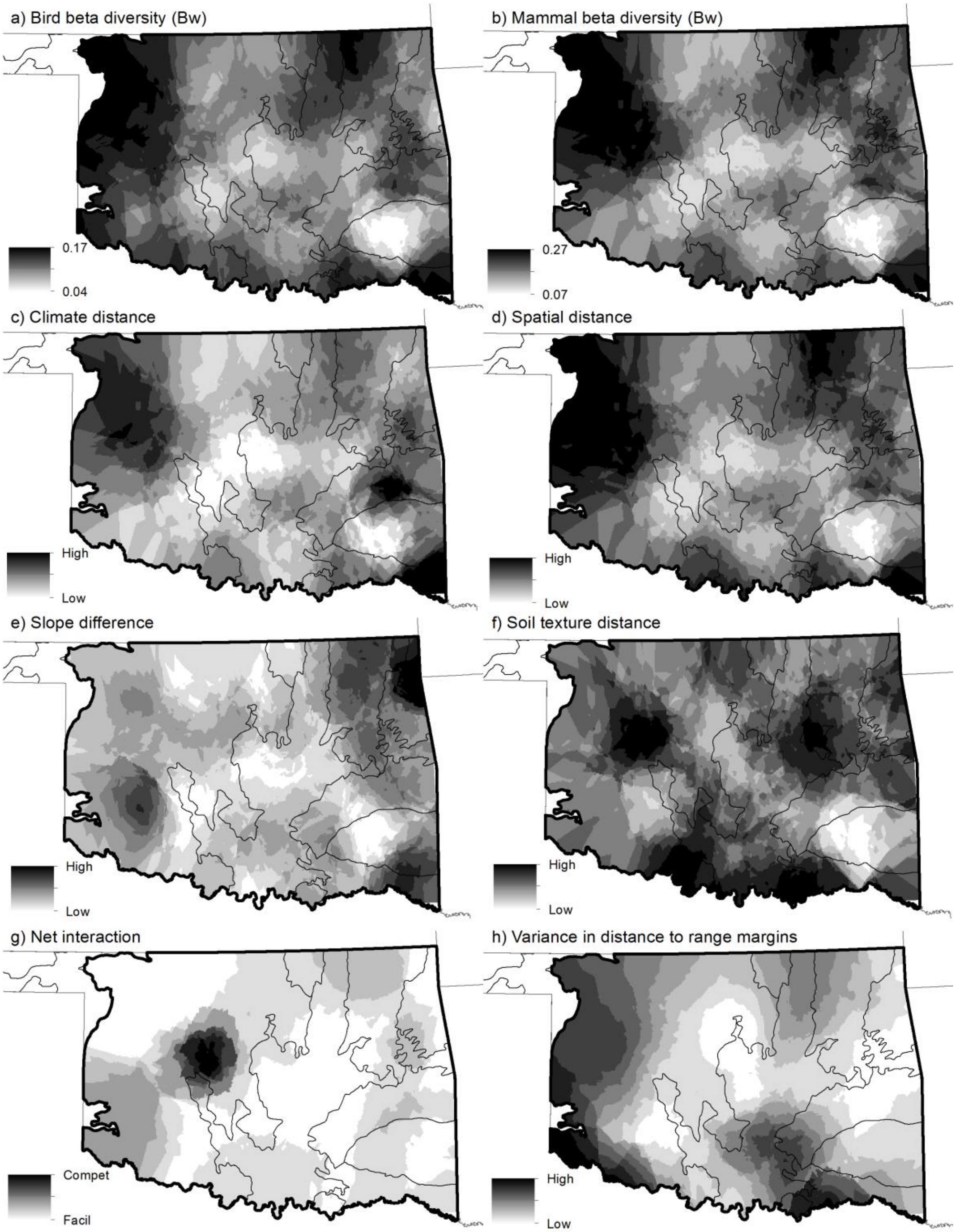


Figure 3.3. Spatial representations of: a) bird and b) mammal beta diversity using Whittaker's beta diversity measurement from Koleff et al. (2003), c) climatic distance, d) spatial distance, e) difference in slope, f) soil texture distance, g) net interspecific interaction, and h) the variance in proximity to range margins between sites

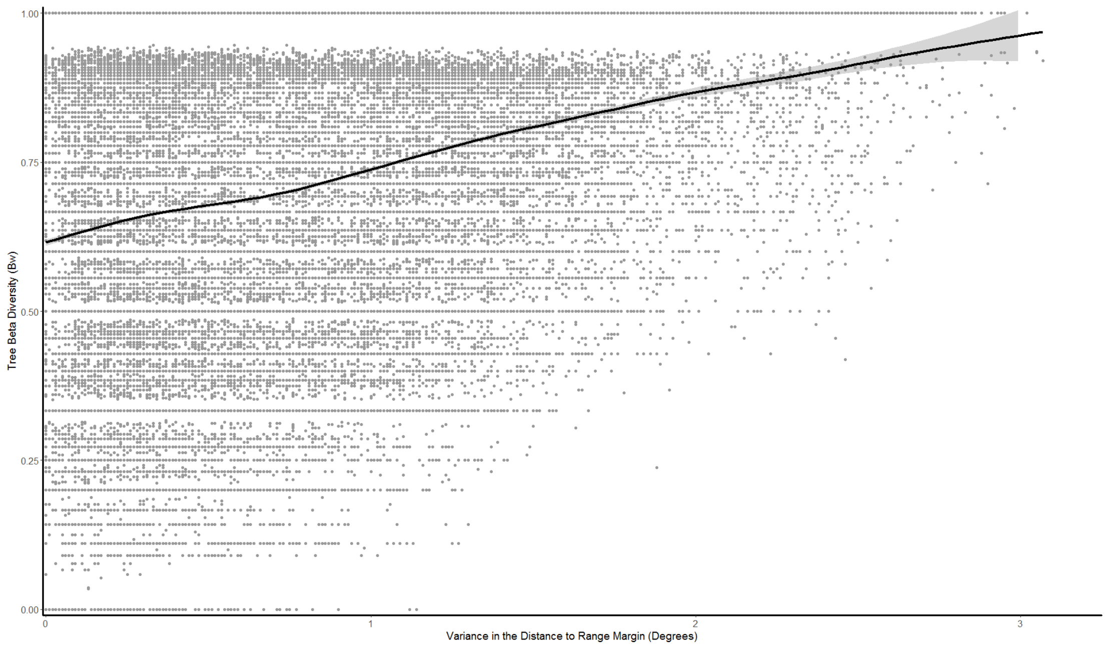


Figure 3.4. The relationship between tree beta diversity and the variance in the proximity to range margins. Regression line was created using a GAM (Hastie and Tibshirani 1990) and the standard error around the regression line is shown in gray

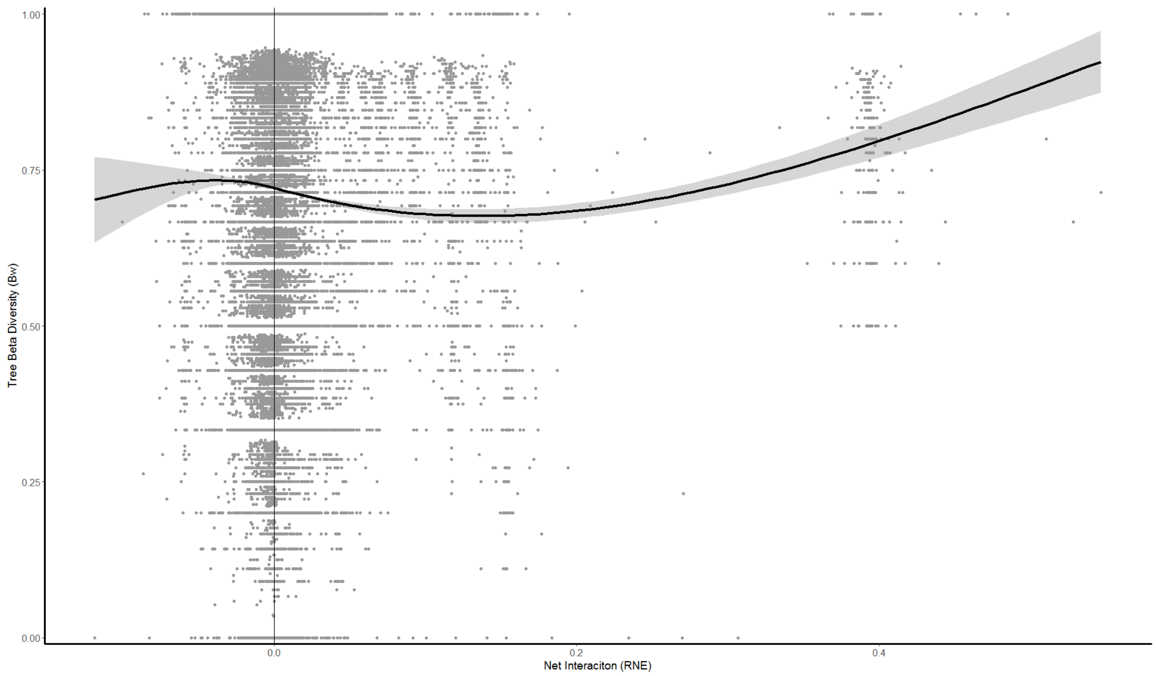


Figure 3.5. The relationship between tree beta diversity and the net interspecific interaction (RNE) between sites. Regression line was created using a GAM (Hastie and Tibshirani 1990) and the standard error around the regression line is shown in gray

Chapter IV

Stress and Species Interactions

Introduction

The Stress Gradient Hypothesis (SGH) states that interactions between organisms monotonically shift from competitive to facilitative with increasing stress (Bertness and Callaway 1994). Bertness and Callaway (1994) proposed two mechanisms for these positive interactions. First, plants can gain associational benefits from other species of plants that can ameliorate consumer pressure. For instance, an edible plant may benefit from reduced consumer pressure by growing in association with inedible plants. The second mechanism is the amelioration of environmental stress, by which presence of neighboring plants can reduce light stress, increase soil moisture, or provide shelter from wind disturbance (Bonanomi et al. 2011).

Recently, the direction of net interactions in extreme stress conditions has been debated. A reshaping of the monotonic curve presented by Bertness and Callaway (1994) has been proposed since some studies have found that net interactions are unimodally distributed across stress gradients, in that they shift back to competitive in extreme stress conditions (see e.g., Castahno et al. 2015). Holmgren and Scheffer (2010) posit three explanations for the unimodal curve along stress gradients. First, there may be a threshold at which the facilitation between two individuals is overcome by competition for a resource when that resource becomes scarce. Second, facilitation alone may not be enough to allow growth under extreme conditions. And third, since organisms in a community are adapted for the abiotic and biotic conditions in that community, facilitative interactions may be more prevalent under conditions that appear less stressful than those that are not. However, studies still find support for the original hypothesis of monotonic shift toward facilitative interactions with increasing stress (see e.g., Rey et al. 2016).

Several meta-analyses have found support for the original formulation of the SGH. For example, facilitative interactions decreased between trees and seedling survival with an increase in precipitation (Derroire et al. 2016). Across salinity and nutrient limited conditions, and with or without consumer pressure, a monotonic shift from competitive to facilitative interactions was found between consumers and producers (Smit et al. 2009). Soliveres and Maestre (2014) found an increase in the frequency of facilitative interactions under more arid conditions in plant communities. Finally, tree-grass interactions shifted from competitive to facilitative under conditions of decreased rainfall (Dohn et al. 2013).

However, several meta-analyses have provided evidence contrary to the SGH. A meta-analysis of plant interactions found a monotonic relationship between interaction and stress but in the opposite direction – competitive interactions under high stress and facilitative interactions under low stress across an aridity gradient (Maestre et al. 2005). That study has been criticized for its selection criteria, such as the incorporation of studies that did not measure or define a stress gradient, did not quantify the length of gradient analyzed, or were not peer-reviewed (Lortie and Callaway 2006). However, Maestre et al. (2006) defended their results by demonstrating that accounting for differences in stress gradient length did not produce different conclusions. Similarly, Soliveres and Maestre (2014) discovered fewer facilitative interactions under higher relative stress measured as increasing elevation in alpine plant communities and larger range in annual temperature in dryland plant communities.

Most previous tests of the SGH have been at small spatial scales, such as an individual plot or across many plots while manipulating the stress gradient. When referring to the scale of tests of the SGH, I define ‘scale’ as the grain at which interactions between organisms has been evaluated. At the smallest scales (< ten m^2), relatively few studies have corroborated the SGH. However, beneficiary plants in extreme arid conditions were found only to exist where heat stress is ameliorated by the shade provided by a nurse plant (Armas et al. 2011). Similarly, performance of alpine plants was higher when grown with neighbors at high elevations (Callaway et al. 2002). Finally, interactions between grass species shifted from competitive to facilitative between sites relatively unaffected and sites and highly affected by anthropogenic disturbance respectively (Villarreal-Barajas and Martorell 2009).

Conversely, a larger number of studies have discovered results that disagree with the SGH at scales less than ten m^2. For instance, interactions shifted back to competitive under water stress (and proximity to the sea) between two rainforest tree species in Brazil (Castanho et al. 2015). Interactions shifted from facilitative to competitive under higher stress between mosses and lichens at a grain of 1.5 m transects in a more arid environment across a 112,000 km^2 area of Spain (Bowker et al. 2010), in interactions at the grain of a single plant between grasses with increasing temperature (Cavieres and Sierra-Almeida 2012), and under increased density of *Elymus nutans* in one m^2 plots (Chu et al. 2008). Similar results were discovered in trees when stem grafting and subsequent decreased mortality occurred in the absence of stress (McIntire and Fajardo 2011), and in shrub-grass interactions at a grain of one m^2 under high grazing pressure (Graff and Aguiar 2011). Lastly, the intensity of competition between shrub seedlings and their herbaceous neighbors increased with lower water availability in 0.1 m^2 plots (Rysavy et al. 2016).

When the scale at which interactions are measured increases, the vast majority of tests agree with the SGH. Facilitative interactions dominated unsheltered sites at high elevation, while competitive interactions characterized sheltered low elevation sites at a grain of 100 to 150 m^2 (Choler et al. 2001). Recently, greater positive spatial associations in shrub species were found in 30 m^2 plots with increasing aridity in the Atacama Desert (López et al. 2016). Additionally, many tests of the SGH analyze nurse plant relations, a situation where a plant is provided a resource by another plant (the nurse plant). On south facing slopes with low precipitation, competitive intensity between a nurse plant and its beneficiary decreased along 100 m^2 transects in the Mediterranean semi-arid region of Spain (Soliveres et al. 2010).

In the columnar cacti forest of the Tehuacán-Cuicatlán valley in southeastern Mexico, Valiente-Banuet and Verdú (2008) analyzed interspecific interactions along one km^2 transects between 102 woody species (761 species pairs). The mean phylogenetic distance between facilitative species pairs was significantly higher than random expectation, while mean phylogenetic distance between competing species was significantly lower. Within a given community, taxonomic differences between species differ based on the taxonomic scale studied. For instance, if focusing on the small taxonomic scale (i.e., a genus) species are often less related than expected while at a larger taxonomic scale (i.e., flowering plants) species are usually more phylogenetically similar than expected (Cavender-Bares et al. 2006). This is likely due to competition between closely related species which often co-exist at smaller spatial scales and have more functional similarity (Swenson et al. 2007).

Stress is an ambiguous term, however, that can be defined as an extreme form of resource limitation, and not simply a departure from what is physiologically optimal for the individual (Körner 2003). Körner (2003) argues that from the position of the plant departure from optimal conditions is anything but stressful, it is normal. Exposure of individuals or a population of a species to environmental heterogeneity drives adaptation to changing conditions through natural selection (Hoffmann and Sgrò 2011). In light of this, researchers should consider stress in two facets: resource (water, nutrients, light, etc.) and non-resource (temperature, altitude, disturbance, etc.) (Maestre et al. 2009). Analysis of these gradients separately have led to disparate results. For example, under high stress the net interaction between two species is expected to be positive and negative for non-resource and resource gradients respectively (Maestre et al. 2009).

In addition to analyzing resource and non-resource gradients separately, several authors suggest that the disparate results when testing the SGH could be caused by only considering a few species and not entire communities as formulated in the original hypothesis (Soliveres and Maestre 2014), or by not analyzing the entire stress gradient (López et al. 2016). I also posit that the commonness or rarity of a species being analyzed should affect the types of interactions they experience since species presence in stressful areas is thought to be determined by a tradeoff between stress tolerance and competitive ability where more broadly tolerant species are affected to greater extent by competition (Grime 1979; Liancourt et al. 2005).

In light of the disparity in results in tests of the SGH, I believe a regional scale approach is appropriate. Therefore, this chapter focuses on a novel stress gradient, the spatial position of an individual within the geographic range of a species, specifically, the proximity to its range margin. The geographic context of an individual or population in relation to its species' geographic range has been absent from tests of the SGH. However, geographic context may better explain stress when a community is the focus of analysis since species within communities respond to environmental stressors individualistically (Gauch and Whittaker 1972; Collins et al. 1993; Hoagland and Collins 1997). Range margins theoretically represent a stressful environment for an individual in the absence of geographic barriers (Gaston 2003). For instance, a species' range margin may exist along a mountain range or the coast of an ocean where environmental conditions are near the species-specific environmental optima. In these scenarios, range margins may not represent stressful environments as much as they represent barriers to dispersal.

Here I am using proximity to range margins as a surrogate for stress on a species and, therefore, make no claims as to what is physiologically stressful to a species. The purpose of utilizing proximity to range margins is to analyze whether the shift in interspecific interactions from competitive to facilitative with increasing stress can be discovered at regional scales and across communities of species. Additionally, whether taxonomic groupings or the rarity of a species affects the type of interactions a species experiences across a stress gradient may assist in explaining the disparity of relationships between interaction type and stress in previous tests of the SGH. Due to the regional scale of this study, I assessed species interactions between sites using nearest neighbor methods at farther distances than direct interactions occur. Therefore, interactions in this chapter are defined as indirect interspecific interactions based on co-occurrence patterns between species and the effect of co-occurrence on species performance, measured as importance.

I assess five research questions to address those goals. 1) How do interspecific interactions change with proximity to a species' range margin? 2) What is the magnitude of interaction change with decreasing distance to a species' range margin compared to the random expectation? 3) What proportion of species exhibit monotonic or unimodal relationships between net interspecific interactions and proximity to range margin? 4) Do interactions between species within a family or genus differ from interactions between species of different families or genera? 5) Do interactions differ based on the commonness or rarity of a species?

Study Area

The study area is located in central United States within the state of Oklahoma at the conjunction of the temperate forests in the east and the Great Plains in the west (Figure 3.1a) and consists of ten ecoregions (Woods et al. 2005). The Ouachita Mountain Ecoregion occurs in the southeast and consists of folded east-west ridges with forests dominated by *Quercus stellata* and *Pinus echinata*. The Cross Timbers ecoregion lies north-south across the center of the region. The topography is of rolling hills and the vegetation is a mosaic of forests, woodlands, and grasslands. *Quercus* *stellata* and *Quercus marilandica* are the dominant tree species in Cross Timbers forest land. The Central Great Plains ecoregion encompasses much of the western part of the study area and is characterized by grasslands with scattered forest land dominated by *Quercus marilandica*, *Quercus stellata* in the uplands, and *Ulmus americana* in the bottomlands. Other ecoregions not described here are used for reference in this chapter.

The climate of the region is predominated by a longitudinal precipitation gradient. Average annual precipitation in this region decreases dramatically from east (1480 mm) to west (520 mm). A latitudinal temperature gradient also exists in the region, with average annual temperature increasing from north (14 deg. C) to south (18 deg. C).

Methods

Data

Two datasets for forest communities across a precipitation gradient were combined and utilized here. The first was collected by Rice and Penfound (1959) for an analysis of upland forest vegetation between 1953-1957. The second was collected by Johnson (1982) to survey woody tree vegetation at bottomland forest sites between 1977-1982. Frequency, density, and basal area were calculated per species. Field data were collected at forty randomly selected points per site using an augmented variable radius technique. A tree was recorded if the distance from a sample point was less than 33 times its diameter. Calculation of basal area is the number of trees selected multiplied by ten (i.e., four trees equals a basal area of forty ft^2/acre). Frequency and density were calculated by forty arm-length transects of approximately forty m^2 each totaling 1600 m^2 at each site. Density was converted to a unit of individuals per acre. This method was standard practice in Germany for expedited forest surveys and introduced to American foresters by Grosenbaugh (1952). These variables were then averaged to generate importance values (IV) for each species at each site. The Rice and Penfound (1959) dataset consists of 70 species over 205 sites, while the Johnson (1982) dataset consists of 69 species across 102 sites, each organized in a site by species matrix. The combined matrix consists of 91 species across 307 sites. Finally, each site had township, range, and section information that were used to identify site locations.

Monthly climate data (PRISM Climate Group 2017) for maximum, minimum, and mean temperature, as well as total precipitation were collected for the thirty years prior to the culmination of Johnson's survey (1953-1982) and utilized to create monthly climate normals. Climate normals were calculated by averaging monthly data over the time period per grid cell. Precipitation data from PRISM contains approximately 5% error annually (between 3 and 6% monthly) in Oklahoma on average, with slightly larger error in western than in eastern Oklahoma (Daly et al. 2008). However, values are not systematically under or over-predicted and therefore, error is random.

Geographic Range Delineation

Range margins were delineated by generating species' geographic ranges using MaxEnt (Elith et al. 2011). Species presence information was downloaded from the Forest Inventory and Analysis dataset (Bechtold and Patterson 2005) for species within the upland and bottomland datasets. FIA data were used to generate range margins to avoid analyzing the relationship between proximity to range margins and interspecific interactions from a single source, the community dataset. The climate variables described above were transformed using Principal Components Analysis (PCA). Six climate axes were chosen based on diminishing returns of variance explained by successive PCA axes using a scree plot. These six climate axes were used as explanatory variables in the distribution models. Species' geographic ranges for the entirety of the contiguous United States were modeled using ten subsample replicates withholding 10% of sample points for testing. Range margins were defined using the average likelihood value of the entire median MaxEnt result (following Cramer 2003). Distance to the range margin of a species was calculated as positive if the site falls within the geographic range and negative if the site is outside of the geographic range.

Spatial Interaction Model

Knowing the relative intensity of the significant interactions between two species is useful to compare across studies and determine changes in interactions across gradients. Markham and Chanway (1996) developed the "relative neighbor effect" (RNE) index to compare the performance of individuals grown in mixture and in monoculture (Eq. 1). The RNE index is symmetric around zero and produces more intuitive scores (if RNE is between -1 and 0 the net interaction is considered facilitative, if RNE is between 0 and 1 the net interaction is considered competitive). The RNE is calculated as follows:

$$RNE = (P_{-N} - P_{+N})/\max(P_{-N}, P_{+N}) \qquad \text{Eq. 1}$$

where P_{-N} is the performance without a neighbor, P_{+N} is performance with a neighbor. The RNE is calculated for pairs of species. P_{+N} is calculated at each site as the average importance value (IV) using a nearest neighborhood of sites when both the focal species and another species are present (Eq. 2). P_{-N} is calculated at each site as the average IV using a nearest neighborhood of sites when the focal species is present, and the other species is absent.

