

ANALYSIS OF MODERN POLLEN DATA FROM THE PRAIRIES OF
CENTRAL NORTH AMERICA

by

CHRISTOPHER J. MORRIS

B.S., Kansas State University, 2011

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF ARTS

Department of Geography
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2013

Approved by:

Major Professor
Kendra K. McLauchlan

Copyright

CHRISTOPHER J. MORRIS

2013

Abstract

Fossil pollen assemblages are widely used in paleoenvironmental reconstruction of vegetation regimes and climate conditions. The modern analog technique (MAT) is a popular method used for analysis of these fossil pollen assemblages, but a large modern pollen dataset, such as the North American Pollen Database (NAPD), is needed to provide modern comparisons for interpretation of analog/no-analog situations. While many climate types are well represented within the NAPD, the climates of the southern and central Great Plains of North America are poorly represented. In this study, I collected 31 sediment samples containing pollen from these underrepresented climate types across the Great Plains in the U.S.A. Analysis of these 31 pollen assemblages, along with 504 samples classified as “prairie” from the NAPD and 24 pollen samples from the Flint Hills of Kansas, U.S.A. was conducted to determine if the three major prairie types (short grass, mixed grass, and tallgrass prairies) could be delineated from pollen records alone. Two different MAT dissimilarity metrics (Squared Chord Distance and Canberra Distance Metric) were assessed for their ability to delineate among prairie types and Squared Chord Distance (SCD) was found to be the better prairie type classifier than Canberra Distance Metric (CDM). Receiver Operator Characteristic (ROC) curve analysis was used to assess the ability of each metric to identify similar pollen assemblages. It has been shown in previous studies that two genera found in this region – *Ambrosia* and *Artemisia* – respond to temperature and moisture availability in different ways. Using the ratio of the proportions of *Ambrosia* and *Artemisia* pollen grains in a pollen assemblage it was found that tallgrass prairies are significantly different from the other two prairie types. The *Ambrosia/Artemisia* ratio is also useful in determining climatic conditions. This ratio provides paleoenvironmental researchers with a simple quantitative tool to quickly assess general climatic conditions and prairie type.

Table of Contents

List of Figures	vi
List of Tables	x
Acknowledgements	xi
Chapter 1 - Literature Review.....	1
Relevance to Geography	1
Pollen	3
North American Pollen Dataset (NAPD).....	4
Modern Analog Technique (MAT).....	9
Prairies of the Great Plains of North America	11
Tallgrass Prairies	13
Short Grass Prairies.....	14
Mixed Grass Prairies	14
Chapter 2 - Introduction.....	16
Chapter 3 - Study Area	22
Chapter 4 - Methods.....	26
Site Selection	26
Sediment Sample Collection.....	29
Data Analysis	31
Squared Chord Distance (SCD)	31
Canberra Metric Distance (CMD)	33
Receiver Operator Characteristic (ROC)	34
<i>Ambrosia/Artemisia</i> Ratio Analysis.....	35
Chapter 5 - Results.....	36
Sample Site Locations and Climate Types	36
<i>Ambrosia</i> and <i>Artemisia</i> Analysis.....	40
MAT Dissimilarity Matrices.....	46
ROC Analysis: Determining Threshold Values.....	53
Chapter 6 - Discussion	56
Sample Site Locations and Climate Types	56

<i>Ambrosia</i> and <i>Artemisia</i> Analysis.....	57
MAT Dissimilarity Matrices.....	60
ROC Analysis: Determining Threshold Values.....	61
References.....	65
Appendix A - Plant Types.....	78
Appendix B - Site Location and Metadata.....	81
Appendix C - Raw Data.....	84
Appendix D - <i>Ambrosia</i> / <i>Artemisia</i> Ratios	89
Appendix E - R Code.....	90

List of Figures

Figure 1.1: Fenneman's (1919) Venn diagram (left) shows his vision of how geography overlaps with various other sciences, including biology. Tuason's (1987) modified Venn diagram (right) shows her vision of how the subdivisions of geography are bridges to the other sciences with no definitive boundaries between them.	2
Figure 1.2 North American Modern Pollen Database (NAPD) sample sites are represented by grey dots, while Commerford's (2010) Flint Hills Tallgrass Prairie sample sites are represented by orange dots.....	6
Figure 1.3: Point Density Map of the NAPD and Commerford's (2010) sites, where the darker the color the higher density of points.....	7
Figure 1.4: Whitmore et al.'s (2005) graph shows the mean annual precipitation and the mean annual temperature of each sample site of the NAPD (black dots). Each color represents a different biome.....	8
Figure 3.1: The study area is shown in green and represents the region of the Great Plains with a climate between 9° C and 21° C mAT and between 400 mm and 1000 mm mAP.	23
Figure 3.2: Panel A shows the location of Great Plains Pollen Sample (GPPS) sites. Panel B shows the location of the GPPS sites (orange) along with Commerford's Flint Hills sample sites (red), and the NAPD sample sites (dark green). Panel C shows the sample sites relative to the chosen climate space. Panel D shows the GPPS sites, Commerford's sites and the three prairie ecoregions as defined by Bailey's (1994) Ecoregions and Subregions Map of the United States.	24
Figure 3.3: This photo shows a windmill used to provide water to cattle and wildlife on the Thelander Ranch in the southwestern part of the Nebraska Sand Hills.....	25
Figure 4.1: This is a modified figure from Whitmore et al. (2005) that shows the climate for each pollen assemblage within the NAPD. The lack of prairie samples sites is evident within the region highlighted by the red quadrangle (the climate space in which most prairies are found).	27
Figure 4.2: This photo demonstrates the use of an 8 foot inflatable raft to deploy an Eckman dredge for sediment collection from the middle of Ray Ranch Pond.	30

Figure 4.3: This figure shows how ROC analysis determines the optimal dissimilarity value (dotted gray line), which is where the true-positive/false positive line (red line) meets the true-negative/false negative (blue line).....	34
Figure 5.1: A) A point density map of all NAPD sites (dark blue dots) along with Commerford's (2010) Flint Hills sample sites (orange dots). B) A point density map that also includes the GPPS sites (red dots). The darker the gray the more sites nearby.	37
Figure 5.2: The climate space of all sample sites classified as "prairie" within the NAPD (shown as open circles), along with the samples collected by Commerford (2010)(orange circles) and the GPPS sites (red circles). The climate space used in this study (maP 9 - 21° C and maT 400 – 1000 mm) is denoted by the dashed line square.	38
Figure 5.3: The NAPD "prairie" and the Flint Hills sample sites are shown here in regional groups to illustrate the different prairie environments all under the generic label "prairie".	39
Figure 5.4: This boxplot of the proportions of <i>Ambrosia</i> (white) and <i>Artemisia</i> (grey) for every pollen assemblage classified as prairie within the NAPD, the Flint Hills Samples (Commerford 2010), and the GPPS sites, further broken down by prairie type.....	41
Figure 5.5: Mean annual precipitation (maP) is the independent variable (x-axis) and the <i>Ambrosia/Artemisia</i> is the dependent variable (y-axis). The vertical dashed line is the maP threshold at 500 mm and the horizontal dashed line is the <i>Ambrosia/Artemisia</i> ratio decision threshold of 10. Each pollen assemblage analyzed in this study is represented by a black dot (n = 559).....	44
Figure 5.6: Mean annual temperature (maT) is the independent variable (x-axis) and the <i>Ambrosia/Artemisia</i> is the dependent variable (y-axis). The vertical dashed line is the maP threshold at 500 mm and the horizontal dashed line is the <i>Ambrosia/Artemisia</i> ratio decision threshold of 10. Each pollen assemblage analyzed in this study is represented by a black dot (n = 559).....	45
Figure 5.7: SCD dissimilarity matrix among all pollen assemblages from grasslands of North America with lower dissimilarity represented by reds and higher dissimilarity represented by yellows. The matrix is broken down by prairie type and the GPPS and Flint Hills sample sets. The numbers represent number of samples starting with 0 in the bottom left corner (n = 559).....	49

Figure 5.8: CDM dissimilarity matrix among all pollen assemblages from grasslands of North America with lower dissimilarity represented by dark grays and black and higher dissimilarity represented by light grays and white. The matrix is broken down by prairie type and the GPPS and Flint Hills sample sets. The numbers represent number of samples starting with 0 in the bottom left corner (n = 559).....	50
Figure 5.9: SCD dissimilarity matrix for only the GPPS set with lower dissimilarity represented by dark grays and black and higher dissimilarity represented by light grays and white. The numbers are the sample numbers from the GPPS dataset. Note sample #9 was left out of the analysis, there only are 31 samples represented in this matrix.	51
Figure 5.10: CDM dissimilarity matrix for only the GPPS set with lower dissimilarity represented by dark grays and black and higher dissimilarity represented by light grays and white. The numbers are the sample numbers from the GPPS dataset. Note sample #9 was left out of the analysis, there only are 31 samples represented in this matrix.	52
Figure 5.11: The receiver operator characteristic (ROC) curves for all prairies for both SCD analysis (left) and CDM analysis (right) are shown. Area under the curve (AUC) indicates the ability of the analysis to classify correctly compared to a completely randomized classification.	53
Figure 5.12: The ROC curves for the three prairie types are shown for both SCD (top) and CDM (bottom) analysis.....	55

List of Equations

Equation 4.1: Squared Chord Distance	32
Equation 4.2: Canberra Metric Distance	33

List of Tables

Table 5.1: The mean proportions of both pollen types (<i>Ambrosia</i> , <i>Artemisia</i>) for all three prairie types (Mixed Grass, Short Grass, and Tallgrass) are shown as well as the ratio of means where (ratio = $Ambrosia/Artemisia$). * Average Ratios do not include values divided by zero.	42
Table 5.2: Mann-Whitney-Wilcox Test compares each prairie type's distribution of the ratios between <i>Ambrosia</i> and <i>Artemisia</i> with the other two prairie types' ratio distributions at a 95% significance level.	42
Table 5.3: List of the 47 pollen types used for MAT analysis in this study	47
Table A.1: List of plant types surveyed in the field.....	78
Table B.1: GPPS site location, elevation, prairie type, land ownership, and the date the sample was collected.	81
Table B.2: Climate data for each GPPS site.	83
Table C.1: The number of pollen grains, spore, and charcoal counted for each site. GPPS 9 had too few pollen grains so was not counted. Pollen analyst: Andrea Nurse, University of Maine	84
Table D.1: The percentiles of <i>Ambrosia/Artemisia</i> ratios for each prairie type.	89

Acknowledgements

I would like to take this space to thank all of those who have helped make this thesis possible, starting with my advisor Dr. Kendra McLauchlan (Kansas State University, Department of Geography). She not only introduced me to pollen assemblages and the world of paleoenvironmental reconstruction, but also provided invaluable support and input throughout this undertaking. I would like to also thank Dr. Douglas Goodin and Dr. Shawn Hutchinson (Kansas State University, Department of Geography) for their input and serving as my committee members. LacCore: National Lacustrine Core Facility (University of Minnesota Limnological Research Center) conducted the pollen preparation and Andrea Nurse (University of Maine, Climate Change Institute) conducted the pollen counts.

I would like to thank all the landowners and contact persons of public lands for allowing me to collect samples (Chris and Flo Ray, Pat Thelander and his sons, Lisa Heiser (Nebraska & Samuel R. McKelvie National Forests, Buffalo Gap, Fort Pierre & Oglala National Grasslands), Mark Koepsel (Crescent Lake National Wildlife Refuge), Rachel Laubhan (Quivira National Wildlife Refuge), Walter Munsterman (Wichita Mountains Wildlife Refuge)).

I gratefully acknowledge a fellowship from the KSU NSF GK-12 EIDRoP program during my 2012-2013 academic year.

I'd like to thank Josh Mueller for not getting "fired" and providing me with laughs when I needed them. A very big thank you to Dani Rouff for all those late night dinners, and so much more. And finally I would like to thank my mother for her endless support and timely boxes of cookies.

Chapter 1 - Literature Review

Relevance to Geography

There are four main traditions of geography: spatial, area studies, human-environment, and Earth science (Pattison 1964). Each of these four traditions could be divided into any number of divisions or topics with many of them being interdisciplinary in nature. The research described in this thesis is an investigation into grassland pollen assemblages from the Great Plains, which encompasses both the spatial and Earth science traditions. Biogeography, as a discipline, exists at the overlap of two other disciplines, biology and geography, as is typical of many geographic topics (Fenneman 1919) (Fig. 1.1). Tuason (1987) insisted that the sub-disciplines of geography are not only a region of overlap between geography and other scientific fields but also a bridge among the sciences which allows for greater flow of ideas and information between disciplines and greater intellectual exploration of research topics. According to Tuason, biogeography is a field of study which occupies an intellectual region of both geography and biology but there are no defining boundaries between biogeography and its parent sciences.

Biogeography, by its very nature, must be multidisciplinary, which is a key strength in fostering creativity in research (Cowell & Parker 2004). Although there is a large diversity of research topics within the field, there are four fundamental issues that have been the main focus of biogeographers over the last century: spatial pattern and process, landscape change, human-environment interface, and linking physical and biological systems (Cowell & Parker 2004). Spatial pattern and process studies can involve determining the distribution of a species, a group

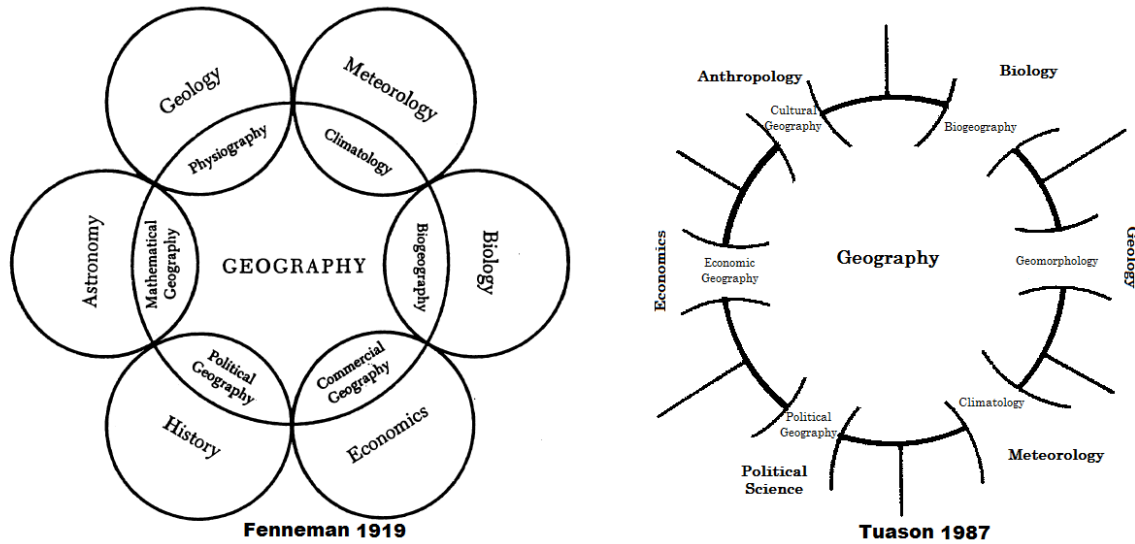


Figure 1.1: Fenneman’s (1919) Venn diagram (left) shows his vision of how geography overlaps with various other sciences, including biology. Tuason’s (1987) modified Venn diagram (right) shows her vision of how the subdivisions of geography are bridges to the other sciences with no definitive boundaries between them.

of species, or an entire ecosystem and the factors that have influenced their distributions (Küchler 1974, Hill et al. 2000, McLauchlan et al. 2007). My research involves the distribution of short, mixed, and tallgrass prairies across the Great Plains of North America which can be categorized in the spatial pattern and process category. My research also addresses issues of linking physical and biological systems, by linking climate to prairie type. Finally this research addresses issues of landscape change, because a main outcome is to contribute to paleoenvironment reconstructions of vegetation, specifically landscape change during the Holocene.

As a subfield of biogeography, paleoenvironmental change is the most relevant subfield to this thesis with its focus on increasing modern surface samples of the North American surface sample dataset. A quick search of “paleobiogeography” on the database search engine Web of

Science (accessed February 2013) brought up over 1,300 results and other keyword searches of paleo-reconstruction related words produced thousands more examples of studies focused on reconstructing past environmental conditions. These studies have connections to all four of Cowell & Parker's (2004) fundamental issues in biogeography. Therefore, paleoecology has a solid foundation in the realm of geography. The basic tool of paleoenvironmental reconstruction is called a proxy record or proxy. A proxy is preserved physical record that is used in place of direct measurement or observation in order to reconstruct conditions of a past environment. Many different proxies are used to recreate past landscapes and processes including but not limited to: charcoal, diatoms, opal phytoliths, packrat middens, chironomids, vertebrate fossils, plant macrofossils, stable isotopes, and pollen. Pollen assemblages have been used to reconstruct past vegetation for almost a century starting with von Post (1916) who found pollen in peat bog deposits which was related to shifts in past vegetation types in southern Sweden.

Pollen

Pollen is produced by plants as part of their reproductive cycle, however most of the pollen produced by these plants does not ever fulfill that intended purpose. Instead, much of the pollen is deposited on to the nearby terrestrial and lacustrine environments, with nearby being a relative term depending on plant taxa (Mazier et al. 2008). The deposition of pollen continues year after year, century after century, and millennia after millennia, and so long as there are plants producing pollen on the landscape and a depositional environment continues to exist, there will be a pollen record of those vegetation types (Yansa et al. 2007). These collections of pollen, known as pollen assemblages, leave a record of what types of vegetation were on the landscape and producing pollen during the time of deposition (e.g. Fredlund 1995). Pollen grains can be isolated from the sedimentary depositional environment and individually identified by taxon and

quantified. The data are then presented as an assemblage of pollen types which can then be used in further analysis (Faegri & Iverson 1989). It is important to remember that pollen assemblages are related to the vegetation surrounding the sample site, but quantitative comparisons are not direct because different taxa produce different amounts of pollen, have different pollen dispersal strategies, and have different pollen preservation rates (Jackson & Williams 2004). These factors can result in the under- or over-representation of certain taxa. For example Hall (1994) found that juniper and pine pollen made up 30% of the total pollen influx into a sample site located in a short grass prairie where the trees made up a very small percentage of the actual vegetation cover. Despite these results it has been found that pollen assemblages provide a relatively accurate assessment of the vegetative cover (Davis & Webb 1975). There are many different depositional environments that have been used to collect pollen samples, including lacustrine sediments (Luly et al. 2006; Grimm et al. 2011), peat bogs (Woillard 1978), and soils (Davis 1995; Bement et al. 2007). Because of the rugged outer casing pollen grains are preserved well in many depositional environs. When conditions are right, anoxic and acidic, pollen can be preserved for over 140,000 years such as found in a peat bog in France (Woillard 1978).

North American Pollen Dataset (NAPD)

Because of the uncertainties in relating pollen assemblages to vegetation cover, one popular analytical approach is the modern analog technique (MAT). A variety of pollen data analysis techniques including the MAT, can be used to assist in paleoenvironment vegetation reconstructions. However, in order for these techniques to be useful in paleo-vegetation and paleoclimate reconstruction, both fossil pollen data and modern pollen data are needed (Sawada et al. 2004; Lytle & Wahl 2005; Ohlwein & Wahl 2012). A modern surface pollen sample is a sample usually acquired from the surface sediments of the 20th and 21st centuries. Therefore, the

climate and vegetation types are known and recorded for the sampling period (Whitmore et al. 2005). Modern surface pollen assemblages have become an essential tool in quantitative vegetation modeling of the paleoenvironment worldwide: (Bonnefille & Chalieu 2000) Africa, (Luo et al. 2010) Asia, (Luly et al. 2006) Australia, (Seppä et al. 2004) Europe, (Tonello & Prieto 2008) South America, and (Williams et al. 2000) North America. Large datasets containing numerous modern and fossil pollen assemblages have also become very important in paleoenvironmental reconstructions, but instead of analyzing a single sample site, regional and continental scale analyses can be implemented. Two such datasets are the European Pollen Database (EPD) and the North American Pollen Database (NAPD). Both these databases are relatively recent phenomena and are necessary for the analysis of spatial patterns of paleovegetation and how they change through time (Fyfe et al. 2009).

In a Herculean effort North American paleoecologists have collected and analyzed and shared data, during the past 40 years (Webb & Bryson 1972; Davis 1995), for thousands of modern pollen samples from all corners of the continent to create the North American Pollen Database (NAPD) within the NAPD (Whitmore et al. 2005). The NAPD surface sample dataset contains samples from core-top samples of sites where fossil data was collected, and from surface samples collected regionally and locally for the specific purpose improving calibration of paleoenvironment reconstructions. Each modern surface sample, from the NAPD, has meta data associated with it: the latitude and longitude of each site, the year the collected, the name of the researcher who collected the sample, depositional environment, and vegetation type derived from either cartographic sources or satellite imagery (AVHRR) (Whitmore et al. 2005). The climate data is also included in the dataset and each site was assigned climatic values, such as mean

annual precipitation and mean annual temperature, based on 10' latitude by longitude grid produced by the Climatic Research Unit (New et al. 2002).



Figure 1.2 North American Modern Pollen Database (NAPD) sample sites are represented by grey dots, while Commerford's (2010) Flint Hills Tallgrass Prairie sample sites are represented by orange dots.

One key use of the NAPD has been paleoclimatic reconstructions, where a modern pollen assemblage is statistically matched with its climate (Fig 1.4). For this technique to be successful there must be adequate representation of pollen assemblages from all climate types.

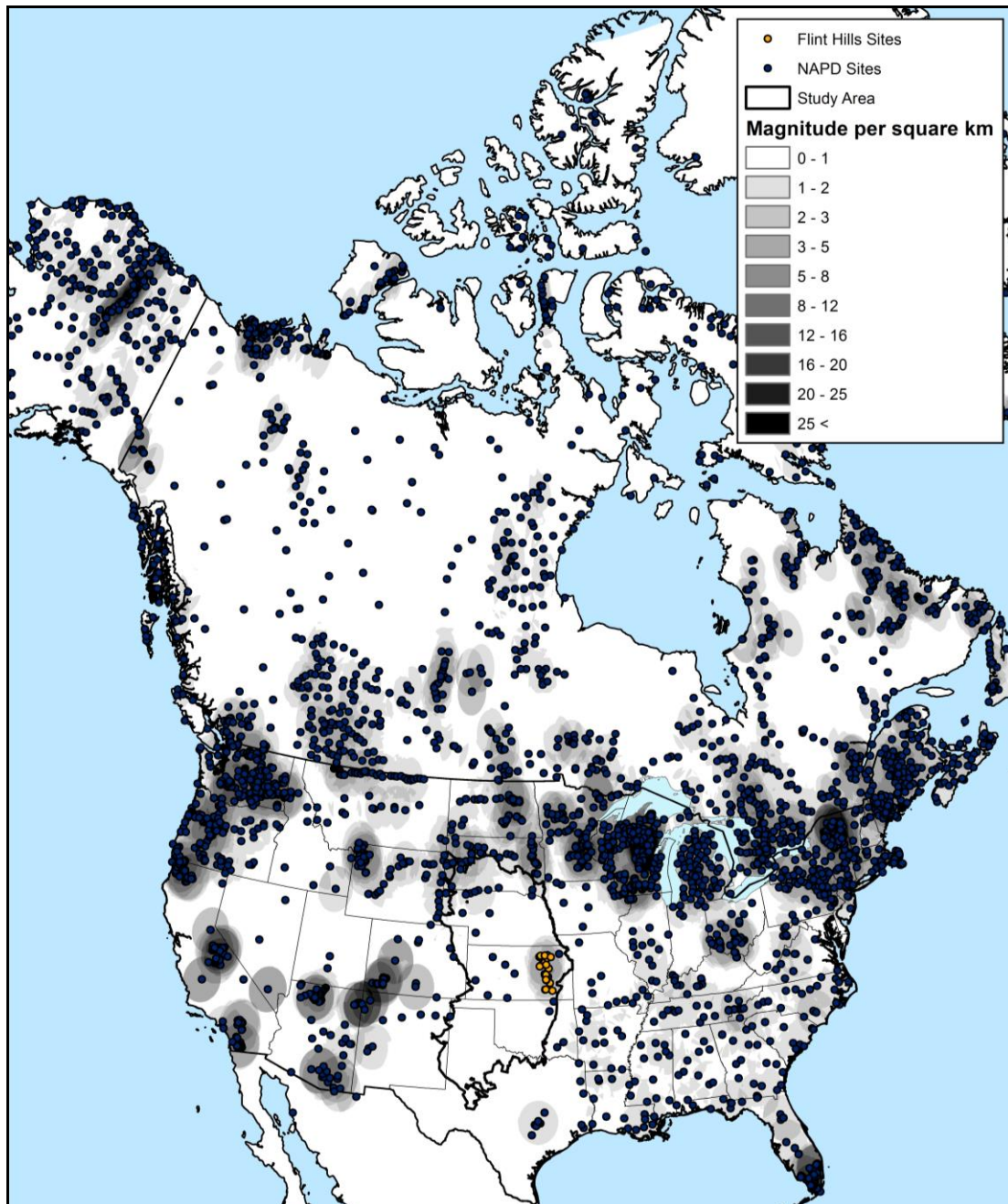


Figure 1.3: Point Density Map of the NAPD and Commerford's (2010) sites, where the darker the color the higher density of points.

