INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Université d'Ottawa • University of Ottawa

Late Quaternary paleoclimates and biogeography of North America

Michael Charles Sawada

Department of Geography
University of Ottawa

A dissertation submitted to the school of Graduate and Postdoctoral Studies
in partial fulfillment for the degree of
Doctor of Philosophy in Geography
June 25th, 2001

Examining Board: **Dr. Gail Chmura** (Professor of Geography and Center for Climate and Global Change Research, *McGill University*), **Dr. L.C. Cwynar** (Professor of Biology, *University of New Brunswick*), **Dr. Marianne S.V. Douglas** (Professor of Geology and Chair of Arctic Working Group, *University of Toronto*), **Dr. George Jacobson, Jr.** (Professor of Biological Sciences and Director of the Institute for Quaternary and Climate Studies, *University of Maine*, Bryand Global Sciences Center), **Dr. K. Gajewski** (Supervisor, *University of Ottawa*).

ABSTRACT

Pollen, spores, and dinoflagellate cysts are used with the modern analog technique (MAT) to provide paleoclimate reconstructions for terrestrial and marine environments in northeastern North America. Multivariate analysis of marine and nearby terrestrial pollen sequences from Hudson Bay, Labrador and the St. Lawrence, differentiate tundra, boreal and deciduous forest assemblages in time and space. These three regions had differing climate histories with respect to deglaciation and air mass boundaries. Prior to 6000 ^{14}C yr BP, cooler temperatures reconstructed along the Labrador margins agree with climate simulations indicating a persistent anticyclone over the Québec-Labrador ice sheet. A late Holocene cooling at forest-tundra sites suggests a recent southern movement in the mean position of the polar front.

The degree to which those critical thresholds of dissimilarity, that are used to identify non-analog pollen assemblages, are due to limitations of the modern pollen database or critical decisions within the MAT are explored by means of stochastic simulation, spatial statistics and graphical techniques. Critical thresholds of SQD, as derived by the expected value under randomization, become greater as the number of taxa in the pollen set increases. Larger pollen sets, with continentally infrequent but regionally abundant taxa, better distinguish between continental vegetation zones. Global and local spatial autocorrelation within climate anomalies indicate where the modern sample network induces biases in the climate reconstruction using the MAT.

The spatial scale of terrestrial climate or vegetation reconstructions from pollen in lake-sediments is investigated through the study of pollen source area in southern Québec. Despite the different regional vegetation, estimated pollen source areas and relevant pollen productivity for *Pinus*, *Picea*, *Abies*, *Fagus*, *Quercus* and *Tsuga* are consistent with studies from Michigan, Wisconsin and Sweden. These estimates are robust with respect to various plant abundance distance-weighting schemes and imply that the same inferences can be made regarding plant abundance from pollen throughout a lake-derived fossil pollen sequence. Stochastic simulations illustrate that the definition of relevant pollen source area requires consistent within-site vegetation heterogeneity within a network of pollen sites.

Underutilized proxy-climate data from wetland taxa are demonstrated to contain climate signals at the continental scale and have the potential to further our climatic and biogeographic picture of North America over the past 21,000 years. Pollen and spores from modern wetland taxa conform to their geographic ranges and allow interpretations of their past range changes. The climatic tolerances that govern their geographic distributions are used to interpret past range changes in climatic terms. *Sphagnum* spore distributions suggest major peatland developments after 9 ka and 5 ka. *Sphagnum, Potamogeton, Isoëtes, Myriophyllum Typha/Sparganium*, and *Menyanthes trifoliata* were in Alaska during the last glacial maximum (LGM) and moved into the ice-free corridor by 13 ka. Since the LGM, four migration routes for aquatic taxa are identified in response to the climate changes of the late Quaternary.

Respectfully and affectionately dedicated
to
Kathie, Michael Sr., Raija
and the greatly missed
Marilyn Sawada

PREFACE

The chapters of this dissertation were written for publication in peer-reviewed scientific journals; however, each contains details and explanations that were not published or submitted. Four of the seven chapters were written with co-authors; including my dissertation supervisor and research partners. I am the primary author on all of the chapters.

CHAPTER 2 was published in the Holocene (Sawada *et al.*, 1999). Dr. Anne de Vernal collected the dinocyst data and provided the temperature reconstructions for these cores as well as a paragraph in the methods section and Dr. Pierre Richard facilitated the terrestrial pollen site selection and was invaluable in the discussion and editing of the published paper.

CHAPTER 3 contains contributions from Dr. Pierre Richard who provided the Base de Données Polliniques et Macrofossiles du Québec, invaluable discussion, editing and encouragement. Dr. Richard negotiated the use of the Québec Ministère Resources Naturelles modern vegetation samples without which this would not have been possible. A large portion of this chapter has been submitted as a paper and is currently under peer review.

CHAPTER 5 represents my contributions to a larger project on reconstructing carbon sequestration in Peatlands of North America since the Last Glacial Maximum. This chapter is in-press and has been published in the July 2001 issue of the Journal *Global Biogeochemical Cycles* (Gajewski *et al.*, 2001) (Appendix F).

CHAPTER 6 Mr. Andre Viau provided the extraction of data from the North American Pollen Database (Grimm, 1999) and verification of this database for consistency and accuracy as well as invaluable discussion and original ideas. This chapter, in a reduced form, has been submitted and is under peer review.

ACKNOWLEDGEMENTS

The longer you're around, the more people there are deserving of thanks. About 9 years ago, I had the opportunity to go on a summer field camp with other undergraduates in the Yukon Ranges of Kluane National Park, Yukon. On the edge of the St. Elias icefieds, Dr. Peter Johnson (U. of O. Geography) shared with us his awe inspired sense of wonder for the processes that shape the natural landscape. I went back to Kluane for two subsequent years and completed a Master's in glacierized basin hydrology in 1996 under Peter's supervision.

In 1997, I begged my current supervisor, Dr. Konrad Gajewski to let me join his team (which at the time consisted of himself and David "Mad Dog" Atkinson (now Dr. Atkinson) and currently two postdoctral and 7 graduate students). I wanted to look at large-scale paleoclimate reconstructions. My current situation is largely due to Konrad's emphasis on high-quality research and publication and I cannot thank him enough for his constant vigilance in that regard.

Dr. Atkinson (of the superfluous publications "Breaking the neatline by D.E. Atkinson" and "A rat among the A's by D.E. Atkinson", among others) has been a mentor, coauthor and friend for many years. He has contributed to my learning and professional development on many occasions. Being a successful author, he is also adept at the art of figure captioning. Andre Viau, Konrad's newest Ph.D student has worked closely with me on many projects and our discussions combined with his database expertise have been invaluable in the development of this thesis and will undoubtedly continue into the future. Konrad has fostered an environment of free exchange and support at the Laboratory for Paleoclimatology and Climatology (LPC), which is most likely unique in academia.

This dissertation is based on large databases, without which it could not be completed. Almost all of the data in the first two chapters was secured by Dr. Pierre J.H. Richard (U. de Montreal). Although he is extremely busy with his own research, Pierre has always found the time to give me outstanding support during the past few years of work and I am fortunate to know him. He has an outstanding passion for the work and a keen eye for the big-picture. Dr. Anne De Vernal (U. de Québec a Montreal) has also provided data and encouragement. Dr. Thompson Webb, III (Brown U.) and Dr. Jack Williams (National Center for Ecological Analysis and Synthesis [NCEAS], Santa Barbara California) were invaluable in the construction of the new modern pollen database and both have reviewed parts of this thesis on an informal basis, providing constructive comments and directions.

Closer to home, Dr. Daniel Lagarec (U. of O. Geography) has read and provided invaluable comments on many parts of this dissertation. My conversations with him have provided precious direction in all circumstances. I would also like to sincerely thank Dr. Antoni Lewkowicz (U. of O. Geography) for giving me the opportunity to enter university teaching and encouraging me to continue in that vein. Dr. Barry Wellar (U. of O. Geography) has provided me precious insight into the methodological underpinnings of the science of geography that encouraged my spatial perspective.

My friends over the past few years have often noted my absence as I worked on this dissertation. I'm glad Pat, Mark, John, Jen, Phil, Steve, Pete and Toni didn't forsake me completely during that time. I'd especially like to thank Tana for the last few months as everything became rather hectic and hurried as the dissertation came to a timely end. I wouldn't have finished on time without her help. Thanks for the proof reading, editing, collating and gum on the floor. My sister Raija also joined "Team Sawada" in the final rush to get things done and worked tirelessly to that end.

This work could not be completed without ample funding for computers, software and basic needs like food and drink. The Government of Canada is acknowledged for the Natural Sciences and Engineering Research Council (NSERC PGS B) postgraduate scholarship and the Ontario Government for the Ontario Graduate Scholarship (OGS). My supervisor also provided research funding by way of his NSERC research grant. Funding was also acquired and gratefully acknowledged as part of the Climate Systems History and Dynamics (R. Peltier, U. of T. Physics) project supported by NSERC and Atmospheric Environment Service (AES), Canada. The Geological Survey of Canada also provided support in the dissemination of research results. J.P. Saucier and the ministère des Ressources naturelle du Québec are gratefully acknowledged for providing the vegetation database used in Chapter 3 of this dissertation. Eric Lalonde also helped with digitizing the range boundaries for some of the aquatic taxa in Chapter 6.

Of course throughout my graduate studies my family has been extremely supportive as I finished this last paper and so it is dedicated to them. Particularly my mother Kathie Sawada and father Michael Sawada Sr. have supported me throughout my Masters and Ph.D. My sister Raija Sawada, an adept mathematician, has even contributed to some of the research I've undertaken while spending many hours deriving mathematical formulae on my behalf, most of which unfortunately I never had the chance to pursue in this dissertation. My late aunt Marilyn Sawada always showed an interest and encouraged me to continue in my studies. Her smile and good cheer were always an inspiration. My aunt Bernadette Candelaro has also bestowed me grand sustenance. My great aunts Alice Widurski and Shirley Pelkey have given me undoubting support and encouragement as well as lively discussion.

LATE QUATERNARY PALEOCLIMATES AND BIOGEOGRAPHY OF NORTH AMERICA

TABLE OF CONTENTS

Appendices

LIST OF FIGURES

LIST OF TABLES

Chapter 1

1.1 INTRODUCTION

The Quaternary Period, spanning approximately the last 1.6 Ma (Monroe and Wicander, 1998), was characterized by climate change at a number of temporal scales. At timescales of 10^5 years, proxy climate records show evidence of glacial to interglacial cycles with periods of 100 ka (*kilo annum* = 1000 years) and implicate earth orbital changes (e.g., eccentricity, tilt and precession) as major forcings (Martinson *et al.*, 1987; Imbrie *et al.*, 1984). At a 10^3 year scale, during the last glacial period, the Wisconsinan, numerous strong warming/cooling signals (Dansgaard-Oeschger events) are found in the ^{18}O record of ice and ocean cores (Dansgaard *et al.*, 1993; Grootes *et al.*, 1993; GRIP Members 1993). These warm interstadial DO events lasted for a few hundred to a few thousand years (~0.4 to 12 ka from first warming to final cooling (Bender *et al.*, 1994)). Subsequent work has identified these DO events in different regions of the northern and southern hemispheres, such as the Antarctic (Bender *et al.* 1994), Pacific Santa Barbara basin (Behl and Kennett, 1996) and mid-Atlantic (Adkins *et al.*, 1997). At scales of 10^2 years, the transition from glacial to interglacial conditions between ~20 - 10 ka was characterized by numerous oscillations, including the Bolling-Allerod, Younger-Dryas (Grootes *et al.*, 1993; Dansgaard *et al.*, 1993; GRIP Members 1993) and changes in North American vegetation (Grimm and Jacobson 1992; Jacobson and Grimm, 1986). Abrupt climate changes at the century scale in the late Quaternary are evident in the transition from cold YD conditions to Holocene conditions which may have taken only 40-50 years and were coincident with increases in methane, water vapour and a decrease in atmospheric dust (Taylor *et al.* 1997). The Holocene, beginning ~ 10 ka, has also experienced climate variability at the millennial and century timescales

(Gajewski, 1987, 1993; Kutzbach and Webb, 1993). Understanding the regional expression of these climate changes over the globe is a major research challenge.

The cyclic nature of vegetation changes during successive interglacial cycles has been observed in European pollen records for the Eemian (~MIS 5e in North America) and Holocene (Field *et al.*, 1994). Vegetation change in North America over the past 21 ka as inferred from pollen and macrofossil records suggests that most populations were displaced during the last glacial maximum and these subsequently acted as spreading centres for large-scale migrations during deglaciation (Jackson *et al.*, 1997; McGlone, 1996; Overpeck et al. 1992; Jacobson *et al.* 1987; Webb, 1987; Webb *et al.*, 1993).

A significant amount of research by Quaternary scientists implicates climate as the primary independent driving force of late Quaternary vegetation change on continental scales (e.g., Webb, 1997; Webb, 1993, 1986; Grimm and Jacobson, 1992; Davis *et al.*, 1991; Prentice *et al.*, 1991; Huntley and Webb, 1989; Gajewski, 1987; Jacobson *et al.*, 1987; Jacobson and Grimm, 1986; Prentice, 1986; Ritchie, 1986; Davis *et al.*, 1986; Davis, 1981, 1976). The degree to which hemispheric and global climate changes of the late Quaternary were experienced regionally and locally within a continent can be determined using fossil pollen and spores as proxy-climate indicators. Quantitative paleoclimate reconstruction techniques employing fossil pollen sequences can be used to show the expression of regional climatic changes over the past 21,000 years in North America.

In the past few years there have been vast improvements in the proxy data that have become available in North America. In particular, the development of the North American Pollen Database (NAPD) (Grimm, 1999; Contributors to the NAPD, 1999) initially by T. Webb, III (Brown University) and now administered by E.C. Grimm (Illinois State Museum)[**] has allowed research into the large-scale vegetation and climate changes in North America over the late Quaternary. The NAPD embodies the amalgamation of over 800 fossil pollen sites

[**] The NAPD also benefits from expert advisory board of palynologists, paleoecologists and paleoclimatologists composed of K.J. Gajewski, G.L. Jacobson Jr., G.M. MacDonald, L.J. Maher, V. Markgraf, P.J.H. Richard, T. Webb III, and C. Whitlock.

originally collected and analyzed by individual researchers (with 2000 additional candidate sites as of June, 2001), and represents many thousands of person hours of work by those individuals over the past few decades. All of these studies are now standardized into a single taxonomically consistent database. Such datasets and better utilization of existing data allows us to advance the methodologies and challenge the assumptions underlying paleoclimatic and biogeographic reconstructions at the continental scale. This dissertation uses fossil pollen, spores, dinoflagellate cysts and modern vegetation and climate observations from the NAPD and other databases in order to contribute to our understanding of the past climates and biogeography of the North American continent during the late Quaternary.

1.2 Overview and Organization of Dissertation Research

Improving paleoclimatic reconstructions at the North American scale is the unifying theme of this dissertation. New quantitative climate reconstructions and syntheses of the marine and terrestrial environments are presented as well as an exploration of the underlying assumptions of using pollen data for climate reconstructions. Moreover, the proxy climate record is explored for underutilized data with climatic signals that show promise for improving and/or constraining paleoclimatic reconstructions.

An ultimate goal of paleoclimatic research is the creation of global maps of marine, ocean and terrestrial climates on the same spatial and temporal scales that can aid us in better understanding the processes behind climate changes. One step towards deriving such hemispheric maps of Holocene climates is the comparison of climate reconstructions within the marine and terrestrial environments across broad regions of North America at the same temporal scales. Chronological control in the marine and terrestrial environments differ (Stuiver and Reimer, 1993) and marine sediments often lack sufficient organic matter to be dated. Chapter 2 presents a new approach to adding confidence to the relative dating of marine sediments by utilizing multivariate analysis. Confidence in the chronological control allowed the reconstruction and comparison of sea surface and terrestrial air temperatures in northeastern North America

using marine dinoflagellate cysts and terrestrial pollen using the Modern Analog Technique (MAT) (method of modern analogs) (Prentice, 1980; Overpeck *et al.*, 1985).

Two major questions arose with the marine-terrestrial quantitative paleoclimate reconstructions during another research project that aimed to produce climate reconstructions for 6 ka (Gajewski *et al.*, 2001). First, the similarity between the marine and terrestrial pollen signals (Chapter 2) raised questions of the spatial scale at which pollen is capable of representing vegetation. Understanding the pollen source area and those pollen types that are significantly related to plant abundance are fundamental to reconstructing past climates and biogeographies using networks of sites on the continental scale. There is only one region (Wisconsin and Michigan) in North America where pollen has been compared to plant abundance, and these studies utilized basal area as the plant abundance measure. The availability of a new dataset in southern Québec collected over a 6-year period by the Québec Government containing areally-projected crown cover is used in Chapter 3 to explore the pollen-plant abundance relation and to determine those pollen types that are best related to plant abundance. This chapter critically explores the various pollen-plant abundance calibration models and underlying assumptions of the techniques.

The second question arising from Chapter 2 was the need to better understand why there is greater dissimilarity between the modern and fossil pollen spectra as we go further back in time. Specifically, critical thresholds of dissimilarity are used to determine whether or not a modern pollen assemblage is a 'good' or 'poor' analog (or non-analog) for a fossil pollen assemblage. Was the increasing dissimilarity between modern and fossil pollen spectra observed for pollen assemblages in the older sediments (so called "down-core") of Chapter 2 caused by past vegetation assemblages that are very different from today [the non-analog problem (Jackson *et al.*, 1997; Webb, 1987; Overpeck *et al.*, 1985; Prentice, 1980)]? Alternatively, are non-analogs and the critical thresholds of squared chord distance (SQD) partially a function of the technique used to measure dissimilarity between pollen spectra? In Chapter 4, using a new modern pollen database and stochastic simulation, the critical limits of the squared chord distance (SQD) coefficient between pollen spectra that determine a good vs. poor analog

are explored at the continental scale. Several other issues in the application of the MAT are explored in Chapter 4, including the effect of the pollen set on the univariate and spatial biases in the climate reconstructions and the use of stochastic simulation and spatial statistics to aid us in better understanding the limitations associated with the application of the MAT. Chapter 4 presents a general methodology that may be applied to any proxy dataset subjected to the MAT.

The first part of this dissertation (Chapters 2, 3, and 4) presents the development of new quantitative paleoclimate reconstructions, explores the spatial scale at which lake basins are capable of representing surrounding vegetation, and analyses the assumptions of the MAT. The second section of this thesis concentrates on the exploration of the fossil pollen and spore record in North America for underutilized information that can be used to increase our knowledge of past climates. The first step to utilizing new elements of proxy climate data for quantitative or qualitative climate reconstructions, requires assurance of data quality and a demonstration of the biological-proxy-climate relations.

Chapter 5 shows that *Sphagnum* spores can be used to estimate past peatland distribution and changes in peatland extent over the past 21,000 years in North America. Moreover, the climatic signal in the distribution of modern pollen and spores is made explicit which is of significance for the use of spore distributions in the past. Chapter 6 explores the continental-scale climate signal in aquatic pollen and spores as a source of new paleoclimate information (Dieffenbacher-Krall, 1998; Dieffenbacher-Krall and Jacobson, 2001). This chapter demonstrates the conformity between the modern ranges of aquatic plants and the pollen and spores they produce. Such information is used to show the climatic signal within aquatic pollen and spore records as well as to interpret maps for the past 21,000 ka that illustrate the continental-scale history of aquatic plant migrations. Consequently, both of these last chapters contribute to our understanding of the biogeographic history of North America.

The chapters in this dissertation are in the form of scientific papers and the relevant literature is reviewed within each chapter. The taxonomic authority used in Chapters 2, 3, 4 and 5, is Gleason and Cronquist (1963) and in Chapter 6, the Flora of North America North of Mexico (Flora of North America Editorial Committee, 1993; 1997; 2000). There are a number of appendices that contain either technical contributions and/or scientific literature that I have co-authored that are related to and/or in support of the methodologies and approaches taken in this dissertation. These appendices are referenced within their relevant chapters.

Chapter 2

COMPARISON OF MARINE AND TERRESTRIAL HOLOCENE CLIMATIC RECONSTRUCTIONS FROM NORTHEASTERN NORTH AMERICA

2.1 INTRODUCTION

Working separately, terrestrial and marine paleoclimatologists have reconstructed past climates of their respective environments and created maps of changing climate conditions through the Holocene (e.g., Wright *et al.*, 1993). An ultimate goal of such work is the creation of hemispheric maps of marine, ocean and terrestrial climates on the same time-scales. A first step towards deriving such hemispheric maps of Holocene climates is the comparison of climate reconstructions in the marine and terrestrial domains. The comparison of sea surface and terrestrial air temperature reconstructions from the marine dinoflagellate cyst record and terrestrial pollen record is the focus of this chapter.

2.1.1 DINOFLAGELLATE CYSTS AND CLIMATE

Dinoflagellates inhabit the marine photic zone and can also be found in brackish estuarine and some fresh water environments. Dinoflagellates are of the botanical group Pyrrhophyta (fire-plants). They are single celled organisms that reproduce predominantly asexually and are generally between 20 and 150 μm long (Brasier, 1994). Dinoflagellates have both plant and animal characteristics but are generally considered plants or algae due to cellulose in the cell walls and the presence of chloroplasts (Brasier, 1994). The dinoflagellate exhibits two life stages: a planktonic motile stage and a benthic cyst stage. The motile plant encysts during sexual reproduction or during less than favourable environmental conditions. The cyst,

made of decay-resistant materials, is shed as the motile stage excysts when conditions are favourable again. The cyst stage may enable the species to survive stressing conditions, allowing repopulation when conditions improve (Wall, 1971). The cysts can often be identified to the species level and are preserved in bottom sediments leaving a record of their changing abundance and taxonomy through time. Dinocysts are easily isolated from marine sediments by palynological methods (Figure 2.1).

The growth of the dinoflagellate is sensitive to the same factors affecting the growth of other phytoplankton: incident and effective radiation, temperature, salinity, nutrients, and hydrodynamic controls (Williams, 1971a). However, it is widely recognized that temperature and salinity are in most cases the primary controls on dinocyst distributions and abundance (Dodge, 1994; Marret, 1994; Versteegh and Zonneveld, 1994; De Vernal *et al.*, 1989, 1993; Williams, 1971a). Early work by Wall (1971) showed that most dinocysts are cosmopolitan but that the abundance centres of individual species are geographically and climatically well-defined. Contemporary studies allow us to assume that the geographic abundance of cysts in bottom sediments corresponds to the areas where the motile organism thrived under favourable environmental conditions, making the study of their compositional changes through time useful for purposes of reconstructing marine surface climatologies.

Qualitative reconstructions of temperature, salinity and circulation patterns in surface waters have been inferred on many temporal scales using sequences of dinocysts in marine sediments (e.g. MacRae *et al.*, 1996; Williams *et al.*, 1995; Marret, 1994; Matsuoka, 1994; Versteegh and Zonneveld, 1994; ; Bilodeau *et al.*, 1990; Duane and Harland, 1990; de Vernal *et al.*, 1989). The relative abundance of organic-walled dinocysts, which are produced after the sexual reproduction of dinoflagellates, relates to productivity in the photic zone. Their modern distribution conforms to large-scale marine temperature gradients and salinity conditions that apparently control the species diversity and assemblages (e.g. de Vernal *et al.*, 1994; Dodge, 1994; Marret, 1994; Versteegh and Zonneveld, 1994, de Vernal *et al.*, 1993). Contemporary dinocyst-climate relations have been used with modern analog techniques (Guiot, 1990; Overpeck *et al.*, 1985) to

quantitatively reconstruct sea surface temperature and salinity conditions of the Late Quaternary and Holocene (Levac and de Vernal, 1997; de Vernal *et al.*, 1996; de Vernal *et al.*, 1994; de Vernal *et al.*, 1993).

2.1.2 TERRESTRIAL POLLEN AND CLIMATE

Pollen grains are the male gametophytes of flowering plants and do not themselves respond to climate parameters, although pollen production is implicitly linked to climate through vegetation. Pollen type and abundance within sedimentary records are functions of the surrounding vegetation composition and abundance that are in turn functions of the prevailing biotic and abiotic conditions. Pollen accumulates in lake sediments providing a record of vegetation change through time at local to regional spatial scales (e.g. Richard, 1996; Prentice, 1986; Prentice, 1985). Analyses of modern and fossil pollen indicate that continental to sub-continental scale vegetation patterns over the Holocene and longer time periods have responded to changes in climate (e.g. Prentice *et al.*, 1991). Pollen from well-dated sediments has been used in quantitative estimates of paleoclimate changes in terrestrial environments (Wright *et al.*, 1993). Although the relation between pollen proportions and relative vegetation abundance is not one-to-one (Prentice, 1985), it is sufficient in eastern North America (Davis and Webb, 1975) to allow modern climate-pollen relations to be used in reconstructing past climates (Arigo *et al.*, 1986).

2.1.3 MARINE AND TERRESTRIAL POLLEN AND PROVENANCE

Pollen abundance within a particular depositional environment is partly a function of the dispersal vector: air, water and/or sediment. Generally, different pollen types will be transported in different ways depending on their hydrodynamic properties and in different amounts depending on production by vegetation (Tauber, 1965). Most pollen in the terrestrial environment is assumed to be deposited from the atmosphere within a short distance from its source. The 'source area' or provenance of a lake collecting basin is assumed to be between 10 km to 100 km (Jacobson and Bradshaw, 1981; Lynch, 1996; Prentice, 1985, 1986; Groot, 1971). However, the study of pollen provenance in lake collecting basins is based largely on a single regional dataset in Michigan and Wisconsin (Prentice and Webb, 1986; Bradshaw and Webb, 1985; Prentice, 1985; Prentice and Parsons, 1983; Webb *et al.*, 1981) and one in Sweden (Prentice

et al., 1987). There are no studies on the regional scale in the mixed boreal forests of eastern North America with the exception of Davis and Webb (1975) and this is a goal of future research. Here the term "terrestrial pollen" is used to denote pollen from lake sediments, and "marine pollen" to denote pollen from marine cores, although terrestrial vegetation is the source of both pollen.

Atmospheric pollen concentrations over the marine environment has been shown to vary depending on season, the amount of pollen produced, the amount shed, distance from the source area, wind direction, and strength and relative timing of these factors (Calleja *et al.*, 1993). That is to say, the transport of pollen in the marine environment will depend on the climatic conditions, such as the direction of mean circulation, air temperature, the given time of season *et cetera*. Further, given the same velocity, pollen will be transported for much greater distances by a dense medium, such as water (Groot, 1971). Groot and Groot (1971) show that the absolute counts of pollen grains generally decrease linearly as the distance from the shore increases regardless of whether the coast is windward or leeward of the dominant atmospheric circulation. Korvena (1971) came to a similar conclusion for pollen counts in Mediterranean bottom sediments. Groot (1971) reviews research that indicates that in estuarine environments and areas influenced by coastal ocean currents, pollen can be transported over great distances and be representative of a mix of vegetation zones. In some cases the influence of nearby vegetation was shown to be of little influence on the neretic (near shore) marine pollen spectra (Groot, 1971). Marine pollen sequences have been used in a number of studies to reconstruct terrestrial paleovegetation changes and paleoclimate over a range of spatial (continental/subcontinental) and temporal scales (> 10^3 years) (e.g., Lézine and Denefle, 1997; Ning and Dupont, 1997; Hooghiemstra *et al.*, 1992; Lézine, 1991; Heusser and Morley, 1990; Korveneva, 1971). The more recent studies (e.g., Hooghiemstra, 1992; Heusser and Morley, 1990) indicate that marine and ocean current transport are less important than were once suspected and that in general, marine pollen records, unlike their terrestrial counterparts, tend to register vegetation patterns on large regional to subcontinental scales rather than local and regional scales because of different basin characteristics (Ning and Dupont, 1997).

2.1.4 Marine and Terrestrial Paleoclimate Reconstructions

As a first step to comparing marine and terrestrial paleoclimate estimates, the chronologies of both must be resolved. Work at individual sites has used pollen in marine cores, in conjunction with well-dated terrestrial pollen sequences, to provide chronologies using relative dating principles for poorly dated marine cores lacking sufficient organic matter (Levac and de Vernal, 1997; de Vernal *et al.*, 1994; de Vernal *et al.*, 1993; Bilodeau, 1990; de Vernal *et al.*, 1989). This has indirectly shown that marine pollen sequences covary with terrestrial sequences on fine temporal (10^3 years) and spatial scales. However, it is not clear if this can be extended to larger spatial scales in northeastern North America because the pollen signatures in the two environments differ. We therefore needed to quantify the agreement between marine and terrestrial pollen assemblages with two goals: (a) to aid in the dating of marine cores with low organic matter and (b) to compare the quantitative climate reconstructions produced from proxy data in the two different environments over broader spatial (regional) and smaller temporal scales than previously examined in northeastern North America.

It can be shown by using principal components analysis (PCA) (Johnston, 1980) that marine dinocysts and terrestrial pollen record similar responses to environmental forcing over the Holocene in eastern Canada. PCA also indicates that pollen assemblages in both terrestrial (lake-sediment) and marine cores off eastern Canada record regional terrestrial environmental changes through time. This observation allows us to confidently use marine-terrestrial pollen correlations to aid in securing chronologies for the marine cores with low organic matter. Increased confidence in the marine age models permits direct comparison of marine/terrestrial climate fluctuations of sea surface temperature and air temperature over the Holocene and the quantitative reconstruction, using the method of modern analogs (modern analog technique - MAT), of August sea surface temperature (SST) and August terrestrial air temperatures derived from dinocysts and pollen, respectively.

2.2 Data and Methods

2.2.1 Marine and Terrestrial Data

In this study we use pollen and dinocyst data from marine cores from Hudson Bay (cf. Bilodeau *et al.*, 1990), the Labrador Shelf (Levac and de Vernal, 1997) and the Gulf of St. Lawrence (de Vernal *et al.*, 1993) (Table 2.1, Figure 2.2).

Table 2.1: ***Location of sites and references for microfossil diagrams.***

Environment	Site	Latitude	Longitude	Depth / Elevation	Reference
Marine	Hudson Bay				
	87-028-001 (TWC&P)	63°01.45' N	81°04.91' W	273 m	Bilodeau *et al.*, 1990
	87-028-069 (TWC&P)	55°28.62' N	77°57.78' W	165 m	Bilodeau *et al.*, 1990
	Labrador coast				
	HU91-045-005 (B)	54°42.26' N	56°27.06' W	530 m	Levac and de Vernal, 1997
	HU91-045-006 (TWC&P)	54°42.27' N	56°26.96' W	534 m	Levac and de Vernal, 1997
	St Lawrence				
	89-007-111 (TWC&P)	47°31.00' N	59°53.06' W	503 m	de Vernal *et al.*, 1993, 1996
Terrestrial	Northwestern Québec				
	LR1	58° 34.98'N	75° 15.00' W	170 m	Gajewski *et al.*, 1993
	LB1	57° 55.20'N	75° 37.20' W	200 m	Gajewski *et al.*, 1993
	BI2	57° 07.00'N	76° 22.70' W	210 m	Gajewski and Garralla, 1992
	GB2	55° 00.17'N	75° 17.00' W	300 m	Gajewski *et al.*, 1993
	Labrador				
	Lake Hope Simpson	52° 27.00'N	56° 20.00' W	295 m	Engstrom and Hansen, 1985
	Gravel Ridge	55° 02.00'N	62° 38.00' W	565 m	Lamb, 1985
	Eagle Lake	53° 14.00'N	58° 33.33' W	400 m	Lamb, 1980
	St Lawrence				
	Lac à l'Ange	47° 37.00'N	70° 41.83' W	648 m	Labelle and Richard, 1981
	Lac Mimi	47° 37.17'N	70° 41.83' W	423 m	Richard and Poulin, 1976
	Lac J'arrive	49° 22.33'N	65° 27.83' W	56 m	Marcoux and Richard, 1995
	Lac Euloge	49° 20.67'N	65° 25.33' W	83 m	Marcoux and Richard, 1995
	Lac à Raymond	48° 14.00'N	65° 59.67' W	50 m	Jetté and Richard, 1992
	Lac à Robin	48° 03.07'N	65° 16.01' W	50 m	Jetté and Richard, 1992

All marine cores were analyzed using standard palynological techniques (de Vernal *et al.*, 1986). The marine cores from Hudson Bay and Labrador are all farther than 200 km from any major estuaries which means that river-borne pollen may not significantly distort the marine pollen spectra (Ning and Dupont, 1997). This is not the case for the St. Lawrence core, which integrates pollen taxa from a large and diverse provenance (Figure 2.2).

Terrestrial pollen data from 13 cores were taken from published sources (Table 2.1). Details of the dating and laboratory techniques used in the terrestrial cores can be found in their respective references. Terrestrial sites to be used in this study were chosen close to the marine cores.

For PCA and analog analysis, the pollen set consisted of 58 types (Table 2.2).

Table 2.2: Pollen set used in present study.[**] Alternative nomanclature in brackets.

TYPE	TYPE
Picea	*Artemisia*
Abies	Liguliflorae
Pinus	Tubuliflorae
Betula	Caryophyllaceae
Larix	Chenopodiaceae
Fraxinus	Cyperaceae
Cupressaceae	Poaceae (Gramineae)
Populus	*Epilobium*
Quercus	Fabaceae (Leguminae)
Tsuga	Ranunculaceae
Ulmus	Saxifragaceae
Fagus	*Thalictrum*
Acer	*Oxyria-Rumex*
Carya	Umbelliferaceae
Ostrya-Carpinus	*Polygonum*
Juglans	Brassicaceae (Cruciferae)
Tilia	Lycopodiaceae
Castanea	*Equisetum*
Alnus cripsa (Ait.) Pursh.	*Selaginella*
Alnus rugosa (DuRoi) Spreng.	*Sphagnum*
Ericaceae	Polypodiaceae
Myrica	*Botrychium*
Salix	Other herbs (OHE)
Corylus	Other trees & shrubs (OTS)
Cornus	*Typha*
Shepherdia	Triletes
Viburnum	*Potamogeton*
Ilex-Nemopanthus	*Isoëtes*
Platanus	Other Aquatics
Rosaceae	

For the terrestrial modern analog analysis, *Ambrosia*, Poaceae and *Rumex* were excluded because of overrepresentation in modern samples due to human disturbance of the landscape since European settlement. *Iva* was also excluded because it is lumped with Ambrosia in the database. Cyperaceae and

[**] Alternative nomenclature in brackets.

Ericaceae pollen were excluded because of their local effects on spectra in forested areas. Aquatic pollen was retained in the dissimilarity measures. However, the climatic signal in aquatic pollen has not been studied at the continental scale and this is the subject of ongoing work that will be discussed in the last chapter. Because of the large number of taxa included in the pollen set, we used a square root transform of the pollen percentages in the PCA. Taking the square root has the effect of increasing the relative importance of less abundant pollen taxa (Prentice, 1980).

To compare the inter-correlations between different taxa in the marine and terrestrial cores, principal components analysis was used (Johnston, 1980). A PCA identifies groups of taxa that have similar inter-correlations. The group of taxa that have stronger inter-correlations are approximated by a component, with each component orthogonal and independent of the previous. Each component is a linear combination of all the variables present. Every taxon is included in each component; however, a given component will explain more of the variance in one taxon than another. Those taxa that have the greatest variance explained by a given component are considered important to the component. It is these important taxa that give each component its unique interpretation. Furthermore, subsequent components explain less of the total variance in the set of variables. Component scores are interpreted as the correlation a given observation has with a given factor. That is to say, an observation's score will be high on a given factor if it has a high value for a variable with a high loading on that factor. In other words, if observations are sequences of time, such as timeslices in a core, then the scores will portray how a factor changes through time.

A database of modern pollen samples from lakes and bogs was obtained from T. Webb (Webb *et al.*, 1993). From this we extracted pollen assemblages between 37° - 82° N and 53° - 96° W. We removed older sites where pollen counts were incomplete, leaving 1573 assemblages to which were added 68 from Gajewski (1991, 1995).

2.2.2 Climate Reconstructions and the Analog Method

Quantitative climate reconstructions of the terrestrial sequences were completed using the method of modern analogs (modern analog technique - MAT) with dissimilarity measures based on squared chord distance (Overpeck *et al.*, 1985). The squared chord distance is given by:

$$d_{ij} = \sum_{k} (p_{ik}^{1/2} - p_{jk}^{1/2})^2 \qquad (2.1)$$

Here, d_{ij} is the dissimilarity coefficient between two pollen spectra i and j whose proportions at a fossil sample site i, p_{ik}, are compared to the proportions at a modern sample site j, p_{jk}, for each pollen type k. (Overpeck *et al.*, 1985). A dissimilarity coefficient, d_{ij}, is computed for each and every modern site for each fossil spectrum and the modern spectrum with the smallest d_{ij} is considered the best modern analog (Figure 2.3).

The analog method compares a fossil pollen spectrum to all modern spectra within a given geographic extent. For a given fossil pollen spectrum, the best analog is the modern site that possesses the least dissimilar modern pollen spectrum – in this study 'least dissimilar' means the smallest squared chord distance (Equation 2.1). The environment that gave rise to the pollen spectra at the best analog site presumably also produced the fossil pollen spectrum (Overpeck *et al.*, 1985). Overpeck *et al.* (1985) tested the responsiveness of eight dissimilarity coefficients and found that signal-to-noise measures like squared chord distance performed best when reconstructing an original pollen spectrum from a choice of modern sites. They found that unweighted coefficients like Manhattan metric, Squared Euclidean Distance, Squared cosine $-\theta$ (Figure 2.3) distance all tend to be heavily influenced by pollen types with large ranges, generally the most common types, and types with low values and ranges tend to have little influence on these measures. Alternatively, equal weight measures like the Canberra Metric and Squared standardized Euclidean distance (Figure 2.3) tend to increase the weight of less common types and decrease the weight of large valued pollen types. These measures may accentuate noisy types at the expense of

types with a strong signal but small values. The third set of coefficients is named Signal to Noise and includes Squared Chord Distance, Information Statistic and Squared Chi Squared Coefficient (Figure 2.3). These tended to emphasize the signal component of the differences between spectra while downplaying the noise component. They are influenced by less common pollen types to a greater degree than the unweighted coefficients but to a lesser degree than the equal weight coefficients making these signal to noise ratios a middle of the road alternative. The dissimilarity coefficients all take a value of zero between identical pollen spectra and positive values otherwise. The squared chord distance measure ranges from zero for a perfect analog to two for a completely dissimilar site.

For the terrestrial climate, the mean August temperatures of the five best analog sites for each pollen spectrum in each core were averaged. Accuracy of the analog method was verified by estimating modern mean August terrestrial air temperatures from the modern pollen assemblages using the average of the 2^{nd} - 6^{th} best modern analogs (coefficient of correlation between simulated and instrumental values: $r = 0.97$).

Sea surface conditions were reconstructed from the dinocyst assemblages using the modern analog method (Guiot, 1990). The August sea surface temperature estimates correspond to the weighted average of values from the ten best analogs in the modern dinocyst database. In both terrestrial and marine reconstructions the averaging method was chosen to minimize the variability of temperature reconstructions in each particular dataset. The modern dinocyst reference database was from 439 modern sites in the middle to high latitudes of the North Atlantic, subpolar basins and adjacent epicontinental seas; it also included reference data points from polynya in the Arctic Ocean, Siberian and Beaufort Seas. The modern dinocyst database is representative of a wide range of sea surface conditions. The validation tests for the reconstruction of modern sea surface temperature in August yielded accurate results: the coefficient of correlation between simulated and instrumental values is 0.96; more than 95% of estimates are included within the range of interannual variability, i.e., the standard deviation around the instrumental data averages.

2.3 RESULTS

2.3.1 PRINCIPAL COMPONENTS ANALYSIS

In order to discern the relations among the taxa and the different groups of sites, four PCAs were performed. With reference to the geographic locations in Figure 2.2, these include: 1) pollen data from all terrestrial sites; 2) pollen data from marine cores; 3) pollen data from the combined marine and terrestrial cores; and 4) dinocysts in the marine cores.

The first PCA was performed on the terrestrial (lake-sediment) sequences (Figure 2.4a). The first component explains 18.75% of the variance and has high positive loadings for tundra taxa (e.g. Cyperaceae, Poaceae, *and Artemisia*) and high negative loadings for tree taxa (*Acer*, *Betula*, *Fagus*, *and Abies)*. Thus, the first component distinguishes between tundra and more southern environments. The second component shows strong positive loadings for boreal forest taxa (*Picea*, *Alnus crispa*) and negative loadings for non-boreal types. This component explains 16.35% of the variance and distinguishes between boreal taxa and those from tundra or southern forest assemblages.

The first two principal components of the pollen taxa from all of the marine cores present a signal similar to that of the terrestrial pollen alone (the sign of the loadings is arbitrary in a PCA), however, the relative importance of the first two components is reversed (Figure 2.4b). The first component explains 13.94% of the variance and separates boreal from southern forest taxa, while the second explains 12.42% of the variance and separates tundra from southern tree taxa. The highest positive loading on the first component is for *Pinus*. *Picea*, which is the most important tree of the boreal forest, does not load significantly on either of the first two components. Despite differences in loadings, taxa with similar environmental responses covary in a manner similar to the terrestrial pollen PCA.

The first component for the combined analysis of the marine and terrestrial pollen explains 16.29% of the variance and the second explains 14.14% of the variance (Figure 2.4c). The component loadings again

distinguish between tundra, boreal and southern forest taxa similar to those seen in the PCA of the terrestrial sites alone.

The dinoflagellate assemblages in the marine cores show similar environmental differentiation. The first component for the dinocyst taxa in the marine cores explains 20.56% of the variation and distinguishes between temperate taxa such as *Spiniferites ramosus*, *Peridinium faeroense* and *Spiniferites elongatus* with high negative loadings (Figure 2.4d). Cool water taxa have high positive loadings on the first component and to some extent so do open ocean taxa. The second component explains 13.93% of the variation and seems to reflect onshore to offshore differences and/or salinity gradients.

The warm taxa - cool taxa signal evident in the terrestrial and combined terrestrial-marine pollen PCA is also evident for the dinocysts. These four analyses indicate similar responses to environmental forcing over the Holocene for taxa within the marine and terrestrial environments. The component scores show the importance through time of the climate signal seen in the component loadings; representative terrestrial sites are illustrated in Figure 2.5. Each site shows a transition from tundra to forest conditions; this transition is earlier at the southern sites (Lac à l'Ange, Lac Euloge) and later at sites in higher latitudes (LB1, Eagle Lake). Lake LB1 shows an abrupt transition from tundra to forest conditions shortly after deglaciation in northwestern Québec. The scores on the second component for these sites (not shown) indicate a time transgressive change from boreal to non-boreal conditions.

2.3.2 Derivation of marine chronologies

Since the component scores differentiate the relative timings of vegetative change by geographic location, we attempted to use these as a chronostratigraphical correlation tool for the marine cores in order to develop an age model consistent for both onshore and offshore time series. The ^{14}C-dated component scores for the nearest terrestrial sites were plotted adjacent to the scores for the marine sites and chronologies were derived based on previously assigned ^{14}C dates, distinctive events or gradients in the two curves' component scores, and/or events in the marine-terrestrial pollen spectra indicative of regional

vegetation changes in eastern Canada over the Holocene. Core 89-007-111 in the St. Lawrence Estuary was dated by ^{14}C (de Vernal *et al.*, 1996; de Vernal *et al.*, 1993), and core HU91-045-006 in the Cartwright Saddle off Labrador had previously been dated by assemblage zones in the original marine and terrestrial pollen diagrams and ^{14}C (Levac and de Vernal, 1997). The covariation in the component scores between the marine and terrestrial sequences for these sites added confidence to the dating derived by previous work (Figure 2.6 & 2.7). However, the Hudson Bay marine and northwestern Québec terrestrial component scores showed no distinct gradients or trends that unequivocally indicated changes from one vegetation type to another. These sequences are short, < 8000 years, with only a brief tundra period followed by the establishment of the boreal forest. Thus, in the case of core 87-028-069, the component scores were not as useful in chronological correlation as were the original pollen diagrams and previous work done with assemblage zones and chronostratigraphy by Bilodeau *et al.* (1990). In general, the consistent patterns in the marine pollen data with respect to the PCA loadings and scores added confidence to the marine core age models derived by previous work.

2.3.3 Quantitative climate reconstructions

With the chronologies established for the marine sequences, climate reconstructions using the dinocysts in the marine cores and the pollen in the lake cores were completed using the method of modern analogs (modern analog technique - MAT). The climate reconstructions of mean August temperature are presented by region (Figures 2.8-2.10).

In northwestern Québec / Hudson Bay, site GB2 shows a warming throughout the Holocene, while site LB1 shows maximum temperatures between 3000 and 1000 yr BP (Figure 2.8). These temperature curves generally parallel the change in spruce proportions over the same time period at these sites (Gajewski *et al.*, 1993). The climate reconstruction from site LR1, in the present-day tundra, shows little variation during the Holocene (Figure 2.8). This is because the pollen percentages change little down-core, although, frequently in tundra, climate signals may instead be expressed by changes in the concentrations or pollen influx (Gajewski, 1995). The samples older than 6000 yr BP at sites BI2, LB1 and LR1 show warmer

temperatures, but these are associated with larger dissimilarities, and arise due to larger relative amounts of tree taxa pollen blown from the south at a time when the local pollen production was low. Marine core 87-028-069, in Hudson Bay near site GB2, shows a flat maximum temperature extending from 5700 yr BP to 2000 yr BP followed by cooling toward present-day conditions (Figure 2.8). The timing of temperature changes in 87-028-069 parallel those of site GB2, except that there is no neoglacial cooling at GB2 (Figure 2.8). However, site LB1, located to the north of GB2, does show a reversion of forest to forest-tundra in the past 1500 years in response to cooling temperatures (Gajewski *et al.*, 1993). There is a difference in the absolute values of the temperature changes prior to 5700 BP where 87-028-069 shows a rise in August SST from 0°C to 6.5°C. Over the same period, GB2's temperature increases by approximately one degree.

The terrestrial sequences from southern Labrador (Lake Hope Simpson and Eagle Lake) cover a longer time period and show cool temperatures prior to 7000 yr BP, maximum warmth between 7000 and 6000 yr BP and cooling in the past 5000 yr (Figure 2.9). The timing of maximum temperatures is different for the three terrestrial sites. Gravel Ridge, the northernmost site, shows maximum temperatures between 4600 and 1800 yr BP. Eagle Lake shows maximum temperatures at 5600 yr BP and decreasing temperatures to present. Lake Hope Simpson, the most southern Labrador site, shows maximum temperatures at 6400 yr BP and decreasing temperatures to present day. Thus, the timing of maximum temperature at the terrestrial sites occurs first in the south and progressively later towards the north. Large dissimilarity values occur for the terrestrial pollen spectra older than 6000 yr BP and they are associated with tundra pollen assemblages, while the temperature maxima are associated with a peak in *Abies* (fir) pollen. The Labrador marine sequence, core HU91-045-006, which is latitudinally between Gravel Ridge and Eagle Lake, shows maximum temperatures at approximately 6200 yr BP (Figure 2.9). The large fluctuations after 7400 yr BP in the marine core may be due to variations in meltwater discharge from the residual ice sheet. These results suggest a time transgressive period of maximum warmth at the terrestrial sites. Alternatively, because of uncertainty in the chronologies of the sequences in this region, a more simultaneous registration

of maximum temperatures around 5000 yr BP could be interpreted. However, since these sites lie 'downwind' of the residual ice sheet, a time transgressive mid-Holocene warming seems more reasonable.

Sites in the St Lawrence region show consistent and stable temperatures for the past 7000 yr BP but have large dissimilarities in assemblages older than 8000 yr BP (Figure 2.10). Indeed, some of the spectra would be considered as "non-analog", that is, there are no modern pollen spectra that are very similar to the fossil spectra (Overpeck *et al.*, 1985). Reasons for this non-analog situation are discussed below. Generally, these older assemblages reconstruct cooler temperatures, whereas there is little change in the past 7000 yr BP. The SST reconstruction from dinocysts in marine core 89-007-111 shows a similar trend (Figure 2.10). Sea surface temperatures recorded by 89-007-111 have been relatively stable since 8000 yr BP.

2.4 Discussion

We have shown using Principal Components Analysis (PCA) that pollen spectra and dinocyst assemblages are correlated on a regional scale consistent with forcing by a common climate signal. The same clusters of taxa - boreal, tundra and southern - are found on the loadings of the first two components of all analyses and indicate a large-scale spatial/temporal signal brought out by the PCA, which is easier to see than by using the pollen diagrams alone. These loadings may be interpreted as being due to climatic forcing on the taxa. The common signal is interpretable in both time and space using the component scores. This signal in the pollen data was found in all regions, and in both the terrestrial and marine environments.

There are some differences between the three PCAs that can be explained by the relative transport ability of the pollen taxa. Most pollen is deposited within a short distance from its source, usually 10-50 km; however, wind-pollinated taxa may undergo significant long-distance transport (e.g. Lynch, 1996; Gajewski, 1995; Prentice, 1986; Prentice, 1985). For example, *Pinus* is over-represented in the marine records, particularly those in Hudson Bay. The overrepresentation of *Pinus* is due to its high pollen productivity and dispersal. In particular, jack pine (*Pinus banksiana* Lamb.) is more abundant in central Canada compared to the Québec-Labrador Peninsula (Richard, 1996) and this may contribute to the abundance of *Pinus* in the Hudson Bay marine spectra. Studies in the Arctic Archipelago show that pine is

significantly over represented in terrestrial assemblages when the local pollen production is reduced (Gajewski, 1995). In the lake cores used in this study, *Pinus* pollen was less important due to the relatively greater contributions of surrounding vegetation to the pollen record.

A second factor causing the differences among these analyses is that the pollen that reaches a marine site undergoes different transport processes. First, transport by wind brings pollen from source areas to the marine continental margin. Atmospheric pollen content over the marine environment varies by season, by pollen productivity, by distance from source area, and by wind direction and strength (Calleja *et al.*, 1993). Once the pollen enters the ocean it can be transported for large distances by ocean currents, gyres or offshore turbidity currents. Owing to the increased density of the transporting medium, pollen that can be transported for long distances within the atmosphere can be transported for even longer distances in the ocean. However, despite such complicating factors, recent paleoenvironmental studies using marine pollen sequences indicate that pollen spectral distortion in marine cores from current and estuarine effects are minimal (Ning and Dupont, 1997; Hooghiemstra, 1992; Heusser and Morley, 1990).

On the time and spatial scales considered in this study, the PCAs illustrate that the marine and terrestrial pollen sequences covaried in a manner that allowed the common pollen signal in the two environments to be used in the development of marine chronologies consistent with terrestrial sites. Absolute dating of marine cores from shelf environments is more difficult than for terrestrial (lake-sediment) cores due to the low amounts of biogenic remains such as carboniferous foraminifers in the sediments. PCA shows that the marine pollen in this study was reflecting the changes in continental vegetation at the spatial and temporal scales under study. Indirectly, the PCAs indicate that marine and terrestrial pollen sequences show similar responses to major changes in the regional vegetation through the Holocene. Demonstrating this allowed the agreement between the marine and terrestrial pollen assemblages to be used in developing a common marine-terrestrial age model and ultimately the intercomparison of the paleoclimate signals in the two environments.

The climate reconstructions exhibit broad scale consistency within any one region. Terrestrial and nearby marine climate reconstructions are also broadly coherent by region, although they differ in some detail. Typically, climate reconstructions tend to show more high frequency variations in the lower sections of the core, where the dissimilarity coefficients are also larger. Increasing dissimilarity down-core is partly due to

the uneven and sparse distribution of modern sites in the north. Such is a particular problem when attempting to reconstruct recent climate history in arctic and subarctic regions or when reconstructing the climate of southern sites whose sedimentary sequence extends to times when boreal and tundra conditions were present. Moreover, many samples in our sites greater than 8000 yr BP are non-analog, using the criteria of Overpeck *et al.* (1985), Anderson *et al.* (1989) and Webb *et al.* (1993), and we are therefore less confident in the climate reconstructions for this time period. These non-analogs prior to 8000 yr BP may be due to lack of modern data from which to choose an analog or to the nature of the vegetation in the late-glacial and early Holocene. Prior to 8000 yr BP, "arctic-subarctic" taxa that are today spread over many degrees of latitude, were located close to the temperate forests at the time. In addition, there was no "boreal forest" prior to 8000 BP anywhere on the continent. Furthermore, pollen concentrations suggest that the density of forest in northern Québec (Richard, 1996) and Labrador (Lamb, 1985) decreased during the late Holocene and are not represented in pollen percentage data.

The three regions had different climate histories associated with their location with respect to deglaciation and air mass boundaries. Prior to 6000 yr BP, an ice sheet persisted in the Québec-Labrador Peninsula and this strongly influenced the climate of the region. Consistent with our results, climate model simulations show an anticyclone over the ice sheet, with resultant northerly flow and cool temperatures in Labrador (COHMAP members, 1988). The time transgressive warming reconstructed for the Labrador terrestrial sites might result from the wasting of this ice sheet; however, dating uncertainty must be considered within this interpretation. As the ice sheet melted, the opening of Hudson Bay affected the Labrador Current (Dyke *et al.*, 1996), thus affecting the dinoflagellates as well. In both Labrador and northwestern Québec, a late Holocene cooling affected sites in the forest-tundra, but is not evident in boreal forest sites, suggesting movements in the polar front. The high frequency changes in the St Lawrence and the cooling reconstructed for the period prior to 8000 yr BP are less reliable due to the larger values of the dissimilarity coefficients.

2.5 Conclusion

This study shows that quantitative climate reconstructions based on marine dinocysts and terrestrial pollen sequences are consistent through the Holocene in northeastern North America. Principal components analysis (PCA) indicates a large-scale climate signal in the dinocysts and pollen. The combined and

separate analyses of marine and nearby terrestrial pollen sequences from Hudson Bay, Labrador and the St. Lawrence estuary differentiate tundra, boreal forest and deciduous forest assemblages in time and space. These analyses indicate that the marine pollen record reflects vegetation changes of the regional terrestrial environment and allows direct correlations between the marine and terrestrial stratigraphies. Sea surface temperatures estimated from dinocysts and terrestrial air temperatures from pollen using the method of modern analogs show that the three regions had differing climate histories associated with their location with respect to deglaciation and air mass boundaries. High frequency climatic changes reconstructed for the St. Lawrence estuary and Gulf, and a cooling reconstructed for the period prior to 8000 yr BP, are less reliable due to the larger values of the dissimilarity coefficients. Prior to 6000 yr BP, cool temperatures reconstructed along the Labrador margins, both in the marine and terrestrial environments are in agreement with climate simulations indicating the persistence of an anticyclone over the Québec-Labrador ice sheet. In both Labrador and northwestern Québec, a late Holocene cooling affects sites in the forest-tundra, but is not evident in boreal forest sites, suggesting movements in the mean position of the polar front. These climate reconstructions provide further evidence that marine and terrestrial climates responded coherently over the Holocene in the different geographic regions considered.

However, there are a number of unanswered questions that this research has brought to light. First, we do not have any firm way to place estimates of uncertainty on the temperature estimates using the method of modern analogs (modern analog technique - MAT). One approach to the non-analog question is to understand which vegetation types are related well to pollen, if some are not then they can cause noise and contribute to dissimilarity down-core. Using pollen types well related to vegetation abundance on the landscape in the pollen set could increase the chances of finding more similarity between modern and fossil pollen and provide better climate reconstructions. However, we need to determine which types of pollen are best related to plant abundance and the spatial scale at which pollen records vegetation in different depositional environments, in particular, there are no studies of this in the boreal regions of North America. Third, the question arises as to whether we can get more climatic information out of the fossil pollen record in North America, perhaps to constrain these paleoclimate reconstructions and assess their reliability. This comparison of terrestrial and marine quantitative climate reconstructions in eastern Canada suggests that the ultimate goal of deriving hemispheric maps of Holocene climates will be attainable as more data are collected and chronological control is maintained.