I analyzed the difference in interaction intensity over space using a nearest neighbor analysis iteratively for each site. The Euclidean distance method selects the k nearest points to the focal point where distance is measured as a straight line between points. Calculating RNE requires that a set of nearest neighbors contain sites where both species being compared are present and where the focal species is present but the other is absent. Simple Euclidean distance will likely not select sites in both categories due to spatial autocorrelation (Condit et al. 2002; Bryant et al. 2008). Therefore, the three closest sites where the two species co-occur and three closest sites where only the focal species occurs were selected. To address autocorrelation of species importance, importance at the focal site was calculated using an inverse distance weighted technique assuming that as neighbors become further separated in space they become less influential with each other.

To my knowledge, no model currently exists to assess spatial change in the direction and intensity of interspecific interactions. Therefore, I created this spatial model in Python programming language (Van Rossum 2007). The code is available in Code A1 in Appendix A. The model takes the following inputs: 1) a species by sites matrix of IV scores, 2) a site-to-site distance matrix, and 3) a species by sites matrix of distances from each site to each species' range margin. The model first creates site presence and absence lists for each species using the species by sites importance matrix.

Next, a focal species is selected and its site presence is compared to presence and absence of all other species in the matrix iteratively. This creates a list of sites in which both species are present (“Shared”) and a second list of sites where only the focal species is present (“Focal Only”). If either of the Shared or Focal Only lists contains less than three sites, the model does not calculate interactions between the two species. At each site in the Shared and Focal Only lists, distances to all other sites in the respective list are gathered from the site-to-site distance matrix and the three closest sites are selected. Importance values are selected for focal species at each of those three sites for both Shared and Focal Only scenarios. Average weighted importance by distance is then calculated using:

$$Avg\ IV = \frac{(\frac{IV1}{d1} + \frac{IV2}{d2} + \frac{IV3}{d3})}{(\frac{1}{d1} + \frac{1}{d2} + \frac{1}{d3})} \qquad \text{Eq. 2}$$

where IV_1 is the importance value for the focal species at the nearest site, d_1 is the distance between the nearest site and the focal site, IV_2 is the importance value for the focal species at the second nearest site, and so forth.

In order to test the effect of phylogeny on species interactions across the stress gradient, I compared interactions between species within the same genus or family with interactions between species in different genera or families. In addition to being in the same genus or family, species were also required to exist in the same community (bottomland or upland). For example, I could not analyze *Fraxinus* L. since the two species in the community matrix exist in different communities; *Fraxinus americana* L. in uplands and *Fraxinus pennsylvanica* in bottomlands and therefore do not co-occur in the community datasets.

The rarity of a species was assessed by classification according to local IV to analyze the differences in the interactions a species experience depending on the dominance of a species as a possible explanation for the disparity of curves found in tests of the SGH. The maximum IV for a species across all sites was used to classify species into three groups: dominant, common, and rare. I classified species with a maximum IV > 50 as dominant, 10 > IV > 50 as common, and IV < 10 as rare. As a result, seven species were classified as dominant (*Fraxinus pennsylvanica*, *Pinus echinata*, *Quercus lyrata* Walter, *Quercus marilandica*, *Quercus stellata*, *Quercus velutina* Lam., and *Ulmus americana*), thirty as common, and ten as rare (*Acer rubrum* L., *Acer saccharum* Marsh., *Carya laciniosa* (Michx. F.) G. Don, *Cercis canadensis* L., *Cornus florida* L., *Diospyros virginiana* L., *Fraxinus americana*, *Gleditsia triacanthos* L., *Ostrya virginiana* (Mill.) K. Koch, and *Prunus serotina* Ehrh.).

Null Interaction Model

The null interaction model calculates the RNE using the same method described in the Spatial Interaction Model section. First, I created 100 null communities (code can be found in Code A2). While generating these null communities the species richness at each site was kept constant, and species were confined to the communities they occur in within the upland and bottomland datasets. For example, a species that only occurs in the bottomland dataset (i.e., *Betula nigra* L.) was not placed in the species pool for upland sites. Confining species to communities is imprecise since some species only found in one of the datasets undoubtedly exist in both communities either outside of the study area or were not sampled within the study area due to random chance. However, confining species to the datasets they occur in allow for comparisons between the actual and null communities.

At each site, the null model selects a species at random and assigns it an IV randomly between 0 and 100 (the maximum total importance at a site) from a uniform distribution. The model subsequently randomly selects another species and assigns it an IV between 0 and the difference between 100 and the first species' IV, and so forth until the total number of species based on the real species richness at that site have been selected. This method of randomly building null communities is akin to MacArthur's broken stick method (MacArthur 1957).

An IV of 0.1 is the minimum possible value for a present species at a site (following Rice and Penfound 1959). Therefore, at every site multiple species would receive an IV of 0.1 even if that value would increase the total importance at that site over 100. To correct for this, after all species were selected at a site the importance greater than 100 was subtracted from the most common species' IV (i.e., if the most common species had an IV of 70 and the site had a total importance of 100.5, these were corrected so that the species' IV was 69.5 and total site importance was 100). The process of generating null communities was performed for bottomland and upland sites separately and subsequently merged into a single matrix (code can be found in Code A3).

A uniform distribution should be expected for null communities since co-occurrence and relative importance are decoupled from environmental preferences in the random community and therefore should not be affected by the level of stress (Gotelli 2001). Furthermore, range margins are kept in place; however, a species can be randomly sampled at any site within the communities from which they occur regardless of the distance beyond its range margin. Therefore, the presence of range margins is not expected to have an effect on the type or intensity of interspecific interactions.

A general additive model (GAM) was utilized to fit relationships between RNE and proximity to range margins or IV. The GAM chosen fits the model by smoothing partial residuals using a locally weighted linear least squares algorithm which estimates weights for data points based on proximity to the fitted mean (Hastie and Tibshirani 1990). No additional smoothing was performed on regression lines derived by the GAM.

Results

The Cross Timbers is situated at the margin of the geographic range for many eastern and western tree species, and the density of tree species' range margins is greater in the Cross Timbers than anywhere else in the study area (Figure 3.1b). In the southern Cross Timbers, range margin density is highest at 5.1 km of margin per km^2. Range margin density is lowest in the southeast part of the Southern Coastal Plains at 1.8 km of margin per km^2.

Interactions in the Real Community

I calculated indirect interspecific interactions between 58 species and 1570 species pairs. Distances from sites where a species is present to its range margin is approximately normally distributed to slightly positively skewed (Figure E1). The median and standard deviation of proximity to range margins is 188.1 km and 128.4

km, respectively (Table E1). The median distance between neighbors used in calculating RNE was 24.1 km (interquartile range = 21.6 km). The range margin gradient is not synonymous with any single environmental gradient since the direction of the range margin was not differentiated (i.e., north, south, east, and west margins were treated the same). For example, the western range margins for species such as *Quercus stellata* (Figure E2a) and *Quercus marilandica* (Figure E2b) are related to the longitudinal precipitation gradient. Conversely, the eastern range margins of western United States species (12% of the species in these datasets) may be located within the state. For example, the geographic range of *Celtis reticulata* (Figure E2c) is also related to the longitudinal precipitation gradient; however, its range margin is located under wetter conditions relative to its geographic range indicating that water availability is likely not limiting at its range margin. Several species, such as *Gymnocladus dioicus* (Figure E2d), have range margins within Oklahoma that respond to multiple gradients.

Over nearly the entire range margin gradient the net interaction is slightly competitive (Figure 4.1). The most competitive net interactions occur farthest away from range margins (mean RNE = 0.05) in addition to a local increase in competitive interactions around 125 km from range margins. The most facilitative net interactions are located beyond range margins (mean RNE = -0.02). Interactions became facilitative on average approximately 100 km beyond range margins (Figure 4.1). Overall, the trend in net interactions is monotonic and shifts to more facilitative with increasing proximity to range margins, and therefore increasing stress.

I tested for the shape of the RNE distribution across the proximity to range margin gradient for each species individually. I expected to find highest proportions of unimodal (with competitive interactions at the center of the gradient) and monotonic distributions. This expectation was only partly realized (Table 4.1). Unimodal distributions were most common (15 out 47 species), followed by monotonic distributions and distributions with multiple modes (12 species each). Four species' RNE distributions possessed no trend. Finally, two species had inverse monotonic (competitive interactions more common near range margins) and two species had inverse unimodal distributions (more facilitative interactions at the center of the gradient).

Taxonomic groupings should affect interactions with increasing stress. I tested for differences within and outside of genera in addition to within and outside families. Five genera in the community matrix met the criteria of having multiple species occurring within the same community: *Acer* L. (four species pairs), *Carya* Nutt. (ten species pairs), *Celtis* L. (two species pairs), *Quercus* L. (82 species pairs), and *Ulmus* L (six species pairs). Interactions within and outside of genera both monotonically shifted to more facilitative with increasing proximity to range margins, becoming facilitative approximately 75 km and 125 km beyond range margins for within genus and between genera respectively (Figure 4.2a). However, at the mild end of the gradient away from range margins, interactions between species within the same genus were significantly more competitive (mean RNE = 0.5) than interactions between species in different genera (mean RNE = 0.05).

I analyzed five families which met the criteria of having multiple genera occurring within the same community: Fabaceae (six species pairs), Juglandaceae (seventeen species pairs), Moraceae (two species pairs), Salicaceae (two species pairs), and Sapindaceae (eight species pairs). As with genera, interactions both within and outside of families monotonically shifted toward more facilitative with increasing proximity to range margins (Figure 4.2b). Interactions between species in different families became facilitative at approximately 100 km beyond range margins, while interactions between species within a single family did not become facilitative on average. Additionally, differences between interactions within a family and outside of families are not significantly different despite within family interactions being consistently more competitive across the gradient.

Finally, I tested to discern whether the shift in interactions across the stress gradient differed between dominant and rare species. At the local scale, rare species (IV < 3) at a site experience greater intensity of competitive interactions (RNE up to 0.38) relative to all other species (Figure 4.3). Competitive interactions (RNE up to 0.05) also affect highly dominant species at a site (IV > 75). Species that are not overly dominant or rare experience net facilitative interactions (RNE down to -0.05).

At the regional scale, difference between the net interactions of dominant and rare species is not significant across a large part of the stress gradient (Figure 4.4). However, interactions tend to be more competitive for dominant species, with the exception of approximately 250 km from the range margin. The difference between dominant (RNE = 0.06) and rare species (RNE = 0) is most apparent at the mild end of the stress gradient. Dominant species experience greater competitive pressure at distances far within range margins, as well as around 100 km within range margins. Beyond range margins, both dominant and rare species experience relatively more facilitative interactions. Rare species experience net facilitative interactions at approximately 75 km beyond range margins and interactions become increasingly facilitative with increasing distance beyond range margins (RNE down to -0.04). Dominant species experience net competitive interactions beyond range margins, however the intensity of competition decreases (RNE down to 0.01).

Comparison to Null Models

Net interactions between species in the null communities are significantly more competitive across the entirety of the stress gradient (Figure 4.5). Mean RNE ranges between 0.15 and 0.25 in the null communities compared to mean RNE between -0.02 and 0.05 in the real community. The real community shifts from competitive to facilitative interactions beyond the range margin, while the null community remains competitive. The slope of the RNE distribution for the real community is steeper than that of the null community due to a shift towards more facilitative interactions in the real community. Finally, the shape of the RNE distribution differs between real and null communities. The real community generally exhibits a monotonic shift from competitive to facilitative interactions with increasing stress. The null community RNE distribution is uniform or slightly unimodal, with lower competitive intensity at both ends and higher competitive intensity near the middle of the stress gradient.

Discussion

Within the context of the SGH, I addressed five research questions: 1) how do interspecific interactions change as distance to a species’ range margin decreases, 2) what is the magnitude of interaction change with decreasing distance to a species’ range margin compared to the expected change in a randomly composed community, 3) what proportion of species exhibit monotonic or unimodal relationships between net interspecific interactions and proximity to range margin, 4) do interactions between species within a family or genus differ from interactions between species of different families or genera, and 5) do interactions differ based on the commonness or rarity of a species?

Aggregating all species in the community matrix, net interactions monotonically became more facilitative with increasing proximity to range margin (increasing stress). This supports the original hypothesis of Bertness and Callaway (1994). However, analyzing species separately produced different results with approximately equal proportions of species exhibiting unimodal, monotonic, and multiple modal distributions. Species respond in an individualistic fashion to environmental gradients (Gauch and Whittaker 1972; Collins et al. 1993; Hoagland and Collins 1997). Therefore, it is also possible that shifts in interspecific interactions between species along stress gradients are also species-specific.

There were two distinct differences in this chapter between how interspecific interactions are distributed along a stress gradient in the actual and null communities: the distribution and intensity (Figure 4.5). The null community exhibited a uniform or slightly unimodal distribution, opposed to a monotonic one. The null community distribution is mostly uniform except for an increase in competitive interactions around 220 km from range margins. A uniform distribution should be expected for null communities since co-occurrence and relative importance are decoupled from environmental preferences in the random community and therefore should not be affected by the level of stress (Gotelli 2001). The intense competitive mode around 220 km from range margins can be explained by the modeling design and edge effects. The model was designed to hold two properties constant to their real life counterparts; range margin placement and species richness at each site. Range margin density is highest in the center of the study area in the Cross Timbers (Figure 3.1b). The distance between the area of high range margin density and the eastern edge of the study area is approximately 220 km. Additionally, the eastern edge of the study area contains sites with higher relative species richness. Therefore, a larger number of random species are assigned to sites leading to greater turnover between sites. Greater turnover equates to less co-occurrence in space and is calculated as higher intensity of competition. Turnover is the cause of greater competition compared to the real community over the entire stress gradient since, in a real community, species are often associated based on similar environmental and niche preferences (Diniz-Filho et al. 2016). In a real community, turnover between adjacent sites is lower and, therefore, net interactions are less competitive.

I have also uncovered effects of stress on community assembly; phylogeny and rarity. Phylogeny alters net interactions between species at the genus level, but not at the family level (Figure 4.2), which is likely due to differentiation between species within these groups. All five families I analyzed contain species of multiple genera. Species of different genera within a single family must therefore not co-occur much differently than species of different genera and families in the bottomland and upland datasets. However, species within a genus are subjected to increased competition intensity at the mild end of the stress gradient far away from range margins (Figure

4.2a) corroborating findings that closely related taxa are more competitive than those that are distantly related (Valiente-Banuet and Verdú 2008). In the columnar cacti forest of the Tehuacán-Cuicatlán valley in southeastern Mexico, Valiente-Banuet and Verdú (2008) analyzed interspecific interactions between 102 woody species (761 species pairs). The mean phylogenetic distance between facilitative species pairs was significantly higher than random expectation, while mean phylogenetic distance between competing species was significantly lower. In the bottomland and upland forest datasets, interactions between species within a genus are approximately equal to interactions between species of different genera under stressful conditions and beyond range margins indicating that competition between all species is less intense under high stress, even amongst similar species. However, as conditions become relatively more benign, species within the same genus that have similar environmental preferences compete more intensely for habitat.

Stress affects dominant and rare species differently, both generally and across stress gradients. At the local scale (a site), species experience more intense competitive interactions at sites where they are locally rare (a species with IV < 3 at a site) (Figure 4.3). These locally rare species would otherwise fill empty habitat in the absence of other species, however, are outcompeted for this habitat by most other species (Tilman 1982). Species also experience net competitive interactions where they are locally dominant (a species with IV > 75 at a site), indicating that at these sites a dominant species could fill all available habitat in the absence of competitive interactions (i.e., environmental conditions are not limiting). Any habitat an additional rare species acquires at the site is habitat that is ceded by the dominant species and, therefore, competition is the primary interaction acting on the dominant species. Species that are not dominant or rare experience net facilitative interactions. Species therefore may be locally dominant due to preferable environmental conditions, higher competitive ability, higher growth rate, lower mortality rate, more time since establishment, or some combination of factors, or species that are not locally dominant. Species that are not locally dominant are either kept locally rare by competition or facilitated by the dominants. Those species locally facilitated above rarity could be provided shade (Jucker et al. 2015), water via hydraulic lift (Dawson 1993), associational defense from herbivory (Bertness and Callaway 1994), release from allelopathy affecting competitors (Callaway 1995), or modification of soil (Bonanomi et al. 2011). Regardless, it is clear that stress structures species importance at a site.

At the regional scale in this study, stress also affects net interactions between species, and these interactions differ based on the local rarity of a species across stress gradients (Figure 4.4). Species that attain local dominance at a location experience more competition regardless of proximity to range margins. This is particularly apparent approximately 100 km from range margins where dominant species experience relatively higher competition. Importance of the dominant focal species decreases and may be replaced with increasing importance of another locally dominant species. At the mild end of the stress gradient dominants experience more intense competitive interactions as well indicating that in their productive habitats dominant species could fill all habitat space at a site in the absence of competitive interactions. Additionally, all seven dominants are species of the eastern deciduous forests of the United States. If we define the eastern part of the study area as the mild end of the stress gradient for the dominants, benign conditions coincide with higher relative species richness, and thus, these eastern dominants experience a greater number of interspecific interactions. Species that remain rare across their entire geographic ranges within the study area experience net facilitative interactions beyond their range margins and relatively neutral (RNE close to zero) interactions elsewhere. Beyond their range margins, presence of these rare species is maintained by facilitative interactions under these stressful conditions. Therefore, broadly tolerant (dominant) species were affected by more intense competition than narrowly tolerant (rare) species which agrees with theory about a tradeoff between stress tolerance and competitive ability (Grime 1979; Liancourt et al. 2005).

There are several methodological concerns that I attempted to address in this chapter. The nearest neighbor analysis utilized compares co-occurrence between two species to the occurrence of a focal species. The requirement to calculate interactions was only three sites where the two species co-occur and only three sites where the focal species occurs but the compared species does not. If a pair of species co-occur infrequently, those three sites could potentially be far apart in space. In this case, interactions would be analyzed over distances in which sites are no longer similar (Condit et al. 2002). Nevertheless, I attempted to address differences in distances between neighbors by using an inverse distance weighted technique assuming that as neighbors become further separated in space they become less influential with each other.

Here I have demonstrated the importance of geographic context in interpreting how interspecific interactions change across stress gradients. I assumed that range margins represent extreme stress to a species. I found monotonic shifts towards net facilitative interactions with increasing proximity to range margins, corroborating results found by many previous tests of the SGH. Therefore, I conclude that geographic context of an individual relative to its species' geographic range is an adequate measure of stress for many species in a community over large spatial areas. Additionally, analyzing entire communities at the regional scale may explain the disparity in previous tests of the SGH. The stress gradient distribution may be monotonic for the community, however, individual species within the community can exhibit a wide range of distributions. Additionally, dominant species experience more intense competition than rare species regardless of level of stress.

I suggest that future studies should explicitly state 1) the geographic location relative to species' distributions, and 2) the relative rarity of a species being analyzed. These two factors may help to illuminate the varying findings when testing the SGH. The majority of range margins tested here (88%) are the leading edges of species' geographic ranges across a drying precipitation gradient. Facilitative interactions at leading edges could "pull" species' geographic ranges towards stressful environments (Jones and Gilbert 2016). Here I discovered net facilitative interactions at the dry range margin when all species were aggregated and for rare species. These results, along with theoretical understanding of how geographic range shifts are altered by interspecific interactions should be utilized to inform and improve species distribution modeling efforts moving forward (Bruno et al. 2003; Michalet et al. 2006; Jones and Gilbert 2016).

Tables and Figures

Distribution	Median Rank	Freq.
Monotonic (inverse)	20	2
Unimodal	20	15
Monotonic	28	12
Multiple	38.5	12
Unimodal (inverse)	44.5	2
No trend	48.5	4

Table 4.1. Summary table for distributions of individual species' net interactions with increasing proximity to its range margin. Species are ranked according to dominance, with 1 = most dominant. Median rank is therefore the median dominance rank of species that exhibit each distribution. Total number of species exhibiting each distribution are also shown (Freq.)