And with over 4,800 surface samples, and with more samples being submitted for inclusion regularly (McLauchlan et al. 2013), the NAPD (Fig.1.2) is an incredible resource for paleoenvironment and paleoclimate researchers (e.g. Shuman et al. 2009; Viau et al. 2012). Unfortunately as extensive and valuable as the NAPD is, it still does not cover the entire continent sufficiently and the biomes where lacustrine environments are rare (i.e. parts of the desert southwest, the Great Basin, and the southern and central Great Plains) are underrepresented in the NAPD (Fig1.3). Not only are the prairies of the southern and central Great Plains spatially underrepresented they are also climatically underrepresented as shown by Whitmore et al.'s (2005) climate space figure (Fig1.4). Figure 1.4 shows that the prairie biome

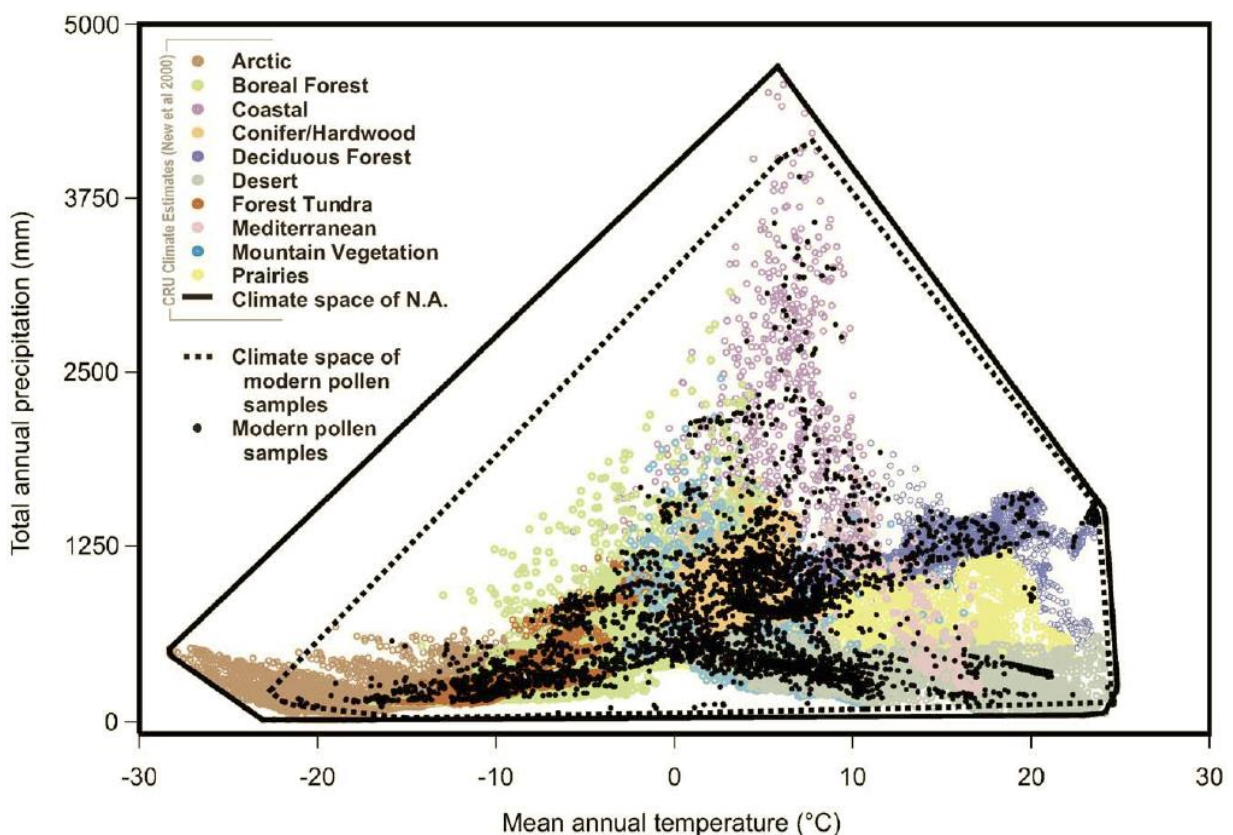


Figure 1.4: Whitmore et al.'s (2005) graph shows the mean annual precipitation and the mean annual temperature of each sample site of the NAPD (black dots). Each color represents a different biome.

(yellow circles) existing in the climate space where mean annual precipitation is between 400 mm and 1000 mm and the mean annual temperature is between 9° C and 21° C. This climate space is noticeably lacking in the number of black dots representing the modern pollen samples. The paucity of modern surface pollen samples within this climate space, that covers much of the southern and central Great Plains, has led to a deficiency in understanding regional vegetation responses to past shifts in climate (Clark et al. 2002).

Modern Analog Technique (MAT)

There are three conventional methods for interpretation of pollen assemblages: qualitative, quantitative, and the MAT. Qualitative analysis uses raw pollen percentages to infer changes to the composition of surrounding vegetation (von Post 1916; Zerniskaya & Mikhailov 2009). Quantitative analysis takes into account catchment basin size and relevant source area and its vegetative characteristics, known as the relevant source area of pollen (Sugita 1994). The MAT is a different type of quantitative method that statistically compares modern and fossil pollen assemblages through the use of a multivariate dissimilarity coefficient (Overpeck et al. 1985; Wahl 2003). The best chance for correct inferences to be made using the MAT is to ensure that there is large array of modern surface pollen samples that is spatially extensive and representative of the modern landscape (Bartlein et al. 1998). In North America the MAT is the most commonly used method for reconstructing past climates because of the availability of such a large array of both fossil and modern pollen assemblages (Whitmore et al. 2005).

The MAT works by calculating the compositional dissimilarity between the fossil pollen assemblages and each modern assemblage. Then modern analogs are determined for the fossil pollen assemblage. Then environmental metrics of the most closely matched modern samples are averaged and then assigned to the fossil sample, thus allowing for the reconstruction of past

vegetation (Williams & Shuman 2008). The relationship between vegetation types and climate has been well established (Thompson et al. 1999), thus allowing for past climate to be inferred from the vegetation present on the past landscape. The decision on what the threshold should be for a match between pollen assemblages is made with the intent to reduce false positive error without over inflating the false negative error (Wahl 2004). Threshold values are commonly chosen through the comparison of paired modern pollen assemblages, with a distance metric, from the same vegetation type to produce analog distance values and between vegetation types to produce no-analog distance values (Overpeck et al. 1985; Davis 1995). A threshold value is then decided upon that splits the analog and no-analog distance values.

Three classes of dissimilarity coefficients were identified by Overpeck et al. 1985; unweighted, equal weight, and signal-to-noise. The unweighted distance coefficients do not adjust for pollen types with large ranges and therefore are heavily influenced by the pollen types with the highest proportions. The equal-weighted distance coefficients down-weight the high proportion pollen types and increase influence of the low proportion pollen types, thus giving the rare pollen taxa greater influence on the coefficient (Prentice 1980). The Canberra Distance Metric (CDM) is an example of this type of coefficient. The signal-to-noise distance coefficients are less influenced by rare pollen types than the equal-weighted metrics, but more influenced by rare pollen types than the unweighted metrics (Gavin et al. 2003). The Squared Chord Distance (SCD) is the most commonly used signal-to noise-metric (Overpeck et al. 1985; Lytle & Wahl 2005).

Typically a threshold value between analog (matches) and no-analog (non-matches) is chosen after the comparison with a distance metric of paired modern pollen assemblages, from the same vegetation type used to produce analog distance values, and between vegetation types

used to produce no-analog distance values and a value between the two sets of dissimilarity coefficients is selected to be the threshold value for that particular study (Overpeck et al. 1985; Davis 1995). Receiver Operator Characteristic (ROC) analysis is widely used in the medical field to help assess which test is best at diagnosing the presence/absence of a disease (Metz 1978; Henderson 1993). Although the ROC method is popular in the medical field, it has been rarely used to compare the dissimilarity metrics used in the modern analog technique. Some preliminary results have been promising (Gavin et al. 2003; Oswald et al. 2003; Wahl 2004). The ROC method can be used to assess how well one distance metric identifies analogs versus no-analogs compared to another distance metric (Oswald et al. 2003). The ROC can also be used to determine a decision threshold for determining the analog/no-analog cutoff value (Wahl 2004).

Although there are many modern plant communities similar to past plant communities not every plant community that has occurred in the past is present in the modern sample set. These past plant communities that are “unknown” in the modern sample set are called no-analog communities (Williams & Jackson 2007). The two most common explanations for no-analog communities are: 1. No modern analog truly exists for that particular past ecological community or 2. A modern analog does exist for that past ecological community but it has yet to be sampled.

Prairies of the Great Plains of North America

The Great Plains of North America is an expansive area of rolling plains located in the central United States and south central Canada. Grasslands have been the dominant vegetation type of the Great Plains of North America for most of, if not all of the Holocene (Brown 1993). They also cover more area in North America than any other biome (Risser et al. 1981), and with such a wide range, spatially and temporally, it is not surprising that grasslands are tolerant of a

wide range of climatic variability on time scales from hours to millennia. Droughts are a common occurrence across the prairie biome and can last for months, years, or even decades (Borchert 1971; Clark et al. 2002). To make growing conditions even more difficult in this region, the soil water content is in a state of perpetual flux because of the ever changing temperature, wind, and precipitation patterns (Risser et al. 1981), meaning that native plants must be hearty and tolerant of rapidly changing conditions. Along with being drought tolerant and temperature adaptable, vegetation of the Great Plains must also be fire tolerant (Changnon et al. 2002; Axelrod 1985). With fire occurrence intervals being estimated as low as 2-3 years in some regions, fire is a major factor in the prevalence and preservation of grasslands by reducing woody plant abundance (Allen & Palmer 2011; Axelrod 1985). Sankaran et al. (2005) found that when mean annual precipitation was over ~650 mm fire becomes an essential process in protecting the grasslands from woody encroachment. Fire and other disturbances (i.e. herbivory) are not only important for maintaining grasslands in North America but also around the world (Briggs et al. 2002; Dalle et al. 2006).

The Great Plains are dominated by three major vegetation types: tallgrass, mixed grass, and short grass prairies. These vegetation types are defined by their composition with different grasses and forbs being dominant in each prairie type (Küchler 1972). When the first Europeans crossed the Great Plains of North America they noticed a change in the grasslands and tried to map the differing grassland types they encountered while exploring the vast sea of grasses. However, while the general east-west, tallgrass to short grass, pattern is easily discernible, defining the exact boundaries between each prairie type is much more difficult (Küchler 1972). Each prairie type is classified by its composition of grasses (Poaceae) and forbs with their

associated heights resulting in shorter grasses growing in drier climates and taller grasses in wetter climates.

The early explorers failed to see the potential of these grasslands including explorer Stephen H. Long who went so far as to label this region the “Great American Desert” in maps of his 1820 expedition (James 1823). This inaccurate opinion of the Great Plains, however, did not last long as European immigrants and their descendants settled, ranched, and farmed the mid-continent, leading to the destruction of a large percentage of the native prairies. Due to the rise of agriculture across the Great Plains over the last 150 years North American prairies have become one of the most endangered biomes on the continent, with as little as 13 percent of the tallgrass prairie’s historical extent still surviving today (Samson et al. 2004). The shorter grass prairies fared better than the tallgrass prairies because of the east-west precipitation gradient with the shorter grasses appearing farther west where it is drier, and the tallgrasses appear in the much wetter east (Lane et al. 2000).

Tallgrass Prairies

The dominant grass species (Poaceae family) of tallgrass prairies are Big Bluestem (*Andropogon gerardii*), Little Bluestem (*Schizachyrium scoparium*), Indian Grass (*Sorghastrum nutans*), and Switch Grass (*Panicum virgatum*) (Küchler 1972). Most of the original tallgrass prairies have been converted to cropland because the organic-rich soils are extremely fertile (Samson et al. 2004). The largest remaining region of tallgrass prairie on the continent is located within the Flint Hills of Kansas. Due to the steep slopes and shallow, rocky upland soils the tallgrass prairies of this region were never plowed under and converted into croplands, as was the case with most other regions of tallgrass prairie. Of the three Great Plains prairie types, tallgrass prairies are found in regions that receive the highest amounts of annual precipitation. According

to Brown (1947) the annual rainfall for tallgrass prairies ranges from about 63.5 cm annually along the western edge to 99 cm annually along the prairie forest boundary in the east. The higher availability of water allows the grasses and forbs of this region to grow taller than the same plant species in the drier regions of the mixed grass and short grass prairies. Another effect of higher moisture availability is that the dominant species of the tallgrass prairies are the species are better-suited to growing taller, with some species like *Andropogon gerardii* that grows up to 8 feet tall (Owsley 2011), in order to compete with other plants for sunlight (Lane et al. 2000).

Short Grass Prairies

The short grass prairies are found in the higher (in elevation) and drier regions of the Great Plains. The dominant grass species are Blue grama (*Bouteloua gracilis*), Buffalo grass (*Bouteloua dactyloides*), and Hairy grama (*Bouteloua hirsute*) (Küchler 1972). The short grass prairies consist of drought-resistant grasses and forbs that generally grow to heights of 30 cm or less (Lane et al. 2000). While the short grass prairie dominates western Great Plains, mixed grass and tallgrass species can be found along riparian zones deep into the short grass dominated regions and during wet years these species will even invade out into the short grass prairies (Küchler 1972). These invasions are always short-lived because as drier conditions return the tallgrass and mixed grass species die back and are replaced by the more xeric tolerant short grasses. The average annual precipitation for this region ranges from 25 cm in the far west to 50 cm in west-central Great Plains (Risser et al. 1981).

Mixed Grass Prairies

The mixed grass prairie region is the most difficult of the three vegetation types to classify because as Risser et al. (1981) states the “mixed prairie dominants are derived from the

two adjacent communities”, with those two communities being the short grass and tallgrass prairies to the west and east respectively. The dominant species within the mixed grass prairie biome are Big Bluestem (*Andropogon gerardii*), Little Bluestem (*Schizachyrium scoparium*), Side-oats grama (*Bouteloua curtipendula*), Sand-dropseed (*Sporobolus cryptandrus*), Western wheat grass (*Agropyron smithii*), Red three-awn (*Aristida longiseta*), June grass (*Koeleria cristata*), Needle-and-thread grass (*Stipa comata*), and Blue grama (*Bouteloua gracilis*) (Küchler 1972).

Many of these species can be found in either tallgrass or short grass prairies leading to the conclusion that there are no unique species within the mixed grass prairie ecoregion. Risser et al. (1981) describes the eastern boundary of this region as the area where, as a result of plant competition, the short grass species are no longer able to compete as a dominant grass type. The western boundary is described much the same way only with the tallgrasses being outcompeted due to drier conditions. The floristic diversity of the mixed grass prairie creates a situation where its geographical boundaries are not constant and instead they shift from year to year and decade to decade depending on climatic conditions (Küchler 1972).

Chapter 2 - Introduction

Fossil pollen can potentially be useful in reconstructing paleoenvironmental vegetation regimes. However, interpreting fossil pollen records can be difficult because there is not a direct relationship between the amount of pollen in a sediment record and the abundance of the vegetation on the landscape. In North America, fossil pollen records are the most spatially extensive paleoenvironmental record available for use in past vegetation and climate reconstructions and they also lend themselves easily to multi-scale analysis, both temporally and spatially (Whitmore et al. 2005.; Fyfe et al. 2009). The modern analog technique (MAT), a statistical method that compares modern and fossil pollen records, has been widely used across North America to reconstruct paleo-vegetation and paleoclimate (Overpeck et al. 1985; Wahl 2003). Despite the popularity of the MAT, most studies have focused on the forested regimes of eastern and western North America (e.g. Webb III 1987; Bartlein et al. 1998; Jackson et al. 2000), leaving the central and southern Great Plains sparsely sampled and studied. As the climate continues to warm across the Great Plains, the underrepresented climates of the southern Great Plains are likely to shift in novel ways across much of central North America (Williams et al. 2007). With a shift in climate, a shift vegetation composition is likely to follow (Jackson & Williams 2004). By examining modern and fossil pollen assemblages from the central and southern Great Plains, we can assess if future climate conditions along with their associated plant communities have occurred in the past, exist today but in another location, or are from a novel climate and/or have a no-analog plant community. In order to test these possibilities modern pollen data is needed from the central and southern Great Plains. Modern pollen data is a necessity because Overpeck et al. (1985) “the analog method is most likely to succeed in regions with extensive, representative collections of modern pollen data”.

The MAT is a popular method for reconstructing paleoenvironments in North America because of the large, spatially and temporally extensive databases containing thousands of fossil and modern pollen assemblages, e.g. the North American Pollen Database (NAPD), available for analysis (Whitmore et al. 2005). The MAT uses a multivariate distance metric to calculate the compositional dissimilarity between the fossil pollen assemblages and modern pollen assemblages. Then the environmental metrics of the most closely matched modern samples are averaged and assigned to the fossil sample, thus allowing for the reconstruction of past climates and vegetation types (Williams & Shuman 2008). The influence of rare and common pollen types on MAT analysis is dependent on which distance metric is used (Prentice 1980; Overpeck et al. 1985).

The most commonly used distance metric is the squared chord distance (SCD) which works best in regions where the most common pollen types are the most important for similarity classification, i.e. most forested regions (Overpeck et al. 1985; Jackson et al. 2000; Viau et al. 2006). A distance metric that allows greater influence of rarer pollen types in calculating dissimilarity coefficients is the Canberra Distance Metric (CDM), which has been used successfully in regions with low vegetational diversity, e.g. the Arctic tundra (Oswald et al. 2003). The decision on what the threshold value should be for a match between fossil and modern pollen assemblages is made with the intent to reduce false positive error without over inflating the false negative error and can be done using receiver operator characteristic (ROC) analysis (Wahl 2004). However, there are no presupposed criteria for determining a threshold between similar (analogs) and dissimilar (no-analogs) pollen assemblages (Gavin et al. 2003). In forested regions using 64 pollen types a cutoff metric of dissimilarity coefficients between 0.2

and 0.3 is recommended (Williams & Shuman 2008), but in grasslands it has been shown that fewer pollen types with a much lower cutoff metric of 0.12 can be effective (Hoyt 2000).

Many recent studies involving paleoenvironment reconstructions in North America rely upon a large array of modern pollen assemblages used to provide a basis in which to compare the fossil pollen assemblages used in each study (Davis 1995; Bartlein et al. 1998; Jackson et al. 2000). Most of the research has focused on biomes where perennial lacustrine environments are common and accessible, and these studies have help contribute to the creation of the NAPD. However, several regions and modern climate types are underrepresented within the NAPD including the central and southern Great Plains (Whitmore et al. 2005). With more modern pollen samples the relationship between climate and grasslands can be better understood and the accuracy of grassland reconstructions will be improved (Tonello & Prieto 2008; Commerford 2010). Sampling these underrepresented regions may lead to fewer no-analog situations because the two most common explanations for no-analog communities are: 1. No modern analog truly exists for that particular past ecological community or 2. A modern analog does exist for that past ecological community but it has yet to be sampled (Williams & Jackson 2007). Acquisition of new modern pollen samples is an active area of research in many regions worldwide, for instance in China numerous studies on the relationship between surface pollen and modern vegetation in underrepresented regions have recently been conducted (Luo et al. 2010; Zhao et al. 2011).

There are three major grassland biomes found in the Great Plains – short grass, mixed grass, and tallgrass prairies – with each defined by different compositions of grass types and forbs (Küchler 1972). Short grass prairies are found in the drier (~25 cm to ~50 cm mean annual precipitation) western Great Plains (Risser et al. 1981) and the dominant grass species are Blue

grama (*Bouteloua gracilis*), Buffalo grass (*Bouteloua dactyloides*), and Hairy grama (*Bouteloua hirsute*) (Küchler 1972). Tallgrass prairies are found in the wetter (~64 cm to ~99 cm mean annual precipitation) eastern Great Plains (Brown 1947) and the dominant grass species are Big Bluestem (*Andropogon gerardii*), Little Bluestem (*Schizachyrium scoparium*), Indian Grass (*Sorghastrum nutans*), and Switch Grass (*Panicum virgatum*) (Küchler 1972). Mixed grass prairies form a buffer between the short and tallgrass prairies and as a result are dominated by a compositional mix of both prairie types. Unfortunately, in paleo-vegetation and paleoclimate reconstruction studies using pollen as a proxy, all three grassland types have been lumped into a generic “prairie” category (Whitmore et al. 2005), despite the differences in composition and climate spaces. This generic prairie category has resulted from the inability to distinguish grass (Poaceae) pollen for the purposes of species classification. Although the lack of morphological differences in grass pollen is a hindrance in prairie type classifications of pollen assemblages, it is by no means the only way prairie types can be classified through analysis of pollen assemblages. Forbs and other rarer pollen types may hold the key to delineating between prairie types (Hoyt 2000).

Objectives:

Through the collection and analysis of 32 modern pollen samples throughout central and southern Great Plains, my objectives were to:

1. Collect new pollen samples from underrepresented climate types to increase the potential for future success in the use of MAT for reconstructing past environments of the Great Plains of North America.
2. Assess the dissimilarity of the pollen assemblages from the new samples to the prairie samples from the NAPD sites and other published records such as Commerford (2010) or Hoyt (2000).
3. Determine what characteristics distinguish grassland pollen assemblages in North America among short grass, mixed grass, and tallgrass prairies.

Hypotheses:

H1: The modern pollen samples collected from the western short grass prairies will have the highest percentages of *Artemisia*, along with the lowest percentages of *Ambrosia* compared with the pollen data from the mixed and tallgrass prairies.