Figure 2.1: Example of microscope slide prepared from a marine core illustrating preserved fossil dinoflagellate cysts (*O. centrocarpum*, *Spinifirites*), pollen (*Picea*, *Pinus*) and spores. *Photo*: Dr. A. deVernal.

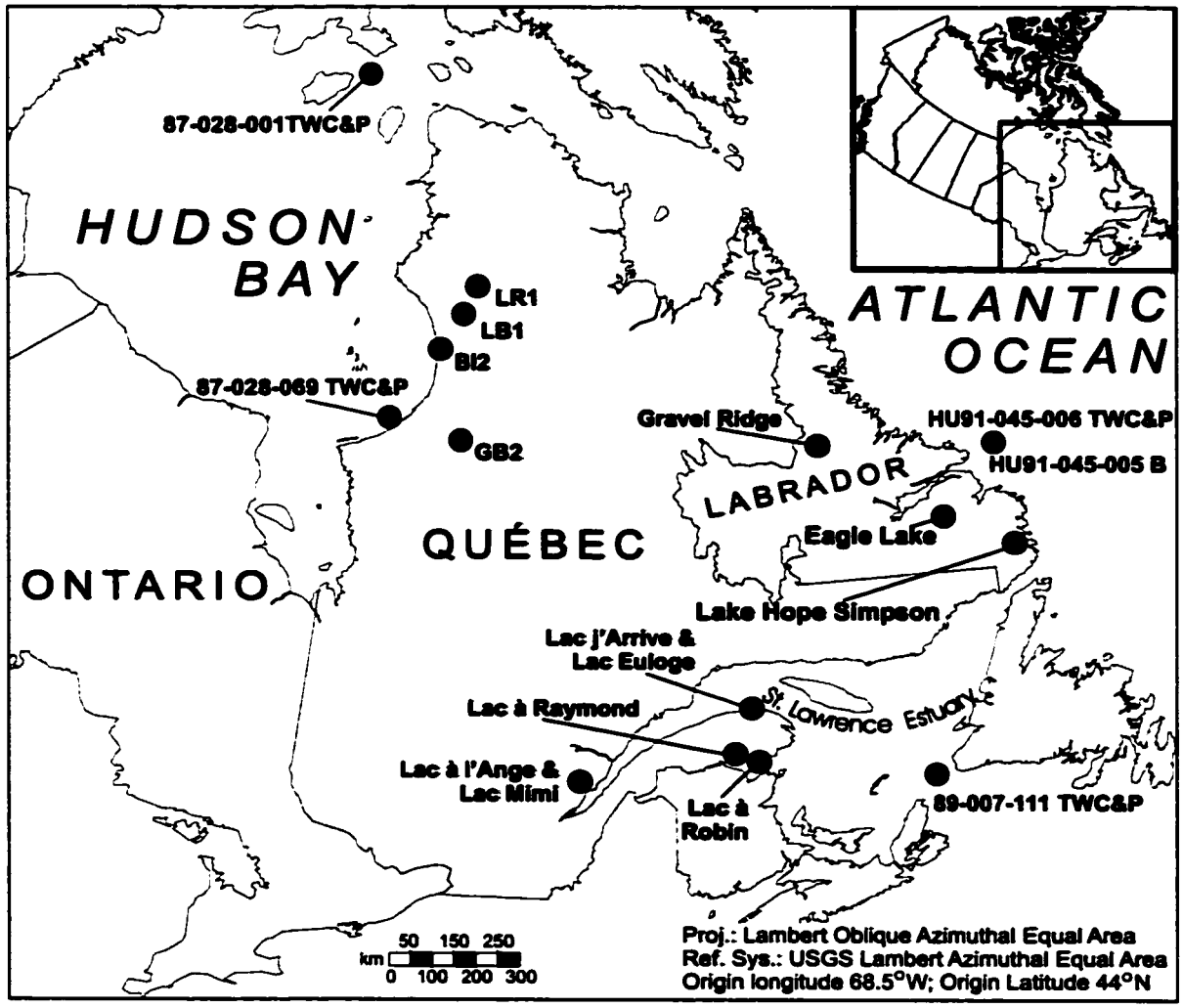


Figure 2.2: Location of sites analyzed in this study, see Table 2.1 for references. *Abbreviations*: TWC – Trigger Weight Core; P – Piston Core; B – Box Core.

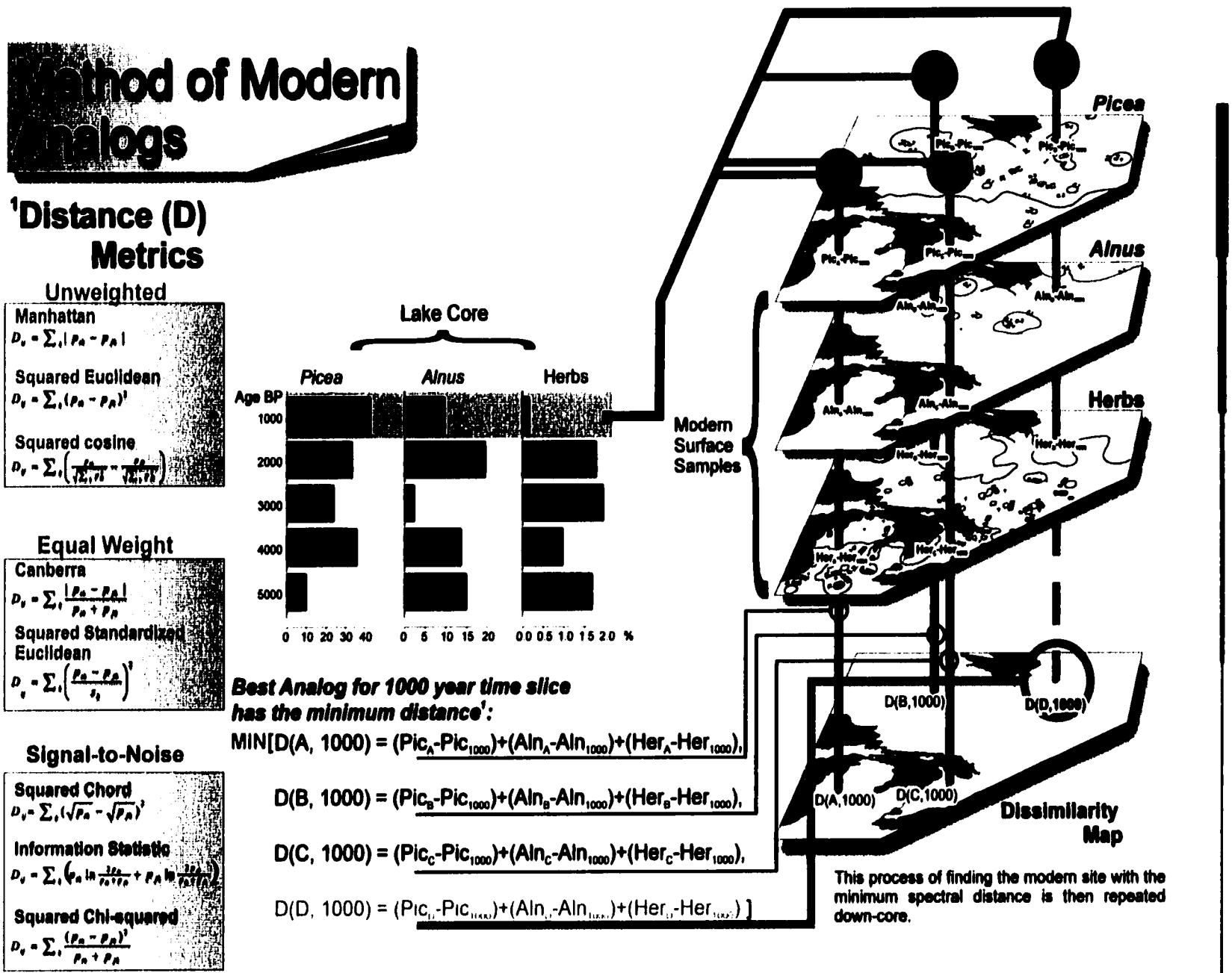


Figure 2.3: The method of modern analogs (modern analog technique – MAT). A fossil pollen spectrum, e.g., at 1000 years ago, is compared with modern spectra from a range of vegetation types and environments using some distance metric, e.g., squared chord distance. The modern pollen spectrum with the minimum distance is considered the best modern analog. From this we infer that the ecosystem around the modern pollen–sampling site is an analog for the past ecosystem cantered on the fossil pollen–sampling site. [1]Distance metrics from Overpeck *et al.* (1985).

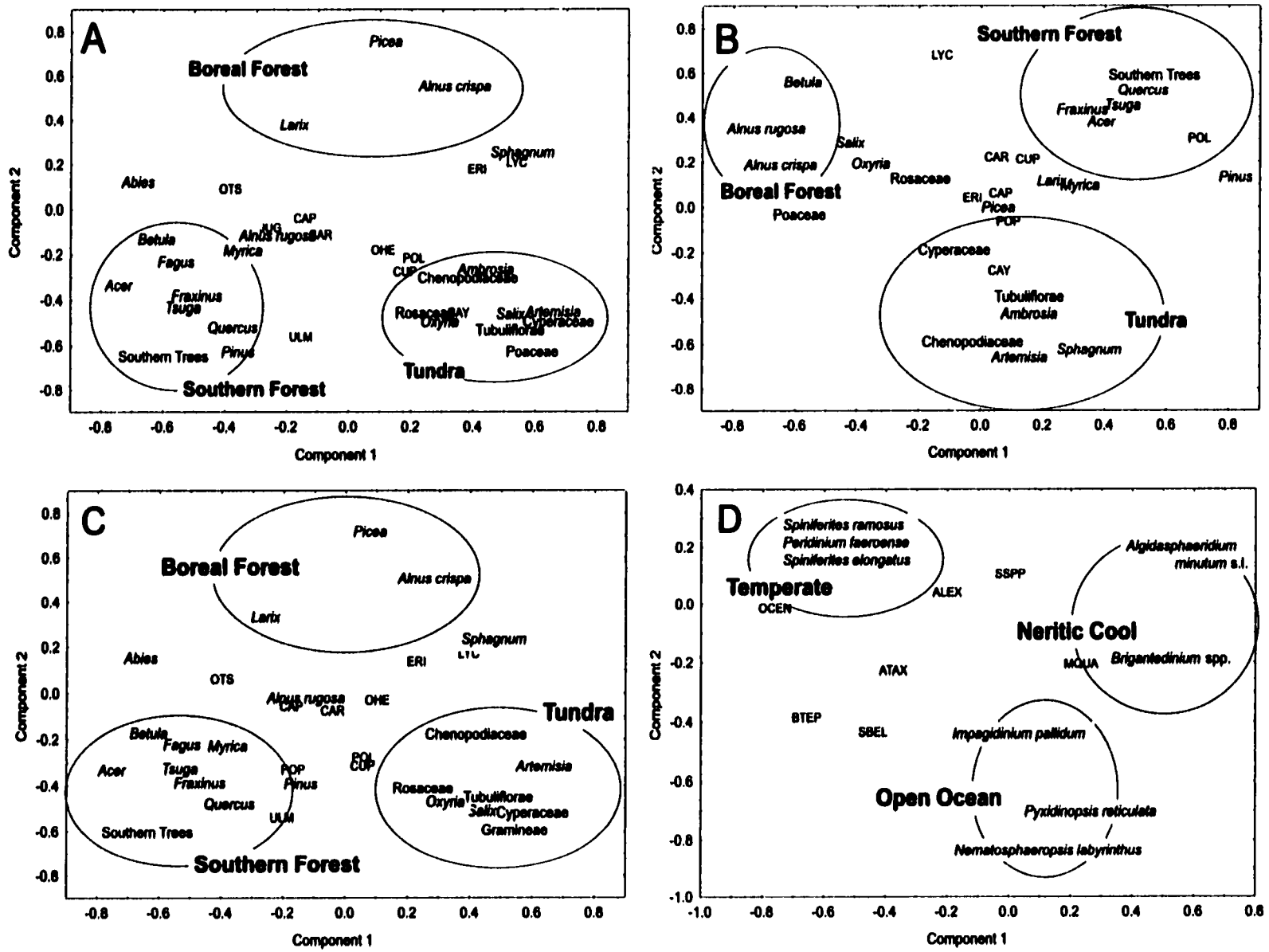


Figure 2.4: Principal component loadings derived from: **A** – Lake-sediment pollen spectra from all terrestrial sites; **B** – pollen from all marine sites; **C** – combined pollen from terrestrial (lake-sediment) and marine cores; **D** – dinoflagellate cysts in marine cores. Abbreviations are as follows: OTS – Other Trees and Shrubs; CAP – *Carpinus*; CAR – *Carya*; CUP – Cupressaceae; CAY – Caryophyllaceae; ERI – Ericaceae; LYC – Lycopodiaceae; POL – Polypodiaceae; OHE – Other Herbs; JUG – *Juglans*; POP – *Populus*; OCEN – *Operculodinium centrocarpum*; ALEX – *Alexandrium excavatum*; ATAX – *Ataxiodinium chaoanum*; BTEP – *Bitectatodinium tepikiense*; SBEL – *Spiniferites belerius*; SSPP – *Spiniferites* sp.; MQUA – *Multispinula quanta*.

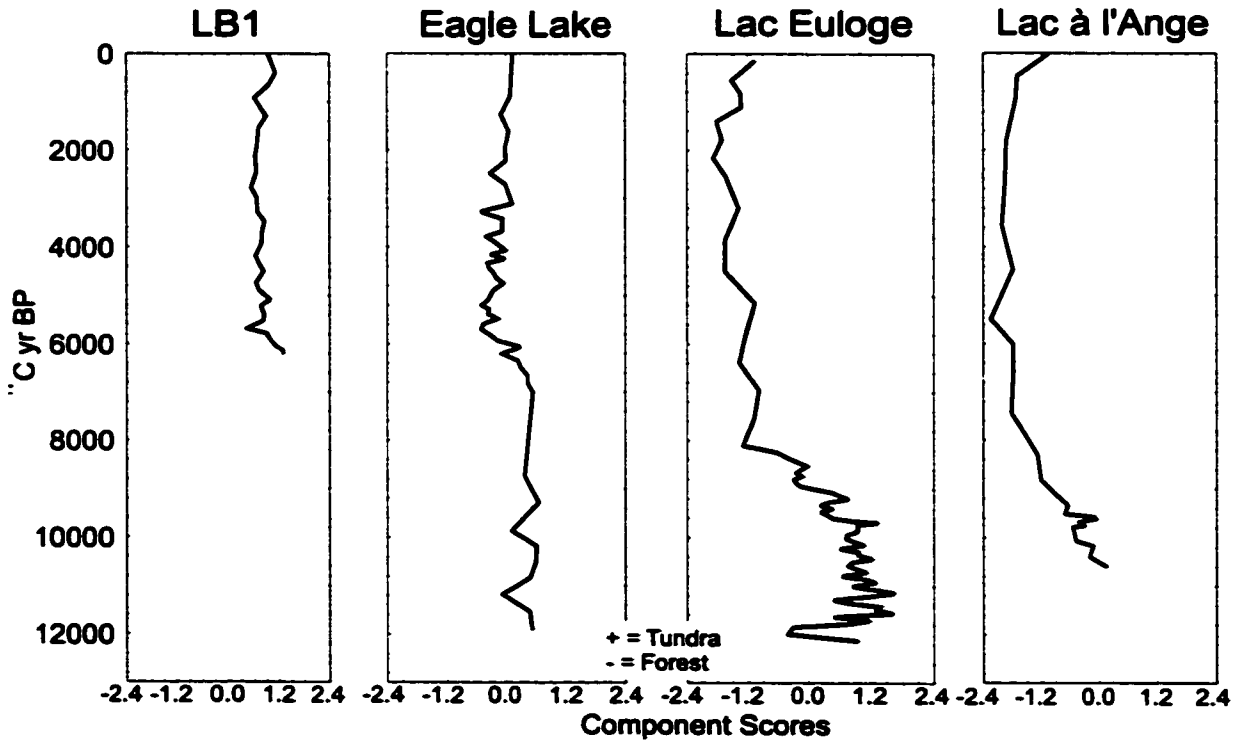


Figure 2.5: Principal component scores from the combined PCA (Figure 2.4c) for several representative lake sequences. Location of sites in Figure 2.2.

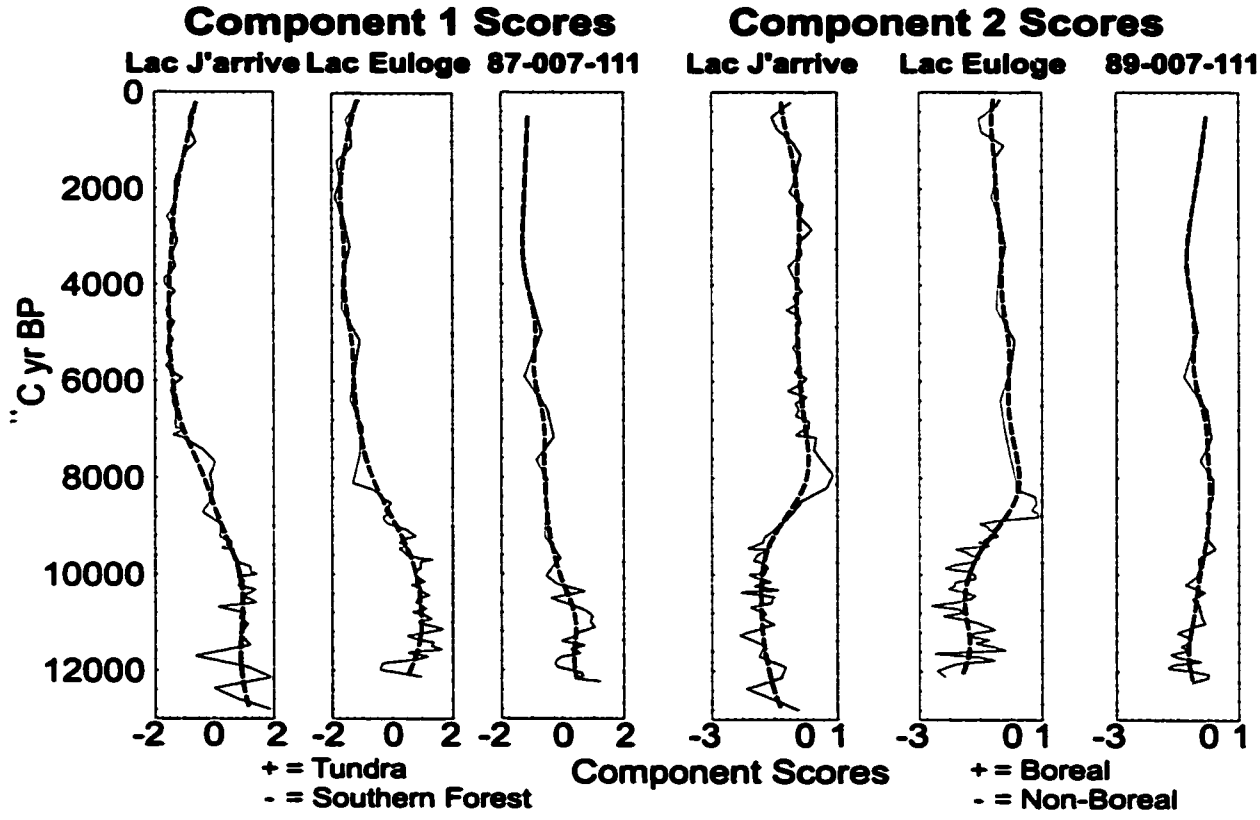


Figure 2.6: Principal component scores for combined marine and lake pollen sequences (Figure 2.4c) in the St. Lawrence region. The fitted curve (dotted-line) is a second order distance–weighted least squares.

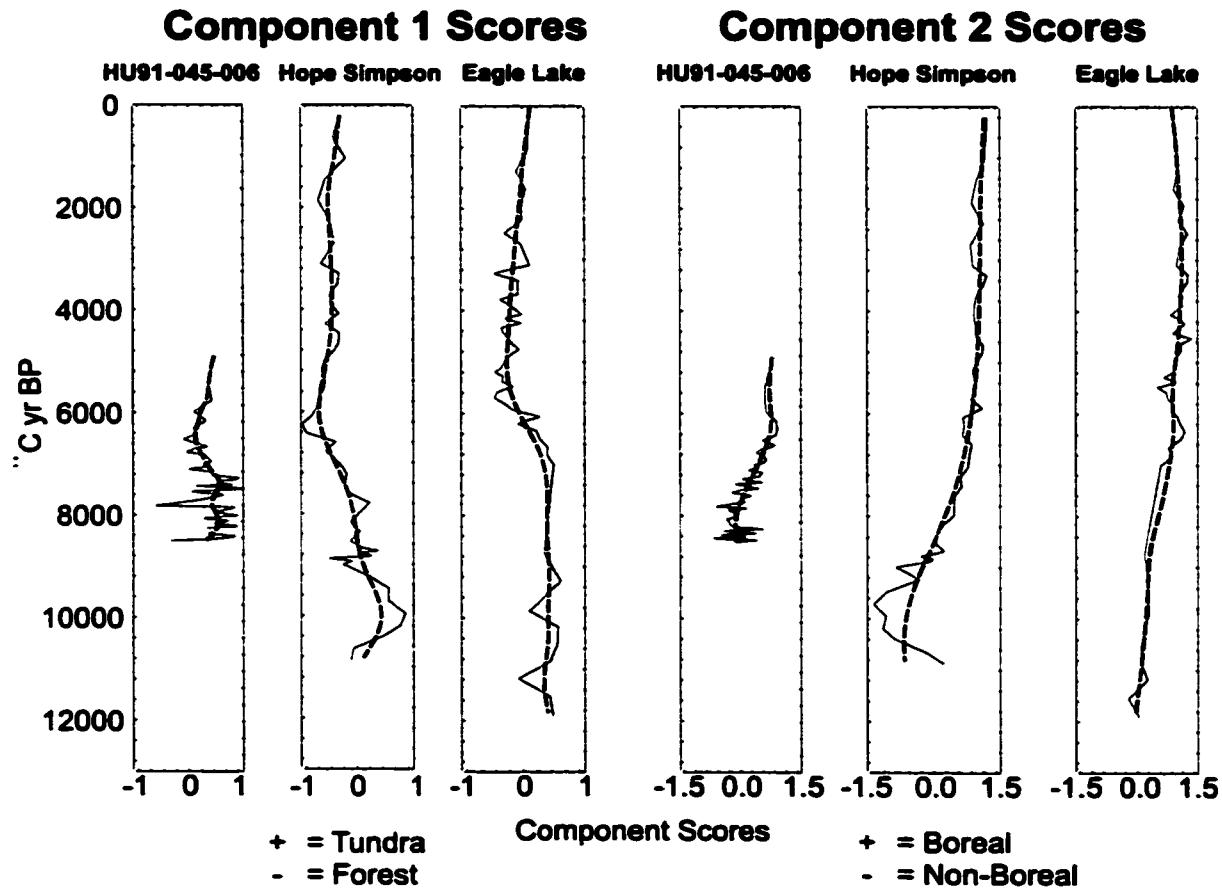


Figure 2.7: Principal component scores for combined marine and lake pollen sequences (Figure 2.4c) in Labrador. *Fitted curve as in Figure 2.6.*

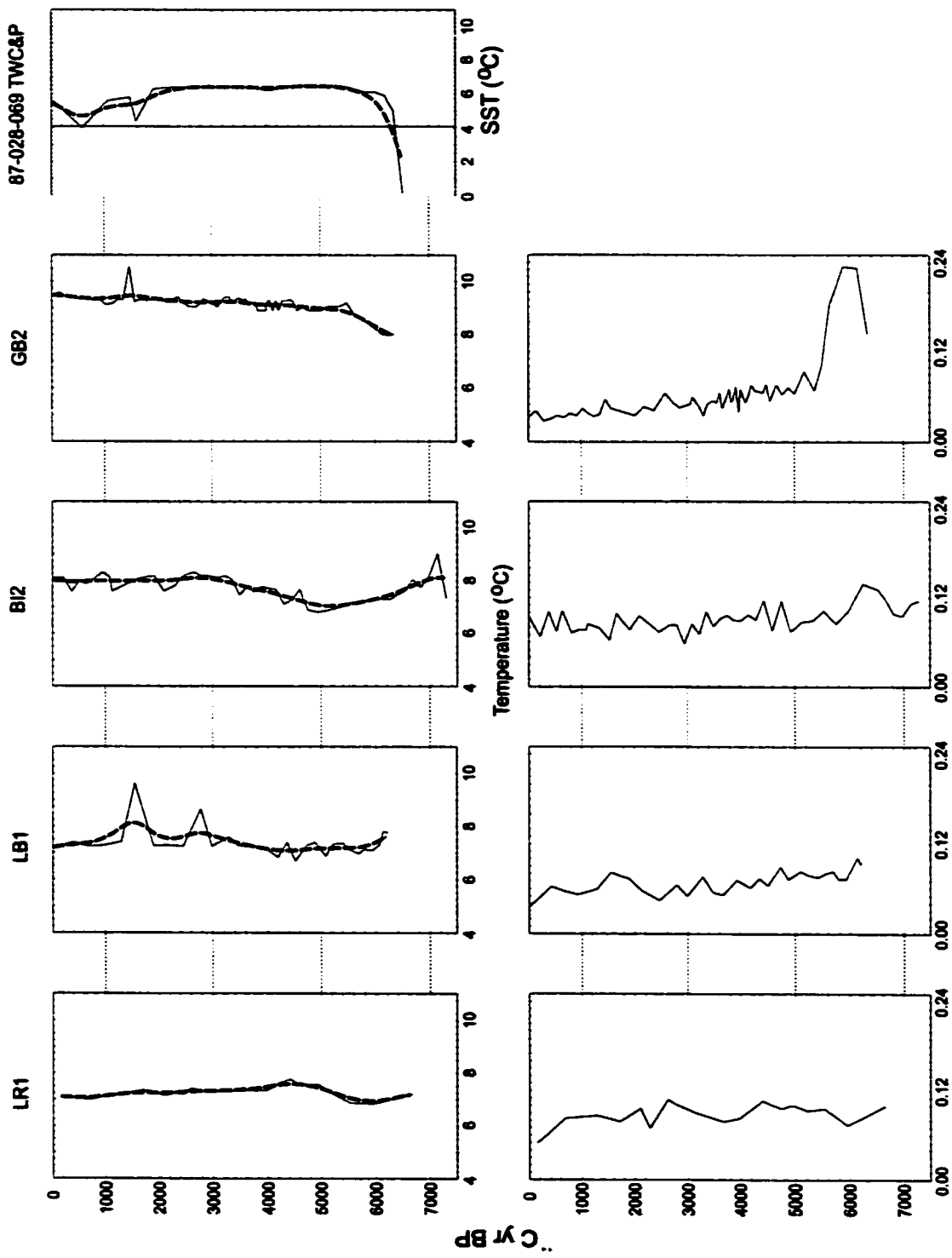


Figure 2.8: Mean August temperature reconstructions and dissimilarity coefficients for terrestrial (lake-sediment) pollen and marine dinocysts in cores from Hudson Bay and northwestern Québec. *Fitted curve as in Figure 2.6.*

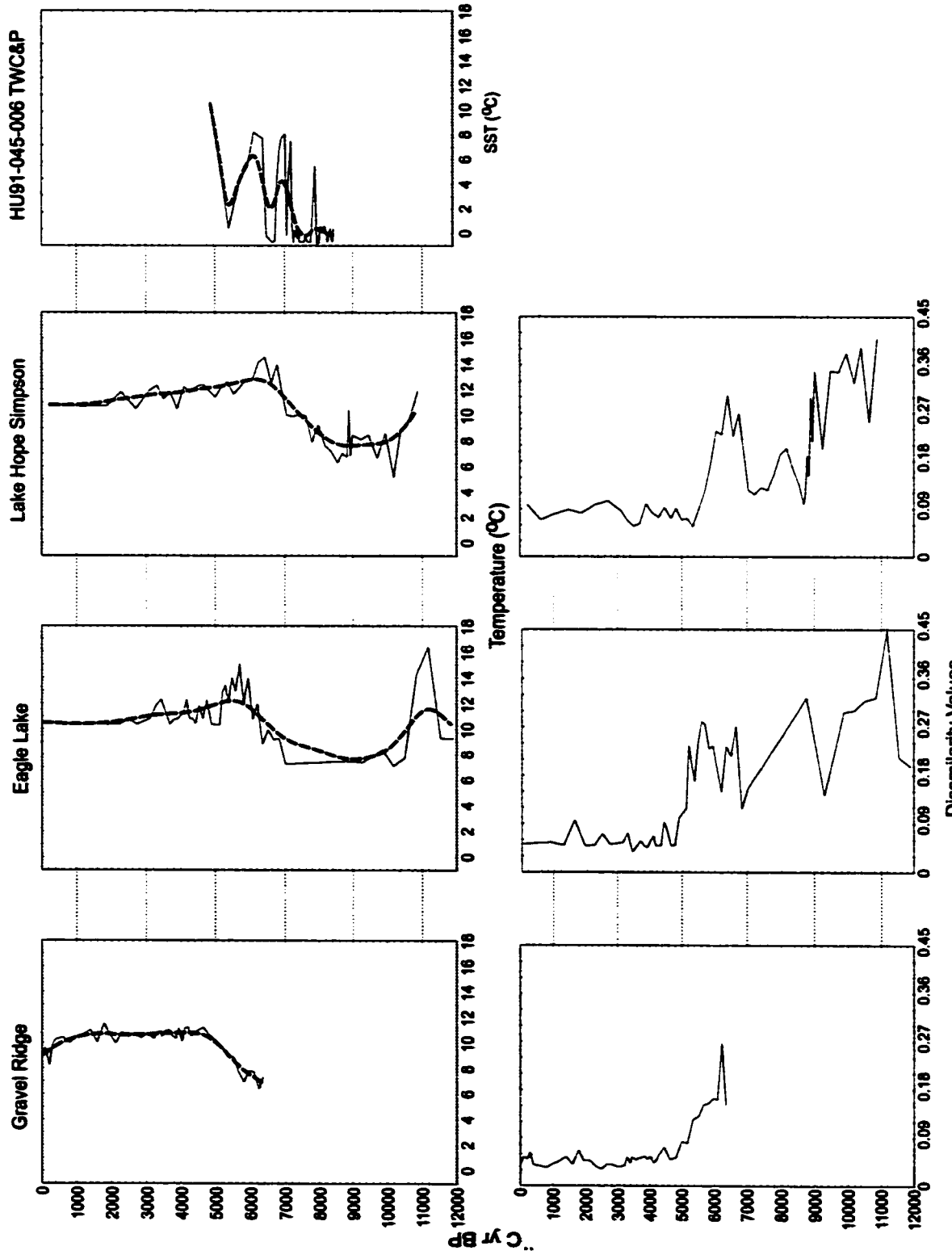


Figure 2.9: Mean August temperature reconstructions and dissimilarity coefficients for terrestrial (lake-sediment) pollen and marine dinocysts in cores from Labrador. *Fitted curve as in Figure 2.6.*

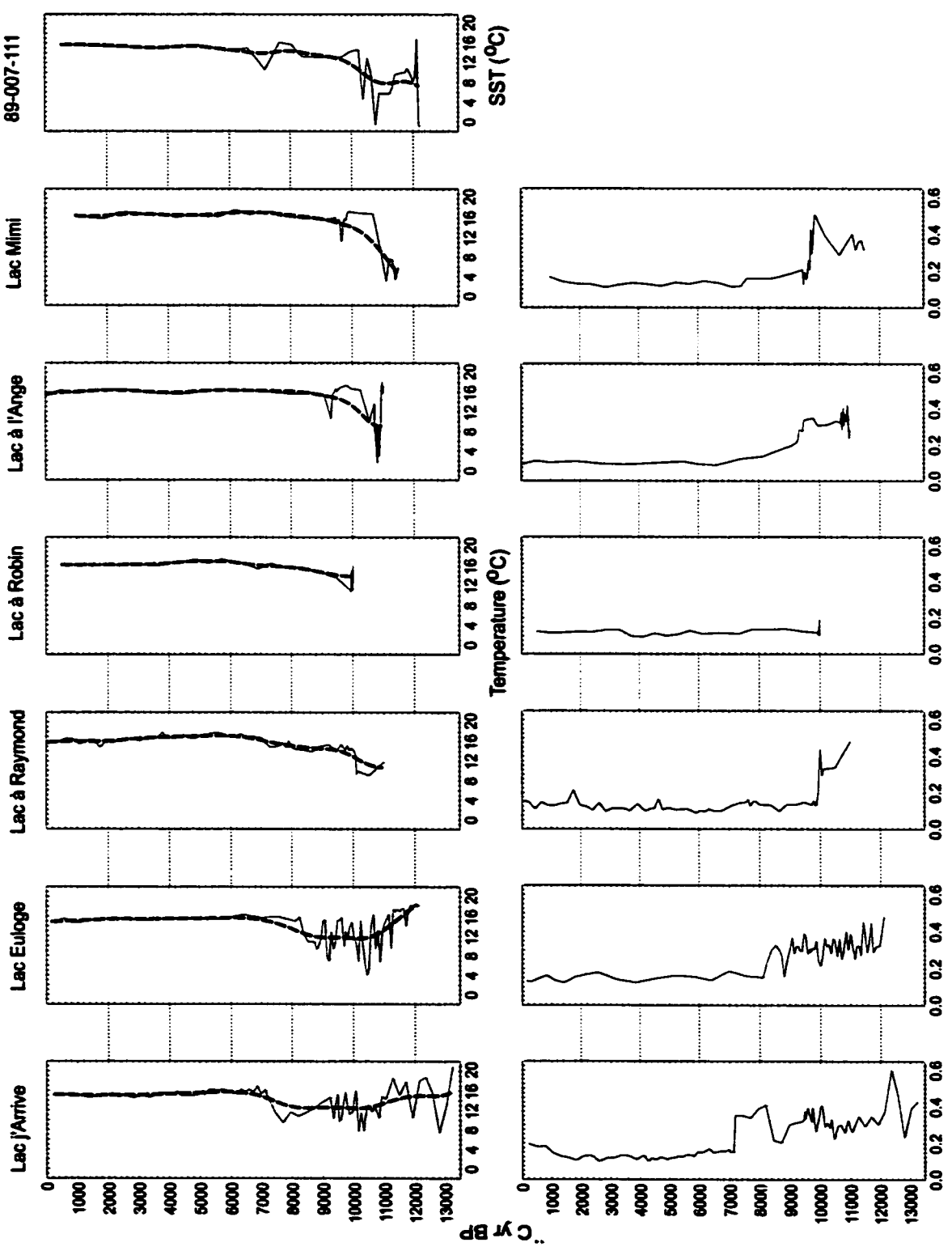


Figure 2.10: Mean August temperature reconstructions and dissimilarity coefficients for terrestrial (lake-sediment) pollen and marine dinocysts in cores from the St. Lawrence region. *Fitted curve as in Figure 2.6.*

Chapter 3

CONTEMPORARY POLLEN-PLANT ABUNDANCE RELATIONS IN SOUTHERN QUÉBEC: POLLEN SOURCE AREA

3.1 INTRODUCTION

An analysis of landscape-level patterns requires data at the appropriate scale, but information of vegetation dynamics at regional (10s to 100s km) scales is difficult to obtain. At the regional scale, pollen diagrams from lake sediments provide appropriate records of vegetation variability through time if sampled at fine temporal intervals. Pollen diagrams may be used to estimate vegetation change if the relation between pollen assemblages and plant abundance is sufficiently understood. This chapter explores the contemporary relations between pollen assemblages and plant abundance on regional scales, as a step toward the goal of calibrating pollen assemblages and reconstructing vegetation dynamics.

Variations through time in the relative proportions of fossil pollen in lake sediments are interpreted as changes in local plant abundance, and networks of fossil pollen are used to infer regional vegetation dynamics during the Holocene (Huntley, 1996; Webb, 1993; Prentice *et al.*, 1991, 1993; Gajewski, 1987; Davis *et al.*, 1986; Webb *et al.*, 1983; Davis, 1976; Davis and Webb, 1975). Plant taxa vary in pollen production and pollen varies in its dispersal ability, so pollen proportions in lake sediments are not directly proportional to plant abundance on the landscape. Interpreting fossil pollen sequences in terms of plant abundance requires determining the relative representation of pollen types and the degree to which these different relations are consistent across different physiognomic regions. Interpreting regional vegetation

dynamics from networks of fossil pollen is aided by understanding the pollen source area or the spatial scale at which a particular pollen taxon represents surrounding plant abundance. This permits pertinent inferences from pollen spectra to be made (Jackson and Lyford, 1999; Jackson and Kearsley, 1998; Calcote, 1995; Jackson, 1990, 1994; Sugita, 1993, 1994; Prentice, 1988, 1986, 1985; Bradshaw and Webb, 1985; Webb *et al.*, 1981; Janssen, 1973) and is a function of the basin size (Jacobson and Bradshaw, 1981).

Our understanding of the pollen-plant abundance relation in forested regions of eastern North America is based on theoretical models of pollen dispersal (Jackson and Lyford, 1999; Sugita, 1993, 1994; Prentice, 1985, 1988; Tauber, 1965), pollen-plant abundance map comparisons (e.g., Delcourt *et al.*, 1984; Webb *et al.*, 1978, 1983; Davis and Webb, 1975; Webb, 1974), pollen rain studies (e.g., Fall 1992), and empirical pollen-plant abundance comparisons (Jackson and Kearsley, 1998; Lynch, 1996; Calcote, 1995; Jackson and Smith, 1994; Jackson and Wong, 1994; Jackson and Dunwiddie, 1992; Jackson, 1990; Prentice *et al.*, 1987; Prentice, 1986; Bradshaw and Webb, 1985; Delcourt *et al.*, 1984; Parsons *et al.*, 1980). However, these latter studies require dense networks of vegetation and pollen observations over large regions which, in North America, are few in number. The majority of regional-scale studies have focused on refining comparison techniques while using the same few datasets in deciduous and mixed deciduous forests of the northeastern United States (e.g., Jackson *et al.*, 1995; Prentice and Webb, 1986; Prentice, 1985; Bradshaw and Webb, 1985; Prentice and Parsons, 1983; Webb *et al.*, 1981). In order to determine the robustness of results, there is a need to extend these studies into the mixed northern hardwood and boreal forests of North America.

An interest in fine-scale ecological processes has recently shifted the focus of many pollen representation studies to smaller spatial scales that utilize small ponds, moss-polsters and forest-hollows (Jackson and Kearsley, 1998; Calcote, 1995; Jackson and Wong, 1994; Jackson, 1990). However, the issues of regional-

scale reconstructions from lake sediments have not been satisfactorily resolved, and these are of relevance to landscape ecology questions.

The comparison of pollen and plant abundance has been approached in two primary ways and for two reasons. First, within a predefined area (e.g., defined by a 30 km radius) around a set of modern pollen sample sites, plant abundances are sampled and compared to pollen proportions in bivariate scatterplots from which correlation, linear regression (OLS-y) and/or extended R-value analysis (ERV) coefficients are derived in order to describe the average relations (e.g., Bradshaw and Webb, 1985; Webb *et al.*, 1981). OLS-y and/or ERV slope terms indicate how small changes in plant abundance affect changes in pollen proportions and illustrate the degree to which pollen are over- or under represented with respect to the surrounding plant abundance. Representation is primarily an issue of pollen production and dispersal: some arboreal taxa produce more pollen than others and/or are better dispersed. The intercept terms indicate the amount of regional background pollen present. Webb *et al.* (1978) discuss the many variables involved in the relative representation of pollen. Some pollen may have no significant relations to plant abundance at the imposed scale of study.

Plant abundance sampled at a fixed radius can only uncover pollen-plant abundance relations at the imposed scale of sampling, negating the possibility of finding stronger relations at other sampling radii. Theoretically, because of differences in the dispersal characteristics among pollen types, the source area from which pollen arrives in a collecting basin differs among taxa (Sugita, 1993, 1994; Prentice, 1985, 1986, 1988; Webb *et al.*, 1981; Faegri and Iversen, 1975; Andersen, 1973). The second type of study attempts to determine the provenance or pollen source area of each pollen type by testing pollen proportions against plant abundances at successively larger radii around the pollen sample (e.g., Jackson and Kearsley, 1998; Calcote, 1995; Prentice *et al.*, 1987). For a given taxon, the radius at which plant abundance is most linearly related to pollen proportions defines pollen source area and the trees within this pollen source area contribute most strongly to changes in the pollen spectra (Jackson, 1990).

In both simulation (Sugita, 1994) and empirical (Jackson and Kearsley, 1998; Calcote, 1995) studies, the explained variance (R^2) between pollen and distance-weighted plant abundance exhibits asymptotic behaviour after reaching a certain plant abundance sampling radius. The distance at which R^2 or likelihood function scores reach an asymptote is the 'relevant' pollen source area (Calcote 1995; Sugita, 1994), beyond which no further information is gained on the pollen-plant abundance relation. Alternatively, Prentice (1985) defined the area surrounding a pollen sample site from which 70% of the pollen originates as the 'effective' pollen source area (Jackson, 1990; Prentice, 1985, 1988). The '70% radius' was intended as a "convenient indicator of the major source area" (Prentice, 1988:27) that accounted for "some reasonably large fraction of pollen input of a particular taxon" (Prentice *et al.*, 1987:52). Simulations suggest that the radius from which 30-45% of the pollen originates contains the vegetation that significantly affects variations in the pollen spectra (Sugita, 1994). The concept of 'relevant' pollen source area is less ambiguous but requires that distance-weighted plant abundance be used.

Pollen proportions should be a distance-weighted function of plant abundance because for a given taxon, trees close to a lake should contribute more pollen than distant trees (Jackson and Lyford, 1999; Sugita, 1993, 1994; Prentice, 1985, 1988). Many studies have not employed distance weighted functions of plant abundance (Jackson and Wong, 1994; Prentice *et al.*, 1987; Bradshaw and Webb, 1985; , Parsons and Prentice, 1981; Parsons *et al.*, 1980) or pollen proportions have been viewed arbitrarily as an isotropic inverse-squared distance function of plant proportions (Prentice and Webb, 1986; Prentice and Parsons, 1983; Webb *et al.*, 1981). Areal weighted averages have also been used (Jackson, 1990). Determining which form of distance weighted plant abundance best approximates the 'pollen-samples view' (Prentice and Webb, 1986) is a critical challenge (Jackson and Lyford, 1999; Jackson and Kearsley, 1998). Unweighted, inverse-squared distance and dispersal weighted methods have been compared at the forest hollow scale (Calcote, 1995). Dispersal weighting is less arbitrary than inverse-distance methods: it provides taxon-specific distance weights based on theoretical models of atmospheric pollen dispersal

(Jackson and Lyford, 1999; Sugita, 1993, 1994; Prentice, 1985, 1988) and should better approximate the pollen samples view of the surrounding vegetation. Nevertheless, at the forest-hollow/moss-polster scale, dispersal weighted and inverse-squared distance weighted plant abundances make little difference (Jackson and Kearsley, 1998; Calcote 1995). At small spatial scales, Jackson and Kearsley (1998) advocate the application of various weighting schemes in order to place boundaries on estimates of pollen source area. There is an apparent need to compare different distance-weightings at larger scales to see how they affect estimates of pollen productivity, background input and pollen source area.

The size of the pollen-collecting basin (e.g., lake, moss polster) can influence pollen source area estimates (Webb *et al.*, 1981). Larger basins integrate pollen from greater areas (Jacobson and Bradshaw, 1981; Prentice, 1985). Locally abundant taxa are better represented using smaller basins and basin size can be selected for the study of ecological processes at a given spatial scale (Calcote, 1995; Jackson, 1990; Bradshaw and Webb, 1985). Consequently, basins of similar size are best for determining pollen-plant abundance relations (Sugita, 1994; Prentice, 1985). Basin size can be incorporated into pollen dispersal models (Prentice, 1985, 1988) which in turn may be used in the distance weighting of plant abundance. Arbitrarily applied inverse-squared distance weighting of plant abundance may not be realistic when the sizes of collecting basins are heterogeneous nor does inverse-squared distance weighting consider the varying nature of dispersal among pollen types.

The effect of vegetation heterogeneity on pollen source area estimates is unclear (Jackson and Kearsley, 1998; Jackson and Wong, 1994; Sugita, 1994). Simulations indicate that pollen productivity and background pollen estimates can be in error if the size of the pollen collecting basin is larger than the average patch size of the plant taxa of interest on the landscape, in which case vegetation may appear homogeneous (Sugita, 1994). Testing of such a hypothesis requires vegetation samples resolved at a spatial scale less than that of the collecting basin but no regional datasets exist with the required resolution.

However, the importance of vegetation heterogeneity surrounding pollen-collecting basins and its effect on relative representation and pollen source area can be explored using simulation.

Thus, the lack of appropriate vegetation data for comparison with pollen from different physiognomic regions and unresolved methodological issues still impede our understanding of the regional scale pollen-plant abundance relations. The primary objectives in this chapter are: a) to determine which of the arboreal taxa in southern Québec exhibit significant linear relations to plant abundance at the regional scale; b) for those taxa with significant linear relations to provide estimates of the relevant pollen source areas, pollen productivity and background pollen input; c) to determine how relevant pollen source area and pollen productivity vary as a function of different distance weighted plant abundances; d) to compare estimates of background pollen input, pollen productivity and relevant source areas to similar studies in the Midwestern United States in order to test the hypothesis of Bradshaw and Webb (1985) that the slope parameters, estimates of relative pollen productivity, are recording differences in the production and dispersal independently of the relative plant abundances of taxa in different physiognomic regions; and e) to determine how within-site vegetation heterogeneity influences the definition of 'relevant' pollen source area under the different plant abundance weighting schemes. This study also uses crown area as a measure of plant abundance, an improved index of plant abundance with relevance to pollen production.

3.2 Study Area

The present-day vegetation of southern Québec (Figure 3.1) consists of northern hardwood forest (Great Lakes Formation) in the south and west and boreal forest (and Taiga) to the north and east. *Picea glauca* (Moench) Voss., *P. mariana* (Mill.) BSP., *Pinus banksiana* Lamb., *Larix laricina* (DuRoi) K. Koch and *Abies balsamea* (L.) Mill. dominate the boreal forest. The dominant conifers of the northern hardwood forest are *Pinus strobus* L., *P. resinosa* Ait. and *Tsuga canadensis* (L.) Carr. and important broadleafs include *Betula lutea* Michx. f., *Acer saccharum* Marsh., *A. rubrum* L., *Tilia americana* L., *Quercus borealis* Michx. f., and *Ulmus americana* L.. Boreal species are of a lesser importance in the northern

hardwood forest but include *Picea mariana*, *P. glauca*, *P. rubens* Sarg., *Pinus banksiana*, *Betula papyrifera* Marsh., *Populus tremuloides* Michx., and *P. balsamifera* L.. The St. Lawrence lowlands have been modified since European settlement through forest harvesting and agricultural clearance. The history of vegetation change in this region of southern Québec has been detailed by a number of authors (e.g., Webb *et al.*, 1983; Webb *et al.*, 1978; Richard, 1977) and in general within eastern North America by Webb *et al.* (1993), Jackson *et al.* (1997) and Delcourt *et al.* (1984). Regional vegetation has been inferred from pollen by Webb *et al.* (1978, 1983) and mapped patterns of vegetation zonation and modern pollen have been compared (Webb *et al.*, 1978; Davis and Webb, 1975).

3.3 DATA AND METHODS

3.3.1 POLLEN AND VEGETATION DATA

Pollen-plant abundance comparison requires a collection of modern pollen spectra and corresponding forest spectra for each pollen site (Jackson, 1994; Sugita, 1994; Prentice *et al.*, 1987). For comparison with pollen, the best measure of plant abundance is pollen source strength (number of grains produced per unit area per unit time). Crown cover (area) is a sufficient proxy for source strength (Jackson, 1994), however most studies have used basal area (Jackson and Kearsley, 1998; Calcote, 1995; Jackson and Wong, 1994; Jackson, 1990; Bradshaw and Webb, 1985) or derivatives (e.g., Growing Stock Volume) based on Canadian/United States forest inventories (Delcourt *et al.*, 1984; Webb *et al.*, 1978; Webb, 1974) to estimate tree abundance. Crown cover is used in this study for the first time at a regional scale in North America.

Crown cover observations were collected over a 6 year period by the Gouvernment du Québec Ministère des Ressources Naturelles - Forêt Québec at approximately 4254 sampling locations (Saucier *et al.*, 1998) (Figure 3.2). A Thiessen tessellation (Boots 1986) shows that 87% of samples represent areas less than 200 km^2. At each sample location there are between 4-6 circular quadrats of 400 m^2 ranging from distances of 0 to 1.5 km which yield 24452 quadrats in the region. For each quadrat, vertically projected absolute

crown cover for all arboreal taxa was measured at two strata, the top strata was used in this study. Relative crown cover is given for each tree species in each quadrat as <1%, 1-5%, 6-25%, 26-40%, 41-60%, 61-80%, or 81-100%, of a 400 m^2 disk that intersects the ground surface. Overlap among crowns in a given quadrat is allowed so that total crown cover in a quadrat can exceed 400 m^2. Using the midpoint of each cover class, the proportional cover was converted to absolute areal crown cover (m^2) per quadrat for each taxon.

The Brown University Modern Pollen Database (Avizinis and Webb, 1985) provided 186 modern pollen samples from the southern Québec region (Figure 3.2). Of these, 114 are bog samples, 16 are from pitcher plants and moss polsters and 57 are from lakes in the southern Québec region. The majority of this study concentrates on lakes because their pollen-collecting properties are of concern. For the lakes, two duplicate samples were averaged and nine pollen sites without sufficient vegetation samples within 20 km were deleted, leaving 46 lakes for the main part of this work (Table 3.1, Figure 3.1).

The median lake radius is 131 m and sizes range between 0.6-730 ha, with 98% between 0.6-200 ha in size and 80% between 2-110 ha. These lake sizes and the scale of study are similar to that of Prentice *et al.* (1987). First-order nearest-neighbor vegetation quadrats are found from 86 m to 13 km from lake sites with a mean distance of 4.1 ± 3.3 km and second-order nearest neighbors have a mean distance of 4.4 ± 3.3 km. Scatterplots of mean plant-abundance proportions within 100 km of each pollen sample site stratified by lake size suggest that variability in lake size does not produce significant biases or outliers in the overall pollen-plant abundance relations (Figure 3.3).

The 17 most common arboreal taxa in the Boreal and mixed Northern Hardwood forests were used in the pollen and plant abundance sets. However, mean plant percentages of *Larix*, *Thuja-Juniperus*, *Fraxinus*, *Tilia*, *Ostrya-Carpinus*, *Ulmus*, *Carya*, and *Juglans* do not exceed 2% within any sampling radius surrounding each modern lake (Figure 3.4).

Table 3.1: ***Pollen sample site descriptions, references and locations corresponding to Figure 3.1.***

SITE #	Site Name	Longitude(W)	Latitude	Area (ha)	Radius* (m)	Elev (m)	Vegetation Zone
109	Yelle Lake	79° 37' 59"	48° 30' 14"	2.7	92	366	Boreal
110	Lac Louis	79° 07' 01"	47° 16' 59"	7.9	158	305	Northern Hardwood
111	NoName	77° 40' 48"	46° 16' 48"	4.9	125	213	Northern Hardwood
112	Pearch Lake	77° 22' 01"	46° 01' 59"	41.7	364	168	Northern Hardwood
113	NoName	77° 39' 00"	50° 55' 01"	4.1	115	274	Boreal
114	NoName	77° 34' 01"	50° 36' 00"	4.0	113	290	Boreal
115	NoName	77° 33' 00"	51° 06' 00"	4.7	122	213	Boreal
116	Lac Ramsay	76° 06' 00"	45° 36' 00"	14.1	212	213	Northern Hardwood
117	NoName	75° 55' 01"	45° 31' 59"	1.8	75	305	Northern Hardwood
118	PinkLake	75° 49' 01"	45° 28' 12"	14.7	217	150	Northern Hardwood
119	NoName	74° 33' 00"	46° 09' 00"	3.4	105	259	Northern Hardwood
120	NoName	74° 31' 01"	46° 18' 00"	0.6	43	366	Northern Hardwood
121	St Agathe	74° 28' 01"	46° 03' 29"	2.0	80	454	Northern Hardwood
122	Aux Quenoilles	74° 24' 00"	46° 10' 19"	2.0	80	403	Northern Hardwood
123	St. Germain	74° 22' 01"	45° 56' 42"	5.0	126	473	Northern Hardwood
124	Gea-1	74° 00' 00"	45° 58' 59"	1.0	56	365	Northern Hardwood
125	St. Calixte	73° 52' 01"	45° 57' 40"	1.7	74	261	Northern Hardwood
126	Tortue	73° 19' 01"	45° 32' 46"	2.5	89	136	Northern Hardwood
127	Atocas	73° 18' 40"	45° 32' 35"	1.0	56	114	Northern Hardwood
128	Romer	73° 19' 59"	45° 58' 01"	6.0	138	18	Northern Hardwood
129	Lac Hertel	73° 09' 00"	45° 31' 59"	43.4	372	170	Northern Hardwood
130	Castor	73° 00' 00"	46° 36' 00"	1.0	56	220	Northern Hardwood
131	Yamaska	72° 52' 01"	45° 28' 01"	2.0	80	265	Northern Hardwood
132	Martini	72° 46' 01"	47° 28' 01"	2.0	80	242	Boreal
133	St Francois de Sales	72° 09' 00"	48° 17' 42"	2.0	80	358	Boreal
134	Lac Marcotte	71° 25' 01"	47° 04' 37"	4.1	115	503	Boreal
135	Lac Joncas	71° 10' 37"	47° 15' 29"	5.5	132	792	Boreal
136	Boundary	70° 40' 48"	45° 34' 12"	24.0	276	603	Northern Hardwood
137	Unknown	70° 37' 59"	45° 37' 12"	7.5	155	489	Northern Hardwood
138	Petit Lac a l'Ange	70° 40' 59"	47° 28' 59"	1.2	61	640	Boreal
139	Lac Therrien	70° 37' 01"	46° 34' 48"	7.4	153	122	Northern Hardwood
140	Lac a Busque	70° 35' 60"	46° 16' 59"	22.1	265	91	Northern Hardwood
141	Dufresne	70° 21' 00"	45° 51' 00"	16.9	232	650	Northern Hardwood
142	Lac Colin	70° 18' 00"	46° 43' 01"	6.9	148	107	Northern Hardwood
143	Lac des Cedres	68° 57' 00"	47° 24' 00"	35.3	335	274	Northern Hardwood
144	Lac Tremblay	68° 55' 59"	48° 52' 59"	7.4	153	122	Boreal
145	Lac Unique	68° 45' 00"	47° 19' 59"	118.8	615	274	Northern Hardwood
146	Lac Beazley	68° 30' 00"	47° 34' 01"	37.3	345	198	Northern Hardwood
147	NoName	68° 28' 01"	49° 21' 00"	6.3	141		Boreal
148	Lac a la Baleine	68° 15' 00"	49° 40' 59"	75.4	490	198	Boreal
149	Lac Fleche	68° 10' 59"	50° 01' 59"	730.4	1525	152	Boreal
150	Lac Jean-Marie	67° 36' 00"	49° 22' 01"	106.0	581	152	Boreal
151	Duffs Lake	66° 46' 01"	47° 58' 01"	7.9	158	91	Northern Hardwood
152	Lac Ste.Anne	66° 01' 59"	48° 48' 00"	141.4	671	381	Boreal
153	Lac Berry	66° 01' 59"	48° 42' 00"	28.3	300	381	Boreal
154	Lac Charles-Cote	65° 57' 00"	48° 58' 01"	19.6	250	991	Boreal

***Lake radius is the average of the semi-major and semi-minor axes measured from 1:50000 NTS topographic sheets.**

Some of these taxa are at their northern limit in the study area (Figure 3.4) and analyses should not be significantly affected by their inclusion in the pollen and plant abundance sets. Although these taxa are

presented and discussed (Figure 3.4), the majority of the results are presented for the most abundant taxa, *Picea*, *Abies*, *Betula*, *Acer*, *Populus*, *Pinus*, *Fagus*, *Quercus* and *Tsuga*, which are most regionally abundant in either the plant abundance or pollen data (Figure 3.4).