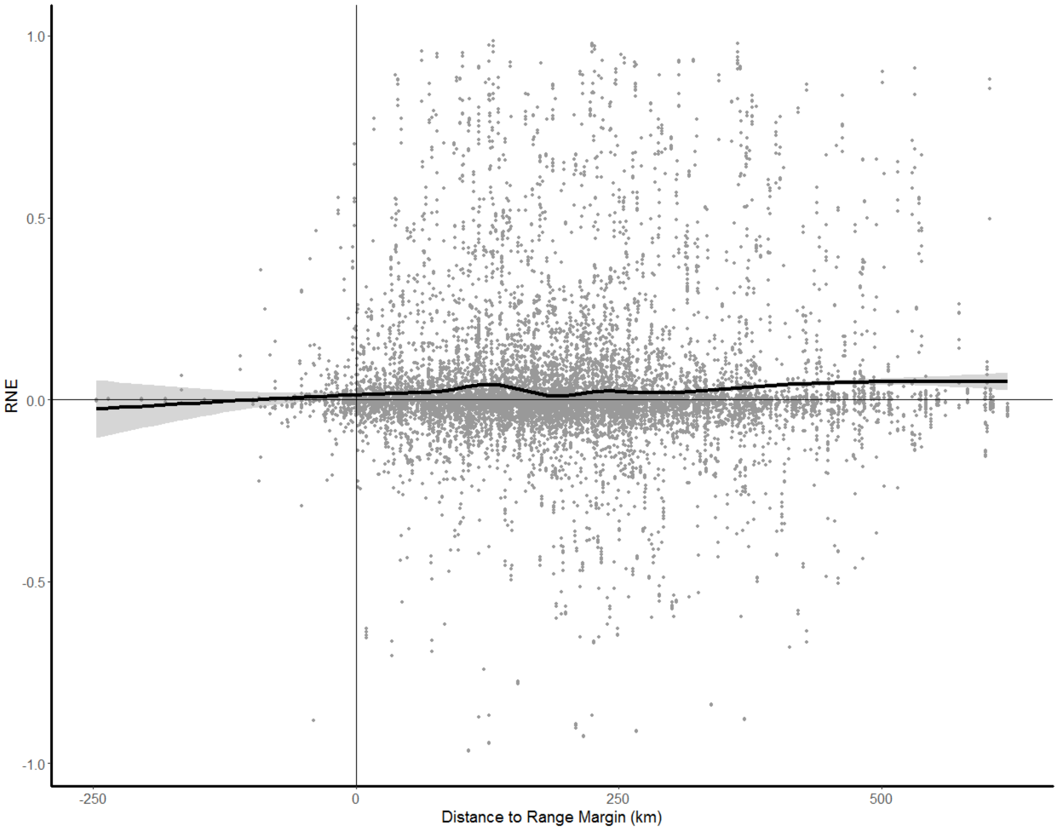


Figure 4.1. Shift in net interactions between all species across a gradient of distance to each species respective range margins. Regression line was created using a GAM (Hastie and Tibshirani 1990) and the standard error around the regression line is shown in gray. Positive distance indicates species occurrence within its range, negative distance indicates species occurrence beyond its range margin

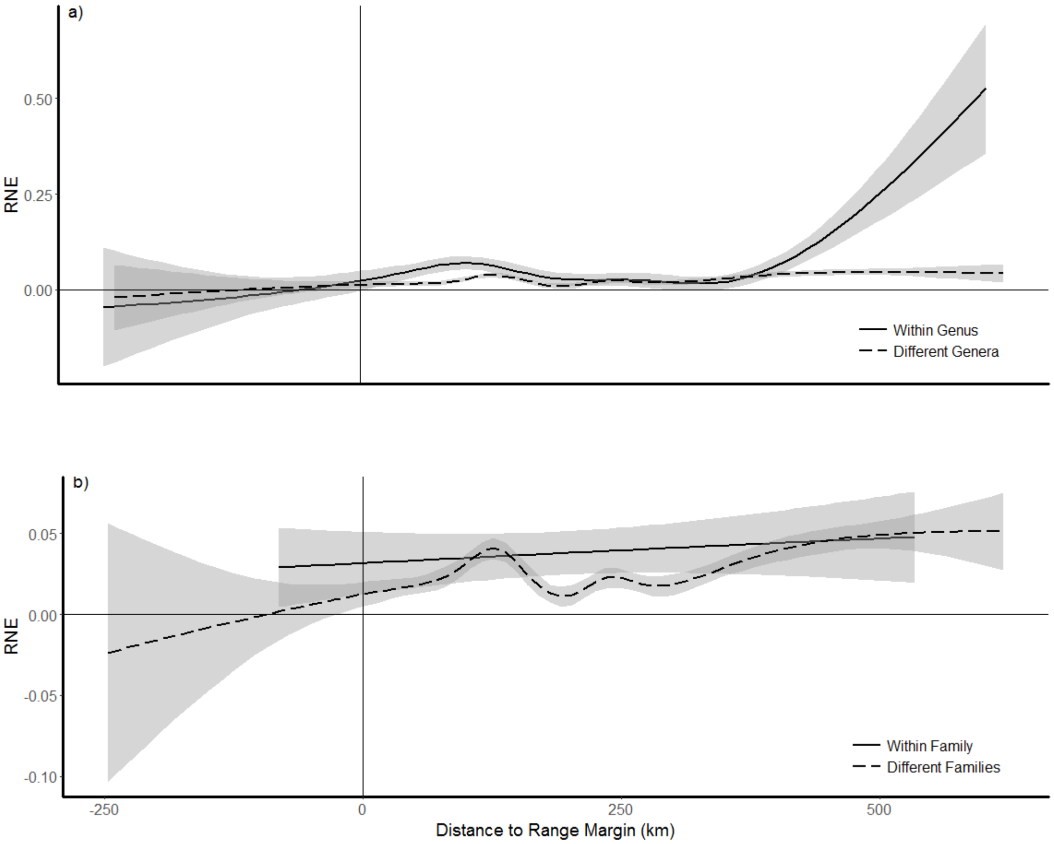


Figure 4.2. Shifts in net interactions across a gradient of increasing proximity to range margins, comparing a) species within a genus to species from multiple genera, and b) species within a family to species from different families (note the difference in y-axis scales). Regression lines were created using a GAM (Hastie and Tibshirani 1990) and the standard errors around the regression lines are shown in gray. Positive distance indicates species occurrence within its range, negative distance indicates species occurrence beyond its range margin

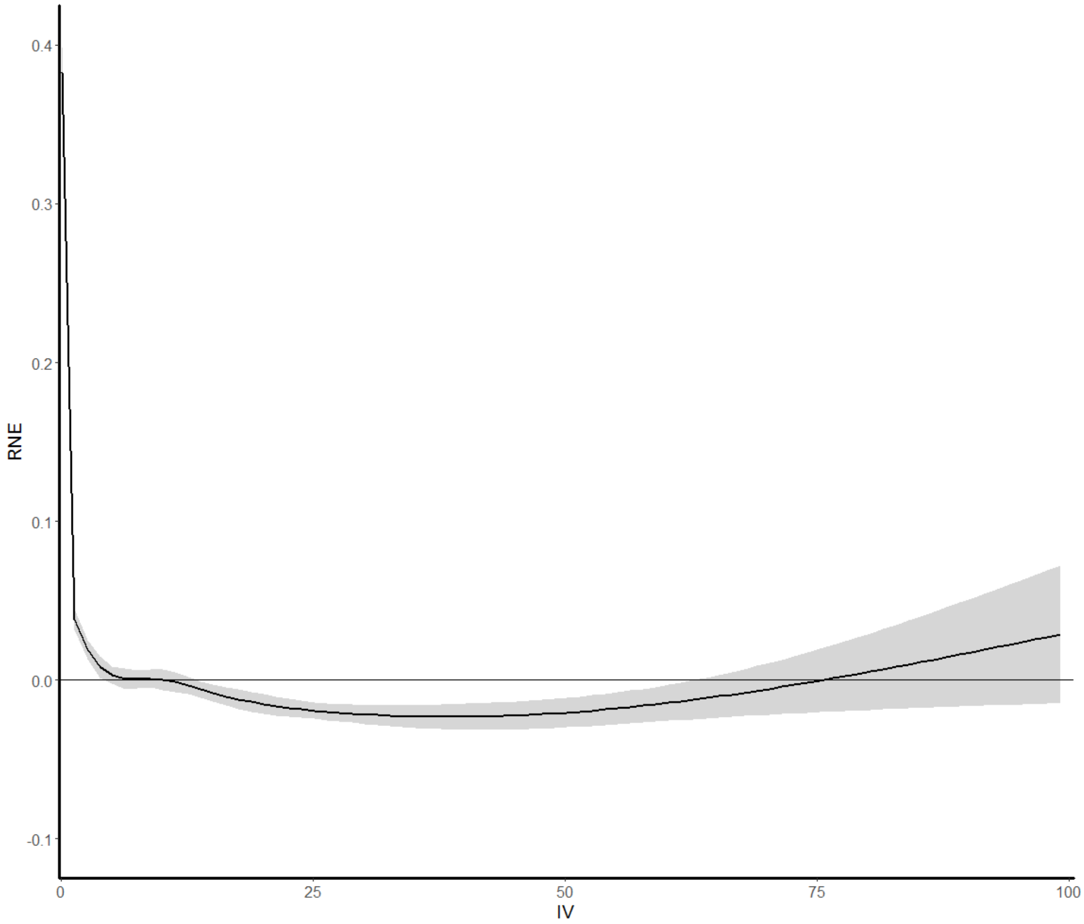


Figure 4.3. Net interactions for local importance of a species at a site (IV = 0 if species is not present at site, IV = 100 if species is the only species at site). Regression line was created using a GAM (Hastie and Tibshirani 1990) and the standard error around the regression line is shown in gray

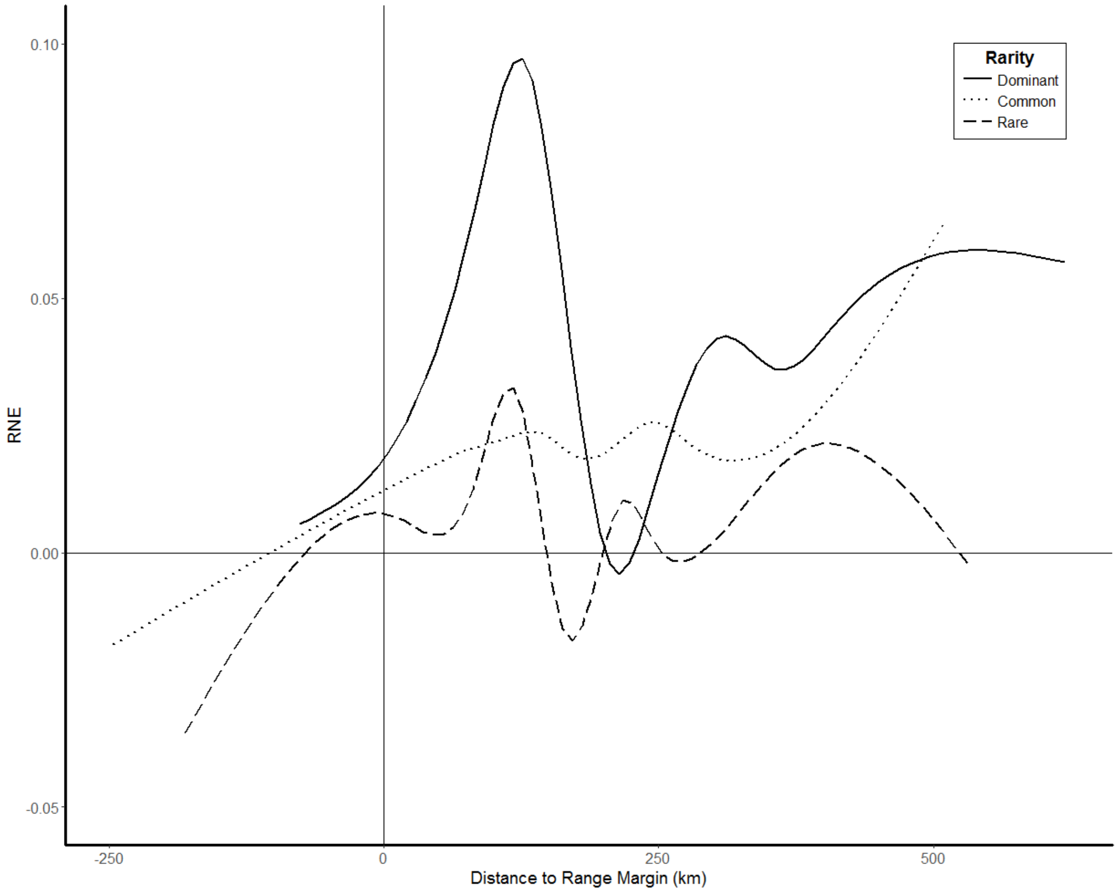


Figure 4.4. Net interactions across a gradient of increasing proximity to range margins for dominant species (maximum IV > 50, solid line), common species (10 > maximum IV > 50, dotted line), and rare species (maximum IV < 10, dashed line). Regression lines were created using a GAM (Hastie and Tibshirani 1990). Positive distance indicates species occurrence within its range, negative distance indicates species occurrence beyond its range margin

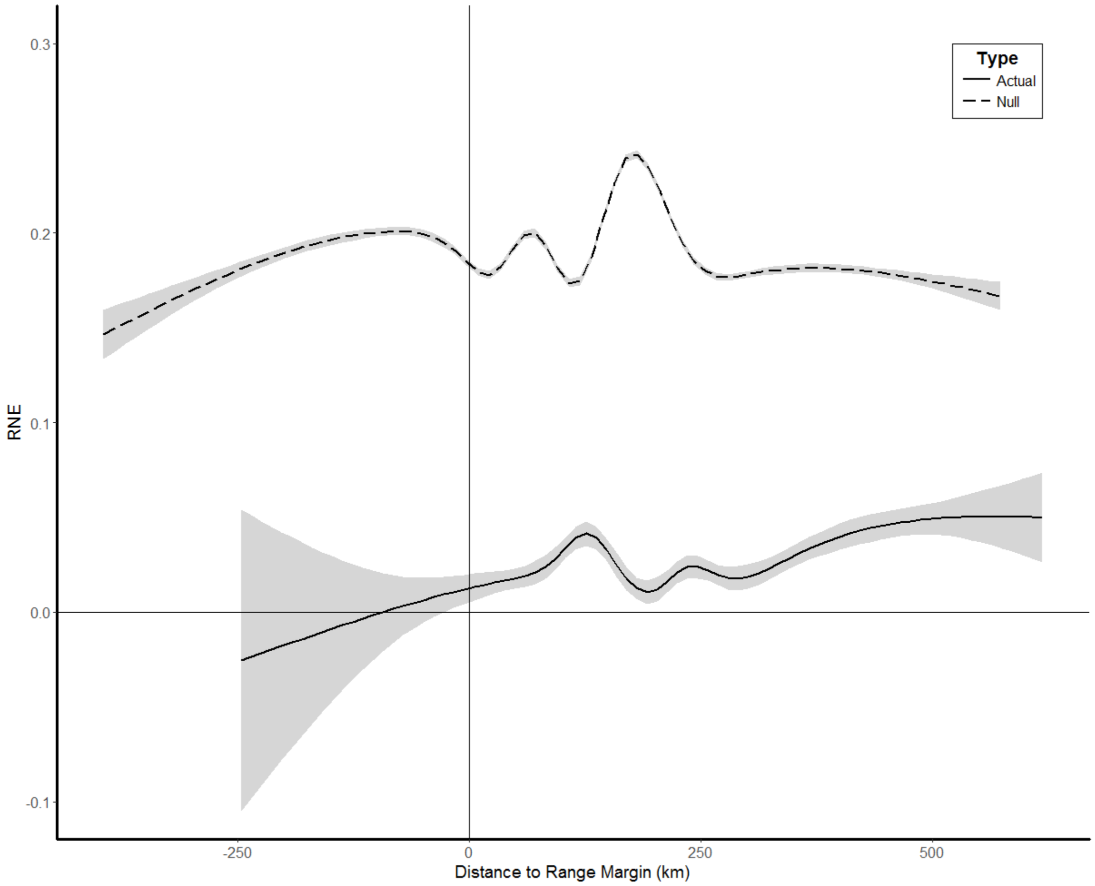


Figure 4.5. Net interactions between all species across a gradient of increasing proximity to range margins for actual (solid line) and null (dashed line) communities. Regression lines were created using a GAM (Hastie and Tibshirani 1990) and the standard errors around the regression lines are shown in gray. Positive distance indicates species occurrence within its range, negative distance indicates species occurrence beyond its range margin

Chapter V

Synopsis

Answering Research Questions

This dissertation has utilized forest survey data within the state of Oklahoma to better explain the relationships between community assembly, beta diversity, and interspecific interactions across species' geographic ranges. Before discussing the broader implications of the findings in this dissertation further, I believe it to be worthwhile to restate the research questions posited within and highlight their respective findings. Full results are located within the chapters individually.

Bottomland Forest Structure

R1. Are bottomland forest species in the Johnson dataset aggregated to a greater extent in physical or environmental space?

Bottomland forest communities are generally spatially clustered and transition longitudinally, responding to the east-west precipitation gradient across the state. Average annual precipitation, as well as monthly precipitation between September and April are most highly correlated with this transition. The spatial clustering of communities at the wettest and driest ends of the precipitation gradient are the most well defined suggesting that species that possess some competitive advantage at either end of the precipitation gradient tend to cluster together spatially. Communities in the middle of the precipitation gradient appear to be more clustered in environmental space than physical space, where gradients in the range in precipitation and minimum/maximum July temperatures partly drive compositional change.

R2. Is variance in community composition better explained by environmental (species sorting) or spatial (mass effects) mechanisms?

Results from an NMS ordination suggest that precipitation is the dominant gradient structuring bottomland communities (Table 2.1). Specifically, average precipitation in the months of December, January, and April in addition to average annual precipitation are most highly correlated to the primary NMS axis; whereas, the standard deviation of monthly precipitation and the range of monthly precipitation are most highly correlated to the secondary NMS axis. Environmental variables explained 22% of the variance in community data between sites, twice as much as either spatial or edaphic variables, providing further evidence that abiotic conditions are the primary drivers of community composition at regional scales (O'Brien 1993; Cottenie 2005; García-Valdés et al. 2015; Jones et al. 2016; Zellweger et al. 2016).

Beta Diversity and Stress

R1. Does variation in stress affect beta diversity between sites?

I tested a multitude of potential stressors for their effects on tree beta diversity. Climate distance between sites was most highly correlated with change in beta diversity. The farther apart in climate space two sites are the higher beta diversity is likely to be (Spearman's $\rho = 0.26$). Several other stressors were significantly positively correlated to beta diversity: mammal beta diversity ($\rho = 0.25$), bird beta diversity ($\rho = 0.23$), variance in the distances to range margins at a site ($\rho = 0.18$), and difference in soil texture between sites (ρ = 0.14). The difference in slope between sites and the difference in net interspecific interactions between sites were also positively correlated with tree beta diversity; however, not significantly. Finally, increase in physical distance between sites is positively correlated with tree beta diversity, as expected (Condit et al. 2002; Byrant et al. 2008). However, due to violations of the underlying assumptions of mantel tests I have refrained from making a significance claim.

R2. Does beta diversity change with net interspecific interaction?

I found no trend in beta diversity with change in net interaction (Figure 3.5). This is likely caused by competition or facilitation acting on both common and rare species simultaneously, rather than facilitative interactions simply benefitting rare species and subsequently increasing beta diversity for instance. However, at highly competitive sites, the variance in beta diversity decreases and is generally higher than beta diversity at facilitative or neutral (RNE close to zero) sites. This finding indicates that under relatively benign environmental conditions, higher levels of competition act on common species rather than on rare species.

Stress and Species Interactions

R1. How do interspecific interactions change with proximity to a species' range margin?

I reaffirmed several prevailing theories related to current stress gradient hypothesis (SGH) research. To accomplish this, I utilized the proximity to species' range margins as a surrogate for stress on a taxon. I found that interspecific interactions shift from competitive to facilitative monotonically with increasing proximity to range margins, corroborating the original formulation of the SGH (Bertness and Callaway 1994). Additionally, I have provided several novel insights related to the SGH. First, I have analyzed the SGH in the context of a species community. Interspecific interactions across entire species communities are rarely studied with regards to the SGH (Soliveres and Maestre 2014). Secondly, I have described an indicator of stress that has yet to be analyzed in tests of the SGH, proximity to a species' range margin. Proximity to range margin may be a more adequate measure of stress on a given taxon and may prove to be a useful indicator in multi-species studies in which individual species are limited by different abiotic or biotic gradients. I discovered that analyzing entire communities across the proximity to range margin gradient yielded results supporting the original formulation of the SGH (Bertness and Callaway 1994).

R2. What is the magnitude of interaction change with decreasing distance to a species' range margin compared to the random expectation?

Two differences between the actual and null communities were discovered. First, although interactions monotonically shifted toward facilitative with increasing proximity to range margins in the actual community, interactions in the null communities were uniformly (or slightly unimodally) distributed across the stress gradient. A uniform distribution across stress gradients should be expected for communities in which co-occurrence is independent of spatial proximity between sites and spatial proximity from range margins. Second, null communities experience significantly more intense competitive interactions than real communities over the entirety of the stress gradient. This is due to greater species turnover between spatially proximal sites in the null communities where species assemblages were randomly generated independent of spatial autocorrelation that exists in real communities.

R3. What proportion of species exhibit monotonic or unimodal relationships between net interspecific interactions and proximity to range margin?

Furthermore, I quantified the proportions of species experiencing unimodal, monotonic, and multimodal distributions of interspecific interaction direction across a stress gradient. Unimodal distributions were most common (32%) followed by monotonic and multimodal (each 26%), highlighting that species are affected by interspecific interactions individualistically and therefore partially explaining the disparity in previous tests of the SGH.

R4. Do interactions between species within a family or genus differ from interactions between species of different families or genera?

Interactions far away from range margins between species within the same genus are more intensely competitive than interactions between species in different genera, agreeing with previous research related to taxonomic groupings (Valiente-Banuet and Verdú 2008). Conversely, interactions between species in a single family were not significantly different from interactions between species in different families. This suggests that species in these datasets within a single family are sufficiently distinct, therefore decreasing competition between them.

R5. Do interactions differ based on the commonness or rarity of a species?

Finally, how interspecific interactions change across stress gradients depending on the dominance or rarity of a species was analyzed for the first time. Species that are locally rare experience more intense competitive interactions. Species that attain local dominance at a site experience more competition across the entire stress gradient. Species that remain rare across their entire geographic ranges within the study area experience net facilitative interactions beyond their range margins and relatively neutral (RNE close to zero) interactions elsewhere.

Community Assembly

This dissertation focused primarily on explaining how abiotic and biotic mechanisms determine community assembly and subsequently influence species' geographic ranges in bottomland and upland forests in the central United States. The biotic-abiotic-migration (BAM) framework provides a method for describing the processes that shape species' geographic ranges (Peterson et al. 2011). The BAM framework suggests that there is an abiotic space in which a species can survive (the fundamental niche) that is constrained by a species' ability to reach a site and persist there amongst the community of organisms (the realized niche). Since these three factors influence species presence at a site, they also determine the assemblage of species at a site when analyzed for all species in a regional species pool. Therefore, as sites become more different their respective assemblages should as well. For example, species assemblage similarity between places decreases with increasing physical distance or elevation difference (Condit et al. 2002; Bryant et al. 2008). I tested variables encompassing all three factors in the BAM framework for their influence on community assembly.

Precipitation gradients were most highly correlated with the first and second axes of a NMS ordination performed on bottomland sites. Difference in community composition between bottomland sites is best explained (NMS axis 1) by the primary gradient, average annual precipitation. A secondary gradient (NMS axis 2), the standard deviation of monthly precipitation, also explains part of the difference in community composition. Many species respond unimodally (or bimodally) across the average annual precipitation gradient; however, several species respond monotonically with increasing or decreasing precipitation (Figure 5.1). It should be noted that species that respond monotonically are those in which their entire species response curve has not been captured due to the study area boundary. For example, *Populus deltoides* increases monotonically with decreasing precipitation and becomes the dominant species in bottomlands at the at the dry end of the precipitation gradient in the study area (Figure 5.2). However, *Populus deltoides* is also present in drier areas north of the study area in the prairies of Kansas, Nebraska, and South Dakota (Figure 5.3). This partitioning of the precipitation resource gradient can be explained from two perspectives: 1) a species may be more competitive or 2) more tolerant at different parts of the gradient in relation to other species given that there is a tradeoff between stress tolerance and competitive ability (Dobzhansky 1950; Grime 1979).

Standard deviation of monthly precipitation is more appropriately classified as a non-resource gradient rather than a resource gradient, and therefore species respond differently. Very few species respond unimodally to the standard deviation of monthly precipitation, rather most increase in importance monotonically with increasing stability or variability in precipitation (Figure 5.4). This could be caused by only a portion of the gradient being analyzed. Most species in these datasets are eastern tree species, and therefore this explanation is plausible at the end of the gradient with consistent monthly precipitation but not at the variable end since monthly precipitation in eastern United States is generally less variable (except for Florida) than monthly precipitation in Oklahoma (Figure 5.5). Additionally, the importance of common species increases with increasing variability in monthly precipitation, with the exception of *Populus deltoides* (Figure 5.6). This could be caused by common species being more broadly physiologically tolerant of stressful conditions. For example, species that are exposed to variable conditions have been found to create more persistent seed banks in which a proportion of seeds survive periods of stressful conditions and germinate when conditions are more amenable (Arroyo et al. 2006). Within the study area, areas with consistent monthly precipitation occur in areas with high annual precipitation. Therefore, species that increase in importance with consistent monthly precipitation are those that are superior competitors for soil moisture (e.g., due to faster uptake of water, or greater ability to access water via deeper root systems) when water is not a limiting resource.