This is based on the findings of Hall (1994) and Hoyt (2000), where pollen data from Tauber traps in tallgrass and short grass prairies was collected and analyzed as well as Minckley et al.'s (2008) comments on how an array of modern pollen samples can identify unique vegetation patterns within a region which can potentially reflect regional climate patterns .

H2: The pollen signatures of the tallgrass, mixed grass, and short grass prairies will allow for the classification of each grassland type at a 0.12 or higher squared chord distance dissimilarity coefficient.

This is based on Hoyt (2000) who found that Tauber trap collected pollen data created distinct pollen signatures for each grassland type at a 0.12 dissimilarity level between mixed and short grasses and a 0.41 dissimilarity level between tallgrasses and other grassland types.

Chapter 3 - Study Area

The Great Plains of North America cover a large region of the mid-continent, extending from Canada to Texas from north to south and from the Rocky Mountains in the west to roughly the Kansas – Missouri border in the east. Grasslands are the dominant vegetation cover across the entire region, though there are a number of small gallery forests that border many of the rivers that dissect the plains. The Great Plains have a strong west to east precipitation gradient with very low annual precipitation (< 400 mm) in the west, to a much higher annual precipitation in the east (> 1000 mm). This precipitation pattern allows for different types of grasslands to form. These grasslands can be broken down into three ecoregions. The tallgrass prairies are found in the east along with higher amounts of precipitation, the short grass prairies are found in the west in areas of much lower amounts of precipitation, and the mixed grass prairies in between the other two prairie types where short grass dominated vegetation cover transitions to tallgrass dominated vegetation cover. With three different grassland types growing under different climatic conditions in relatively close proximity to one another makes the Great Plains an ideal location for studying differences in pollen assemblages sampled from different prairie types.

The area of interest within this study consists of a sub-region of the Great Plains defined by specific climatic parameters (Fig 4.1). This study area is a climate space defined by a mAT between 9° C and 21° C and a mAP between 400 mm and 1000 mm as well as being bound in the east by Iowa and Missouri (Fig 3.1). The climate space that represents my study area covers over 796,000 sq. km, and all three prairie types are represented within this climate space that extends from central South Dakota to northern Texas, north to south, and from approximately 95° W - 104° W, east to west.



Figure 3.1: The study area is shown in green and represents the region of the Great Plains with a climate between 9° C and 21° C mAT and between 400 mm and 1000 mm mAP.

Much of the natural prairies have been altered for agricultural uses because the organic rich soils of the Great Plains are well suited to the production of annual crops. The tallgrass prairies have been the most affected by the spread of agriculture across the Great Plains, with only 13.4% of its historical range remaining intact today (Samson et al. 2004). Only 29.1% of mixed grass prairies remain today and only 51.6% of short grass prairies remain (Samson et al. 2004). These percentages indicate that there is a higher likelihood of finding larger tracts of unaltered or minimally alter prairies in the mixed and short grass prairies compared to the tallgrass prairies. This is confirmed when comparing the amount of public lands in each

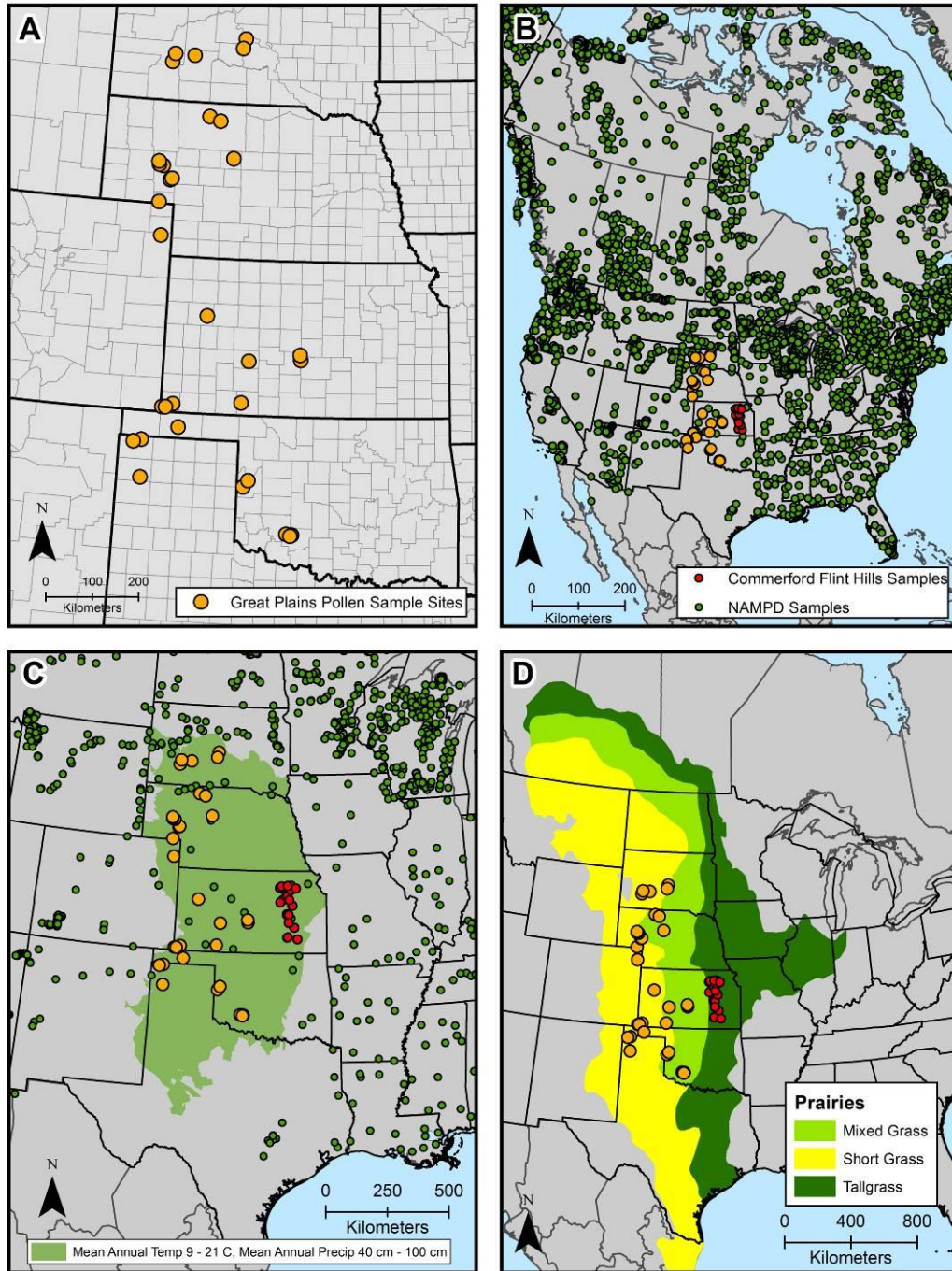


Figure 3.2: Panel A shows the location of Great Plains Pollen Sample (GPPS) sites. Panel B shows the location of the GPPS sites (orange) along with Commerford's Flint Hills sample sites (red), and the NAPD sample sites (dark green). Panel C shows the sample sites relative to the chosen climate space. Panel D shows the GPPS sites, Commerford's sites and the three prairie ecoregions as defined by Bailey's (1994) Ecoregions and Subregions Map of the United States.

prairie type, with every national grassland in the study area existing in the mixed or short grass prairie. Most of the samples collected in this study were obtained from public lands in mixed or short grass prairie. Since perennially standing water is at a premium in the drier climates of the mixed and short grass prairies many of the sites were supplied with water from windmills (Fig. 3.3).



Figure 3.3: This photo shows a windmill used to provide water to cattle and wildlife on the Thelander Ranch in the southwestern part of the Nebraska Sand Hills.

Chapter 4 - Methods

As previously stated in Chapter 2, the two main goals of this research are to increase the number of modern prairie pollen assemblages available for paleoenvironmental reconstruction, and to determine if it is possible to discern different prairie types through pollen assemblage analysis. The first goal is accomplished by selection of appropriate sites from data collection and the second is accomplished by a number of statistical methods including two variants of the MAT.

Site Selection

The sample sites were selected for a set of attributes to ensure a representative sample of the underrepresented prairie types. The attributes used in this selection process were: a mean annual temperature (mAT) between 9° C and 21° C; mean annual precipitation (mAP) between 400mm and 1000mm, the distance from a cultivated field of crops must be at least 1 km, the sites need to be at least 10 km from any previously collected site reported within the NAPD. The climate metrics were decided upon with the use of Whitmore et al.'s (2005) figure graphing the mAT and the mAP of each sample site within the NAPD. This figure shows the paucity of samples within the prairie biome, particularly the warmer, drier prairies (Fig 4.1). The 1 km distance from cultivated field crops was chosen to attempt to prevent the prairie vegetation regime's pollen signature from being overwhelmed by the anthropomorphic vegetation pattern. The 10 km buffer from previously sampled NAPD sites was used to ensure that new areas were sampled.

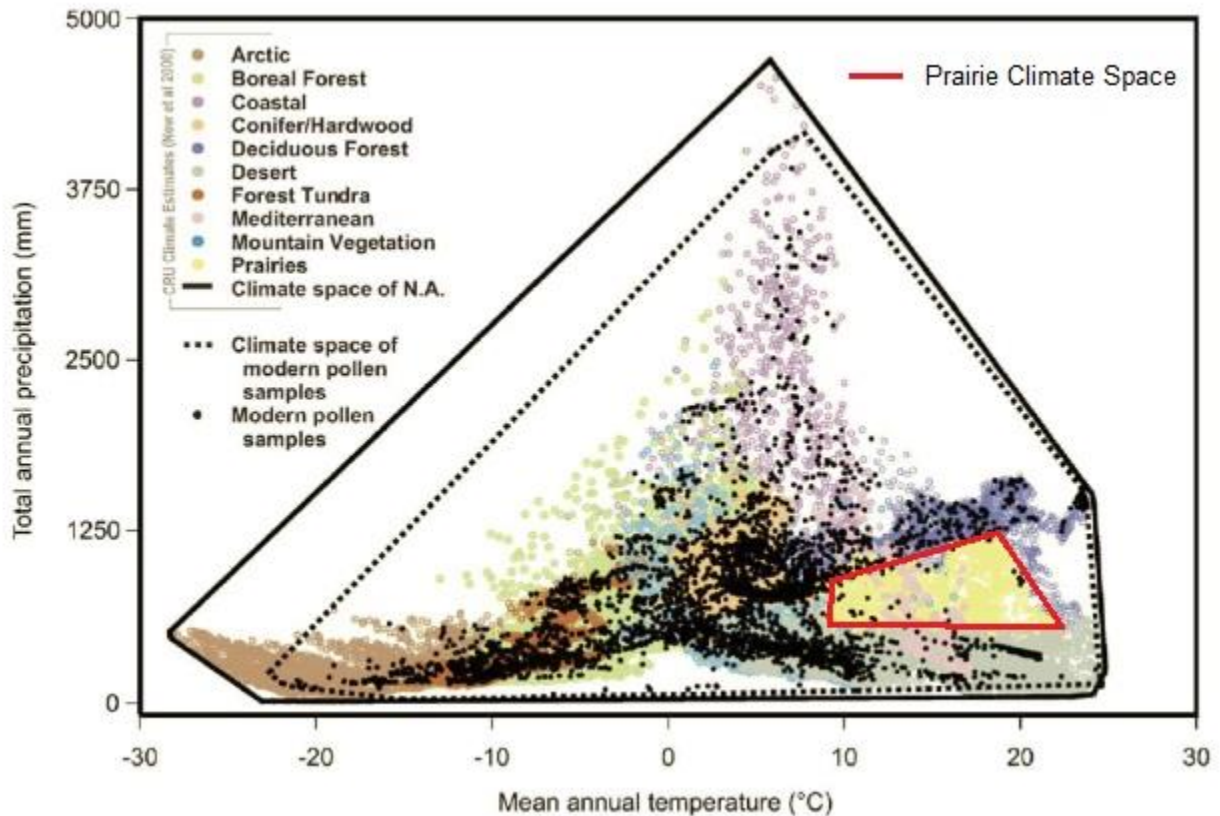


Figure 4.1: This is a modified figure from Whitmore et al. (2005) that shows the climate for each pollen assemblage within the NAPD. The lack of prairie samples sites is evident within the region highlighted by the red quadrangle (the climate space in which most prairies are found).

The climate data was obtained from the WorldClim 2.5-arcmin (4km) dataset (Hijmans et al. 2005). The WorldClim, Daymet, and PRISM climate models varied little when predicting climate parameters for any region with low topographic variability (Daly et al. 2008). Therefore, little difference is shown between models when defining the spatial extent of a climate space within the Great Plains. The data was then imported into ESRI's ArcMap for analysis, where the mean monthly temperatures were added together and divided by 12 to produce a predicted mAT and the mean monthly precipitation values were added together to produce a predicted mAP.

The climate data was then selected for regions that have a predicted mAT between 9° C and 21°

C and a predicted mAP between 400mm and 1000mm. The climate data was then further filtered to select only selected climate space found within the central and southern Great Plains of North America (Fig. 3.2).

Since a vast majority of the natural grasslands across the Great Plains have been converted to row crops, and with most of the remaining grasslands being privately held, public lands were identified first, for the easiest access. The public lands existing within the study area were assessed for parcels of open grassland at least one kilometer from any intensive cultivation and also containing a body of water to act as a pollen catchment. The national grasslands within the study area were the first to be surveyed for possible sites. Each national grassland was evaluated through the use of motor vehicle maps available to the public where each recorded water body was visually assessed for easy access to roads as well as nearness to cultivated fields. National wildlife refuges and state lands were also surveyed. The free downloadable program, Google Earth, was used to confirm the existence of the water bodies as well as the distance from roads and agricultural fields. Google Earth has a plethora of aerial and satellite imagery taken at multiple dates, even in remote areas, which helped determine whether or not a water body was ephemeral, seasonal, or permanent. All sites were visually assessed on Google Earth using the 'Historical Imagery' slider tool to ensure that the water body is not only perennial but also that there were no major changes to land use or land cover over the last 20 years or more. The potential sites were again checked for proximity to existing sites in the NAPD and Commerford's (2010) sites. The final sites were then chosen from areas that appeared to be representative of the grasslands biome, meaning that if a potential site was surrounded by a gallery forest, or some other disruption of the prairie, it was not used in this study. The sites

selected for this study came from a variety of land holders which included: Private, Non-profit, State lands, National Wildlife Refuges, National Grasslands, and National Forests.

Sediment Sample Collection

Surface sediment samples were collected from the 32 selected sites scattered across the Great Plains in October 2011 and in May and June of 2012 (Fig. 3.1). Of the 32 sites used in this study 28 of them were created through human activity, with the main function of these man made water bodies being to provide accessible, drinkable water for livestock. Three natural water bodies used in this study came from the Crescent Lake National Wildlife Refuge where the Ogallala Aquifer is very near the surface and provides a continuous source of water to the lakes and ponds of the western Nebraska Sand Hills. The other natural water body was Horseshoe Lake in Quivira National Wildlife Refuge in central Kansas. Of the 28 anthropomorphic water bodies 26 of them were created by earthen or concrete dams on small drainage basins with ephemeral streams. The remaining two sites were collected from cattle tanks with a buildup of at least 15 cm of sediment on the bottom of the tank.

Through the use of an 8 foot inflatable raft, a sit-on-top 8 foot kayak, or chest waders, an Eckman dredge was deployed to the approximate middle of the water body (Fig. 4.2). Sampling the sediment in the middle of a water body is common practice (e.g. Oswald et al. 2003; Seppa et al. 2004), because the sediment there is considered to be a good representation of the all the pollen that falls on the surface of the water body (Prentice 1985; Sugita 1994). An Eckman dredge is an open-bottom 6" x 6" metal box with spring loaded scoops which snap shut across the bottom opening of the box when triggered. Using the Eckman dredge, the top 2 centimeters of sediment were collected; 1 - 2 cm is standard for the collection of modern surface pollen



Figure 4.2: This photo demonstrates the use of an 8 foot inflatable raft to deploy an Eckman dredge for sediment collection from the middle of Ray Ranch Pond.

samples (e.g. Webb 1974; Oswald et al. 2003; Seppa et al. 2004; Zhao et al 2009). It has been found that sedimentation rates for some prairie sites were between 0.18 cm/year and 0.24 cm/year (Dean & Schwalb 2000). At those rates a 2cm surface sediment sample would represent 4 - 7 years of accumulation. Each site's position was recorded with a handheld global positioning system unit. The area around each pond, out to approximately 50 m, was also visually assessed for predominant species of prairie type, i.e. short grass, mixed grass, or tallgrass prairie. The indicator species chosen for this quick classification included grasses and forbs. Nearby tree types were also noted (Appendix A).

The sediment samples were transported from each site in an ice filled cooler and stored in a refrigeration unit until all the samples were collected. Then, 1.5 cubic centimeters of each sediment sample was placed into an individual 15 mL plastic centrifuge tube. The samples were then shipped to the Limnological Research Center which houses the National Lacustrine Core Facility (LacCore) at the University of Minnesota in Minneapolis, MN. There the samples

underwent LacCore's standard pollen processing procedure which involves a number of acid baths to dissolve away the unwanted materials but leaving the pollen grains intact and identifiable (Faegri & Iverson, 1989). The samples were then visually assessed in order to identify individual pollen grains taxonomically and counted to a sum of 100 through the use of a high-powered light microscope.

Data Analysis

After the pollen samples were counted, 47 pollen types were selected for use in the statistical analysis of the pollen assemblages. The 47 pollen types were chosen to improve the resolution of non-arboreal taxa to allow for a more in depth analysis of prairie biomes (McLauchlan et al. 2013). These pollen assemblages were then analyzed through the use of two different dissimilarity metrics as part of the MAT: squared chord distance (SCD) and Canberra metric distance (CMD). The SCD is a signal-to-noise metric versus the CMD which is an equal weighted metric (Overpeck et al. 1985). The use of receiver operating characteristic (ROC) analysis was used to determine which of the two metrics' values are better at distinguishing between vegetation types (Oswald et al. 2003). The ROC analysis was also used to produce a threshold value used in interpreting pollen assemblage "matches" and "non-matches" (Wahl 2004). The ratio of *Ambrosia* to *Artemisia* pollen counts within each pollen assemblage was also calculated to further investigate the use of this ratio as a simple diagnostic for prairie type.

Squared Chord Distance (SCD)

The SCD is a signal-to-noise metric meaning that the major pollen types are given slightly less influence and the minor (rarer) pollen types are given slightly more influence on the metric when compared to equal weighted or unweighted dissimilarity measures (Overpeck et al.

Equation 4.1: Squared Chord Distance

$$SCD = \sum_i^n (\sqrt{x_i} - \sqrt{y_i})^2$$

Where:

x_i = the proportion of pollen type i in sample x

y_i = the proportion of pollen type i in sample y

n = number of pollen types used in the analysis

1985). SCD is calculated by taking the square root of the proportion of a pollen type from the first pollen assemblage (x_i), and subtracts it from the square root of the proportion of the same pollen type from a second pollen assemblage (y_i) being compared with the first, and then the value is squared to ensure a positive value. This step is repeated for every chosen pollen type in the pollen assemblage and these values are summed up to give a final dissimilarity value. The entire process is repeated until every pollen assemblage is compared to every other pollen assemblage in the study producing a dissimilarity value for every comparison. The closer the dissimilarity value is to zero the less dissimilar, i.e. more similar, the two pollen assemblages are to each other. A threshold value, determined through ROC analysis, is then used to interpret pollen assemblage “matches” (less than the threshold value) and “non-matches” (greater than the threshold value).

With the use of R, an open source free statistical computer software program, each pollen assemblage from this study, the Great Plains pollen sample (GPPS) set, was compared using the SCD technique to every other pollen assemblage within the sample set. The GPPS assemblages were also compared, using SCD, to the Flint Hills tallgrass surface samples collected by Julie Commerford (2010), as well as being compared against all pollen assemblages classified as “Prairie” in the NAPD.

Canberra Metric Distance (CMD)

The CMD was also used to compare the various pollen assemblage sets. As an equal weight metric the CMD gives every pollen type the same amount of influence. This means that the rarer types have been up weighted and the most common types have been down weighted (Overpeck et al. 1985). CMD has been used successfully for biomes with low vegetational diversity such as the Alaskan tundra (Oswald et al. 2003). This is likely because the rarer pollen types were able to exert enough influence on the metric value to allow for classification. While the Great Plains prairies have a highly diverse vegetation regime, the rarer plant types may hold the key to classifying separate prairie types from one another, i.e. tallgrass to short grass.

Equation 4.2: Canberra Metric Distance

$$CMD = \sum_i^n \frac{|x_i - y_i|}{x_i + y_i}$$

Where:

x_i = the proportion of pollen type i in sample x

y_i = the proportion of pollen type i in sample y

n = number of pollen types used in the analysis

The CMD is calculated by finding the absolute value of the difference of the proportions of a single taxon between two pollen assemblages (x_i & y_i) and then this value is divided by the sum of these same proportions, this step is repeated for all pollen types in the pollen assemblages and then these values are added up to produce the final dissimilarity value. With the use of R every pollen assemblage was compared to every other pollen assemblage in the study so that all “matches” and “non-matches” could be identified with the help of a threshold value determined by ROC.

Receiver Operator Characteristic (ROC)

ROC analysis was used in this study to determine a discrimination threshold by finding the optimal dissimilarity value, or the value at which the minimum number of both the false positive (false matches) and false negative (false non-matches) classifications intersect (Wahl 2004). Both the SCD and CMD values were submitted to ROC analysis (performed in R statistical software) to produce discrimination thresholds that allowed for classification of like and un-like assemblages for each distance metric (Fig. 4.3). Each pollen assemblage was assigned a prairie type (short grass, mixed grass, or tallgrass) with the use of Bailey's (1994) Ecoregions and Subregions Map of the United States. This allowed for the comparison of how well each distance metric distinguished between prairie types using the optimal dissimilarity threshold value.

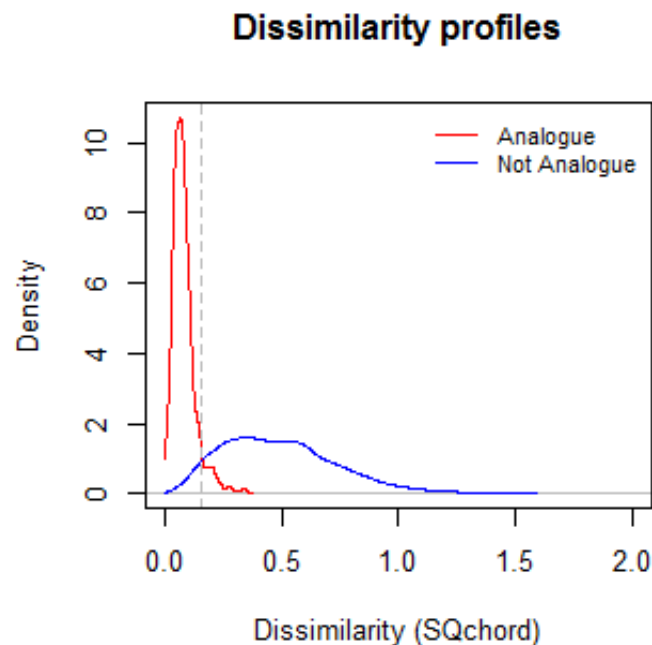


Figure 4.3: This figure shows how ROC analysis determines the optimal dissimilarity value (dotted gray line), which is where the true-positive/false positive line (red line) meets the true-negative/false negative (blue line).