3.3.2 POLLEN-PLANT ABUNDANCE COMPARISON METHODS

To compare pollen proportions to plant abundances, crown cover for all 17 arboreal taxa were summed around each pollen sample site within successively increasing 10 km radii from 10 km to 100 km. Within each sampling radius all crown cover observations are totalled and converted to proportions. For each sampling radius and taxon, scatter plots of plant proportions against pollen proportions were made and the explained variance (R^2) and linear regression (OLS-y) slope and intercept coefficients were calculated. All regression analyses were undertaken utilizing the *lm* procedure of SPLUS 5.0 (S-Plus 1998).

Only the OLS-y derived parameters for the estimation of pollen source area and the interpretation of the pollen-plant abundance relations are presented here. However, they are compared with parameters estimated from geometric mean regression (GM) (Webb *et al.*, 1981) and Extended R-Value (ERV) analysis using the FAGERLND program (Jackson *et al.*, 1995) for the overall relation between pollen and plant abundance in order to illustrate that they all provide expected and similar results to OLS-y regression parameters. ERV models are generally preferred over OLS-y or GM regression (Prentice and Webb, 1986; Parsons and Prentice, 1981, Webb *et al.*, 1981; Riggs *et al.*, 1978; Kermack and Haldane, 1950) because the latter do not correct for nonlinearities imposed by compositional data. These nonlinearities, known as the Fagerlind effect in palynology (Prentice, 1985, 1986, 1988; Prentice and Webb, 1986; Prentice and Parsons, 1983; Fagerlind, 1952), can be pronounced when the pollen or plant abundances of a taxon exceed 30-40%, an effect observed in regional (Prentice and Webb, 1986) and local scale studies (Calcote, 1995). ERV models correct for this non-linear relation by making bivariate scatterplots more linear (Jackson *et al.*, 1995; Sugita, 1993; Prentice, 1986; Bradshaw and Webb, 1985; Prentice and Parsons, 1983; Parsons and Prentice, 1981). Ordinary linear regression, however, provides a suitable approximation to ERV parameters in most cases (Prentice and Webb, 1986). In the Québec comparisons, ERV analysis showed

strong similarities to OLS-y derived explained variance, slope and intercept coefficients at all sampling radii for the different weighting plant abundance weighting schemes.

In the present study, *Picea* has the most proportional relation to vegetation at all sampling radii, and therefore this taxon is used as the standard for the ERV analysis in the manner suggested by Parsons and Prentice (1981). The FAGERLND program (Jackson *et al.*, 1995) is used to calculate ERV Model I parameters. The program was run with default values until parameter calculation converged normally and/or until the log-likelihood function was minimized. Negative alpha values were encountered for *Larix;* these were forced to zero and the model was then reinitialized with one less parameter until convergence (Jackson *et al.*, 1995; Prentice and Parsons, 1983). ERV is the maximum likelihood solution for percentages and is thus preferred over OLS-y regression. However, ERV analysis is computationally burdensome and existing programs (Jackson *et al.*, 1995) do not lend themselves to simulation studies.

The derivation of accurate slopes and intercepts from linear regression requires that background pollen inputs be invariant for each taxon over the study area, an assumption that is more likely to be violated if data are collected in regions of varying physiognomy (Sugita, 1994). Multivariate statistical techniques can be applied in order to verify physiognomic variability before pollen and plant abundances are compared (Sugita, 1994). A Principal Components Analysis (PCA) (Johnston, 1980) of 24453 vegetation quadrats using all 17 taxa showed that the first component had Boreal taxa loading negatively and Northern Hardwood taxa loading positively, thus distinguishing the two zones (Figure 3.1b). The component scores were plotted and contoured to see which of the 46 pollen sample sites fell into one of the three sets for analysis: boreal (18 sites), northern hardwood (28 sites) and combined (all 46 sites) (Table 3.1, Figure 3.1a). Using unweighted plant abundance, it was found that the background components, as measured by OLS-y intercept parameters were, when significant and within the ranges of the standard errors derived, effectively identical to the combined dataset and this is illustrated in the **Results** section. The combined

dataset does not significantly violate the assumptions of regression approach with regards to a constant background pollen input.

3.3.3 VEGETATION WEIGHTING

Plant abundance is weighted at each of the 10 sampling radii in 3 different ways. The first is by distance: unweighted relative areal crown cover (*RCC*) is calculated by adding the crown cover (m^2) of all quadrats within a given sampling radius and converting to proportions, inverse distance $RCC(1/d)$ and inverse-squared distance $RCC(1/d^2)$ weights (where d is distance in kilometers from the center of pollen site) are used directly on crown cover (m^2) before conversion to proportions. Because of the nature of inverse-squared distance weighting, any plant abundance sample within 1 km of a basin center was assigned a weight of unity (e.g., Figure 3.6). In the Québec dataset, the number of vegetation quadrats increases proportionally with the areas defined by the increasing sampling radii and thus no corrections were necessary for differences in sample size between sites or absolute areal differences between each radius (Figure 3.5).

Step-wise distance weighting of vegetation (Jackson and Kearsley, 1998) requires absolute crown cover ($m^2\ ha^{-1}$) *ACC* calculated within mutually exclusive rings (at 10 km radial intervals 0-10, 10-20, 20-30,...) around each pollen site. The unweighted sum of all plant abundance samples within each ring is weighted using the distance to the midpoint of each ring and estimates are derived by summing across all rings within a given sampling radius. This can be described by an equation of the form:

$$ACC_i = \frac{(r_2^2 - r_1^2)\sum_{i=1}^{i=n} a_i}{\phi^2 n(r_2 - p/2)^b}, \quad where \ \ r_2 > r_1 \qquad (3.1)$$

where, ACC_i is the unweighted, distance weighted or inverse-squared distance weighted absolute crown cover of taxon i at a given site (m^2), within a vegetation sampling ring between radius r_2 and r_1 (starting at 10 km and increasing at 10 km intervals to 100 km), a_i is the absolute crown cover estimate for taxon a ($m^2\ ha^{-1}$) within the ring, n is the number of vegetation samples within the ring, ϕ is the radius of each

vegetation sample quadrat (11.28 m), p is the minimum vegetation sampling radius, in our case 10 km, b = 0 for unweighted ACC, b = 1 for inverse distance weighted ACC($1/d$), and b = 2 for inverse-squared distance weight vegetation, ACC($1/d^2$), where the distance is the midpoint to each vegetation sampling ring, implicit in the equation. For each pollen sample site, ACC_i was summed across all rings and percentage ACC was based on the sum of all arboreal taxa within a given radius.

A third approach to plant abundance weighting is dispersal weighting. Calcote (1995) used dispersal weighting utilizing the Sutton (1953) model for atmospheric diffusion of particulate from a ground level source. This study utilizes dispersal weighting of relative areal crown cover (RCC) based on Sutton's equations as modified by Prentice (1985) and refers to this as the Sutton-Prentice (SP) weighting model. Because larger basins intercept pollen from larger distances, the Prentice (1985) derivation allows the weight given to a taxon's abundance to increase with increasing basin radius. This equation strictly describes the proportion of pollen that is still airborne at r (Prentice, 1988), the basin radius, which in the present study is unique for each basin according to the radii presented in (Table 3.1) and fall velocities, v_g, of each pollen type (Figure 3.6). The proportion of pollen of a given taxon, i, w_i, reaching the center of the basin of radius r is a function of the distance d of a tree (source) from the basin center and its pollen's fall velocity v_g (m/s):

$$w_i(v_g, r, d) = e^{a(v_g/u)(r^{n/2} - d^{n/2})} \tag{3.2}$$

where, $d > r$, and u (ms^{-1}) is the wind speed above the canopy, held constant at 3 ms^{-1} (Jackson and Lyford, 1999; Calcote, 1995; Prentice, 1985). Thus, w_i, is the weight given to a trees cover value in m^2 at distance d from the basin. The constant $a = 4/n\sqrt{\pi}C_z$ where n is a dimensionless turbulence parameter and C_z is a vertical diffusion coefficient (m$^{1/8}$), both of which are a function of atmospheric stability above the forest canopy (Jackson and Lyford, 1999). Since unstable conditions may better

approximate conditions above the forest canopy (Jackson and Lyford, 1999) dispersal weighting is compared under both neutral ($a = 75.2$ when $n = 0.25$ and $C_z = 0.12$) and unstable ($a = 32.2$ when $n = 0.2$ and $C_z = 0.35$) atmospheric conditions. Representative dispersal weighting curves and inverse-squared distance curves are presented in Figure 3.6.

3.3.4 Vegetation Heterogeneity

The importance of within-site vegetation heterogeneity on parameter estimates under all vegetation weightings was tested using conditional randomization. The geographic coordinates of the n vegetation quadrats within the 100 km radius are permuted 100 times. From each permutation, plant abundance proportions are determined at each sampling radius using RCC under all weighting methods and OLS-y coefficients are derived. This randomization is conditional on the location of each pollen sample site and removes within-site heterogeneity of vegetation without changing the between-site heterogeneity.

3.4 Results

3.4.1 Calibration Model and Vegetation Weighting Effect on Overall Parameter Estimates

RCC, RCC($1/d$) and RCC($1/d^2$) under ERV, OLS-y and GM regression are employed in order to illustrate how different vegetation weightings influence the slope and intercept parameters of different calibration models. Slope and intercept, respectively, are proxy measures of pollen productivity and background pollen.

For OLS-y and ERV under all weightings, intercepts decrease and slopes increase with increasing vegetation sampling radius. GM, however, exhibits a decreasing trend in intercepts and slopes with increasing sampling radius (Figure 3.7). Within any model, unweighted vegetation produces the lowest slopes and intercepts at all sampling radii, whereas inverse-square distance weighted vegetation produces the highest slopes and intercepts at all radii (Figure 3.7). OLS-y, however, shows the least variation in the absolute values of slope parameters among different weightings (Figure 3.7).

Among the different calibration models, GM regression produced the largest slopes (Figure 3.7*abc*) and smallest intercepts (Figure 3.7*def*), a predictable effect (Webb *et al.*, 1981) noted in other studies (Prentice and Webb 1986). OLS-y has the lowest slopes and highest intercepts (Figure 3.7*abc*) under any weighting but the values are similar to ERV and confirm that OLS-y analysis is a suitable approximation (Prentice and Webb, 1986).

Individual taxa follow the trends in the aggregate parameters. The higher slopes and lower intercepts of OLS-y over ERV can be partly accounted for by Fagerlind induced curvature in the scatter plots which is most evident when vegetation or pollen proportions are excessive (over 40%) as observed in regional (Prentice and Webb, 1986) and local scale studies (Calcote, 1995). For example, *Picea* and *Pinus* have ERV alpha parameters consistently greater than the slopes from OLS-y. Prentice and Webb (1986) show that Fagerlind-induced curvature in the scatter plots 'pulls-down' or 'pulls-up' the OLS-y beta parameters. In this study, *Picea* pollen percentages increase proportionally with vegetation percentages but at high percentages (> 40%), *Picea* percentages plateau and pull down OLS-y beta values (Figure 3.8*a-k*). ERV corrected vegetation proportions reduce the Fagerlind induced curvature resulting in more linear scatter, lower intercepts and higher slopes under an OLS-y model (Figure 3.8*g,h,i*). Generally, in the case of underrepresented types like *Acer, Abies, and Populus,* whose ERV alpha parameters are less than those of OLS-y, the opposite is true.

3.4.2 Regional Comparison

The Boreal and Northern Hardwood forest and Combined datasets are compared using unweighted vegetation under OLS-y (Tables 3.2, 3.3 and 3.4).

Table 3.2 ***Ordinary Least Squares Regression estimates of slope parameters and their standard errors for relation between pollen and tree proportions at vegetation-sampling radii from 10-100 km. Only significant slope estimates (Table 3.4) are presented. ALL – combined Boreal and Northern Hardwood dataset (Table 3.1). B – Boreal forest sites only. NH – Northern Hardwood forest sites only. Dashes indicate non-significant coefficients and are not presented.***

Taxon	Zone	10 km	20 km	30 km	40 km	50 km	60 km	70 km	80 km	90 km	100 km
Picea	ALL	0.54 ± 0.08	0.66 ± 0.04	0.70 ± 0.05	0.70 ± 0.05	0.72 ± 0.05	0.77 ± 0.05	0.80 ± 0.05	0.83 ± 0.05	0.86 ± 0.05	0.86 ± 0.06
	B	0.41 ± 0.15	0.63 ± 0.08	0.66 ± 0.10	0.68 ± 0.10	0.72 ± 0.11	0.79 ± 0.10	0.81 ± 0.10	0.84 ± 0.11	0.88 ± 0.11	0.86 ± 0.12
	NH	-	0.97 ± 0.24	1.10 ± 0.23	1.15 ± 0.26	1.20 ± 0.26	1.14 ± 0.25	1.03 ± 0.23	0.97 ± 0.23	0.95 ± 0.23	0.96 ± 0.23
Abies	ALL	0.17 ± 0.03	0.18 ± 0.03	0.22 ± 0.04	0.27 ± 0.04	0.30 ± 0.04	0.34 ± 0.05	0.36 ± 0.05	0.38 ± 0.05	0.38 ± 0.05	0.39 ± 0.05
	B	0.20 ± 0.04	0.19 ± 0.04	0.23 ± 0.04	0.29 ± 0.05	0.33 ± 0.05	0.36 ± 0.06	0.38 ± 0.06	0.38 ± 0.07	0.38 ± 0.08	0.38 ± 0.08
	NH	-	-	-	-	0.36 ± 0.13	0.38 ± 0.12	0.45 ± 0.12	0.46 ± 0.11	0.46 ± 0.10	0.44 ± 0.10
Betula	ALL	0.59 ± 0.18	0.90 ± 0.21	0.92 ± 0.22	1.04 ± 0.25	1.04 ± 0.26	1.05 ± 0.28	1.05 ± 0.28	1.04 ± 0.29	1.03 ± 0.30	1.01 ± 0.31
	B	-	0.67 ± 0.30	-	0.78 ± 0.36	0.79 ± 0.37	-	0.87 ± 0.37	0.89 ± 0.37	0.95 ± 0.37	0.99 ± 0.37
	NH	0.67 ± 0.28	1.28 ± 0.28	1.37 ± 0.29	1.50 ± 0.34	1.53 ± 0.36	1.61 ± 0.40	1.66 ± 0.45	1.66 ± 0.50	1.66 ± 0.55	1.50 ± 0.59
Pinus	ALL	1.90 ± 0.28	1.18 ± 0.37	1.81 ± 0.45	2.60 ± 0.40	2.72 ± 0.42	2.22 ± 0.43	2.04 ± 0.44	1.87 ± 0.43	1.68 ± 0.43	1.74 ± 0.44
	B	1.92 ± 0.75	-	-	2.10 ± 0.78	2.01 ± 0.77	-	-	-	-	-
	NH	1.90 ± 0.26	1.14 ± 0.41	2.29 ± 0.48	3.06 ± 0.43	3.53 ± 0.45	3.51 ± 0.47	3.70 ± 0.55	3.91 ± 0.57	4.14 ± 0.63	4.13 ± 0.70
Acer	ALL	0.05 ± 0.01	0.06 ± 0.01	0.06 ± 0.01	0.06 ± 0.01	0.06 ± 0.01	0.07 ± 0.01	0.07 ± 0.01	0.07 ± 0.01	0.07 ± 0.01	0.07 ± 0.01
	B	0.08 ± 0.03	0.03 ± 0.01	0.03 ± 0.01	0.04 ± 0.01	0.04 ± 0.01	0.04 ± 0.01	0.04 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01
	NH	-	-	-	-	-	-	-	-	-	-
Populus	ALL	0.04 ± 0.01	0.04 ± 0.01	0.04 ± 0.01	0.05 ± 0.01	0.06 ± 0.01	0.06 ± 0.02	0.07 ± 0.02	0.07 ± 0.02	0.07 ± 0.02	0.08 ± 0.02
	B	0.04 ± 0.01	0.05 ± 0.01	0.05 ± 0.01	0.06 ± 0.02	0.07 ± 0.01	0.07 ± 0.02	0.07 ± 0.02	0.09 ± 0.02	0.10 ± 0.03	0.11 ± 0.03
	NH	0.03 ± 0.01	-	-	-	-	-	-	-	-	-
Fagus	ALL	0.23 ± 0.08	0.24 ± 0.09	0.34 ± 0.11	0.44 ± 0.12	0.51 ± 0.14	0.52 ± 0.16	0.63 ± 0.17	0.73 ± 0.17	0.78 ± 0.18	0.78 ± 0.18
	B	-	0.32 ± 0.08	0.21 ± 0.05	0.23 ± 0.06	0.26 ± 0.06	0.23 ± 0.06	0.28 ± 0.07	0.34 ± 0.07	0.39 ± 0.07	0.40 ± 0.08
	NH	-	-	-	-	-	-	-	-	-	-
Quercus	ALL	1.01 ± 0.22	1.02 ± 0.21	1.18 ± 0.23	0.95 ± 0.23	0.85 ± 0.22	1.03 ± 0.24	0.92 ± 0.24	0.99 ± 0.25	1.02 ± 0.27	1.12 ± 0.26
	B	-	-	-	-	-	-	-	-	-	7.89 ± 3.21
	NH	0.93 ± 0.28	0.88 ± 0.28	1.02 ± 0.32	0.77 ± 0.32	0.68 ± 0.30	0.85 ± 0.33	0.72 ± 0.33	0.79 ± 0.36	0.80 ± 0.37	0.92 ± 0.37
Tsuga	ALL	0.79 ± 0.11	0.91 ± 0.15	1.09 ± 0.13	1.04 ± 0.14	1.08 ± 0.16	1.25 ± 0.16	1.29 ± 0.17	1.26 ± 0.16	1.20 ± 0.17	1.32 ± 0.18
	B	-	4.48 ± 1.02	8.90 ± 2.02	4.10 ± 0.98	2.97 ± 0.71	1.05 ± 0.40	-	-	-	-
	NH	0.70 ± 0.12	0.76 ± 0.21	0.98 ± 0.19	0.93 ± 0.20	0.95 ± 0.23	1.16 ± 0.24	1.20 ± 0.26	1.18 ± 0.25	1.09 ± 0.26	1.27 ± 0.28

Picea, *Pinus* and *Betula*, have larger slopes and smaller intercepts in the Northern Hardwood and Combined dataset when compared to the Boreal (Tables 3.2-3.3). Where significant, slopes and intercepts for *Abies* are effectively identical in all regions. Other taxa lack a sufficient number of significant linear

relations for any comparison between the Boreal and Northern Hardwood forests alone, however, when compared individually to the Combined dataset, slopes and intercepts, with the exception of *Tsuga*, are identical. *Tsuga* has larger slopes in the Boreal subset because it is over-represented in the pollen of that

Table 3.3 ***Ordinary Least Squares Regression estimates of intercept parameters and their standard errors for relation between pollen and tree proportions at vegetation-sampling radii from 10-100 km. Only significant slope estimates (Table 3.4) are presented. Zone coding same as Table 3.2.***

Taxon	Zone	10 km	20 km	30 km	40 km	50 km	60 km	70 km	80 km	90 km	100 km
Picea	ALL	0.08 ± 0.02	0.06 ± 0.01	0.05 ± 0.01	0.05 ± 0.01	0.05 ± 0.01	0.04 ± 0.01	0.04 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01
	B	0.17 ± 0.07	0.07 ± 0.04	0.07 ± 0.05	0.06 ± 0.05	0.05 ± 0.05	0.03 ± 0.04	0.03 ± 0.04	0.03 ± 0.04	0.02 ± 0.04	0.03 ± 0.05
	NH	-	0.03 ± 0.02	0.02 ± 0.02	0.02 ± 0.02	0.02 ± 0.02	0.02 ± 0.02	0.02 ± 0.02	0.02 ± 0.02	0.03 ± 0.02	0.02 ± 0.02
Abies	ALL	0.02 ± 0.01	0.02 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01
	B	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.00 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.02	-0.01 ± 0.02
	NH	-	-	-	-	0.00 ± 0.01	0.00 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.01
Betula	ALL	0.35 ± 0.04	0.28 ± 0.05	0.27 ± 0.05	0.25 ± 0.05	0.25 ± 0.06	0.25 ± 0.06	0.25 ± 0.06	0.25 ± 0.06	0.25 ± 0.06	0.25 ± 0.07
	B	-	0.28 ± 0.07	-	0.25 ± 0.08	0.25 ± 0.08	-	0.23 ± 0.09	0.22 ± 0.08	0.20 ± 0.09	0.19 ± 0.09
	NH	0.36 ± 0.06	0.24 ± 0.06	0.21 ± 0.06	0.19 ± 0.07	0.19 ± 0.07	0.18 ± 0.08	0.17 ± 0.09	0.17 ± 0.10	0.17 ± 0.11	0.20 ± 0.12
Pinus	ALL	0.11 ± 0.02	0.13 ± 0.02	0.12 ± 0.02	0.09 ± 0.02	0.08 ± 0.02	0.09 ± 0.02	0.09 ± 0.02	0.10 ± 0.02	0.10 ± 0.02	0.10 ± 0.02
	B	0.12 ± 0.03	-	-	0.09 ± 0.04	0.09 ± 0.04	-	-	-	-	-
	NH	0.11 ± 0.02	0.13 ± 0.03	0.10 ± 0.02	0.08 ± 0.02	0.07 ± 0.02	0.07 ± 0.02	0.06 ± 0.02	0.05 ± 0.02	0.04 ± 0.02	0.04 ± 0.03
Acer	ALL	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
	B	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
	NH	-	-	-	-	-	-	-	-	-	-
Populus	ALL	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
	B	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	-0.01 ± 0.00	-0.01 ± 0.00	-0.01 ± 0.00
	NH	0.00 ± 0.00	-	-	-	-	-	-	-	-	-
Fagus	ALL	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01
	B	-	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
	NH	-	-	-	-	-	-	-	-	-	-
Quercus	ALL	0.02 ± 0.00	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01
	B	-	-	-	-	-	-	-	-	-	0.01 ± 0.00
	NH	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	0.03 ± 0.01
Tsuga	ALL	0.02 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00
	B	-	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	-	-	-	-
	NH	0.02 ± 0.01	0.02 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01

region, and a few points of low vegetation (< 1%) and relatively high pollen (~2.5%) proportions cause higher slope estimates. Within the ranges indicated by the standard error of the estimates, background pollen estimates between regions are identical. This supports the results of the combined analysis since the assumption of constant background pollen seems reasonable.

Table 3.4 ***Explained variance (R^2) between pollen and tree proportions at vegetation-sampling radii from 10-100 km. All other coefficients are significant at the $p < 0.05$. Zone codes same as Table 3.2.***

Taxon	Zone	10 km	20 km	30 km	40 km	50 km	60 km	70 km	80 km	90 km	100 km
Picea	ALL	0.56	0.85	0.82	0.83	0.83	0.86	0.86	0.85	0.85	0.84
	B	0.34	0.79	0.72	0.74	0.73	0.80	0.79	0.78	0.79	0.76
	NH	-	0.39	0.47	0.43	0.46	0.44	0.44	0.40	0.40	0.40
Abies	ALL	0.41	0.40	0.45	0.52	0.54	0.56	0.58	0.57	0.56	0.56
	B	0.67	0.55	0.62	0.73	0.71	0.69	0.69	0.65	0.60	0.59
	NH	-	-	-	-	0.23	0.29	0.36	0.38	0.42	0.43
Betula	ALL	0.21	0.30	0.28	0.28	0.27	0.24	0.24	0.22	0.21	0.19
	B	-	0.24	-	0.23	0.22	-	0.25	0.26	0.30	0.31
	NH	0.20	0.45	0.46	0.43	0.41	0.38	0.35	0.30	0.26	0.20
Pinus	ALL	0.53	0.19	0.27	0.49	0.49	0.37	0.33	0.30	0.26	0.26
	B	0.30	-	-	0.31	0.30	-	-	-	-	-
	NH	0.68	0.23	0.46	0.66	0.71	0.68	0.64	0.65	0.62	0.57
Acer	ALL	0.30	0.40	0.35	0.32	0.34	0.36	0.35	0.35	0.35	0.36
	B	0.28	0.31	0.37	0.44	0.46	0.48	0.46	0.41	0.41	0.41
	NH	-	-	-	-	-	-	-	-	-	-
Populus	ALL	0.35	0.22	0.25	0.23	0.28	0.28	0.26	0.23	0.21	0.22
	B	0.45	0.62	0.53	0.50	0.56	0.52	0.49	0.50	0.49	0.48
	NH	0.20	-	-	-	-	-	-	-	-	-
Fagus	ALL	0.17	0.14	0.18	0.22	0.23	0.20	0.24	0.29	0.31	0.30
	B	-	0.50	0.55	0.51	0.51	0.46	0.54	0.57	0.63	0.62
	NH	-	-	-	-	-	-	-	-	-	-
Quercus	ALL	0.34	0.35	0.36	0.27	0.25	0.30	0.25	0.26	0.25	0.29
	B	-	-	-	-	-	-	-	-	-	0.27
	NH	0.32	0.28	0.28	0.19	0.16	0.20	0.15	0.16	0.15	0.19
Tsuga	ALL	0.57	0.44	0.60	0.56	0.53	0.59	0.57	0.58	0.52	0.56
	B	-	0.55	0.55	0.52	0.52	0.30	-	-	-	
	NH	0.57	0.34	0.50	0.45	0.40	0.47	0.45	0.47	0.40	0.44

Generally, with the exception of the 10 km vegetation-sampling radius, explained variance for all types is constant at all vegetation-sampling radii (Table 3.4).

The highest explained variance is found for *Picea*, *Abies*, *Pinus* and *Tsuga*. In the Boreal forest, where they are more abundant, explained variance is greater for *Picea* and *Abies*. *Pinus* has greater explained variance in the Northern Hardwood forest. *Tsuga* and *Betula* are comparable in both regions. *Acer*, *Fagus* and *Populus* have non-significant relations at most vegetation sampling radii in the Northern Hardwood forest. This is not surprising for *Acer* and *Populus* since they are severely underrepresented in the pollen spectra with slopes and intercepts near zero. *Fagus* on the other hand has no significant relation to plant abundance in the Northern Hardwood forest where it is most abundant. The minor discrepancies between the 10 km sampling radius and the other sampling radii are due to the smaller sample size at the 10 km radius.

3.4.3 Pollen Source Areas

The different vegetation weightings produce only minor changes in the absolute value of R^2 (Figure 3.9*a-i*). No single weighting technique produced clearly better-explained variance estimates for all taxa.

For all plant abundance weightings, explained variance does not exhibit clear asymptotical behavior with increasing sampling radius for *Fagus*, *Acer*, *Populus*, *Quercus*, and *Tsuga* (Figure 3.9c-f, 3.9*i*), and 'relevant' pollen source areas cannot be defined. *Picea* has no improvements in explained variance after 20 km, suggesting a source area less than 20 km. Explained variance in *Abies* improves for all sampling radii when plant abundance is weighted using inverse-distance techniques but improvements are very small. However, unweighted plant abundance shows maximum explained variance between 40 to 80 km (Figure 3.9*b*). An estimate of *Abies* relevant source area is approximately 40 km where the inverse-squared distance weighted curves improve no further. Unweighted *Betula* plant abundances show no improvements in explained variance after 20 km and weighted *Betula* plant abundance exhibits asymptotical behavior at 30 km (Figure 3.9*g*), consequently, the relevant source area for *Betula* is less than or equal to 30 km (Figure

3.9*g*). Unweighted and dispersal weighted curves for *Pinus* (Figure 3.9*h*) are maximal between 40-50 km (also where the lowest intercepts and highest slopes are found (Figure 3.10h and 3.11h)), and suggest a 'relevant' source area under inverse-squared distance weighting between 50 to 60 km (Figure 3.9*h*).

Dispersal-weighted plant abundance shows trends in explained variances that are most similar to unweighted plant abundance. Inverse-squared distance weights either show asymptotical behavior or no substantial variability in explained variance as sampling radius increases. The asymptotes in R^2 using inverse-squared distance weighting are found where the unweighted curves are globally maximal for *Betula* and *Pinus* - *Pinus* at 50 km and *Betula* at 30 km and both are ± 10 km from where the unweighted curves reach a global maximum (Figure 3.9).

The scale of study is similar to that of Bradshaw and Webb (1986) in the United States Midwest and that of Prentice *et al.* (1987) in Sweden and so their source area estimates can be compared to those of this present study (Table 3.5). The source area estimates in Québec are similar to previous studies for those taxa that indicated 'relevant' source area estimates (Table 3.5). Prentice *et al.* (1987) found a 5-10 km source area for *Picea*, whereas the Québec results lacked any change in R^2 after 20 km and there was no sampling at finer spatial resolutions than 10 km.

Table 3.5: ***Estimated source areas (in km) from studies of similar spatial scale, except Jackson (1990) who used small lakes.***

Taxon	Prentice *et al.* (1987) Source Radius (km)	Bradshaw & Webb (1986) Source Radius (km)	Jackson (1990) Source Radius (km)	Present Source Radius (km)
Picea	5 < x < 10	-	x = .1	x < 20
Pinus	50 < x < 100	x > 30	x > 1	50 < x < 60
Fagus	5 < x < 100	x < 4.5	x = 1	Indifferent
Betula	Indifferent	4.5 < x < 30	x > 1	x ≤ 30
Quercus	Indifferent	x > 30 km	x ≥ 1	Indifferent
Acer	-	x < 2.3	*x ≤ .1	Indifferent
Tsuga	-	4.5 < x < 30	x ≤ .5	Indifferent
Ulmus	-	4.5 < x < 30	-	-
Sample Radius	5-100	2.5, 4.5 & 30	.2-1	10-100
Lake Size (ha)	2-220	30-150	0.1-0.5	0.6-730

**Acer rubrum.*

3.4.4 EFFECTS OF PLANT ABUNDANCE WEIGHTING ON ESTIMATES OF POLLEN PRODUCTIVITY AND BACKGROUND POLLEN

There are few substantial changes in the magnitude of the OLS-y parameters with increasing sampling radius for any single plant abundance weighting technique with one exception: Sutton-Prentice dispersal weighted plant abundance. Generally, for all plant abundance weightings, slope estimates increased and intercepts decreased with increasing sampling radius (Figure 3.10 and 3.11). For equal weighting methods RCC($1/d^2$) and ACC($1/d^2$) produced the smallest slopes and largest intercepts for all taxa. Trends in slope and intercept are asymptotical for most taxa under weighting and more variable for unweighted plant abundances. *Betula* (Figure 3.10*g*) exhibits asymptotical behavior for all plant abundance weightings. The slope estimates for *Quercus* (Figure 3.10*f*) and *Tsuga* (Figure 3.10*i*) are essentially invariant with increasing sampling radius.

When neutral and unstable Sutton-Prentice dispersal weightings of plant abundance are used, *Abies* (Figure 3.10*b*), *Fagus* (Figure 3.10*c*) and *Tsuga* (Figure 3.10*c*) exhibited slope estimates whose magnitudes and trends at any given sampling radius were very different from the equal-weighting schemes. Alternatively, trends in the slopes and intercepts for *Pinus* (Figure 3.10*h*) and *Quercus* (Figure 3.10*f*) reflected those of the unweighted vegetation. For the other taxa, trends under all weightings were effectively identical.

Among the plant abundance weighting techniques, the magnitude of the slope estimates at any given plant abundance sampling radius are variable (Figure 3.10). This variability in slope values means that plant abundance weightings would produce different estimates of pollen productivity (slopes), particularly for *Betula*, *Pinus*, *Abies*, *Fagus*, *Quercus* and *Tsuga*. There is less variability among the intercept (background pollen estimates) terms between the different plant abundance weighting schemes (Figure 3.11). The most pronounced differences in estimated slopes and intercepts in all cases are between the unweighted plant abundances and the Sutton-Prentice dispersal weighting under neutral atmospheric conditions and the inverse-squared distance weighting.

3.4.5 POLLEN PRODUCTIVITY AND BACKGROUND POLLEN INPUT

Bradshaw and Webb (1985) analyzed datasets in Wisconsin and Upper Michigan (WI-UM), a subset of these for moderate sized lakes (MOD) and a dataset in Lower Michigan (LM) using OLS-y and unweighted basal area sampled within a 30 km radius (Table 3.6). There have also been analyses of this same regional dataset using ERV and GM regression (Jackson *et al.*, 1995; Prentice and Webb, 1986; Prentice, 1985; Prentice and Parsons, 1983; Webb *et al.*, 1981). Here the dataset of Bradshaw and Webb (1985) is compared to the Québec dataset sampled within 30 km of each pollen site utilizing unweighted plant abundance.

Table 3.6. ***Comparison of Midwestern Wisconsin and Michigan pollen productivity (Slope), Background Pollen (Intercept) and Correlation (R) estimates to Southern Québec (Present).***

	*OLS-y Slope				*OLS-y Intercept (%)				[2]Pearson's R				
Taxon	MOD	WI-UM	LM	[3]Present	MOD	WI-UM	LM	[3]Present	WI-UM	MOD	LM	[3]Present	p (present)
Picea	0.39	0.34	0.22	0.7±0.05	0.70	0.80	0.40	5.05±1.38	0.60	0.73	0.84	0.91	0
Abies	0.07	0.09	0.07	0.22±0.04	0.20	0.20	0.10	1.34±0.7	0.65	0.56	0.74	0.67	3.37E-07
Larix	-0.01	0.05	0.00	-	0.40	0.20	0.00	-	-	-	-	-	-
Cupressaceae	0.14	0.17	0.07	-	1.30	0.90	0.80	-	-	-	-	-	-
Pinus	1.36	1.59	1.15	1.81±0.45	25.90	23.80	21.20	11.57±2.16	0.43	0.41	0.45	0.52	2.10E-04
Betula	1.79	2.09	0.85	0.92±0.22	5.90	5.30	5.20	27.19±4.9	0.65	0.76	0.80	0.53	1.39E-04
Populus	0.05	0.05	0.02	0.04±0.01	0.00	0.20	1.80	0.05±0.14	-	-	-	0.5	3.81E-04
Fraxinus	0.17	0.17	0.32	0.34±0.04	0.50	0.70	1.60	0.3±0.12	0.57	0.59	0.75	0.77	4.86E-10
Quercus	1.29	1.08	1.02	1.18±0.23	7.20	7.50	13.00	1.85±0.51	0.90	0.94	0.89	0.6	8.78E-06
Ulmus	0.45	0.48	0.41	1.17±0.54	1.70	1.80	4.00	1.06±0.25	0.66	0.64	0.54	0.31	3.65E-02
Acer	0.04	0.06	0.09	0.06±0.01	1.10	1.10	1.80	0.23±0.38	0.35	0.24	0.52	0.59	1.43E-05
Tilia	0.06	0.11	0.04	0.23±0.04	0.20	0.00	0.20	0±0.07	-	-	-	0.71	3.80E-08
Juglans	0.28	1.05	0.87	-	0.40	0.30	0.70	-	0.60	0.25	0.63	-	-
Carya	0.16	0.42	0.33	1.05±0.13	0.90	0.40	1.00	0.05±0.03	0.76	0.45	0.76	0.78	1.51E-10
Tsuga	2.32	0.85	2.31	1.09±0.13	0.50	1.80	1.70	0.88±0.37	0.52	0.71	0.66	0.77	3.18E-10
Fagus	0.22	0.34	0.56	0.34±0.11	0.20	0.10	3.20	1.57±0.57	-	-	-	0.43	2.86E-03
Ostrya-Carpinus	-	-	-	0.81±0.12	-	-	-	0.24±0.13	-	-	-	0.71	4.49E-08

*MOD – Modified dataset of Prentice and Webb, 1986, WI-UM – Combined Wisconsin-Upper Michigan dataset (ibid), LM – Lower Michigan dataset (ibid); [2]From Webb *et al.*, 1981, [3]Derived via linear regression of pollen proportions vs. unweighted plant abundance within a 30 km radius.

Picea is the most equally represented taxon and has the strongest linear relation between pollen and plant-abundance of all taxa in Québec (Figure 3.3a, Figure 3.12). *Picea* pollen and plant abundances are between 60-80% and mapped patterns are spatially identical (Figure 3.3a). In the Midwestern United States datasets

(e.g., Figure 4d in Webb *et al.*, 1981), however, *Picea* basal areas did not exceed 15% and the scatterplots suggest greater under-representation than the present study.

Pinus, *Quercus*, *Tsuga*, *Betula* (Figure 3.12) and *Ulmus* are over-represented in the Québec pollen spectra and in the Midwest (Table 3.6). However, *Pinus* intercepts are lower and slopes are higher in Québec than in the Midwest (Table 3.6) where *Pinus* is more abundant in the vegetation. At numerous sites, *Pinus* comprises as much as 16% of the pollen spectra but is not present in the surrounding vegetation (Figure 3.12). The over representation of *Pinus* is clearly evident in the isopoll and isoabundance maps where *Pinus* pollen is as much as 60% where tree abundance is only 10% (Figure 3.3c). Although *Betula* in Québec exhibits significantly larger intercepts, the slopes are within the range of variability observed in the Midwestern datasets (Table 3.6). Low *Betula* plant abundances are associated with a range of high pollen proportions (Figure 3.12; Figure 3.3d). A similar scatter for *Betula* was found in Sweden (Prentice *et al.*, 1987) and in Wisconsin and Upper Michigan (Bradshaw and Webb, 1985). The differences in background pollen between *Betula* and *Pinus* in Québec versus the Midwest is consistent with their regional abundance differences in the two regions.

Both the pollen and plant abundances for *Quercus* in Québec are low, yet this taxon remains over represented in the pollen spectra similarly to the Midwest (e.g., Figure 3.8*a* in Bradshaw and Webb, 1985). *Ulmus* is over represented in Québec with a higher slope compared to the Midwest, however, the Québec data only reach abundances less than 1%, which easily fit into the pollen-plant abundance scatter as great as 25% in the Midwest (e.g., Figure 9*a* in Bradshaw and Webb, 1985). *Ostrya-Carpinus* is over represented in this study but was not analyzed in the Midwest.

Abies, *Acer*, *Populus*, *Fagus*, *Tilia*, and *Fraxinus* are under represented in the present study and in the Midwestern data (Table 3.6). There are higher abundances of *Abies* in Québec and it is not as under represented as in the Midwest (e.g., Figure 5*b* in Webb *et al.*, 1981). Although the Québec data exhibit

slope estimates for *Fraxinus* that are slightly higher than those in the Midwest (Table 3.6), the Québec scatter fits into the Midwest data up to approximately 6.5%. Most of the *Fagus* sites in Québec exhibit more under representation when compared to the Midwest. *Acer* and *Populus* are severely under represented in the pollen spectra (Figure 3.12) of Québec and in the Midwest (Table 3.6).

For the average relations between pollen and plant abundance, *Tilia*, *Fraxinus*, and *Ostrya-Carpinus* tend to be under represented and *Juglans*, *Larix* and *Cupressaceae* have insignificant linear relations in this study (Table 3.6). Although *Carya* indicates larger slopes than in the Midwest, its scatter is clearly questionable considering its low pollen and plant abundances (Figure 3.3, Figure 3.4).

The correlations between pollen and plant abundance in Québec are all within the range of variability observed in the Midwest (Table 3.6). In the Midwest, *Betula* exhibits a stronger relation than in Québec (Table 3.6), however, an influence analysis indicates that if Lac Louis (Table 3.1) is removed then the correlation between *Betula* pollen and plant abundance increases to $r = 0.65$ which is similar to the Midwest. *Larix* and *Cupressaceae* are not significantly related to vegetation in southern Québec.

3.4.6 Effect of Vegetation Heterogeneity on Parameter Estimates

The patchiness of plant abundance on the landscape may be explored by looking at the mean proportions of a taxon at different vegetation sampling radii surrounding all modern pollen sites. If plant abundance changes little with increasing vegetation sampling radius then the proportion of a given taxon, e.g., *Picea*, at one sampling radius, e.g., 10 km, would be equivalent to the proportions of that taxon at a larger sampling radius, e.g., 30 km. If the proportions at 10 km are less than those at 30 km then that particular taxon may represent a 'hole' in the regional matrix, or alternatively, if the proportions at 10 km are greater than at 30 km then this may indicate that *Picea* patches are found within a more heterogeneous vegetation matrix. In southern Québec, *Acer*, *Picea* and *Betula* exhibit consistent proportionality at two different scales (Figure 3.13ade). *Pinus* and *Fagus* exhibit some patchiness at the 10 km scale such that proportions are generally greater at 10 km than at 30 km on the landscape (Figure 3.13fi).

By removing within-site variation in plant abundance patterns through randomization, different plant abundance weighting methods produce identical trends in explained variance (Figure 3.14), slopes and intercepts (not shown) with increasing plant abundance sampling radius. These R^2 values are those expected due only to between site variations in plant abundances within the region. Exceptionally, Sutton-Prentice Neutral weighting produced slightly higher absolute explained variance (Figure 3.14g) and smaller slopes (not shown) for *Betula* and larger overall slopes for *Abies* and *Tsuga* (not shown) at all sampling radii.

3.5 Discussion

Issues involving the effects of different calibration models have been dealt with in previous research (Jackson *et al.*, 1995; Prentice, 1986; Prentice and Webb, 1986; Bradshaw and Webb, 1985; Prentice and Parsons, 1983; Parsons and Prentice, 1981; Webb *et al.*, 1981), with the exception of the models' behavior under different vegetation weightings. All calibration models exhibit increasing slopes and decreasing intercepts with increasing vegetation sampling radius. As sampling radius increases, more of the pollen contributing area, or 'effective' source area, is contained and so background components decrease towards zero and slopes increase commensurately. A decreasing background pollen component with increasing sampling radius is consistent with studies in the Midwestern United States (Bradshaw and Webb, 1986) and simulations (Sugita, 1994).

Some taxa in southern Québec exhibit significant linear relations to plant abundance. For the abundant regional taxa, the overall agreement of the slopes, intercepts and correlations between pollen and plant abundances in Québec and the Midwest suggest that regional variations in the relative abundance of different taxa (e.g., *Picea* is a dominant in Québec whereas *Quercus* is a dominant in the Midwest) have only minor effects on the relative representation of a given taxon in the pollen spectra as estimated using OLS-y slope parameters. This observation supports the hypothesis of Bradshaw and Webb (1985) that the

slope parameters, estimates of relative pollen productivity, are recording differences in the production and dispersal independently of the relative plant abundances of different physiognomic regions. Although Bradshaw and Webb did not suppose that their regression parameters would be applicable to widely different physiognomic regions like the Boreal forest, the Québec results are encouraging in that regard. However, background pollen inputs are clearly a function of the dominant types in a region which is why the Québec data shows higher background estimates for *Picea*, *Betula* and *Abies* and lower background estimates for *Pinus* and *Quercus* than in the Midwestern datasets.

Larix, *Cupressaceae* and *Juglans* had no significant linear relations to pollen proportions in Québec. Although *Carya*, *Ostrya-Carpinus*, and *Tilia* had significant linear relations, they tend to show positive pollen proportions and zero plant abundances or vice-versa. This is because they are at their northern limit in the Québec region and their occurrence on the landscape is too sparse to effectively produce a signal in the pollen spectra on this large regional scale when compared to the dominant boreal taxa in the region. Nothing definitive about these taxa can be concluded.

Moreover, estimates of both pollen productivity and background pollen are partially dependant on the type of distance-weighting applied to plant abundance and on the plant abundance sampling radius used for comparison with pollen proportions. For consistency, the Québec estimates are compared to the Midwestern ones using unweighted relative plant abundance sampled at 30 km around each pollen site. However, this comparison is specific to this plant abundance weighting and sampling scale. As is shown, pollen productivity and background pollen estimates can change with increasing sampling radius and as a function of different distance-weighting models. No single plant-abundance weighting, however, produces clearly better relations for all taxa concerned.

It was hypothesized that if pollen is a distance-weighted function of plant abundance, then distance weighting should yield greater explained variance when compared to unweighted vegetation, allowing

unequivocal definitions of 'relevant' pollen source area. No single weighting scheme, however, produced a superior explanation of variance for all taxa when compared to unweighted vegetation. Most taxa were significantly correlated with vegetation at most sampling radii with invariant trends so inferences of 'relevant' source area were not clarified by alternate weighting schemes. Estimates of pollen productivity, background pollen and source area are a function of plant abundance weighting technique. However, these differences are primarily artifacts of the nature of the individual weighting methods.

Only *Betula*, *Pinus* and *Abies* had clear asymptotical behavior in R^2 under inverse-squared distance weighting with increasing sampling radius, which allowed estimates of the 'relevant' pollen source area to be made. Moreover, for these taxa, the asymptote occurs where the maximum explained variance in unweighted vegetation is found. Jackson and Kearsley (1998) suggest that under inverse-squared distance weighting, as distance from a pollen sample site increases, the plant abundance weighting terms, w_i's, become smaller (e.g., $\lim_{d \to \infty} w(d) = 1/d^2 = 0$) and so any variations in plant abundance, no matter how extreme, eventually will no longer contribute substantially to changes in the plant abundance sum (e.g., Figure 3.6). The plant abundance sampling radius at which R^2 stabilizes is not pre-determined by a weighting model. However, asymptotical behavior and the definition of 'relevant' pollen source area are artifacts imposed on the pollen-plant abundance relation as weights become smaller with distance. Although the Québec data indicate asymptotes of R^2 to occur in the vicinity (±10 km) of the R^2 maxima in the unweighted plant abundances, the use of inverse-squared distance weighted plant abundances may not detect variation in the pollen proportion plant abundance relations that occur at larger sampling radii (Jackson and Kearsley, 1998). In addition, what should be considered a "significant" improvement in R^2 from one sampling radius to another? If one looks at R^2 in *Quercus*, it shows asymptotical behavior but over a range of only 0.02 R^2; such a change is essentially invariant.

Inverse-distance weighting of plant abundance also affects inferences regarding background pollen estimates and pollen productivity. Inverse-squared distance weighted plant abundance caused larger intercept terms and smaller slopes at all sampling radii. As such, when inverse-squared distance weighted plant abundance is used, the Québec data indicates that slopes increase at a decreasing rate and intercepts decrease at a decreasing rate with increasing sampling radius since effectively only plant abundances close to a pollen sample site contribute to substantial variation in the plant abundance sum. Inverse-squared distance weighted plant abundance imposes an effectively smaller sampling radius in which pollen and plant abundance are compared, such that the relations uncovered at the smallest sampling radii are maintained at greater distances.

Sutton-Prentice dispersal weighting describes the proportion of pollen originating (remaining airborne) from a tree at a given distance from the basin center. When used for weighting plant abundance by distance, the technique did not improve pollen plant abundance relations as measured by R^2 compared to unweighted plant abundance except for *Betula*. Calcote (1995) observed that the dispersal weighted model provided results similar to the inverse-squared model at the forest-hollow scale. However, Calcote did not incorporate variability in basin size, whereas this was attempted here. Under the unstable (Jackson and Lyford, 1999) formulation of the Sutton-Prentice model, the trends with increasing sampling radius and the magnitudes of R^2 mirrored those in the unweighted plant abundance - with no asymptotical behavior unless this was present in the unweighted curves themselves such as with *Betula*. This effect is produced under the unstable formulation because distance weights decay with distance at a rate less than the equal weighting and neutral formulation methods and so the plant abundances at large distances still contribute substantially to changing the plant abundance sum. Under the neutral model, results for R^2 fell between the inverse-distance and unweighted plant abundances.

Dispersal weighted plant abundances for *Abies*, *Fagus* and *Tsuga* resulted in significantly larger slope estimates at all sampling radii when compared to equal weighted plant abundances. This effect is due to

the interaction between the magnitude of dispersal weights as a function of pollen fall velocity among the 17 taxa in this study and the spatial resolution of Québec vegetation quadrats. The mean distance between all pollen sample sites and their nearest vegetation quadrat is 4.1 km. *Abies* pollen has the greatest fall velocity of the taxa under consideration (Figure 3.6). In a basin with a radius of 130 m, an *Abies* plant abundance sample within 4.1 km from the basin center would only receive a weight of ~8.9% of its total cover, as compared to say ~50% for *Pinus* or 56% for *Betula* at the same distance. Thus, beyond a few kilometers *Abies*, *Tsuga*, and *Fagus* plant abundances are given such a small weight, compared to taxa whose pollen have slower fall velocities. Thus, the variation of plant abundances at even small distances does not contribute much to changing the dispersal weighted plant abundance sum. In other words, because most vegetation quadrats are found at a distance from pollen sample sites, taxa whose pollen types possesses fast fall velocities are forced to become over represented in the pollen spectra even if they are regionally very abundant. The effect is greater for Neutral conditions because weights decrease faster with distance than in the Unstable formulation. Unless vegetation samples are finely resolved so that taxa whose pollen has fast fall velocities are abundant in the vegetation adjacent to a lake, the dispersal weighted plant abundance will lead to overestimates of pollen productivity at the regional scale regardless of the regional abundance of such taxa in the surrounding vegetation.

Within-site variation in plant abundance (within 100 km of a lake site) affects how the weighting of plant abundance influences the pollen-plant abundance relation and the definition of pollen source area. If a taxon is randomly distributed in the vegetation surrounding a pollen sample site, then its proportions are relatively equivalent at all sampling radii. Therefore, only between site variations in plant-abundance will affect the pollen-plant abundance relation. All distance-weighting techniques produce the same results when there is no within-site variation of plant abundance. As such there is no apparent asymptotical behavior or variation in R^2. Consequently, the empirical definition of 'relevant' pollen source area is dependent on consistent and measurable within-site vegetation heterogeneity within a network of sites.

In addition, comparing parameter estimates from the randomized data to those of the actual data illustrates the degree to which the within-site variation in plant abundance controls the definition of 'relevant' source area estimates. For example, R^2 curves for *Picea* under randomization are the same as those produced by the non-randomized data (Figures 3.9*a* and 3.14*a*) indicating that *Picea* proportions are distributed similarly in both cases. *Picea* shows little to no patchiness on the landscape at different vegetation sampling scales (e.g., Figure 3.13d). However, Sugita's (1994) simulations suggested that if a taxon is found on the landscape in patch sizes smaller than the collecting basin then that taxa may appear 'homogenous' when measured by the pollen spectra. Therefore, the average patch size of *Picea* on the landscape in Québec may occur at scales smaller than the average basin size, but Québec vegetation data is not sufficiently resolved to test this hypothesis. Sugita (1994) showed that consistent patch sizes in simulated landscapes were required to produce asymptotical behavior in R^2 with increasing vegetation-sampling radius. However, Sugita only used three taxa with patches randomly distributed over space with a constant proportion and the simulated landscapes did not possess significant gradients in vegetation abundance or between site variations. The results with randomization, illustrate the importance of vegetation heterogeneity to the definition of pollen source area, but more work is required with vegetation data at a finer spatial resolution at the regional scale to test Sugita's simulation results.

Relative crown cover should provide a better proxy to relative pollen source strength, as crown area is a proxy itself for the crown volume, which is proportional to the surface area where pollen is produced. However, the Québec results do not show improved relations between pollen and plant abundance over previous pollen-vegetation studies in other regions of North America that utilize basal areas as a plant abundance measure.

This study extends those in the Midwestern United States by providing pollen-plant abundance comparisons of the mixed Northern Hardwood and Boreal forests of North America. Estimates of pollen productivity and relative representation among taxa in the different physiognomic regions are reasonably

consistent. However, different vegetation weightings will produce different estimates of pollen productivity, background pollen and pollen source area on the same dataset. Therefore, plant abundance weighting is a critical decision and should be considered when calibrating pollen proportions to plant abundance for the purpose of paleovegetation reconstruction. Further work is required to determine the effects of vegetation heterogeneity on source area estimates but will require vegetation data sampled at fine spatial resolutions across broad regions.

3.6 Conclusion

Understanding the relation between modern pollen and plant abundance is fundamental to inferring vegetation change from fossil pollen sequences and for inferring the spatial extent of past vegetation communities using networks of fossil pollen sites. Because pollen loading in a basin is a distance-weighted function of the surrounding plant abundance, the implications of different distance weighting techniques on the definition of relevant pollen source area and estimates of pollen productivity and background pollen input were tested. Modern pollen samples at 46 lakes in southern Québec were compared for the first time to surrounding plant abundance derived from a detailed survey of crown area. The comparison to similar studies in Wisconsin and Michigan revealed that the regional variations in vegetation composition (e.g., Boreal vs. Mixed Deciduous Forest) do not significantly affect relative pollen productivity estimates but do affect background pollen input. The relevant pollen source areas for major arboreal taxa are consistent with estimates in the Midwestern United States and Sweden. However, explained variance estimates between plant abundance and pollen vary in both magnitude and trend as a function of different vegetation weightings and this has implications for the definition of relevant pollen source area, pollen productivity and background pollen estimates. The randomization approach illustrated that the variation in parameter estimates were due to within-site vegetation heterogeneity and its effects on distance-weighted plant abundance. The variability introduced into the pollen-plant abundance relation from different plant

abundance weightings is a critical issue in our understanding of the degree to which pollen represents plant abundance on the landscape.

Only the major arboreal taxa in this study were shown to have significant relations to plant abundance at the regional scale when lakes are used as the sensors for the surrounding vegetation. These types should show a strong signal in the fossil record. However, less abundant types in this region, like *Ulmus* and *Juglans*, also show consistent relations to the Midwestern analysis. Other types like *Larix* and *Tilia* show no signal at these regional scales and may only contribute noise when they are included in pollen sets for paleoclimatic analysis. It is the regional scale vegetation that is controlled by the macroclimate. The noise occurs because some sites contain these taxa in the surrounding vegetation but their pollen may be absent or not proportional to their representation within the pollen spectra and this can affect the squared chord distance coefficient and choice of best modern analog. In other words, pollen proportions are not consistent with the macroscale vegetation and so pollen proportions may be large or small with no defined relation to the vegetation that produced the pollen.