Variance in site conditions drive beta diversity in the study area. I found that beta diversity increases with both the greater the difference in climate and soil texture between sites. Difference in climate between sites increases beta diversity due to differing environmental optima for individual species and changes in competitive outcomes along environmental gradients. For example, consider a pair of sites in which one is located under relatively warm, dry conditions and the other under relatively cool, wet conditions. The identities of the species that are able to survive at each site is likely to be different. Additionally, species that are able to outcompete other species at the dry site (perhaps due to higher water use efficiency) will likely not outcompete species at the wet site where water is not limiting due to tradeoffs between stress tolerance and competitive ability (Dobzhansky 1950; Grime 1979). Some species specialize on or are restricted by certain soil orders or textures, either due to the physical and chemical properties or water holding capacity. For example, *Styphnolobium affine* (Torr. & A. Gray) Walp. is a rare species confined to moist soils or limestone outcrops (Little 1981). Additionally, many upland species, such as *Quercus stellata*, cannot persist in waterlogged poorly drained soils due to lack of oxygen exposure to the roots.

As noted above, spatial distance between sites also drives beta diversity. Beta diversity increases with increasing distance between sites, and is perhaps the most ubiquitous property of beta diversity (Whittaker 1960). One cause of this relationship is the increasing likelihood of two places being environmentally different with increasing distance between them such as what was discussed in the proceeding few paragraphs. However, another cause of the positive relationship between beta diversity and distance is dispersal limitation. Large geographic barriers to dispersal, such as mountain ranges or large bodies of water, can prevent propagules of a tree species from reaching sufficient habitat for germination. Although the study area within the state of Oklahoma does not have large barriers to dispersal, smaller scale barriers do exist. For example, species occurrence is heterogeneous within a species' geographic range partly caused by heterogeneity of sufficient habitat due to changes in microsite site conditions such as shading or habitat fragmentation. Furthermore, habitat is temporally heterogeneous. A site may be suitable at one time and unsuitable at another due to changing community composition and environmental change at the site. Therefore, in the absence of large geographic barriers, a propagule finding a site with suitable conditions in space and time act as limitations to migration.

Tree species in the study area are dispersed via gravity, water, wind, and animals. Dispersal by gravity is strictly a spatial process in that new propagules are found within a relatively short distance from the parent tree and therefore the likelihood of saplings being present decreases with distance from adults. Other dispersal mechanisms, such as by animals, are not as straightforward. I found both mammal and bird beta diversity to be positively correlated with tree beta diversity, suggesting that change in community composition of tree species is related to change in the composition of these dispersers. In an analysis not described in this dissertation, beta diversity at sites within a given watershed was no different than randomly chosen sites. Therefore, dispersal by water may not be a common dispersal mechanism relative to other mechanisms.

Greater variation of stress at a site, measured as the variation in the distances to range margins for species present at the site, increases beta diversity of tree species. Both beta diversity (Figure 3.2) and the density of species' range margins (Figure 3.1b) are highest near the center of the study area in the Cross Timbers ecoregion. Beta diversity in the Cross Timbers is likely driven by species turnover of eastern species and the introduction of a few western tree species. However, variation in the proximity to range margins is not high throughout the Cross Timbers, rather it spatially correlates to high beta diversity in the west and southeast in addition to areas of low beta diversity in the southwest and southeast (Figure 3.3h). Areas of high variation in the proximity to range margins correspond to xeric tolerant species (such as *Quercus marilandica*, *Quercus stellata*, and *Juniperus virginiana* L.) in the west and gulf coastal plains species (such as *Taxodium distichum*) in the southeast. These xeric tolerant species in the west are outcompeted by less tolerant, faster growing species as precipitation increases to the east. The gulf coastal plains species in the southeast tolerate waterlogged soils other species cannot, but are outcompeted where soils are well drained. Both of these scenarios result in a relatively large number of species' range margins.

Finally, interspecific interactions drive community assembly as well. Rarity, both locally (Figure 4.4) and regionally (Figure 4.5) affects the direction and intensity of interactions a species is likely to encounter. Species that are locally rare experience more intense competitive interactions. These locally rare species that would otherwise fill empty habitat in the absence of other species are outcompeted for this habitat by most other species (Tilman 1982). Species that attain local dominance at a location experience more competition across the entire stress gradient. These dominant species are those that could fill all available habitat if grown in monoculture. Any habitat an additional rare species acquires at a site is habitat that is ceded by the dominant species and, therefore, competition is the primary interaction acting on the dominant species. Species that remain rare across their entire geographic ranges within the study area experience net facilitative interactions beyond their range margins and relatively neutral (RNE close to zero) interactions elsewhere.

Therefore, there appears to be a dichotomy insofar as species are concerned. Dominant species face a higher intensity of competitive interactions and attain dominance by being both broadly tolerant and avoiding interactions with other species.

Conversely, rare species experience less intense competition across their geographic range. This finding corroborates theory of a tradeoff between stress tolerance and competitive ability for regionally rare species (Dobzhansky 1950; Grime 1979). In addition to having an advantage in tolerance over other species in stressful environments, regionally rare species benefit from a relaxation in competitive interactions. Regardless of the level of rarity, range margins of species are characterized either by more facilitative interactions or by a lower intensity of competitive interactions.

Species Geographic Ranges and Modeling

Facilitative interspecific interactions may more frequently occur under more stressful environmental conditions (Bertness and Callaway 1994; Villarreal-Barajas and Martorell 2009; Armas et al. 2011; Derroire et al. 2016; Rey et al. 2016). This dissertation has demonstrated that facilitative interactions also more often predominate near species' range margins. Spatially, net facilitative interactions occur where they would be expected based on those facts (Figure 3.3g), over a large area of the Cross Timbers corresponding to the highest density of range margins within the study area (Figure 3.1b). Moreover, net facilitative interactions also characterize the far northwest part of the study area where the lowest amounts of precipitation (Figure 2.1b) and high variance in the proximity to range margins (Figure 3.3h) occur. Contrary to areas of stress and high density of range margins, net facilitative interactions also occur along the eastern border of the study area. This is partly caused by several bottomland species

(*Diospyros virginiana*, *Fraxinus pennsylvanica*, and *Platanus occidentalis* L.) establishing in low numbers at the upland sites in those regions.

Therefore, analyzing the proximity to range margins as a surrogate for stress provides a spatial framework for utilizing interactions in the modeling of species' geographic ranges. This is not the first framework for incorporating interactions into species distribution modeling (SDM) efforts, and furthermore is not much unlike previous frameworks. For example, Jones and Gilbert (2016) combined the Cold-Tolerance Competition Hypothesis (CTCH) (MacArthur 1972) and the SGH to describe how interactions might shape range margins. The CTCH states that there is a tradeoff between tolerance of cold temperatures and competitive ability. Jones and Gilbert (2016) posited that at leading range margins (colder or higher elevation, and more stressful) facilitative interactions increase growth rate, while at trailing range margins (warmer or lower elevation, and less stressful) competitive interactions decrease growth rate. Furthermore, they suggested that an aridity gradient could be substituted for the temperature or elevation in proper circumstances, such as I have done in this dissertation.

Here I have demonstrated that interactions shift to facilitative toward all range margins, not only at the drier, stressful, leading margins. Only these leading margins were analyzed for all but two species in the bottomland and upland communities since the large majority of species are eastern United States species with their western range margins occurring within the study area. Range margins at wet end of the precipitation gradient fall within the study area for *Celtis reticulata* and *Sapindus saponaria* var. *drummondii*. *Celtis reticulata* is not represented at sites near or beyond its range margin in the bottomland and upland datasets and therefore cannot be studied for interaction change at the margin. However, a shift to facilitative interactions with increasing proximity to its range margin was found for *Sapindus saponaria* var. *drummondii*.

It is crucial to consider change in environmental conditions when modeling species' geographic ranges, particularly when projecting geographic ranges in space and time. Accounting for environmental change is the reason Jones and Gilbert (2016) distinguished between 'leading' and 'trailing' range margins. As temperature is projected to increase, populations of tree species are expected to shift range margins to the north or to higher elevations to track this change. Part of the difficulty of substituting an aridity gradient into this framework is that, unlike temperature, precipitation is spatially heterogeneous. Therefore, predicting change in precipitation over time spatially is more difficult. However, in Oklahoma soil moisture is expected to decrease significantly (~7%) and drought conditions are predicted to be more frequent and of higher intensity by the year 2100 (Dai 2012). For tree species in Oklahoma, the environmental change described above should contract leading range margins further east of their present location due to decreasing soil moisture beneath physiological tolerances of individual species. Soil moisture stress will increase over the entire region, however, the increase in stress will likely have larger effects in the west where water is already limiting to many species.

Therefore, the main purpose of this dissertation is to discern the importance of drivers on community assembly and how these drivers are altered across spatial, environmental, and stress gradients. Stress gradients are species specific and species' geographic ranges are partly the manifestation of these stress gradients. Viewing 'stress' through this framework, stress for a species will likely increase to a greater extent beyond, at, and just within its range margin. Therefore, the importance of existing facilitative interactions between species near range margins might assist in ameliorating future increase in stress. At a species' range margin, existing trees of any species may provide shade which reduces evaporation of soil moisture or provide an increase in soil moisture through hydraulic redistribution (Dawson 1993). Environmental change might be mediated for species that benefit from existing facilitative interactions at their range margin. Identifying species that do not benefit from facilitation at their range margin may therefore be crucial for modeling their future geographic ranges. In the bottomland and upland datasets, *Acer negundo*, *Carya texana* Buckley, *Celtis laevigata*, *Diospyros virginiana*, *Quercus macrocarpa*, *Salix nigra* Marshall, and *Ulmus alata* all experience increases in competitive interactions at their respective range margins. Therefore, these species may likely be most adversely affected by environmental change.

Modeling species assemblages spatially has proven a difficult endeavor; however, the spatially explicit species assemblage modelling (SESAM) framework uses SDMs and three filtering processes to predict assemblage (Guisan and Rahbek 2011). SDMs have often been utilized to relate species occurrence to underlying environmental data (Peterson et al. 2011). Individual species can be modeled for likelihood of occurrence in multidimensional environmental space, and subsequently these models can be overlaid to generate a list of species that are likely to occur at a site, termed stacked SDMs (S-SDMs, Dubuis et al. 2011). However, S-SDMs always over-predict observed species richness at a location (Dubuis et al. 2011).

Due to this over-prediction, SESAM predicts species assemblage in four steps: 1) the global species pool is defined (e.g., all species in both upland and bottomland forests datasets), 2) species are filtered based on the habitat of a site (e.g., only bottomland species if site occurs in a riparian area) aided by the creation of S-SDMs, 3) sites are constrained by macroecological properties such as limitations to species richness by abiotic gradients, and 4) species are filtered by ecological assembly rules (EARs) such as competitive interactions. Several recent studies have developed novel techniques to better inform species assemblage predictions. Pellissier et al. (2010) used the frequency of occurrence of a dominant species as a predictor variable in SDMs of subordinate taxa and found that model performance improved. Evidence suggests that dominant species exclude each other under favorable environmental conditions (e.g., higher precipitation); however, geographic ranges are controlled by climatic factors under stressful environmental conditions (Meier et al. 2011). Finally, in the first experimental test of SESAM, D'Amen et al. (2015) used a probability ranking system where n species with the highest likelihood of presence were selected up to species richness = n to constrain species assemblages at a site.

The results of this dissertation can assist in informing the macroecological constraints and EARs portions of the SESAM. Beta diversity of tree communities was demonstrated to be higher when the difference in climate, difference in soil texture, bird beta diversity, mammal beta diversity, and the variance in species' range margins between sites are higher (Chapter 3). These results on beta diversity could be used to constrain the differences in species assemblages between sites when the change in those variables over space is known. Furthermore, if relationships between beta diversity and the variables above are temporally consistent, these findings could additionally be utilized as macroecological constraints when projecting S-SDMs under future climate scenarios.

The fourth step of SESAM, the creation of EARs, is the most difficult to achieve due to lack of spatial research on the effect of interactions on species' geographic ranges (Guisan and Rahbek 2011; D'Amen et al. 2015). Here I have quantified the indirect interspecific interactions between species, described how those interactions change across species' geographic ranges, and demonstrated that dominant species experience interactions different from rare species (Chapter 4). I believe these findings could improve upon the probability ranking system used by D'Amen et al. (2015). One could use overall likelihood of species presence across a study area to determine the relative dominance or rarity of individual species in the regional species pool. When stacking SDMs, one could alter the likelihood of occurrence values to account for competition or facilitation depending on the rarity of the species and proximity to its range margin, thereby implicitly incorporating interspecific interactions in SESAM.

Results on how interactions affect species performance across their geographic range additionally inform broader discussion of the species niche. The intersection of abiotic preferences, the ability to reach sites where those preferences exist, and the ability to persist at those sites amongst the community through facilitative or competitive interactions is the realized species niche (Peterson et al. 2011). Due to more frequent research dedicated to facilitative interspecific interactions, Bruno et al. (2003) have proposed reshaping conceptions of realized versus fundamental niches. If individuals benefit from facilitative interactions via amelioration of adverse environmental conditions, the realized niche of the species may extend beyond that of the fundamental niche. Here I have provided spatial evidence that the realized niche of a species is expanded through net facilitative interactions beyond species' range margins (Chapter IV). Determining mechanisms for why some species face an increase in competitive interactions at range margins, despite the community at large experiencing net facilitative interactions, may be a worthwhile endeavor. Moreover, elucidating the differences between species that respond monotonically versus unimodally to stress gradients would provide better predictions of species assemblages across space and through time.

Tables and Figures

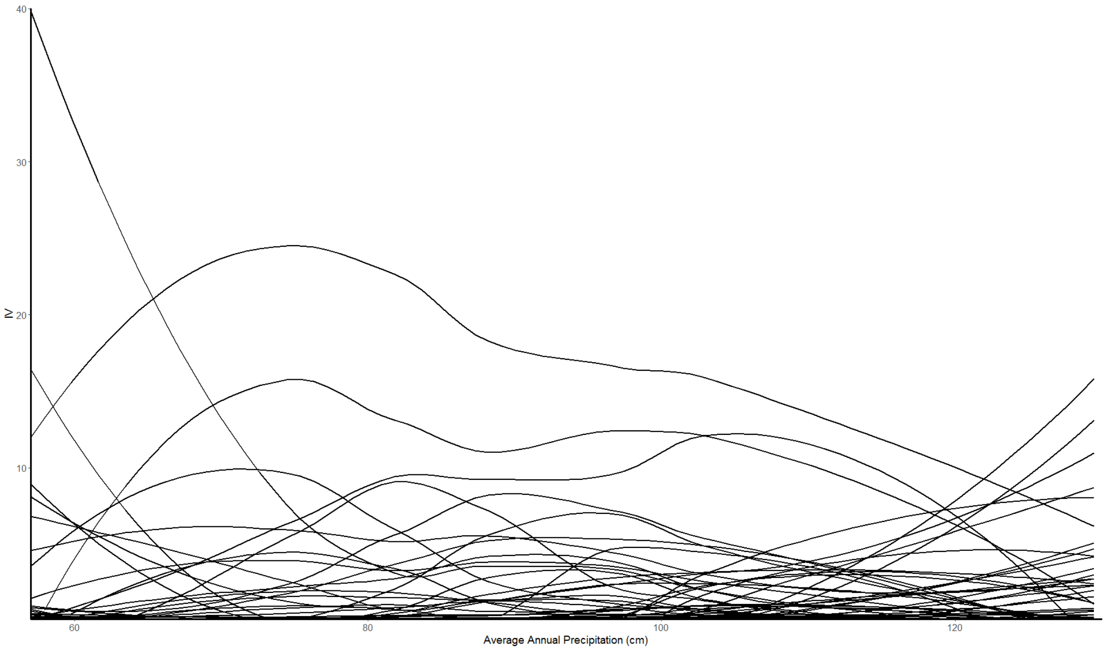


Figure 5.1. Coenoclines of all bottomland species across the average annual precipitation gradient. Regression lines were created using a GAM (Hastie and Tibshirani 1990)

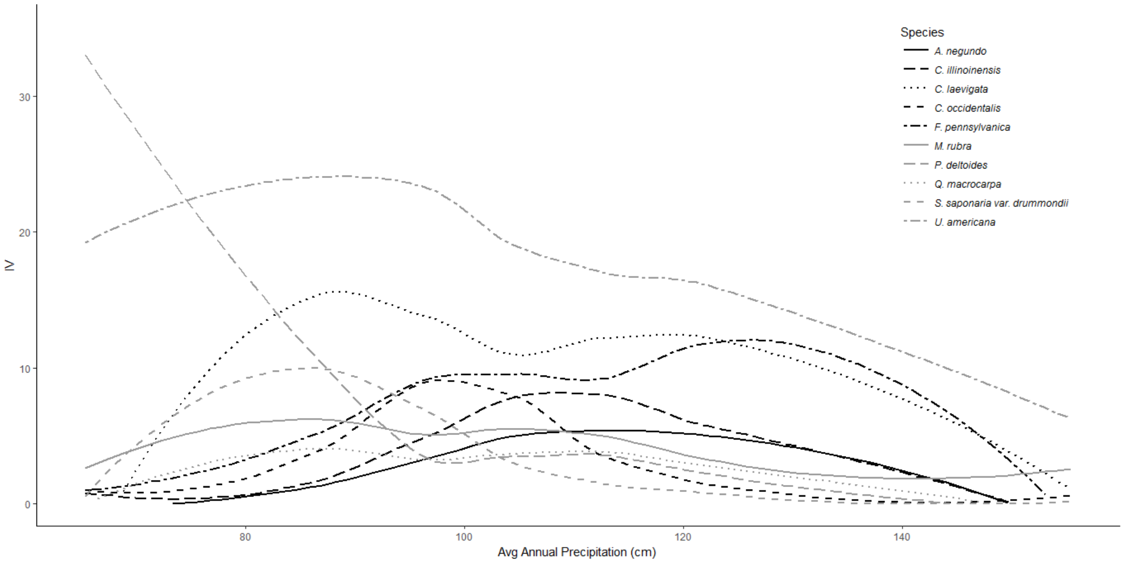
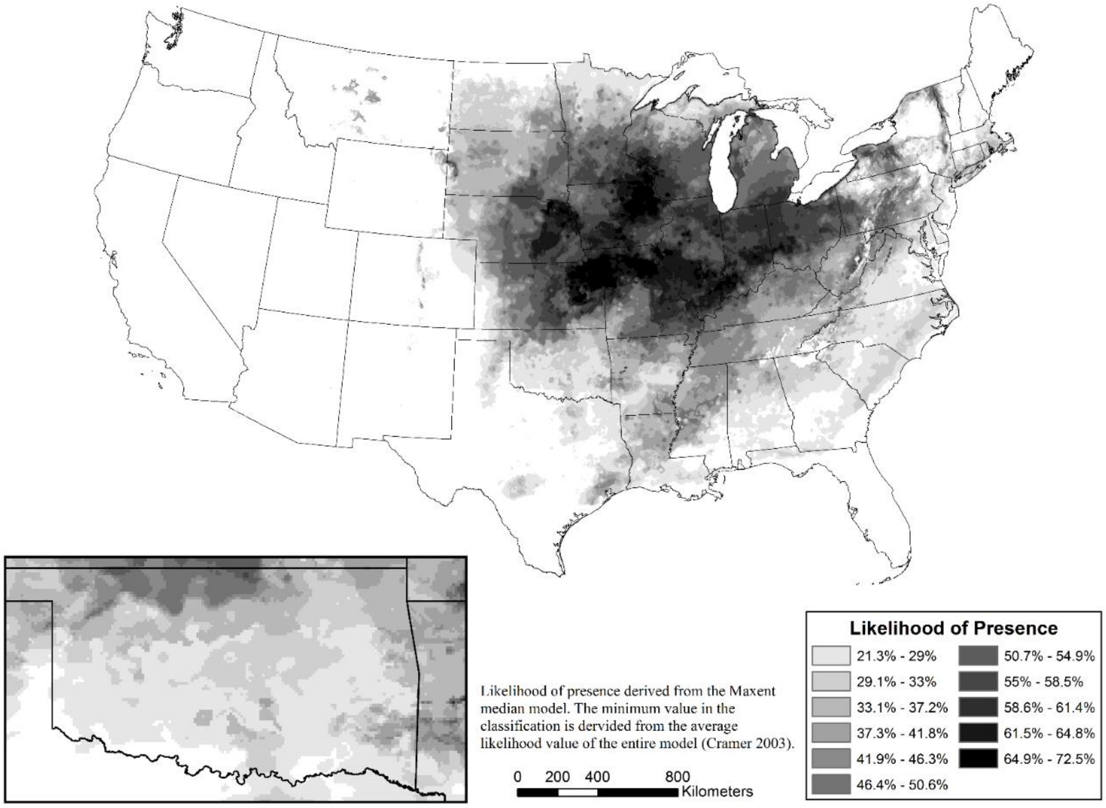


Figure 5.2. Coenoclines of the ten most common bottomland species across the average annual precipitation gradient. Regression lines were created using a GAM (Hastie and Tibshirani 1990)



Likelihood of presence derived from the Maxent median model. The minimum value in the classification is dervided from the average likelihood value of the entire model (Cramer 2003).