Ambrosia/Artemisia Ratio Analysis

The ratio of the proportion of *Ambrosia* and the proportion of *Artemisia* in each prairie pollen assemblage was calculated. These ratios were then submitted to analysis of variance (ANOVA) and regression analysis to determine if the ratio between the two can be used as an indicator for prairie type in order to try and replicate the findings of Hall (1994), where the tallgrass prairies had a higher *Ambrosia/Artemisia* ratio than the drier mixed and short grass prairies.

The 559 prairie pollen assemblages were separated into 3 categories (Short Grass, Mixed Grass, and Tallgrass) based on Bailey's Ecoregions and Subregions Map of the United States (Bailey 1994). The Mann-Whitney-Wilcoxon Test was used to determine if there was a statistically significant difference between each of the prairie types' *Ambrosia/Artemisia* ratio distributions at a 95% confidence interval (e.g. Tallgrass - Mixed Grass; Tallgrass -Short Grass; Mixed Grass - Short Grass).

Chapter 5 - Results

Sample Site Locations and Climate Types

Samples were collected from 32 sites across the central and southern Great Plains, from locations not represented in the NAPD. The Great Plains Pollen Sample (GPPS) sites have increased the number and density of surface pollen sites in the western and southern Great Plains (Fig. 5.1). All but one site yielded enough pollen for rigorous statistical analysis. This achieved the objective collecting pollen assemblages from climate types underrepresented in the NAPD in order to increase the potential for future success in the use of the MAT in paleovegetation and paleoclimate reconstructions of the Great Plains of North America.

The climate space identified as underrepresented in the NAPD is a zone where maP is between 400 mm and 1000 mm and maT is between 9° C and 21° C. The climate types of the GPPS collected in this study are all within the specified climate zone (Fig. 5.2). Two distinct climate spatial patterns emerged from the sample sites: a general east-west precipitation gradient and a general north-south temperature gradient (Appendix B). For example the precipitation gradient is strongly demonstrated as maP values for sites west of 102° W are lower than 450 mm and maP values for sites east of 99° W are over 600 mm, regardless of latitude. The temperature gradient ranged from the highest maTs at ~19° C found at the southernmost sites in southwest Oklahoma to the lowest maTs at ~ 9.7° C found in the northern sites of central South Dakota. The warmer GPPS sites were collected from prairies with climate types that have not been previously sampled within the NAPD (Fig. 5.2).

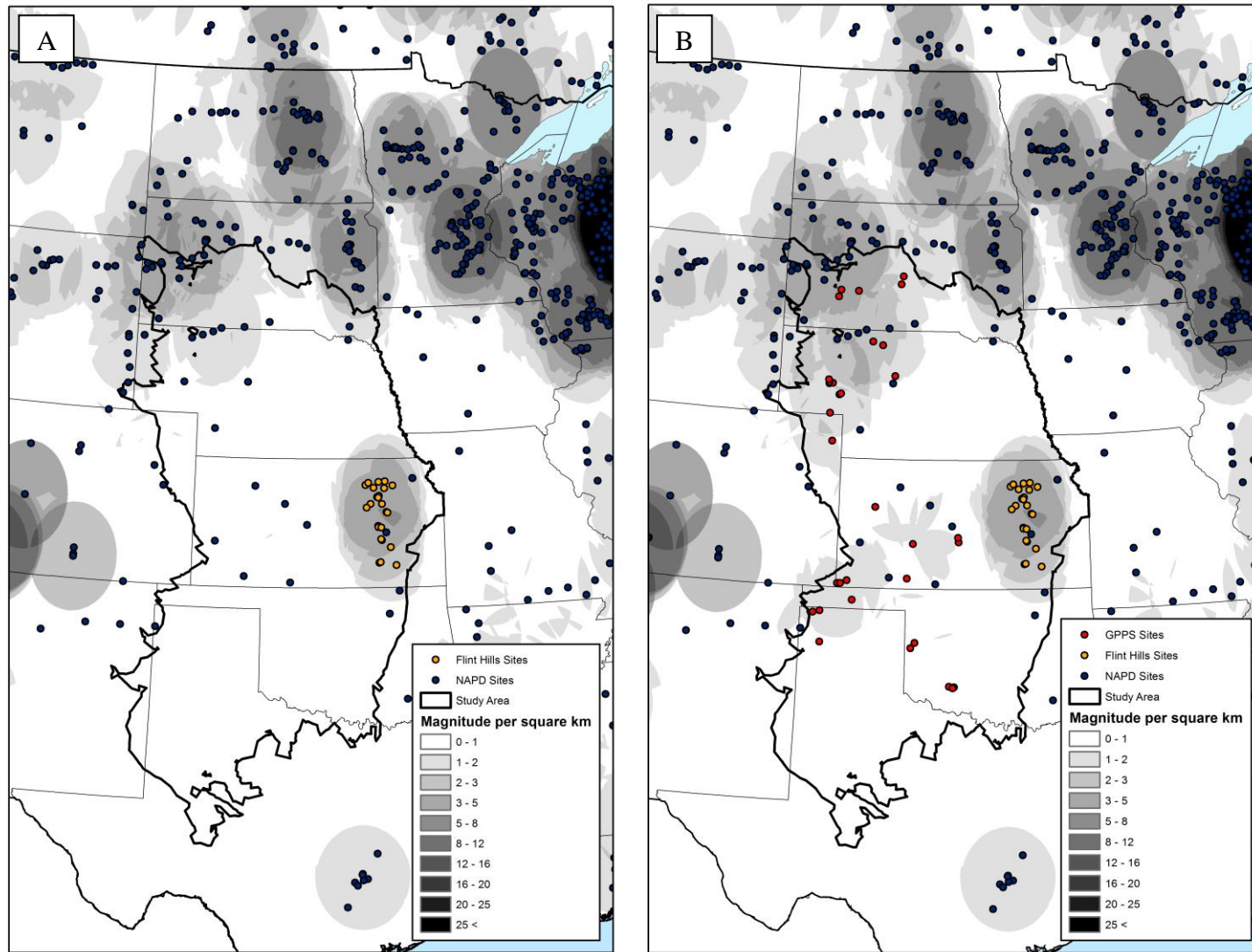


Figure 5.1: A) A point density map of all NAPD sites (dark blue dots) along with Commerford's (2010) Flint Hills sample sites (orange dots). B) A point density map that also includes the GPPS sites (red dots). The darker the gray the more sites nearby.

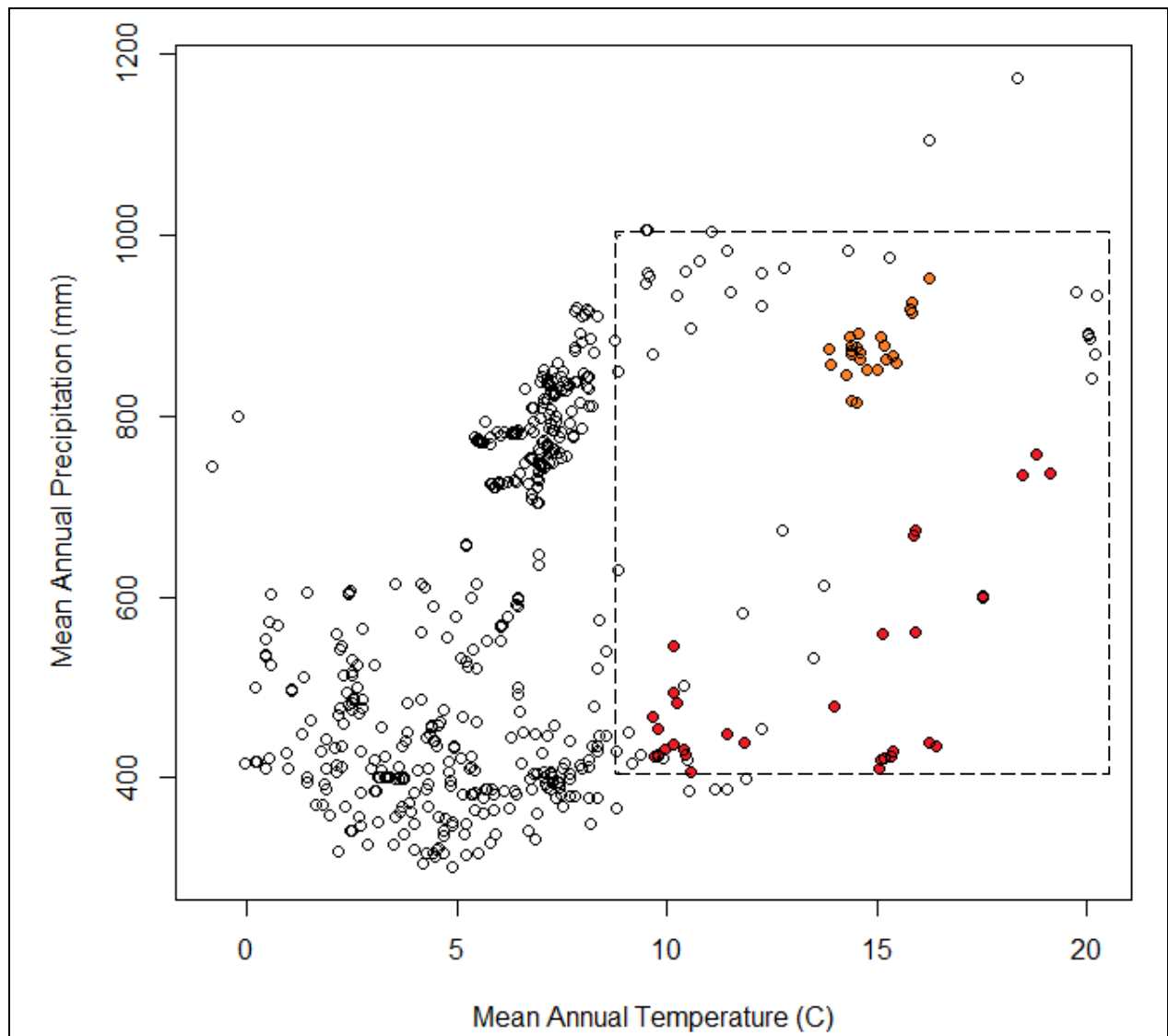


Figure 5.2: The climate space of all sample sites classified as “prairie” within the NAPD (shown as open circles), along with the samples collected by Commerford (2010)(orange circles) and the GPPS sites (red circles). The climate space used in this study (maP 9 - 21° C and maT 400 – 1000 mm) is denoted by the dashed line square.

The prairies of central North America exist in a large climatic zone but also are spatially extensive ranging from southern Texas to central Alberta and from Wyoming to Wisconsin (Fig. 5.3). The NAPD “prairie” sites can be separated into geographic regions for inter- and intra-regional comparisons. For example arboreal pollen is much more prevalent in pollen assemblages of prairies with neighboring forests of the upper Midwest than the vast treeless expanses of the western plains.

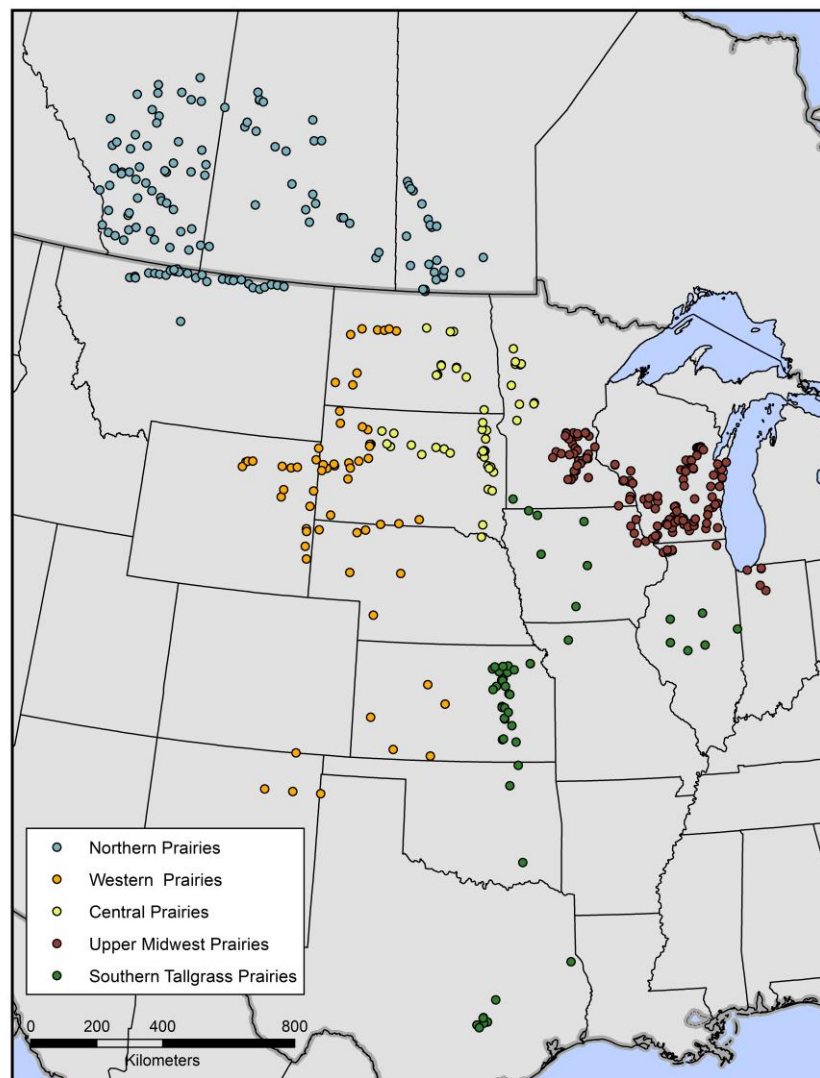


Figure 5.3: The NAPD “prairie” and the Flint Hills sample sites are shown here in regional groups to illustrate the different prairie environments all under the label “prairie”.

***Ambrosia* and *Artemisia* Analysis**

There are differences in the bioclimatic preferences of two common North American Grassland taxa: *Ambrosia* (ragweed) and *Artemisia* (sage). Therefore, the ratios between the proportions of *Ambrosia* and *Artemisia* pollen grains the same assemblage were compared for every sample within the NAPD classified as prairie, the Flint Hills Sample set (Commerford 2010), and the GPPS set collected during this study. Samples were reclassified into three prairie types short, mixed, and tallgrass prairies (Fig. 5.4). The pollen type with largest range of pollen assemblage proportions is short grass prairie *Artemisia* which varies from 0 to 0.739, but tallgrass prairie *Ambrosia* has the highest mean proportion of the all three prairie types and both pollen types (Table 5.1). Tallgrass prairies also have the lowest mean proportion of *Artemisia* along with the smallest range of the group at 0 to 0.256. The short grass prairies have the highest mean of proportions for *Artemisia* and the lowest mean of proportions for *Ambrosia*. The tallgrass prairies are the inverse of the short grass prairies with having the highest mean proportion of *Ambrosia* and the lowest mean proportion of *Artemisia*. The mixed grass prairies have mean proportions of both pollen types in between the means of the other prairie types, with the mean proportion of *Ambrosia* being slightly higher than that of *Artemisia*.

Ambrosia/Artemisia ratios ranged from 0 to 165 across all samples. The ratios of *Ambrosia/Artemisia* for each prairie type was compared to the other prairie types using the Mann-Whitney-Wilcoxon Test in order to test whether the distributions of *Ambrosia/Artemisia* ratios are different between prairie types without having to assume they have normal distributions.

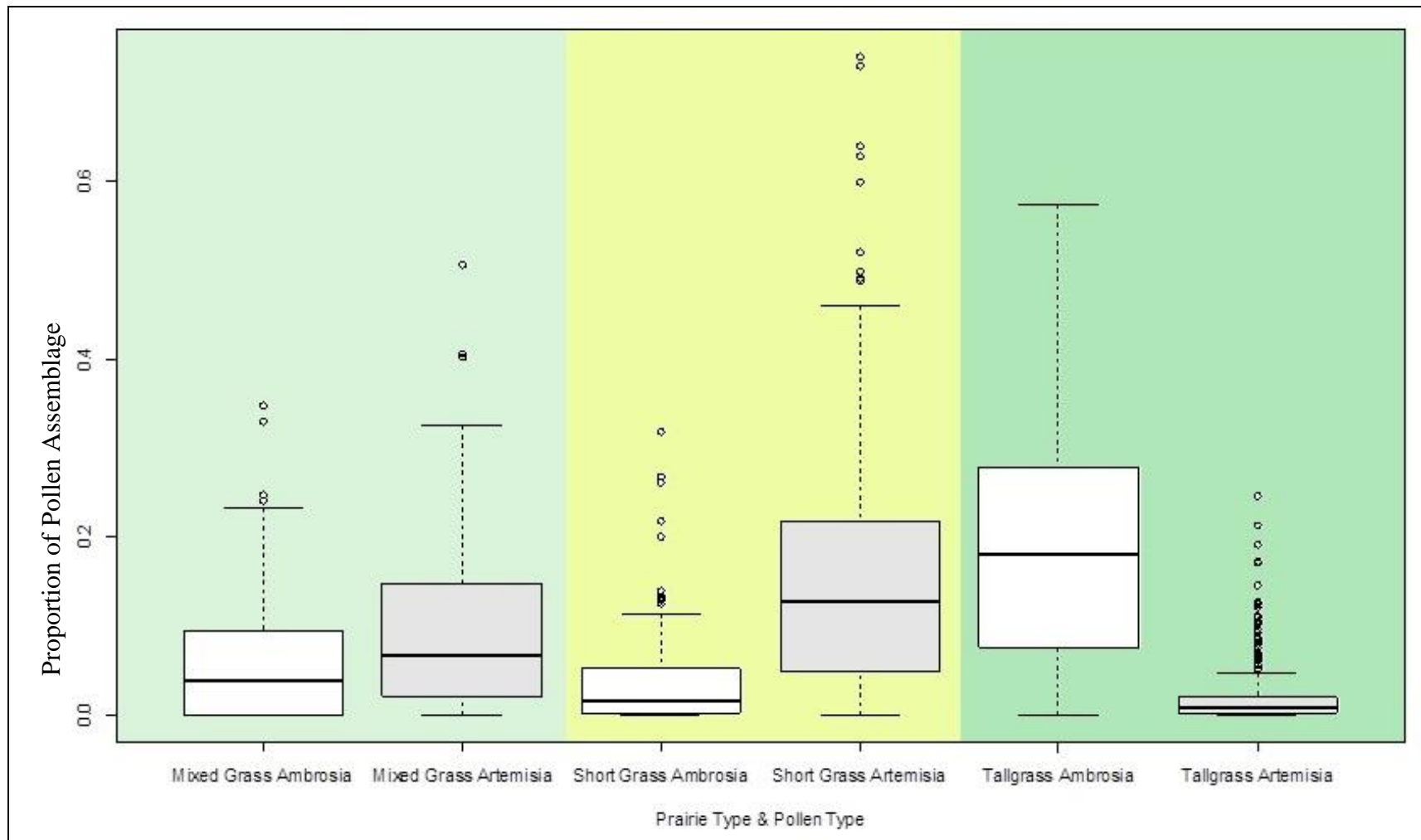


Figure 5.4: This boxplot of the proportions of *Ambrosia* (white) and *Artemisia* (grey) for every pollen assemblage classified as prairie within the NAPD, the Flint Hills Samples (Commerford 2010), and the GPPS sites, further broken down by prairie type.

Table 5.1: The mean proportions of both pollen types (*Ambrosia*, *Artemisia*) for all three prairie types (Mixed Grass, Short Grass, and Tallgrass) are shown as well as the ratio of means where (ratio = *Ambrosia*/*Artemisia*). * Average Ratios do not include values divided by zero.

Mean <i>Ambrosia</i> Proportion	Mean <i>Artemisia</i> Proportion	Ratio of Mean Proportions	Average Ratio*
Mixed Grass	Mixed Grass	Mixed Grass	Mixed Grass
0.067	0.055	1.21	3.81
Short Grass	Short Grass	Short Grass	Short Grass
0.037	0.161	0.229	0.628
Tallgrass	Tallgrass	Tallgrass	Tallgrass
0.185	0.022	8.27	22.09

Table 5.2: Mann - Whitney-Wilcoxon Test compares each prairie type's distribution of the ratios between *Ambrosia* and *Artemisia* with the other two prairie types' ratio distributions at a 95% significance level.

Prairie Type	W - value	p-value
Short Grass – Mixed Grass	7220.5	0.02397
Short Grass – Tallgrass	4165	< 0.000001
Mixed Grass – Tallgrass	3742	< 0.000001

This analysis shows that there is a significant difference in the distributions of *Ambrosia*/*Artemisia* ratios between tallgrass prairies and mixed grass ($p < 0.000001$) and between tallgrass and short grass prairies ($p < 0.000001$). There is also a significant difference in the distribution of *Ambrosia*/*Artemisia* ratios between short grass and mixed grass ($p = 0.024$).

Each pollen assemblage has a mean annual temperature (maT) and a mean annual precipitation (maP) associated with it, and these climate metrics were compared with the *Ambrosia*/*Artemisia* ratio in each pollen assemblage. There is an increase in variability of values of the *Ambrosia*/*Artemisia* ratio as maP increases over 500 mm (Fig. 5.5). Thus, samples with

low maP also have low variability of ratios. To further illustrate this pattern, a subset of the 559 pollen assemblages analyzed in this study, sites with maP of less than 500 mm were examined (n = 243) and only one site has a ratio value over 10 at 15.33. The site is GPPS 14 Aermo Pond, collected during this study, located in the north central Nebraska Sand Hills and has a maP of approximately 494 mm. The *Ambrosia/Artemisia* mean ratio for all sites below 500 mm maP is 0.717 and the range is from 0 – 15.33. The mean ratio for all sites above 500mm maP is 23.23 and the range is from 0 – 169. This ratio could potentially be used as a new tool for researchers involved in paleoenvironmental reconstructions of the Great Plains to help determine if a fossil pollen assemblage was produced during a period when the maP was wetter than 500mm.

There appears to be a similar pattern with temperature. The maT to *Ambrosia/Artemisia* ratio graph shows little variability in ratio values below 5 °C maT (Fig. 5.6). The range of ratios for the 162 sites below 5 °C maT is 0 to 11.43 and the mean ratio is 0.446. Above 5 °C maT the variability of ratio values increases dramatically with a range of 0 – 169 and a mean ratio of 17.96. As maT increases above 9 °C there appears to be a slight decline in variability.