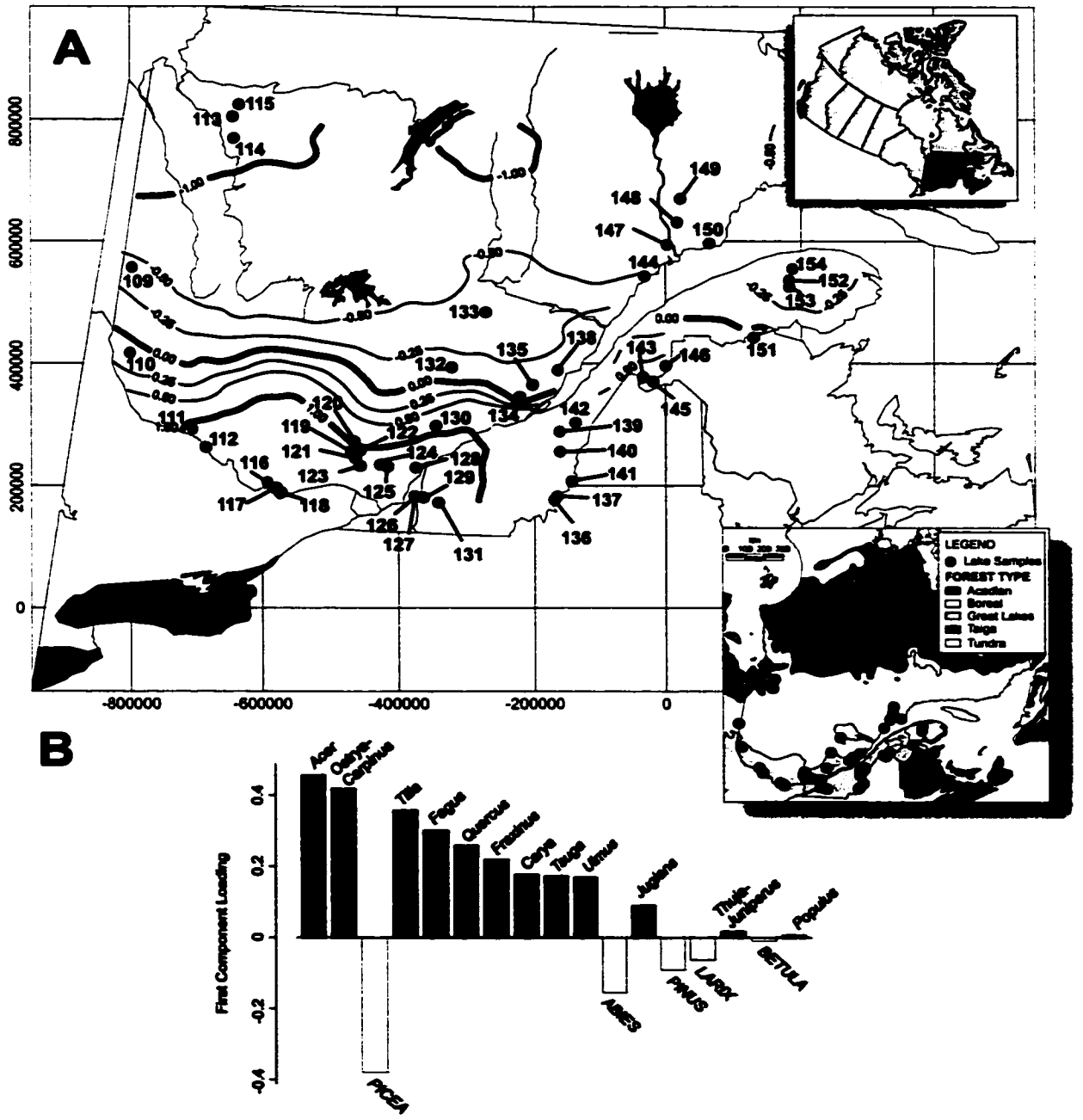


Figure 3.1: A – Study area with lake pollen–sampling sites and Site # corresponding to Table 3.1. and Mapped principal component scores for first component illustrate the gradation from Northern Hardwood Forest in the South to Boreal Forest in the North. B – First component loadings for the 17 taxa in this study. Forest regions inlay is modified from NRCAN (2001).

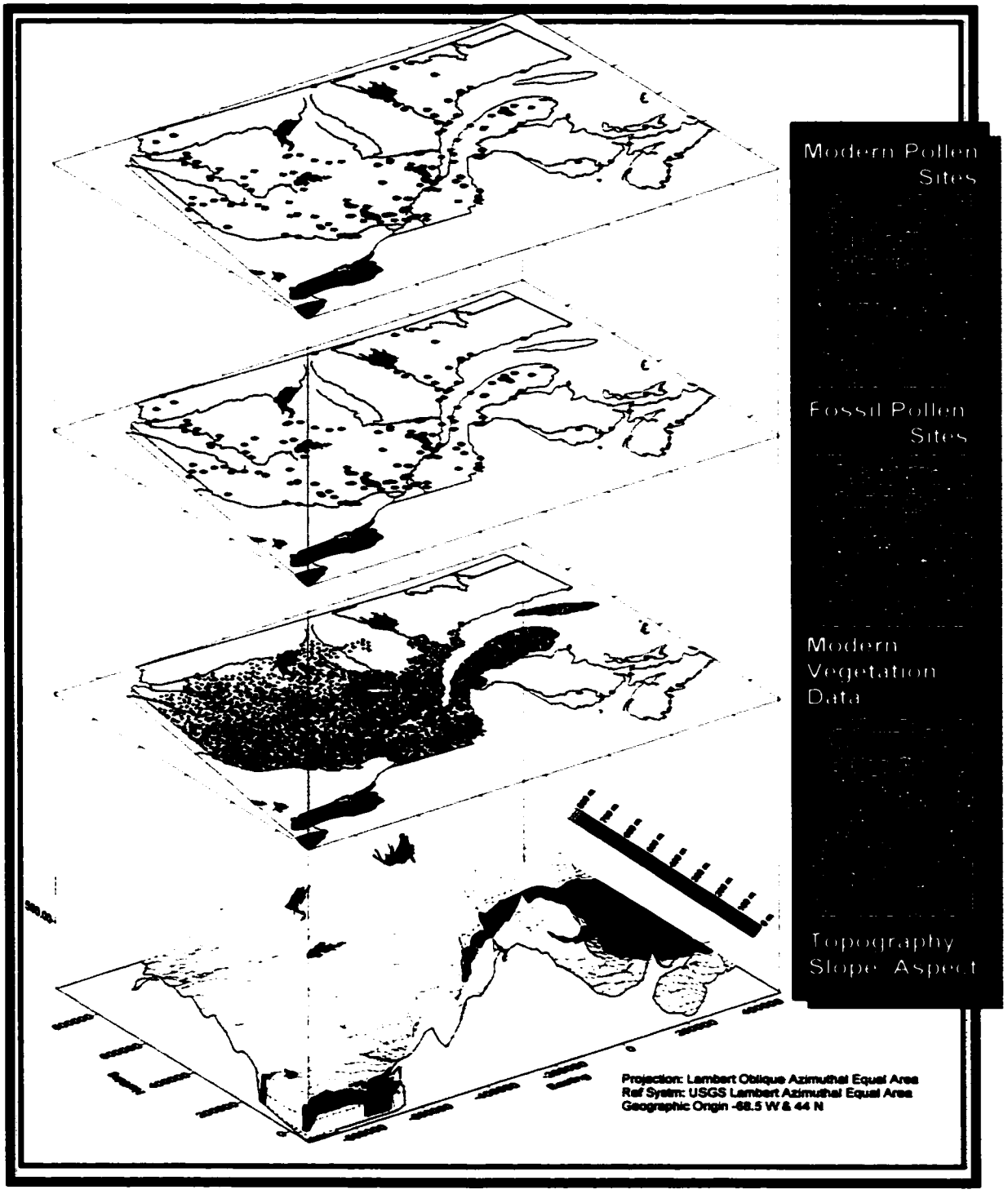


Figure 3.2: Database schematic illustrating various data coverages in the GIS database for this research.

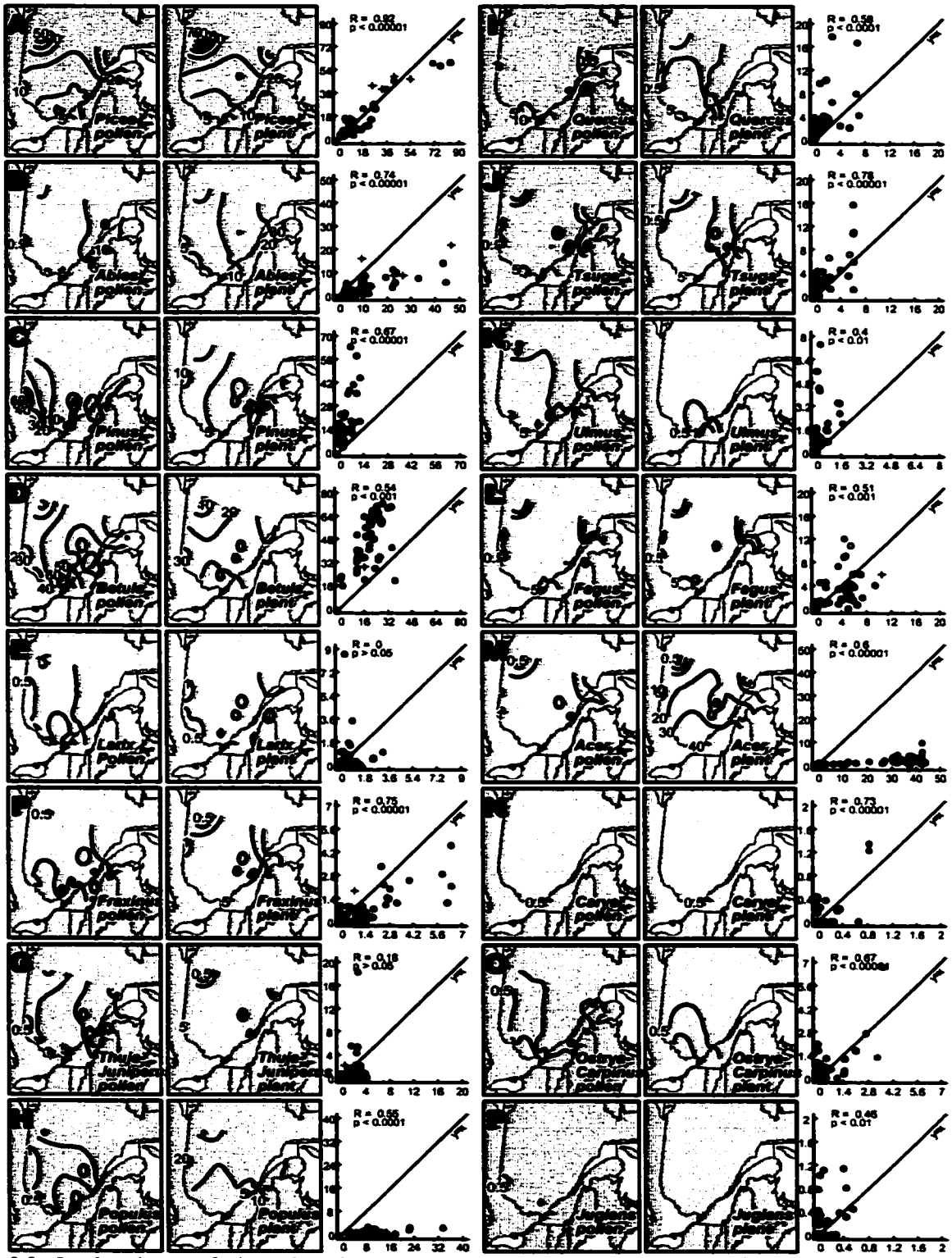
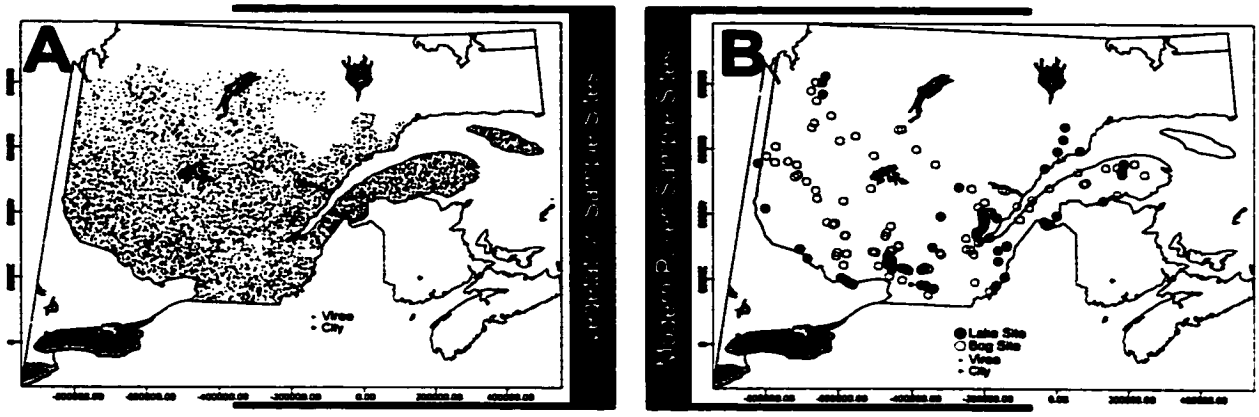


Figure 3.3: Isoabundance of plants based on plant abundance sampled at the 30 km radius around each pollen sample site and Isopoll maps with contours showing pollen and plant proportions. Scatter plots illustrate the average plant percentages within 100 km of each pollen sample site (x–axis) vs. pollen percentages (y–axis) – Red open circles denote lakes from 0–20 ha; Closed blue circles denote lakes from 20–50 ha; Green crosses denote lakes greater than 50 ha. ***Note*** the different scales for each taxon.



C

Lakes & Bogs	10 km	20 km	30 km	40 km	50 km	60 km	70 km	80 km	90 km	100 km	Pollen	Vegetation
	21.36	21.04	20.69	20.62	20.34	19.93	19.46	19.03	18.83	18.59	24.39	29.66
	19.66	20.93	21.12	21.19	21.11	21.41	21.57	21.96	22.11	22.47	1.38	13.32
	19.63	20.44	20.77	20.68	21.04	21.43	21.71	21.91	22.05	22.10	32.92	21.88
	15.01	14.18	14.60	14.29	13.67	13.10	12.60	12.09	11.86	11.64	6.69	12.17
	9.71	9.51	9.45	9.53	9.94	9.98	10.35	10.50	10.51	10.47	0.34	0.66
	4.44	4.18	3.90	4.01	4.12	4.31	4.38	4.53	4.59	4.56	24.67	6.40
	2.72	2.78	2.69	2.68	2.66	2.60	2.60	2.59	2.56	2.55	1.62	1.11
	1.68	1.02	0.80	0.74	0.76	0.79	0.79	0.78	0.78	0.80	0.38	1.66
	1.59	1.64	1.58	1.69	1.72	1.77	1.79	1.84	1.91	1.98	0.43	2.10
	0.94	1.07	1.22	1.34	1.32	1.36	1.37	1.37	1.38	1.40	0.93	0.87
	0.92	0.93	0.80	0.80	0.82	0.81	0.84	0.83	0.81	0.80	2.40	0.46
	0.81	1.06	1.13	1.18	1.22	1.23	1.22	1.24	1.26	1.27	1.94	0.66
	0.48	0.63	0.66	0.64	0.62	0.61	0.61	0.61	0.61	0.62	0.13	0.26
	0.17	0.34	0.32	0.32	0.33	0.35	0.37	0.36	0.36	0.36	0.36	0.19
	0.15	0.17	0.16	0.18	0.19	0.21	0.22	0.23	0.24	0.25	1.26	0.16
	0.05	0.04	0.04	0.05	0.05	0.06	0.07	0.07	0.07	0.06	0.15	0.03
	0.01	0.04	0.04	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.00	0.04
LAKES												
	24.86	25.56	24.65	24.50	24.06	24.71	24.68	25.09	25.06	25.33	1.85	13.32
	19.68	19.46	20.11	19.84	19.68	19.65	19.61	19.57	19.81	19.96	46.57	21.88
	18.11	18.14	17.95	18.03	17.94	17.74	17.58	17.19	17.00	16.74	17.46	29.56
	15.11	13.25	12.93	13.04	12.87	12.24	12.18	12.04	12.04	12.03	4.87	12.17
	7.79	8.85	9.60	9.37	9.81	9.75	9.80	9.93	9.87	9.77	0.44	0.66
	3.61	3.89	3.59	3.44	3.32	3.18	3.04	2.95	2.91	2.89	2.85	1.11
	2.59	2.88	2.89	3.15	3.21	3.46	3.61	3.72	3.82	3.77	16.35	6.40
	1.20	0.58	0.56	0.51	0.55	0.64	0.73	0.72	0.69	0.70	0.81	1.66
	1.14	1.35	1.60	1.81	1.96	2.13	2.23	2.28	2.37	2.43	1.27	2.10
	0.99	1.33	1.46	1.60	1.57	1.62	1.59	1.64	1.67	1.64	2.43	0.66
	0.75	1.27	1.59	1.72	1.67	1.74	1.76	1.75	1.75	1.74	0.90	0.87
	0.71	1.31	1.13	1.17	1.27	1.23	1.28	1.24	1.19	1.16	3.11	0.46
	0.67	1.15	1.05	0.92	0.89	0.85	0.82	0.82	0.80	0.80	0.24	0.26
	0.35	0.75	0.55	0.51	0.53	0.56	0.56	0.54	0.51	0.50	0.73	0.19
	0.08	0.07	0.08	0.08	0.09	0.11	0.11	0.10	0.11	0.11	0.12	0.03
	0.06	0.13	0.19	0.24	0.28	0.32	0.34	0.32	0.32	0.34	1.91	0.16
	0.01	0.04	0.05	0.06	0.09	0.09	0.09	0.09	0.09	0.09	0.17	0.04
BOGS												
	22.72	22.33	21.92	21.77	21.41	20.90	20.29	19.85	19.65	19.42	27.34	29.56
	19.61	20.87	21.06	21.05	21.56	22.21	22.65	22.96	23.05	23.04	27.39	21.88
	17.48	18.88	19.55	19.73	19.79	19.94	20.19	20.57	20.80	21.20	1.17	13.32
	14.97	14.60	15.35	14.85	14.02	13.49	12.79	12.11	11.77	11.47	7.86	12.17
	10.51	9.81	9.39	9.60	10.00	10.08	10.59	10.76	10.79	10.79	0.29	0.66
	5.21	4.76	4.35	4.39	4.53	4.69	4.73	4.89	4.94	4.91	28.36	6.40
	2.34	2.28	2.29	2.35	2.37	2.34	2.40	2.42	2.41	2.40	1.87	1.11
	1.88	1.22	0.91	0.85	0.86	0.85	0.81	0.81	0.82	0.84	0.32	1.66
	1.78	1.76	1.57	1.63	1.61	1.62	1.60	1.65	1.71	1.78	0.86	2.10
	1.03	0.98	1.06	1.17	1.16	1.18	1.20	1.20	1.22	1.25	0.94	0.87
	1.01	0.76	0.65	0.63	0.62	0.63	0.65	0.65	0.65	0.64	2.89	0.46
	0.73	0.94	0.98	0.99	1.06	1.06	1.05	1.06	1.06	1.10	1.72	0.66
	0.39	0.40	0.49	0.52	0.51	0.50	0.52	0.52	0.52	0.53	0.66	0.26
	0.19	0.19	0.15	0.16	0.16	0.16	0.17	0.18	0.20	0.21	0.87	0.16
	0.10	0.15	0.21	0.24	0.24	0.26	0.28	0.29	0.29	0.30	0.28	0.19
	0.04	0.02	0.02	0.04	0.04	0.04	0.05	0.05	0.05	0.06	0.17	0.03
	0.01	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.05	0.05	0.06	0.04
	10 km	20 km	30 km	40 km	50 km	60 km	70 km	80 km	90 km	100 km	Pollen	Vegetation

Figure 3.4: A – 4254 Modern vegetation sample sites in Québec; B – Modern lake and bog pollen sites from the Brown University Modern Pollen Database (Avizinis and Webb, 1985); C – Comparison of the average percentages of each taxon from 10 – 100 km surrounding each lake sample site and pollen percentages.

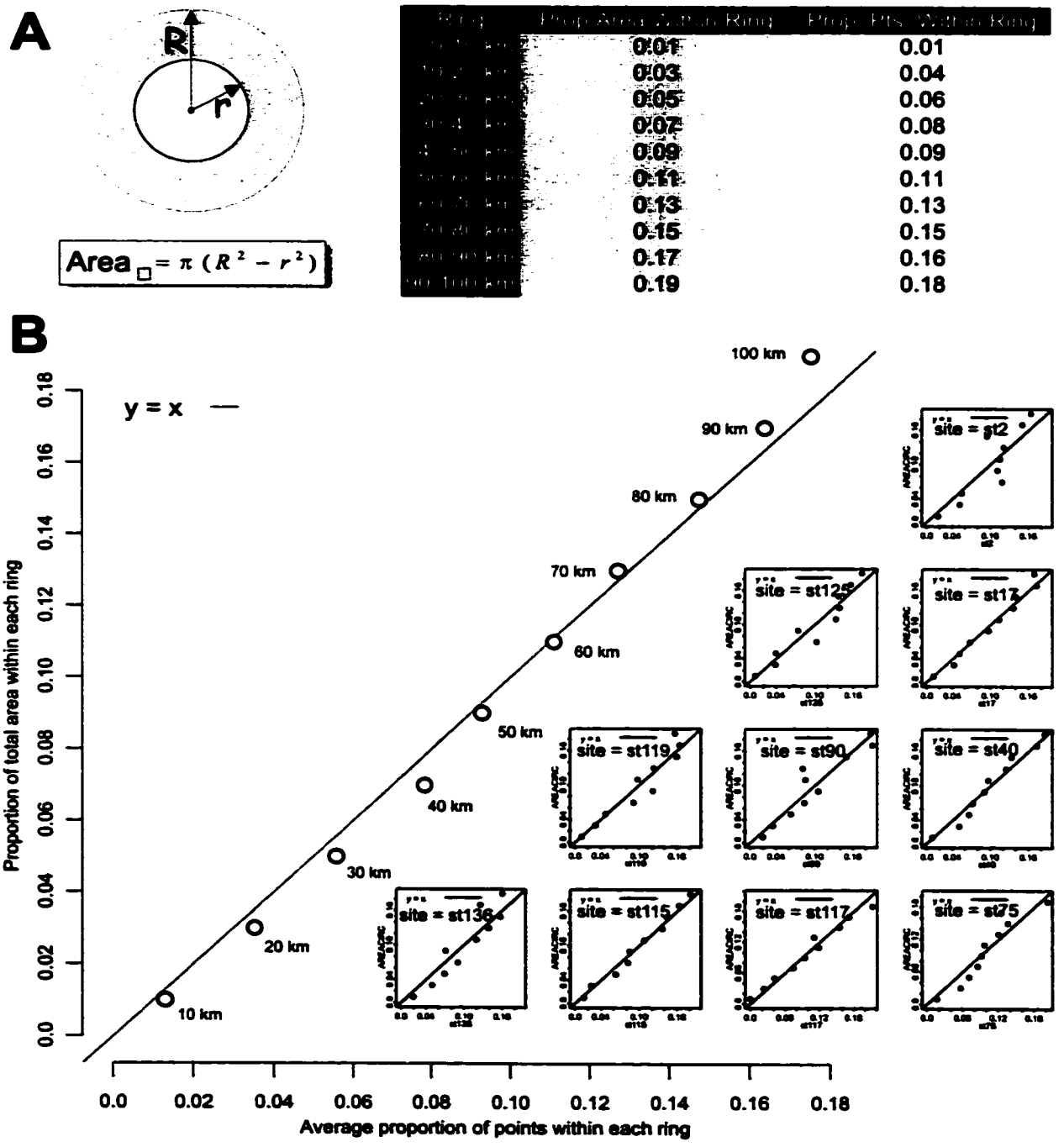
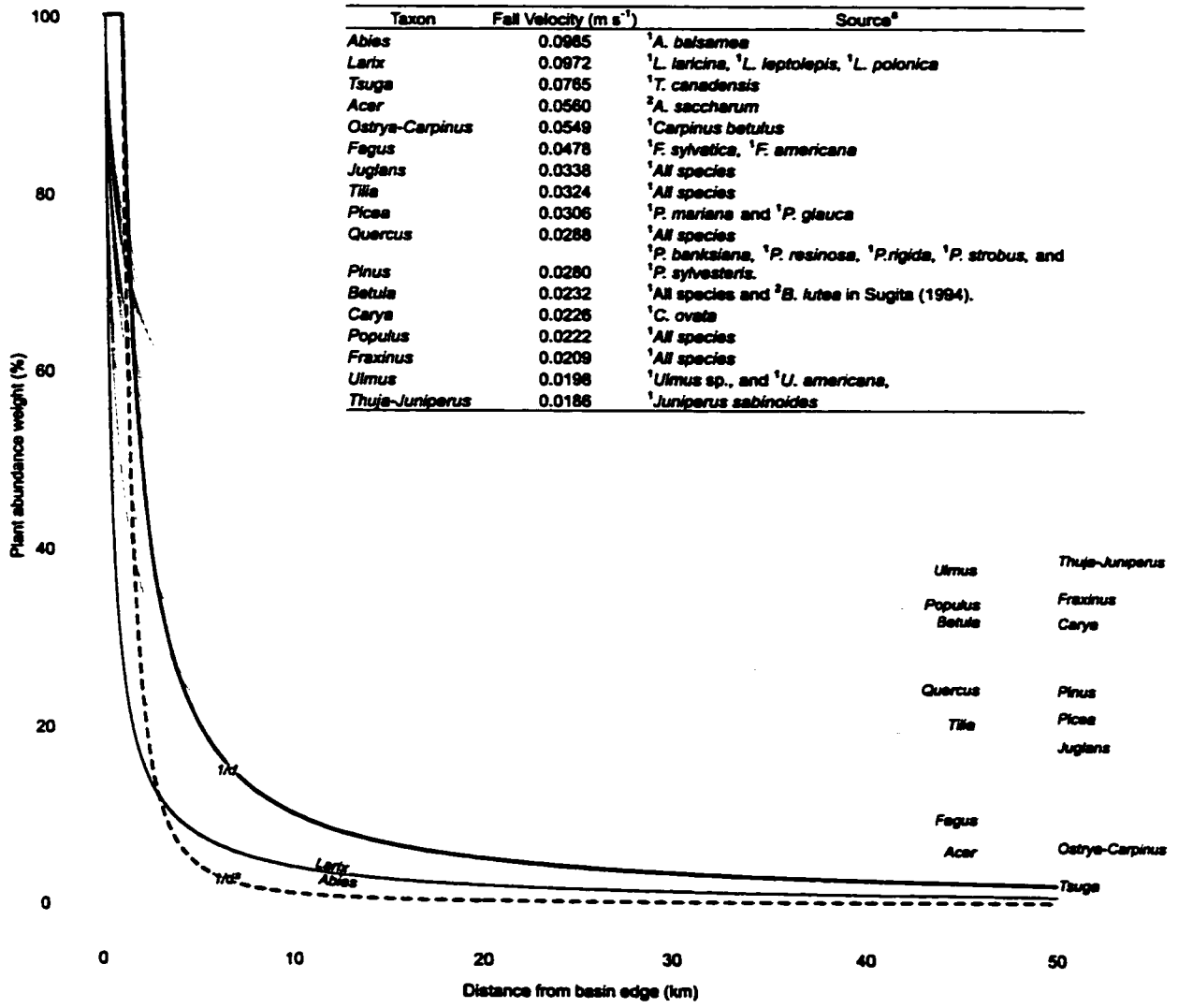


Figure 3.5: Illustration of proportionality between number of vegetation quadrats within each plant abundance sampling radius and the area of each sampling radius. **A** – Definition of sampling radius "rings" and area calculation within each ring with table of values; **B** – Proportion of points within each ring versus the proportion of total area within each ring with representative samples for individual sites. ***Note*** the different *x,y* scales for each site.



Taxon	Fall Velocity (m s^{-1})	Source[a]
Abies	0.0985	[1]*A. balsamea*
Larix	0.0972	[1]*L. laricina*, [1]*L. leptolepis*, [1]*L. polonica*
Tsuga	0.0765	[1]*T. canadensis*
Acer	0.0560	[2]*A. saccharum*
Ostrya-Carpinus	0.0549	[1]*Carpinus betulus*
Fagus	0.0478	[1]*F. sylvatica*, [1]*F. americana*
Juglans	0.0338	[1]*All species*
Tilia	0.0324	[1]*All species*
Picea	0.0306	[1]*P. mariana* and [1]*P. glauca*
Quercus	0.0288	[1]*All species*
Pinus	0.0280	[1]*P. banksiana*, [1]*P. resinosa*, [1]*P.rigida*, [1]*P. strobus*, and [1]*P. sylvesteris*.
Betula	0.0232	[1]All species and [2]*B. lutea* in Sugita (1994).
Carya	0.0226	[1]*C. ovata*
Populus	0.0222	[1]*All species*
Fraxinus	0.0209	[1]*All species*
Ulmus	0.0196	[1]*Ulmus* sp., and [1]*U. americana*.
Thuja-Juniperus	0.0186	[1]*Juniperus sabinoides*

Figure 3.6: Representative distance weighting curves using Sutton–Prentice Dispersal weights (Eqn. 3.2) for a lake size of 130 meters and inverse–squared distance and inverse–distance weights for comparison. Inlay table: [a]Average fall velocities used for Eqn. 3.1 and taxa used to determine these fall velocities and their sources: [1] Jackson and Lyford (1999 - Appendix 1), [2]Sugita (1993 - Table 3.1).

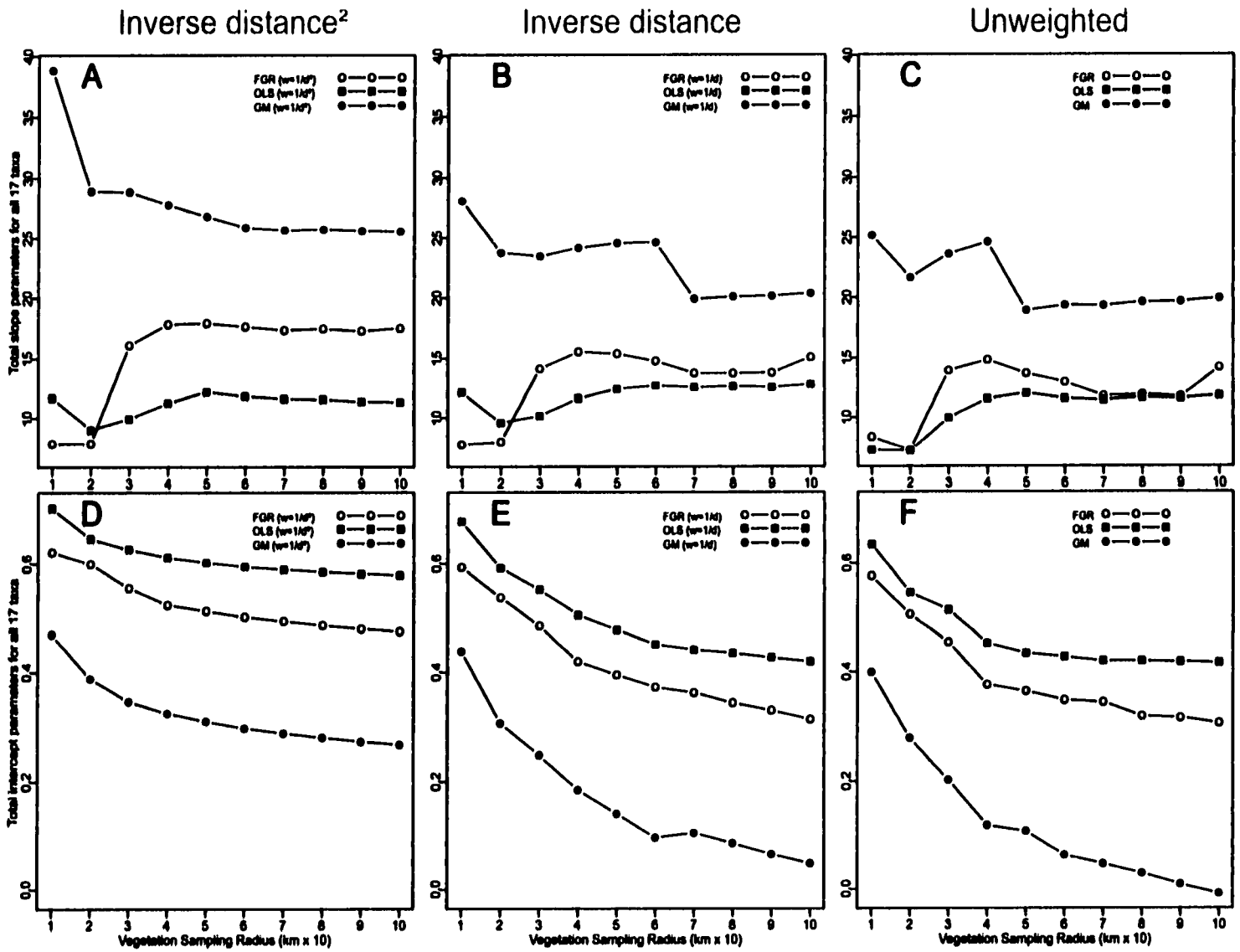


Figure 3.7: A, B & C – illustrate the differences between each calibration technique for a given vegetation weighting on total estimated slopes of all 17 taxa at each vegetation sampling radius. **D,E & F** – illustrates the effect of vegetation weighting on the total estimated background pollen of all 17 taxa at each vegetation sampling radius. FGR – FAGERLIND ERV *z* parameters; OLS-*y* – Ordinary Least Squares intercept parameters; GM – Geometric Mean Regression intercept parameters; *w* – weight function; *d* – distance in km.

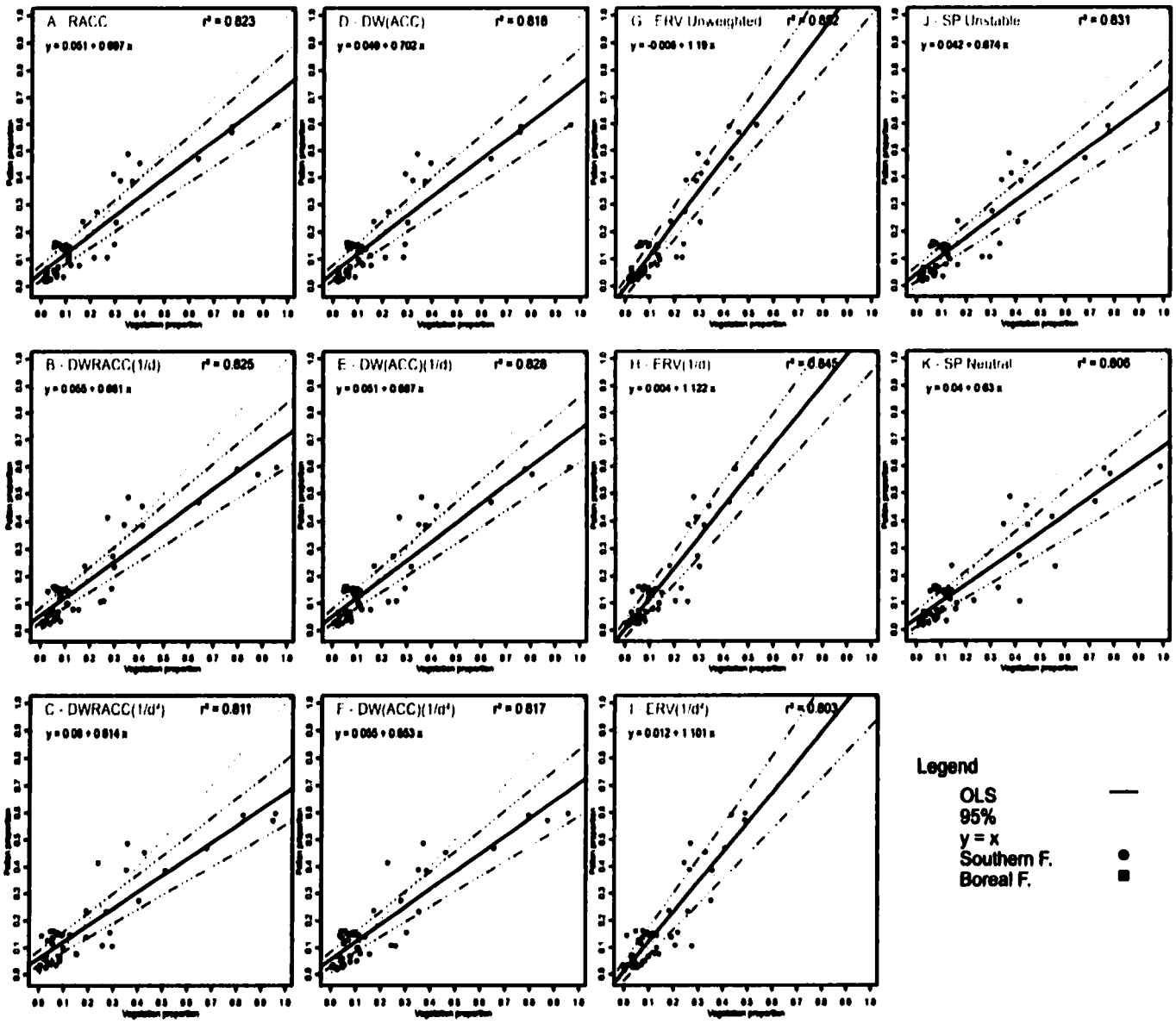


Figure 3.8: Illustration of various weighting schemes and their effect on *Picea* proportions when compared to pollen for OLS–*y* and ERV calibration models. *Picea* proportions at 30km vegetation sampling radius for all weightings and calibration techniques for Picea. ERV vegetation is in the form of corrected vegetation percentages and the best–fit is based on OLS–*y* slope and intercept in order to illustrate how corrected vegetation proportions become linear after calibration using ERV site factors (*f*).

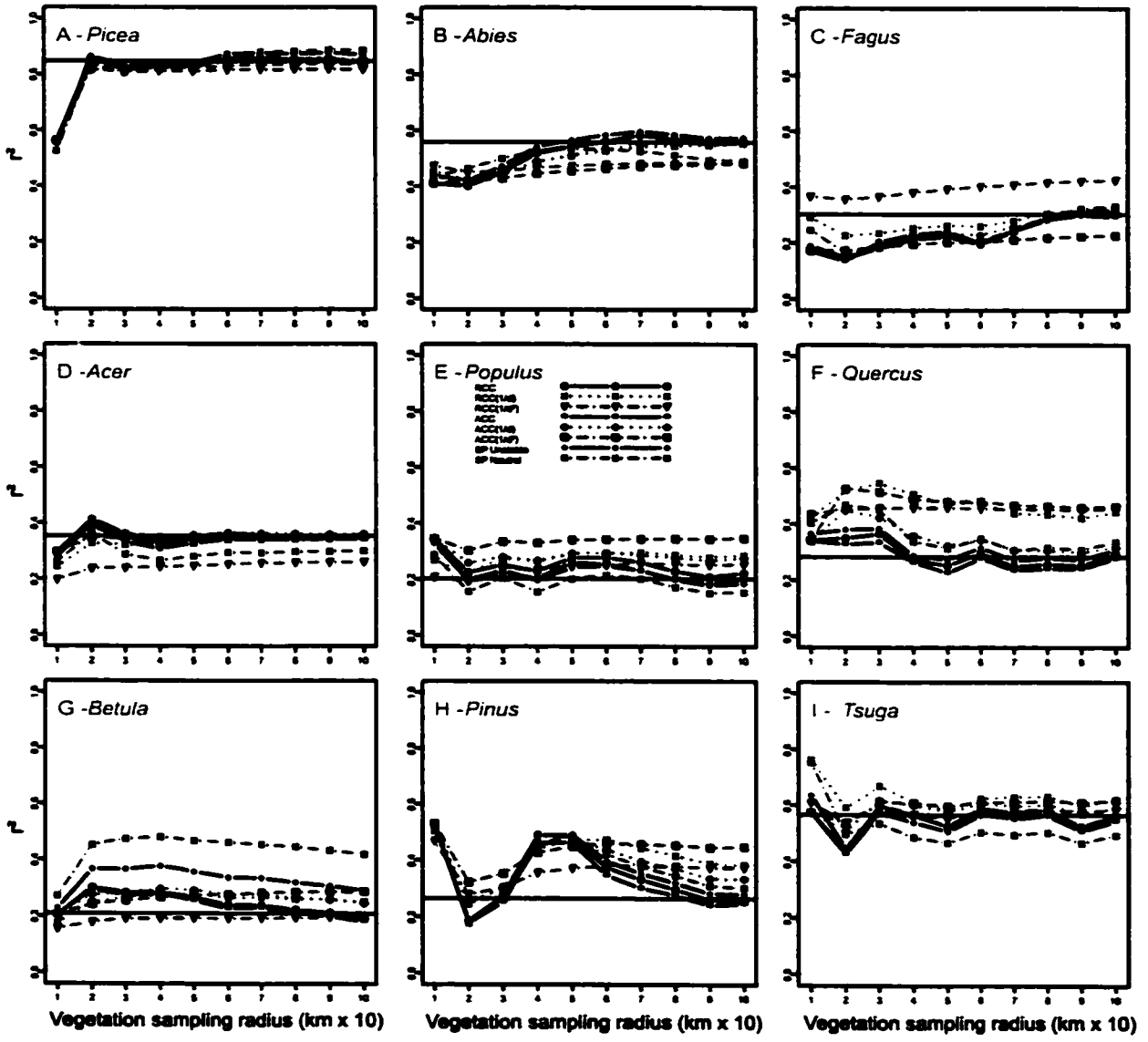


Figure 3.9: Plots of R^2 for each of 9 taxa against vegetation sampling radius for all weighting schemes. Straight line represents the average parameter value for all weighting schemes at all vegetation sampling radii derived from randomization of plant abundance quadrats within 100 km of each pollen site. Legend Abbreviations: RCC – relative areal crown cover unweighted, RCC($1/d$) inverse distance weighted and RCC($1/d^2$) inverse–squared distance weighted vegetation, ACC – absolute areal crown cover, equation (1) with $b = 0$ for unweighted, $b = 1$ for inverse distance weighted ACC($1/d$), and $b = 2$ for inverse–squared distance weighted ACC($1/d^2$) vegetation, SP – Sutton–Prentice dispersal weighting under neutral and unstable atmospheric conditions.

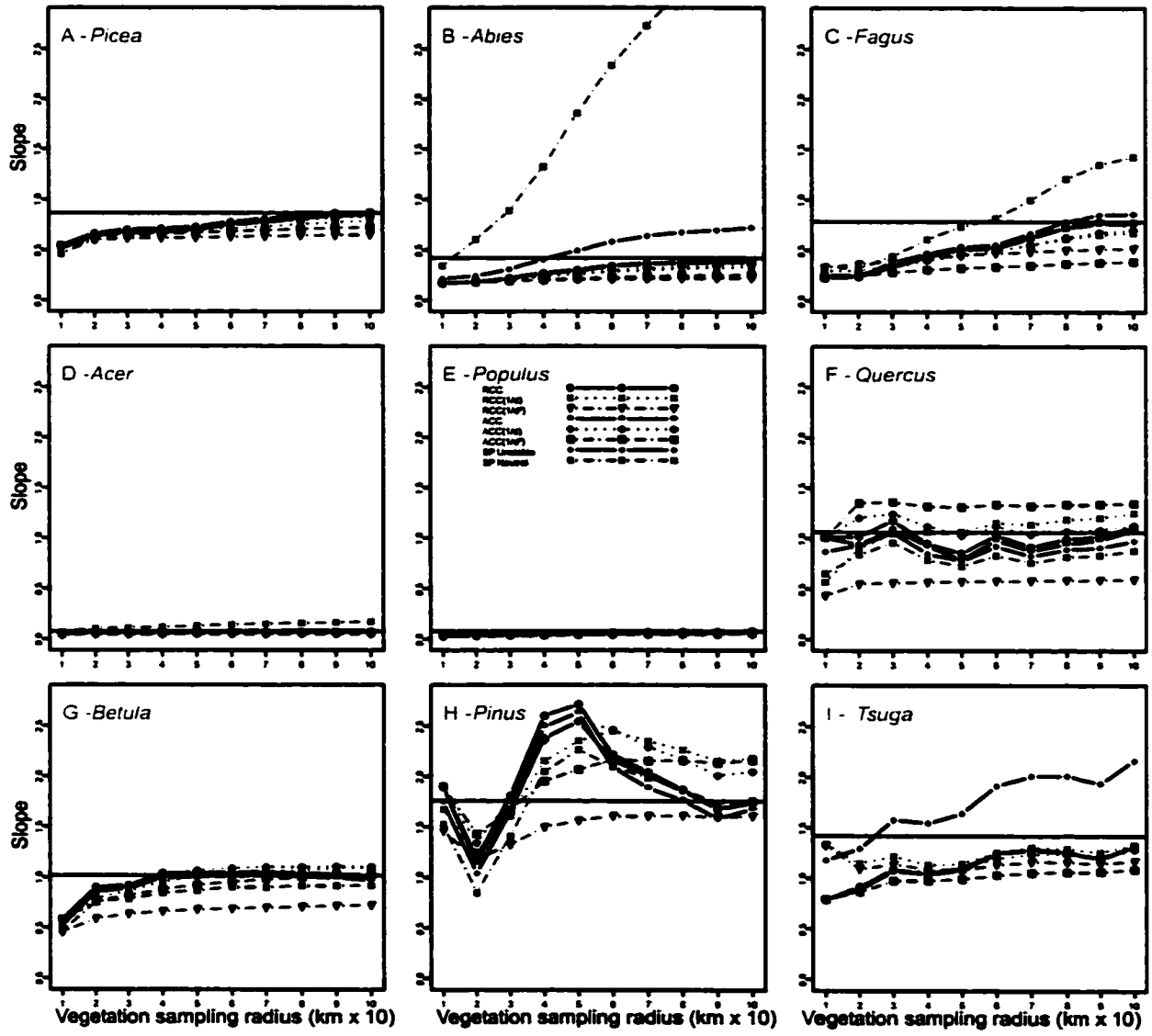


Figure 3.10: Plots of slope parameters for 9 taxa against vegetation sampling radius for all weighting schemes. *Straight line and legend abbreviations same as Figure 3.9.*

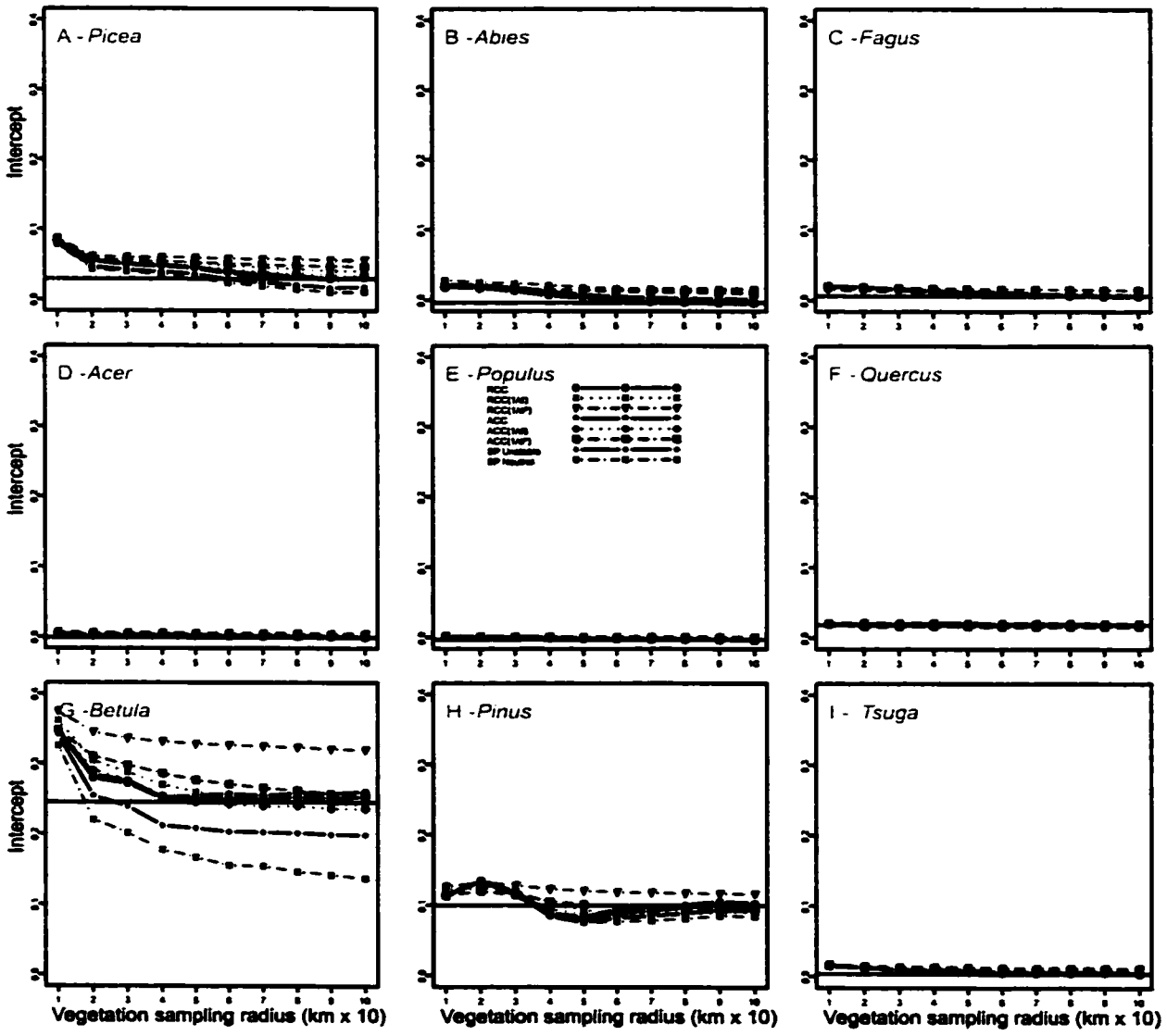


Figure 3.11. Plots of intercept parameters for 9 taxa against vegetation sampling radius for all weighting schemes. *Straight line and legend abbreviations same as Figure 3.9.*

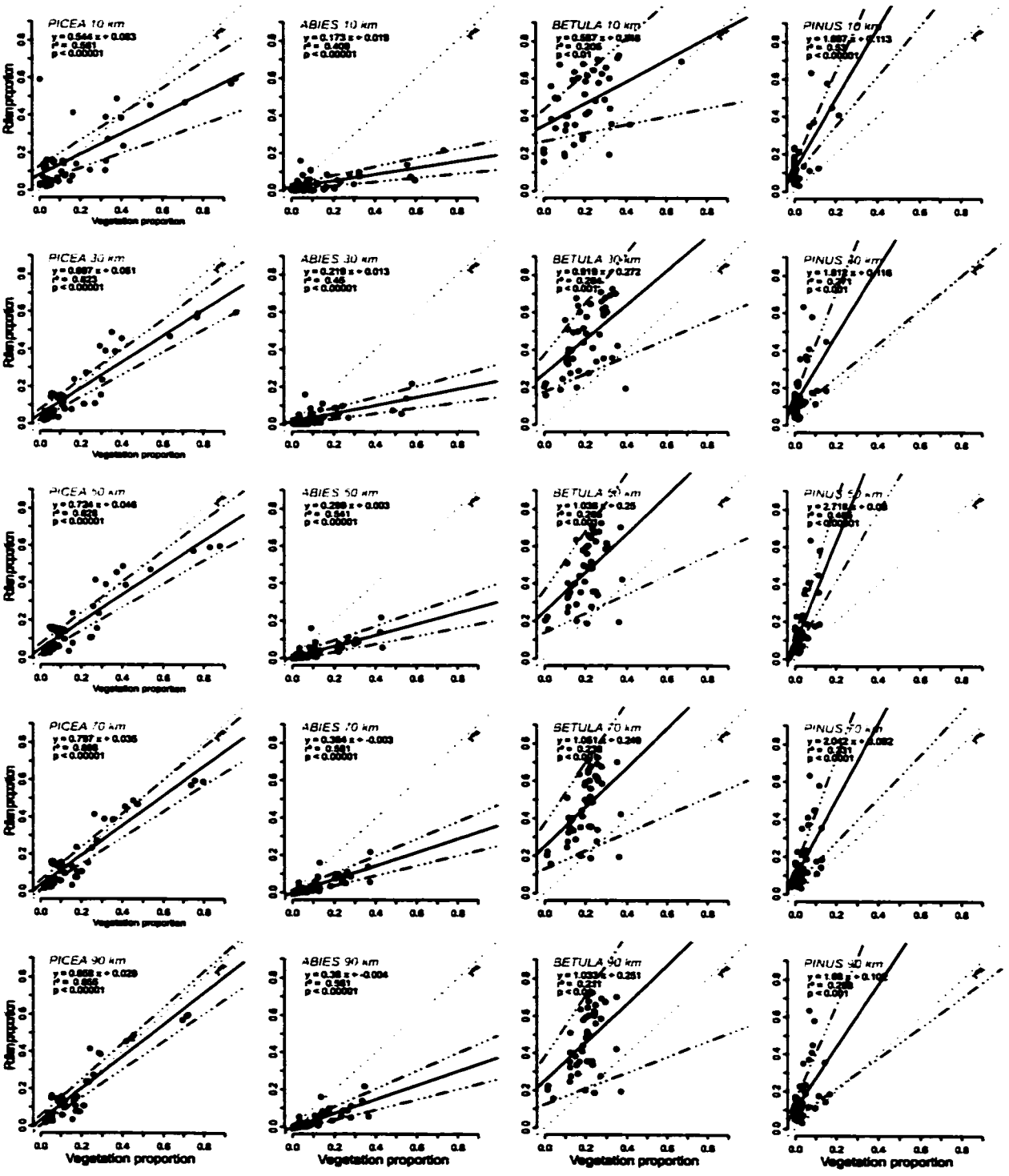


Figure 3.12: Scatterplots of pollen proportions and vegetation proportions using RCC vegetation at increasing vegetation–sampling radii. Fitted curve (solid line) is linear using equation $y = mx+b$ with parameters m and b determined using linear regression (OLS–y). Dashed–dotted lines denote the 95% confidence limits on this line. Dotted line represents $y=x$.