Figure 5.3. Modeled species geographic range of *Populus deltoides* using the median likelihood of presence output from MaxEnt

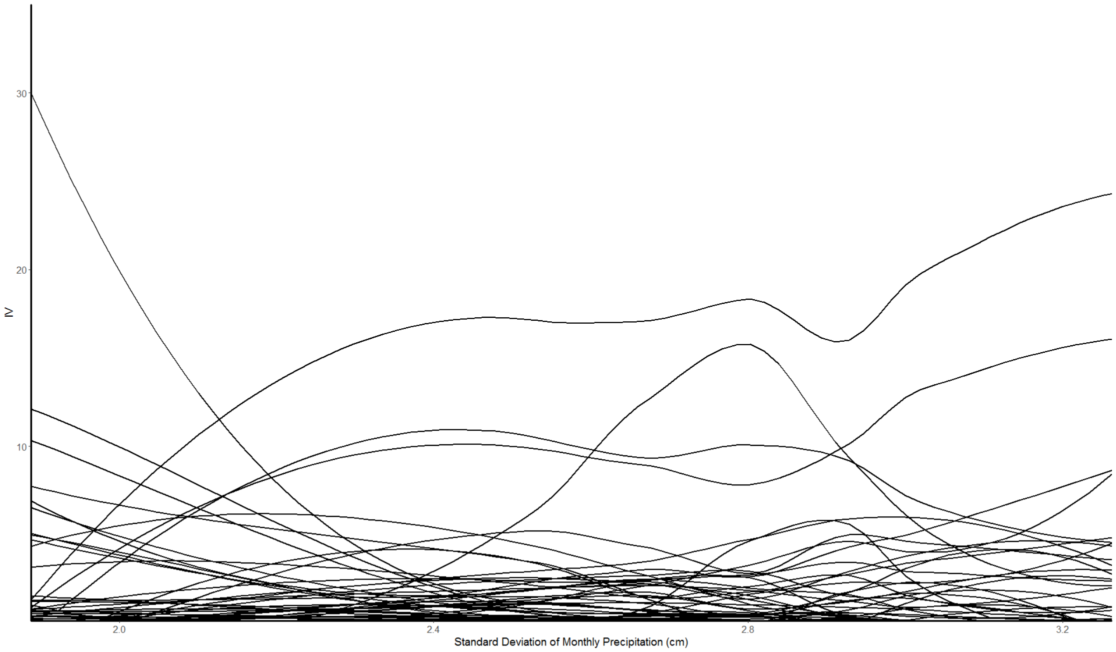


Figure 5.4. Coenoclines of all bottomland species across the standard deviation of monthly precipitation gradient. Regression lines were created using a GAM (Hastie and Tibshirani 1990)

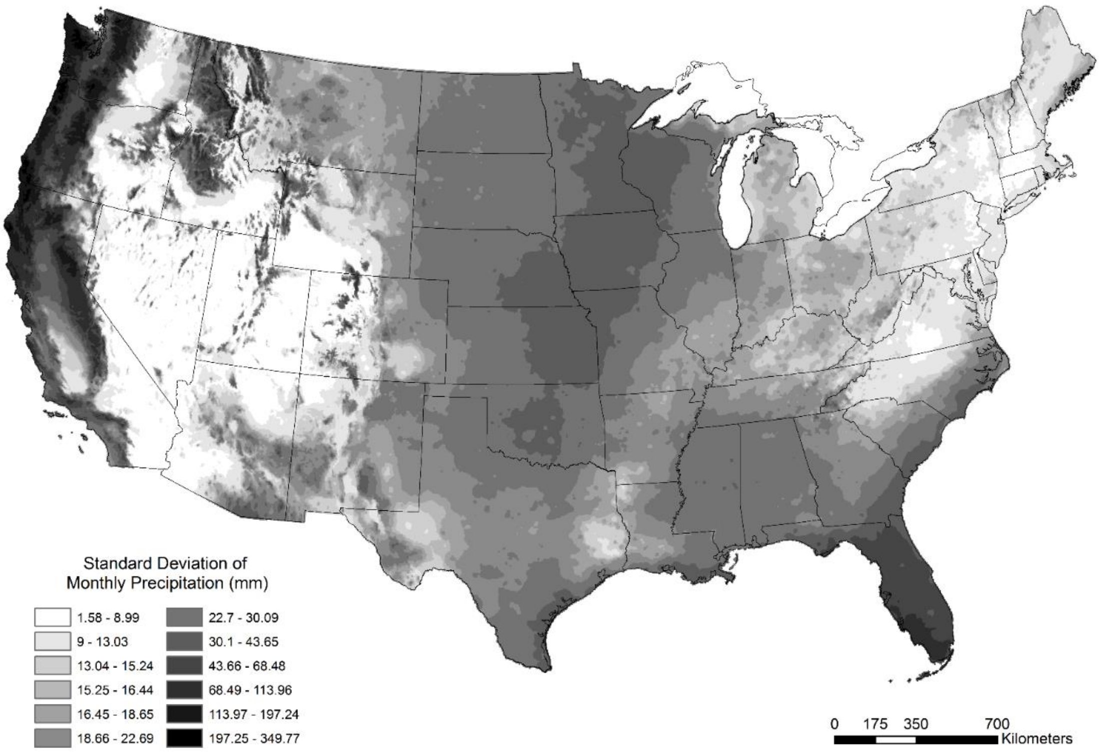


Figure 5.5. Standard deviation of monthly precipitation for the period of 1953-1982 (PRISM data)

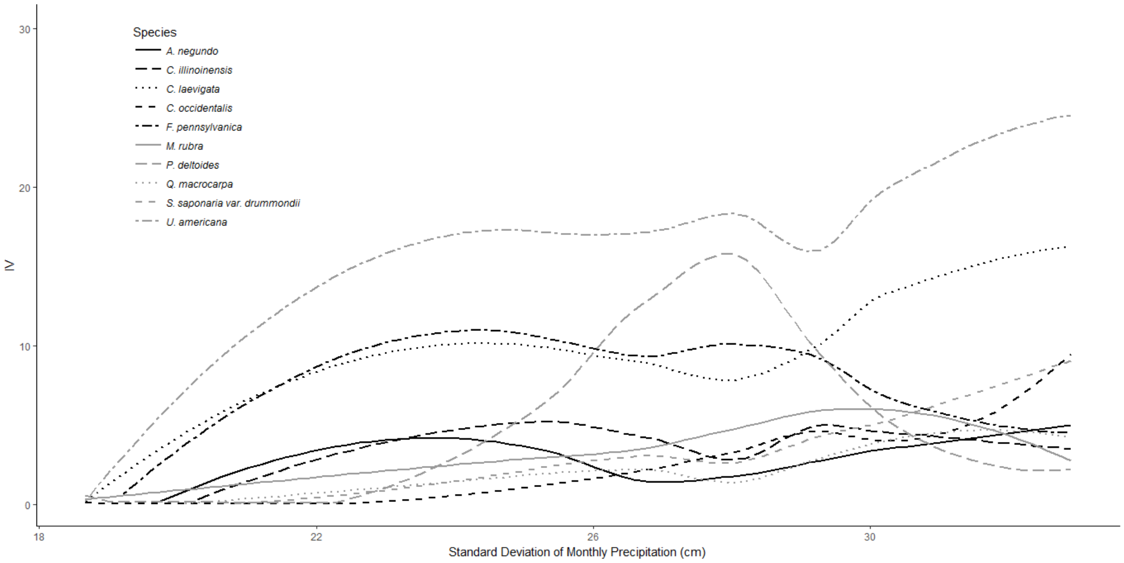


Figure 5.6. Coenoclines of the ten most common bottomland species across the standard deviation of monthly precipitation gradient. Regression lines were created using a GAM (Hastie and Tibshirani 1990)

References Cited

Anderson MJ, Crist TO, Chase JM, Velland M, Inouye BD, Freestone AL, Sanders NJ, Cornell HV, Comita LS, Davies KF, Harrison SP, Kraft NJB, Stegen JC, Swenson NG (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology Letters 14: 19-28.

Arellano G, Tello JS, Jørgensen PM, Fuentes AF, Loza MI, Torrez V, Macía MJ (2016) Disentangling environmental and spatial processes of community assembly in tropical forests from local to regional scales. Oikos 125: 326-335.

Armas C, Rodríguez-Echeverría S, Pugnaire FI (2011) A field test of the stress-gradient hypothesis along an aridity gradient. Journal of Vegetation Science 22: 818-827.

Arroyo MTK, Chacon P, Cavieres LA (2006) Relationship between seed bank expression, adult longevity and aridity in species of Chaetanthera (Asteraceae) in Central Chile. Annals of Botany 98: 591-600.

Arroyo-Rodríguez V, Rös M, Escobar F, Melo FPL, Santos BA, Tabarelli M, Chazdon R (2013) Plant β-diversity in fragmented rain forests: testing floristic homogenization and differentiation hypotheses. Journal of Ecology 101: 1449-1458.

Beals EW (1984) Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. Advances in Ecological Research 14: 1-55.

Bechtold WA, Patterson PL (2005) The enhanced forest inventory and analysis program-national sampling design and estimation procedures. US Forest Service General Technical Report SRS-80.

Bell G (2001) Neutral Macroecology. Science 293: 2413-2418.

Bell G (2003) The interpretation of biological surveys. Proceedings of the Royal Society of London 270: 2531-2542.

Bertness MD, Callaway R (1994) Positive interactions in communities. TRENDS in Ecology and Evolution 9: 191-193.

Bonanomi G, Incerti G, Mazzoleni S (2011) Assessing occurrence, specificity, and mechanisms of plant facilitation in terrestrial ecosystems. Plant Ecology 212: 1777-1790.

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. Ecology 73: 1045-1055.

Bowker MA, Soliveres S, Maestre FT (2010) Competition increases with abiotic stress and regulates the diversity of biological soil crusts. Journal of Ecology 98: 551-560.

Bradshaw AD (1965) Evolutionary significance of phenotypic plasticity in plants. Advances in Genetics 13: 115-155.

Brooker RW, Maestre FT, Callaway RM, Lortie CL, Cavieres LA, Kunstler G, Liancourt P, Tielbörger K, Travis JMJ, Anthelme F, Armas C, Coll L, Corcket E, Delzon S, Forey E, Kikvidze Z, Olofsson J, Pugnaire F, Quiroz CL, Saccone P, Schiffers K, Seifan M, Touzard B, Michalet R (2008) Facilitation in plant communities: the past, the present, the future. Journal of Ecology 96: 18-34.

Brown JH (1984) On the Relationship between Abundance and Distribution of Species. The American Naturalist 124: 255-279.

Brown JH (1995) Macroecology. University of Chicago Press, Chicago, IL.

Bruno JF, Stachowicz JJ, Bertness MD (2003) Inclusions of facilitation into ecological theory. TRENDS in Ecology and Evolution 18: 119-125.

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL (2008) Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. Proceedings of the National Academy of Sciences 105: 11505-11511.

Buchanan KL (2000) Stress and the evolution of condition-dependent signals. TRENDS in Ecology and Evolution 15: 156-160.

Burns RM, Honkala BH (1990) Silvics of North America: 1. Conifers; 2. Hardwoods. Agriculture Handbook 654. US Department of Agriculture, Forest Service, Washington DC. Vol 2, 877 p.

Callaway RM (1995) Positive interactions among plants. Botanical Review 61: 306-349.

Callaway RM, Brooker RW, Choler P, Kikvidze Z, Lortie CJ, Michalet R, Paolini L, Pugnaire FI, Newingham B, Aschehoug ET, Armas C, Kikodze D, Cook BJ (2002) Positive interactions among alpine plants increase with stress. Nature 417: 844-848.

Castanho CT, Oliveira AA, Prado PIKL (2015) Does extreme environmental severity promote plant facilitation? An experimental field test in a subtropical coastal dune. Oecologia 178: 855-866.

Cavender-Bares J, Keen A, Miles B (2006) Phylogenetic Structure of Floridian Plant Communities Depends on Taxonomic and Spatial Scale. Ecology 87: S109-S122.

Cavieres LA, Sierra-Almeida A (2012) Facilitative interactions do not wane with warming at high elevations in the Andes. Oecologia 170: 575-584.

Chalcraft DR, Williams JT, Smith MD, Willig, MR (2004) Scale Dependence in the Species-Richness-Productivity Relationship: The Role of Species Turnover. Ecology 85: 2701-2708.

Chase JM (2010) Stochastic Community Assembly Causes Higher Biodiversity in More Productive Environments. Science 328: 1388-1391.

Chase JM, Leibold MA (2003) Ecological Niches. University of Chicago Press, Chicago, IL.

Choler P, Michalet R, Callaway RM (2001) Facilitation and Competition on Gradients in Alpine Plant Communities. Ecology 82: 3295-3308.

Chu C, Maestre FT, Xiao S, Weiner J, Wang Y, Duan Z, Wang G (2008) Balance between facilitation and resource competition determines biomass-density relationships in plant populations. Ecology Letters 11: 1189-1197.

Clark JS, Silman M, Kern R, Macklin E, HilleRisLambers J (1999) Seed dispersal near and far: patterns across temperate and tropical forests. Ecology 80: 1475-1494.

Clements FK (1916) Plant Succession: An Analysis of the Development of Vegetation. No. 242. Carnegie Institution of Washington.

Collins SL, Glenn SM, Roberts DW (1993) The hierarchical continuum concept. Journal of Vegetation Science 4: 149-156.

Condit R, Pitman N, Leigh Jr EG, Chave J, Terborgh J, Foster RB, Núñez P, Aguilar S, Valencia R, Villa G, Muller-Landau HC, Losos E, Hubbell SP (2002) Beta-Diversity in Tropical Forest Trees. Science 295: 666-669.

Connell JH, Tracey JG, Webb LJ (1984) Compensatory Recruitment, Growth, and Mortality as Factors Maintaining Rain Forest Tree Diversity. Ecological Monographs 54: 141-164.

Cottenie K (2005) Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8: 1175-1182.

Cramer JS (2003) Logit models: from economics and other fields. Cambridge University Press.

D’Amen M, Dubuis A, Fernandes RF, Pottier J, Pellissier L, Guisan A (2015) Using species richness and functional traits to constrain assemblage predictions from stacked species distribution models. Journal of Biogeography 42: 1255-1266.

Dai A (2012) Increasing drought under global warming in observations and models. Nature Climate Change 3: 52-58.

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology 28: 2031-2064.

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

Dawson TE (1993) Hydraulic lift and water use by plants: implications for water balance, performance and plant-plant interactions. Oecologia. 95: 565-574.

Derroire G, Tigabu M, Odén PC, Healey JR (2016) The effects of established trees on woody regeneration during secondary succession in tropical dry forests. Biotopica 48: 290-300.

Diniz-Filho JAF, Barbosa ACOF, Collevatti RG, Chaves LJ, Terribile LC, Lima-Ribeiro MS, Telles MPC (2016) Spatial autocorrelation analysis and ecological niche modelling allows inference of range dynamics driving the population genetic structure of a Neotropical savanna tree. Journal of Biogeography 43: 167-177.

Dobzhansky T (1950) Evolution in the tropics. American Scientist 38: 209-221.

Dohn J, Dembélé F, Karembé M, Moustakas A, Amévor KA, Hanan NP (2013) Tree effects on grass growth in savannas: competition, facilitation and the stress-gradient hypothesis. Journal of Ecology 101: 202-209.

Dubuis A, Pottier J, Rion V, Pellissier L, Theurillat JP, Guisan A (2011) Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. Diversity and Distributions 17: 1122-1131.

Dunning JB, Danielson BJ, Pulliam HR (1992) Ecological Processes That Affect Populations in Complex Landscapes. Oikos 65: 169-175.

Edwards W, Westoby M (1996) Reserve mass and dispersal investment in relation to geographic range of plant species: phylogenetically independent contrasts. Journal of Biogeography 23: 329-338.

Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17: 43-57.

Forster JR (1778) Observations made during a voyage round the world, on physical geography, natural history, and ethic philosophy. G. Robinson, London.

García-Valdés R, Gotelli NJ, Zavala MA, Purves DW, Araújo MB (2015) Effects of climate, species interactions, and dispersal on decadal colonization and extinction rates of Iberian tree species. Ecological Modelling 309-310: 118-127.

Gaston KJ (2003) The structure and dynamics of Geographic Ranges. Oxford University Press, New York.

Gaston KJ, Blackburn TM, Spicer JI (1998) Rapoport's rule: time for an epitaph? TRENDS in Ecology and Evolution 13: 70-74.

Gauch Jr HG (1982) Multivariate Analysis in Community Ecology. Cambridge University Press.

Gauch Jr HG, Wentworth TR (1976) Canonical correlation analysis as an ordination technique. Vegetatio 33: 17-22.

Gauch Jr HG, Whittaker RH (1972) Coenocline Simulation. Ecology 53: 446-451.

Gibson N, Prober S, Meissner R, van Leeuwen S (2017) Implications of high species turnover on the south-western Australian sandplains. PLoS ONE 12: e0172977.

Gleason HA (1926) The individualistic concept of the plant association. Bulletin of the Torrey Botanical Club 7-26.

Gotelli, NJ (2001) Research Frontiers in null model analysis. Global Ecology and Biogeography 10: 337-343.

Graff P, Aguiar MR (2011) Testing the role of biotic stress in the stress gradient hypothesis. Processes and patterns in arid rangelands. Oikos 120: 1023-1030.

Grime JP (1979) Plant Strategies and Vegetation Processes. John Wiley & Sons.

Grime JP (1989) The stress debate: symptom of impending synthesis? Biological Journal of the Linnean Society 37: 3-17.

Grosenbaugh LR (1952) Plotless timber estimates new, fast, easy. Journal of Forestry 50: 32-37.

Guillot G, Rousset F (2013) Dismantling the Mantel tests. Methods in Ecology and Evolution 4: 336-344.

Guisan A, Rahbek C (2011) SESAM – a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. Journal of Biogeography 38: 1433-1444.

Hanski I (1998) Metapopulation dynamics. Nature 396: 41-49.

Hargreaves AL, Samis KE, Eckert CG (2014) Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. The American Naturalist 183: 157-173.

Hassell MP, Comins HN, May RM (1994) Species coexistence and self-organizing spatial dynamics. Nature 370: 290-292.

Hastie TJ, Tibshirani R (1990) Generalized additive models. Chapman and Hall, London.

Hoagland BW, Buthod A, Butler I, Callahan-Crawford P, Elisens W, Tyrl R (2012) Oklahoma vascular plants database. Oklahoma Natural Heritage Inventory, Oklahoma Biological Survey, Norman, Oklahoma. http://www.oklahomaplantdatabase.org. Accessed 25 January 2017.

Hoagland BW, Collins SL (1997) Gradient models, gradient analysis, and hierarchical structure in plant communities. Oikos 78: 23-30.

Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. Nature 470: 479-485.

Holmgren M, Scheffer M (2010) Strong facilitation in mild environments: the stress gradient hypothesis revisited. Journal of Ecology 98: 1269-1275.

Hubbell SP (2001) The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press.

Hutchinson GE (1957) Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology 22: 415-457.

Johnson F (1982) Field survey of bottomland forests in Oklahoma. Unpublished raw data.

Johnson KS (2008) Geologic history of Oklahoma. Norman, OK: Oklahoma Geological Survey, Educational Publication 9.

Jones MM, Gibson N, Yates C, Ferrier S, Mokany K, Williams KJ, Manion G, Svenning J (2016) Underestimated effects of climate on plant species turnover in the Southwest Australian Floristic Region. Journal of Biogeography 43: 289-300.

Jones NT, Gilbert B (2016) Biotic Forcing: the push-pull of plant ranges. Plant Ecology 217: 1331-1344.

Jones E, Oliphant E, Peterson P (2014) SciPy: Open Source Scientific Tools for Python. http://www.scipy.org/ [Online; accessed 2016-6-1].

Jonsson M, Snäll T, Asplund J, Clemmensen KE, Dahlberg A, Kumordzi BB, Lindahl BD, Oksanen J, Wardle DA (2016) Divergent responses of β-diversity among organism groups to a strong environmental gradient. Ecosphere 7: e01535.10.1002/ecs2.1535.

Jucker T, Bouriaud O, Coomes DA (2015) Crown plasticity enables trees to optimize canopy packing in mixed species forests. Functional Ecology 29: 1078-1086.

Karron, JD (1987) A comparison of levels of genetic polymorphism and self-compatibility in geographically restricted and widespread plant congeners. Evolutionary Ecology 1: 47-58.

Koleff P, Gaston KJ, Lennon JJ (2003) Measuring beta diversity for presence-absence data. Journal of Annual Ecology 72: 367-382.

Körner C (2003) Limitation and stress – always or never? Journal of Vegetation Science 14: 141-143.

Legendre P, Fortin M, Borcard D (2015) Should the Mantel test be used in spatial analysis? Methods in Ecology and Evolution 6: 1239-1247.

Leibold MA, Holyoak M, Monquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Shurin JB, Law R, Tilman D, Loreau M, Gonzalez A (2004) The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7: 601-613.

Lemly JM, Cooper DJ (2011) Multiscale factors control community and species distribution in mountain peatlands. Botany 89: 689-713.

Liancourt P, Callaway RM, Michalet R (2005) Stress tolerance and competitive-response ability determine the outcome of biotic interactions. Ecology 86: 1611-1618.

Little EL (1981) Forest Trees of Oklahoma. Oklahoma Department of Agriculture Forestry Services.

López RP, Squeo FA, Armas C, Kelt DA, Gutiérrez JR (2016) Enhanced facilitation at the extreme end of the aridity gradient in the Atacama Desert: a community-level approach. Ecology 97: 1593-1604.

Lortie CJ, Callaway RM (2006) Re-analysis of meta-analysis: support for the stress-gradient hypothesis. Journal of Ecology 94: 7-16.

MacArthur RH (1957) On the relative abundance of bird species. Proceedings of the National Academy of Science 43: 293-295.

MacArthur RH (1972) Geographical Ecology: patterns in the distribution of species. Princeton University Press.

Maestre FT, Callaway RM, Valladares F, Lortie CJ (2009) Refining the stress-gradient hypothesis for competition and facilitation in plant communities. Journal of Ecology 97: 199-205.

Maestre FT, Valladares F, Reynolds JF (2005) Is the change of plant-plant interactions with abiotic stress predictable? A meta-analysis of field results in arid environments. Journal of Ecology 93: 748-757.

Maestre FT, Valladares F, Reynolds JF (2006) The stress-gradient hypothesis does not fit all relationships between plant-plant interactions and abiotic stress: further insights from arid environments. Journal of Ecology 94: 17-22.

Markham JH, Chanway CP (1996) Measuring Plant Neighbor Effects. Functional Ecology 4: 548-549.

McCune B, Mefford MJ (2011) PC-ORD. Multivariate Analysis of Ecological Data. Version 6. MjM Software, Gleneden Beach, Oregon, USA.

McIntire EJB, Fajardo A (2011) Facilitation within Species: A Possible Origin of Group-Selected Superorganisms. The American Naturalist 178: 88-97.

McIntire EJB, Fajardo A (2014) Facilitation as a ubiquitous driver of biodiversity. New Phytologist 201: 403-416.

Meier ES, Edwards Jr. TC, Kienast F, Dobbertin M, Zimmermann NE (2011) Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. Journal of Biogeography 38: 371-382.

Meier ES, Kienast F, Pearman PB, Svenning J, Thuiller W, Araújo MB, Guisan A, Zimmermann NE (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. Ecography 33: 1038-1048.

Michalet R, Brooker RW, Cavieres LA, Kikvidze Z, Lortie CJ, Pugnaire FI, Valiente-Banuet A, Callaway RM (2006) Do biotic interactions shape both sides of the humped-back model of species richness in plant communities? Ecology Letters 9: 767-773.

Mouquet N, Loreau M (2002) Coexistence in Metacommunities: The Regional Similarity Hypothesis. The American Naturalist 4: 420-426.

Mouquet N, Loreau M (2003) Community Patterns in Source-Sink Metacommunities. The American Naturalist 162: 544-557.

Nathan R, Muller-Landau HC (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. TRENDS in Ecology and Evolution 15: 278-285.

Nathan R, Schurr FM, Spiegel O, Steinitz O, Trakhtenbrot A, Tsoar A (2008) Mechanisms of long-distance seed dispersal. TRENDS in Ecology and Evolution 23: 638-647.

Nguyen DT, Gómez-Zurita J (2016) Subtle ecological gradient in the tropics triggers high species-turnover in a local geographical scale. PLoS ONE 11: e0156840.

O’Brien EM (1993) Climatic Gradients in Woody Plant Species Richness: Towards an Explanation Based on an Analysis of Southern Africa’s Woody Flora. Journal of Biogeography 20: 181-198.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2017) vegan: Ecology Package. R package version 2.4-1. https://CRAN.R-project.org/package=vegan.

Omernik JM, Griffith GE (2014) Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management 54: 1249-1266.