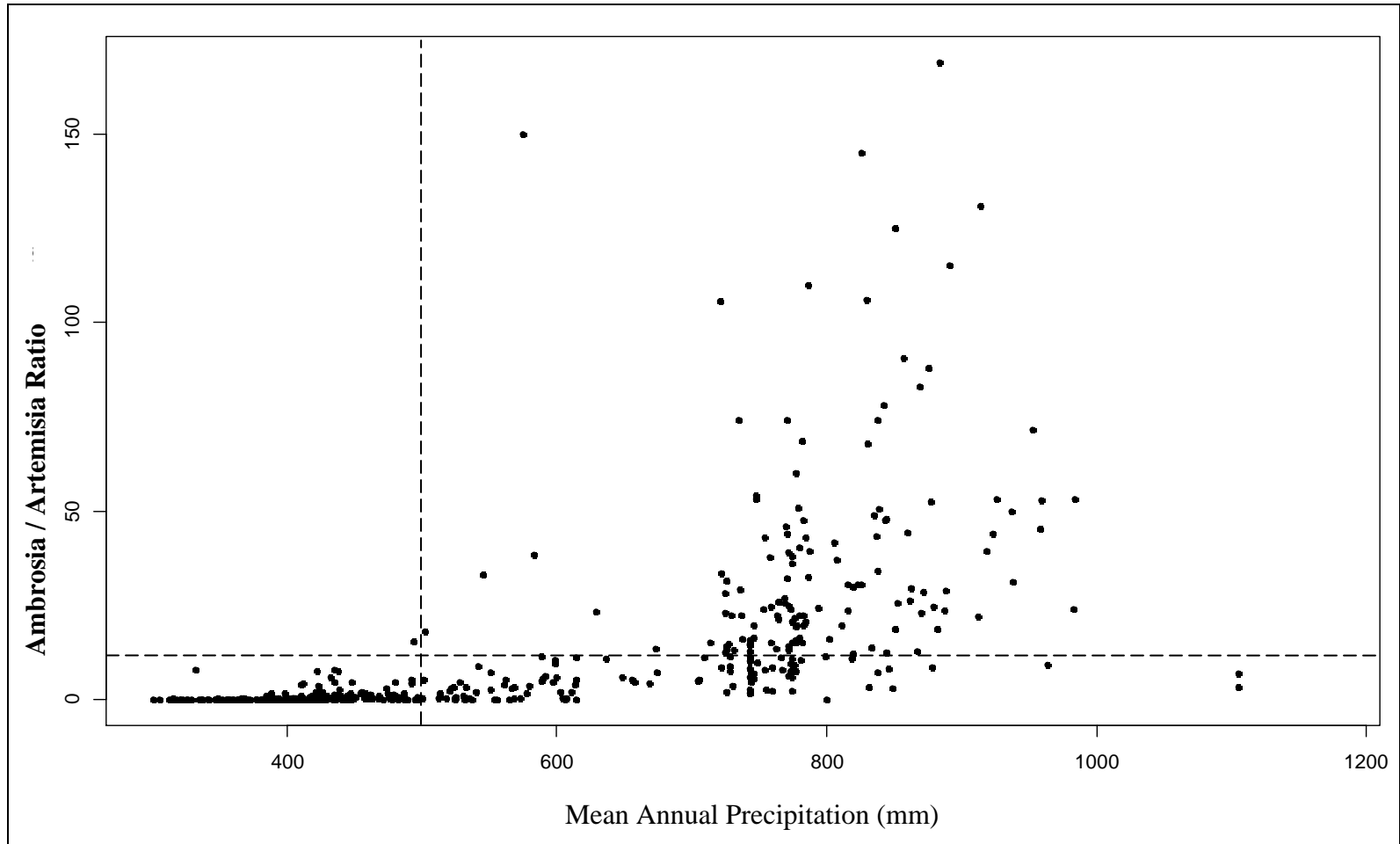


Figure 5.5: Mean annual precipitation (maP) is the independent variable (x-axis) and the *Ambrosia/Artemisia* is the dependent variable (y-axis). The vertical dashed line is the maP threshold at 500 mm and the horizontal dashed line is the *Ambrosia/Artemisia* ratio decision threshold of 10. Each pollen assemblage analyzed in this study is represented by a black dot (n = 559).

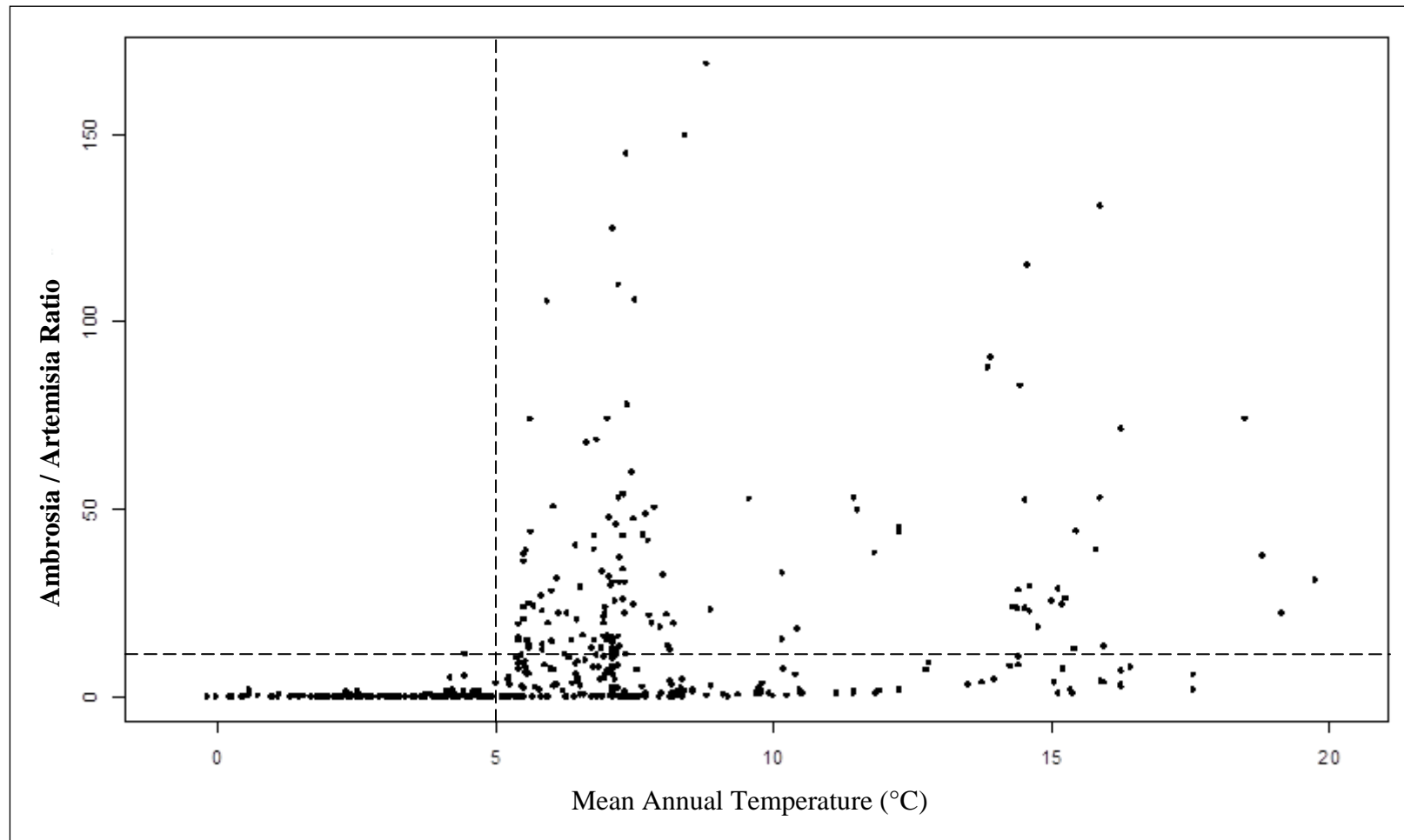


Figure 5.6: Mean annual temperature (maT) is the independent variable (x-axis) and the *Ambrosia/Artemisia* is the dependent variable (y-axis). The vertical dashed line is the maP threshold at 500 mm and the horizontal dashed line is the *Ambrosia/Artemisia* ratio decision threshold of 10. Each pollen assemblage analyzed in this study is represented by a black dot (n = 559).

MAT Dissimilarity Matrices

The GPPS pollen assemblages were compared to the NAPD samples and the Flint Hills samples using two different dissimilarity metrics, SCD and CDM , using 47 selected pollen types (Table 5.4). The GPPS pollen assemblages were also compared to themselves using the entire pollen assemblage to assess similarity between the newly collected sites. Dissimilarity matrices were created from each assessment and color coded with black/dark gray representing the least dissimilar values scaling to light gray/white representing the most dissimilar values. The SCD dissimilarity matrix shows that the GPPS set placed in between the mixed and short grass prairies and the Flint Hills sample set placed in between the mixed grass and tallgrass prairies (Fig. 5.7). A CDM dissimilarity matrix was created with the same layout; short grass, GPPS, mixed grass, Flint Hills samples, and then tallgrass (Fig. 5.8). The GPPS pollen assemblages have a number of samples with high similarity to each prairie type, and each sample is similar to at least one other sample within the set which can be observed in the SCD dissimilarity matrix (Fig. 5.9). A dissimilarity matrix using CDM was created as well and it shows that the values created using this metric have a larger range and will lead to higher dissimilarity values (Fig. 5.10). Each dissimilarity matrix compares every pollen assemblage to every other pollen assemblage within the dataset. To read the graphs follow a sample number vertically until the sample is compared to itself (where the line hits the diagonal cut-off), this line shows the dissimilarity values between the chosen sample and every sample with a lower sample number. To compare the chosen sample to every sample with a higher sample number follow the line horizontally from the diagonal to the right. The darker the color the more similar the two samples are to each other.

Table 5.3: List of the 47 pollen types used for MAT analysis in this study

<u>Scientific Name</u>	<u>Common Name</u>
<i>Abies</i>	Fir
ACER	Maple
ALNUS	Alder
<i>Ambrosia</i>	Ragweed
<i>Artemisia</i>	Sagebrush
ASTERACEAE	Daisy Family
<i>Betula</i>	Birch
Brassicaceae	Mustard Family
<i>Carya</i>	Hickory
Caryophyllaceae	Chickweed Family
<i>Celtis</i>	Hackberry
Chenopodiaceae/Amaranthaceae	Goosefoot/Pigweed Families
<i>Cornus</i>	Dogwood
<i>Corylus</i>	Hazel
Cupressaceae (<i>Juniperus</i>)	Cedar Family
Cyperaceae	Sedge Family
<i>Ephedra</i>	Jointfir
Fabaceae	Pea Family
<i>Fagus</i>	Beech
<i>Fraxinus</i>	Ash
<i>Iva</i>	Marsh Elder
JUGLANS	Walnut
Larix	Larch
<i>Liquidambar</i>	Sweetgum
Minor Forbs	Minor Forbs
Moraceae	Mulberry Family
<i>Ostrya/Carpinus</i>	Hophornbeam/Hornbeam
PICEA	Spruce
PINUS	Pine

Plantaginaceae

Plantanus

Poaceae

Polygonaceae

Populus

Quercu

Ranunculaceae

Rhamnaceae/Vitaceae

Rosaceae

Rumex

Salix

Sarcobatus

SAXIFRAGA

Shepherdia

Thalictrum

Tilia

Ulmus

Xanthium

Plantain Family

Sycamore

Grass Family

Buckwheat Family

Poplar, Aspen

Oak

Buttercup Family

Buckthorn/ Grape Families

Rose Family

Sorrel

Willow

Greasewood

Saxifrage

Buffaloberry

Meadow Rue

Basswood

Elm

Cocklebur

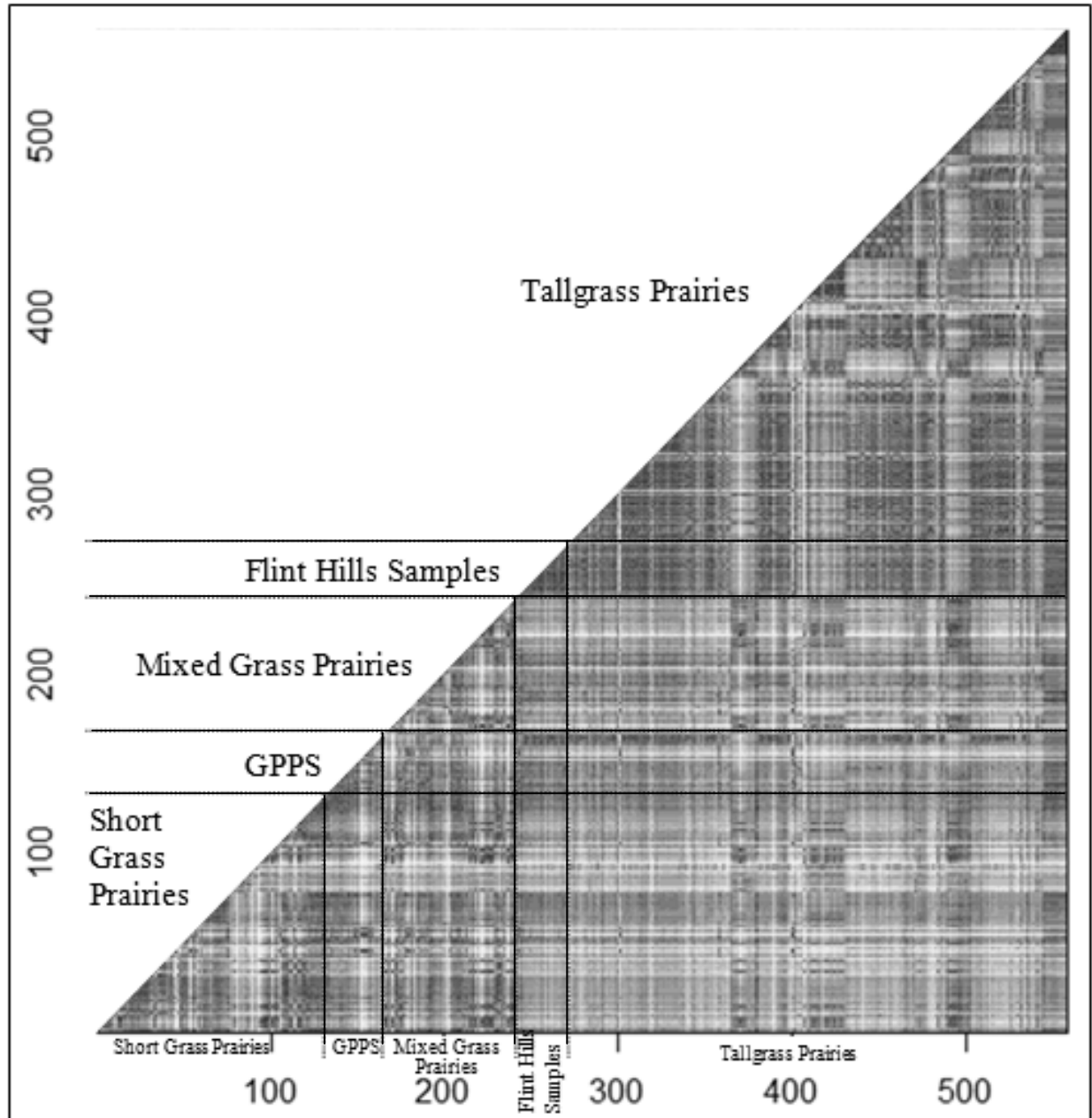


Figure 5.7: SCD dissimilarity matrix among all pollen assemblages from grasslands of North America with lower dissimilarity represented by reds and higher dissimilarity represented by yellows. The matrix is broken down by prairie type and the GPPS and Flint Hills sample sets. The numbers represent number of samples starting with 0 in the bottom left corner (n = 559).

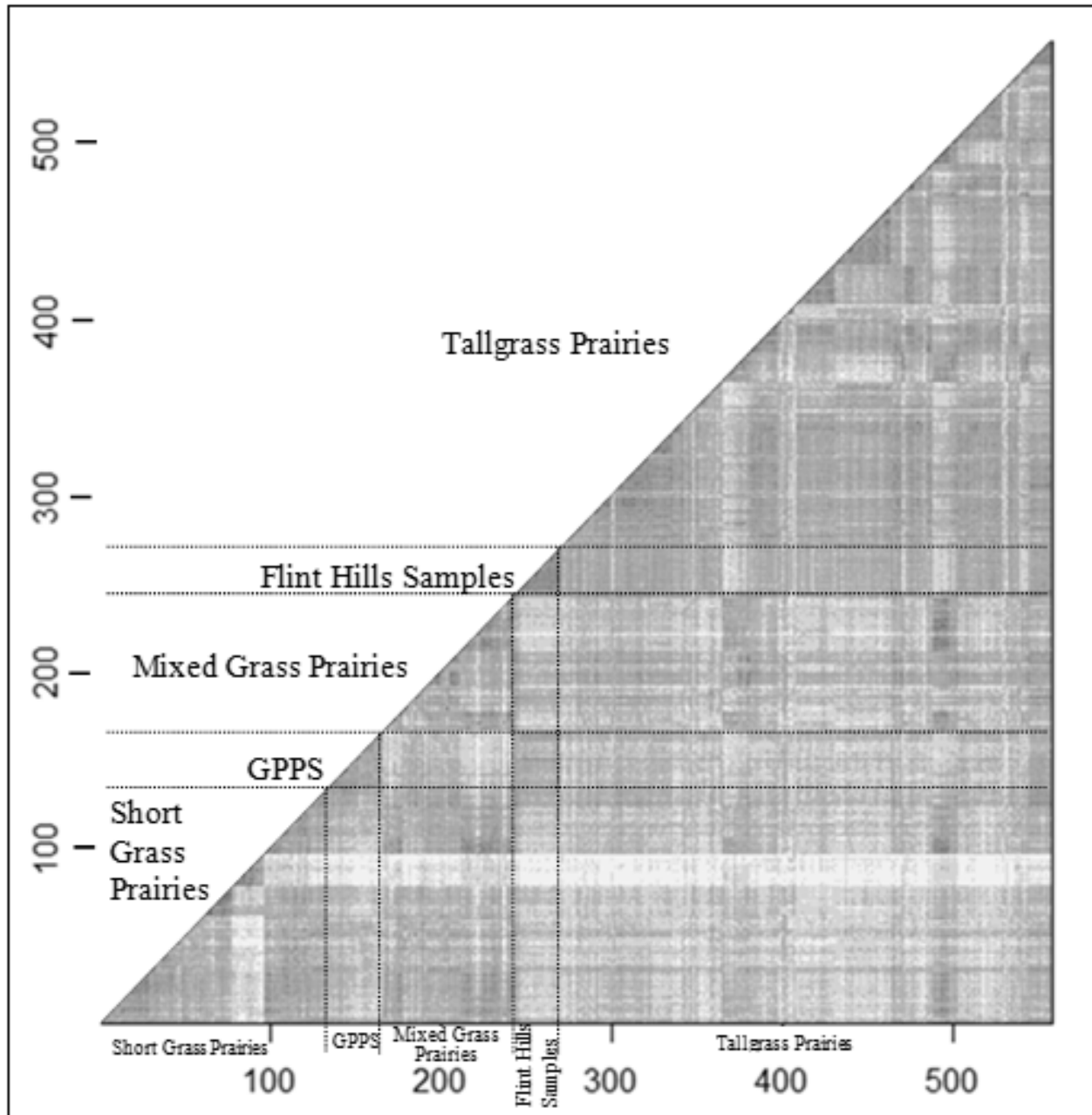


Figure 5.8: CDM dissimilarity matrix among all pollen assemblages from grasslands of North America with lower dissimilarity represented by dark grays and black and higher dissimilarity represented by light grays and white. The matrix is broken down by prairie type and the GPPS and Flint Hills sample sets. The numbers represent number of samples starting with 0 in the bottom left corner ($n = 559$).

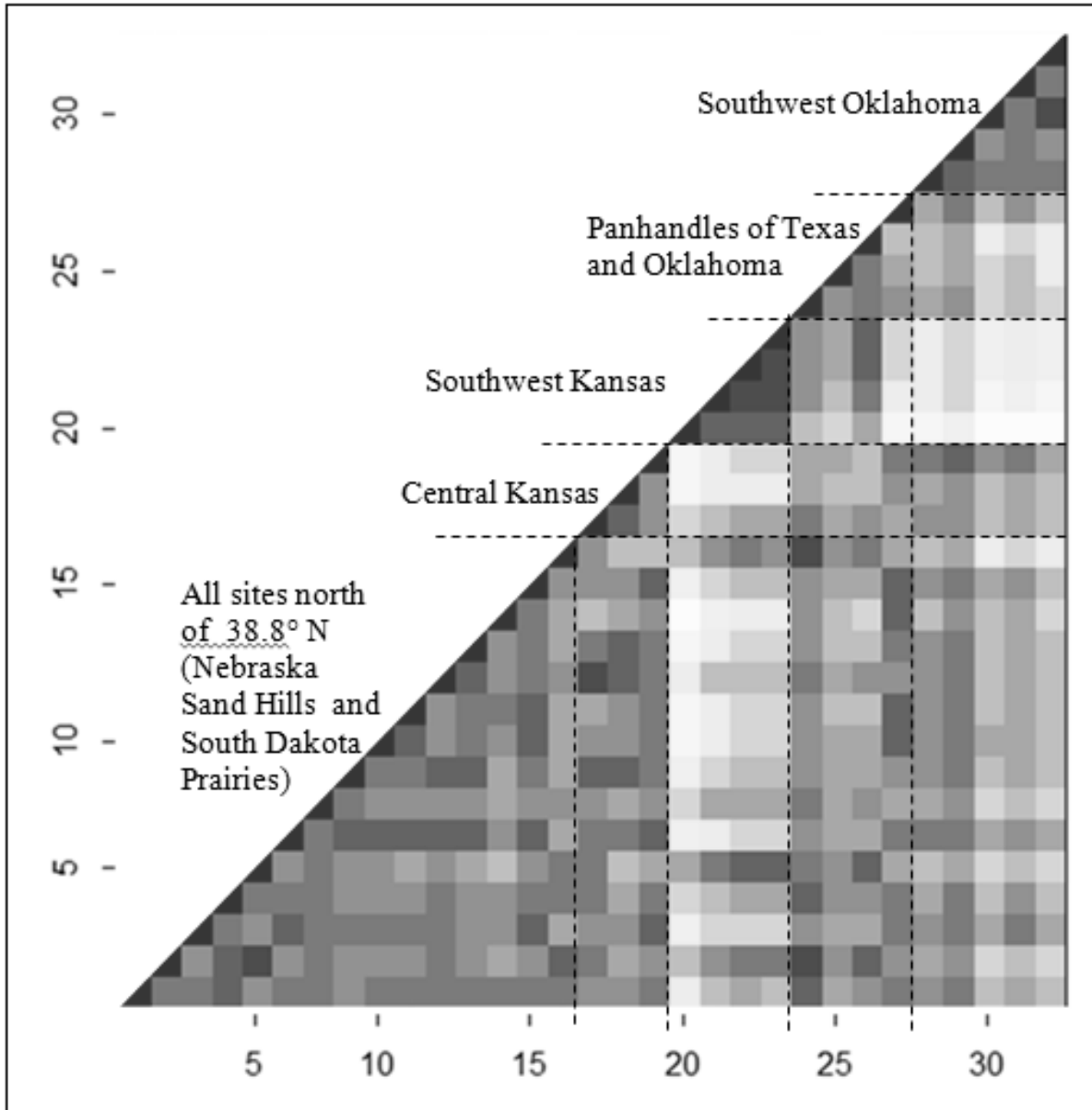


Figure 5.9: SCD dissimilarity matrix for only the GPPS set with lower dissimilarity represented by dark grays and black and higher dissimilarity represented by light grays and white. The numbers are the sample numbers from the GPPS dataset. Note sample #9 was left out of the analysis, there only are 31 samples represented in this matrix.