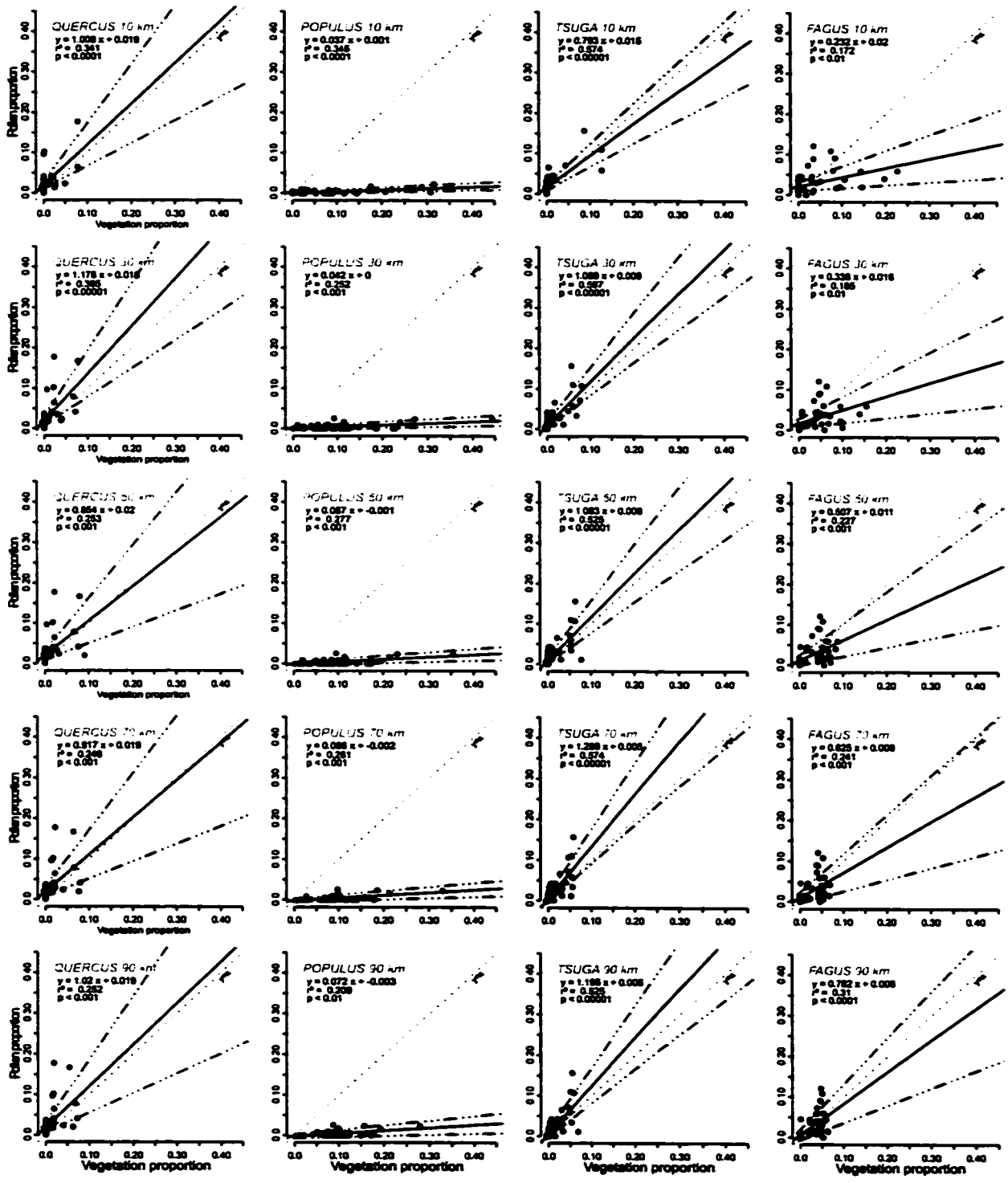


Figure 3.12. *Continued...*

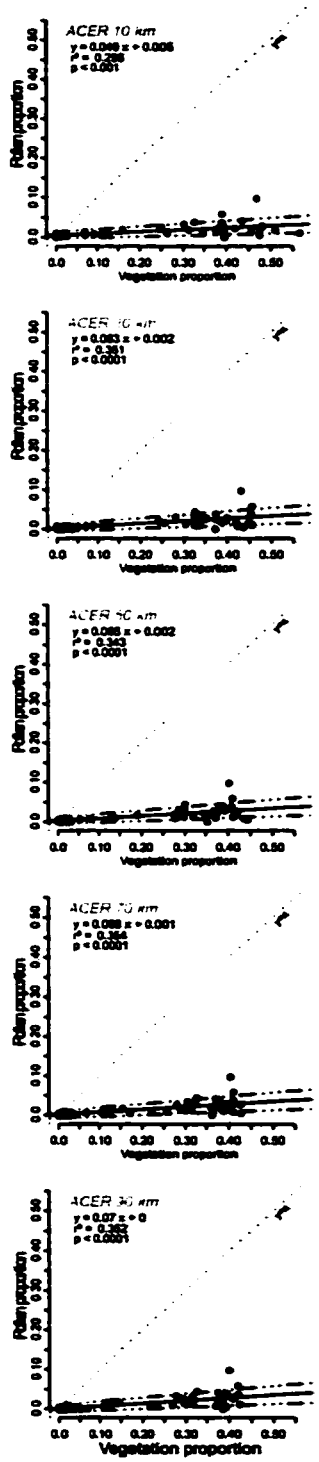


Figure 3.12. *Continued...*

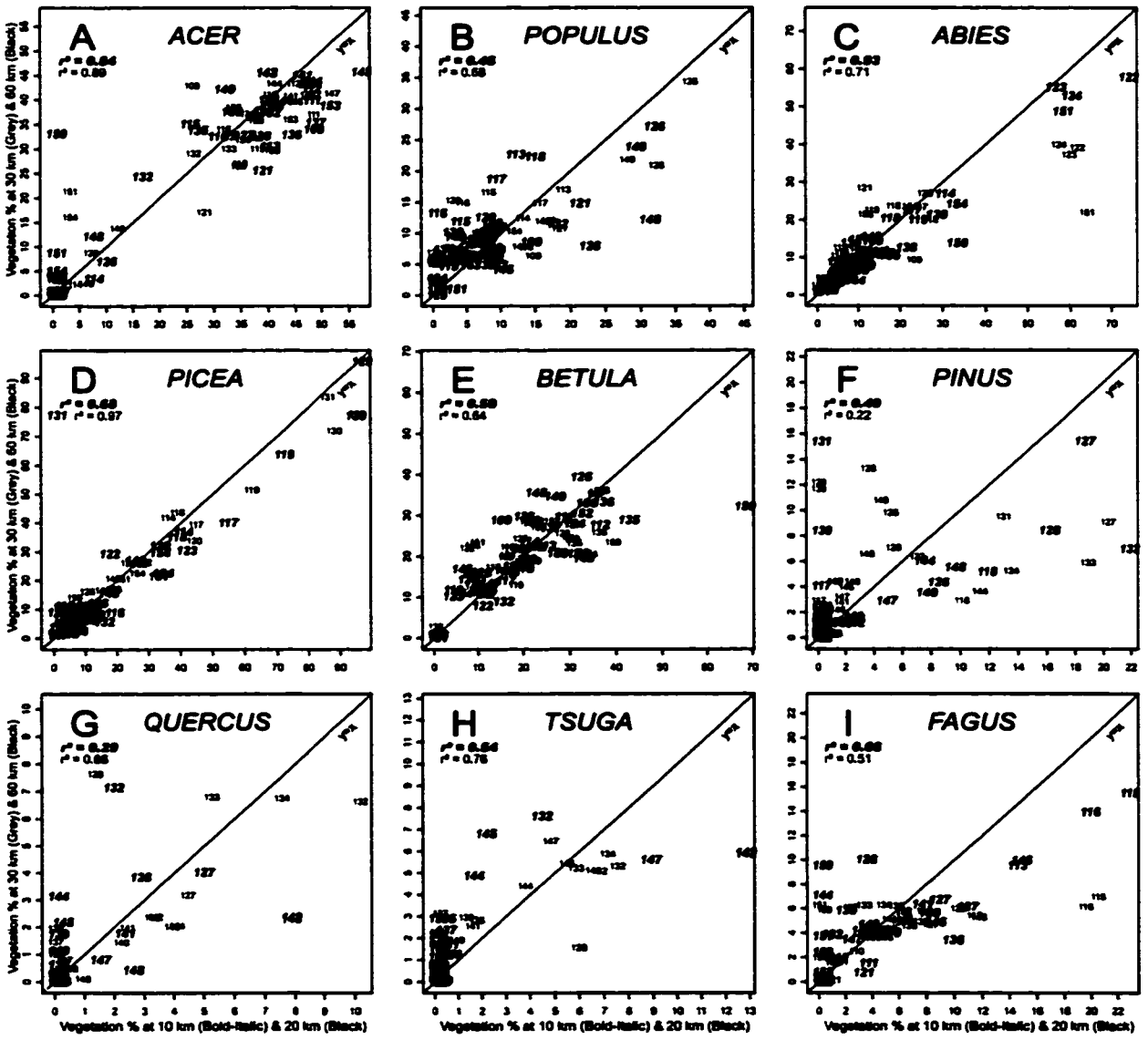


Figure 3.13: Patchiness of vegetation surrounding pollen–sampling sites as a function of two sampling radii. Site numbers correspond to those in Figure 3.1 and Table 3.1. Site numbers in the color gray indicate the degree to which the 10 km sampling radius represents distinct patches of a taxon within the 30 km sampling radius and site numbers in the color black indicate the 20 km sampling radius within the larger 60 km radial scale. *Note the different x,y scales used for each graph.*

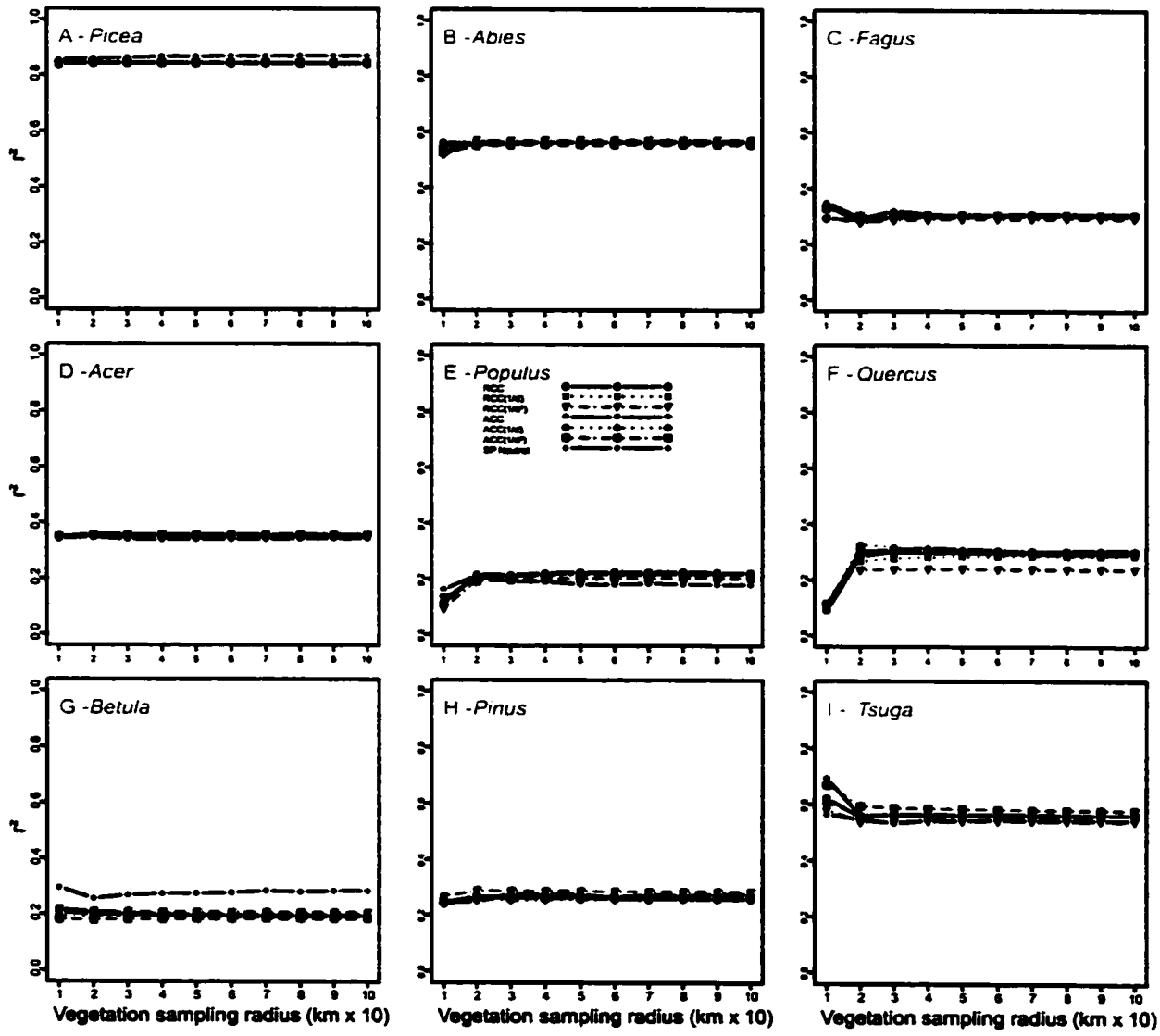


Figure 3.14: Plots of R^2 parameter under all weighting schemes for each taxa using randomized vegetation surrounding each pollen–sampling site. *Legend abbreviations same as Figure 3.9.*

Chapter 4

ASSESSING THE SUITABILITY OF A MODERN DATASET FOR PALEOCLIMATIC RECONSTRUCTION USING THE MODERN ANALOG TECHNIQUE (MAT)

4.1 INTRODUCTION

Techniques for quantitative paleoclimate reconstruction are numerous (Birks, 1995; Jacobson and Grimm, 1986) and the choice of technique often depends on the assumptions present regarding the particular type of proxy data in question. Holocene and full-glacial temperature and precipitation of several regions have been reconstructed using the modern-analog and plant-functional-type methods (Peyron *et al.*, 2000; Williams *et al.*, 2000; Prentice *et al.*, 1986). Alternative quantitative paleoclimate reconstruction methods such as response surfaces have been explored in eastern (Webb *et al.*, 1998) and western North America (Anderson *et al.*, 1991). The BIOME project has provided large-scale syntheses of present and past land-surface conditions that can also be used to evaluate climate model simulations (Prentice *et al.*, 2000; Jolly *et al.*, 1998). When large amounts of modern samples are available from regions representing the range of climates of interest, the modern analog technique (MAT) (Overpeck *et al.*, 1985) is appropriate and frequently used in paleoenvironmental reconstructions during both the Holocene (Sawada *et al.*, 1999; Anderson *et al.*, 1989) and for earlier periods (Dowsett and Robinson, 1997; Thompson *et al.* 1999). However there are a number of unanswered questions regarding this technique and how to assess the suitability of a modern dataset for use in paleoclimate reconstruction. This chapter explores the sensitivity of the climatic reconstruction to the various assumptions of the method.

Climate reconstructions based on fossil pollen records are calibrated using the relation between modern pollen and climate. For this, an extensive set of modern pollen samples spanning the range of modern vegetation and climate conditions are essential. In the past few decades great efforts have been made to assemble the available modern pollen data across North America. This modern pollen dataset permits the quantitative analysis of past climates using networks of fossil pollen sites. This chapter assesses the suitability of the current North American modern pollen database and provides the necessary assessment of where more modern data are needed to improve paleoclimate work.

Overpeck *et al.* (1985) concluded that the pollen set used is of little practical importance in the reconstruction of original pollen spectra downcore when using squared chord distance in the MAT. They suggested that a squared chord distance (SQD) with a value of 0.15 or less determined a poor vs. good analog. They based this finding on a pollen set of 15 types and explored this in two ways: using the latitudinal variability of squared chord distance between modern samples in the database and using the ability of the modern database to reconstruct pollen spectra downcore. Others have studied the SQD distribution and suggested different cutoff values of SQD to distinguish a good analogs from non-analogs based on zonal vegetation differentiation (Anderson *et al.*, 1989). However, all of these studies have used limited datasets and only explored specific regions of North America. The degree to which the pollen set affects the criteria for what may be considered a good analog and the degree to which different pollen sets affect the robustness of climate reconstructions using the MAT are explored in detail in this chapter.

Pollen spectra contain random noise and the degree to which this component of a pollen spectrum affects the choice of best modern analog has yet to be determined. Consequently, researchers often take the average climatic value of the top *n* best modern analogs (Sawada *et al.*, 1999). Sawada *et al.* (1997) examined the effects of the weighting of the top *n* best modern analogs using inverse Euclidean distance, inverse squared-chord distance, inverse rank-order distance and equal weighted averages on the climatic prediction from the MAT for a subset of fossil pollen sites in eastern North America. These decisions can

make critical differences in the climate estimate by as much as 2°C for a single site. Further work is necessary in order to determine the degree to which the weighting of the top *n* analogs affects the accuracy of the climate reconstruction. The new modern dataset provides the opportunity to determine the variation in reconstructed modern values using these various weighting schemes.

This study takes two complimentary approaches to determining the critical limits of a good analog. The first tests two pollen sets and their ability to reconstruct modern and 6ka temperature and precipitation in North America under the hypothesis that using those types best related to plant abundance (Chapter 3) will provide climate reconstructions that are more robust with respect to the assumptions and decisions in the MAT. The second is based on the ability of the pollen set used to differentiate between pollen from the same vegetation zone and those from different zones (Overpeck *et al.*, 1985; Anderson *et al.*, 1989). Here, the approach is further tested using Monte Carlo (Davison and Hinkley 1998) methods and stochastic simulation. Both approaches are shown to produce complimentary and converging results which provide confidence in the use of the modern dataset. Next, the question of the optimal pollen set to use is taken up again. Previous work has been based on regional studies- eastern or northeastern North America –for continental scale climate reconstruction a larger set may be more appropriate.

4.2 DATA AND METHODS

4.2.1 POLLEN AND CLIMATE DATA

Modern pollen data were assembled from 4590 sites across North America (Figure 4.11). This database represents the most extensive set of modern pollen samples in North America to date; the details of the construction of this new database and the taxa and variables contained therein are found in Appendix A.

Table 4.1: ***Large 89-type pollen set used for climate reconstruction***[***]

Taxon	Percent	Taxon	Percent
Pinus undiff.	19.80	*Sarcobatus*	0.05
Betula undiff.	11.76	*Ephedra*	0.05
Picea undiff.	9.97	*Cephalanthus*	0.05
Poaceae	6.71	Brassicaceae	0.04
Alnus undiff.	6.39	Ranunculaceae	0.04
Ambrosia	5.18	*Polygonum*	0.04
Quercus	5.07	Apiaceae	0.03
Cyperaceae	4.96	Ericaceae *Empetrum*	0.03
Artemisia	3.54	Caryophyllaceae	0.03
Chenopodiaceae/Amaranthaceae	3.00	*Thalictrum*	0.03
Sphagnum	2.95	Fabaceae	0.03
Tsuga	2.49	*Pteridium*	0.03
Asteraceae	1.75	*Celtis*	0.02
Salix	1.43	*Lycopodium annotinum* L.	0.02
Ericaceae undiff.	1.34	Rhamnaceae	0.02
Cupressaceae	1.32	*Cercocarpus*	0.02
Alnus crispa (Ait.) Pursh.	1.27	*Clethra*	0.01
Pinus strobus L.	1.21	*Picea glauca* (Moench) Voss.	0.01
Abies	1.17	*Dryas*	0.01
Ulmus	0.90	*Selaginella*	0.01
Acer	0.88	*Potentilla*	0.01
Larix/Pseudotsuga	0.67	Saxifragaceae	0.01
Fagus	0.65	*Eriogonum*	0.01
Fraxinus	0.55	Rubiaceae	0.01
Myrica	0.53	*Prosopis*	0.01
Ostrya/Carpinus	0.45	Euphorbiaceae	0.01
Populus	0.41	*Lycopodium selago* L.	0.01
Carya	0.35	Magnoliaceae	0.01
Polypodiaceae	0.34	Caprifoliaceae	0.01
Alnus rugosa (DuRoi) Spreng.	0.31	*Chrysolepis*	0.01
Lycopodium undiff.	0.29	*Shepherdia*	0.00
Rumex	0.26	*Larrea*	0.00
Corylus	0.24	Onagraceae	0.00
Picea mariana (Mill.) BSP.	0.13	Malvaceae	0.00
Juglans	0.12	*Lycopodium complanatum* L.	0.00
Plantaginaceae	0.11	*Arceuthobium*	0.00
Lycopodium clavatum L.	0.11	Cactaceae	0.00
Taxodium	0.11	Liliaceae	0.00
Nyssa	0.10	Ericaceae *Chamaedaphne/Ledum*	0.00
Liquidambar	0.09	Ericaceae *Vaccinium*	0.00
Aquifoliaceae	0.08	Campanulaceae	0.00
Oxyria	0.08	*Amorpha*	0.00
Tilia	0.08	*Ceanothus*	0.00
Platanus	0.07	Papaveraceae	0.00
Castanea	0.07		

To test the effects of the pollen set on the climate reconstructions using the MAT, pollen percentages were computed using modern pollen[**] sets: 1) a large set containing 89 pollen taxa (Table 4.1) and a smaller set

[***] *Abbreviation "undiff." means the pollen has not been identified to the species level and is therefore undifferentiated from a taxonomic viewpoint. Grains in the "undiff." category do not include differentiated grains.*

containing only 18 arboreal taxa (Table 4.2), most of which were found to have a good relation with plant abundance in Chapter 3 of this dissertation or in previous work. Note, however, that the determination of pollen taxa with good relation to plant abundance was based on a small area of southern Québec.

Modern climate data at each of the 4950 modern pollen sample sites and 768 fossil pollen sample sites were extracted from Leemans and Cramer (1991). These data consist of a 0.5 x 0.5 latitude/longitude grid for all the continental landmasses. The data were imported into a Geographic Information System (GIS), where the value of temperature and precipitation for each pollen site was extracted using nearest neighbor interpolation. That is, the climatic value of the nearest climate grid node was assigned to each modern pollen sample site.

Table 4.2: ***Small 18-type pollen set used for climate reconstructions.***

Taxon
Abies
Acer
Alnus crispa
Alnus rugosa
Alnus undiff.
Betula undiff.
Carya
Castanea
Fagus
Fraxinus
Picea glauca
Picea mariana
Picea undiff.
Pinus strobus
Pinus undiff.
Quercus
Tsuga
Ulmus
OTS (other trees and shrubs)

** *A modern pollen set consists of pollen spectra sampled from surface sediments in a lake or otherwise in a fossil core where the age of the spectrum is approximately 0 years before present.*

A third data set included fossil pollen samples for 6ka. In order to reconstruct the climate at 6ka in North America, calendar year dating control on all fossil sites in the North American Pollen Database (NAPD) (Grimm, 2000) was required. The age-depth relations of all sites in the NAPD were remodeled using linear age-depth models based on the authors preferred radiocarbon chronologies, the details of which are given in (Appendix B). Over 98% of the sites in the NAPD contain linear age-depth models, a few in calendar years but the majority in ^{14}C years BP. Age-depth models were derived using the most recent radiocarbon calibration curve (INTCAL 98; Stuiver and Reimer, 2000). New age-depth models were computed and new chronologies in calendar years before present (ka) were derived for each of the 768 fossil sample sites. The calibration program code used can be found in Appendix C.

Anomaly maps of total annual precipitation and July temperatures were created for modern (predicted – actual) and fossil (6 ka – 0ka) time periods. Anomaly maps show the deviation of the predicted climate from the contemporary climate at each site, and allow the identification of regions where temperatures/precipitations were greater or less than present for 6ka. In the modern vs. modern reconstructions, large regions of positive or negative anomalies allow the determination of bias in the ability of the modern pollen database to distinguish similar climatic regimes.

4.2.2 Effect of the Pollen Set on the Critical Limits of a Good Modern Analog

A good analog must be one that is unlikely to occur by chance alone, in other words, one which is significant in a statistical sense. Determining if an observed coefficient is likely to be observed by chance requires determining the empirical distribution that governs the range of variation of the coefficient under randomization.

The most common method used to determine analogs is the squared chord distance. The squared chord distance, a signal-to-noise measure, has been shown to provide the best reconstructions of original pollen spectra (Prentice, 1980: Overpeck *et al.*, 1985) and has been used in other studies from different environments (Dowsett and Robinson, 1997; Thompson *et al.*, 1999; de Vernal *et al.*, 1996, 1994). The

distributional properties of the squared chord distance coefficient under complete randomization have not been detailed. It is hypothesized here that the distribution of the coefficient is partly a function of pollen set utilized for a particular dataset.

Because a given modern pollen or other proxy database will always contain a limited number of observations, the true distribution of a dissimilarity measure may not be observed when using the database itself. Consequently, stochastic simulation techniques are appropriate to develop an empirical distribution of possible outcomes that may be expected by chance alone. The squared chord distance coefficient can take on values between 0 and 2 (Prentice, 1980), however, what is the most likely value of dissimilarity between two randomly chosen pollen spectra? Does the number of types in the pollen set affect this expected value? How does this relate to the critical limits of a good analog?

Two approaches are used in this analysis to explore the effect of the pollen set on the critical limits of dissimilarity between two pollen spectra: 1) stochastic simulation and 2) within vs. between-zone differentiation in the modern pollen database tested using a Monte Carlo approach.

At which point does adding additional pollen taxa to the set not make a difference in the expected value of the dissimilarity coefficient? The expected value of squared chord distance as a function of the number of types in the pollen set has not been systematically explored. To explore this, 3000 pollen spectra were created. Each spectrum was composed of n types, where n ranged from 2 types to 100, and where the proportion of each type was determined by chance using a uniform random number generator. One random spectrum was then compared to another random spectrum and the squared chord distance was recorded. This was repeated 3000 times for each set of n types, the distribution of squared chord distance coefficients were plotted and descriptive statistics derived. This simulation provides the empirical distributions of squared chord distance as a function of the number of types in the pollen set. In other words, the expected value is derived for squared chord distance between pollen spectra of n types that differ randomly.

In reality, large pollen sets have many taxa, some of these taxa will be present in some geographic regions and not in others. Thus, large pollen sets frequently have numerous taxa with zero counts. The median number of zero counts across all 4590 modern sites for the expanded set of 89 types is 75 and for the reduced set of 18 types is 13. If two spectra are compared from regions of very different vegetation, then the zeros or positive observations in one region will only add to the dissimilarity and such sites would not be chosen as best analogs. However, even similar regions can have zero observations for one or more types. Consequently, in larger pollen sets the probability that any given taxon will take on a value of zero (because it is not present in the vegetation near the pollen sample site) increases.

4.2.3 Within vs. Between-Vegetation Zone Differentiation and the Critical Limits of a Good Modern Analog

The utility of the modern pollen sample database for purposes of quantitative paleoclimate reconstructions will partly depend on the ability of the pollen spectra to distinguish between the different vegetation zones in North America and the ability of the modern pollen database to reconstruct modern temperature and precipitation.

Each modern pollen sample site was assigned a vegetation zone from a world vegetation map by Fedorova and Volkova (1990). This map was then sampled to a grid of 1.5 km within North America and the frequency of vegetation classes for all grid cells within 30 km of each modern pollen sample site were sampled. The most frequent (modal) class was considered the vegetation zone for each site. This map was chosen over alternative vegetation classifications because it presents the natural primary vegetation cover utilizing the structural and floristic characteristics of vegetation and eco-geographic criteria (Fedorova and Volkova, 1990). It should be the natural vegetation patches within the landscape mosaic where most of the modern surface pollen samples are collected. This map is more appropriate than some of the alternative vegetation classifications based on Advanced Very High Resolution Radar (AVHRR) derived imagery (Loveland *et al.*, 2000) which present significant regions of North America as cropland. All tundra classes were aggregated before analysis.

Pollen spectra within a given vegetation zone (within-zone) should be more similar to each other than to those of other vegetation zones (between-zone) in North America. The within-zone average squared chord distance should be significantly different from the between-zone average squared chord distance. To test this, the distribution of squared chord distance (Overpeck *et al.*, 1985) for all pollen samples within and between zones was calculated. Within a given zone (e.g., Forest-tundra vs. Forest-tundra) the second best modern analog for each sample was recorded (within a single zone, the best modern analog is between a site and itself so needs to be excluded). Between zones, the best modern analogs for a given zone (e.g., Forest-tundra) were selected from all other zones (Grasslands, Tundra, Deserts, etc...) excluding from the selection process those samples for the zone being considered.

The degree to which inappropriate vegetation zonation in North America will affect these zonal comparisons is explored via Monte-Carlo simulation (Davison and Hinkley, 1998). A random sample of 100 modern sites was taken from the set of 4590 samples in North America using a discrete random number generator following a discrete uniform distribution. This random sample was considered a simulated vegetation zone. It then had the squared chord distance of its second best modern analog for within the "random" zone and the best modern analog for between the "random" zone and all other zones recorded. This simulation was repeated 100 times. Since the zone is chosen randomly from a large sample there is a greater chance that a modern site will find its best modern analog outside of this zone. Therefore the between-zone average squared chord distance should be significantly smaller than the within-zone. Moreover, the simulated average within-zone squared chord distance should approximate the expected value of SQD derived in the randomization trials above in Section 4.2.2. The simulated average SQD values were then compared to the actual within vs. between-zone comparisons.

4.2.4 Ability of the Database to Represent the Range of Modern Climates in North America

If the modern database is to be suitable for reconstructing modern climate, assemblages must span the modern climatic range. The modern climatic range must also encompass the possible range of climate

values that may have occurred in the past and the relation between the proxy indicator and its source must remain constant through time (at the time-scale of interest). The ability of the modern database to reconstruct modern temperatures is a function of: 1) the degree to which the modern pollen site network represents modern climate space; and 2) the degree to which pollen spectra are capable of reconstructing this modern climate space. The degree to which the modern climate is represented by the modern pollen sampling network will affect the ability to produce accurate paleoclimates reconstructions.

The ability of the site network and pollen set to distinguish the climate of the network is an inherent limitation of the database, however, climate estimates from the MAT can differ as a function of the weighting of the top *n* best modern analogs and this requires elaboration.

To determine which climates in North America are capable of being reproduced by the modern database, modern climate values of July temperature and total annual precipitation were extracted at each of the modern pollen sample sites. The locations of these climatic values along the axes of July temperature and total annual precipitation with respect to all possible values in North America north of Mexico indicates those climates that are capable of being adequately represented by the modern pollen dataset and specifies where more data is needed.

If the modern pollen database is to be suitable for paleoclimate reconstruction it must adequately reconstruct the modern climate. For each pollen spectrum, the second best modern analog in the database was determined and the ability of the modern pollen spectra to reconstruct July temperature and total annual precipitation was examined. This was repeated for both the large and small pollen sets to determine if one performed better than the other.

To further assess the ability of the modern database to reproduce modern climate it can be hypothesized that anomalies (PREDICTED – ACTUAL) for July temperature should be normally distributed and lack

significant spatial structure. The existence of spatial structure in the anomalies would suggest that the ability of the modern pollen database to reconstruct modern climatic parameters is regionally or globally biased. This is the spatial analysis of the search for lack of structure in the residuals of a regressions analysis.

The test for spatial structure in the modern temperature anomalies was based on Moran's *I*, a standard measure of spatial autocorrelation. Spatial autocorrelation deals simultaneously with both location and attribute information (Goodchild, 1987; Griffith, 1987; Cliff and Ord, 1973). A lack of spatial autocorrelation should be found in a mapped pattern that does not significantly deviate from a map where each value x_{ij} was assigned randomly with equal probability to each location (i,j) on the map. Moran (1950) introduced the first measure of spatial autocorrelation in order to study stochastic phenomena that are distributed in space in two or more dimensions. Moran's *I* has been subsequently used in almost all studies employing spatial autocorrelation (reviewed by Upton and Fingleton, 1985). It is analogous to the conventional correlation coefficient because its numerator is a product moment coefficient (Sokal and Oden, 1978) and ranges from +1 (strong positive spatial autocorrelation) to $-1/(n-1)$ (signifying a pattern with values of a variable assigned randomly to each location) and finally to −1 (indicating strong negative spatial autocorrelation). The *I* statistic can be applied to ordinal, interval or ratio level data. Moran's *I* is given as,

$$I = \frac{n}{\sum_{i=1}^{i=n}\sum_{j=1}^{j=n} W_{ij}} \bullet \frac{\sum_{i=1}^{i=n}\sum_{j=1}^{j=n} W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{i=n}(x_i - \bar{x})^2} \qquad (4.1)$$

which is the product of the distance of the value of x_i at location i and x_j at location j from the mean. The denominator $\sum_{i=1}^{i=n}(x_i-\bar{x})^2$ will be undefined if all x_i are equal to the mean of the dataset, as would be the case in a field of constant value. The only case where Moran's *I* equals zero is when $(x_i - \bar{x})(x_j - \bar{x})$ equals zero. The null hypothesis is that there is no spatial autocorrelation in the modern vs. modern anomaly

dataset for both temperature and precipitation. The program used for this analysis is ROOKCASE (Appendix D; Sawada, 1999).

Because spatial dependence can vary in a complex manner across space, spatial autocorrelograms provide a means of measuring spatial dependence between regions or points separated by arbitrary distances. Spatial autocorrelograms are computed on a given dataset by using a number m of either mutually exclusive or cumulative distance classes. The correlogram consists of a set of correlated coefficients and any technique that assesses the significance of the correlogram must take into account the covariances between the correlogram coefficients (Oden, 1984). This means that each subsequent correlogram value is not independent of the previous, a requirement of standard statistical tests of significance.

The null hypothesis of spatial autocorrelation analysis is that a mapped pattern has its z-values assigned randomly across space. If 20 tests have been performed on the same network of sampling sites (i.e., a correlogram with 20 values) and 1 of these tests is significant at the 5% level, should the single significant value be considered a significant departure from randomness? The problem is that we might expect that 1 in 20 of the multiple tests from the same population would be significant just by chance alone. To remedy this question of multiple tests, before the individual correlogram coefficients can be considered significant it must be exhibited that the overall correlogram is significant (Legendre and Fortin 1989; Oden, 1984). This is most often accomplished by considering the smallest p-value in a spatial autocorrelogram and determining if this is less than a pre-specified level which is defined by Bonferroni (Cooper, 1968) or Sidak (Miller, 1977) bounds that are a function of the number of tests undertaken. The Bonferroni test is given by Cooper (1968) as:

$$\alpha' = \alpha / m \tag{4.2}$$

where α' is the Bonferroni corrected significance level, α is the desired significance level (the probability of rejecting the null hypothesis when it is actually true), and m is the number of tests undertaken. For

example, if 20 tests are undertaken and the desired significance level is $\alpha_{0.05}$, then at least one of the correlogram coefficients must be significant at the $p = \alpha' = {}^{0.05}/_{20} = 0.0025$. That is to say, at least one correlogram coefficient must have a p-value less than or equal to 0.0025 before the correlogram can be interpreted and individual coefficients considered significant at the accepted $\alpha_{0.05}$. Oden (1984) tested four correction procedures and found that the Sidak and Bonferroni methods are preferable to other tests when there are few distance classes and a weak spatial pattern. The differences between the modified *p*-values for the Sidak and Bonferroni are very slight and the latter test is adopted here because of its simplicity.

Local indicators of spatial autocorrelation (LISA) (Getis and Ord, 1992; Anselin, 1995; Ord and Getis, 1995) can be used to search for regional biases in the modern anomaly patterns in North America. LISA can be used to uncover spatially coherent regions of unusually similar anomalies in temperature or precipitation where the predictions are biased. Two important interpretations of LISA are:

1. A LISA provides an indication of the statistical significance of spatial clustering around a given location. However, the notion of clusters or 'hot spots' relates in the case of Local Moran's *I* to positive spatial autocorrelation while the notion of spatial outliers (non-clusters) regards negative spatial autocorrelation (Tiefelsdorf and Boots 1997; Anselin 1995).
2. The indication of pockets of local instability or non-stationarity that suggest outliers, which contribute more than the average share to the global statistic (Anselin 1995).

Globally there may be no spatial autocorrelation present within the mapped anomaly patterns, however, local pockets of positive association may exist that identify regions where the analog reconstruction is biased. Local Moran's I (I_i) was chosen because it can be used to decompose a global indicator in order to examine the contribution of each individual observation (Anselin 1995). Local Moran's I (I_i) is given as:

$$I_i = (x_i - \bar{x}) \sum_j w_{ij}(x_j - \bar{x}) \quad j \in J, w_{ii} = 0 \tag{4.3}$$

where J is the neighborhood of x_i, and w_{ij} is equal to 1 if x_i and x_j are adjacent by some measure of adjacency and zero otherwise. Local Moran's I_i provides a measure of the extent of significant spatial clustering of similar values around each observation in a dataset (Anselin, 1995). The expected value if I_i is given by Anselin (1995) and in another form by (Sokal *et al.*, 1998ab) and allow the computation of standard z-values and significance tests to be carried out. Because Local Moran's I is a LISA, the sum of all local I_i values for all observed anomalies are proportional to global Moran's I at the same lag distance,

$$\sum_{i=1}^{n} I_i(d) = \gamma I(d) \tag{4.4}$$

Which says that the sum of all of the n local indicators, I_i, at a given lag distance, d, are directly proportional to the global SA coefficient, I, at that same distance, with a factor of proportionality, γ. Thus, maps can be produced that allow the assessment of the contribution that each climatic anomaly value to the global statistic (Anselin, 1995; Tiefelsdorf and Boots, 1997, Sokal *et al.* 1998ab), thereby identifying regional biases in the ability of the modern pollen database to reproduce modern climates. Local values, which are very different from the mean, would indicate locations that contribute more than their expected share to the global statistic (outliers or high leverage points).

The degree to which climate estimates may differ as a function of the weighting of the top n best modern analogs is tested by comparing actual climate values at all modern sites to those derived from the best modern analog, and a set of weighted averages of the top five best modern analogs. The weighting of the top five best modern analogs is done using inverse Euclidean distance ($1/d_{ij}$ – where the d_{ij} represents the distance between site i and its j^{th} best modern analog), inverse squared-chord distance, inverse rank-order distance weights the estimated temperatures by the position of the nth best analog, e.g., 1/1, ½, 1/3, etc.., and equal weighted averages. These decisions can make critical differences in the climate estimate by as much as 2 °C for a single site. Further work is necessary in order to determine the degree to which the

weighting of the top n analogs affects the accuracy of the climate reconstruction. The new modern dataset provides the opportunity to determine the variation in reconstructed modern values using these various weighting schemes.

4.2.5 PALEOCLIMATE RECONSTRUCTIONS OF THE 6KA TIME SLICE

Fossil pollen sites for the 6ka time slice were extracted from the NAPD. All sites extracted had calendar year dates spanning the interval between 7 ka and 5 ka. A total of 501 fossil samples were available. Pollen proportions were linearly interpolated to the 6ka time slice and the resultant dataset used for reconstruction of the average July temperature and total annual precipitation in North America. The MAT (Appendix E) was applied to both pollen sets at 6ka and anomaly maps were created for average July temperature and total annual precipitation in order to determine the effect of the pollen set and interpolation method on the robustness of the climate reconstruction.

Vector maps were created that illustrate the distance and direction from each fossil site to its best modern analog at 6ka. Showing the distance and direction to the best modern analog can aid in the explanation of the regional anomalies.

4.3 RESULTS

4.3.1 EXPECTED VALUE OF SQUARED CHORD DISTANCE UNDER RANDOMIZATION AND THE EFFECT OF THE POLLEN SET

The number of types in the pollen set affects the expected value of squared chord distance under randomization (Figure 4.2ab). The distribution of squared chord distances begins to approximate a Gaussian distribution when the number of taxa in the pollen set is >20 (Figure 4.3). When each taxon has an equal probability of being zero or some positive value, the expected value of squared chord distance stabilizes rapidly to approximately 0.23 after a set of thirty taxa is reached. However, when the probability of any given taxon in the pollen set assuming a zero value increases, so does the expected value of squared chord distance (Figure 4.2b). This increase in expected value of squared chord distance is a linear function of the increasing probability of any one taxa in the set being zero.

4.3.2 WITHIN VS. BETWEEN-ZONE

For all vegetation zones, the within-zone variation in squared chord distance between a modern site and its second best modern analog is smaller than when a given zone finds its best modern analog outside of its zone (Figure 4.4 and Figure 4.5). This is true for both the 89-type (Figure 4.4) and 18-type (Figure 4.5) pollen sets. The distributions, however, are not normal and tend to be negatively skewed.

The average within-zone squared chord distance for the large pollen set of 89 types for all zones is larger than the average within-zone squared chord distance for the pollen set of 18 types (Table 4.3). The group mean over all zones between the 18 and 89-type pollen set are significantly different using a standard t-test.

This significant group difference is true for the minimum, maximum, median and standard deviations presented in Table 4.3. The relative variability in squared chord distance between zones, as measured by the coefficient of variation ($\bar{x}/\sqrt{\sigma}$), is greater for the 89-type set (Table 4.3). For some zones, specifically, the forest-tundra, hemixerophytic, northern taiga for the 89-type set, there were no significant differences observed between the mean values of the within vs. between-zone squared chord distances as measured by a standard t-test (Figure 4.4). The 18-type pollen set for these same zones as well as mountain vegetation, grasslands, deserts, central taiga, wet savannas, and semideserts had no statistically significant differences in the means between the within vs. between-zone squared chord distance distributions (Figure 4.5).

The Monte Carlo simulations of the within vs. between-zone variation in squared chord distance (Figure 4.6), suggest that the between-zone average squared chord distance is significantly smaller than the within-zone average squared chord distance (Figure 4.6) when pollen spectra are compared. Moreover, the average simulated within-zone squared chord distance for the 89-type pollen set approximated the expected value of the randomized trials (e.g., Figure 4.2 and Figure 4.3). The range of variation in squared chord distance for the 89-type pollen set is more than twice that of the 18-type pollen set (Figure 4.6)

Table 4.3: ***Summary statistics for comparison of squared chord distance of best modern analog***

		Within-zone vs. Between-zone												
		89-Type Pollen set						18-Type Pollen set						
Vegetation Zone	N	$\bar{x}_{89}$	SD	Median	Max.	Min.	$\bar{x}_{89}/SD$	$\bar{x}_{18}$	SD	Median	Max.	Min.	$\bar{x}_{89} - \bar{x}_{18}$	$\bar{x}_{18}/SD$
Central taiga	192	0.10	0.08	0.08	0.45	0.001	***0.79***	0.04	0.04	0.02	0.20	0.000	0.06	***0.92***
Deserts	408	0.06	0.06	0.05	0.43	0.004	***0.89***	0.00	0.01	0.00	0.17	0.000	0.06	***0.21***
Forest-steppes	296	0.08	0.07	0.07	0.36	0.000	***0.89***	0.02	0.02	0.02	0.19	0.000	0.05	***0.92***
Forest-tundra	130	0.06	0.06	0.05	0.52	0.014	***0.97***	0.02	0.04	0.01	0.36	0.000	0.04	***0.60***
Grasslands	222	0.08	0.07	0.07	0.76	0.000	***0.91***	0.02	0.02	0.01	0.14	0.000	0.07	***0.68***
Hemixerophytic forests	44	0.13	0.07	0.11	0.40	0.018	***0.56***	0.03	0.03	0.02	0.16	0.000	0.11	***1.00***
Humid forests	197	0.12	0.09	0.10	0.63	0.004	***0.75***	0.02	0.02	0.01	0.09	0.000	0.10	***1.13***
Mesophytic forests	598	0.09	0.08	0.08	0.70	0.000	***0.80***	0.04	0.03	0.03	0.34	0.000	0.06	***1.09***
Mountain	1027	0.09	0.10	0.07	1.11	0.005	***1.01***	0.02	0.03	0.01	0.41	0.000	0.07	***0.63***
Northern taiga	190	0.08	0.07	0.06	0.40	0.003	***0.91***	0.02	0.04	0.01	0.33	0.000	0.05	***0.62***
Semideserts	279	0.07	0.08	0.05	0.79	0.000	***1.17***	0.01	0.04	0.00	0.37	0.000	0.06	***0.40***
Southern taiga	248	0.09	0.06	0.08	0.52	0.001	***0.71***	0.04	0.04	0.03	0.40	0.000	0.05	***1.00***
Tundra	360	0.07	0.07	0.05	0.49	0.002	***0.96***	0.01	0.02	0.01	0.25	0.000	0.06	***0.63***
Wet savannas	95	0.09	0.06	0.07	0.39	0.028	***0.67***	0.00	0.01	0.00	0.06	0.000	0.09	***0.50***
CV		***0.23***	***0.15***	***0.29***	***0.37***	***1.471***		***0.54***	***0.39***	***0.73***	***0.48***	***0.001***	***0.29***	

		Within-zone vs. Between-zone												
Vegetation Zone	N	$\bar{x}_{89}$	SD	Median	Max.	Min.	$\bar{x}_{89}/SD$	Mean	SD	Median	Max.	Min.	$\bar{x}_{89} - \bar{x}_{18}$	$\bar{x}_{18}/SD$
Central taiga	192	0.13	0.09	0.11	0.51	0.014	***0.75***	0.04	0.03	0.04	0.24	0.001	0.09	***1.29***
Deserts	408	0.12	0.06	0.11	0.40	0.018	***0.50***	0.00	0.01	0.00	0.10	0.000	0.12	***0.29***
Forest-steppes	296	0.12	0.05	0.11	0.43	0.003	***0.44***	0.03	0.02	0.03	0.15	0.000	0.09	***1.61***
Forest-tundra	130	0.07	0.06	0.06	0.43	0.023	***0.76***	0.02	0.03	0.02	0.26	0.001	0.05	***0.78***
Grasslands	222	0.13	0.08	0.12	0.73	0.003	***0.61***	0.02	0.02	0.01	0.10	0.000	0.11	***0.89***
Hemixerophytic forests	44	0.14	0.05	0.13	0.32	0.062	***0.40***	0.02	0.01	0.02	0.04	0.006	0.12	***2.38***
Humid forests	197	0.19	0.10	0.17	0.75	0.008	***0.52***	0.03	0.02	0.02	0.16	0.000	0.16	***1.17***
Mesophytic forests	598	0.14	0.08	0.12	0.51	0.002	***0.54***	0.06	0.04	0.05	0.38	0.000	0.09	***1.45***
Mountain	1027	0.16	0.11	0.13	1.06	0.006	***0.70***	0.04	0.06	0.02	0.40	0.000	0.12	***0.73***
Northern taiga	190	0.08	0.05	0.07	0.40	0.006	***0.68***	0.02	0.02	0.02	0.13	0.000	0.06	***1.10***
Semideserts	279	0.12	0.08	0.10	0.40	0.015	***0.64***	0.02	0.02	0.01	0.13	0.000	0.10	***0.90***
Southern taiga	248	0.13	0.07	0.13	0.45	0.002	***0.50***	0.05	0.04	0.05	0.33	0.000	0.08	***1.43***
Tundra	360	0.12	0.08	0.10	0.53	0.007	***0.69***	0.02	0.01	0.02	0.08	0.000	0.10	***1.29***
Wet savannas	95	0.25	0.11	0.23	0.55	0.027	***0.44***	0.00	0.01	0.00	0.05	0.000	0.24	***0.50***
CV		***0.32***	***0.27***	***0.35***	***0.36***	***1.140***		***0.61***	***0.59***	***0.73***	***0.66***	***2.806***	***0.44***	

4.3.3 THE ABILITY OF THE MODERN POLLEN DATABASE TO REPRESENT THE RANGE OF MODERN CLIMATE IN NORTH AMERICA

Approximately 78% of the modern climate space in North America north of Mexico is represented by the modern pollen site network (Figure 4.7). The degree to which the modern pollen is representative of possible climates is a function of the geographic locations of the sample sites themselves. Our new modern pollen dataset (Figure 4.1) samples the majority of North American July temperatures. Sample coverage needs to be improved in areas of moderate summer temperatures (< 14°C) and high precipitation (> 900 mm y^{-1}) such as vegetation communities in the mountains of the west coast, and in areas of high summer temperatures (> 22°C) and low total annual precipitation (< 900 mm y^{-1}) such as in the southwestern United States.

The univariate distributions of modern temperature (Figure 4.8) and precipitation anomalies (Figure 4.9) (predicted – actual) for both the reduced and expanded pollen set are approximately normal. However, the anomalies produced by the pollen set of 89 types have smaller standard deviations (Figure 4.8) and a greater ability to reconstruct modern temperature and precipitation (Figure 4.8; Figure 4.9; Figure 4.10ab). For example, the intercept of the regression of modern actual July temperature vs. predicted July temperature based on the 89-type set is approximately 1°C (Figure 4.10a), whereas the 18-type set is 2°C (Figure 4.10b). This relation is also observed for precipitation estimates from both sets (Figure 4.10cd). The regression residuals of the 18-type set are not normally distributed (Figure 4.10a-c). There are no significant differences in the means of the anomalies (°C 89 types, $[\bar{x}_{predicted-actual}]$, vs. 18 $[\bar{x}_{predicted-actual}]$) for July temperature or precipitation as estimated between the two pollen sets as measured by a standard t-test.

The global spatial autocorrelogram of the modern July temperature anomalies for the 18-type pollen set exhibits significant positive spatial autocorrelation until a distance of between three to four hundred kilometers (Figure 4.11ab). The local contributions to this are strongly found in the southwest and Florida that show regions with the greater contribution to the global autocorrelation. For the 89-type pollen set and

throughout western North America there are small clusters of regions that contribute strongly to the global autocorrelation in the July temperature anomalies. Generally, the range of spatial autocorrelation and the consequent spatial bias in the ability of the modern pollen set to reconstruct modern temperature is on the order of 300-400 km in the 18-type set and much smaller if present in the 89-type pollen set. Both pollen sets reconstruct negative precipitation anomalies in Texas, southeastern Alaska and coastal British Columbia and around the James Bay in Northern Ontario. These regions are well sampled but the intersite distances are larger. There are few anomalies in the Great Lakes area where intersite distances are small and the region has a dense network of fossil pollen samples. These patterns are strongly reflected in the spatial analysis of the modern precipitation anomalies (Figure 4.12ab) for both pollen sets. Again however, the 89-type pollen set exhibits smaller anomalies and spatial autocorrelation to 600 kilometers whereas the 18-type pollen set exhibits far more anomalies for the same regions and these are greater in magnitude and spatially autocorrelated to a distance of 1700 kilometers.

Table 4.4 ***Correlation between actual July temperatures at 4050 modern pollen sites and those reconstructed from the 2nd best modern analog at those sites where the top 5 (two to six) best modern analog estimates are weighted in different ways (See text). All correlations are significant at p < 0.05.***

July Temperatures (Estimated – Actual)						
Weighting	Actual	Best Analog	Inverse Rank	Inverse SQD	Inverse Distance	Equal Weight
Actual	1.000	0.943	0.958	0.957	0.270	0.954
Best Analog	0.943	1.000	0.986	0.976	0.273	0.970
Inverse Rank	0.958	0.986	1.000	0.998	0.286	0.996
Inverse SQD	0.957	0.976	0.998	1.000	0.288	0.999
Inverse Distance	0.270	0.273	0.286	0.288	1.000	0.289
Equal Weight	0.954	0.970	0.996	0.999	0.289	1.000

Precipitation (Estimated – Actual)						
Weighting	Actual	Best Analog	Inverse Rank	Inverse SQD	Inverse Distance	Equal Weight
Actual	1.000	0.926	0.941	0.939	0.866	0.932
Best Analog	0.926	1.000	0.977	0.960	0.857	0.951
Inverse Rank	0.941	0.977	1.000	0.997	0.886	0.994
Inverse SQD	0.939	0.960	0.997	1.000	0.887	0.998
Inverse Distance	0.866	0.857	0.886	0.887	1.000	0.885
Equal Weight	0.932	0.951	0.994	0.998	0.885	1.000

The variation reconstructed by weighting the top *n* analogs is similar in all cases to the actual July temperatures at the modern pollen sites. (Table 4.4). Weighting the climate estimates by the inverse of the distance shows a noticeably different performance. Generally, the variability in the reconstructed values is similar among sites as evidenced by the strong correlations between weighted temperature and precipitation compared to actual values. However, the differences between the actual climate and those estimated using the different weighting schemes suggest that using the best analog leads to the highest variation in estimated values (Table 4.5).

Table 4.5: ***Summary statistics across all 4590 pollen sites for the anomalies between actual temperature and temperature estimated using the MAT under different weightings of the top n analogs.***

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD
Best Analog	-17.600	-0.200	0.000	0.032	0.200	19.700	3.932
Inverse Rank	-11.430	-0.463	0.000	0.029	0.429	20.200	2.832
Inverse SQD	-8.860	-0.473	0.000	0.032	0.435	20.240	2.860
Inverse Distance	-8.584	-0.300	0.000	0.047	0.258	20.230	2.325
Equal Weight	-8.800	-0.535	0.000	0.027	0.480	20.240	3.049

That some noise is introduced in the estimated values when only the best modern analog is selected is clearly evident from the greater variation in the anomalies using only the best modern analog when compared to inverse square chord distance. The regional expression of these differences is the focus of ongoing experimentation.

4.3.4 PALEOCLIMATE RECONSTRUCTIONS FOR THE 6KA TIME SLICE

The 6ka reconstructions exhibit positive temperature anomalies (Figure 4.13a) and little to no change in precipitation (Figure 4.14a) in the central and prairie regions of North America for the 89-type pollen set. However, this signal is ambiguous in the 18-type pollen set where a couple of sites in the prairies produce a strong cool center in southern Saskatchewan and southern Alberta (Figure 4.13b). These sites are finding their best modern analog north of their present position although they do not show high dissimilarity (Figure 4.15). Likewise a wetter and warmer Canadian Arctic archipelago and a cooler and drier Texas/New Mexico region are reconstructed for both pollen sets.

Generally, for the 89-type pollen set, sites in western North America find their best modern analogs along similar latitudes towards the east. In northeastern North America, sites find their best modern analogs in a southwards direction. Northern Quebec is distinct in the vector plots between the two pollen sets: the 18-type pollen set finds the best modern analogs in the western forest-tundra, while the 89-type pollen set differentiates between the eastern and western boreal forests and finds analogs southwards. Despite differences in where modern analogs are found there are similar temperature and precipitation anomalies in these regions

The distribution of anomalies is very different for both sets: the temperature (Figure 4.16a) and precipitation anomalies (Figure 4.17a) for the 89-type set are positively skewed whereas the opposite is true for the 18-type set (Figure 4.16b and Figure 4.17b).

4.4 DISCUSSION

The randomization experiment suggests that when there are even small proportions of zero values between two spectra, the expected value of squared chord distance increases. The expected value of squared chord distance under randomization is approximately 0.23 and is consistent with the average squared chord distance of the simulated within-zone mean value. This shows the convergence between the results of the stochastic simulation, Monte Carlo results and the observed between-zone squared chord distances within the modern pollen dataset.

The average squared chord distance of the best analog between and within vegetation zones illustrates that the smaller pollen set gives a smaller squared chord distance between a given site and its best modern analog. With a smaller pollen set, chances are less of having pollen types with a value of zero. The randomizations suggest that when there are even small proportions of zero's between two spectra, the expected value of squared chord distance under randomization increases. Moreover, the mean within-zone variation in the SQD of the best modern analogs is significantly less for the small pollen set. Consequently, the critical SQD value for what is considered a good analog is a function of the pollen set.

For the majority of vegetation zones (e.g., semideserts, northern taiga, grasslands, etc.) the small pollen set showed no significant differences between the mean of the within-zone vs. between-zone comparisons. This is due to the lack of arid, northern and western taxa within the 18-type pollen set when compared to the 89-type set. The 18-type pollen set represents mainly taxa found in eastern North America that have

significant relations to modern plant abundance. The exclusion of the other taxa from the reduced pollen set is largely a consequence of the lack of studies on pollen-plant abundance relations in other regions of North America. The ability of a large pollen set to differentiate between different vegetation assemblages is more effective. However, the critical value of SQD should be set higher in the case of a larger set, as the simulations and data have shown.

The search for modern analogs must take into account the effects of the pollen set and the range of vegetation and climates that may have been encountered in the past. If the modern database is not representative of the range of modern climates then there will be restrictions on the ability of the database to reconstruct past climates. The climate space diagrams suggest that the modern database represents most of the available climates in North America north of Mexico; however, very wet and dry regions need further sampling.

Because of the geographic sampling density of the modern network, in addition to the limitations this imposes on the ability of the database to reconstruct modern climate, the modern pollen database also exhibits regional biases in its ability to reconstruct modern temperatures in western and southwestern North America, and the northeastern coastal regions of British Columbia and Alaska. There is a smaller spatial bias using the 89-type pollen set and this is consistent with the univariate distribution of residuals. The univariate distribution of the temperature anomalies for the 89-type pollen set is narrower than the 18-type pollen set and this suggests that the 89-type set is more successful in reconstructing temperature in the modern database. Moreover, the regions that contribute positively to the spatial autocorrelation in the expanded set tend to be spatially random. Regions with large anomalies in the modern database are Texas, southeastern Alaska and the James Bay region. Although they contain sample sites, the intersite distances are on the order of 100 km or more as compared to the Great Lakes region (e.g., Figure 4.1), which exhibits few anomalies in temperature under both the large and small pollen set. It is in these regions of larger intersite distances where the contribution to the global spatial autocorrelation is significant. Consequently, the spatial biases in the modern vs. modern reconstruction of temperature are largely due to this intersite distance effect in the modern database.

The west is still a significant problem when reconstructing temperature and precipitation. If pollen spectra similar to western environments are encountered in the fossil record they will have fewer modern samples

from these regions to choose from and may show larger dissimilarities. However, the southwest, for the 18-type pollen set, contributed strongly to the global spatial autocorrelation of the modern anomalies. The greater intersite distances and the bias of the reduced set towards eastern deciduous taxa contribute to the poor performance of this pollen set. The improved performance of the 89-type pollen set is clearly evident in the smaller range of spatial autocorrelation, narrower range of anomalies for both climatic parameters and better overall relation between actual and reconstructed temperatures.