Pardieck KL, Ziolkowski Jr DJ, Hudson MAR, Campbell K (2016) North American Breeding Bird Survey Dataset 1966 - 2015, version 2015.1. U.S. Geological Survey, Patuxent Wildlife Research Center. <www.pwrc.usgs.gov/BBS/RawData/>; doi:10.5066/F7C53HZN.

Patterson BD, Ceballos G, Sechrest W, Tognelli MF, Brooks T, Luna L, Ortega P, Salazar I, Young BE (2007) Digital Distribution Maps of the Mammals of the Western Hemisphere, version 3.0. NatureServe, Arlington, Virginia, USA.

Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology and Biogeography 12: 361-371.

Pellissier L, Bråthen A, Pottier J, Randin F, Vittoz P, Dubuis A, Yoccoz NG, Alm T, Zimmermann NE, Guisan A (2010) Species distribution models reveal apparent competitive and facilitative effects of a dominant species on the distribution of tundra plants. Ecography 33: 1004-1014.

Peterson AT, Soberón J, Pearson RG, Anderson RP, Martínez-Meyer E, Nakamura M, Araújo MB (2011) Ecological Niches and Geographic Distributions. Princeton University Press.

Pielou EC (1984) The interpretation of ecological data: a primer on classification and ordination. John Wiley & Sons.

Poos MS, Jackson DA (2012) Addressing the removal of rare species in multivariate bioassessments: The impact of methodological choices. Ecological Indicators 18: 82-90.

Preston FW (1948) The commonness, and rarity, of species. Ecology 29: 254-283.

PRISM Climate Group (2017) Oregon State University. http://www.prism.oregonstate.edu/ created 1 Jun 2016.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rapoport, EH (1982) Aerography: geographical strategies of species. Pergamon, Oxford.

Rey PJ, Alcántara JM, Manzaneda AJ, Sánchez-Lafuente AM (2016) Facilitation contributes to Mediterranean woody plant diversity but does not shape the diversity-productivity relationship along aridity gradients. New Phytologist 211: 464-476.

Rice EL, Penfound WT (1955) An evaluation of the variable-radius and paired-tree methods in the Blackjack-Post Oak forest. Ecology 36: 315-320.

Rice EL, Penfound WT (1959) The upland forests of Oklahoma. Ecology 40: 593-608.

Richardson PJ, MacDougall AS, Stanley AG, Kaye TN, Dunwiddie PW (2012) Inversion of plant dominance-diversity relationships along a latitudinal stress gradient. Ecology 93: 1431-1438.

Ricklefs RE (1987) Community diversity: relative roles of local and regional processes. Science 235: 167-171.

Rysavy A, Seifan M, Sternberg M, Tielbörger K (2016) Neighbour effects on shrub seedling establishment override climate change impacts in a Mediterranean community. Journal of Vegetation Science 27: 227-237.

Schmida A, Wilson MV (1985) Biological determinants of species diversity. Journal of Biogeography 12: 1-20.

Schoener TW (1987) The geographical distribution of rarity. Oecologia 74: 161-173.

Simard S, Vyse A (2006) Trade-offs between competition and facilitation: a case study of vegetation management in the interior cedar-hemlock forests of southern British Columbia. Canadian Journal of Forest Research 36: 2486-2496.

Smit C, Rietkerk M, Wassen MJ (2009) Inclusion of biotic stress (consumer pressure) alters predictions from the stress gradient hypothesis. Journal of Ecology 97: 1215-1219.

Snyder SA (1991) Cephalanthus occidentalis. In: Fire Effects Information System [online]. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory. Available online at http://www.fs.fed.us/database/feis/. Accessed Mar 25 2017.

Soil Survey Staff (2017) Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/. Accessed June 1 2016.

Soininen J (2016) Spatial structure in ecological communities – a quantitative analysis. Oikos 125: 160-166.

Soininen J, McDonald R, Hillebrand H (2007) The distance decay of similarity in ecological communities. Ecography 30: 3-12.

Soliveres S, DeSoto L, Maestre FT, Olano JM (2010) Spatio-temporal heterogeneity in abiotic factors modulate multiple ontogenetic shifts between competition and facilitation. Perspectives in Plant Ecology, Evolution and Systematics 12: 227-234.

Soliveres S, Maestre FT (2014) Plant-plant interactions, environmental gradients and plant diversity: a global synthesis of community-level studies. Perspectives in Plant Ecology, Evolution and Systematics 16: 154-163.

Stevens, GC (1989) The latitudinal gradient in geographical range: how so many species co-exist in the tropics. The American Naturalist 133: 240-256.

Svenning JC, Gravel D, Holt RD, Schurr FM, Thuiller W, Münkemüller T, Schiffers KH, Dullinger S, Edwards TC, Hickler T, Higgins SI, Nabel JEMS, Pagel J, Normand S (2014) The influence of interspecific interactions on species range expansion rates. Ecography 37: 1198–1209.

Swenson NG, Enquist BJ, Thompson J, Zimmerman JK (2007) The Influence of Spatial and Size Scale on Phylogenetic Relatedness in Tropical Forest Communities. Ecology 88: 1770-1780.

Tahvanainen JO, Root RB (1972) The influence of vegetational diversity on the population ecology of a specialized herbivore, *Phyllotreta cruciferae* (Coleoptera: Chrysomelidae). Oecologia 10: 321-346.

Tálamo A, Barchuk A, Cardozo S, Trucco C, Marás G, Trigo C (2015) Direct versus indirect facilitation (herbivore mediated) among woody plants in a semiarid Chaco forest: a spatial association approach. Austral Ecology 40: 573-580.

Thomas JA, Bourn NAD, Clarke RT, Stewart KE, Simcox DJ, Pearman GS, Curtis R, Goodger B (2001) The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. Proceedings of the Royal Society B 268: 1791-1796.

Tilman D (1982) Resource competition and community structure. Princeton University Press.

Tilman D (1994) Competition and Biodiversity in Spatially Structured Habitats. Ecology 75: 2-16.

Urban MC, Tewksbury JJ, Sheldon KS (2012) On a collision course: competition and dispersal differences create no-analogue communities and cause extinctions during climate change. Proceedings of the Royal Society B 279: 2072–2080.

Valiente-Banuet A, Verdú M (2008) Temporal shifts from facilitation to competition occur between closely related taxa. Journal of Ecology 96: 489-494.

Van Rossum G (2007) Python Programming Language. In USENIX Annual Technical Conference. 41: 36.

Villarreal-Barajas T, Martorell C (2009) Species-specific disturbance tolerance, competition and positive interactions along an anthropogenic disturbance gradient. Journal of Vegetation Science 20: 1027-1040.

von Humboldt A, Bonpland A (1807) Essay on the geography of plants. Paris: Fr. Schoell.

Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs 30: 279-338.

Whittaker RH (1962) Classification of Natural Communities. Botanical Review 28: 1-239.

Whittaker RH (1965) Dominance and diversity in land plant communities. Science 147: 250-260.

Whittaker RH (1967) Gradient Analysis of Vegetation. Biological Reviews 42: 207-264.

Woods AJ, Omernik JM, Butler DR, Ford JG, Henley JE, Hoagland BW, Arndt DS, Moran BC (2005) Ecoregions of Oklahoma [color map, descriptive text, summary tables, and photographs]. Reston, Virginia, US Geological Survey (map scale 1: 1,250,000).

Zellweger F, Baltensweiler A, Ginzler C, Roth T, Braunisch V, Bugmann H, Bollmann K (2016) Environmental predictors of species richness in forest landscapes: abiotic factors versus vegetation structure. Journal of Biogeography 43: 1080-1090.

Appendix A: Python Scripts

```
# Name: CompareIV.py
# Author: Daryn Hardwick
# Description: This script chooses nearest neighbors from sites where both the focal
# species and another species occur together and where only the focal species occurs,
# calculates the average IV using an inverse distance weighted algorithm, and prints
# the site, where or not the other species is present, the average IV value, and the
# distance from the site to the focal species range margin.

## Variable List ##
# dwa_iv() - the inverse distance weighted algorithm
# Lists - a python script that contains lists of presence/absence sites per species
# IV - the site x species matrix of importance values
# Distance - the site x site matrix of distances between sites
# Range - the species x site matrix of distances from sites to species range margins
# SharedPres - list that will contain all sites in which the focal species and other species
# co-occur
# sitedf - subset of "Distance" dataframe that contains distances from a focal site to all
# other sites in "SharedPres" or "FocalOnly"
# sitelist - list of site to site distances from "sitedf", sorted from shortest to longest
# distance
# s1,s2,s3 - focal site, nearest neighbor site, and 2nd nearest neighbor site
```

```
# d1,d2,d3 - distance to focal site (always 0), distance to nearest neighbor site, distance
# to 2nd nearest neighbor site
# IVdf - subset of "IV" dataframe that contains the IV's of the focal species at the 3
# nearest neighbor sites
# iv1,iv2,iv3 - IV of focal species at focal site, IV of focal species at nearest neighbor
# site, IV of focal species at 2nd nearest neighbor site
# IVval - variable that calls the inverse distance weighted calculation
# Rangedist - subset of "Range" dataframe that contains the distance to the focal
# species range margin from the focal site
# FocalOnly - list that will contain all site in which only the focal species occurs

#Import module - pandas for dataframe manipulation
import pandas as pd

#Distance Weighted Average function
#Takes 3 IV's and 3 distances as inputs
def dwa_iv(iv1,iv2,iv3,d1,d2,d3):
    dwa = (iv1*(1/d1)+iv2*(1/d2)+iv3*(1/d3))/((1/d1)+(1/d2)+(1/d3));
    return dwa;

#Import species presence and absence lists
#For ACNE, ACNEPres is a list of sites where ACNE is present, ACNEAbse is a list of
#site where ACNE is absent
```

```
from Lists import *


#Import species x sites IV matrix, site x site distance matrix, and site x species margin
#matrix
IV = pd.read_csv('.../IVMatrix.csv',index_col=0)
Distance = pd.read_csv('.../DistanceMatrix.csv',index_col=0)
Range = pd.read_csv('.../SiteMatrix_toRangeMargins.csv',index_col=0)


#Create list of shared presence between species
#ACNE is the focal species, ACSA is the other species in this run
SharedPres = []
for firstsite in ACNEPres:
    for secondsite in ACSAPres:
        if firstsite == secondsite:
            SharedPres.append(firstsite);


#Returns site, distance weighted average IV, and distance to nearest 2 neighbors for
#sites where both species exist
if len(SharedPres) > 2:
    for site in SharedPres:
        sitedf = Distance.loc[site,SharedPres]
        sitedf = sitedf.sort_values()
        sitedf = sitedf[0:3]
```

```
        sitelist = sitedf.index.tolist()
        s1 = sitelist[0];s2 = sitelist[1];s3 = sitelist[2]
        d1 = float(sitedf[0])+1; d2 = float(sitedf[1]); d3 = float(sitedf[2])
        IVdf = IV.loc["ACNE",sitelist]
        iv1 = float(IVdf[0]); iv2 = float(IVdf[1]); iv3 = float(IVdf[2])
        IVval = dwa_iv(iv1,iv2,iv3,d1,d2,d3)
        Rangedist = Range.loc[site,"ACNE"]
        print("ACNE","ACSA","Both Present",site,s1,s2,s3,IVval,Rangedist)

#Create list of sites where focal species exists and other species does not
FocalOnly = [ ]
for firstsite in ACNEPres:
    for secondsite in ACSAAbse:
        if firstsite == secondsite:
            FocalOnly.append(firstsite);

#Returns site, distance weighted average IV, and distance to nearest 2 neighbors for
#sites where focal species exists and other species does not
if len(FocalOnly) > 2:
    for site in FocalOnly:
        sitedf = Distance.loc[site,FocalOnly]
        sitedf = sitedf.sort_values()
        sitedf = sitedf[0:3]
```

```
sitelist = sitedf.index.tolist()

s1 = sitelist[0]; s2 = sitelist[1]; s3 = sitelist[2]

d1 = float(sitedf[0])+1; d2 = float(sitedf[1]); d3 = float(sitedf[2])

IVdf = IV.loc["ACNE",sitelist]

iv1 = float(IVdf[0]); iv2 = float(IVdf[1]); iv3 = float(IVdf[2])

IVval = dwa_iv(iv1,iv2,iv3,d1,d2,d3)

Rangedist = Range.loc[site,"ACNE"]

print("ACNE","ACSA","Only ACNE",site,s1,s2,s3,IVval,Rangedist)
```

Code A1. Python code to create distance weighted IV’s for species pairs where both species exist and only the focal species exists for the real community

```
# Name: CreateNullModels.py
# Author: Daryn Hardwick
# Description: This script creates null communities by randomly selecting species up to
# the real world species richness at a site, generating random importance values, and
# iterating this process over all sites in a community matrix. This is then iterated to
# create as many null communities as the user wants.

## Variable List ##
# SiteRichnessCSV - CSV file of species richness at each site, contains 2 fields named
# Site and Richness
# iterations - the number of null models to be created
# path - the location where the null models will be saved to
# UplandSpecies - the species pool for the upland dataset
# BottomlandSpecies - the species pool for the bottomland dataset
# iternum - counter variable to stop script after number of iterations specified has been
# reached
# siteidx - counter variable that keeps track of which site species and IV's are being
# selected for
# richness_df - variable to read in "SiteRichnessCSV"
# sites - list of sites in "SiteRichnessCSV"
# richness - list of species richness values in "SiteRichnessCSV"
# richness_dict - Dictionary that relates "sites" to "richness"
# df - a starting (mostly empty dataframe) with two columns: Species - where rows are
```

```
# the species in the species pool, and Delete - where all rows contain zeros. Dataframes
# in pandas need to have at least 2 columns. The Delete column will be deleted after the
# null model is finished
# IValuesList - list that contains all of the randomly selected IV's
# spList - a list of species randomly selected from the species pool
# d - dictionary that will contain species randomly sampled and IV's randomly selected
# for those species
# x - counter variable that keeps track of how many IV's have been randomly selected
# for a site
# maxval - the maximum importance value that can be randomly selected
# samplelen - the number of species sampled from the species pool
# n - the randomly selected IV
# df2 - a new dataframe generated from dictionary "d" and will be joined to dataframe
# "df"

## -------------------------- Edit these variables -------------------------- ##
SiteRichnessCSV = 'BottomlandRichness.csv' # CSV with sites and their richness
iterations = 100 # The number of iterations you wish to perform
path = "C:/Temp/NullMatrices/Bottomland/" # The output location
## ------------------------------------------------------------------------- ##

# Import modules - random for drawing random numbers, pandas for dataframe
# manipulation
```

```
import random
import pandas as pd

# Create species pools
UplandSpecies = ["ACRU","ACSA","CACO","CAIL","CATE","CATO","CELA",
"CEOC","CERE","CECA","COFL","CRVI","DIVI","FRAM","FRPE","GLTR",
"GYDI","JUNI","JUVI","LIST","MAPO","MORU","NYSY","OSVI","PIEC","PLOC",
"PRME","PRSE","QUAL","QUFA","QUHA","QUMA","QUMR","QUMU","QUNI",
"QUPH","QUSH","QUST","QUVE","RHCA","ROPS","SANI","SASA","SILA",
"ULAL","ULAM","ULRU"]
BottomlandSpecies = ["ACNE","ACRU","ACSA","ACSI","BENI","CACO","CAIL",
"CALA","CATE","CATO","CELA","CEOC","CERE","CEP_OC","CECA","COFL",
"CRVI","DIVI","FRPE","GLTR","GYDI","ILDE","JUNI","JUVI","LIST","MAPO",
"MORU","NYSY","PIEC","PLOC","PODE","PRME","PRSE","QUAL","QUFA",
"QULY","QUMA","QUMU","QUNI","QUPH","QUSH","QUST","RHCA","ROPS",
"SANI","SASA","SILA","TAGA","ULAL","ULAM","ULRU","VIPR"]

# Initialize three loops
iternum = 1
# This while loop is for the number of iterations to be performed
while iternum<=iterations:
    siteidx = 0
    richness_df = pd.read_csv(SiteRichnessCSV)
```

```
sites = richness_df.Site.tolist()
richness = richness_df.Richness.tolist()
richness_dict = {'Site':sites,
            'Richness':richness}
df = pd.read_csv('empty_bland_df.csv')
df = df.set_index('Species')
# This for loop works through each site and samples species based on the richness
for SR in richness:
  d = {}
  x=1
  maxval=100
  IValuesList = []
  spList = random.sample(BottomlandSpecies,SR)
  d['Species'] = spList
  samplelen = len(spList)
  # This while loop generates an IV for each species sampled in the above for loop
  while x<samplelen:
    n = float(random.uniform(0,maxval))
    if n < 0.1:
      n = 0.1
    IValuesList.append(n);
    x=x+1
    maxval = maxval - n
```

```
        else:
            maxval = round(maxval,1)
            if maxval < 0.1:
                maxval = 0.1
            IValuesList.append(maxval);
        d[str(sites[siteidx])] = IValuesList
        df2 = pd.DataFrame(d)
        df2 = df2.set_index('Species')
        df = df.join(df2)
        siteidx = siteidx+1
    # Converts all NaN values to 0, rounds all numbers to one decimal, removes an
    # unnecessary field
    df = df.fillna(0)
    df = df.round(1)
    df = df.drop('Delete',axis=1)
    # Exports data frame to CSV
    df.to_csv(path + "run" + str(iternum) + ".csv",sep=',')
    print("Species matrix " + str(iternum) + " complete!")
    iternum = iternum+1
```

Code A2. Python code to create 100 null communities

```
# Name: CompareIV_NullModels.py
# Author: Daryn Hardwick
# Description: This script chooses nearest neighbors from sites where both the focal
# species and another species occur together and where only the focal species occurs,
# calculates the average IV using an inverse distance weighted algorithm, and prints
# the site, where or not the other species is present, the average IV value, and the
# distance from the site to the focal species range margin. It does so using a folder of
# null community data. The null communities must be set up in sites x species matrices
# of importance values in CSV file format.

## Variable List ##
# dwa_iv() - the inverse distance weighted algorithm
# NullLists - a python script that contains lists of IV's for each species and for each null
# model and "SpLists"
# PA_NullLists - a python script that contains lists of presence/absence sites per species
# and for each null model, "P_SpLists", and "A_SpLists"
# Distance - the site x site matrix of distances between sites
# Range - the species x site matrix of distances from sites to species range margins
# Sites - List of all sites in the bottomland and upland datasets
# y - counter variable for indexing the lists of species IV's in "SpLists"
# SpLists - a list containing lists of species IV's for each null model
# x - counter variable for indexing the lists of presence/absence sites for each species in
# "P_SpLists" and "A_SpLists"
```

```
# P_SpLists - a list containing lists of sites where each species is present for each null
# model
# A_SpLists - a list containing lists of sites where each species is absent for each null
# model
# counter1 - counter variable used for indexing focal species presence lists and the
# other
# species presence lists
# path - the file path of the folder that contains the null model communities in CSV file
# format
# filenames - lists the file name of each null model in the specified folder "path"
# IV - the site x species matrix of importance values
# SharedPres - list that will contain all sites in which the focal species and other species
# co-occur
# sitedf - subset of "Distance" dataframe that contains distances from a focal site to all
# other sites in "SharedPres" or "FocalOnly"
# sitelist - list of site to site distances from "sitedf", sorted from shortest to longest
# distance
# s1,s2,s3 - focal site, nearest neighbor site, and 2nd nearest neighbor site
# d1,d2,d3 - distance to focal site (always 0), distance to nearest neighbor site, distance
# to 2nd nearest neighbor site
# IVdf - subset of "IV" dataframe that contains the IV's of the focal species at the 3
# nearest neighbor sites
# iv1,iv2,iv3 - IV of focal species at focal site, IV of focal species at nearest neighbor
```

```
# site, IV of focal species at 2nd nearest neighbor site
# IVval - variable that calls the inverse distance weighted calculation
# Rangedist - subset of "Range" dataframe that contains the distance to the focal
# species range margin from the focal site
# counter2 - counter variable used for indexing focal species presence lists and the
# other species absence lists
# FocalOnly - list that will contain all site in which only the focal species occurs

#Import modules - pandas for dataframe manipulation, glob for file structure
#manipulation
import pandas as pd
import glob

#Distance Weighted Average function
#Takes 3 IV's and 3 distances as inputs
def dwa_iv(iv1,iv2,iv3,d1,d2,d3):
    dwa = (iv1*(1/d1)+iv2*(1/d2)+iv3*(1/d3))/((1/d1)+(1/d2)+(1/d3));
    return dwa;

#Import species presence and absence lists of null models
from NullLists import *
from PA_NullLists import *
```

```
#Import site x site distance matrix, and site x species margin matrix
Distance = pd.read_csv('.../DistanceMatrix.csv',index_col=0)
Range = pd.read_csv('.../SiteMatrix_toRangeMargins.csv',index_col=0)

#Site list
Sites=['B1','B2','B3','B4','B5','B6','B7','B8','B9','B10','B11','B12','B13','B14','B15','B16',
'B17','B18','B19','B20','B21','B22','B23','B24','B25','B26','B27','B28','B29','B30','B31',
'B32','B33','B34','B35','B36','B37','B38','B39','B40','B41','B42','B43','B44','B45','B46',
'B47','B48','B49','B50','B51','B52','B53','B54','B55','B56','B57','B58','B59','B60','B61',
'B62','B63','B64','B65','B66','B67','B68','B69','B70','B71','B72','B73','B74','B75','B76',
'B77','B78','B79','B80','B81','B82','B83','B84','B85','B86','B87','B88','B89','B90','B91',
'B92','B93','B94','B95','B96','B97','B98','B99','B100','B101','B102','RP001','RP002',
'RP003','RP004','RP005','RP006','RP007','RP008','RP009','RP010','RP011','RP012',
'RP014','RP015','RP016','RP017','RP018','RP019','RP020','RP021','RP022','RP023',
'RP024','RP025','RP026','RP028','RP029','RP030','RP032','RP033','RP034','RP035',
'RP036','RP037','RP043a','RP043b','RP044a','RP044b','RP045','RP046a','RP046b',
'RP047a','RP047b','RP049a','RP049b','RP050','RP052a','RP052b','RP053','RP054a',
'RP054b','RP055','RP057','RP058','RP059','RP060','RP061','RP062','RP063','RP064',
'RP066','RP067','RP069','RP070','RP071','RP072','RP073','RP074','RP075','RP077',
'RP078','RP079','RP081','RP082','RP083','RP085','RP086','RP087','RP088','RP089',
'RP090','RP091','RP092','RP093','RP094','RP095','RP096','RP097','RP098','RP099',
'RP100','RP102','RP103','RP104','RP105','RP106','RP107','RP108','RP109','RP110',
```

```
'RP111a','RP111b','RP112','RP113','RP114a','RP114b','RP115','RP116','RP117','RP118',
'RP119','RP120','RP121','RP122','RP123','RP124','RP125','RP126','RP127','RP128',
'RP129','RP130','RP131','RP132','RP133','RP134','RP135','RP136','RP137','RP139',
'RP140','RP141','RP142','RP143','RP144','RP145','RP146','RP147','RP148','RP149',
'RP150','RP151a','RP151b','RP152','RP153','RP154a','RP154b','RP155','RP156','RP157',
'RP158','RP159','RP160','RP162','RP163a','RP163b','RP164','RP165','RP166','RP167',
'RP168','RP169','RP170','RP171','RP172','RP173','RP174','RP175','RP176','RP177',
'RP178','RP179','RP180','RP181','RP182','RP183','RP184','RP185','RP187','RP188',
'RP189','RP190','RP191','RP192','RP193','RP194','RP195','RP196','RP197','RP199',
'RP200','RP201','RP202','RP203','RP204','RP205','RP206','RP207','RP208','RP209']

# Create Presence and Absence Site Lists
y = 0
for splist in SpLists:
    x = 0
    for val in splist:
        if val > 0:
            P_SpLists[y].append(Sites[x]);
        else:
            A_SpLists[y].append(Sites[x]);
        x = x + 1
    y = y + 1
```

```
#Create list of shared presence between species
counter1 = 0
path =r'G:\Dissertation\Scripts\Data\NullModels_ForLists'
filenames = glob.glob(path + "/*.csv")
for filename in filenames:
    IV = pd.read_csv(filename,index_col=0)
    SharedPres = []
    for firstsite in P_ACNE[counter1]:
        for secondsite in P_ACSA[counter1]:
            if firstsite == secondsite:
                SharedPres.append(firstsite);

#Returns site, distance weighted average IV, and distance to nearest 2 neighbors for
#sites where both species exist
    if len(SharedPres) > 2:
        for site in SharedPres:
            sitedf = Distance.loc[site,SharedPres]
            sitedf = sitedf.sort_values()
            sitedf = sitedf[0:3]
            sitelist = sitedf.index.tolist()
            s1 = sitelist[0]; s2 = sitelist[1]; s3 = sitelist[2]
            d1 = float(sitedf[0])+1; d2 = float(sitedf[1]); d3 = float(sitedf[2])
```

```
            IVdf = IV.loc["ACNE",sitelist]
            iv1 = float(IVdf[0]); iv2 = float(IVdf[1]); iv3 = float(IVdf[2])
            IVval = dwa_iv(iv1,iv2,iv3,d1,d2,d3)
            Rangedist = Range.loc[site,"ACNE"]
            print(filename,"ACNE","ACSA","Both Present",site,s1,s2,s3,IVval,Rangedist)
    counter1 = counter1 +1

#Create list of sites where focal species exists and other species does not
counter2 = 0
path =r'G:\Dissertation\Scripts\Data\NullModels_ForLists'
filenames = glob.glob(path + "/*.csv")
for filename in filenames:
    IV = pd.read_csv(filename,index_col=0)
    FocalOnly = []
    for firstsite in P_ACNE[counter2]:
        for secondsite in A_ACSA[counter2]:
            if firstsite == secondsite:
                FocalOnly.append(firstsite);

#Returns site, distance weighted average IV, and distance to nearest 2 neighbors for
# sites where focal species exists and other species does not
    if len(FocalOnly) > 2:
        for site in FocalOnly:
```

```
        sitedf = Distance.loc[site,FocalOnly]

        sitedf = sitedf.sort_values()

        sitedf = sitedf[0:3]

        sitelist = sitedf.index.tolist()

        s1 = sitelist[0]; s2 = sitelist[1]; s3 = sitelist[2]

        d1 = float(sitedf[0])+1; d2 = float(sitedf[1]); d3 = float(sitedf[2])

        IVdf = IV.loc["ACNE",sitelist]

        iv1 = float(IVdf[0]); iv2 = float(IVdf[1]); iv3 = float(IVdf[2])

        IVval = dwa_iv(iv1,iv2,iv3,d1,d2,d3)

        Rangedist = Range.loc[site,"ACNE"]

        print(filename,"ACNE","ACSA","Only ACNE",site,s1,s2,s3,IVval,Rangedist)

    counter2 = counter2 + 1
```