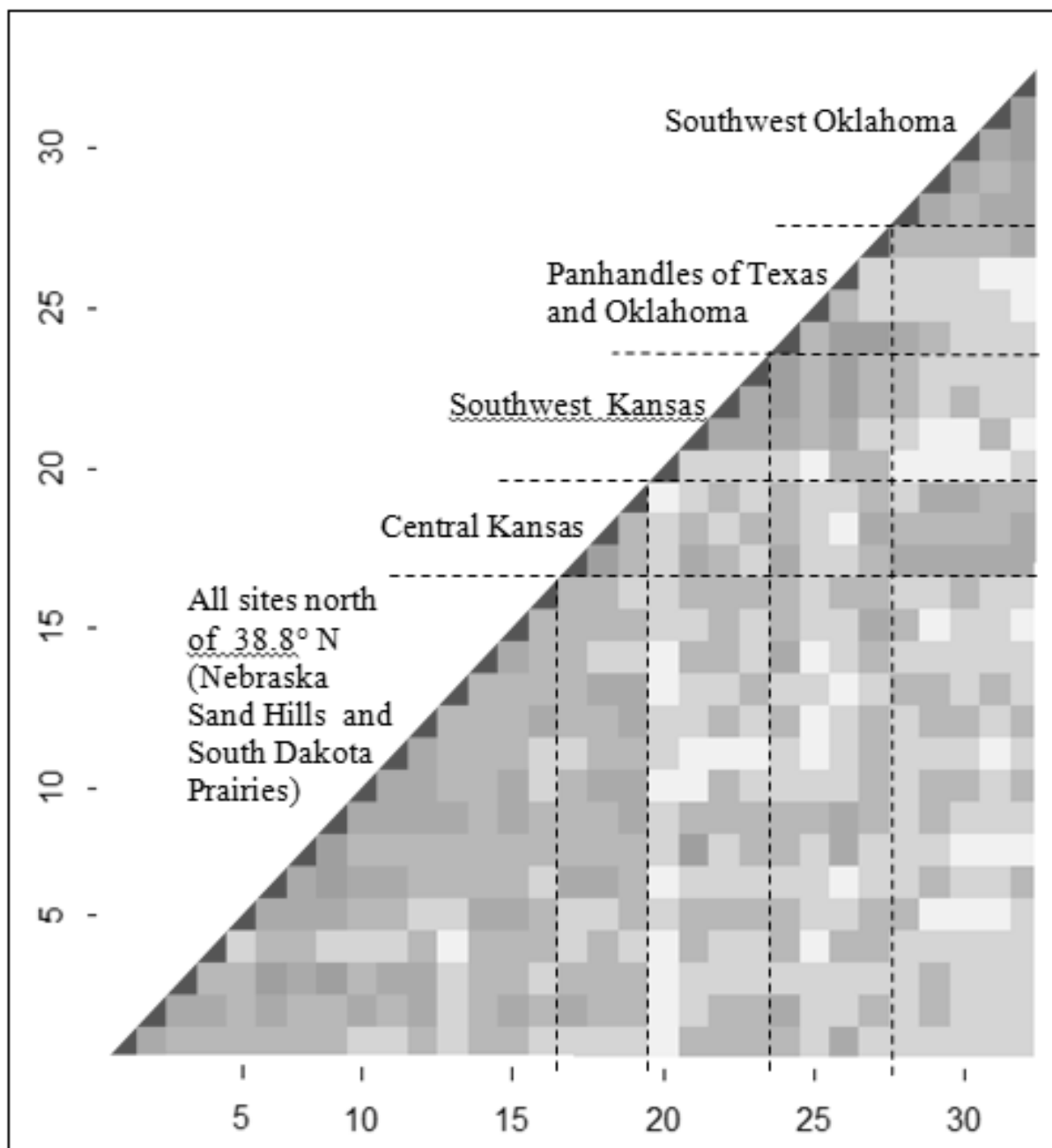


Figure 5.10: CDM dissimilarity matrix for only the GPPS set with lower dissimilarity represented by dark grays and black and higher dissimilarity represented by light grays and white. The numbers are the sample numbers from the GPPS dataset. Note sample #9 was left out of the analysis, there only are 31 samples represented in this matrix.

ROC Analysis: Determining Threshold Values

Receiver operator characteristic curve analysis was used to compare the SCD and the CDM dissimilarity metrics ability to assess analog/no analog situations among prairie types. Area under the curve (AUC) indicates the ability of the analysis to classify vegetation type correctly compared to a completely randomized classification. The optimal dissimilarity value (ODV) is a decision threshold where values above the ODV are classified as no-analogs and the values below the ODV are classified as analogs (Gavin et al. 2003; Wahl 2004). Comparing SCD to CDM for all prairie sites shows the AUC for SCD at 0.991 is slightly higher than that of CDM at 0.976 (Fig. 5.11). The AUCs and ODVs were calculated for both distance metrics for

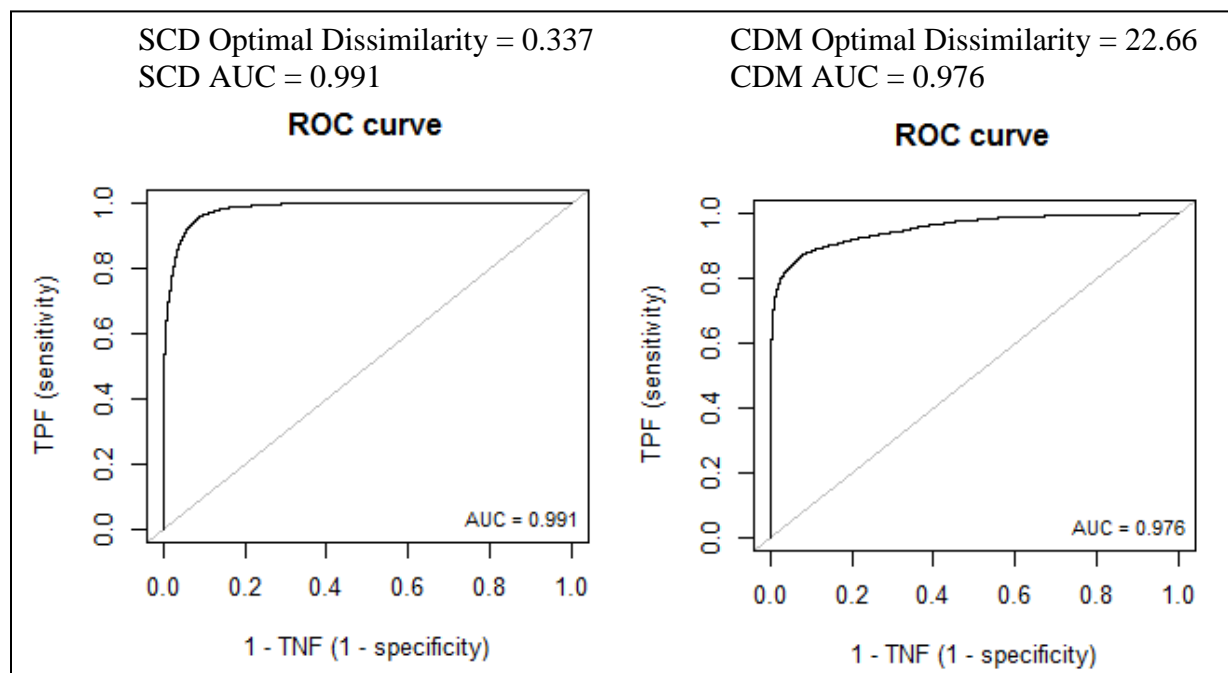


Figure 5.11: The receiver operator characteristic (ROC) curves for all prairies for both SCD analysis (left) and CDM analysis (right) are shown. Area under the curve (AUC) indicates the ability of the analysis to classify correctly compared to a completely randomized classification.

all three prairie types with the AUCs of the SCD being higher for each prairie type than the AUCs of the CDM (Fig. 5.12). The lowest AUC was calculated for the CDM analysis of mixed grass prairies (0.896) and the highest AUC was calculated for the SCD analysis of tallgrass prairies (.975). The least symmetrical ROC curve is the curve for the CDM short grass prairie analysis, suggesting that the CDM analysis is affected by a variance in the pollen proportions that is contained in the SCD metric. The same phenomenon is visible for the ROC curve of CDM analysis of mixed grass prairies as well, but to a slightly lesser extent. The SCD ODV was lowest for tallgrass prairies (0.223) and highest for mixed grass prairies (0.291). The CDM ODV was highest for mixed grass prairies (30.72) but lowest for short grass prairies (21.07). Directly comparing the ODVs of SCD and CDM is of little use because each metric produces a different range of values.

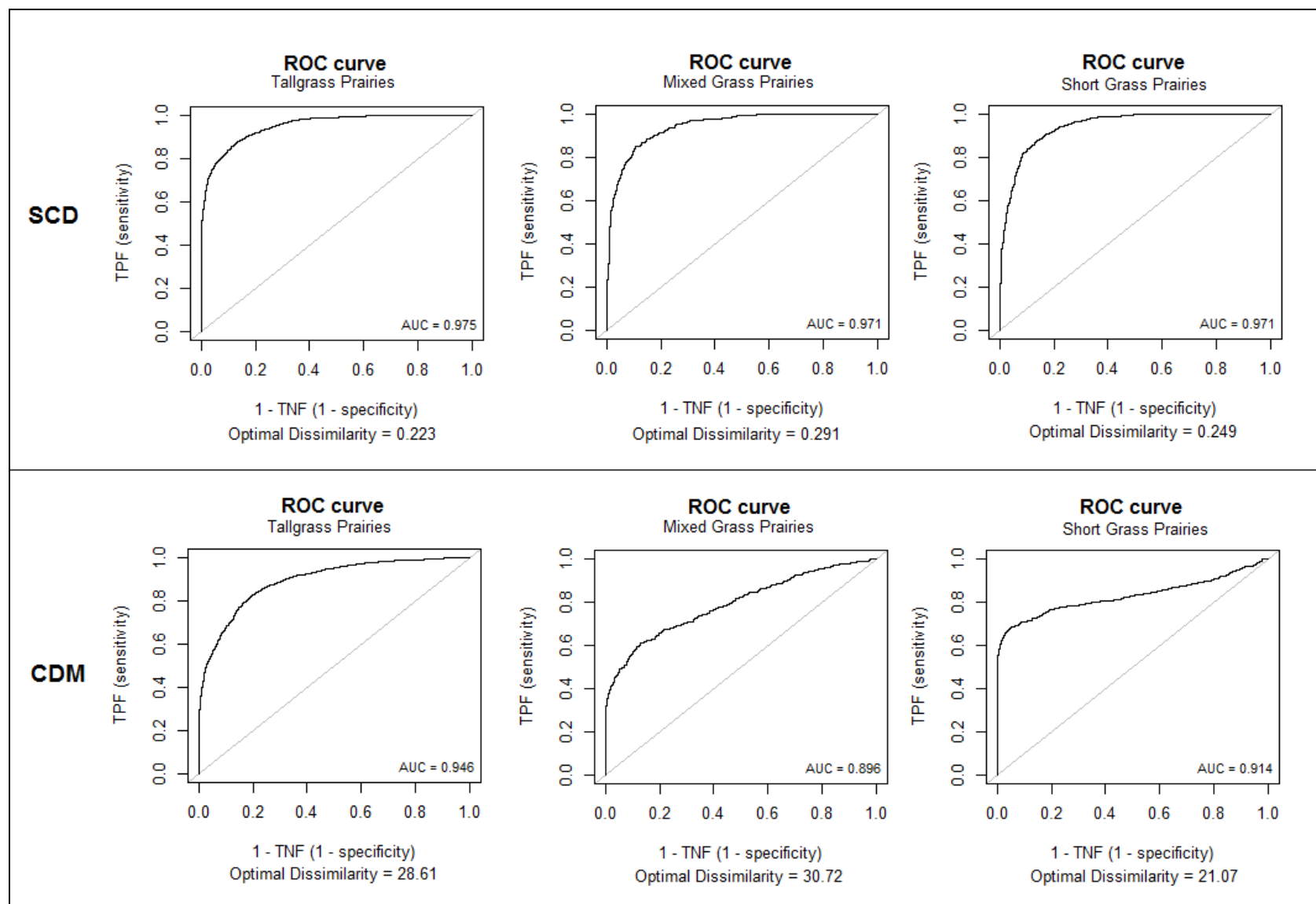


Figure 5.12: The ROC curves for the three prairie types are shown for both SCD (top) and CDM (bottom) analysis.

Chapter 6 - Discussion

Sample Site Locations and Climate Types

The samples collected in this study represent a range of climate types that are underrepresented in the NAPD. The NAPD previously had only 34 pollen assemblages out of 504 samples classified as “prairie” that met both the maT and maP parameters (Whitmore et al. 2005). All 24 Flint Hills samples (Commerford 2010) are also within the climate space bringing the total number of prairie sites up to 58 available for analysis in the specified climate space. This study adds another 31 pollen assemblages, further increasing the representation of the selected climate space to 89 available sites. The NAPD has 460 “prairie” sites that have a maT between -1°C and 9°C. The total number of prairie pollen assemblages with maTs > 9°C is only 99 sites with 44 NAPD “prairie” sites, 24 Flint Hills samples and 31 GPPS sites. The warmer prairie sites went from being less than 9% of the available prairie pollen assemblages to fewer than 18% with the addition of both the Flint Hills and the GPPS. This sampling bias towards cooler climates is likely because there are more perennial water bodies available for sampling due to both past glaciations leaving kettle ponds and lakes and the lower evaporation rates compared to the more southerly prairies (Grimm et al. 2011).

The GPPS sites represent a wide range of climatic conditions from ~9.7 °C maT in the Fort Pierre National Grassland of South Dakota to ~19 °C maT in southwestern Oklahoma and from 410 mm maP in the northwest panhandle of Texas to ~750 mm maP in southwestern Oklahoma. In addition to having a wide range of climate types, of all the pollen assemblages used in analysis only 37 had a maT > 15 °C, and 16 of those sites are from the GPPS. Many of the sample sites are spatially exclusive as well, with many 10's of kilometers between them and any previously sampled sites. The sites in southwestern Oklahoma are even more distant from

the NAPD sites with over 200 km to the nearest NAPD site (Figure 5.1). The GPPS sites not only sampled areas that were previously unsampled in the NAPD but also sampled numerous sites with unrepresented climate types as well.

***Ambrosia* and *Artemisia* Analysis**

The ability of pollen assemblages to distinguish among tallgrass, mixed grass, and short grass prairie vegetation has been an ongoing question. Determining prairie type through pollen assemblage analysis has been somewhat successful in a previous study when using low taxonomic resolution (8 taxa) and small sample set (20 sites) collected from Tauber traps in the southern Great Plains (Hoyt 2000). On a continental scale, however, most investigations have focused on differences among forest types or between grassland/forest boundaries (Williams et al. 2009; Williams & Shuman 2008). Here we find the of the proportions of two pollen types *Ambrosia* and *Artemisia* provide a simple but powerful tool for distinguishing prairie types among the 504 prairie samples in North America. Samples from tallgrass prairies indicate that *Ambrosia* pollen, with a larger distribution and higher mean proportion, is much more likely to be prevalent than *Artemisia* pollen. Hall (1994) found similar results of *Ambrosia* and *Artemisia* pollen influx in Tauber traps for prairies in Oklahoma, Texas, and New Mexico. However, the opposite is true for short grass prairies where *Artemisia* pollen has a larger distribution of proportions and a higher mean proportion than *Ambrosia* pollen. This by itself indicates, but does not confirm, that there is a high likelihood of a significant and measurable difference between the two prairie types' proportions of *Ambrosia* and *Artemisia*.

The ratio of *Ambrosia*/*Artemisia* pollen can statistically distinguish tallgrass from the other two prairie types (Table 5.3). Specifically, pollen assemblages with a *Ambrosia*/*Artemisia* ratio of greater than 20 are tallgrass prairie. The significant differences in ratio means between

tallgrass and the other prairie types indicates that once a fossil pollen assemblage is classified as prairie using MAT then the *Ambrosia/Artemisia* ratio can be analyzed and assigned a prairie type of either tallgrass or short/mixed grass, thus adding one more tool into the paleoenvironmental researcher's toolbox. This quantitative result improves previous research indicating spatial differences in *Ambrosia* pollen abundance possibly related to vegetation type or climate (Hall 1994; Williams et al. 2006; Grimm et al. 2011).

The threshold ratio value would vary from 2.5 to 74. A conservative threshold value is an *Ambrosia/Artemisia* ratio of 20 because all samples from short grass prairies and over 95% samples from mixed grass prairies used in this study are below 20 (Appendix D). A slightly less conservative threshold value would be an *Ambrosia/Artemisia* ratio of 10, which is still greater than all observed short grass ratios and still greater than 90% of observed mixed grass ratios. Researchers will have to make a decision as to how much error is acceptable for their study. It also must be recognized that any threshold value used in this method cannot provide any positive or negative classifications of prairie type if the sample ratio is below the threshold. This technique can only give a positive classification of tallgrass prairie if the sample ratio is above the threshold value.

Climatic information can also be interpreted from using the ratio of pollen assemblage proportions of *Ambrosia* to *Artemisia*. Both genera contain several species that are relatively sensitive to temperature and moisture availability (Williams et al. 2006). This sensitivity occurs at both a continental and regional scale. McLauchlan et al. (2011) found that early spring precipitation led to high *Ambrosia* pollen counts in Kansas City. In this study both maP and maT have a relationship with the *Ambrosia/Artemisia* ratio (Fig. 5.4 and Fig. 5.5), although using

the ratio as a predictor of a specific value of maP is not possible due to the large variance of ratio values for high precipitation values.

In this sample, an appropriate maP threshold value is 500 mm, because of the low variability of the *Ambrosia* / *Artemisia* ratio in all samples with less than 500 mm maP and the much larger amount of variability in the *Ambrosia*/*Artemisia* ratio above 500 mm maP. With the maP threshold set at 500 mm, a conservative ratio threshold value is 10 because only one of the 246 pollen assemblages with a maP of less than 500mm has an *Ambrosia*/*Artemisia* ratio above 10. The interpretation of this ratio threshold leads to two conclusions: 1) If the ratio of proportions of *Ambrosia* and *Artemisia* of a pollen assemblage classified as a prairie is higher than 10, then there is a very high likelihood that the maP is greater than 500mm; 2) If the ratio of proportions of *Ambrosia* and *Artemisia* of a pollen assemblage classified as a prairie is lower than 10, then nothing definitive can be said about maP.

A temperature gradient on the Great Plains also indicated temperature sensitivity of *Ambrosia* pollen (Ziska et al. 2011). The *Ambrosia*/*Artemisia* ratio's relationship to maT is similar to that of maP where the variability of ratio values increases dramatically above a climate threshold, in this case 5 °C maT. The ratio threshold for climate classification is again set at 10 because of all samples with maT of < 5 °C (n = 169) only one has a ratio greater than 10.

Since the threshold ratio of the proportions of *Ambrosia*/*Artemisia* in a pollen assemblage can be set at 10 for both maP and maT, samples above this threshold probably had maT greater than 5 °C and maP greater than 500 mm. Additionally, if the *Ambrosia*/*Artemisia* ratio is greater than 20, it is likely that pollen assemblage was created by a tallgrass prairie. The implications of the use of the *Ambrosia*/*Artemisia* ratio in paleoenvironmental reconstructions is that if after MAT analysis a fossil pollen assemblage is classified as prairie, a quick assessment of the ratio

of *Ambrosia/Artemisia* proportions can be done as well. This could be very useful for the Holocene pollen assemblages of central North America (Fredlund 1995), where the severity of the mid-Holocene warm period has been debated (McLauchlan et al 2013). One drawback to this method is that no definitive statements can be made for any ratio below the threshold value, meaning that a pollen assemblage with a low ratio can still have a maT higher than 5 °C and/or a maP higher than 500 mm.

MAT Dissimilarity Matrices

MAT analysis has been used successfully many times to interpret analogs for many different vegetation types (e.g. Jackson et al. 2000), however very little attention has been given to using MAT to delineate between prairie types. Overall, grassland pollen assemblages have been considered very similar to one another with an average SCD value greater than 0.15 (Gavin et al. 2003). The SCD is the most commonly used distance metric in North American paleoenvironment reconstruction studies, because it gives more influence to the most common pollen types (Overpeck et al. 1985). The CDM has been used to assess pollen assemblages in biomes with low variability of vegetation in order to allow the rarer pollen types greater influence on the analysis (Oswald et al. 2003). By comparing the differences of dissimilarity values between prairie types for each distance metric, insight into which metric may be more successful for interpreting analogs different prairie type analogs.

There is a visible difference between the tallgrass prairie sites and the mixed grass and short grass prairie sites as seen by a dissimilarity matrix of SCD (Fig. 5.6). However, there appear to be many mixed grass and short grass prairie pollen assemblages that are very similar to some tallgrass prairie pollen assemblages. The CDM dissimilarity matrix produces a similar result: the tall grass prairies as a whole have less dissimilarity than the other two prairie types

(Fig. 5.7). Since CDM produces a much larger range of values than SCD, fewer pollen assemblage comparisons are shown to have low dissimilarity, which may indicate that CDM is more conservative when assigning match/no-match interpretations. Low dissimilarity between different prairie type pollen assemblages is seen in both dissimilarity matrices and may be because of improper classification of prairie type in this analysis or it may be because of limitations placed on the distance metrics through the use of the 47 pollen types. The 47 pollen types used in this analysis were chosen because the pollen types are present in the GPPS and NAPD sample sets. It is important to keep in mind that the pollen types chosen for this analysis are appropriate for all prairie pollen assemblages but this pollen set may not be optimal for regional scale prairie pollen analysis (e.g. the upper Midwest).

Dissimilarity matrices were also created for only the GPPS pollen assemblages show that the sites north of 38° N (sites 1-19) have relatively low dissimilarities to each other when compared to the southern samples (Fig. 5.8). The region with the most unique pollen signature is southwestern Kansas (sites 20-23), which are also unlike almost all NAPD prairie sites as well. These sites, having warm, dry climates, could provide insight into future shifts in grassland composition as the Great Plains continue to warm over the next 50 years (IPCC 2007).

ROC Analysis: Determining Threshold Values

Using the MAT to interpret fossil pollen assemblages can provide great insight into past vegetation regimes and climate types, but unfortunately there is no set cut-off or threshold value, for determining analog/no-analog situations, which is dependent on sample set size and pollen types used in analysis (Williams & Shuman 2008). ROC analysis has been recently used to determine cut-off values for tundra types in the Arctic (Oswald et al. 2003) and montane forests (Wahl 2004). However, ROC analysis has not previously been used on grassland pollen

assemblages. ROC analysis does not only produce an optimized dissimilarity (threshold) value (ODV) but also an area under the curve (AUC) which can be used to determine the overall power of the data to distinguish between matches and non-matches (Wahl 2004). In this study the AUC is used to assess the ability of two MAT distance metrics (SCD and CDM) to delineate between different prairie type pollen assemblages.

The first comparison is between the ability of SCD and CDM ability to distinguish prairie and non-prairie vegetation types, using all the prairie pollen assemblages as a training set ($n = 559$). The ROC curve analysis found that the ODV for SCD is 0.337, so any SCD value below 0.337 for two prairie samples can be considered a match. The ODV for CDM was 22.66 and can be interpreted the same way. Since SCD and CDM produce distance metric values that are not directly comparable, another method of determining which MAT method has a more robust analysis and comparing AUC values can provide that method. The AUC for the ROC curve analysis of SCD is 0.991 compared to an AUC for CDM of 0.976, indicating that SCD is more accurate in delineating between matches and non-matches within the prairie sample set because its value is closer to 1.