The southwest and west coast of North America is a region poorly represented in climate space by the modern database. Improved reconstructions and reductions in the spatial biases would be achieved by more modern samples in those regions. How these biases transfer to paleoclimates reconstructions would depend on the pollen set.

Boundary conditions at 6ka were similar to the present – no major ice-sheets were present and vegetation was essentially consistent with its modern configuration. However, paleoclimate records from many regions of the Northern Hemisphere suggest terrestrial surface conditions were warmer (Kutzbach and Ruddiman, 1993). By 6 ka, the Laurentide ice sheet had almost completely wasted (Kutzbach and Webb, 1993; Dyke and Prest, 1987) and CO_2 had returned to interglacial levels. However, solar radiation inputs at 6ka were about 6% greater than present, altering the seasonal balance of Northern Hemisphere insolation and producing the maximum warming in the continental interior of North American (Kutzbach and Ruddiman, 1993). Biome distributions at 6000 ka as reconstructed from fossil pollen indicate a warmer and drier western continent (Williams *et al.*, 2000).

Consistent with a warmer continental interior, the 6ka reconstructions exhibit positive temperature anomalies in the central states and southern provinces, more evident for the 89-type set than the 18-type set. There is little to no change in precipitation reconstructions. However, using the 18-type pollen set, two sites produce a strong cool center in southern Saskatchewan and southern Alberta. These sites are finding their best modern analog north of their present position. Consequently, the choice of pollen set can be critical in observing patterns of anomalies in regions of sparse data coverage such as in western North America. A cooler southern Texas and New Mexico are reconstructed because there are no modern pollen samples to the south from which to choose analogs from and so a warming cannot be reconstructed for this region. Likewise, warmer anomalies in the Canadian Arctic archipelago are due to the choice of modern

samples from the south for both pollen sets. Essentially there is an edge effect present in the reconstructions to the south and north of the continent that biases any reconstructions that may be made for the 6ka period.

In the expanded pollen set of 89-types, these sites find their best modern analog in the east. The 18-type pollen set produces overall very small values of SQD and the modern sites show low dissimilarity values and would by the definition of Overpeck *et al.* (1985) be considered good analogs. However, this chapter shows that the small dissimilarities are to be expected in small pollen sets within this database and the threshold of non-analog dissimilarity must be considered within the context of the size and nature of the pollen set utilized.

The differentiation in where the two different pollen sets find their best modern analogs is critical and reflected in the better performance of the 89-type pollen set in the modern vs. modern reconstructions and the within vs. between-zone comparisons: The 18-type pollen set has difficulty in distinguishing between the different vegetation zones, particularly those in the north and west, where the average squared chord distance within-zone is not significantly different from that between the zones.

4.5 CONCLUSIONS

The pollen set is a critical decision in the reconstruction of continental scale paleoclimates. Larger pollen sets can better distinguish between different vegetation zones in North America. The simulation results indicate that the critical values of squared chord distance must be adjusted for the larger pollen sets, and so the SQD limit of what may be considered a good analog increases with the number of pollen types. The spatial analysis of the modern temperature and precipitation anomalies have contributed to our understanding of the degree to which spatial biases are introduced into the MAT because of the varying density of the modern sampling network. They also identify areas that require more pollen samples, those same regions that lacked adequate representation in modern climate space. The climatic reconstructions for the 6ka time slice are consistent with model simulations for the continental interior. However, under two different pollen sets the spatial distribution of anomalies is variable. The application of the methodology presented in this chapter will help to improve paleoclimatic reconstructions through a thorough understanding of the limitations of the underlying database, whether it is a pollen based reconstruction or utilizing some other proxy data.

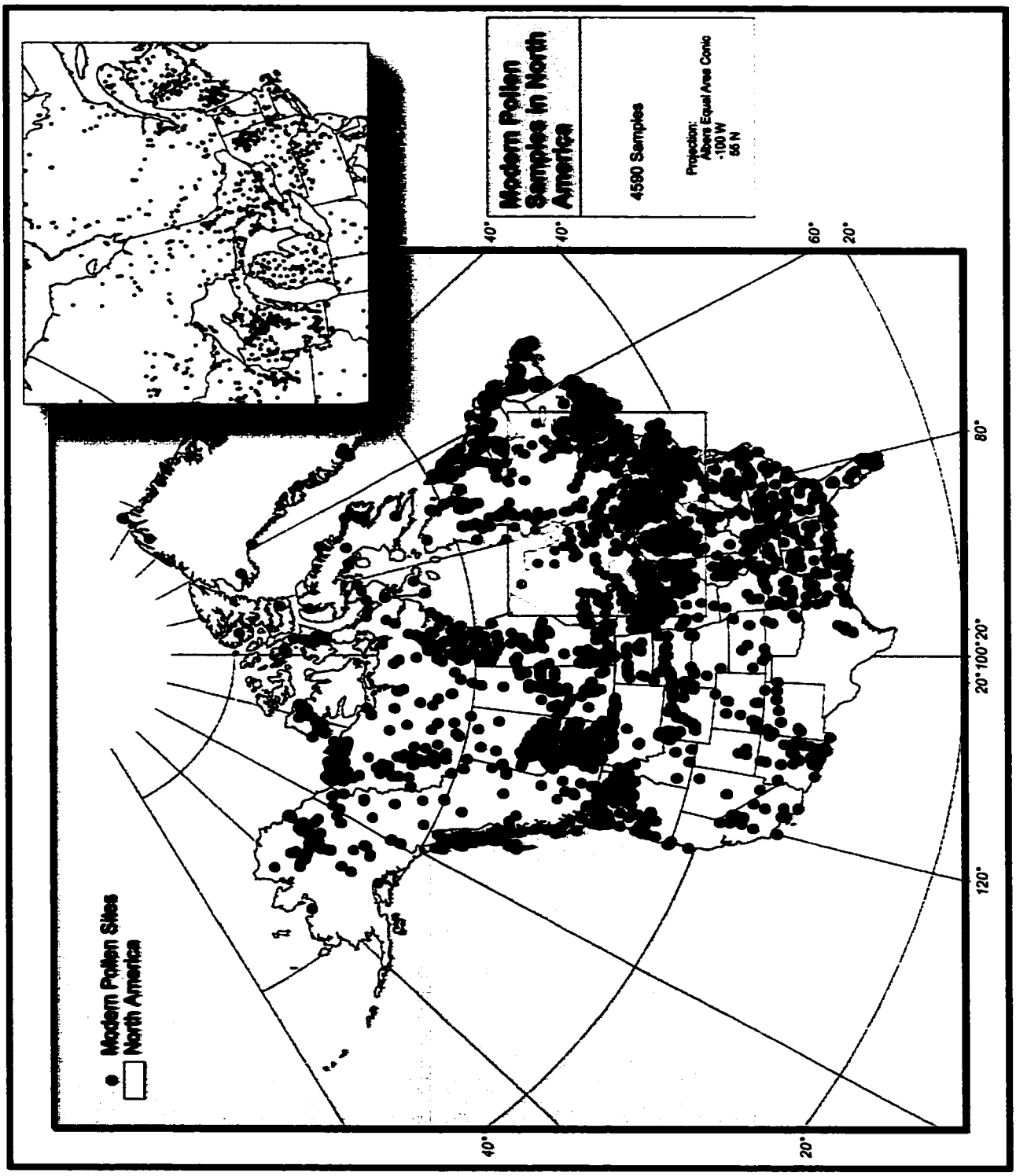


Figure 4.1: Distribution of modern pollen sample sites from the new modern pollen database in North America. ***Inlay***: Enhanced eastern North America where sample density is greatest.

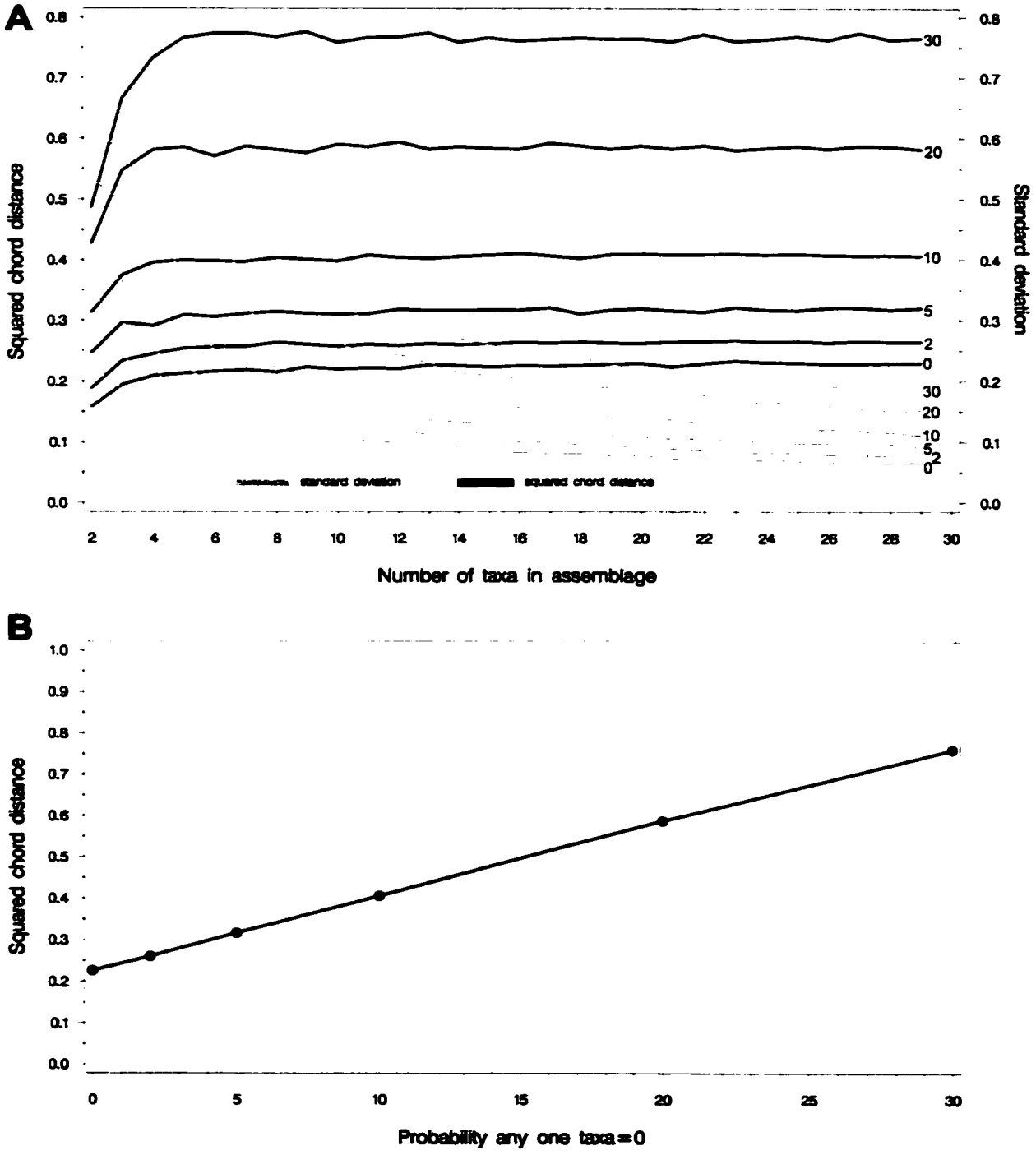


Figure 4.2: A – Black Lines: Mean value of squared chord distance (y–axis) between 3000 random pollen spectra as a function of the number of types in the pollen sum (x–axis). Numbers to the right of the curves indicate the probability for each pollen sum that a given taxa can assume a value of zero; Grey lines – standard deviations of Black Lines. **B** – Mean value of squared chord distance (y–axis) as a function of the probability that any given taxon can take a value of zero. Curve is for a pollen sum of 30 types.

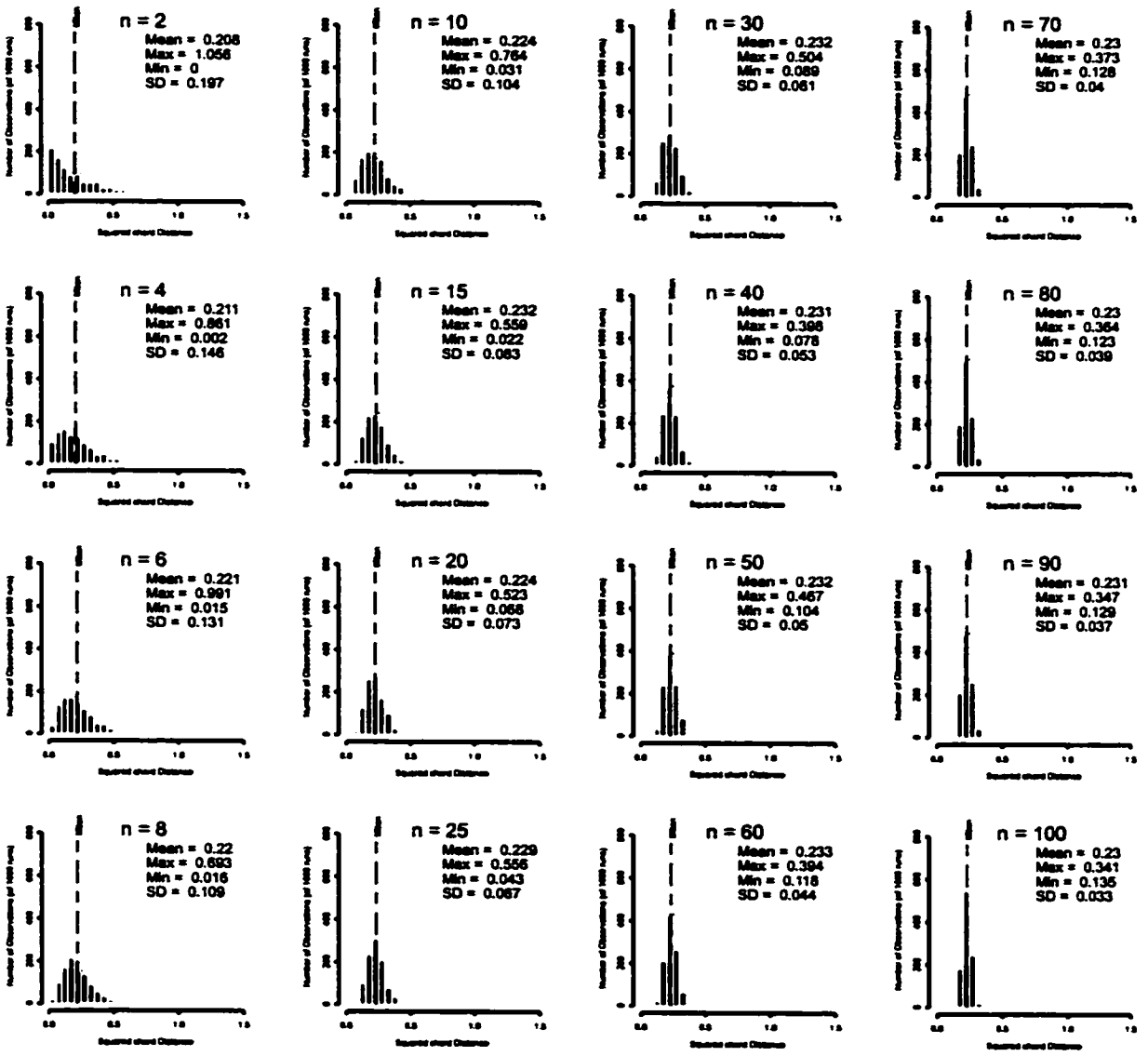


Figure 4.3: Sample of the distributions of squared chord distance between two random pollen spectra for 1000 runs for *n* taxa in pollen set.

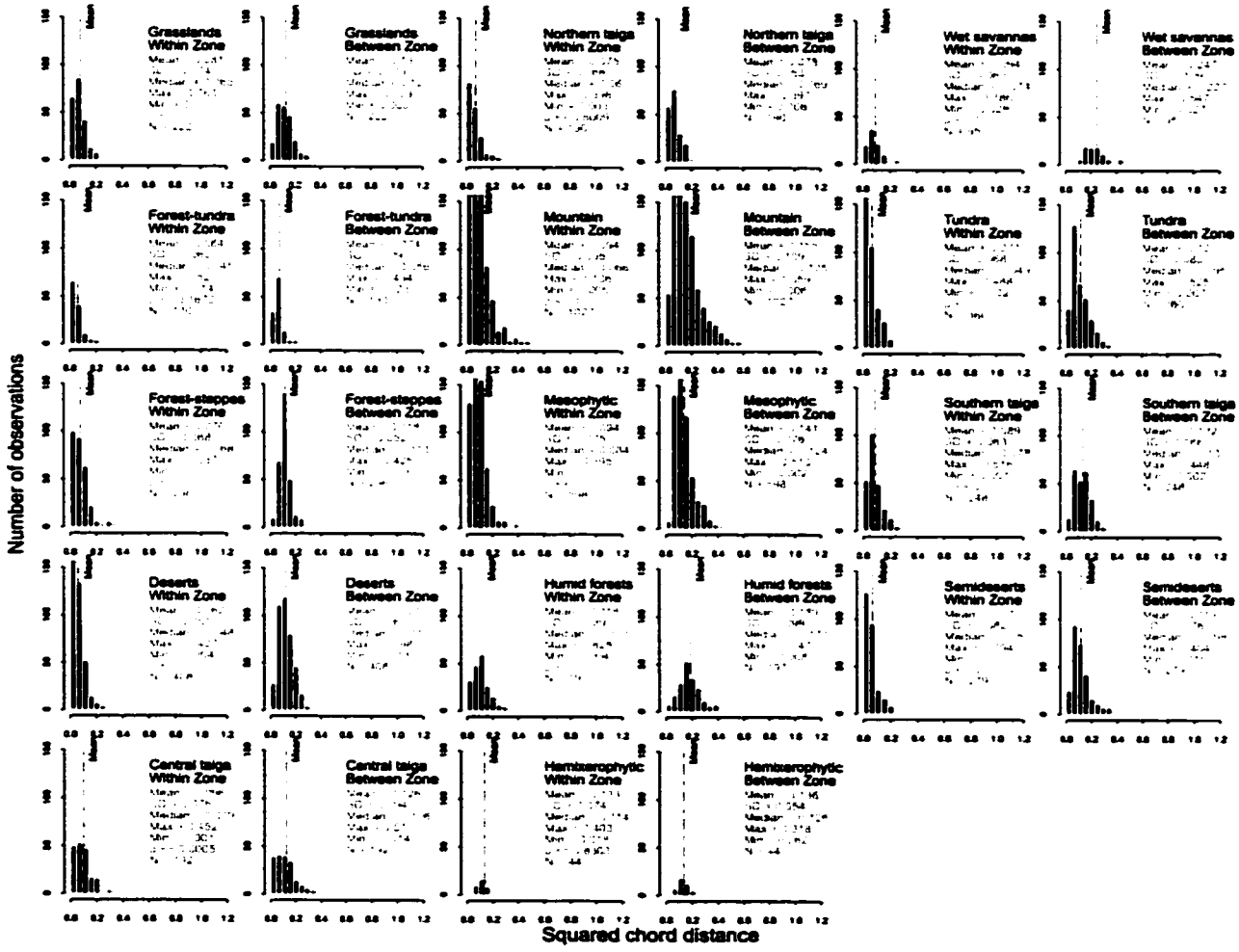


Figure 4.4: Squared chord distance for the best modern analog for the modern pollen dataset for the within–zone (2^{nd} best modern analog, see text) vs. between–zone (1^{st} best modern analog) pollen spectra comparisons with a pollen set of 89-types. The *p*–values reported for the within–zone represent the probability that the difference between the within–zone mean squared chord distance and between–zone mean squared chord distances are due to chance alone.

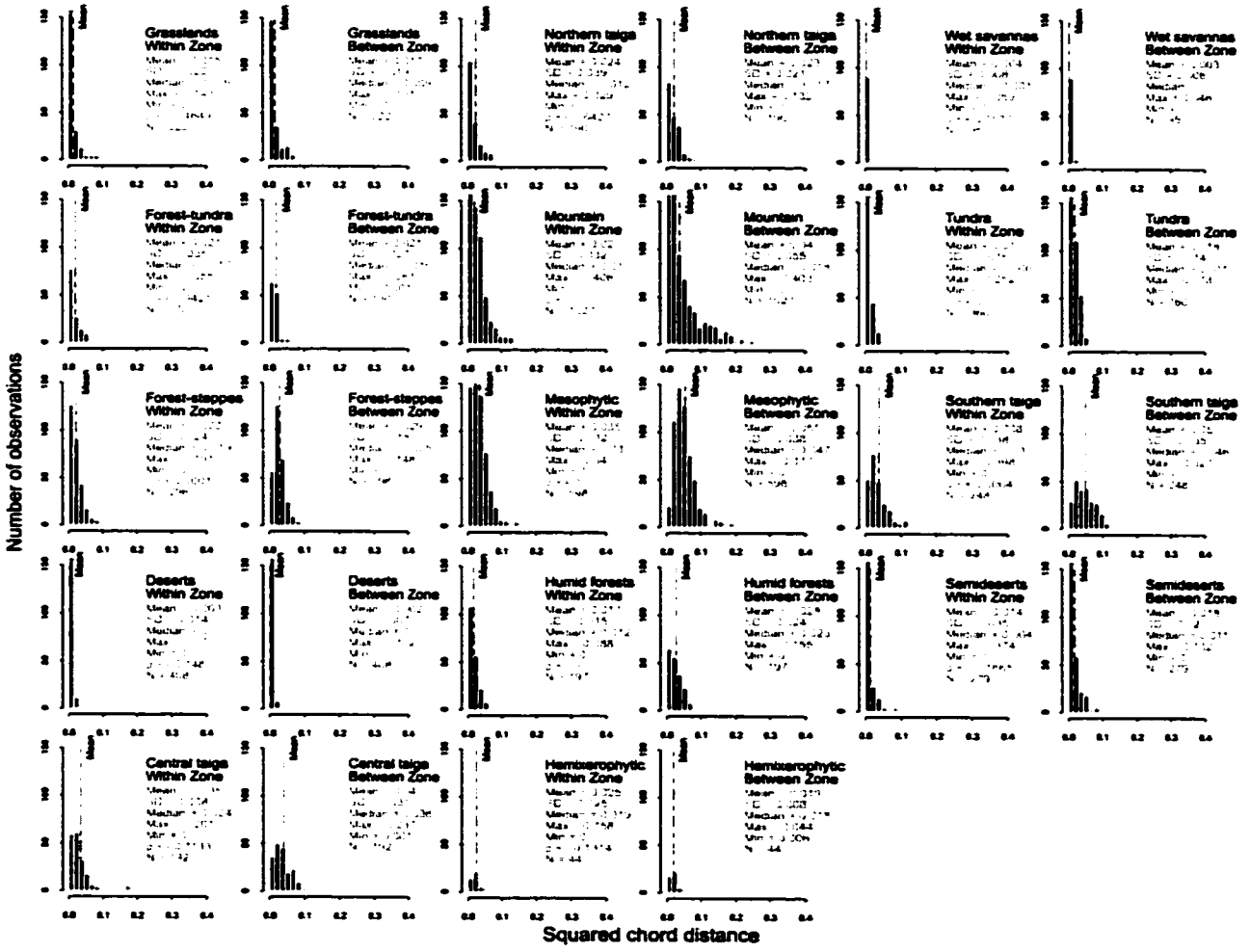


Figure 4.5: Same as Figure 4.4 but for the pollen set of 18-types.

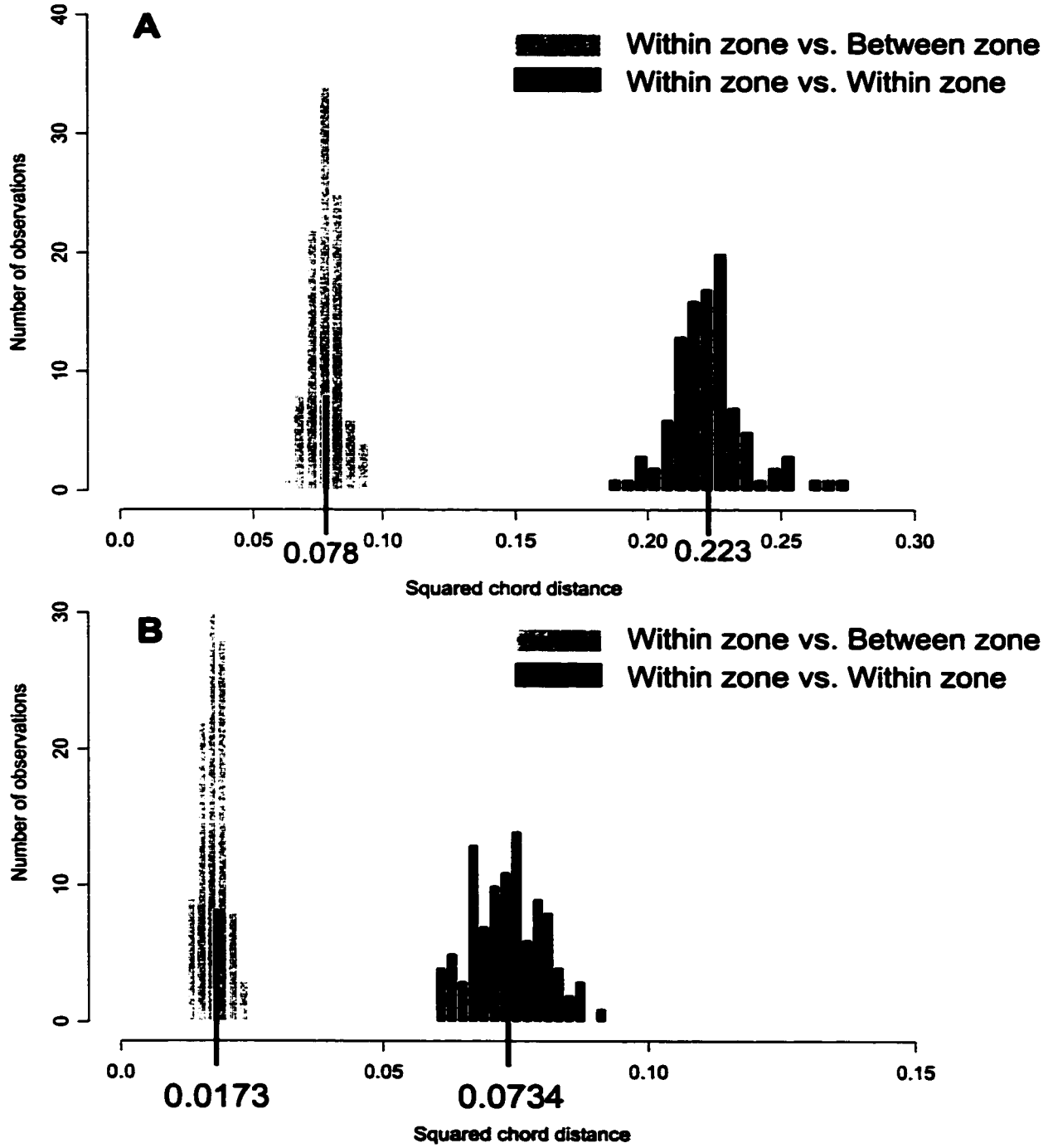


Figure 4.6. **A** – Simulated within vs. between–zone distribution of squared chord distance for the best modern analog for a random sample of 100 modern pollen sites from the modern pollen dataset for the within–zone (2nd best modern analog, see text) vs. between–zone (1st best modern analog) pollen spectra comparisons with a pollen set of 89-types. **B** – Same but for pollen set of 18-types. *Note the different x-axis scales between A & B.*

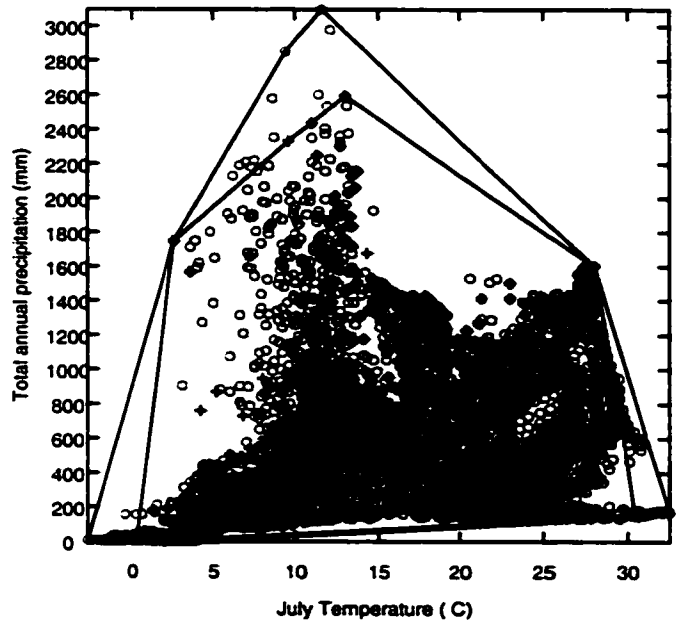


Figure 4.7: Degree of overlap between 4950 modern pollen sampling sites and North American climate space. Solid outer line is the convex hull containing North American climate; the shaded region is the convex hull for pollen sampling sites in North America. Small gray open circles are individual climate combinations from all grid points from Leemans and Cramer (1991) in North America north of Mexico. Black crosses are the climate combinations at the pollen sample locations in North America.

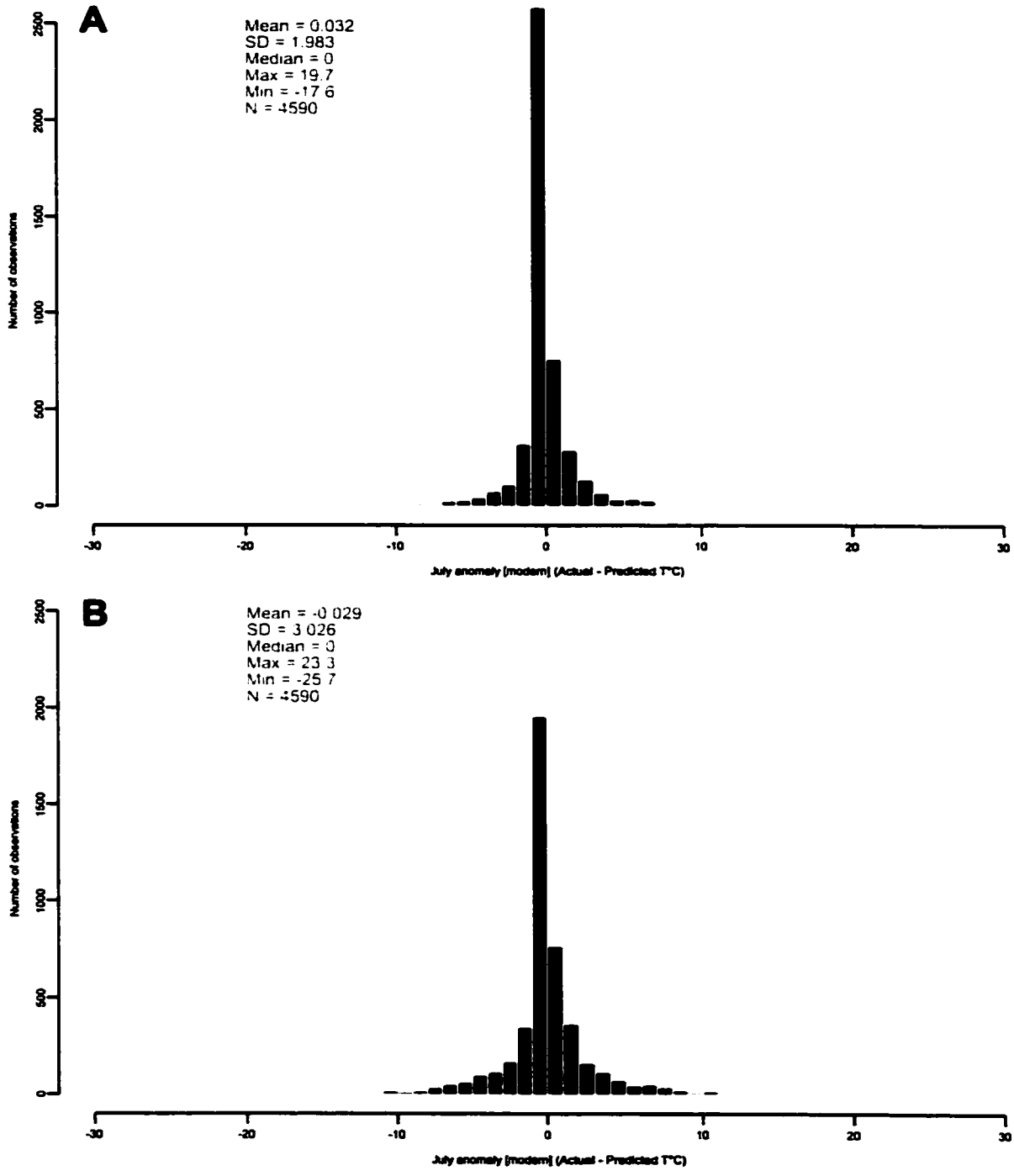


Figure 4.8: A –July temperature anomalies for the modern vs. modern climate reconstruction using the modern analog technique (MAT) for a pollen set of 89-types. B – Same but for pollen set of 18-types.

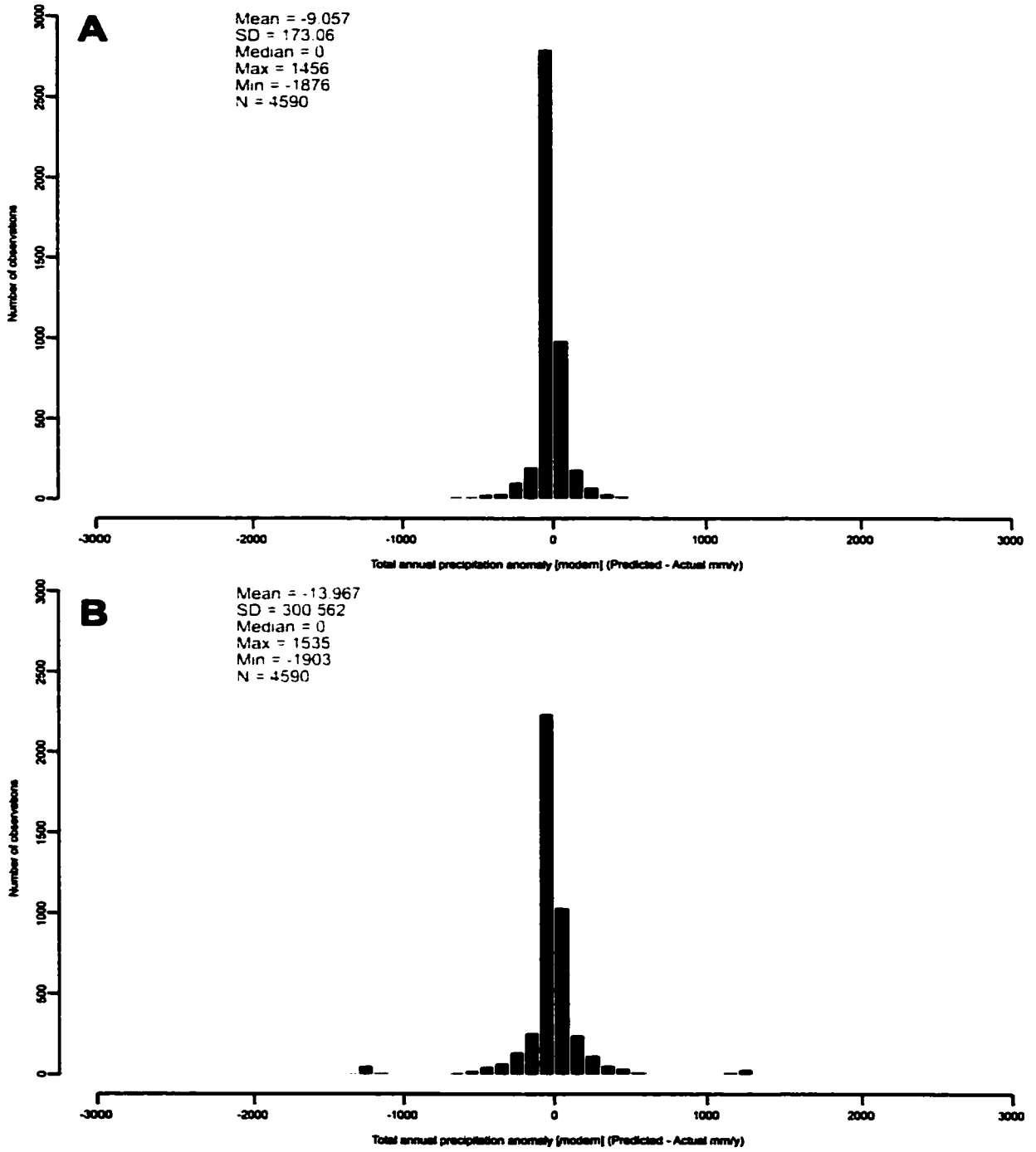


Figure 4.9: Same as Figure 4.8 but for total annual precipitation anomalies.

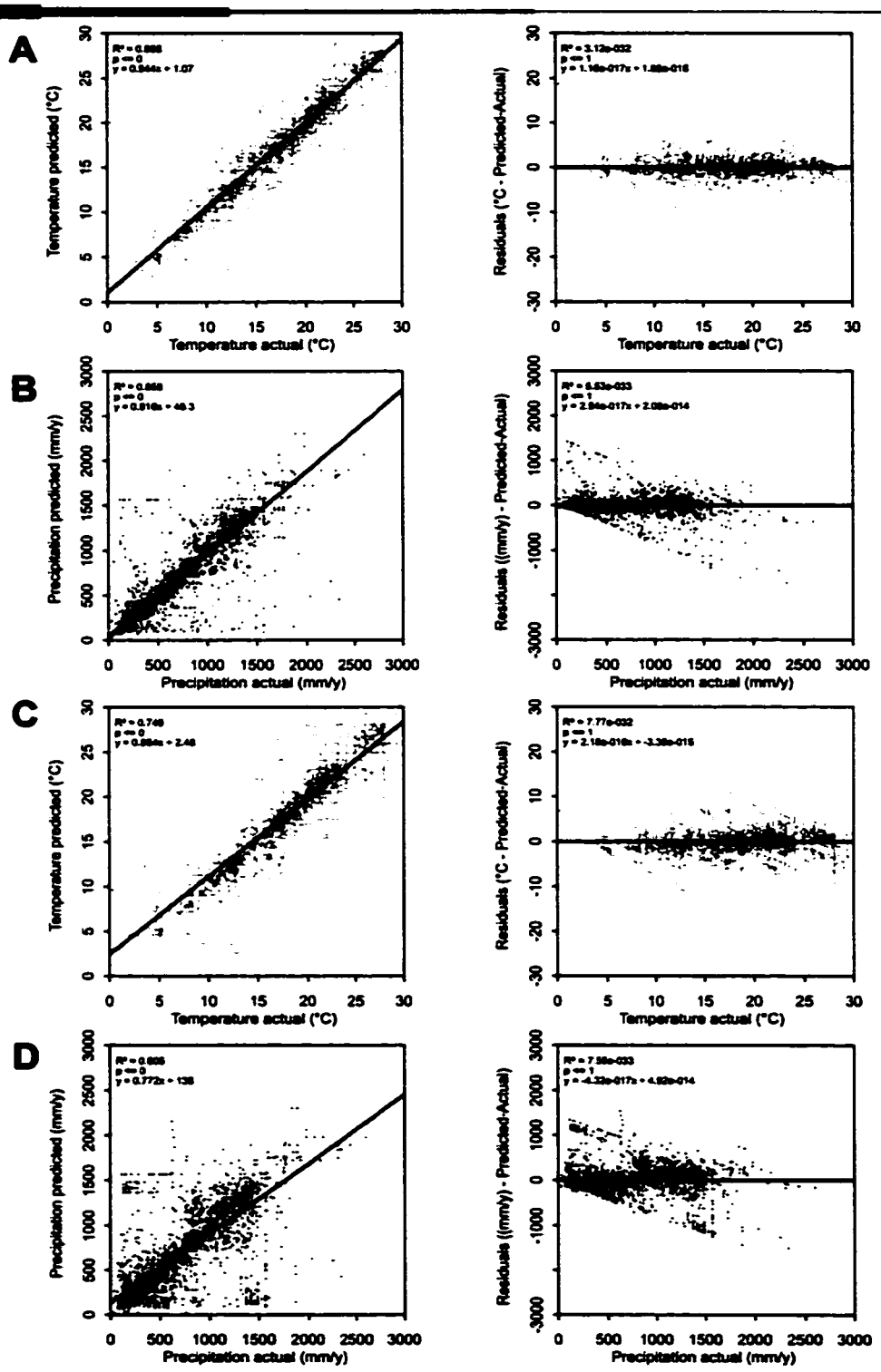


Figure 4.10: **A** – The left panel shows the linear regression between modern temperature at modern pollen sites in North America and the modern temperature as reconstructed at these sites from the 2nd best modern analog using a pollen set of 89-types. The line is a linear regression fit and the adjacent right plot shows the regression residuals. The *p*–values represent the probability that the linear regression slope is equal to zero. **B** – Same as above but for precipitation. **C** & **D** – Same as above but for a pollen set of 18-types.

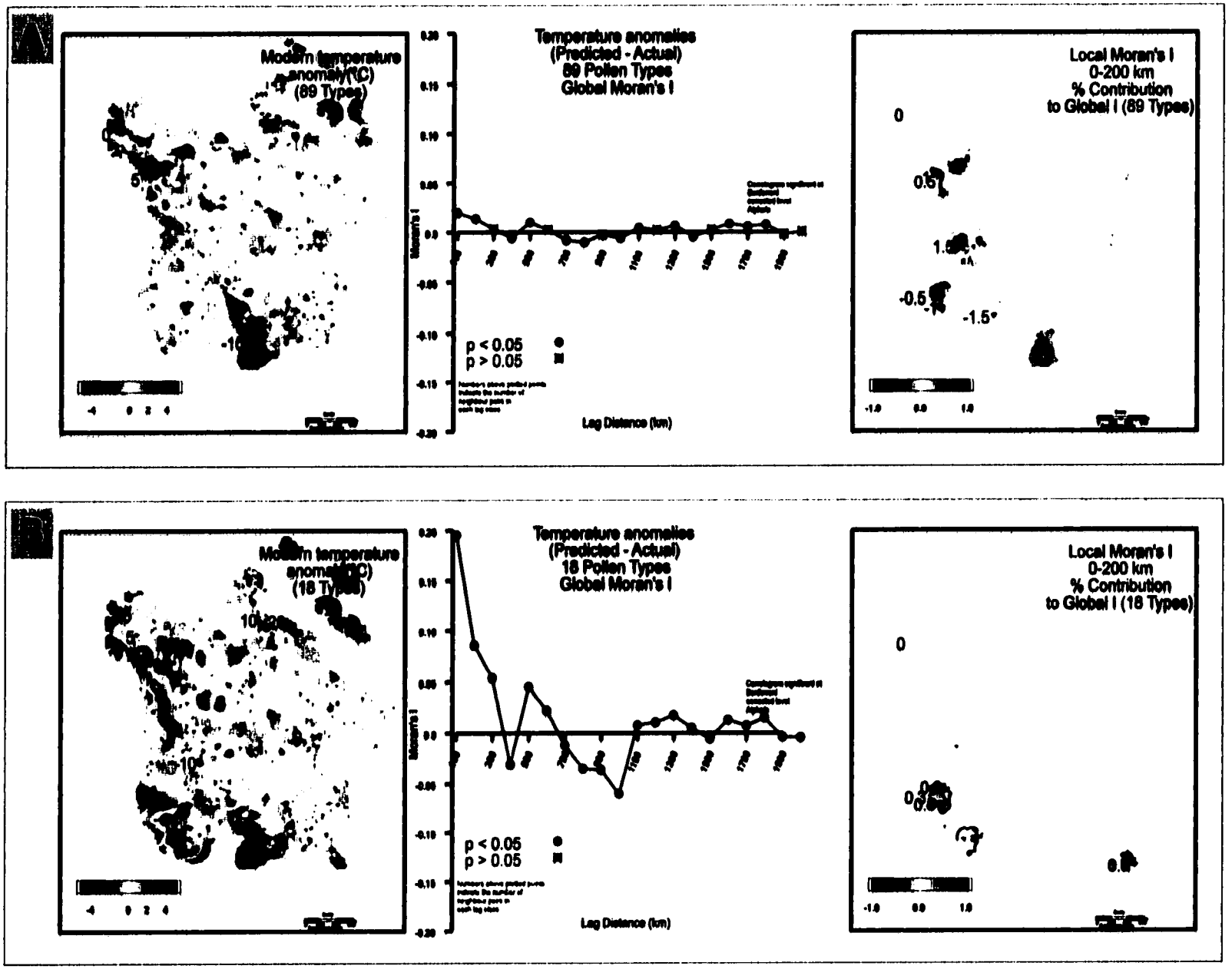


Figure 4.11: A – ***Left Panel.*** Map of temperature anomalies for the modern vs. modern reconstruction using the 2nd best modern analog with a sum of 89-types; ***Middle Panel.*** Spatial autocorrelogram showing Moran's I at mutually exclusive distance classes ranging from 0–100, 100–200, 200–300,...,1900–2000 km for the temperature anomalies; ***Right Panel.*** Map of Local Moran's I illustrating the proportion of global spatial autocorrelation contributed to the global Moran's I at a distance from 0–200 km. **B** – Same as above but for a pollen set of 18-types.

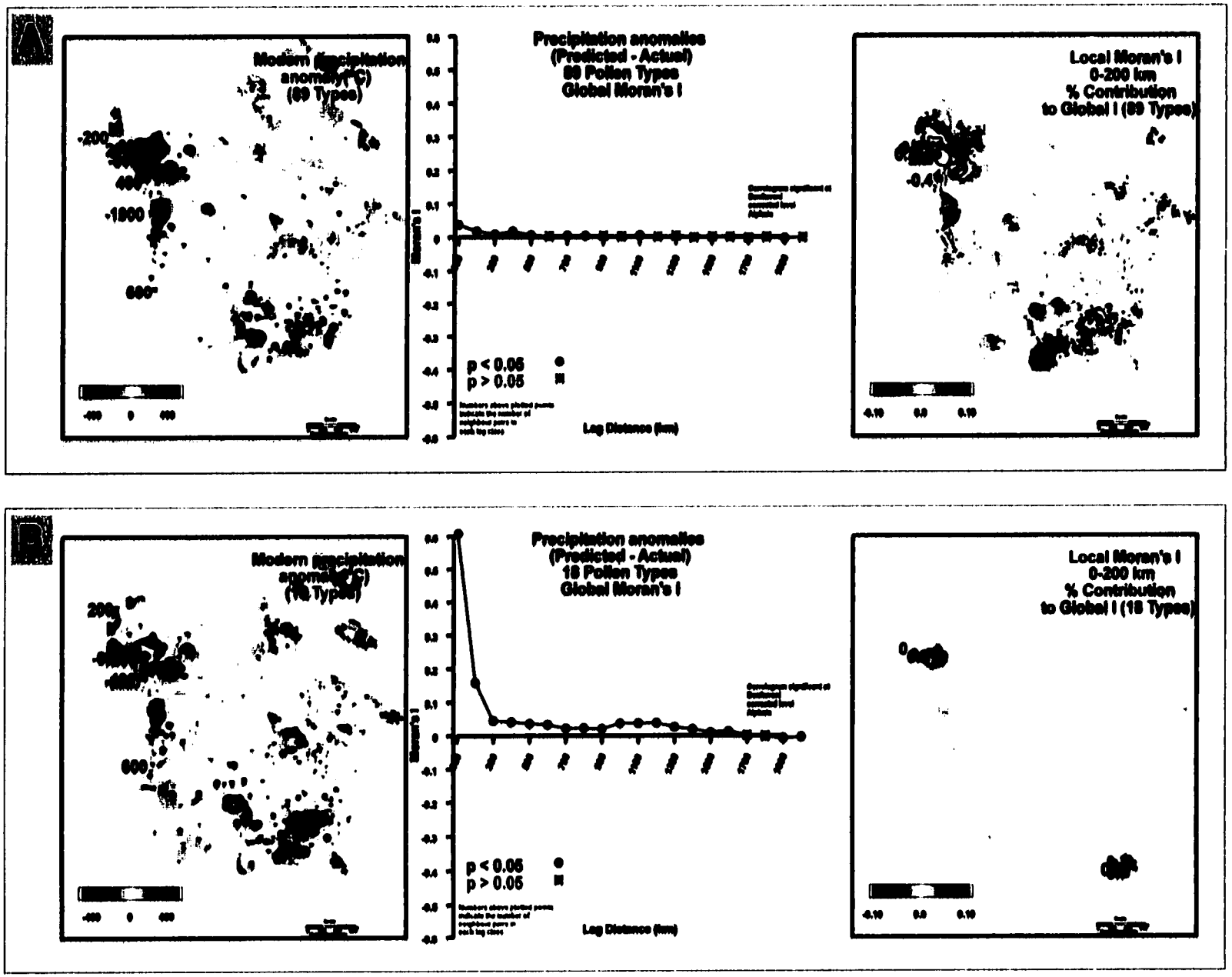


Figure 4.12: **A** – Same as Figure 4.11a but illustrating total annual precipitation anomalies. **B** – Same as Figure 4.11b but illustrating total annual precipitation anomalies.

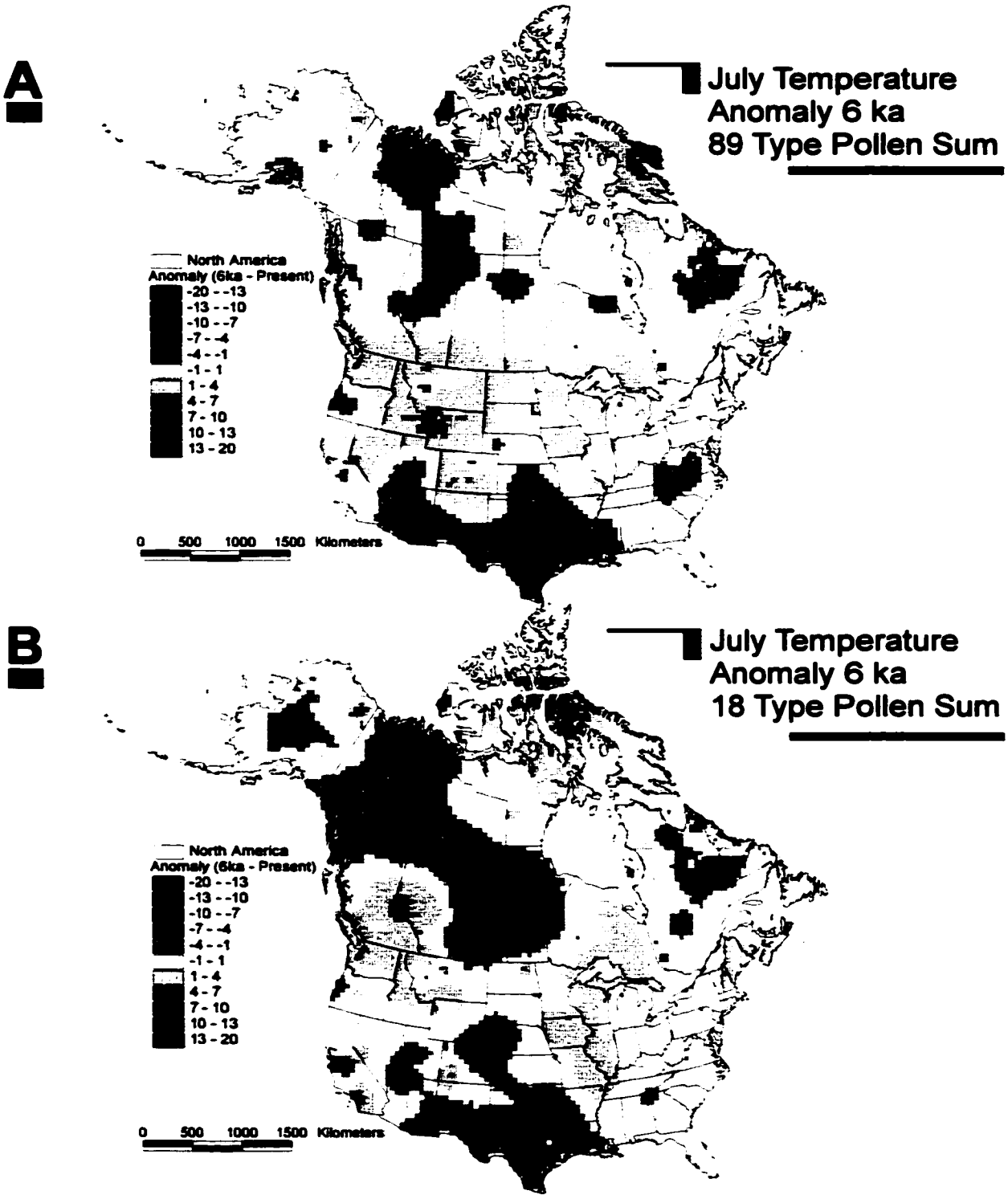


Figure 4.13: A – July temperature anomalies for the the 6 ka - Present from the 89-type pollen set. B – Same as above but for the 18-type pollen set.

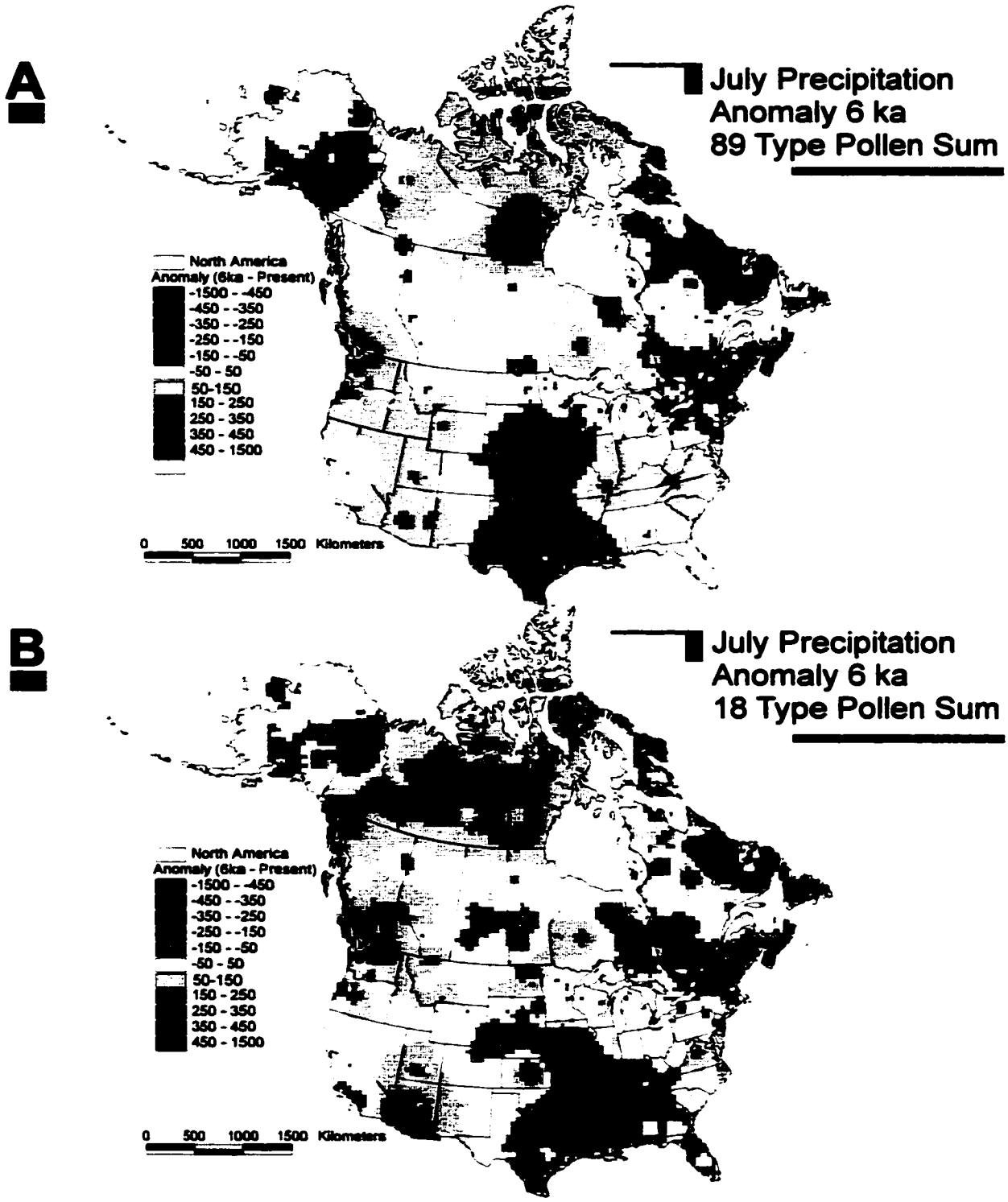


Figure 4.14: Same as Figure 4.13 but illustrating total annual precipitation anomalies at 6ka.