Code A3. Python code to create distance weighted IV's for species pairs where both species exist and only the focal species exists for the 100 null communities

Appendix B: Species Lists

Species	Family
Acer negundo L.	Sapindaceae
Acer saccharinum L.	Sapindaceae
Betula nigra L.	Betulaceae
Carpinus caroliniana Walter	Betulaceae
Carya cordiformis (Wangenh.) K. Koch	Juglandaceae
Carya illinoinensis (Wangenh.) K. Koch	Juglandaceae
Carya laciniosa (Michx. f.) G. Don	Juglandaceae
Catalpa speciosa (Warder) Warder ex Engelm.	Bignoniaceae
Cercis canadensis L.	Fabaceae
Celtis laevigata Willd.	Cannabaceae
Celtis occidentalis L.	Cannabaceae
Cephalanthus occidentalis L.	Rubiaceae
Celtis reticulata Torr.	Cannabaceae
Cornus florida L.	Cornaceae
Crataegus viridis L.	Rosaceae
Diospyros virginiana L.	Ebenaceae
Fraxinus pennsylvanica Marsh.	Oleaceae
Gleditsia triacanthos L.	Fabaceae
Gymnocladus dioicus (L.) K. Koch	Fabaceae
Ilex decidua Walter	Aquifoliaceae
Juglans nigra L.	Juglandaceae
Juniperus virginiana L.	Cupressaceae
Liquidambar styraciflua L.	Altingiaceae
Maclura pomifera (Raf.) C.K. Schneid.	Moraceae
Morus rubra L.	Moraceae
Nyssa sylvatica Marshall	Nyssaceae
Platanus occidentalis L.	Platanaceae
Populus deltoides W. Bartram ex Marshall	Salicaceae
Prunus mexicana S. Watson	Rosaceae
Quercus falcata Michx.	Fagaceae
Quercus lyrata Walter	Fagaceae
Quercus macrocarpa Michx.	Fagaceae
Quercus muehlenbergii Engelm.	Fagaceae
Quercus nigra L.	Fagaceae
Quercus palustris Münchh.	Fagaceae
Quercus phellos L.	Fagaceae
Quercus shumardii Buckley	Fagaceae
Quercus stellata Wangenh.	Fagaceae
Rhamnus caroliniana Walter	Rhamnaceae

Species	Family
Salix nigra Marshall	Salicaceae
Sapindus saponaria L. *var. drummondii* (Hook. & Arn.) L.D. Benson	Sapindaceae
Sideroxylon lanuginosum Michx. ssp. *lanuginosum* Michx.	Sapotaceae
Tamarix gallica L.	Tamaricaceae
Ulmus alata Michx.	Ulmaceae
Ulmus americana L.	Ulmaceae
Viburnum prunifolium L.	Adoxaceae

Table B1. Species list and their respective families. Includes only those species present at five or more sites in the bottomland dataset

Species	# of Sites	RF	Mean IV	Rel. # of Stems	Mean RD	Mean RBA
Ulmus americana	97	95.1%	18.9	17.0	19.9	19.9
Morus rubra	92	90.2%	4.5	6.4	4.2	2.8
Celtis laevigata	76	74.5%	11.4	11.3	12.5	10.3
Fraxinus pennsylvanica	75	73.5%	7.9	7.4	8.3	8.0
Sapindus saponaria var. drummondii	60	58.8%	4.2	3.9	5.8	2.9
Populus deltoides	56	54.9%	7.1	5.1	5.6	10.7
Quercus macrocarpa	55	53.9%	2.9	3.2	2.0	3.7
Sideroxylon lanuginosum ssp. lanuginosum	54	52.9%	1.7	2.4	1.5	1.2
Acer negundo	53	52.0%	3.1	3.6	3.0	2.6
Juglans nigra	52	51.0%	2.2	2.7	1.6	2.2
Gleditsia triacanthos	47	46.1%	0.8	1.2	0.6	0.6
Carya illinoinensis	46	45.1%	4.0	3.6	3.4	5.1
Salix nigra	45	44.1%	1.6	1.8	1.3	1.8
Platanus occidentalis	42	41.2%	1.5	1.6	1.1	1.8
Quercus shumardii	37	36.3%	2.6	2.4	2.2	3.2
Maclura pomifera	37	36.3%	1.1	1.5	1.1	0.8
Carya cordiformis	31	30.4%	1.6	1.8	1.6	1.5
Juniperus virginiana	26	25.5%	1.8	2.2	2.0	1.1
Cercis canadensis	26	25.5%	0.5	0.9	0.5	0.2
Gymnocladus dioicus	25	24.5%	0.9	1.2	0.8	0.7
Diospyros virginiana	25	24.5%	0.4	0.7	0.4	0.2
Acer saccharinum	19	18.6%	0.9	0.9	0.8	1.0
Quercus muehlenbergii	18	17.6%	1.0	0.9	0.9	1.1
Crataegus viridis	17	16.7%	0.2	0.4	0.2	0.1
Celtis occidentalis	16	15.7%	3.4	2.7	4.0	3.4
Quercus nigra	15	14.7%	1.5	1.2	1.4	1.9
Tamarix gallica	12	11.8%	1.0	1.4	1.3	0.4
Ilex decidua	12	11.8%	0.3	0.5	0.4	0.0
Nyssa sylvatica	10	9.8%	0.3	0.5	0.3	0.3
Prunus mexicana	10	9.8%	0.1	0.1	0.1	< 0.1
Carya laciniosa	9	8.8%	0.3	0.5	0.3	0.3
Quercus falcata	9	8.8%	0.3	0.3	0.2	0.3
Quercus lyrata	8	7.8%	0.9	0.7	1.0	1.1
Quercus phellos	8	7.8%	0.8	0.6	0.8	1.1
Ulmus alata	7	6.9%	0.9	0.8	1.0	0.8
Celtis reticulata	7	6.9%	0.7	1.1	0.8	0.4
Betula nigra	7	6.9%	0.4	0.4	0.5	0.4
Cephalanthus occidentalis	7	6.9%	0.1	0.2	0.1	< 0.1
Cornus florida	7	6.9%	0.1	0.1	0.1	< 0.1

Species	# of Sites	RF	Mean IV	Rel. # of Stems	Mean RD	Mean RBA
Viburnum prunifolium	7	6.9%	< 0.1	0.1	< 0.1	< 0.1
Quercus stellata	6	5.9%	0.9	0.6	1.0	1.0
Liquidambar styraciflua	6	5.9%	0.7	0.6	0.7	0.9
Rhamnus caroliniana	6	5.9%	0.1	0.1	0.1	< 0.1
Quercus palustris	5	4.9%	1.4	0.9	1.4	1.9
Carpinus caroliniana	5	4.9%	0.5	0.5	0.7	0.3
Catalpa speciosa	5	4.9%	0.1	0.2	0.1	0.1
Robinia pseudoacacia	4	3.9%	0.5	0.4	0.7	0.5
Ilex opaca	4	3.9%	0.2	0.2	0.3	0.2
Acer saccharum	3	2.9%	0.1	0.1	0.1	0.1
Juglans microcarpa	3	2.9%	0.1	0.1	0.1	0.1
Carya aquatica	3	2.9%	0.1	0.1	0.1	0.1
Quercus alba	3	2.9%	0.1	0.1	0.1	0.1
Forestiera acuminata	3	2.9%	0.1	0.1	0.1	< 0.1
Planera aquatica	2	2.0%	0.1	0.1	0.2	0.1
Carya texana	2	2.0%	< 0.1	< 0.1	< 0.1	< 0.1
Carya tomentosa	2	2.0%	< 0.1	< 0.1	< 0.1	< 0.1
Asimina triloba	2	2.0%	< 0.1	< 0.1	< 0.1	< 0.1
Ulmus rubra	1	1.0%	0.1	0.1	0.1	0.1
Ulmus crassifolia	1	1.0%	0.1	0.1	0.1	0.1
Crataegus spp.	1	1.0%	< 0.1	0.1	< 0.1	< 0.1
Taxodium distichium	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Prosopis juliflora	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Elaeagnus angustifolia	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Styphnolobium affine	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Pinus echinata	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Tilia americana	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Quercus bicolor	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Prunus serotina	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1
Acer rubrum	1	1.0%	< 0.1	< 0.1	< 0.1	< 0.1

Table B2. Species list for bottomland sites. For each species, the number of sites they are present at, their relative frequency, mean importance value, mean relative number of stems, mean relative density, and mean relative basal area are shown

Species	Family
Acer negundo L.	Sapindaceae
Acer rubrum L.	Sapindaceae
Acer saccharinum L.	Sapindaceae
Acer saccharum Marsh.	Sapindaceae
Aesculus glabra Willd.	Sapindaceae
Amelanchier arborea (F. Michx.) Fernald	Rosaceae
Asimina triloba (L.) Dunal	Annonaceae
Betula nigra L.	Betulaceae
Carpinus caroliniana Walter	Betulaceae
Carya aquatica (F. Michx.) Elliott	Juglandaceae
Carya cordiformis (Wangenh.) K. Koch	Juglandaceae
Carya illinoinensis (Wangenh.) K. Koch	Juglandaceae
Carya laciniosa (Michx. f.) G. Don	Juglandaceae
Carya ovata (Mill.) K. Koch	Juglandaceae
Carya texana Buckley	Juglandaceae
Carya tomentosa (Lam. ex Poir.) Nutt.	Juglandaceae
Castanea ozarkensis Ashe	Fagaceae
Catalpa speciosa (Warder) Warder ex Engelm.	Bignoniaceae
Celtis laevigata Willd.	Cannabaceae
Celtis occidentalis L.	Cannabaceae
Celtis reticulata Torr.	Cannabaceae
Cephalanthus occidentalis L.	Rubiaceae
Cercis canadensis L.	Fabaceae
Cornus florida L.	Cornaceae
Crataegus crus-galli L.	Rosaceae
Crataegus L.	Rosaceae
Crataegus viridis L.	Rosaceae
Diospyros virginiana L.	Ebenaceae
Elaeagnus angustifolia L.	Elaeagnaceae
Forestiera acuminata (Michx.) Poir.	Oleaceae
Fraxinus americana L.	Oleaceae
Fraxinus pennsylvanica Marsh.	Oleaceae
Fraxinus quadrangulata Michx.	Oleaceae
Gleditsia triacanthos L.	Fabaceae
Gymnocladus dioicus (L.) K. Koch	Fabaceae
Ilex decidua Walter	Aquifoliaceae
Ilex opaca Aiton	Aquifoliaceae
Juglans microcarpa Berland.	Juglandaceae
Juglans nigra L.	Juglandaceae
Juniperus virginiana L.	Cupressaceae
Liquidambar styraciflua L.	Altingiaceae

Species	Family
Maclura pomifera (Raf.) C.K. Schneid.	Moraceae
Nyssa sylvatica Marshall	Nyssaceae
Ostrya virginiana (Mill.) K. Koch	Betulaceae
Pinus echinata Mill.	Pinaceae
Pinus taeda L.	Pinaceae
Planera aquatica J.F. Gmel.	Ulmaceae
Platanus occidentalis L.	Platanaceae
Populus deltoides W. Bartram ex Marshall	Salicaceae
Populus deltoides W. Bartram ex Marshall ssp. *monilifera* (Aiton) Eckenw.	Salicaceae
Prosopis glandulosa Torr.	Fabaceae
Prosopis velutina Wooton	Fabaceae
Prunus americana Marshall	Rosaceae
Prunus mexicana S. Watson	Rosaceae
Prunus serotina Ehrh.	Rosaceae
Quercus alba L.	Fagaceae
Quercus bicolor Willd.	Fagaceae
Quercus falcata Michx.	Fagaceae
Quercus havardii Rydb.	Fagaceae
Quercus lyrata Walter	Fagaceae
Quercus macrocarpa Michx.	Fagaceae
Quercus margarettae (Ashe) Small	Fagaceae
Quercus marilandica Münchh.	Fagaceae
Quercus muehlenbergii Engelm.	Fagaceae
Quercus nigra L.	Fagaceae
Quercus palustris Münchh.	Fagaceae
Quercus phellos L.	Fagaceae
Quercus rubra L.	Fagaceae
Quercus shumardii Buckley	Fagaceae
Quercus stellata Wangenh.	Fagaceae
Quercus texana Buckley	Fagaceae
Quercus velutina Lam.	Fagaceae
Rhamnus caroliniana Walter	Rhamnaceae
Robinia pseudoacacia L.	Fabaceae
Salix nigra Marshall	Salicaceae
Sapindus saponaria L. var. *drummondii* (Hook. & Arn.) L.D. Benson	Sapindaceae
Sassafras albidum (Nutt.) Nees	Lauraceae
Sideroxylon lanuginosum Michx. ssp. *lanuginosum* Michx.	Sapotaceae
Styphnolobium affine (Torr. & A. Gray) Walp.	Fabaceae
Tamarix gallica L.	Tamaricaceae
Taxodium distichum (L.) Rich.	Cupressaceae

Species	Family
Tilia americana L.	Malvaceae
Tilia americana L. *var. americana* L.	Malvaceae
Ulmus alata Michx.	Ulmaceae
Ulmus americana L.	Ulmaceae
Ulmus crassifolia Nutt.	Ulmaceae
Ulmus rubra Muhl.	Ulmaceae
Viburnum prunifolium L.	Adoxaceae
Viburnum rufidulum Raf.	Adoxaceae

Table B3. Species list and their respective families for all species in both upland and bottomland datasets

Appendix C: RDA Plots

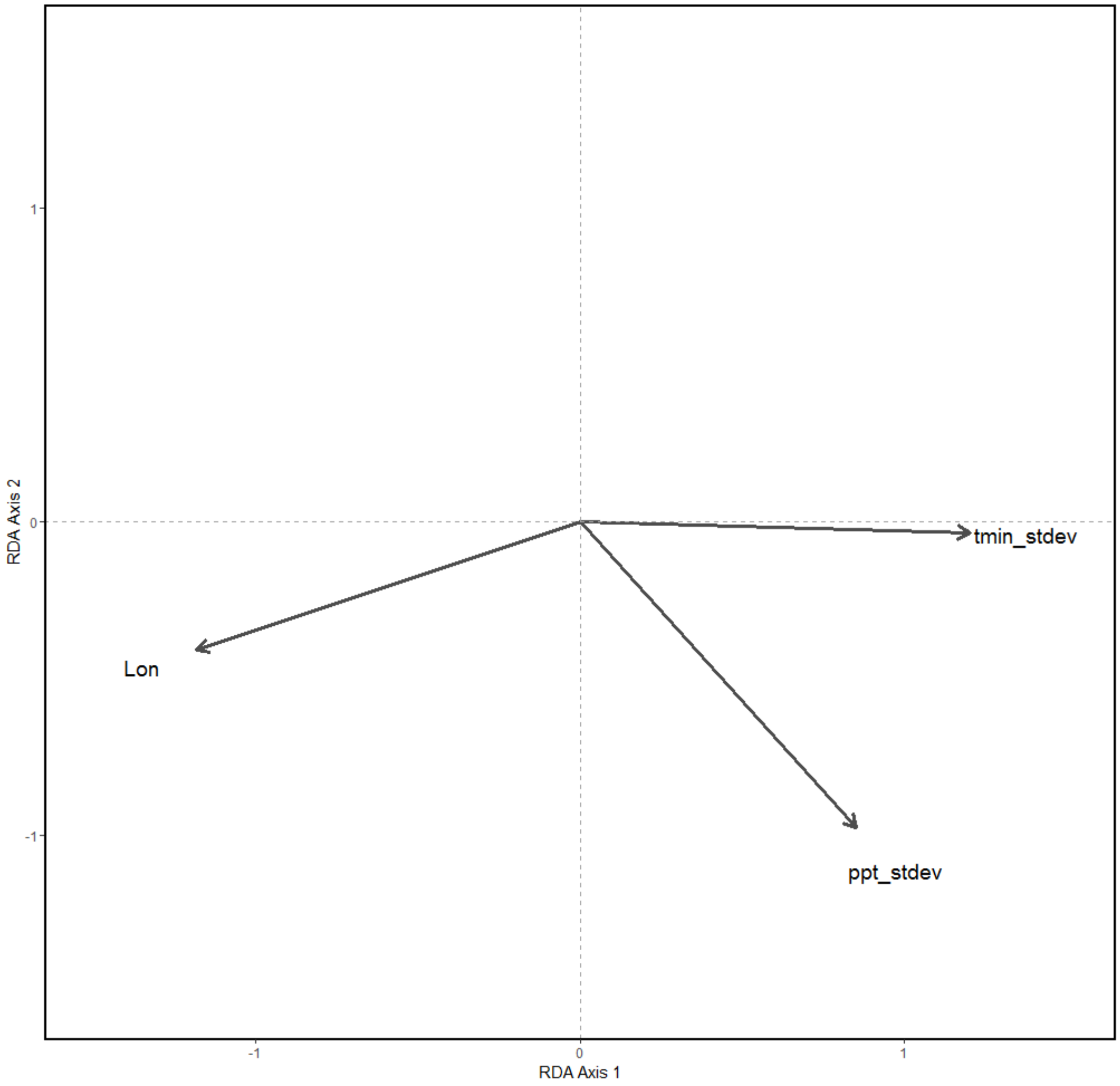


Figure C1. RDA model with all variables included. Variables shown indicate those that explain the most variance in community composition and the length of the arrow describes the relative contribution of that variable