ROC curve analysis was also performed for both SCD and CDM results using training sets of the three individual prairie types, which were compared to all prairie pollen assemblages (Fig. 5.11). This analysis was done to compare the ability of SCD and CDM to delineate between short grass, mixed grass, and tallgrass prairies using pollen assemblages. The ODVs for SCD analysis were all below 0.337 (the ODV for similar vegetation type for any prairie) which indicate that with the use of SCD analysis it may be possible to indicate matches/no-matches to a specific prairie type. The ODVs for CDM analysis were both above and below 22.66 showing that CDM does not delineate between prairie type pollen assemblages

as well as SCD does. Another way to compare the ability of SCD and CDM to differentiate between prairie types is to compare the AUCs of each ROC curve. And for each prairie type the AUC is higher for the ROC curves of the SCD analysis and indicating that SCD may be a more appropriate MAT method than CDM for determining a specific prairie type.

Comparing AUCs across prairie types for SCD shows that there is almost no difference in SCD's ability to determine matches/no-matches, meaning that no prairie type is any easier to classify than any other within the dataset. The shapes of the ROC curves for SCD analysis are all fairly symmetrical and again showing that SCD has no preference for a specific prairie type. The AUCs for CDM analysis of different prairie types shows that each prairie type has a slightly different influence on CDM values. The ROC curves are also influenced differently depending on prairie type with the tallgrass prairies having the most symmetrical curve and the short grass prairies having the most asymmetrical curve. It is unclear as of yet as to what influences have created such differences in analysis of CDM for different prairie types.

One major implication of these results is it will be possible for future paleoenvironmental reconstructions of the Great Plains that use pollen assemblages to gain additional information about prairie type. While other proxies such as paleosols (Miao et al. 2005), dune movement (Forman et al. 2001), and diatom (Hobbs et al. 2011) have been important for documenting the propensity of the region to recurring drought, the biotic response of plant communities has been relatively unknown. This is important for assessing past and future responses of North American grasslands to warmer and drier climate conditions than present (IPCC 2007). With numerous multi-decadal to centennial scale droughts occurring over the Holocene (Schneider et al. 2011), understanding how grassland plant communities transitioned into these long term droughts

would provide landowners (e.g. cattle ranchers and land/natural resource managers) a head start in enacting beneficial conservation methods.

References

- Allen, M. S., and M. W. Palmer. 2011. Fire history of a prairie/forest boundary: more than 250 years of frequent fire in a North American tallgrass prairie. *Journal of Vegetation Science* 22:436-444.
- Axelrod, D.I. 1985. Rise of the grassland biome, central North America. *Botanical Review* 51:163-201.
- Bement, L. C., B. J. Carter, R. A. Varney, L. S. Cummings, and J. B. Sudbury. 2007. Paleo-environmental reconstruction and bio-stratigraphy, Oklahoma Panhandle, USA. *Quaternary International* 169:39-50.
- Bailey RG (revised 1994) Ecoregions of the United States. USDA Forest Service.
- Bartlein, P. J., K. H. Anderson, P. M. Anderson, M. E. Edwards, C. J. Mock, R. S. Thompson, R. S. Webb, and C. Whitlock. 1998. Paleoclimate simulations for North America over the past 21,000 years: Features of the simulated climate and comparisons with paleoenvironmental data. *Quaternary Science Reviews* 17:549-585.
- Bonnefille, R., and F. Chalieu. 2000. Pollen-inferred precipitation time-series from equatorial mountains, Africa, the last 40 kyr BP. *Global and Planetary Change* 26:25-50.
- Borchert, J. 1971. Dust Bowl in 1970s. *Annals of the Association of American Geographers* 61:1-22.

- Briggs, J., G. Hoch, and L. Johnson. 2002. Assessing the rate, mechanisms, and consequences of the conversion of tallgrass prairie to *Juniperus virginiana* forest. *Ecosystems* 5:578-586.
- Brown, D.A. 1993. Early nineteenth-century grasslands of midcontinent plains. *Annals of the Association of American Geographers* 83:589-612.
- Brown, L. 1947. Grasslands. Chanticleer Press, Inc. NY.
- Clark, J., E. Grimm, J. Donovan, S. Fritz, D. Engstrom, and J. Almendinger. 2002. Drought cycles and landscape responses to past aridity on prairies of the northern Great Plains, USA RID G-6331-2011. *Ecology* 83:595-601.
- Changon, S.A., K.E. Kunkel, and D. Winstanley. 2002. Climate factors that caused the unique tall grass prairie in the central United States. *Physical Geography* 23: 259-280.
- Clark, J., E. Grimm, J. Donovan, S. Fritz, D. Engstrom, and J. Almendinger. 2002. Drought cycles and landscape responses to past aridity on prairies of the northern Great Plains, USA RID G-6331-2011. *Ecology* 83:595-601.
- Commerford, J. L. 2010. Calibrating vegetation cover and pollen assemblages in the Flint Hills of Kansas, U.S.A. Master's Thesis, Kansas State University
- Cordova, C. E., W. C. Johnson, R. D. Mandel, and M. W. Palmer. 2011. Late Quaternary environmental change inferred from phytoliths and other soil-related proxies: Case studies from the central and southern Great Plains, USA. *Catena* 85:87-108.
- Cowell, C., and A. Parker. 2004. Biogeography in the Annals. *Annals of the Association of American Geographers* 94:256-268.

- Dalle, G., B. Maass, and J. Isselstein. 2006. Encroachment of woody plants and its impact on pastoral livestock production in the Borana lowlands, southern Oromia, Ethiopia. *African Journal of Ecology* 44:237-246.
- Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. Pasteris. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology* 28:2031-2064.
- Davis, O. K. 1995. Climate and vegetation patterns in surface samples from arid western U.S.A.: Application to Holocene climatic reconstructions. *Palynology* 19:95-117.
- Davis, R., and T. Webb III. 1975. Contemporary Distribution of Pollen in Eastern North-America - Comparison with Vegetation. *Quaternary Research* 5:395-434.
- Dean, W. E. & A. Schwalb. 2000. Holocene environmental and climatic change in the northern Great Plains as recorded in the geochemistry of sediments in Pickerel Lake, South Dakota. *Quaternary International* 67:5-20.
- Fægri, K. and Iversen, J. 1989: Textbook of pollen analysis. John Wiley.
- Fenneman, N. M. 1919. The circumference of geography. *Annals of the Association of American Geographers* 9:3-11.
- Feurdean, A., V. Mosbrugger, B. P. Onac, V. Polyak, and D. Veres. 2007. Younger Dryas to mid-Holocene environmental history of the lowlands of NW Transylvania, Romania. *Quaternary Research* 68:364-378.

- Forman, S. L., R. Oglesby, and R. S. Webb. 2001. Temporal and spatial patterns of Holocene dune activity on the Great Plains of North America: megadroughts and climate links. *Global and Planetary Change* 29:1-29.
- Fredlund, G. 1995. Late Quaternary Pollen Record from Cheyenne Bottoms, Kansas. *Quaternary Research* 43:67-79.
- Fyfe, R. M., J. Beaulieu, H. Binney, R. H. W. Bradshaw, S. Brewer, A. L. Flao, W. Finsinger, M. Gaillard, T. Giesecke, G. Gil-Romera, E. C. Grimm, B. Huntley, P. Kunes, N. Kühl, M. Leydet, A. F. Lotter, P. E. Tarasov, and S. Tonkov. 2009. The European Pollen Database: past efforts and current activities. *Vegetation History and Archaeobotany* 18:417-424.
- Gavin, D. G., W. W. Oswald, E. R. Wahl, and J. W. Williams. 2003. A statistical approach to evaluating distance metrics and analog assignments for pollen records. *Quaternary Research* 60:356-367.
- Grimm, E. C., J. J. Donovan, and K. J. Brown. 2011. A high-resolution record of climate variability and landscape response from Kettle Lake, northern Great Plains, North America. *Quaternary Science Reviews* 30:2626-2650.
- Hall, S. 1994. Modern Pollen Influx in Tallgrass and Shortgrass Prairies, Southern Great-Plains, Usa. *Grana* 33:321-326.
- Henderson, A. 1993. Assessing Test Accuracy and its Clinical Consequences - a Primer for Receiver Operating Characteristic Curve Analysis. *Annals of Clinical Biochemistry* 30:521-539.

- Hill, M., W. Willms, and R. Aspinall. 2000. Distribution of range and cultivated grassland plants in southern Alberta. *Plant Ecology* 147:59-76.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology* 25:1965-1978.
- Hobbs, W. O., S. C. Fritz, J. R. Stone, J. J. Donovan, E. C. Grimm, and J. E. Almendinger. 2011. Environmental history of a closed-basin lake in the US Great Plains: Diatom response to variations in groundwater flow regimes over the last 8500 cal. yr BP. *Holocene* 21:1203-1216.
- Hoyt, C. 2000. Pollen signatures of the arid to humid grasslands of North America. *Journal of Biogeography* 27:687-696.
- IPCC. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Jackson, S., and J. Williams. 2004. Modern analogs in Quaternary paleoecology: Here today, gone yesterday, gone tomorrow? *Annual Review of Earth and Planetary Sciences* 32:495-537.

- Jackson, S., R. Webb, K. Anderson, J. Overpeck, T. Webb, J. Williams, and B. Hansen. 2000. Vegetation and environment in Eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews* 19:489-508.
- Jackson, S., and C. Weng. 1999. Late quaternary extinction of a tree species in eastern North America. *Proceedings of the National Academy of Sciences of the United States of America* 96:13847-13852.
- James, E. 1823. Account of an expedition from Pittsburgh to the Rocky Mountains performed in the years 1819 and '20, by order of the Hon. J.C. Calhoun, sec'y of war: under the command of Major Stephen H. Long. From the notes of Major Long, Mr. T. Say, and other gentlemen of the exploring party. H.C. Carey and I. Lea, Philadelphia.
- Küchler, A. W. 1974. A New Vegetation Map of Kansas. *Ecology* 55:586-604.
- Küchler, A. W. 1972. The oscillations of the mixed prairie in Kansas. *Erkunde* 26:120-129.
- Lane, D., D. Coffin, and W. Lauenroth. 2000. Changes in grassland canopy structure across a precipitation gradient. *Journal of Vegetation Science* 11:359-368.
- Luly, J. G., J. F. Grindrod, and D. Penny. 2006. Holocene palaeoenvironments and change at Three-Quarter Mile Lake, Silver Plains Station, Cape York Peninsula, Australia. *Holocene* 16:1085-1094.

- Luo, C., Z. Zheng, P. Tarasov, T. Nakagawa, A. Pan, Q. Xu, H. Lu, and K. Huang. 2010. A potential of pollen-based climate reconstruction using a modern pollen-climate dataset from arid northern and western China. *Review of palaeobotany and palynology* 160:111-125.
- Lytle, D., and E. Wahl. 2005. Palaeoenvironmental reconstructions using the modern analogue technique: the effects of sample size and decision rules. *Holocene* 15:554-566.
- Mazier, F., A. Brostrom, M. Gaillard, S. Sugita, P. Vittoz, and A. Buttler. 2008. Pollen productivity estimates and relevant source area of pollen for selected plant taxa in a pasture woodland landscape of the Jura Mountains (Switzerland). *Vegetation History and Archaeobotany* 17:479-495.
- McLauchlan, K. K., J. L. Commerford, and C. J. Morris. 2013. Tallgrass prairie pollen assemblages in mid-continental North America. *Vegetation History and Archaeobotany* DOI: 10.1007/s00334-012-0369-8
- McLauchlan, K., C. S. Barnes, and J. M. Crane. 2011. Interannual variability of pollen productivity and transport in mid-North America from 1997 to 2009. *Aerobiologia* 27:181-189.
- McLauchlan, K. K., A. J. Elmore, W. W. Oswald, and S. Sugita. 2007. Detecting open vegetation in a forested landscape: pollen and remote sensing data from New England, USA. *Holocene* 17:1233-1243.
- Metz, C. 1978. Basic Principles of Roc Analysis. *Seminars in nuclear medicine* 8:283-298.

- Miao, X., J. A. Mason, R. J. Goble, and P. R. Hanson. 2005. Loess record of dry climate and aeolian activity in the early- to mid-Holocene, central Great Plains, North America. *The Holocene* 15:339-346.
- Minckley, T. A., P. J. Bartlein, C. Whitlock, B. N. Shuman, J. W. Williams, and O. K. Davis. 2008. Associations among modern pollen, vegetation, and climate in western North America. *Quaternary Science Reviews* 27:1962-1991.
- New, M., D. Lister, M. Hulme, and I. Makin. 2002. A high-resolution data set of surface climate over global land areas. *Climate Research* 21:1-25.
- Ohlwein, C., and E. R. Wahl. 2012. Review of probabilistic pollen-climate transfer methods. *Quaternary Science Reviews* 31:17-29.
- Oswald, W., L. Brubaker, F. Hu, and D. Gavin. 2003. Pollen-vegetation calibration for tundra communities in the Arctic Foothills, northern Alaska. *Journal of Ecology* 91:1022-1033.
- Overpeck, J. T., R. S. Webb, and T. Webb. 1992. Mapping Eastern North-American Vegetation Change of the Past 18 Ka - No-Analogs and the Future. *Geology* 20:1071-1074.
- Overpeck, J. T., T. Webb, and I. C. Prentice. 1985. Quantitative Interpretation of Fossil Pollen Spectra - Dissimilarity Coefficients and the Method of Modern Analogs. *Quaternary Research* 23:87-108.
- Owsley, M. 2011. Plant fact sheet for Big Bluestem (*Andropogon gerardii* Vitman). USDA-Natural Resources Conservation Service, Jimmy Carter PMC Americus, GA.

- Pattison, W. 1964. The 4 Traditions of Geography. *Journal of Geography* 63:211-216.
- Prentice, I. 1985. Pollen Representation, Source Area, and Basin Size - Toward a Unified Theory of Pollen Analysis. *Quaternary Research* 23:76-86.
- Prentice, I.C. 1980. Multidimensional scaling as a research tool in Quaternary Palynology: A review of theory and methods. *Review of Palaeobotany and Palynology* 31:71-104.
- Risser, P.G., Birney, E.C., Blocker, H.D., May, S.W., Parton, S.W., and Wiens, J.A. 1981. *The True Prairie Ecosystem*. Hutchinson Ross Publishing Company, Stroudsburg, PA.
- Samson, F., F. Knopf, and W. Ostlie. 2004. Great Plains ecosystems: past, present, and future. *Wildlife Society Bulletin* 32:6-15.
- Sankaran, M., N. Hanan, R. Scholes, J. Ratnam, D. Augustine, B. Cade, J. Gignoux, S. Higgins, X. Le Roux, F. Ludwig, J. Ardo, F. Banyikwa, A. Bronn, G. Bucini, K. Caylor, M. Coughenour, A. Diouf, W. Ekaya, C. Feral, E. February, P. Frost, P. Hiernaux, H. Hrabar, K. Metzger, H. Prins, S. Ringrose, W. Sea, J. Tews, J. Worden, and N. Zambatis. 2005. Determinants of woody cover in African savannas. *Nature* 438:846-849.
- Sawada, M., A. Viau, G. Vettoretti, W. Peltier, and K. Gajewski. 2004. Comparison of North-American pollen-based temperature and global lake-status with CCCma AGCM2 output at 6 ka. *Quaternary Science Reviews* 23:225-244.

- Schmieder, J., S. C. Fritz, J. B. Swinehart, A. L. C. Shinneman, A. P. Wolfe, G. Miller, N. Daniels, K.C. Jacobs, and E.C. Grimm. 2011. A regional-scale climate reconstruction of the last 4000 years from lakes in the Nebraska Sand Hills, USA. *Quaternary Science Review* 30:1797-1812.
- Seppa, H., H. Birks, A. Odland, A. Poska, and S. Veski. 2004. A modern pollen-climate calibration set from northern Europe: developing and testing a tool for palaeoclimatological reconstructions. *Journal of Biogeography* 31:251-267.
- Shuman, B. N., P. Newby, and J. P. Donnelly. 2009. Abrupt climate change as an important agent of ecological change in the Northeast US throughout the past 15,000 years. *Quaternary Science Reviews* 28:1693-1709.
- Sugita, S. 1994. Pollen Representation of Vegetation in Quaternary Sediments - Theory and Method in Patchy Vegetation. *Journal of Ecology* 82:881-897.
- Thompson, R. S., Anderson, K. H., and Bartlein, P. J. 1999. Atlas of relations between climatic parameters and distributions of important trees and shrubs in North America. Volume A: Conifers. Volume B: Hardwoods. US Geological Survey Professional Paper 1650A & B, Denver, CO, 269pp., 423pp.
- Tonello, M. S., and A. R. Prieto. 2008. Modern vegetation-pollen-climate relationships for the Pampa grasslands of Argentina. *Journal of Biogeography* 35:926-938.
- Tuason, J. 1987. Reconciling the Unity and Diversity of Geography. *Journal of Geography* 86:190-193.

- Viau, A. E., M. Ladd, and K. Gajewski. 2012. The climate of North America during the past 2000 years reconstructed from pollen data. *Global and Planetary Change* 84-85:75-83.
- Viau, A. E., K. Gajewski, M. C. Sawada, and P. Fines. 2006. Millennial-scale temperature variations in North America during the Holocene. *Journal of Geophysical Research-Atmospheres* 111:D09102.
- von Post, L. (1916). Forest tree pollen in south Swedish peat bog deposits. *Pollen Et Spores*, 9, 375-401.
- Wahl, E. R. 2004. A general framework for determining cutoff values to pollen analogs with dissimilarity metrics in the modern analog technique. *Review of Palaeobotany & Palynology* 128:263-280.
- Wahl, E. R. 2003. Assigning climate values to modern pollen surface sample sites and validating modern analog climate reconstructions in the southern California region. *Madrono* 50:271-285.
- Webb, T. 1987. The Appearance and Disappearance of Major Vegetational Assemblages - Long-Term Vegetational Dynamics in Eastern North-America. *Vegetatio* 69:177-187.
- Webb, T. 1974. Corresponding Patterns of Pollen and Vegetation in Lower Michigan - Comparison of Quantitative Data. *Ecology* 55:17-28.
- Webb III, T., and R. A. Bryson. 1972. Late Glacial and Postglacial Climatic Change in the Northern Midwest Usa Quantitative Estimates Derived from Fossil Pollen Spectra by Multi Variate Statistical Analysis. *Quaternary Research* (Orlando) 2:70-115.

- Whitlock, C. 1993. Postglacial Vegetation and Climate of Grand-Teton and Southern-Yellowstone-National-Parks. *Ecological Monographs* 63:173-198.
- Whitmore, J., K. Gajewski, M. Sawada, J. Williams, B. Shuman, P. Bartlein, T. Minckley, A. Viau, T. Webb, S. Shafer, P. Anderson, and L. Brubaker. 2005. Modern pollen data from North American and Greenland for multi-scale paleoenvironmental applications RID A-3912-2008 RID E-4643-2011. *Quaternary Science Reviews* 24:1828-1848.
- Williams, J. W., B. Shuman, and P. J. Bartlein. 2009. Rapid response of the prairie-forest ecotone to early Holocene aridity in mid-continental North America. *Global and Planetary Change* 66:195-207.
- Williams, J. W., and B. Shuman. 2008. Obtaining accurate and precise environmental reconstructions from the modern analog technique and North American surface pollen dataset. *Quaternary Science Reviews* 27:669-687.
- Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment* 5:475-482.
- Williams, J. W., B. Shuman, P. J. Bartlein, J. Whitmore, K. Gajewski, M. Sawada, T. Minckley, S. Shafer, A. E. Viau, T. Webb III, P. Anderson, L. Brubaker, C. Whitlock, and O. K. Davis. 2006. An atlas of pollen-vegetation-climate relationships for the United States and Canada. *American Association of Stratigraphic Palynologists Contributions Series Number 43*.

Woillard, G. 1978. Grande Pile Peat Bog - Continuous Pollen Record for Last 140,000 Years.

Quaternary Research 9:1-21.

Yansa, C. H., W. E. Dean, and E. C. Murphy. 2007. Late Quaternary paleoenvironments of an ephemeral wetland in North Dakota, USA: relative interactions of ground-water hydrology and climate change. *Journal of Paleolimnology* 38:441-457.

Zernitskaya, V., and N. Mikhailov. 2009. Evidence of early farming in the Holocene pollen spectra of Belarus. *Quaternary International* 203:91-104.

Zhao, Y., Q. Xu, X. Huang, X. Guo, and S. Tao. 2009. Differences of modern pollen assemblages from lake sediments and surface soils in arid and semi-arid China and their significance for pollen-based quantitative climate reconstruction. *Review of palaeobotany and palynology* 156:519-524.

Ziska, L., K. Knowlton, C. Rogers, D. Dalan, N. Tierney, M. A. Elder, W. Filley, J. Shropshire, L. B. Ford, C. Hedberg, P. Fleetwood, K. T. Hovanky, T. Kavanaugh, G. Fulford, R. F. Vrtis, J. A. Patz, J. Portnoy, F. Coates, L. Bielory, and D. Frenz. 2011. Recent warming by latitude associated with increased length of ragweed pollen season in central North America. *Proceedings of the National Academy of Sciences of the United States of America* 108:4248-4251.

Appendix A - Plant Types

Table A.1: List of plant types surveyed in the field.

Taxa		Sample Site (GPPS #) – Presence (X) or Absence ()																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Poaceae	<i>Agropyron smithii</i>		X			X	X	X		X	X	X	X			X	X			X	X	X	X	X	X	X
	<i>Andropogon gerardii</i>																	X	X							
	<i>Bouteloua curtipendula</i>		X	X	X	X				X				X		X	X			X	X				X	
	<i>Bouteloua dactyloides</i>		X	X			X	X	X	X	X				X			X			X	X		X	X	
	<i>Bouteloua gracilis</i>						X				X		X	X			X	X	X							
	<i>Koeleria macrantha</i>														X			X	X							
	<i>Panicum virgatum</i>						X					X		X	X	X	X		X							
	<i>Schizachyrium scoparium</i>			X	X		X	X	X						X	X	X	X								
	<i>Sorghastrum nutans</i>				X		X	X	X				X		X		X	X	X	X						
	<i>Sporobolus cryptandrus</i>		X	X	X			X	X		X	X	X	X				X	X	X		X	X	X		
	<i>Stipa comata</i>		X		X	X	X	X		X	X	X		X	X		X	X	X	X		X			X	X
Herbaceous	<i>Amaranthaceae</i>				X		X		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	<i>Ambrosia</i>		X	X			X					X	X	X	X	X	X	X	X	X	X	X	X	X		X
	<i>Apiaceae</i>				X		X				X		X	X	X		X	X		X						
	<i>Artemisia</i>		X	X	X	X	X	X			X	X				X	X	X		X	X	X	X	X	X	X
	<i>Asclepiadaceae</i>		X	X			X	X							X		X	X		X	X					
	<i>Asteraceae</i>		X	X	X	X	X	X	X	X	X				X	X	X	X	X	X	X	X	X	X	X	X
	<i>Chenopodiaceae</i>						X								X	X	X		X	X	X	X	X	X	X	X
	<i>Cornus</i>								X							X		X								
	<i>Fabaceae</i>		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Taxa		Sample Site (GPPS #) – Presence (X) or Absence ()																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Non-herbaceous	<i>Betula</i>																									
	<i>Celtis</i>																									
	<i>Juniperus</i>		X	X									X			X		X		X	X					
	<i>Maclura</i>																			X						
	<i>Morus</i>																									
	<i>Plantanus</i>																									
	<i>Populus</i>		X				X						X					X		X	X	X				
	<i>Quercus</i>												X			X		X								
	<i>Salix</i>		X	X			X		X				X					X		X						
	<i>Ulmus</i>																	X								

Taxa		Sample Site (GPPS #) – Presence (X) or Absence ()						
		26	27	28	29	30	31	32
Poaceae	<i>Agropyron smithii</i>						X	
	<i>Andropogon gerardii</i>							
	<i>Bouteloua curtipendula</i>			X	X			
	<i>Bouteloua dactyloides</i>	X	X	X	X	X		X
	<i>Bouteloua gracilis</i>		X					
	<i>Koeleria macrantha</i>							
	<i>Panicum virgatum</i>			X			X	X
	<i>Schizachyrium scoparium</i>			X	X		X	
	<i>Sorghastrum nutans</i>							
	<i>Sporobolus cryptandrus</i>			X	X	X	X	X
	<i>Stipa comata</i>	X	X			X	X	
Herbaceous	<i>Amaranthaceae</i>	X		X	X		X	
	<i>Ambrosia</i>	X	X	X	X	X	X	X
	<i>Apiaceae</i>			X	X	X	X	X
	<i>Artemisia</i>	X	X	X	X	X	X	X
	<i>Asclepiadaceae</i>		X	X	X	X	X	
	<i>Asteraceae</i>	X	X	X	X	X	X	X
	<i>Chenopodiaceae</i>			X	X			
	<i>Cornus</i>			X	X	X		
	<i>Fabaceae</i>	X	X	X	X	X	X	X
	<i>Betula</i>							
Non-herbaceous	<i>Celtis</i>							
	<i>Juniperus</i>			X	X	X		X
	<i>Maclura</i>							
	<i>Morus</i>							
	<i>Plantanus</i>							
	<i>Populus</i>			X	X			X
	<i>Quercus</i>			X	X	X	X	X
	<i>Salix</i>			X	X			
	<i>Ulmus</i>				X			

Appendix B - Site Location and Metadata

Table B.1: GPPS site location, elevation, prairie type, land ownership, and the date the sample was collected.