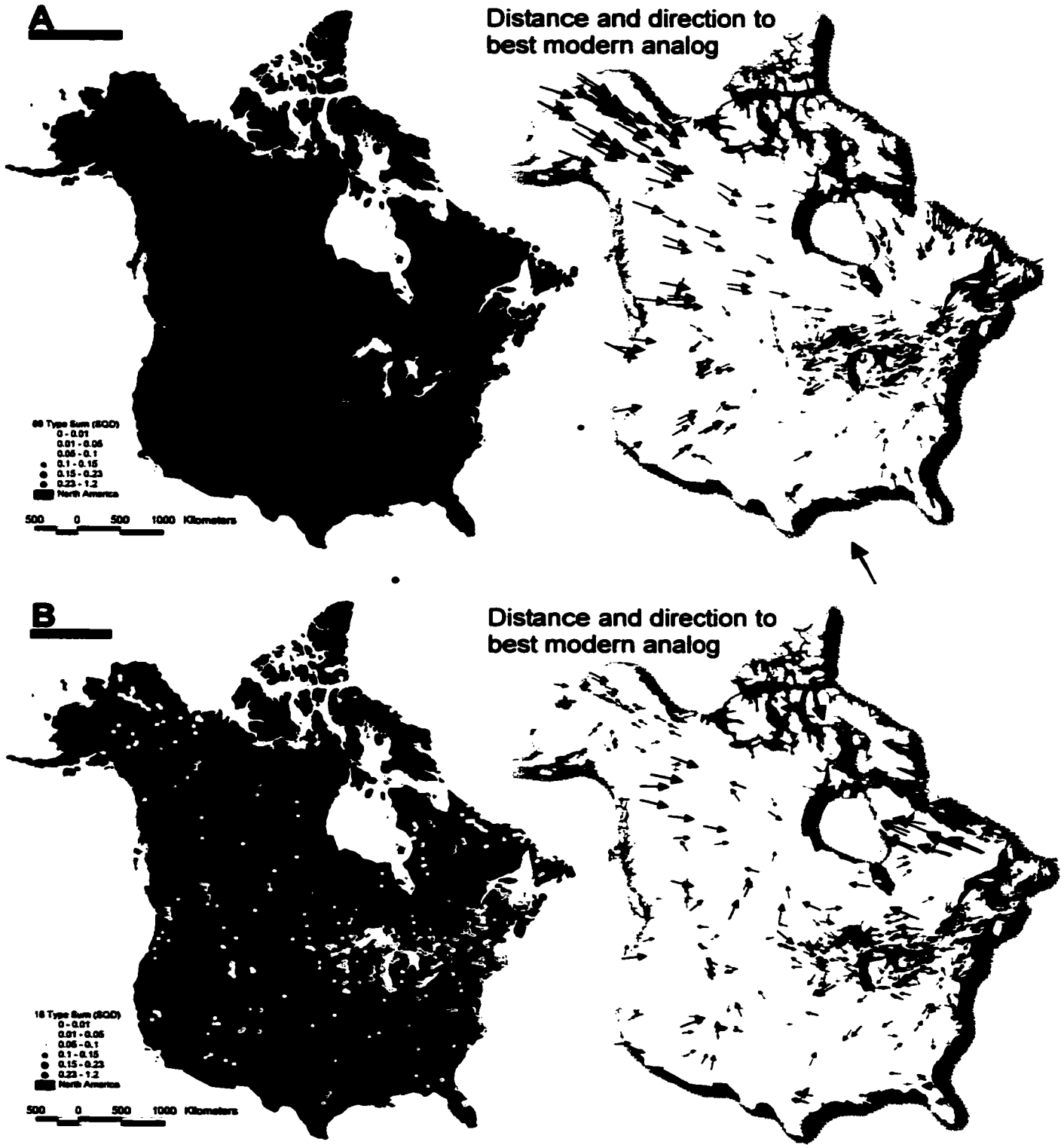


Figure 4.15: A – *Left Panel*: Distribution of squared chord distance coefficients in North America at 6 ka for a pollen set of 89 -types; *Right Panel:* Vectors illustrate the direction at which each fossil pollen site at 6ka finds its best modern analog. The vectors are scaled linearly from the minimum distance to the maximum distance between all sites and are proportional to the distance to the best modern analog. B – Same as above but for a pollen set of 18-types.

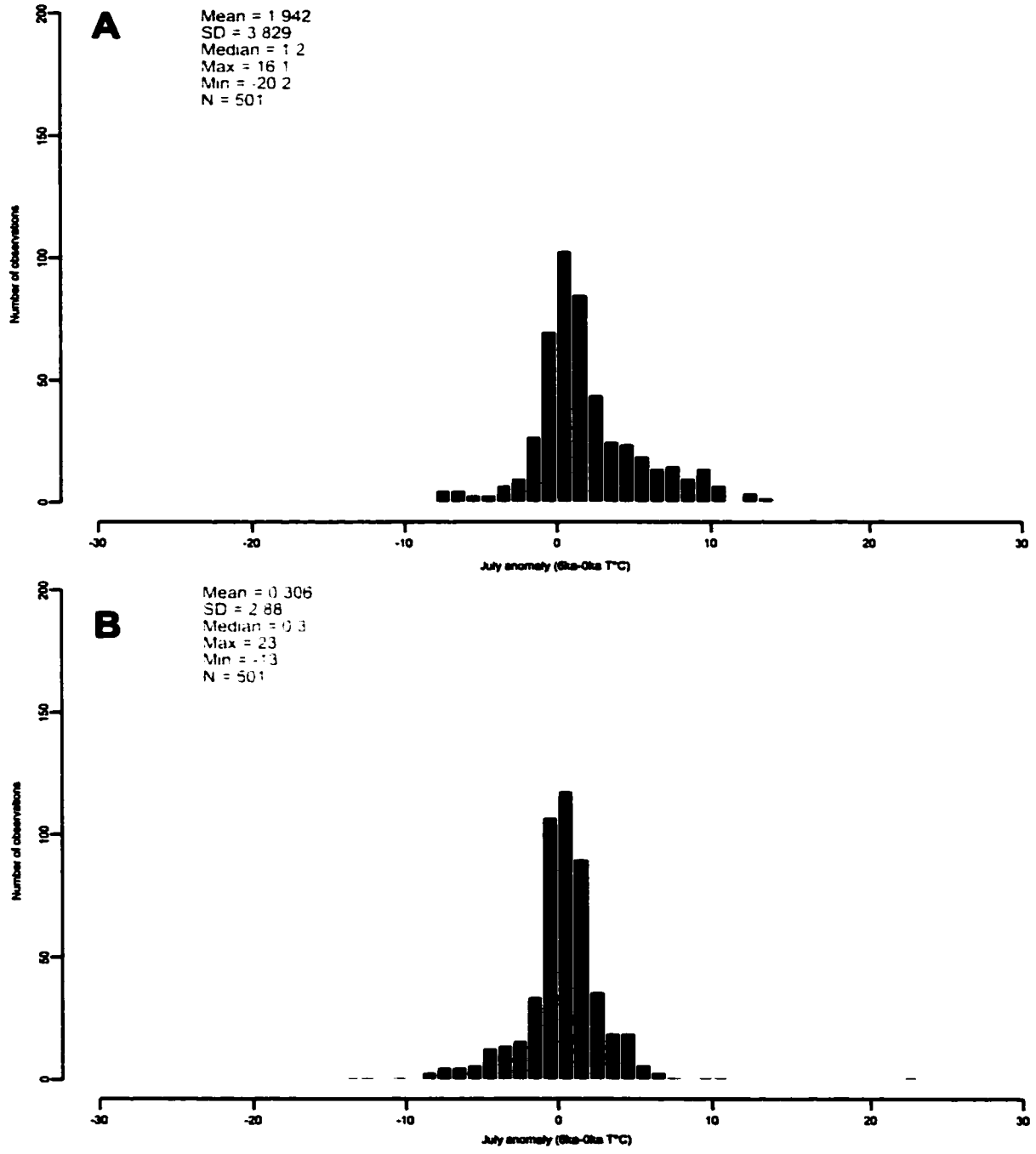


Figure 4.16: A – July temperature anomalies for the 6 ka - Present climate reconstruction using the modern analog technique (MAT) for a pollen set of 89-types. B – Same but for pollen set of 18-types.

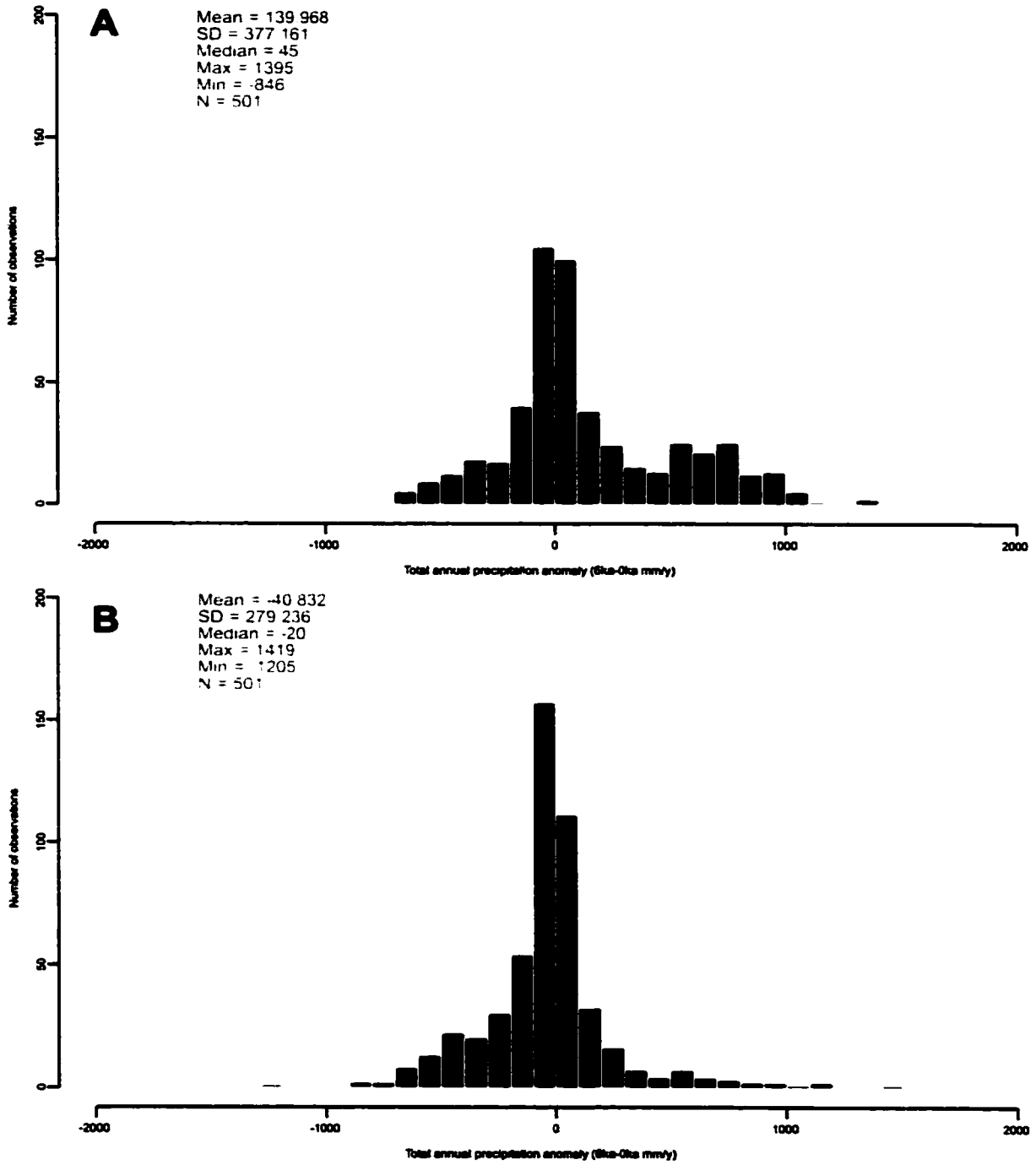


Figure 4.17: Same as Figure 4.16 but illustrating total annual precipitation anomalies for 6 ka – Present.

Chapter 5

EXPLORING THE PROXY CLIMATE RECORD FOR ANCILLARY CLIMATIC INFORMATION: *Sphagnum* PEATLAND DISTRIBUTION OVER THE PAST 21,000 YEARS IN NORTH AMERICA

5.1 INTRODUCTION

This chapter and the next explore the proxy climate record of the North American Pollen Database (NAPD – June 1999 version) (Grimm, 1999; Contributors to the NAPD, 1999) and Global Pollen Database (GPD – June 1999 version) (Grimm, 1999; Contributors to the GPD, 1999), furnished by the NOAA Paleoclimatology Data Center, for underutilized proxy pollen/spore data that can contribute more to our understanding of the paleoclimatic picture of North America over the past 21,000 years. Fossil pollen from arboreal taxa have been utilized extensively for paleoclimate reconstructions in North America and Europe (Sawada *et al.* 1999; Gajewski *et al.* 2000; Prentice *et al.*, 1998; Webb *et al.*, 1998). However, there are underutilized parts of the pollen/spore record that can provide climatic information on a continental scale through the late Quaternary. Specifically, two underutilized parts of the fossil pollen/spore record are *Sphagnum* spores and aquatic pollen (Dieffenbacher-Krall, 1998; Dieffenbacher-Krall and Jacobson, 2001). The exploration of these data for their climatic signal is the focus of this (*Sphagnum*) and the following chapter (Chapter 6).

5.1.1 *Sphagnum* PEATLANDS

Mires are peat-forming systems with high water tables, classified as either bogs or fens (Aerts, 1999). *Sphagnum* mosses dominate bogs and interstitial waters have low pH and alkalinity with nutrients coming

solely from precipitation and dryfall (Aerts, 1999). Fens have nutrient inputs from ground and surface water, have less *Sphagnum* mosses and higher primary productivity – usually dominated by grasses, sedges and rushes (Aerts, 1999; Jonasson and Shaver, 1999). Bogs are most abundant in North America in the Boreal zone (Aerts, 1999; Vitt, 1990; Moore and Bellamy, 1974) and cover 12% of the land area in Canada (Gorham, 1988). *Sphagnum* moss is the dominant component of bogs and low-pH fens (Jonasson and Shaver, 1999). The extensive horizontal growth and slow rate of decay of *Sphagnum* are two of the primary reasons for its large volume in bogs (Aerts, 1999; Clymo 1984). *Sphagnum* spores should be most abundant in regions dominated by bogs and low-pH fens. Because *Sphagnum* plants often establish after peat formation has already begun (Vitt *et al.*, 1994), the distribution of *Sphagnum* dominated peatlands as inferred from *Sphagnum* spores can provide a conservative estimate of those regions capable of supporting peatland growth. Here "peatlands" are defined as *Sphagnum* dominated bogs and low-pH fens.

Delimiting past peatland distribution can contribute to understanding of the global carbon cycle. Past peatland distributions have been used to estimate carbon accumulation through the late Quaternary (Appendix F) (Gajewski *et al.*, 2001; Halsey *et al.*, 2000). Biomass production is greater than the rate of decomposition in the anoxic and acidic water-saturated bog environment, thus *Sphagnum* dominated peatlands sequester carbon through net accumulation (Aerts, 1999; Jonasson and Shaver, 1999; Kuhry and Vitt, 1996; Vitt, 1990; Moore and Bellamy, 1974). Boreal and subarctic peatlands store an estimated 455 Pg [1 Pg = 10^{15} g] of carbon - approximately one third of the global soil carbon pool - and remove an estimated 0.076 Pg of carbon from the atmosphere annually (Gorham, 1991). For example, measurements in Manitoba indicate that the typical carbon storage in moss covered spruce forests is ~40 tons C/ha in live and dead aboveground spruce and ~45 tons C/ha in live/dead moss (as much as ~200 tons C/ha in soil of *Sphagnum* areas) (Goulden *et al.*, 1998). On a global scale, there is more carbon locked up in *Sphagnum* (dead and living) than is fixed by all terrestrial vegetation in a single year (Clymo and Hayward, 1982). The carbon stored in peatlands could be important under global warming due to possible loss of carbon from soil caused by increased depth of the active layer in mid-summer in permafrost dominated Boreal and

subarctic regions. Increased thaw-depth in summer will lead to enhanced rates of decomposition that may not be offset by gains in storage from wood or moss and lead to a net carbon efflux (Goulden *et al.*, 1998). The change in carbon balance from accumulation to loss may be indirectly linked to temperature via enhanced soil drainage, aeration and decreasing water tables (Oechel *et al.*, 1993).

5.1.2 *Sphagnum* PEATLANDS AND CLIMATE

Because climate plays an important role in the regional distribution of peatland types (Gignac *et al.*, 1998; Gignac and Vitt, 1990, 1994; Moore and Bellamy, 1974) the study of past peatland distributions can lead to a better understanding of paleoclimate in North America. Peat stratigraphies have been utilized for the reconstructions of paleoclimates, in particular peat accumulation has been linked with past changes in temperature and precipitation (Blackford, 2000). The presence of *Sphagnum* rich levels in stratigraphic sequences have been associated with wetter and/or cooler conditions (Blackford, 2000). In particular, water table fluctuations that respond to climate affect the growth of *Sphagnum* species (Blackford, 2000; Gerdol, 1995). Contemporary peatland development is strongly linked to climate, with bogs (and fens) occurring in areas where precipitation exceeds evapotranspiration and where total annual precipitation is typically greater than 500 mm (Gignac and Vitt, 1994). However, because temperature is also a fundamental control on peatland development, wide expanses of peatlands occur where precipitation may be less than 400 mm annually (Moore and Bellamy 1974). The extent of peatlands also depends on topography, and flat terrain provides an ideal setting for vast expanses of peatlands (Zoltai and Pollett, 1983). The reconstruction of *Sphagnum* peatland distribution in North America can be an important contributor to constraining paleoclimate reconstructions of precipitation and temperature through the past 21,000 years.

Response surfaces have been used extensively with modern pollen assemblages to estimate vegetation responses along climatic gradients (Anderson *et al.*, 1991; Prentice *et al.*, 1991; Huntley *et al.* 1989; Bartlien *et al.*, 1986;). Response surfaces have also been used with samples of modern mire bryophyte abundance in western Canada to estimate their behavior along climatic and other environmental gradients

(Gignac *et al.*, 1991). Response surfaces have not, however, been applied to *Sphagnum* spore data on a North American scale. The work presented here utilizes response surfaces in order to determine the modern climatic ranges in which 1) *Sphagnum* occurs and, 2) *Sphagnum* is most abundant. Constructing the modern climatic ranges of the spore data allows us to make inferences about past climate conditions in those areas where *Sphagnum* is found.

This chapter shows the relation between the modern spore abundance and peatland limit in North America. Second, it demonstrates the relation between modern climate and spore abundance and distribution using a response surface approach. Finally, it presents maps of the past abundance of *Sphagnum* peatland limits in North America for the past 21,000 years and interprets these in the context of changing climate.

5.2 Data and Methods

5.2.1 Spore Data

Sphagnum percentages were computed within a sum composed of pollen from all upland plants and bryophyte spores taken from the NAPD (June 1999 version) (Grimm, 1999; Contributors to the NAPD, 1999). The radiocarbon equivalents of target calendar years spanning 0ka to 21ka by 1ka intervals were determined using CALIB version 4.2 (Stuiver and Reimer, 1993, 2000). *Sphagnum* percentages were then linearly interpolated to these radiocarbon equivalents to produce the associated maps of *Sphagnum* abundance over the past 21,000 years.

From the GPD (June 1999 version) (Grimm, 1999; Contributors to the GPD, 1999), 3005 modern pollen sample sites in North America (Figure 5.1) were extracted. Using Leemans and Cramer (1991), mean annual temperature and total annual precipitation were extracted via linear interpolation for each of the 3005 sites using a Geographic Information System (GIS). In western Canada, Gignac *et al.* (1991) used the length of growing season and total precipitation falling as rain from 30 yr records at nearby climate stations

for their bryophyte response surfaces. Mean annual temperature and total annual precipitation will be strongly correlated with growing season and pluvial precipitation and are equally justified for use here. Moreover, working at the continental scale requires continental scale climatic data, which was not readily available in the form used by Gignac *et al.* (1991). Some sites on small coastal islands did not fall within the climate grid cells; climate data for these sites were extracted via nearest-neighbor interpolation such that the climate at the nearest grid node was assigned to each spore sample site. For the initial climate surfaces, 830 sites had identical climatic values and the percent of *Sphagnum* at these was averaged, leaving 2174 sites to be used in the construction of a response surface.

5.2.2 Response Surface

Because of the large differences in the absolute values between the two climate axes (temperature *vs.* total annual precipitation), climate axes were scaled to unity prior to interpolation (Gignac *et al.*, 1991):

$$z = \frac{\left[x_i - \min(x_i \ldots x_n)\right]}{\left[\max(x_i \ldots x_n) - \min(x_i \ldots x_n)\right]} \tag{5.1}$$

where z is the scaled variable in the set of observations $x_i \ldots x_n$. To estimate the response of *Sphagnum* spore percentages in climate space, a kernel density interpolation (Bailey and Gatrell, 1995, p. 159; Cressie, 1995, p.599) utilizing a quartic distribution function was employed. Cubic functions have been used for surface interpolation (Gignac *et al.*, 1991), however, the form of the kernel is less important than the choice of bandwidth (Cressie, 1995) which will determine the amount of smoothing and extrapolation in the estimated surface. The quartic function is appropriate because grid nodes with no samples within the bandwidth are forced to zero, have zero weight, thus limiting extrapolation beyond the sample space, that is the quartic kernel is defined as:

$$w_i(d) = \begin{cases} \dfrac{3}{\pi\tau^2}\left[1 - \left(\dfrac{d_{ij}}{\tau}\right)^2\right]^2 & d_{ij} \le \tau \\ 0 & otherwise \end{cases} \tag{5.2}$$

Where $w_i(d)$ is the weight for the *Sphagnum* percentage at grid node *i* where the distance between grid node *i* and a given sample point in climate space, *j*, is d_{ij}. Here, τ, is the bandwidth of the function and represents the radius within which points significantly contribute to the estimate (Figure 5.2). In other words, the bandwith is the range within which points will significantly contribute to the estimated intensity at a given grid node. This quartic kernel is then used as the weight for the distance weighted interpolation for each grid node:

$$\hat{y}_{ij} = \frac{\sum_{j=1}^{j=n} w_i(d_{ij}) Z_j}{\sum_{j=1}^{j=n} w_i(d_{ij})} \tag{5.3}$$

where Z_j is some sample point in space and *i* is a given grid node. Spore percentages were interpolated onto a grid of 100 by 100 (0.46 °C by 19.7 mm) with a kernel bandwidth three times this resolution. The choice of resolution for the climate surface is appropriate for spore data because it limits extrapolation beyond the convex hull defined by the original samples in climate space, e.g., extrapolation will only be at most 59.1 mm by 1.38°C away from any of the original samples; a coarser grid would equate to larger extrapolation in climate space.

The estimated surface (Figure 5.3) was then smoothed with a moving window which was 5 x 5 grid nodes in *x* and *y* using inverse distance squared weighting centered on each of the estimated original grid nodes. Extrapolation was a concern considering that spores can be transported over larger distances than macrofossils and therefore the climatic space defined by positive spore percentages may be larger than the climate space defined by the modern distribution of the *Sphagnum* genera itself. Consequently, because the response surface is estimated using spore data, it is more likely to have sites included in the response surface that have positive *Sphagnum* percentages even though *Sphagnum* might not be present in the immediate geographic region. However, smoothing was necessary because of: 1) local overrepresentation of percentages in bog samples; and 2) overrepresentation of percentages of *Sphagnum* spores in the Arctic

regions that lack significant arboreal taxa within the pollen sum. Smoothing increases our extrapolation to 2.3°C and 98.5 mm. Our smoothed *Sphagnum* response surface (Figure 5.3) explains $R^2 = 0.37$ ($p < 0.05$) of the variance in the original sample percentage values. The mean estimated and predicted percentages of *Sphagnum* are not significantly different using a *t*-test.

5.2.3 Sampling Intensity and Response Surface Robustness

In order to test how well the response surface estimated the *Sphagnum* centre of abundance in climate space, the response surface robustness was tested in three ways:

5.2.3.1 Degree to which *Sphagnum* sample sites represent North American Climate Space

It is necessary to rule out the possibility that the response surface maximum is a consequence of limited sampling in North American climate space. To test the degree to which the response surface is representative of the North American climate space, all temperature and precipitation (Leemans and Cramer, 1991) grid nodes within North America (25-83°N and 49.5-168°W) were plotted in climate space and overlayed with the climate combinations sampled at the modern *Sphagnum* spore sample sites that were used in the construction of the response surface. Using a GIS, convex hulls were constructed around the climate space points for all of North America and around a subset of these at which *Sphagnum* spore sampling sites occurred. An intersection overlay of these convex hulls allowed us to estimate that our *Sphagnum* sites sample approximately 55% of the total climate space within North America (Figure 5.4).

5.2.3.2 Degree to which sampling at sites with non-zero *Sphagnum* affects the maximum of the response surface in climate space

It is also necessary to address the possibility that the response surface maximum may be an artifact of limited sampling of those sites with positive *Sphagnum* observations in climate space. In other words, what is the distribution of samples in climate space that have positive *Sphagnum* observations and how well do they sample the possible climates of North America where *Sphagnum* is likely to occur. All points with *Sphagnum* percentages of zero were deleted from the 2174 sites. The remaining 774 sites were used for the creation of a sampling intensity surface in climate-space (Figure 5.5). The sampling intensity surface utilized a two-dimensional Gaussian kernel:

$$\hat{\lambda}_i(d) = \sum_{j=1}^{j=n} \frac{1}{2\tau^2\pi} e^{-dij^2/2\tau^2} \times n \left(\sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \frac{1}{2\tau^2\pi} e^{-dij^2/2\tau^2} \right)^{-1} \quad (5.4)$$

where the point intensity at grid node i ($n \times unitarea^{-1}$), is a function of τ is the bandwidth and d_{ij}, the distance between grid node i and a given point in the point pattern j containing n points. The Gaussian kernel provides a smoother surface using the same bandwidth as that in the response surface. The Gaussian kernel will provide a more liberal extrapolation of sampling intensity and therefore provide a more conservative comparison with the abundance center.

5.2.3.3 Degree to which the type of sampling site affected the estimated abundance maximum of *Sphagnum* in climate space

Consideration must also be given to whether the type of sampling site, lake vs. bog, unduly affects the location of the response surface maximum. Overrepresentation of *Sphagnum* spore percentages in bog samples may have affected the estimated climatic optimum of the *Sphagnum* response surface. To determine the degree to which the surface optimum may be biased, a response surface was created using the above procedures on *Sphagnum* spore percentages taken only from 664 modern lake samples in North America and a second response surface taking all other sites including bogs (Figure 5.6).

5.2.4 Mapping of *Sphagnum*

Sphagnum spores are most abundant in those regions where *Sphagnum* species are present, with trace amounts found away from the source. Thus, the distribution of *Sphagnum* spores from lake sediments can be used to indicate regions where *Sphagnum* peatlands are present, in the same way that the distribution of pollen of upland taxa can be used to indicate the distribution of trees. Plotting the modern distribution of *Sphagnum* spore percentages illustrates the clear relation between spores and peatland extent (Figure 5.1).

Maps were then drawn of the abundance of *Sphagnum* spores at 2000-year intervals. Geodetic coordinates of sample sites were projected using the forward spherical solution for the Albers Equal-Area Conic projection (Snyder, 1987) and *Sphagnum* percentages were interpolated using a standard inverse-distance

squared algorithm (Lam, 1983) with a grid resolution of 50 km^2. In order to limit any further extrapolation within and outside the sampling space, a search radius of 750 km in North America was imposed, beyond which the observations at sample sites contribute zero weight to the estimate at a given grid node, that is, for each grid node:

$$\hat{y}_{ij} = \begin{cases} \dfrac{\sum_{j=1}^{j=n} z_j / d_{ij}^2}{\sum_{j=1}^{j=n} d_{ij}^2} & d_{ij} \le 750\,km \\ 0 & otherwise \end{cases} \qquad (5.5)$$

where the estimated value of *Sphagnum*, $\hat{y}_{ij}$, at grid node i is a function of the values at sample sites $j_1 \ldots j_n$, weighted inversely by the distance between grid node i and site j, d_{ij}^2.

Based on the modern distribution of peatlands and *Sphagnum* percentages, the 0.5% isopoll corresponds to the present day peatland limits in North America. Grid cells with *Sphagnum* proportions greater than or equal to 0.5% were selected if they fell within modern-day continental boundaries and outside of ice sheet boundaries for each time-slice. The extent of the ice sheets was taken from Peltier (1993, 1994). However, the modern sea-level and coastline were retained for simplicity.

5.3 RESULTS

5.3.1 MODERN PEATLAND DISTRIBUTION AND *Sphagnum* SPORES

In North America, bogs are found primarily in the Boreal zone of Canada and Alaska and the northeastern and northern central United States, all of which exhibit positive values of *Sphagnum* spores (Figure 5.1). The southern modern-day distribution of positive values of *Sphagnum* spores corresponds closely to the distribution of peatlands, but not to the distribution of *Sphagnum* plants. The zone with maximum *Sphagnum* percentages is found in a band from north of the Great Lakes through to the Gaspé Peninsula. Peatland extent in northwestern Ontario and northeastern Manitoba is considerable but we have no modern

spore samples in these regions. *Sphagnum* spore percentages along the Atlantic Seaboard to Florida and at altitudes in the Appalachians reflect local peatlands in these regions (Hofstetter, 1983; Weider *et al.*, 1981). *Sphagnum* development has been observed on Axel Heiberg Island (Kuc, 1973). However, there are estimates in excess of 10% *Sphagnum* spores in the modern samples that are unrepresentative of the actual amount of *Sphagnum* abundance in this region. The lack of arboreal species and sparse ground cover encourages transport from ground level particulate sources like *Sphagnum*. *Sphagnum* plant abundances in the high Arctic may also be controlled more by local physiographic factors that produce favorable microclimates and there are no extensive peatlands in these regions. Many of these spores are likely transported from the boreal peatlands to the south (Bourgeois, 2000). Generally, however, the modern distribution of positive *Sphagnum* spore percentages is well confined to the regions of North America where *Sphagnum* peatlands are most abundant. At the North American scale of study, the errors in range boundaries that may be caused by long-distance transport of *Sphagnum* should not be of any primary concern.

5.3.2 THE RESPONSE SURFACE

Our response surface is not capable of representing climates in eastern Alaska, coastal Oregon and coastal southern British Columbia as there are no sample sites in these regions (Figure 5.3b; Figure 5.1). These areas have moderate mean annual temperatures and very high precipitation. However, on the North American scale these points are few and unlikely to produce significant *Sphagnum* plant abundance due to the dominance of temperate rainforest in these areas. Additionally, there are few samples where mean annual temperatures exceed 11°C and precipitation is less than 1000 mm y^{-1}. Specifically, this area of climate space is found in southern California, Arizona, New Mexico and Texas where precipitation is less than evaporation and *Sphagnum* is unlikely to occur on a large scale. In these arid regions all samples present contain zero *Sphagnum* spore percentages. Within the 55% of total North American climate space sampled, there are considerable climate spaces conducive to the occurrence of *Sphagnum* and it is in these areas that *Sphagnum* spore percentages are most abundant. The remaining 45% of the North American climate space convex hull contains spatially infrequent climates such as those in the western coastal

rainforests where *Sphagnum* is not abundant. Therefore, the spore sample locations capture the majority of available climate space within North America where *Sphagnum* can occur.

The sampling intensity (samples per 0.46°C by 19.6 mm) is used to indicate the degree to which sampling biases may have affected the estimated response of the *Sphagnum* optimum in climate space (Figure 5.5). While the estimated peak abundance of *Sphagnum* does coincide with an area of strong sampling intensity in climate space, it is constrained between three regions with considerably stronger clustering, none of which produce strong peaks in the response surface (Figure 5.5; Figure 5.6). This suggests that our estimated response optimum is 'real' rather than an artifact of a limited cluster of samples with positive *Sphagnum* percentages within one region of climate space.

The position of the center of activity in the response surface using only lake samples is found in the same region of climate space as peak abundance center in the response surface utilizing the entire dataset (Figure 5.6c). Using all spore sample sites (lakes, bogs etc.) leads only to larger absolute percentage estimates within the center of abundance in climate space (Figure 5.6ab). However, the abundance center of *Sphagnum* in climate space is largely a function of the non-lake modern spore samples (Figure 5.6d). *Sphagnum* may not be immediately local but will be regionally evident. The lake-only response surface (Figure 5.6c) provides evidence that lakes are detecting *Sphagnum* in those regions of climate space where *Sphagnum* is regionally present in peatlands. This is an important finding because most of the fossil pollen cores in this study are from lakes, which can therefore be used for reconstructing the past limits of peatlands.

5.3.3 *Sphagnum* AND CLIMATE

The center of *Sphagnum* abundance in climate space ranges from approximately 630 - 1300 mm of precipitation and between –1.6 to 6 °C mean annual temperature (Figure 5.3). Bryophyte studies in western Canada have suggested that peatlands will only occur where precipitation exceeds evaporation and such areas are generally found where precipitation is greater than 500 mm/yr in North America (Halsey *et al.*

1998). The sites responsible for the *Sphagnum* response maximum are mainly from northern Ontario, mid-Québec, the Maritime provinces and the New England states (Figure 5.3b). Our surfaces also have a few areas greater than 12 °C and 1200 mm precipitation (Figure 5.3). These represent samples from the southeastern United States (Figure 5.3b). Another cluster of high *Sphagnum* values is found at approximately 3 °C and above 1440 mm, which represents a few sample sites in Labrador and Newfoundland (Figure 5.3b). The *Sphagnum* response surface would suggest that mires could be found in areas of low mean annual temperature and precipitation such as in Somerset Island within the Canadian Arctic Archipelago. However, for reasons discussed above, high proportions of *Sphagnum* in the high Arctic are due to overrepresentation. Although *Sphagnum* can occur along a wide-range of precipitation and temperature, it is most abundant where precipitation is in excess of 500 mm/yr and mean annual temperatures are cool to moderately cool.

5.3.4 *Sphagnum* DISTRIBUTION IN NORTH AMERICA OVER THE PAST 21,000 YEARS

There was little *Sphagnum* at 21ka south of the ice sheet in North America (Figure 5.7). Scattered sites in eastern North America record a few spores indicating a degree of *Sphagnum* presence, perhaps in small bogs. Almost all sites in Alaska had *Sphagnum* spores, with high percentages registered in a few of these. Around 15 ka the number of sites recording *Sphagnum* increased, especially around the edge of the retreating ice sheet. In the Atlantic Provinces of Canada there were high values in many sites. By 13 ka, the number of available sites in eastern Canada and the United States increases, and *Sphagnum* was recorded in many of these. In western Canada, *Sphagnum* spread eastward on land made available as the ice sheet retreated. During the Holocene *Sphagnum* spread eastward from Alaska and westward from the north eastern United States to cover all of boreal Canada by 5 ka. In the late Holocene, *Sphagnum* spread slightly southward, as evidenced by the increased number of sites in the western Great Lakes region, which recorded large values.

5.4 DISCUSSION

The response surfaces and maps of the modern distribution show that *Sphagnum* spores in sediments can be used to illustrate regions of peatland presence that are consistent with their modern geographic and climatic ranges (Figure 5.1 and Figure 5.3). Moreover, the *Sphagnum* response surface utilizing only lakes illustrates that lakes sufficiently detect the regional presence of *Sphagnum* peatlands and can be used for the reconstruction of past *Sphagnum* peatland distributions. This adds validity to the current fossil reconstructions that utilize lake sediment cores. The northern limit of *Sphagnum* peatlands on the landscape is not well captured by the 0.5% isopoll due to efficient long-distance transport of spores and low local production of pollen, which inflates the *Sphagnum* percentages. Long-distance transport of spores and the low local pollen productivity in Arctic regions in North American lead to significant overrepresentation of *Sphagnum* ranges in this area, in particular over the past 9 ka. However, the area with abundant spores corresponds to the boreal zone, which has extensive peatlands. The arid southwest shows no *Sphagnum* spores in the modern data and the eastern United States, with smaller and less extensive peatlands, exhibits spore proportions where peatlands are local but not extensive. The southern limit of the 0.5% isopoll corresponds to the southern limit of *Sphagnum* peatlands (Figure 5.1). This effect is evident in all of the maps. The response surface constructed concords with previous work based on the distribution of *Sphagnum* plants (Gignac *et al.*, 1998; Halsey *et al.*, 1998).

In this chapter, the closed 0.5% *Sphagnum* contour is interpreted as the area capable of supporting *Sphagnum* peatlands, that is, bogs and low-pH fens. There may be a lag between initial peat formation and the establishment of *Sphagnum* in a particular peatland, and this lag time is longer in continental peatlands than in coastal ones (Vitt *et al.*, 1994). As such, peatlands and the climatic conditions suitable for peatland can occur in regions where *Sphagnum* is not dominant or abundant. Therefore, the estimates of peatland extent encompassed by the 0.5% *Sphagnum* contour in this study are somewhat conservative. The 0.5% contour is a function of the pollen set used; for example, if fewer types were included in the pollen set then some other *n*% contour would correspond to the modern distribution of peatlands in North America. The

presence or absence of peat accumulation in peatlands is under the influence of the macroclimate and it is this, which is the focus of interpretation in this study.

Sphagnum peatlands first developed during deglaciation (15-11 ka) in eastern North America. The area of peatland initiation and development moved eastward from Alaska and westward from the St. Lawrence region. This spread was related partly to the availability of land after deglaciation but initial establishment at any site is also dependent on climatic changes. Several studies have illustrated the importance of climatic changes in initiating peat formation, or in causing a transition from marsh to fen or fen to bog (Vardy *et al.*, 1998; Zoltai, 1995; Vitt, 1994; Janssens *et al.*, 1992; Zoltai and Vitt, 1990; Winkler, 1988). The *Sphagnum* response surface illustrates the strong climatic signal in the spore data. This signal could be used for the interpretation of Late Quaternary reconstructions of temperature and precipitation.

Generally, these reconstructions agree with previous studies by Halsey *et al.* (1998) indicating peatland initiation between 8000-9000 ^{14}C yrs BP in the mountain and northern Alberta regions. As the climate became cooler and moister during the mid- to late Holocene, peatlands formed in the western Great Lakes region with the development of the Red Lake Peatlands and smaller bogs (Janssens *et al.*, 1992; Griffin, 1977; Heinselman, 1970; Winkler, 1988). The maps of Halsey *et al.* (2000), based on presence or absence of spores at a selected subset of North American sites also agree broadly with this study. This study, however, expands these results to the continental scale.

5.5 Conclusion

This work presents several new conclusions. First, the distribution of *Sphagnum* spores like the plants themselves and peatlands in general exhibit a strong climatic signals (e.g., Moore and Bellamy, 1974; Vitt *et al.*, 1994). Thus the *Sphagnum* range boundary maps indicate that portion of the Northern Hemisphere that supports *Sphagnum* growth and how it has changed through time. At the spatial scale and resolution of

global climate models (GCMs), and continental climate reconstructions of North America, these maps and climate response surfaces can aid in understanding large regional scale climates and in constraining climate reconstructions during the Late Quaternary in North America.

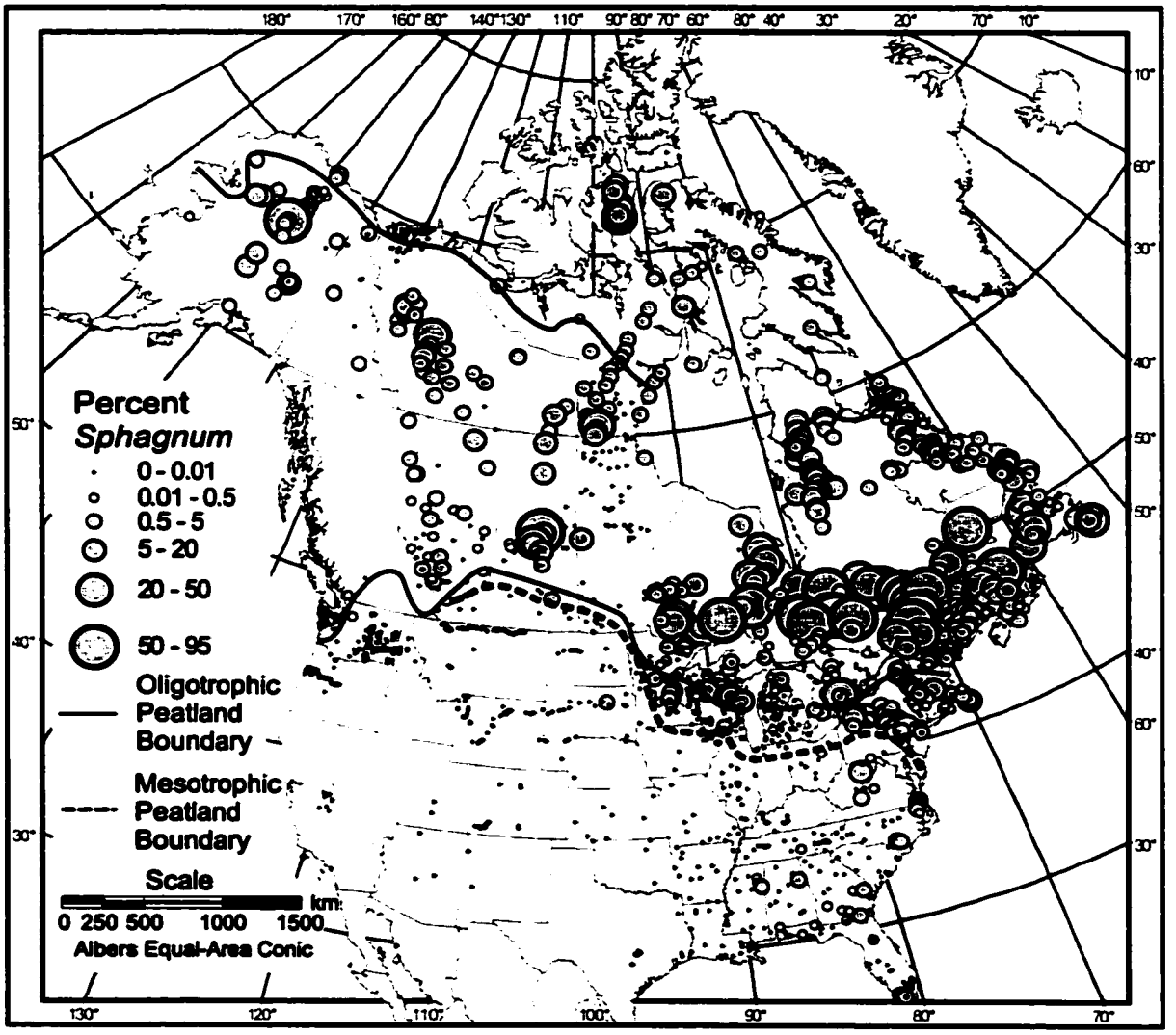


Figure 5.1: Modern *Sphagnum* spore distribution in North America. From 3005 pollen sample sites in the GPD, sites with identical geographic coordinates had their *Sphagnum* percentages averaged to produce this map (from Gajewski *et al.* 2001)

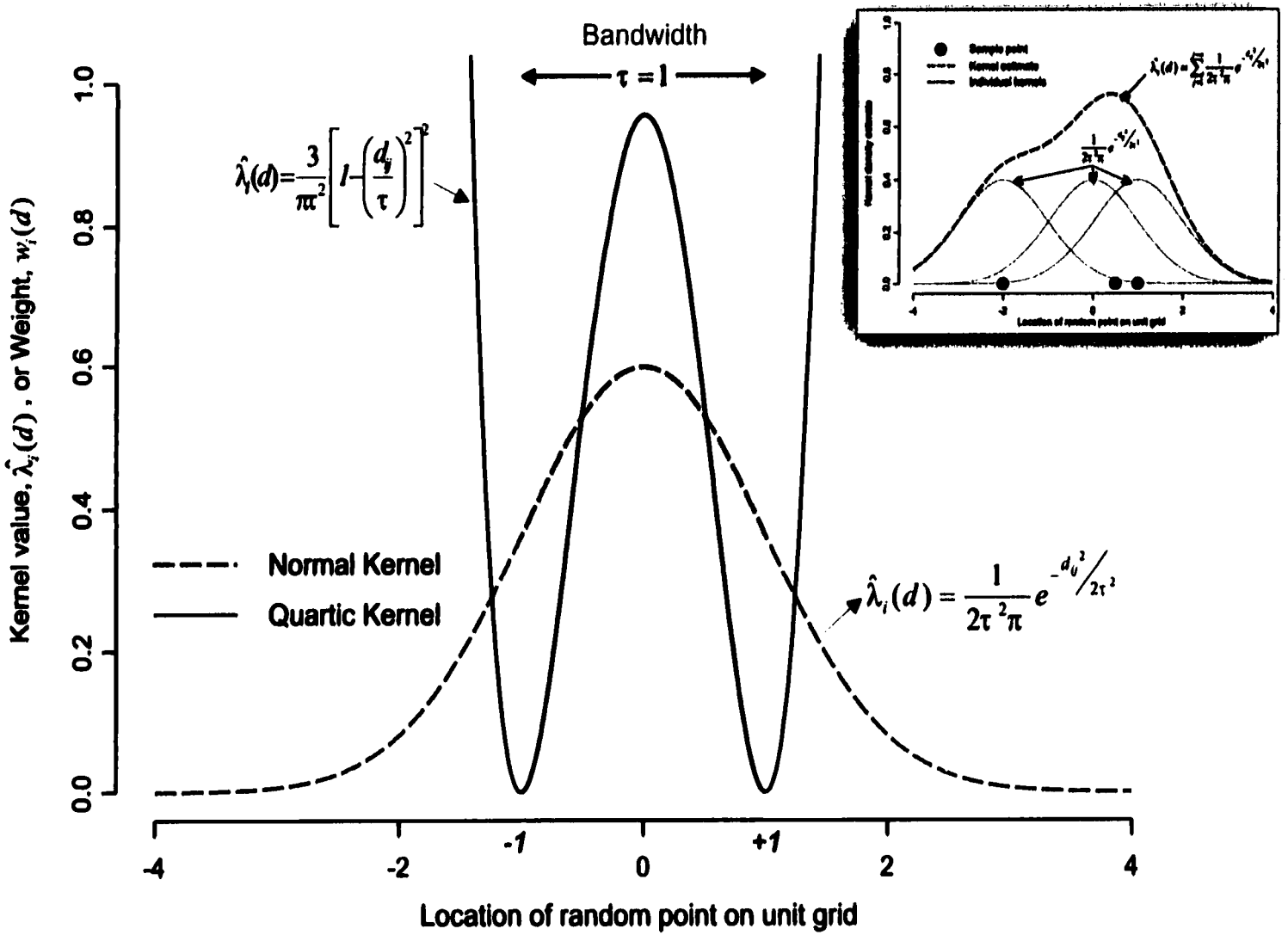


Figure 5.2: Quartic and normal kernel in 1–*d* illustrating bandwidth and form. The density estimate for the quartic kernel is only applicable within the bandwidth and zero otherwise. ***Inlay:*** Illustrates the kernel density estimate in 1–*d* for three points. A normal kernel is placed atop each point location and the distributions are summed vertically to provide a continuous probability density function. In the present dissertation this is done in two dimensions.

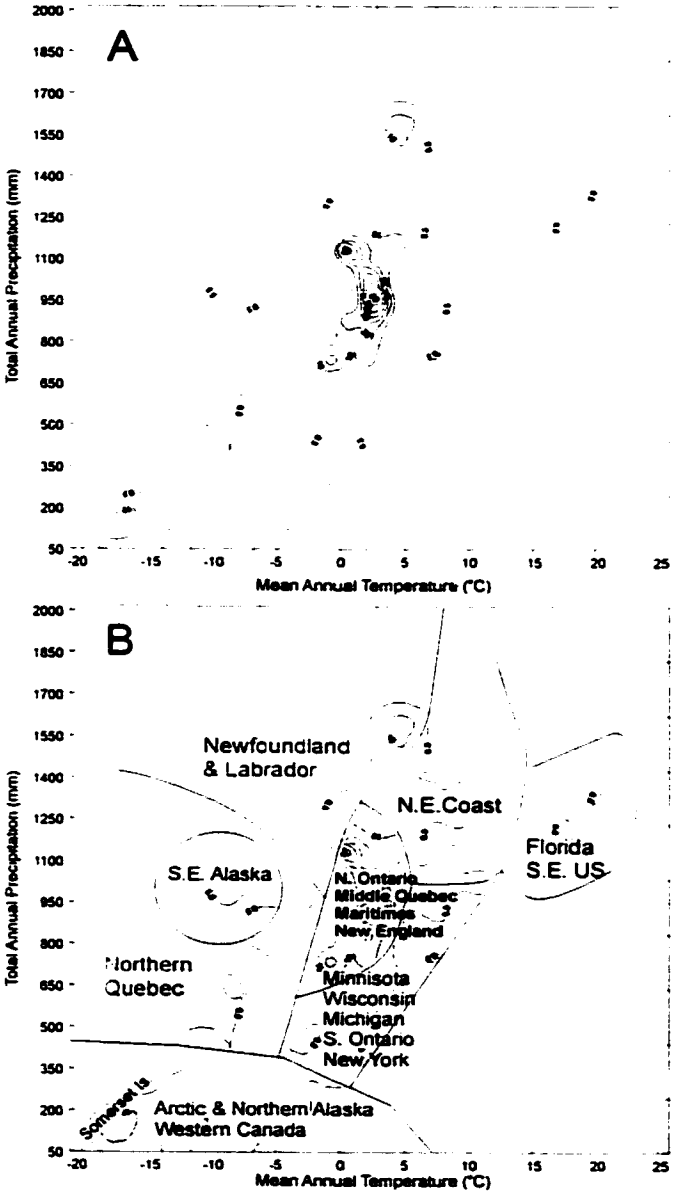


Figure 5.3: **A** – *Sphagnum* response surface in climate space; **B** – Geographic partitioning of *Sphagnum* response surface.

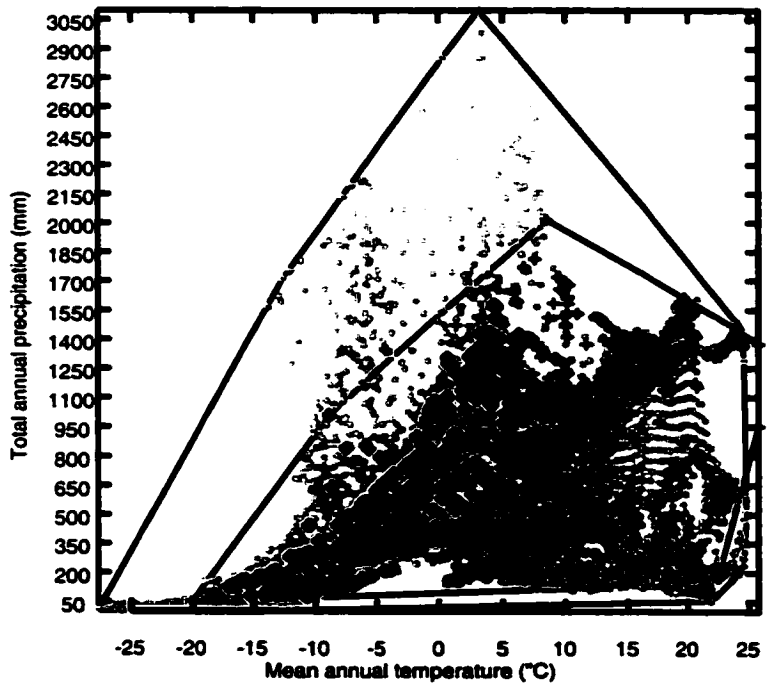


Figure 5.4: Degree of overlap between *Sphagnum* spore sampling sites and North American climate space. Solid outer line is the convex hull containing North American climate; the dotted line is the convex hull for *Sphagnum* sampling sites in North America. Small dots are individual grid points from Leemans and Cramer (1991) between 258 ±838N and 49.58 ±1688W. Black diamonds are sediment sample locations with positive *Sphagnum* percentages, and small crosses are sites with zero *Sphagnum*.

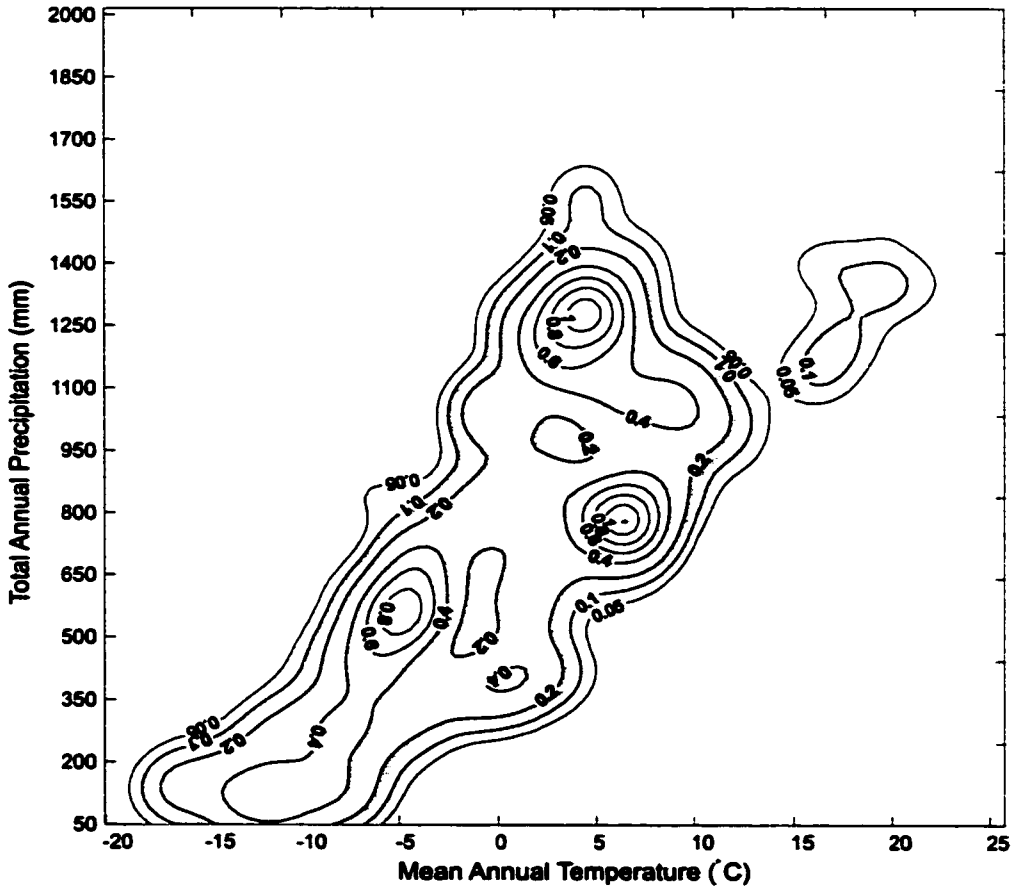


Figure 5.5: Sampling intensity surface for non–zero *Sphagnum* sites in climate space (contours represent points per 0.46 °C by 19.66 mm).