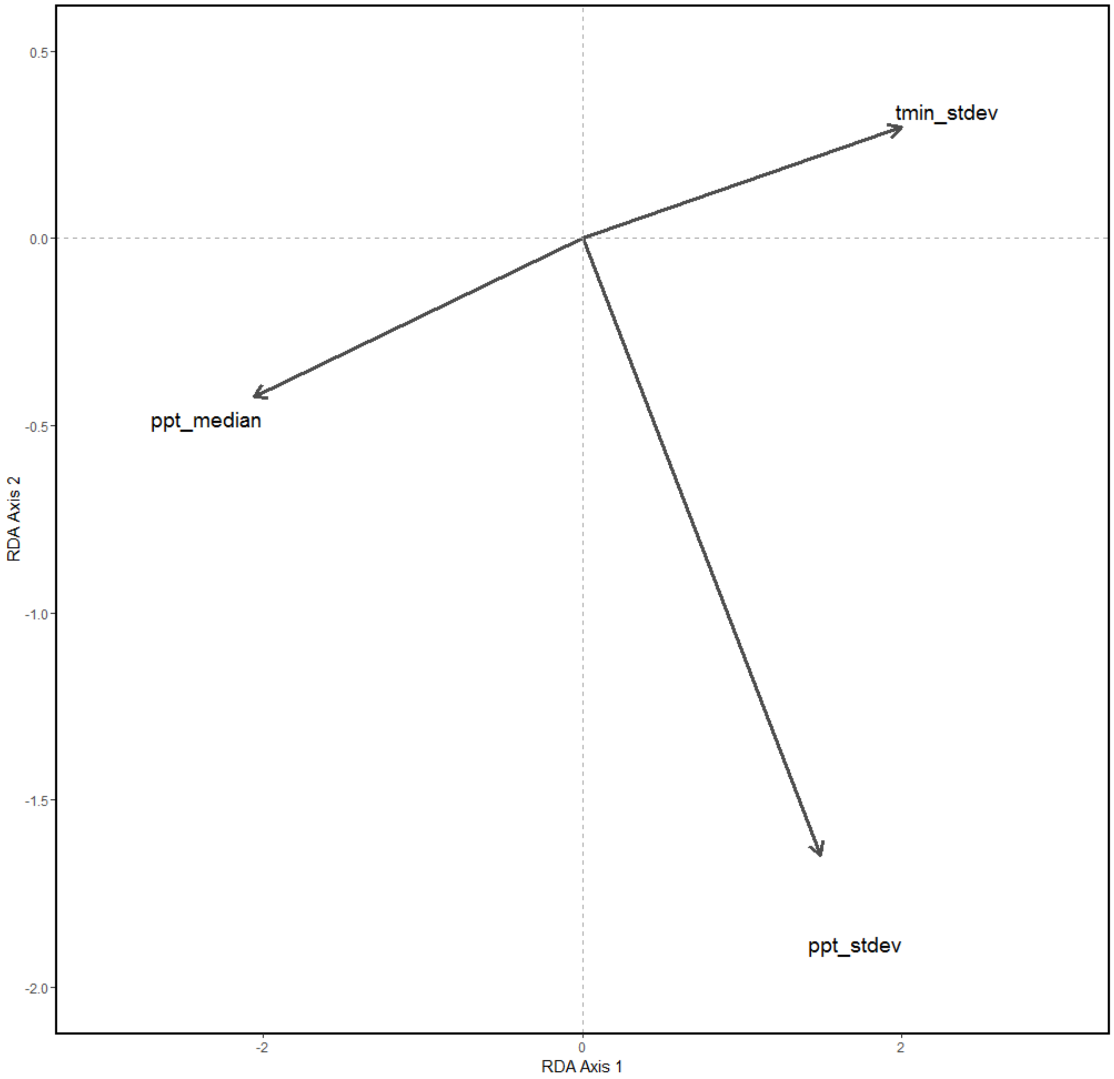


Figure C2. RDA model with only environmental variables included. Variables shown indicate those that explain the most variance in community composition and the length of the arrow describes the relative contribution of that variable

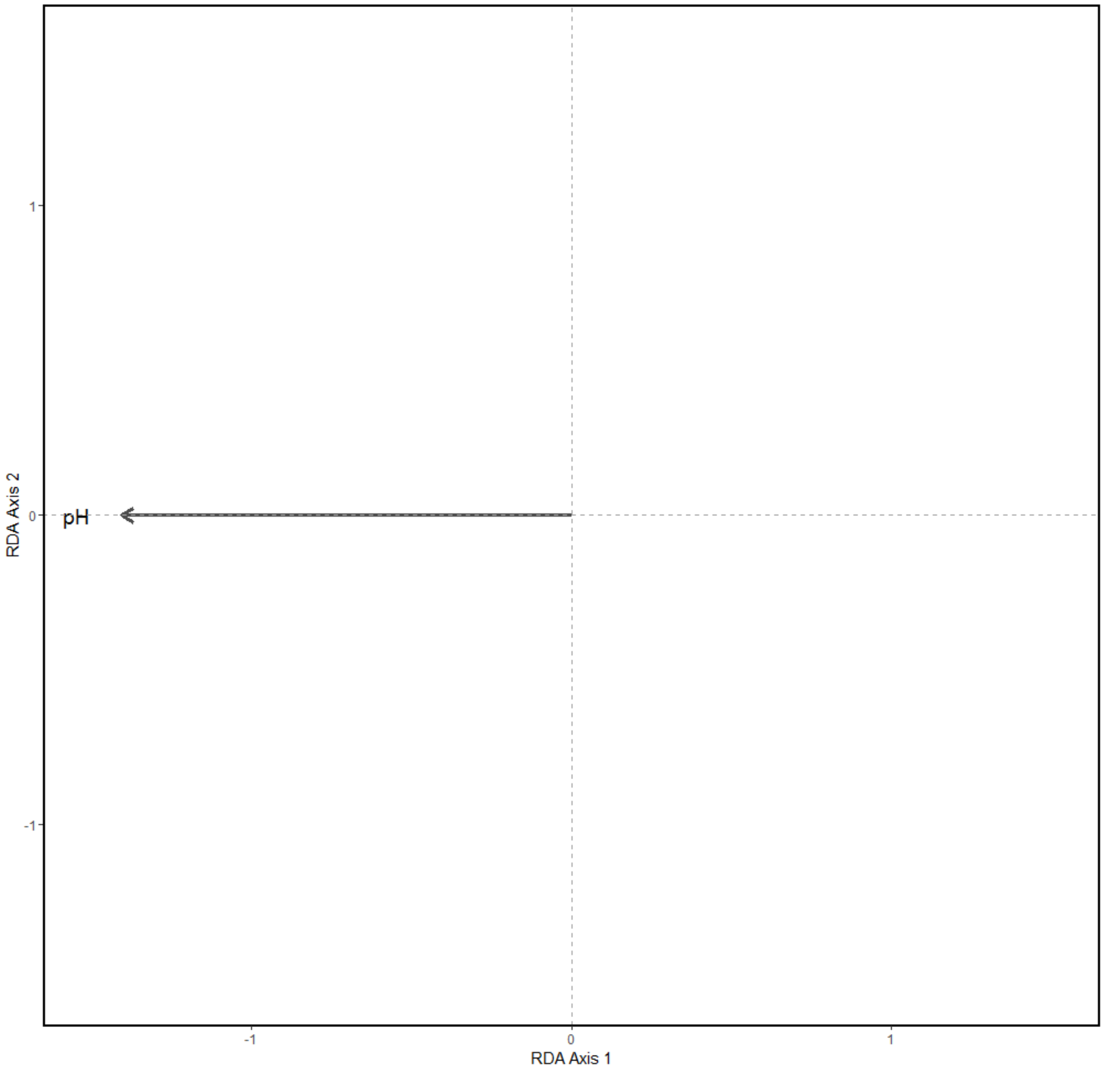


Figure C3. RDA model with only edaphic variables included. The variable shown indicates that pH explains the most variance in community composition

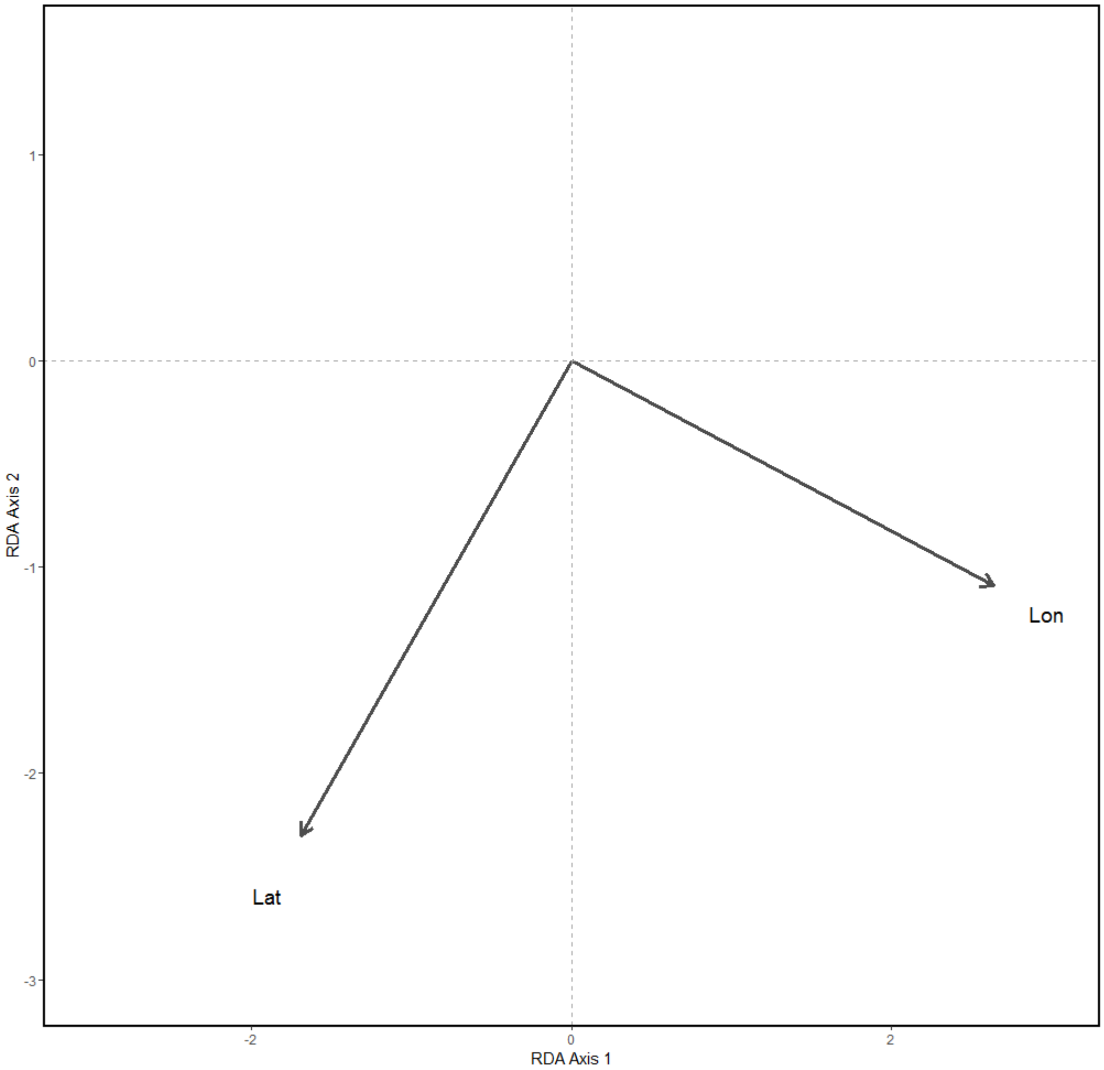


Figure C4. RDA model with only spatial variables included. Variables shown indicate those that explain the most variance in community composition and the length of the arrow describes the relative contribution of that variable

Appendix D: Additional beta diversity maps

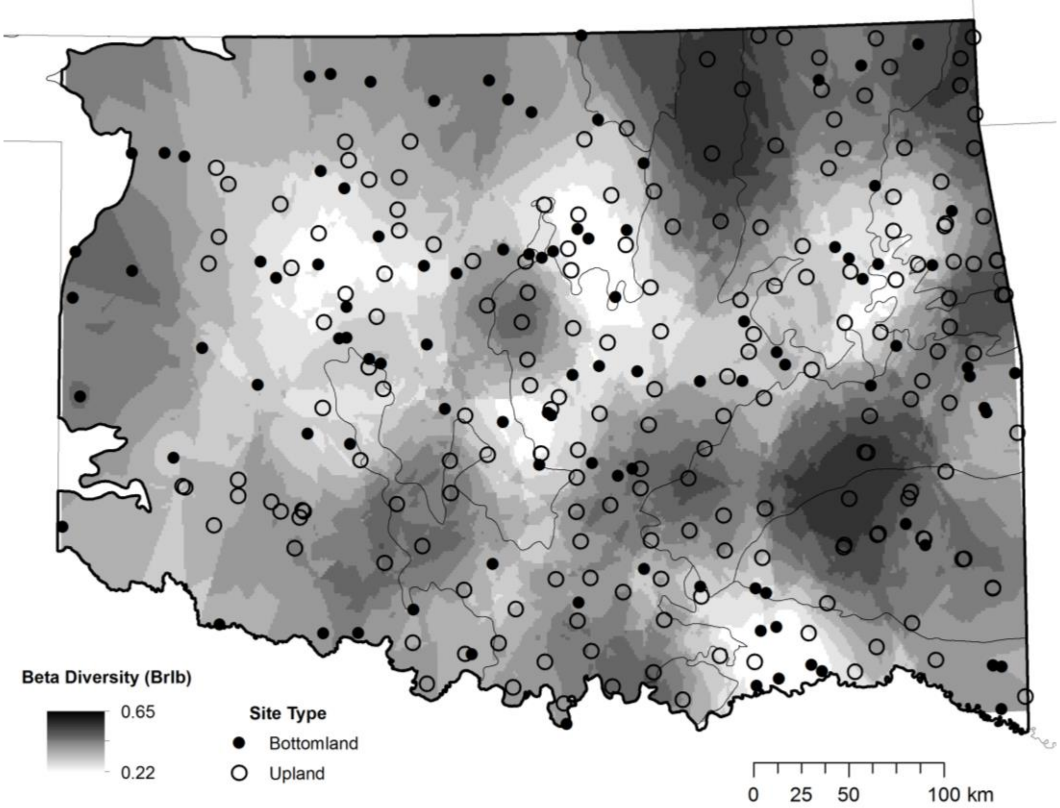


Figure D1. Spatial patterns in tree beta diversity using the 'rlb' beta diversity measuring turnover from Koleff et al. (2003)

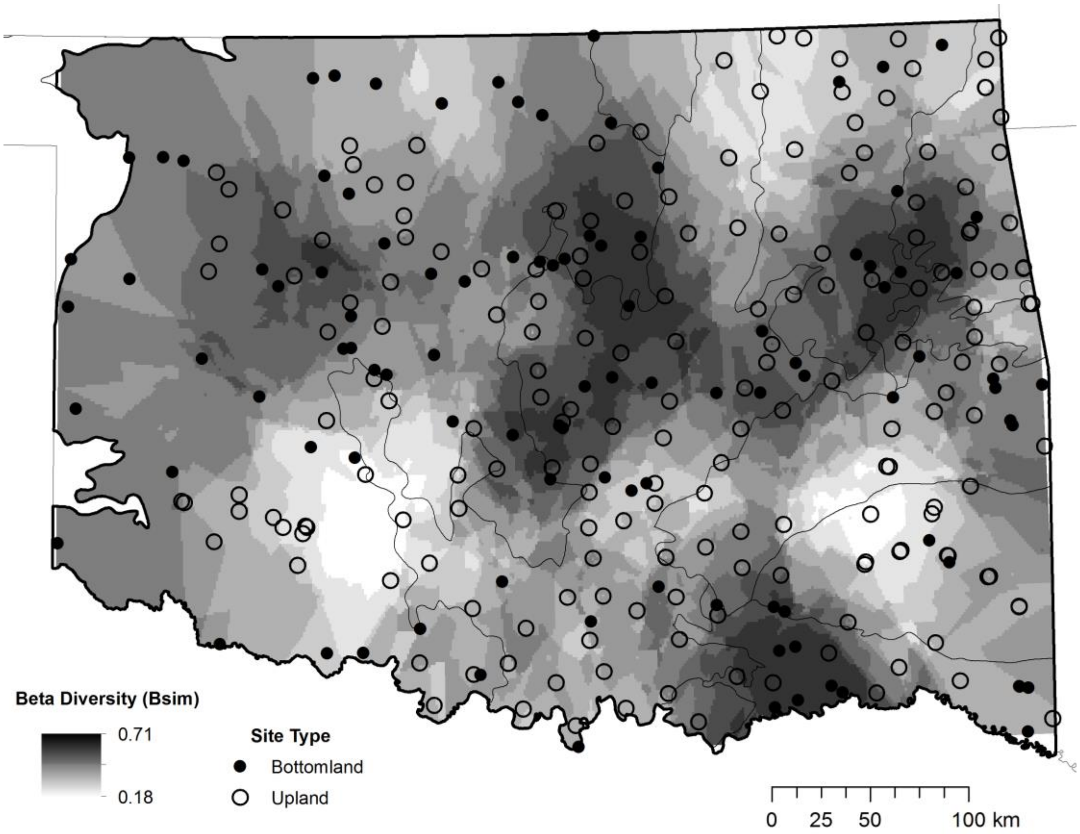


Figure D2. Spatial patterns in tree beta diversity using Simpson's beta diversity measuring gain and loss in species from Koleff et al. (2003)

Appendix E: Proximity to Range Margins

	Distance (km)
Mean	199.5
Median	188.1
Minimum	-246.6
Maximum	619.6
Range	866.3
Standard Deviation	128.4

Table E1. Descriptive Statistics for the distance between sites where a species is present and its range margin

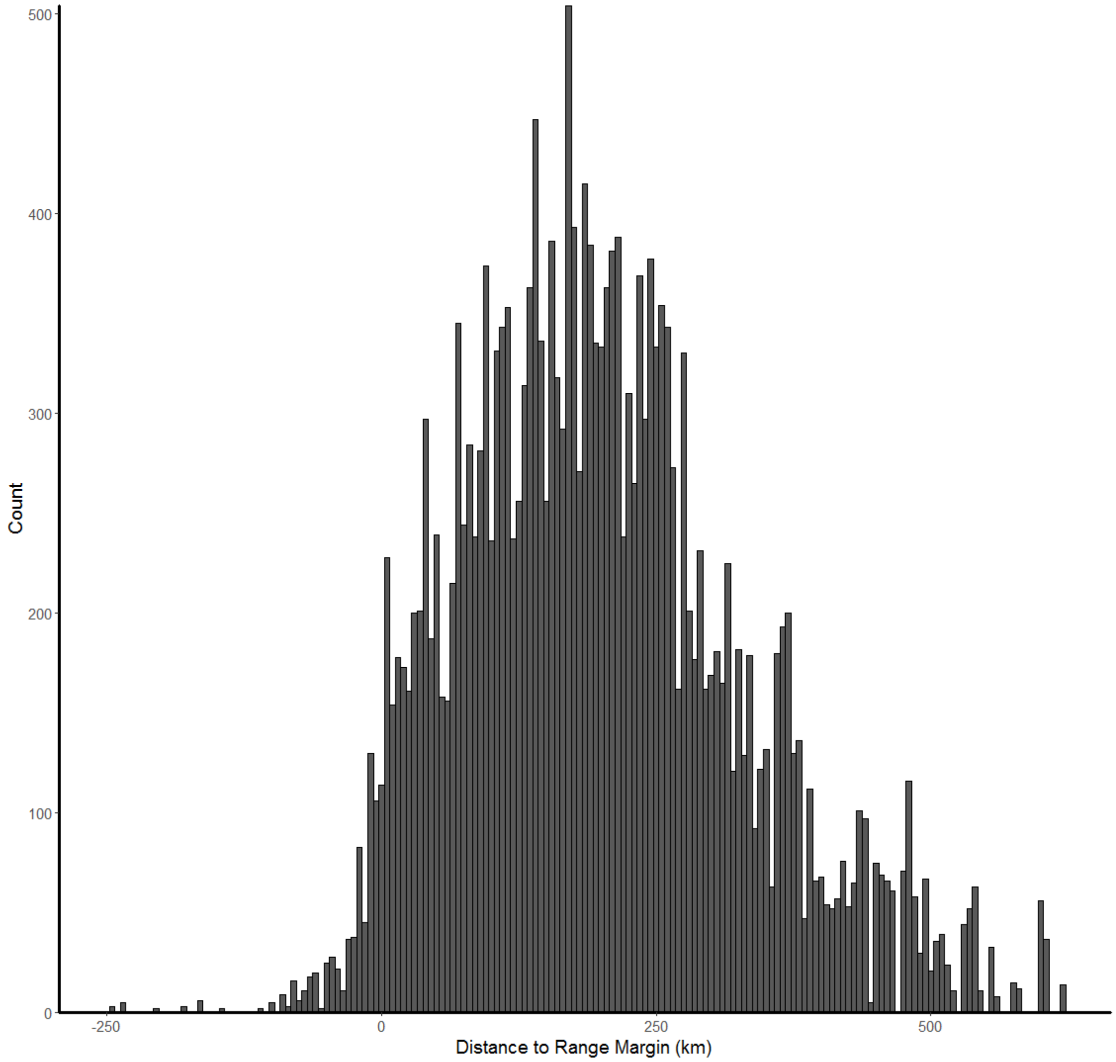


Figure E1. Histogram of distances from sites where a species is present to its range margin

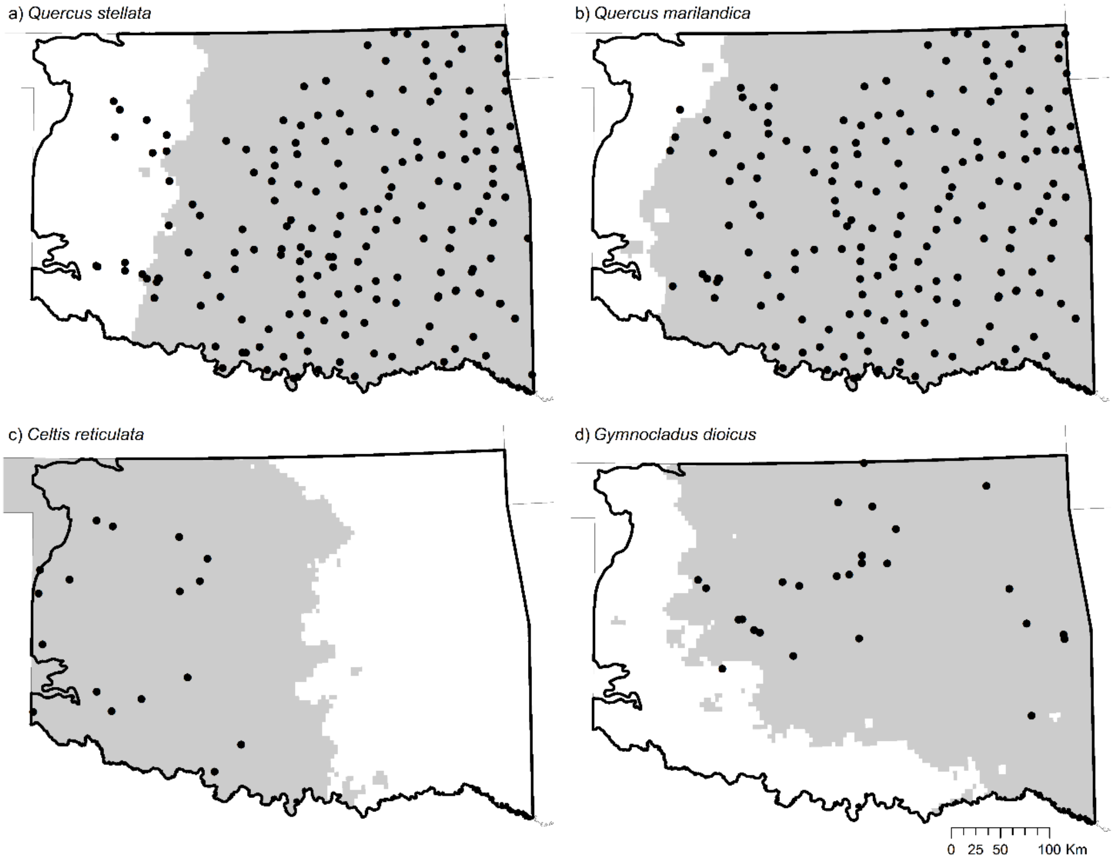


Figure E2. Examples of range margins and site presence for four species. a) *Quercus stellata,* b) *Quercus marilandica,* c) *Celtis reticulata,* and d) *Gymnocladus dioicus*. Range margins were generated using MaxEnt, species presence was obtained for the bottomland (Johnson 1982) and upland (Rice and Penfound 1959) datasets