Site Number	Site Name	Latitude	Longitude	Elevation (m)	Prairie Type	Date Collected	Ownership
GPPS1	Smokey Valley Ranch Pond 1	38.888024	-100.965496	880.26	Short	10/22/2011	Non-Profit
GPPS2	Ray Ranch Pond	41.035259	-102.445187	1144.22	Short	5/14/2012	Private
GPPS3	Snake Pond	41.470058	-102.182389	1069.54	Mixed	5/15/2012	Private
GPPS4	West Brenmann Pond	41.496791	-102.135089	1135.08	Mixed	5/15/2012	Private
GPPS5	Maybe Pond	40.38854	-102.337193	1153.06	Short	5/20/2012	Private
GPPS6	Island Lake CLNWR	41.734132	-102.396995	1154.58	Mixed	5/22/2012	Public
GPPS7	Salt Pond CLNWR	41.75427	-102.508238	1165.86	Mixed	5/22/2012	Public
GPPS8	Martin Lake CLNWR	41.812923	-102.524682	1172.87	Mixed	5/22/2012	Public
GPPS9	View Pond BGNG	43.747754	-102.339599	800.10	Short	6/12/2012	Public
GPPS10	Cowboy Pond BGNG	43.894367	-101.713142	761.70	Short	6/13/2012	Public
GPPS11	Shell Shock Pond BGNG	43.904695	-102.271471	893.67	Short	6/13/2012	Public
GPPS12	Smith Dam FPNG	44.262904	-100.285177	539.80	Mixed	6/13/2012	Public
GPPS13	South Lake Flat Dam FPNG	44.075655	-100.348327	628.19	Mixed	6/13/2012	Public
GPPS14	Aermo Pond SRMNF	42.730425	-101.194532	930.25	Mixed	6/14/2012	Ownership
GPPS15	Oasis Pond SRMNF	42.652625	-100.883994	891.24	Mixed	6/14/2012	Non-Profit
GPPS16	Cattle Tank 4 NNF	41.94247	-100.472814	880.26	Mixed	6/14/2012	Private
GPPS17	Horseshoe Lake QNWR	38.109674	-98.476381	542.85	Mixed	6/21/2012	Private
GPPS18	Sunflower Pond QWNR	38.200898	-98.490061	530.96	Mixed	6/21/2012	Private
GPPS19	Hodgeman Co. Fishing Lake	38.048497	-99.823279	690.68	Mixed	6/21/2012	Private
GPPS20	St. Jacob's Tank	37.241381	-99.980865	704.09	Short	6/21/2012	Public
GPPS21	Gordo Pond CNG	37.160631	-101.726851	1001.57	Short	6/22/2012	Public
GPPS22	Sticker Melon Pond CNG	37.087163	-102.003202	1068.63	Short	6/22/2012	Public

GPPS23	College Pond CNG	37.090325	-101.913257	1054.61	Short	6/22/2012	Public
GPPS24	Mayor Ranch Pond	36.712605	-101.553916	938.48	Short	6/23/2012	Public
GPPS25	Coyote Pond RBNG	36.433014	-102.473828	1218.59	Short	6/23/2012	Public
GPPS26	Ibis Pond RBNG	36.389057	-102.668878	1306.68	Short	6/23/2012	Public
GPPS27	Mesquite Pond	35.704668	-102.435528	1106.73	Short	6/23/2012	Public
GPPS28	Half Boiled Pond BKNG	35.621848	-99.837897	655.62	Mixed	6/25/2012	Ownership
GPPS29	Dead Indian Pond BKNG	35.746173	-99.720336	636.42	Mixed	6/25/2012	Non-Profit
GPPS30	Caddo Lake WMWR	34.738064	-98.727227	523.34	Mixed	6/26/2012	Private
GPPS31	Bottle Pond WMWR	34.72628	-98.585268	498.65	Mixed	6/26/2012	Private
GPPS32	Crater Lake WMWR	34.704947	-98.625405	446.23	Mixed	6/26/2012	Private

GPPS Site Climate Data

Table B.2: Climate data for each GPPS site.

Site Number	Site Name	Mean Annual Precipitation (mm)	Mean Annual Temperature (°C)
GPPS1	Smokey Valley Ranch Pond 1	480	13.96
GPPS2	Ray Ranch Pond	448	11.43
GPPS3	Snake Pond	432	10.39
GPPS4	West Brenmann Pond	438	10.17
GPPS5	Maybe Pond	439	11.83
GPPS6	Island Lake CLNWR	424	9.72
GPPS7	Salt Pond CLNWR	423	9.80
GPPS8	Martin Lake CLNWR	425	9.77
GPPS9	View Pond BGNG	407	10.56
GPPS10	Cowboy Pond BGNG	426	10.45
GPPS11	Shell Shock Pond BGNG	431	9.97
GPPS12	Smith Dam FPNG	454	9.78
GPPS13	South Lake Flat Dam FPNG	468	9.68
GPPS14	Aermo Pond SRMNF	494	10.14
GPPS15	Oasis Pond SRMNF	502	10.41
GPPS16	Cattle Tank 4 NNF	545	10.14
GPPS17	Horseshoe Lake QNWR	673	15.92
GPPS18	Sunflower Pond QWNR	669	15.89
GPPS19	Hodgeman Co. Fishing Lake	560	15.13
GPPS20	St. Jacob's Tank	561	15.93
GPPS21	Gordo Pond CNG	429	15.37
GPPS22	Sticker Melon Pond CNG	419	15.10
GPPS23	College Pond CNG	424	15.32
GPPS24	Mayor Ranch Pond	439	16.23
GPPS25	Coyote Pond RBNG	422	15.19
GPPS26	Ibis Pond RBNG	410	15.03
GPPS27	Mesquite Pond	435	16.40
GPPS28	Half Boiled Pond BKNG	600	17.53
GPPS29	Dead Indian Pond BKNG	602	17.53
GPPS30	Caddo Lake WMWR	735	18.47
GPPS31	Bottle Pond WMWR	758	18.78
GPPS32	Crater Lake WMWR	737	19.13

Appendix C - Raw Data

Table C.1: The number of pollen grains, spore, and charcoal counted for each site. GPPS 9 had too few pollen grains so was not counted. Pollen analyst: Andrea Nurse, University of Maine

GPPS #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Charcoal>500um	226	53	48	44	15	122	93	125		67	88	33	56	54	104	29	66	125	18	17	21	43
Sporormiella 20x11														2								
Spor18x15											1											
Spor17x12				1																	1	
Spor 17x17												1										
Spor15x7	43	3			1												1					
Marker Bead	361	197	180	59	147	159	234	215	many	188	161	214	255	83	108	25	190	104	49	474	70	58
<i>Pinus</i> undiff	14	23	60	13	26	36	77	66		29	42	11	14	30	27	10	2	9	4	7	7	9
Haplox <i>Pinus</i>	2	3	5	3		9	5	8		11	6	6		5	1		1	1			1	2
Diplox <i>Pinus</i>	3	12	13	2		9	14	12		17	18	14	5	17	13	3	1	6		4	1	2
<i>Picea</i> undiff	1	1	3		2	3	5	1			3			1								1
<i>Abies</i>			1		1	1		1														
<i>Larix</i>	6		2	4		2					1							1				
Cupressaceae	6	13	19	4	1	18	3	13		49	15	24	26	15	9	3	8	28	2	14	4	3
<i>Populus</i>	10	3	2		7	4		3		8	11	2	8	2	18	2	3	7		2	1	1
<i>Betula</i>						2		2		5	1						1					
<i>Frax</i> amer/pennC4		1				1				3		4			1		1	2		1		4
<i>Quercus</i>	2	1	6	1	3	8	4	8		7	8	9	8	4	1	4	10	30	4	2	6	1
<i>Ulmus</i>		2	2		1	3				2	2	3	4	1	1	2	14	18		1		
<i>Acer</i>				1														1				
<i>Acer negundo</i>											9											
<i>Platanus</i>			4													1						
<i>Juglans</i>													1				4	2		1	1	
<i>Carya</i>													1				4			1		1
<i>Liquidambar</i>											1											
<i>Nyssa</i>			1																			1
<i>Celtis</i>						1												1				
<i>Cornus</i>					1						3	1										
<i>Maclura</i>	1			1	1	1				1	3			1								
<i>Morus</i>										1		2					13					1
<i>Salix</i>		5	2			2	1			2	1						4	2	2			

<i>Alnus</i>		1									2				1						
<i>Rhamnus</i>		1	4		1	1				1	4					1					
<i>Sarcobatus</i>		1	1		2	2	1				1							1			
<i>Sambucus</i>				6	1	1	1					1	1		1	1	2			1	
Caprifoliaceae					3	7	5	5		1	3	3	2	4	1	2		2	1	1	2
Rosaceae		1	4	1				1		1	3	1	1	6	1	2	6	2		1	
<i>Ephedra</i>							2						2	1							
<i>Vitis</i>	1		1																		
<i>Shepherdia</i>			1									1									
<i>Cephalanthus</i>				2													1				
<i>Arceuthobium</i>				8			2				1										
<i>Typha</i> (single)			8			40	7	87		4		75	54				91	205		1	1
<i>Sagittaria</i>															2						
<i>Sparganium</i>		2		1	1						1		3	1	1	1	2				
Cyperaceae	3	5	21	6	35	53	10	52		95	79	6	20	43	89	4	4	47	1	8	1
<i>Lycopus</i>														4							
<i>Aracea</i> -like			1								2				1				1	3	
<i>Iva</i>	6	5	4	4	11	6	11	7		2	1	1	1	8	18	7	2	3		1	12
<i>Crassula</i>				1	1									1	7		3		1		1
<i>Ambrosia</i>	14	33	71	84	35	50	14	21		24	13	25	9	92	72	33	40	39	12	24	9
<i>Xanthium</i>													2			1		11	1	4	1
<i>Rumex</i>		1	1	1						7	4	2		2			2	1	1	6	
Chenopodiaceae/Amaranthaceae	51	200	52	104	202	48	100	87		61	43	87	30	65	84	200	120	75	400	9	301
<i>Plantago</i>	6	22	6	4	10	2	6	6		5	4	3	1		6	3	7	8	4	4	11
<i>Artemisia</i>	3	36	12	11	31	23	4	24		12	31	15	9	6	4	1	3	9	0	6	9
Asteraceae low spine	12	8	3	5	20	6	4	4		71	2	5	1	3	5	5	11	16	5	5	11
Asteraceae high spine	4	6	3	15	3	6	3	3		5	1	4	1	5	4	15	4	6	14	2	9
<i>Taraxacum</i>		1		1			1					2		1							1
<i>Thalictrum</i>																		1			
Fabaceae		3	1		1	3	1	2			1		2	3	3		1	1	1		1
Ranunculaceae						1		1		6	1	1				1	1			1	2
Linaceae	9												1	1							1
Caryophyllaceae	3	3	6	3	6	5	2	9		3	3	1		1	2		1	1	1	1	15
Papaveraceae		1								1					2						
Polygonaceae					2						2	1		10	3	4			1		
<i>Oxyria</i>										1					1						1
<i>Polygonum lapathifolium</i>	1	10	3	6		2								2			1				1
<i>Polyg. bistortoides</i>				5																	
Saxifragaceae	2	2	1	1	2	2	1					1			2	1	9	4			1
Liliaceae	1	1			4	1															
Brassicaceae			3		2			1		1	1		2		1			1	1		
Orchidaceae		1													2						
Scrophulariaceae		2				5															
Malvaceae																					
Clusiaceae																					

<i>Viola</i>	1	1	4							1	3	2										
<i>Primula</i>				1		1																
<i>Urtica</i>						1				1			1				3	4				
<i>Euphorbia</i>										1				1								
<i>Humulus</i>											1											
<i>Geranium</i>											1	1							1			
<i>Pedicularis</i>																	3					
<i>Valerianella</i>																			4			
Poa <20um	2	5	1	11	4	5	8	7		2	5	1	3	59	8	9	1	31		1	1	5
Poa 20-25	19	7	11	23	15	18	13	14		24	21	16	20	111	38	37	9	39	5	4	9	13
Phrag pore<2.5			1	1	1	3				2		1		3		1					1	1
Festuca pore=2.5		5		6	1	4	1	3		5	6	2	3	8	1	3	2	4	1			2
Poa 26-30	24	23	7	16	8	9	4	7		16	29	19	21	40	17	25	5	9	3	2	13	9
Poa 31-35	13	7	3	2		3	1	1		5	10	16	9	2	4	17	10	6	3	2	4	7
Poa 36-40	1		3		1	1	2				3		4	9	2	14	2	1		1	1	1
Poa 45-60		1								1			1	7	1	9	3	1		4		
<i>Andropogon</i>	4	33	1	5		1		3		3	11	1	1	9		13		3				
<i>Zea</i>		1			1					1												
<i>Lemna</i>	4	3	1	1	1	16		9		5	10		3	2		2		6	2			1
<i>Isoetes</i>							1			1					3							
Hydrocharitaceae	1					1					4	30	6			1		1	1		1	5
<i>Potamogetum</i>					1	2	2			25	18	9	34				2					
<i>Nuphar</i>		1	2	3							1	1	1	1		2		2	1			1
<i>Myriophyllum</i>					2		6			3	8		3		1			1				
<i>Nymphaea</i> 22-28	4	1					1			2												
<i>Nymphaea</i> 40	8						2			4												4
<i>Callitriche</i>	3	1	2	3		1	1			4	3	1	6	9	7	1		6		1		6
<i>Brassenia</i>											2	2			1	1						
<i>Equisetum</i>	8	1			1	5	2			8	4	2	1				3		3			
Polypodiaceae						1				4	1			1			1					
<i>Woodsia</i>					5	1				1												
<i>Pteridium</i>	1										1											
<i>Lycopodium</i>											1	1		1		2						
Monolet spore										1	3		1	3			1	1				
<i>Tilletia</i> (smut)												3	15		1							
Unknowns	11	3	2	2		4	3			4	1	5	1	3	3		1		1			
Indeterminates	28	26	21	22	24	15	16	16		20	15	9	9	10	18	15	10	30	7	9	5	13
Total cell count	293	529	386	394	472	458	341	497	0	580	487	434	352	615	487	459	426	694	486	138	440	395

GPPS #	23	24	25	26	27	28	29	30	31	32
Charcoal>500um	24	25	20	36	31	21	48	66	57	43
Sporormiella 20x11										
Spor18x15				30						
Spor17x12			2		1					
Spor 17x17										
Spor15x7				11	1					
Marker Bead	167	90	74	87	152	224	373	79	87	162
<i>Pinus</i> undiff	9	10	3	15	8	5	5	6	3	4
Haplox <i>Pinus</i>	2	1		2		1	1	2		
Diplox <i>Pinus</i>	6	4	4	7	2	4	6	5	1	4
<i>Picea</i> undiff	1	1		1	1	1				1
<i>Abies</i>			1							
<i>Larix</i>										1
Cupressaceae	1	5	5	5	7	26	13	11	6	18
<i>Populus</i>	3	8	3		3	8	1	9	8	3
<i>Betula</i>										
<i>Frax</i> amer/pennC4								2		6
<i>Quercus</i>		1		4	2	45	23	134	54	157
<i>Ulmus</i>		1				38	22	9	17	23
<i>Acer</i>										
<i>Acer negundo</i>						3		1		
<i>Platanus</i>										
<i>Juglans</i>					1	23	12		1	1
<i>Carya</i>						3	2	7	4	6
<i>Liquidambar</i>								1		
<i>Nyssa</i>										
<i>Celtis</i>						1	1		1	1
<i>Cornus</i>										
<i>Maclura</i>										1
<i>Morus</i>						4		3	2	6
<i>Salix</i>					1	2	2			3
<i>Alnus</i>										
<i>Rhamnus</i>		1	1		1	2	1		2	1
<i>Sarcobatus</i>										
<i>Sambucus</i>	1	1				2				
Caprifoliaceae					1					
Rosaceae	2				1	11	4	4	4	3
<i>Ephedra</i>			1				1			
<i>Vitis</i>							1			
<i>Shepherdia</i>										
<i>Cephalanthus</i>										
<i>Arceuthobium</i>										
<i>Typha</i> (single)		2	1			2	4			
<i>Sagittaria</i>		1			1	1				
<i>Sparganium</i>						1	2			
Cyperaceae	1	1		1	67	2	1	5	10	4
<i>Lycopus</i>										
Aracea-like				1						
<i>Iva</i>	1	1	2	5		8	2	4	5	1
<i>Crassula</i>		2			2	3		2		2
<i>Ambrosia</i>	12	23	15	16	48	23	17	74	113	67
<i>Xanthium</i>	6	10	2	1	9	2	4	5	20	18
<i>Rumex</i>		2			1	1	1	5	1	1
Chenopodiaceae/Amaranthaceae	239	151	111	186	40	15	24	8	20	9
<i>Plantago</i>	6	16	21	8	4	4		2	1	1
<i>Artemisia</i>	6	9	2	4	6	4	9	1	3	3
Asteraceae low spine	23	12	10	40	21	19	4	18	20	7
Asteraceae high spine	11	14	114	50	19	1	1	13	7	6
<i>Taraxacum</i>		1		1	2					
<i>Thalictrum</i>									1	
Fabaceae	4	2	3		4	7			1	

Ranunculaceae		4		2	1		1	1	4	4
Linaceae									1	
Caryophyllaceae	16	12	3	13	1		2	2		
Papaveraceae										
Polygonaceae		2		1	5	1				
<i>Oxyria</i>			1			1				
<i>Polygonum lapathifolium</i>		1								
<i>Polyg. bistortoides</i>										
Saxifragaceae					3	7		6	3	6
Liliaceae	12									
Brassicaceae					2	2				1
Orchidaceae			1						1	
Scrophulariaceae										
Malvaceae	1	1			2					
Clusiaceae			1		1		1	3		4
<i>Viola</i>										
<i>Primula</i>										
<i>Urtica</i>							1	1		1
<i>Euphorbia</i>										
<i>Humulus</i>					1		1	3	4	2
<i>Geranium</i>										1
<i>Pedicularis</i>										
<i>Valerianella</i>										
Poa <20um		2	1		4	2	2	1	5	
Poa 20-25	3	36	18	4	43	1	9	2	7	2
Phrag pore<2.5										
Festuca pore=2.5	2	1	2		1					
Poa 26-30	2	54	12	11	43	3	5	13	12	10
Poa 31-35	6	6	1	2	12	4	1	7	7	4
Poa 36-40	2	4			2		3	1	3	1
Poa 45-60			1	1			5			
<i>Andropogon</i>		1		1		1	2			1
<i>Zea</i>			2							
<i>Lemna</i>	2					5	1	1		
<i>Isoetes</i>			1					29	36	6
Hydrocharitaceae		2								
<i>Potamogetum</i>					1	8		1		1
<i>Nuphar</i>	1	1			1			1		
<i>Myriophyllum</i>								5	15	45
<i>Nymphaea</i> 22-28										
<i>Nymphaea</i> 40					1	2				
<i>Callitriche</i>	2	6	1		4	6		6	1	4
<i>Brassenia</i>		1								
<i>Equisetum</i>			2					2	1	1
Polypodiaceae										
<i>Woodsia</i>										
<i>Pteridium</i>										
<i>Lycopodium</i>										
Monolete spore			1				1	2		1
<i>Tilletia</i> (smut)		3	2				2		1	2
Unknowns			6					1	1	
Indeterminates	9	10	5	7	10	26	5	8	16	11
Total cell count	392	427	360	389	390	341	206	427	423	466

Appendix D - *Ambrosia/Artemisia* Ratios

Table D.1: The percentiles of *Ambrosia/Artemisia* ratios for each prairie type.

	<i>Ambrosia/Artemisia Ratios</i>		
	Tallgrass	Mixed Grass	Short Grass
Min	0	0	0
Percentile 25%	3.32	0.015	0.026
Median	12.5	0.33	0.12
Percentile 75%	28.75	2.06	0.53
Max	169	73.99	8.00
Percentile 95%	78	19.73	3.69

Appendix E - R Code

```
#Importing all prairie pollen data
myprop<-read.table("Allgrassprop.csv",header=T,sep=",")

#Converting to matrix
myprop.matrix<-as.matrix(myprop[,1:47])

#Use analog package to calculate SCD
all.analog<-analog(allpollen,method="SQchord")

#Canberra distance metric example
dist.Canberra=dist(myprop.matrix,method="canberra",diag=TRUE,upper=TRUE)

#convert to a matrix for exporting and graphing.
candist.matrix<-as.matrix(dist.Canberra,rownames.force=NA,
nrow=559,ncol=559,byrow=FALSE,dimnames=NULL)

# display the distance matrix
image(1:559, 1:559, candist.matrix)

# Testing SCD's ability to delineate mixed grass prairie with a mixed grass prairie training set
#need to keep the training set dissimilarities
allTSMixed.ana <- analog(allpollen,trainingmixed, method = "SQchord",keep.train = TRUE)

# fit the ROC curve
clust <- hclust(as.dist(allTSMixed.ana$train), method = "ward")
grps <- cutree(clust,5)
allTSMixed.roc <- roc(allTSMixed.ana, groups = grps)
allTSMixed.roc
```

```
# draw the ROC curve
```

```
plot(allTSMixed.roc, 1)
```

```
#to plot Ambrosia/Artemisia Ratio versus maT (tave) and maP (annp)
```

```
plot(annp, AMBROSIA/ARTEMISIA, main="Ambrosia to Artemisia Ratio Compared to Mean  
Annual Precipitation", xlab="Mean Annual Precipitation (mm)", ylab="Ratio of Ambrosia to  
Artemisia", pch=20)
```

```
#Mann-Whitney-Wilcoxon Test do for each prairie type comparison
```

```
wilcox.test(Ratio ~ GrassTYPE, data= aaratio)
```