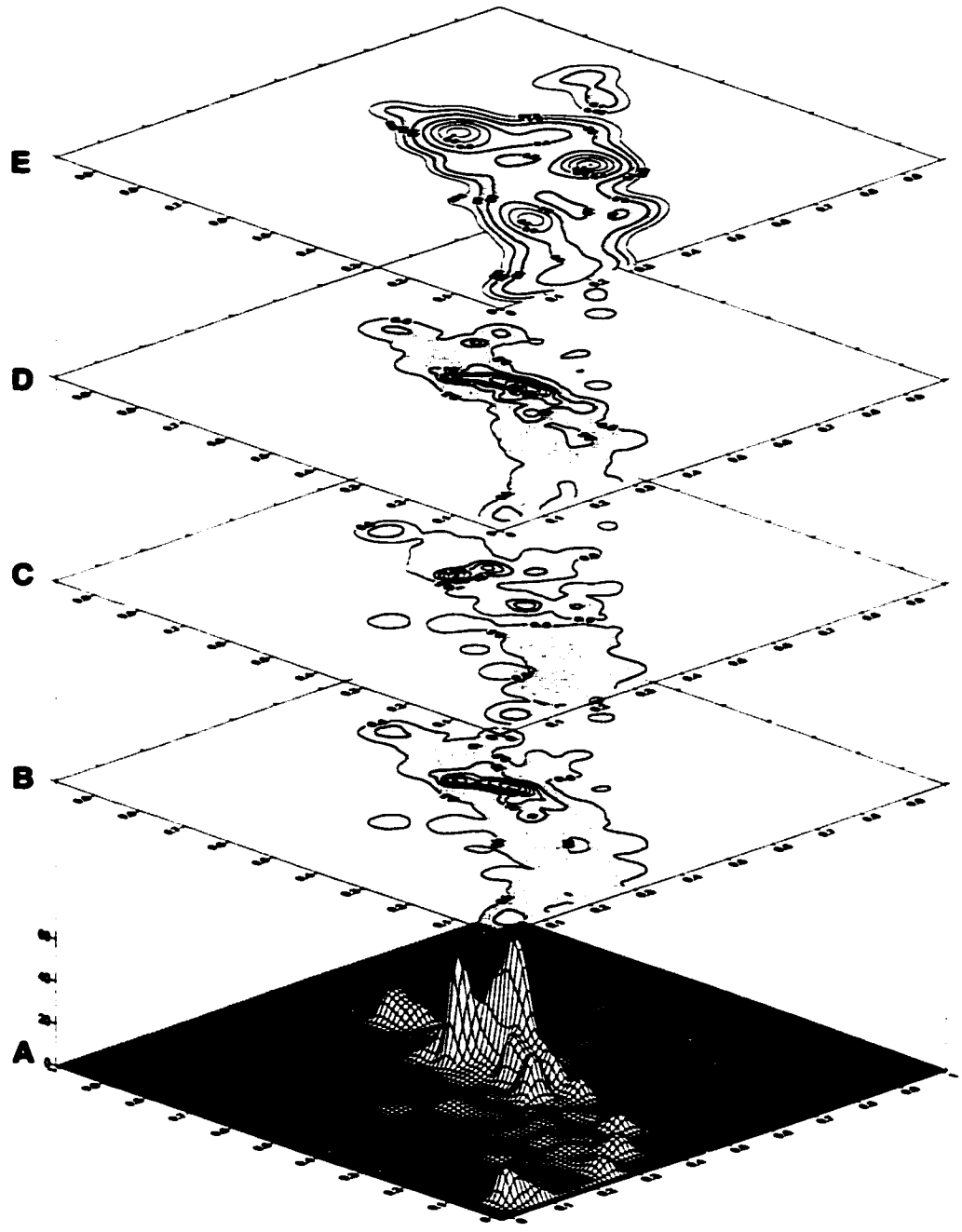


Figure 5.6: **A** – 3-*d* view of *Sphagnum* spore response surface. **B** – Same as Figure 5.3a. **C** – *Sphagnum* response surface from lakes only. **D** – *Sphagnum* response surface from bogs and other samples excluding lakes. **E** – Same as Figure 5.5.

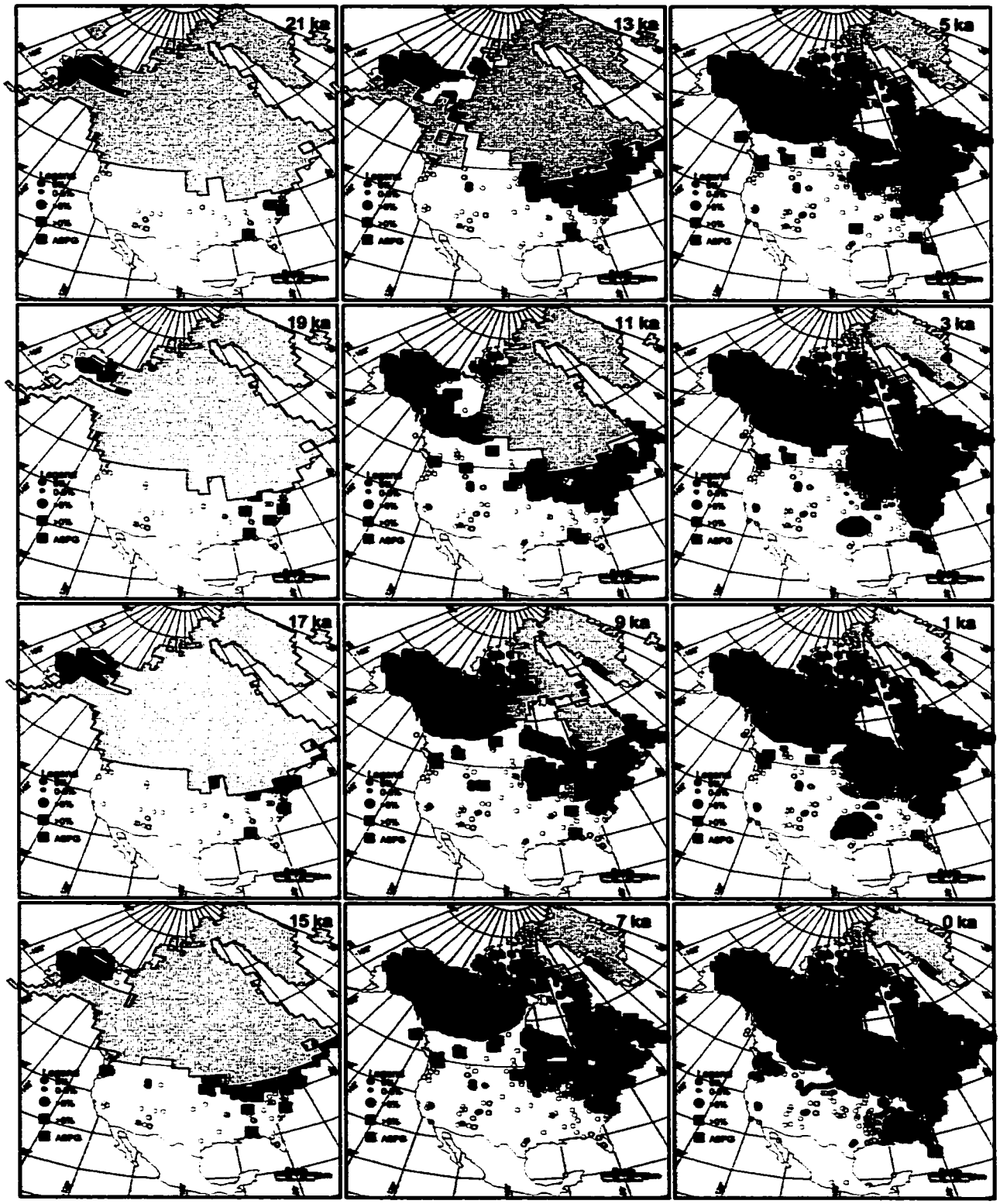


Figure 5.7. *Sphagnum* distribution for the past 21,000 ka at 2000–year intervals for North America. ASPG – Area supporting *Sphagnum* growth within the 0.5% contour is shaded. Ice sheets are grey. A 250 km^2 grey box is placed for each positive *Sphagnum* observation in order to visualize presence/absence patterns. Projection central meridian is 100°W, origin latitude 50°N; standard lines 35° and 80°N.

Chapter 6

THE BIOGEOGRAPHY AND PALEOCLIMATIC IMPLICATIONS OF AQUATIC MACROPHYTES IN NORTH AMERICA SINCE THE LAST GLACIAL MAXIMUM

6.1 INTRODUCTION

In recent years, many terrestrial pollen data have become available and these are now archived in public databases such as the North American Pollen Database (NAPD) (Grimm, 1999; Contributors to the NAPD, 1999). Aquatic pollen is in many cases also included in this database, and can serve for mapping and analysis of the past distributions if it can be shown that the modern aquatic pollen conforms to the modern range boundaries of the species/genera in question (Gajewski *et al.* 2001 – Appendix F).

This chapter demonstrates that the modern pollen of aquatic macrophytes conforms to the modern geographic ranges of aquatic plants. At the continental scale, this conformity provides evidence that changes in the past distributions of aquatic pollen and spores can be interpreted as changes in the range boundaries of the plants. Through mapping the distribution of modern and fossil pollen and spores in North America, this chapter explores the late-glacial and postglacial history of the most abundant aquatic macrophytes and provides an integrative interpretation of their biogeographic history.

At this large scale, the changes in abundance and distribution of aquatics are a response to continental scale climate changes during the late Quaternary. The past changes in range boundaries themselves are evidence

that aquatic pollen and spores have potential as proxy climate indicators. The congruence between the modern pollen and plant ranges allows the climatic tolerances that govern the geographic distribution of aquatic plants in North America today, to be interpretable within the fossil pollen/spore record. As such, aquatic pollen and spores can potentially strengthen our understanding and capability to reconstruct past climates at the continental scale.

6.1.1 AQUATIC MACROPHYTES

Many aquatic macrophytes are cosmopolitan in their modern distribution but their late Quaternary history is poorly understood. The proportion of aquatic taxa confined to a single continent is less than that of terrestrial vascular plants and this indicates their broad climatic tolerances (Stuckey, 1993; Sculthorpe, 1967, p.370). The moderating effects of the aquatic environment on air temperature and on edaphic variation allow longer growing seasons for aquatic plants than for their terrestrial counterparts (Sculthorpe, 1967, p.365). For example, in North America, *Typha* ranges from arctic to sub-tropical latitudes and *Myriophyllum, Potamogeton* and *Sagittaria*, among others, have circumpolar distributions (Johnson *et al.*, 1995, pp.215-226; Sculthorpe, 1967, p.366). The seeds, rhizomes or aquatic plants themselves can be quickly transported along waterways or by waterfowl and established far from parent populations under favorable environmental conditions (Stuckey, 1993; Ridley, 1930).

Aquatic macrophytes can be classified into four broad categories (c.f. Golterman *et al.*, 1988; Sculthorpe, 1967; Penfound, 1952): 1) emergent aquatic macrophytes such as *Typha* (Cattail) and *Sparganium* (Bur-reed) and marginally emergent taxa like *Sagittaria* (Arrowhead) and *Menyanthes trifoliata* L. (Buck-bean) which are found within shallow waters and saturated soils; 2) floating-leaved macrophytes, such as genera of Nymphaceae (Water-lilies – *Nuphar* and *Nymphacaea*), and Cabombaceae (Water-shields – *Brasenia schreberi*); 3) submerged rooted macrophytes such as *Myriophyllum* (Water milfoil), *Ruppia* (Ditch grass), *Isoetes* (Quillwort) and *Potamogeton* (Pondweed); and 4) free-floating rootless plants and mosses.

In most aquatic species, sexual reproduction is often secondary to vegetative reproduction from plant fragments or rhizomes, however, aquatic macrophytes shed pollen or spores as part of their reproductive cycle. These microfossils accumulate in lake and bog sediments producing a record of the regional/local presence of these taxa throughout a site's sedimentation history.

6.1.2 Biogeographic History of Aquatic Macrophytes in North America

There is a currently a lack of clear understanding of how aquatic plants with different life-history traits recolonized North America after the last glacial period. Aquatic macrophytes require water-saturated soil conditions and the study of their late Quaternary history may provide indications of moisture conditions in North America over the past 21,000 years (cf. Edwards *et al.*, 2000) and aid in wetland management strategies that aim to mitigate the effects of range changes in native aquatic plants under continued global warming. Maps of past distributions can lead to a better understanding of those areas capable of supporting paleo-wetlands and ultimately aid in providing better estimates of carbon sequestration during the glacial maximum as has been done for arboreal taxa (Peng *et al.*, 1998) and *Sphagnum* peatlands (Gajewski *et al.* 2001; Halsey *et al.*, 2000).

The continental scale distributional patterns of aquatic plants have not been studied in detail (Stuckey, 1993) until very recently (Dieffenbacher-Krall, 1998; Dieffenbacher-Krall and Jacobson, 1991). The ranges and migrational patterns of aquatic plants during the late Quaternary are little understood, which raises a number of questions. During the Wisconsinin glaciation, were these plants widely distributed in the ice-free areas of North America? Since the Last Glacial Maximum (LGM), how did these taxa migrate to their present ranges and at what rate?

Vesper and Stukey (1977) mapped the first occurrence of aquatic pollen at 32 sites in North America for *Myriophyllum*, *Nymphaea*, *Nuphar*, *Potamogeton*, *Sagittaria* and *Typha*. At the time of their study, they had few pollen data available west of the Appalachians or south of North Carolina and only one site in the west of North America (New Mexico). They found the earliest recorded aquatic pollen in North Carolina

and Virginia, but little evidence for aquatic plants along the Laurentide ice-margins during the LGM. They suggested that aquatic plants survived during the last glacial along the Atlantic coastal plain. From there they migrated up the east coast into Massachusetts and southwestern New York and subsequently into the Great Lakes region and Québec. The Appalachians were considered a significant barrier to westward migration. They speculated that a southern migration route existed from the lower Mississippi Basin via the Mississippi and Ohio River valleys into southern Indiana, Illinois and Ohio. They were not able to examine any western contributions to recolonization and their analysis suggests that additional sites would lead to a better understanding of the Late-Quaternary migration of aquatic macrophytes.

Dieffenbacher-Krall (1998) and Dieffenbacher-Krall and Jacobson (2001) looked at presence/absence patterns of aquatic macrophyte pollen and spores over the past 12,000 ^{14}C years in North America at 1000-year intervals. They found that aquatics showed rapid migrations into ice-free areas - partially corroborating Iversen's (1954) hypothesis to that effect. However, subsequent to their arrival at the northern-most limits, aquatics showed little response Holocene climate changes.

Using the record of aquatic pollen and spores in sediments, this study explores the patterns and rates of migration of aquatic macrophytes in North America, which in turn can lead to improved paleoclimate reconstructions.

6.2 Methods

6.2.1 Pollen and Spore Data

Data on aquatic pollen were extracted from all 782 fossil sample sites available in the North American Pollen Database (NAPD – June 1999 Version) (Grimm, 1999; Contributors to the NAPD, 1999). In the database, sites are classified according to their present environment, e.g., lake, bog, fen, moss-polster, etc. All sites that were not currently classified as lakes and bogs or fens were removed and this left 722 sites. Bogs and fens were included in the analyses because these sites may have evolved from lakes/marshes over

the postglacial. All sites with zero aquatic observations throughout their records were removed. This produced a database of 600 sites with fossil pollen samples.

The focus of most palynological work is the description of the upland vegetation and aquatic pollen are often recorded only for completeness, if at all. There are numerous sites in the NAPD with no aquatic pollen counts. A zero count can be recorded because aquatic pollen was not identified by the original author, aquatic pollen was not published or entered in the database, or aquatic pollen was absent from the site in question throughout the core sequence. Because these unknowns could adversely affect the analysis, any pollen diagram that did not have at least one observation of at least one of the aquatic pollen types during its record was excluded from further analysis. This reduced the number of remaining sites to 330.

Aquatic pollen percentages were computed within a set based on pollen and spores of all upland and aquatic plants. The 10 most abundant aquatic taxa in the NAPD were analyzed: *Brassenia schreberi* Gen. Pl., *Menyanthes trifoliata* L., *Ruppia maritima L.*, and the genera *Isoetes*, *Nuphar*, *Nymphaea*, *Sagittaria*, *Potamogeton*, *Myriophyllum*, and *Typha/Sparganium*. There were too few numbers of pollen grains of other aquatic taxa in the database to permit mapping and analysis. Most pollen cannot be identified to species, and only the genus could be analyzed in these cases.

Typha and *Sparganium* are genera of anemophilous (Dugle and Copps, 1972), emergent macrophytes that are closely allied in leaf morphology and floral organization (Smith, 1967; Sculthorpe, 1967, p.23). The pollen of *Typha* and *Sparganium* are morphologically identical (Nilsson and Praglowski, 1992, p.137; Punt, 1976; Kapp, 1969, p.203). *T. angustifolia* L. and *Sparganium* both occur as indistinguishable single grains (Grace and Harrison, 1986; Lewis *et al.*, 1983, p.125) and *T. latifolia* Linnaeus can only be distinguished when it is observed in a tetrad or dyad (Kapp, 1969, p.203), although such detail is not given in the NAPD. Moreover, *Typha* species hybridize prolifically (e.g., *T. x glauca* [*sensu lato*]) (Grace and Harrison, 1986; Smith, 1967; Hotchkiss and Dozier, 1949) and hybrids and backcrossed populations of *T. latifolia* have

been observed to shed pollen in monads, dyads and tetrads (Smith, 1967). If hybridization was an issue historically, then a tetrad might not indicate *T. latifolia*. However, *T. angustifolia* Linnaeus may be an introduced species (Grace and Harrison, 1986), but attempts to distinguish *Typha* in the pollen record may be premature. *Typha* and *Sparganium* are considered together since it is not clear that palynologists separate them consistently.

Radiocarbon chronologies were obtained from the NAPD. The radiocarbon equivalents of target calendar years spanning 0ka to 21ka by 1ka intervals were determined using CALIB version 4.2 (Stuiver and Reimer, 1993). Percentages of aquatic taxa were then linearly interpolated to these radiocarbon equivalents to produce a series of maps at 2 ka (calibrated radiocarbon years) intervals. For the 0 ka timeslice, 1161 surface samples were extracted from the Global Pollen Database (GPD) (Grimm, 1999) to which were added 90 core-top samples from the NAPD not present in the GPD.

6.2.2 MAPPING OF AQUATIC POLLEN

Maps were then drawn of the abundance of each pollen taxon at 2000-year intervals. On each map, proportional circles were plotted for pollen percentages based on the distribution of the non-interpolated observations over all sites and all time periods for each taxon. To aid in the delineation of past range boundaries, if a taxon is present at a given site, then a 250 km^2 rectangle was plotted with the proportional circle. The extent of the ice sheets was taken from model output by Peltier (1993, 1994) and the modern sea level and coastline were retained for simplicity.

Contemporary range boundaries for all genera with the exception of *Ruppia maritima* were digitized and brought into a Geographic Information System (GIS) in order to delimit the climatic ranges of the taxa along axes of mean annual air temperature (MAAT) and total annual precipitation (TAP). Climate data were obtained from a 0.5 x 0.5 degree longitude/latitude grid of Leemans and Cramer (1991) for all grid nodes within North America north of Mexico. Using GIS, all climate values within the modern range

boundaries for each taxon were extracted and these are used to depict the climate range that governs the geographic distribution within North America. Climatic values were extracted for all species of *Typha* and *Nuphar*, which illustrate the stronger climatic signal of individual aquatic species.

To explore the general pattern and rates of aquatic macrophyte migrations in North America, an isochrone map illustrating the dates of aquatic macrophyte initiation was constructed. However, many records in the NAPD are not complete if basal sediments were not reached during the coring process. Therefore, it was necessary to define what could be considered as a valid date of aquatic "initiation" or time when aquatics first arrive at a site. First, because the mapping begins at the time of the Last Glacial Maximum, for each of the 330 sites any record whose oldest date was 21 ka or older was retained for mapping. Second, any site for which the oldest level in a core was older than the date at which aquatics first appeared was also retained. This provided a dataset containing only those sites where the oldest date in the core was equal to the date of the first aquatic pollen/spore observation. For the remaining sites, if at the oldest dated level in a core (e.g., 14 ka) the site was within the ice sheet boundary (e.g., the 14 ka ice sheet) or was within the boundary of the ice sheet of the previous 1000 years (e.g., the 15 ka ice sheet) then the site was retained. The 268 remaining sites had their initiation times interpolated onto a 50 x 50 km^2 grid producing a map illustrating those areas in North America at which the pollen/spores of aquatic plants simultaneously appeared.

6.3 Results

6.3.1 Modern Climatic Tolerances

Maps of the contemporary range boundaries of the aquatic taxa show consistency with their modern pollen and spore distribution and thereby indicate the climate range supportive of each taxon. *Nuphar*, *Potamogeton and Sparganium* have the largest climatic range of the floating leaved, the submerged and the emergent macrophytes respectively (Figure 6.1). *Isoetes*, *Menyanthes trifoliata*, *Nymphaea*, *Nuphar*, *Brasenia schreberi*, and *Myriophyllum* tend to be excluded from warm dry environments, specifically

temperatures exceeding 20°C and precipitation less than 600 mm annually. *Menyanthes trifoliata* occurs in abundance only where temperatures are less than 20°C and greater than −15°C. Most of the genera do not occur below the −15°C isotherm. *Typha* and *Sagittaria* are not found at temperatures lower than −5°C and precipitation exceeding 1600 mm but *Sparganium* is absent from the warmer and drier environments where *Typha* and *Sagittaria* are more abundant. *Potamogeton* is the most cosmopolitan within climate space.

Generally the climatic ranges of the aquatic genera are broad, however, individual species of aquatic taxa exhibit precise climatic tolerances. For example, neither *T. angustifolia* and *T. domingensis* occurs in North America below the 0°C isotherm (Figure 6.2). Species such as *N. microphylla* and *N. robrodisca* are limited by the 0°C and 10°C isotherm while others such as *N. orbiculata* and *N. ulvacea* are found in very restricted regions of North American climate space (Figure 6.3).

6.3.2 First Arrival of Aquatics

Aquatic pollen were evident during the Last Glacial Maximum (LGM) in western Alaska, in the northwest and southeast United States (Figure 6.4). The migration of the aquatic taxa generally followed the pattern of deglaciation in North America. Aquatic pollen is first found in the ice-free areas south of the Laurentide ice sheet and in western Alaska. This is followed by range extensions along the ice-free corridor in Alberta, then into the Great Lakes and finally into Northern Québec after 7 ka. The time-varying distribution of aquatic initiations indicates that there was an increase in the frequency of aquatic initiations between 14-10 ka, which coincides with the most rapid wasting of the Laurentide ice sheet and the exposure of new environments.

6.3.3 Migration Patterns of Individual Taxa

Menyanthes trifoliata is a marginally emergent species that grows in shallow water on the edges of lakes, marshy ground and bogs (Mason, 1957). It has a circumboreal modern distribution and extends south to the central United States (Hultén, 1968). *M. trifoliata* has no pollen record during the glacial maximum (21ka) but is present in the Mississippi basin at 19ka (Figure 6.5). Only occasional pollen are recorded at 17, 13

and 11 ka in Alaska. The number of sites to the east and west of the Appalachians and in the southwest increased during the period 13-7 ka but remained within or very close to the modern range boundary of the species. During the past 5 ka, sites across the present-day range recorded pollen grains of this species. The modern species distribution in the boreal zone of Canada is poorly represented in the pollen record.

Sagittaria is an emergent found in saline and fresh waters, with species tolerant of a wide range of acidities, alkalinities and climates (Crow and Hellquist, 1981a). At 21ka, *Sagittaria* pollen was present at the Ohio-Mississippi Rivers confluence, in the Alabama/Chattahoochee basins and two sites in Oregon (Figure 6.6). In the east, *Sagittaria* first moved up the Ohio River basin and the east coast by 19 ka and by 13 ka into the upper Mississippi basin. Between 13 and 9 ka *Sagittaria* remained generally south of the ice sheet margin east and west of the Appalachians. It was also found in Alaska from 9-3 ka. During the past 5 ka, there was little change in its distribution. Modern day *Sagittaria* pollen, however, does not conform to its contemporary geographic range boundary in the west.

Typha/Sparganium: *Typha* is an emergent cosmopolitan and found in the subarctic, temperate and tropical climates of the Northern Hemisphere (Crow and Hellquist, 1981b; Mason, 1957). *Sparganium* is also rhizomatous and emergent in shallow waters across North America.

During the glacial maximum, *Typha-Sparganium* was present at one site in Alaska and across the continent south of the ice sheet east and west of the Appalachians (Figure 6.7). By 15 ka, pollen was found throughout Alaska, was abundant in ice-marginal environments in the south, and still found at sites across the remainder of the continent. Between 13-9 ka the taxon quickly moved into newly exposed ice-marginal environments in the ice-free Cordilleran region. Its absence in Manitoba, central and northern Canada, Keewatin and Nunuvut simply reflects the lack of samples. Today *Typha* is only abundant at temperatures 10°C greater than *Sparganium* (Figure 6.1) and many of the ice-marginal sites in the late glacial showing presence may be *Sparganium*.

Nuphar* and *Nymphaea: *Nuphar* and *Nymphaea* are floating-leaved rhizomatous macrophytes of the family *Nympheaceae* with relatively large and heavy seeds dispersed by waterfowl (Barrat-Segretain, 1996; Helsop-Harrison, 1955) and flowers adapted for insect pollination (Langanger *et al.*, 2000). Today, *Nuphar* has a larger geographic (Figure 6.8) and climatic range (Figure 6.1) than *Nymphaea* (Figure 6.9).

During the glacial maximum, *Nuphar* pollen was present at one southwest site, in the Chattahoochee/Alabama basins and the eastern seaboard (Figure 6.8). Between 19 and 15ka, *Nuphar* moved up the east coast and into the upper Mississippi and Ohio basins, including sites along the southern margin of the ice sheet. Throughout this period, *Nuphar* was recorded at two sites in Texas and New Mexico, although all other sites surrounding these report no pollen. By 7ka, *Nuphar* was found throughout its modern range and most of its western range was occupied by 11 ka. Nuphar was present in Alaska at 19 ka and 13 ka but did not become abundant until after the western part of the continent was ice-free.

Nymphaea was found in the Mississippi, Ohio, Alabama and Chattahoochee basins at the LGM but was not found along the east coast until 19 ka (Figure 6.9). The number of sites recording *Nymphaea* increased between 19 and 15 ka along the lower Mississippi and Ohio rivers, although there was a band just to the south of the ice sheet with no pollen at 13-11 ka. *Nymphaea* did not become abundant in the upper Mississippi until 11 ka. *Nymphaea* pollen did not appear in Alaska until 13-11ka and then disappeared until 1 ka. *Nymphaea* pollen is conspicuously absent within its modern range in Alaska and the west coast of North America.

Brasenia schreberi is the only species of this genus found today in North America. It has the most restricted climatic range of all the macrophytes under consideration (Figure 6.1). Its pollen is wind-dispersed (Wiersema, 1997) and the plant is likely dispersed by waterfowl that use it as a food source (Ridley, 1930, pp.491-2). The modern distribution of *B. schreberi* pollen conforms well to the range

boundary of the taxon, except for the absence of pollen from eastern Texas to Louisiana (Figure 6.10). *B. schreberi* was found in southern Florida and the lower Chattahoochee at the LGM and between 19-15 ka moved into the lower Mississippi and Ohio basins and along the east coast, extending north to just south of the Great Lakes. By 13 ka, *B. schreberi* was present throughout the Ohio basin and east coast but did not become abundant in the upper Mississippi until 11ka, after which it was found throughout the Great Lakes. Between 11 to 9 ka it arrived at the modern range boundary in the east and was present in its western range in Washington and Oregon beginning at 13 ka.

Myriophyllum is a submerged perennial of lakes, ponds and shallow water, especially in calcareous regions (Marsh, 1957). Sites in Alaska reported *Myriophyllum* between 21-15 ka, and the number of sites reporting this taxon increased substantially at 13 ka as the ice-free corridor in the west opened (Figure 6.11). South of the ice sheet there were many sites reporting *Myriophyllum* along the east coast and in the Mississippi and Ohio basins, including many ice-marginal sites by 17 ka. By 15 ka, numerous sites in the upper Mississippi recorded pollen. By 13 ka, despite the presence of the eastern and western Laurentide ice-lobes, *Myriophyllum* was present across most of its modern range as soon as it was ice-free, quickly responding and colonizing ice-marginal environments.

Potamogeton is a submerged anemophilous/hydro-anemophilous plant and in some species has floating leaves. It is widespread in lakes, streams, and coastal waters and is generally considered cosmopolitan (Mason, 1957). *Potamogeton* was found in Alaska at 21ka and abundant there by 15 ka prior to the separation of the Laurentide ice sheet from the Cordillera (Figure 6.12). During the LGM, it was also present in the Ohio basin and east of the Appalachians. By 15 ka, *Potamogeton* was abundant in the upper Mississippi and Ohio basins and the northeastern coastal United States. This taxon was present in all areas immediately after deglaciation and remained abundant throughout the Holocene, reaching its present range as soon as it was ice-free.

Isoetes is a pioneer that roots in sand and gravel a few cm to a few dm beneath the water surface, especially where there is low nutrient content (Taylor *et al.*, 1993). Between 21 ka and 15 ka, *Isoetes* spores were found along the Atlantic and Pacific coasts but not in the Great Plains or in the Mississippi and Ohio basins (Figure 6.13). Between 15-13 ka, *Isoetes* was found simultaneously in Alaska and the Cordilleran region. By 17 ka many sites in the southeastern United States had *Isoetes* spores. Beginning at 13ka, *Isoetes* appeared south and west of the Great Lakes region in the Mississippi and Ohio basins. It quickly followed the retreating ice sheet in eastern North America with abundant spores recorded at the edges of the ice sheet. By 7ka, areas throughout the modern range were occupied, although it was generally excluded from warm and dry environments in North America (Figure 6.1; Figure 6.13). *Isoetes* spores are poorly represented within the modern range boundary in western Canada.

Ruppia maritima is a robust and strongly rhizomatous submersed aquatic that can form extensive grasslike mats in saline and alkaline waters. Its modern range is mainly along the coast in the east from Newfoundland to Florida and in the west from Washington to Mexico (Hultén, 1968). Inland, *Rupppia maritimia* is found primarily in saline lakes (Hammer and Heseltine, 1988; Pip, 1979) where it forms, along with *R. occidentalis* S. Wats, the dominant submerged macrophytes in hypersaline lakes (> 60 g l^{-1}) in the Canadian prairies (Hammer and Heseltine, 1988).

The distribution of *Ruppia maritima* differs from all the other pollen types plotted in this paper and its continental distribution seems limited more by edaphic factors than climatic controls. There were no grains of pollen found in sediments older than 13 ka (Figure 6.14). Only occasional grains were found during the past 13 ka, and these only in saline lake sites in Minnesota, the Dakotas and Montana.

6. 4 Discussion

The distribution of modern pollen sufficiently delimits the geographic range boundaries for most of the aquatic macrophytes. In the present maps, taxa such as *Brasenia schreberi*, *Sagittaria*, *Menyanthes*

trifoliata, and *Isoetes*, are not observed everywhere within their modern ranges but pollen for these taxa are rarely found outside the range boundaries. Lack of samples or the depth at which surface sediment samples or lake cores were taken may both be responsible for some discrepancies (Edwards *et al.*, 2000). However, the general agreement between the modern pollen distribution and modern range boundaries in the present maps, suggests that they are suitable for delimiting the ranges of these taxa during the late Quaternary. The modern range boundaries of the aquatic macrophytes are also shown to be influenced by the macroclimate, suggesting potential for the use of fossil aquatic pollen and spore distributions in interpretation and constraint of past climate reconstructions.

The early presence of aquatic macrophytes in the southwestern United States during the LGM could suggest that the region was wetter at that time. Aquatic macrophytes may be used as a general indication of lake and/or wetland development at a continental scale. For example, the appearance of *Ruppia maritima* in the mid-west at 13 ka and its continued presence there may indicate the onset of aridity at that time. The presence of *R. maritima* in the fossil record has been used in a number of paleolimnological studies as an indicator of historical saline conditions (e.g., Laird *et al.*, 1996; Vance and Mathewes, 1994; Hickman and Schweger, 1993; Husband and Hickman, 1984).

Beal *et al.* (1982) suggested that broad-scale climatic patterns for some aquatic species control the conditions necessary for germination and establishment of aquatic plants. Winter temperatures may be important in the over-wintering of rhizomes in lake and pond substrates. Though climatic controls may be significant in the establishment and/or reproduction of some aquatic species, genera in this study have wide climate tolerances with respect to temperature in North America and this would allow them to more easily establish in ice-marginal environments. Temperature is a stronger limiting factor, such that most genera, with the exception of *Brasenia schreberi*, are limited by the −15°C isotherm. The arrival at modern range limits for many of the taxa early in the Holocene suggests that many of these taxa reached their climate limits at this time. Subsequent changes were smaller and movements of the range limits during the middle

and late Holocene would not be recorded at the scale and resolution of the maps presented. A lack of response to mid- to late-Holocene climate variations by aquatics in North America is consistent with findings made by Dieffenbacher-Krall (1998) and Dieffenbacher-Krall (2001) who worked on a finer temporal scale of 1000-year intervals.

Vesper and Stuckey (1977) had available sites only as far south as North Carolina. They found that the Appalachians were an effective barrier to the migration of aquatics into the Great Lakes region. This meant that there were two possible migration routes. The first was along the Atlantic seaboard from North Carolina/Virginia and then moving westward into Québec and the Great Lakes region around 11 ka after reaching Massachusetts and southwestern New York. The second route, to explain the occurrence of aquatic pollen in Ohio around 6 ka, they conjectured that aquatic taxa in the southern Mississippi could have moved north along the Ohio and Mississippi Rivers into Ohio, Indiana, and Illinois around 13-11.5 ka. However, their lack of data could not corroborate this idea.

The current and more extensive dataset employed here suggests that there may have been up to four migration routes prior to the separation of the east and west Laurentide lobes: a southward migration from Alaska between 14-13 ka; a northern migration in the west at the same time into the ice-free Cordilleran region; and two movements east and west of Appalachia as early as 19 ka for some taxa into the lower Mississippi and into the upper Mississippi and Great Lakes by 11 ka.

In the east, most aquatic taxa were present in the Alabama or Chattahoochee river basins at glacial maximum. Aquatics moved along the east coast and west of the Appalachians to areas near the Ohio-Mississippi River confluence by 17 ka. Taxa such as *Sagattaria*, *Typha/Sparganium*, *Nuphar*, *Nymphaea* and *Brasenia schreberi* then moved up the Ohio basin by 15 ka, after which they moved into the upper Mississippi basin by 13 ka as soon as the ice had retreated from that region. Submersed aquatics such as *Myriophyllum* and *Potamogeton* were found in the lower Mississippi/Ohio River basins at LGM and in the

upper Mississippi basin by 15 ka at numerous sites as well as in ice-marginal areas. *Isoetes*, however, stayed more or less east of the Appalachians until 13 ka. This may be due to its slightly different climatic tolerances when compared to the other submersed aquatic taxa. From these regions in the Mississippi/Ohio basins, areas of the east coast, midwest, Great Lakes and St. Lawrence lowlands were rapidly occupied by 11 ka as the ice sheet melted. From the lowlands in the Alabama and Chattahoochee River basins, the mountains would not have presented a barrier to migration since taxa could move both east and west of the Appalachians from this region.

In the west, numerous taxa (e.g., *Potamogteon*, *Menyanthes trifoliata*, *Isoetes*, *Myriophyllum*, and *Typha/Sparganium*) were present in ice-free Alaska during the LGM and migrated southward into the ice-free Cordilleran region by 13 ka. Pollen of *Nuphar* and *Menyanthes trifoliata* are rare in Alaska prior to 13ka but the presence of sporadic observations suggests that these genera may have followed a similar route. South of the Cordilleran ice sheet, *Menyanthes trifoliata*, *Typha/Sparganium*, *Myriophyllum*, *Isoetes* and *Potamogeton* were present at full glacial and/or prior to the opening of the ice-free Cordilleran and these moved northward during deglaciation to join the populations migrating southward out of Alaska. The observations of Vesper and Stuckey (1977) are valid but the current and more spatially extensive data suggests that there was a migration route out from the Mississippi basin and a further migration out of Alaska. The timing of these are also thousands of years earlier than their data suggested.

Early observations of aquatic macrophytes in Alaska may represent *Typha latifolia* and *Sparganium natans* L., *S. hyperboreum* Beurling ex. Laestadius, *S. angustifolium* Michaux., *T. angustifolia and T. domengensis* are more southern species (e.g., Figure 6.2) and the former may be an invasive introduced to North America in the early part of this century (Grace and Harrison, 1986; Hotchkiss and Dozier, 1949). The species common in Alaska and the west were likely *Nuphar polysepala* Engelm. and possibly *N. variegatum* Engelm. whereas species such as *N. robrodisca* Morong, *N. microphylla* (Persoon) Fernald. and *N. Advena* (Aiton) W. T. Aiton were involved in the east coast and Mississippi Basin movements into the

eastern United States. However, considering the climatic tolerances of *N. variegatum* (Figure 6.3), it too may have been involved in the east and spread westward. In the Alaska, *Nymphaea tetragona* Georgi. may have been important during the early periods of the Holocene. However, in the east, *N. ordata* Aiten. and *N. leibergii* Morong would have been implicated in occupying their modern ranges by 9 ka. For *Myriophyllum*, the Alaskan observations likely represent *M. exalbescens* Fern. and *M. verticillatum* L., whereas species associated with *M. heterophyllum* Michx. and *M. exalbescens* may have dominated the southern migration routes. Other genera, such as *Potamogeton*, have too many species and it would be difficult at present to determine which may have been dominant in the different regions in the past. To analyze the species involved in the various migrations will require the use of macrofossil data (cf. Jackson *et al.* 1997; Birks 1980) and this is a future goal when such databases are complete.

The submersed aquatics more quickly occupied ice-marginal environments followed by *Sagittaria*, *Typha/Sparganium*, and *Nuphar*. These taxa tolerate the lowest temperatures (Figure 6.1) and have the largest geographic ranges of the taxa considered in this paper (Figures 6.5 to 6.7). Such temperature tolerances and the moderating effects of the aquatic environment would allow them to quickly occupy ice-free areas. Alternatively, taxa like such as *Brasenia schreberi* and *Nymphaea* are not as temperature tolerant and are less frequently found in ice-marginal environments. In general, the rates of migration shown from the maps produced in this study for most taxa are less than or equal to the sampling resolution of 1000 yrs and reflect the rate of deglaciation in North America. Most taxa are found throughout their modern ranges between 9-7 ka.

The results demonstrate rapid dispersal of these aquatic taxa. Early Holocene drainage patterns were similar to today; the Mississippi flows south and the Mackenzie and Yukon Rivers flow north, yet aquatic plants rapidly moved south from Alaska and north throughout the Great Lakes region. The rapid movement of aquatics into ice-free areas of North America is analogous to the many observed modern

invasive aquatic movements, which can partly be explained by their life history traits, and the dispersal vectors.

Today, many aquatic plants are aggressively invasive, easily naturalized, or have undergone range extensions when habitats were modified. For example, *Typha* has rapidly colonized disturbed environments producing monospecific stands (Grace and Harrison, 1986; Morton, 1975). The South American *Eichhornia crassipes* (Mart.) Solms. (water hyacinth) (Penfound and Earle, 1948) and the European *Myriophyllum spicatum* L. have aggressively invaded aquatic systems since the late 1800's due in part to their effective reproduction via vegetative fragmentation (Hartleb *et al.*, 1993; Trebitz *et al.*, 1993). The European purple loosestrife (*Lythrum salicaria* L.), while not in a strict sense an aquatic plant, has spread rapidly in eastern North America (McCaughey and Stephenson, 2000), establishing itself over competitive native species like *Typha latifolia* (Weihe and Neely, 1997). Numerous eastern North American coastal marshes have reduced biodiversity due to range extensions of monospecific stands of the common reed *Phragmites australis* (Cav.) Steud. (Chambers *et al.*, 1999) and at the same time, this reed is disappearing in Europe (van der Putten, 1997). Cook (1985) cites numerous other examples of aquatic plant invasions and range extensions over the past century. It is clear that aquatic plants are capable of making rapid range extensions at the sub-century scale.

Seed dispersal can be effective over large distances in some taxa such *Typha*. Individual inflorescences can produce between 20,000-700,000 wind-dispersed seeds (Grace and Harrison, 1986; Smith, 1967). Once established, the taxon reproduces prolifically due to its investment in long spreading rhizomes (Grace and Harrison, 1986). A single *Typha* seed, once germinated, can produce a colony that can cover an area of 58 m^2 in two years (Grace and Wetzel, 1981). *Typha* is an early successional taxon and is tolerant of continuous inundation and seasonal drawdowns (Bedish, 1967). It is likely that their spread to ice-free environments occurred instantaneously considering that in recent history it has been observed to rapidly invade newly disturbed environments (Grace and Wetzel, 1998).

Other aquatic plants like *Potamogeton*, *Sparganium*, *Nymphaea*, *Nuphar* and *Brasenia schreberi* do not have airborne seeds and rely on water borne vectors or airborne vectors for long distance dispersal. For example, the seeds of *Nuphar lutea* can range from 0.46 g to 5 mm in size (Heslop-Harrison, 1955) with specific weights exceeding water (Barrat-Segretain, 1996). These species spread by rhizomes and fragments and their seeds are also dispersed by waterfowl (Helsop-Harrison, 1955; Barrat-Segretain, 1996). In general, aquatic plants act as primary food sources and shelter for waterfowl and other aquatic animals (Turner and McCarty, 1998; Esselink *et al.*, 1997; Mitchell and Wass, 1996; Søndergaard *et al.*, 1996; Smith, 1967). Ridley (1930) notes numerous observations made on the feeding habits of various North American and European waterfowl. Seeds attached to waterfowl (e.g., *Sparganium*) and seeds eaten by waterfowl that pass through the digestive tract intact would also permit rapid dispersal over long distances. Airborne dispersal vectors, e.g., waterfowl, may have been responsible for the rapid movement during the late Quaternary as was observed in the maps produced in this study. Some aquatic species have economic and cultural importance and may have been used by the Native Americans (Turner, 1981; Morton, 1975).

Finally, the reproductive strategies of many aquatic plants allow plant fragments, e.g., *Myriophyllum*, *Typha*, and *Nuphar* to successfully establish. These can be carried downstream or on the feet and feathers of waterfowl and lead to the establishment of populations far from the parent plants. This would be particularly important in the late-glacial and postglacial as large proglacial lakes were found along the retreating ice-margin and plants could have easily dispersed between major water bodies. The use of macrofauna databases (FAUNMAP Working Group, 1996) in future work could aid in the delineation of specific dispersal vectors thereby improving our understanding of rapid plant migration.

6.5 Conclusions

The modern distribution of pollen and spores from aquatic plants in North America conforms to their modern geographic ranges and this provides confidence in the use of fossil pollen/spore observations for

reconstructing past range boundaries of aquatic taxa in North America. The conformity also suggests that the climatic ranges defined by the modern geographic distributions of the taxa hold promise for the use of fossil aquatic pollen and spores in understanding climatic conditions during the late Quaternary.

The climatic signal of aquatic pollen and spores is demonstrated at the generic level and tends to be greater in the individual species, a consequence of the smaller geographic ranges of some species. Those taxa with the widest modern climatic ranges, mainly the submersed aquatics, migrated most rapidly into ice-marginal areas as the Laurentide ice sheet wasted, first in the west and then in the east of North America. The ice-marginal affinity of many submersed aquatic taxa may provide impetus for improved ice-marginal climate reconstructions. Climatic interpretation of past range changes in the fossil record would be greatly aided by species identification of aquatic plants using macrofossil evidence in conjunction with the pollen record. The narrow climatic tolerances of individual species (e.g., *Typha* spp. and *Nuphar* spp. Figure 6.2 and Figure 6.3) indicate the potential of fossil pollen and spores of aquatic plants to contribute to paleoclimate research, particularly if constraints can be found within independent macrofossil evidence.

The only previous study of aquatic plant migration over the late Quaternary was by Vesper and Stukey (1977) and recently by Dieffenbacher-Krall and Jacobson (2001). This current work corroborates the latter studies observations of rapid aquatic plant migration into ice-free regions of North America and substantiates Vesper and Stukey's eastern migration route while adding considerable new insights into aquatic plant histories in North America. The migration of aquatic taxa in North America presented here compliments these previous works and suggests that prior to the east-west separation of the Laurentide ice sheet there were up to four regions from which aquatic taxa migrated into ice-free regions. Genera of *Potamogeton*, *Isoetes*, *Myriophyllum Typha/Sparganium*, and *Menyanthes trifoliata* were present in ice-free Alaska during the Last Glacial Maximum (LGM) and moved into the ice-free Cordilleran region by 13 ka. Simultaneously, these same taxa could have moved northwards from Oregon and Washington. During the LGM, aquatic pollen and spores of submersed and emergent taxa like *Sagittaria*, *Nympheaceae* and *Brasenia schreberi* as well as submersed macrophytes were found south of the Laurentide ice sheet in Florida, the Alabama and Chattahoochee River basins, along the Mississippi and on the east coast of North America. By 15-13 ka most of the aquatic taxa had spread to the upper Mississippi and Ohio River basins, along the east coast, and by 11 ka into the Great Lakes, southern Ontario and Québec. Though the directions of migrations are in opposition to the hydrological drainage in North America, the rapid occupation of ice-free areas by aquatics may be explained by their wide climatic tolerances, life-history traits and dispersal vectors.

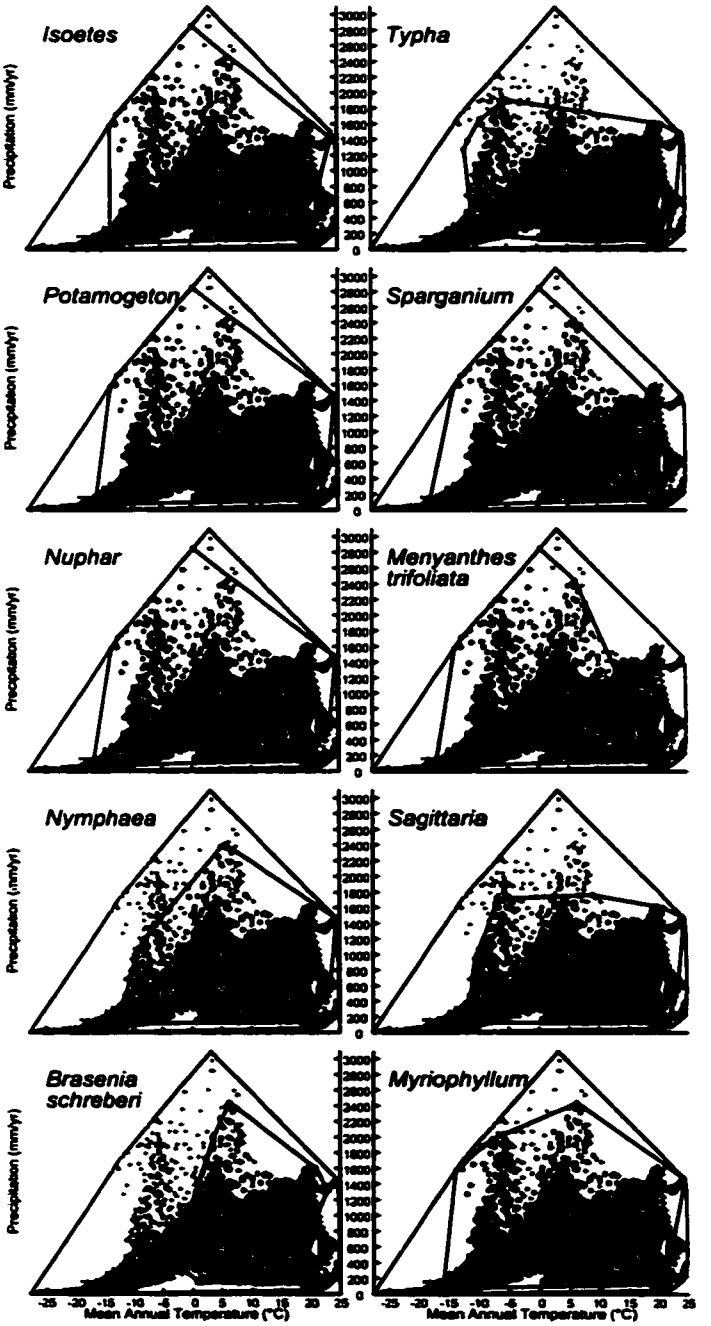


Figure 6.1: Climate range of North American aquatic taxa: Grey and black together define all combinations of mean annual temperature and total annual precipitation in North America north of Mexico. Black dots represent only those combinations within the range boundaries for each taxon. Convex hulls are illustrated for North American climate space (large hull) and individual taxon climate spaces (inside hull).

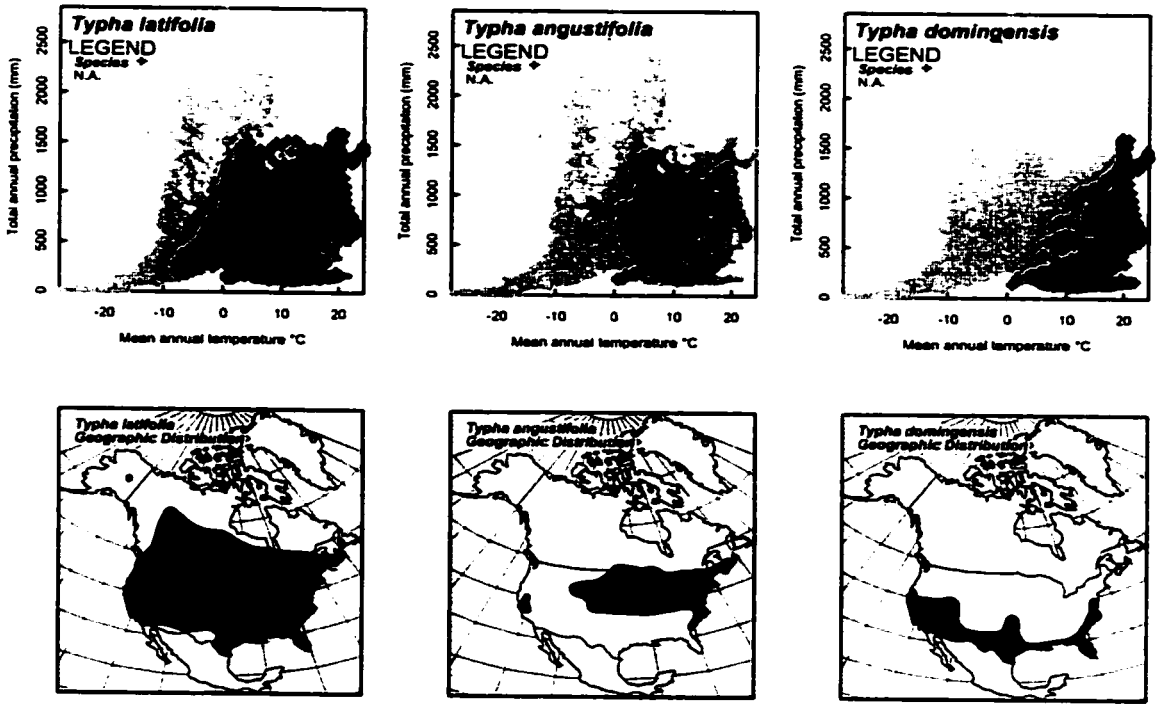


Figure 6.2: *Typha latifolia* L., *T.angustifolia* L., *T. domingensis* Persoon individual climatic and geographic ranges. Geographic ranges were taken from Grace & Harrison 1986, Porslid & Coody 1980, Lewis *et al.* 1983, Hotchkiss & Dozier 1949.

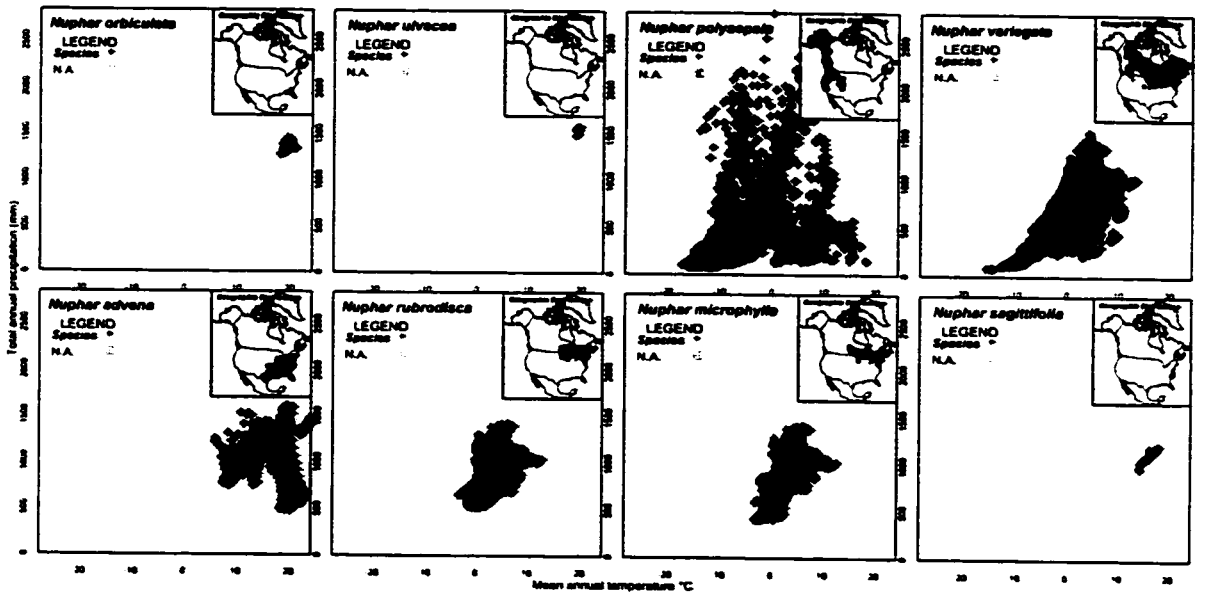


Figure 6.3: *Nuphar advena* Aiton., *N. orbiculata* (Small) Standley, *N. polysepala* Engel., *N. variegata* Durand, *N. sagittifolia* (Walter) Pursh, *N. ulvacea* (G. S. Miller & Standley) Standley, *N. rubrodisca* Morong, *N. microphylla* (Persoon) Fernald, Geographic ranges were taken from Wiersema & Hellquist 1997.

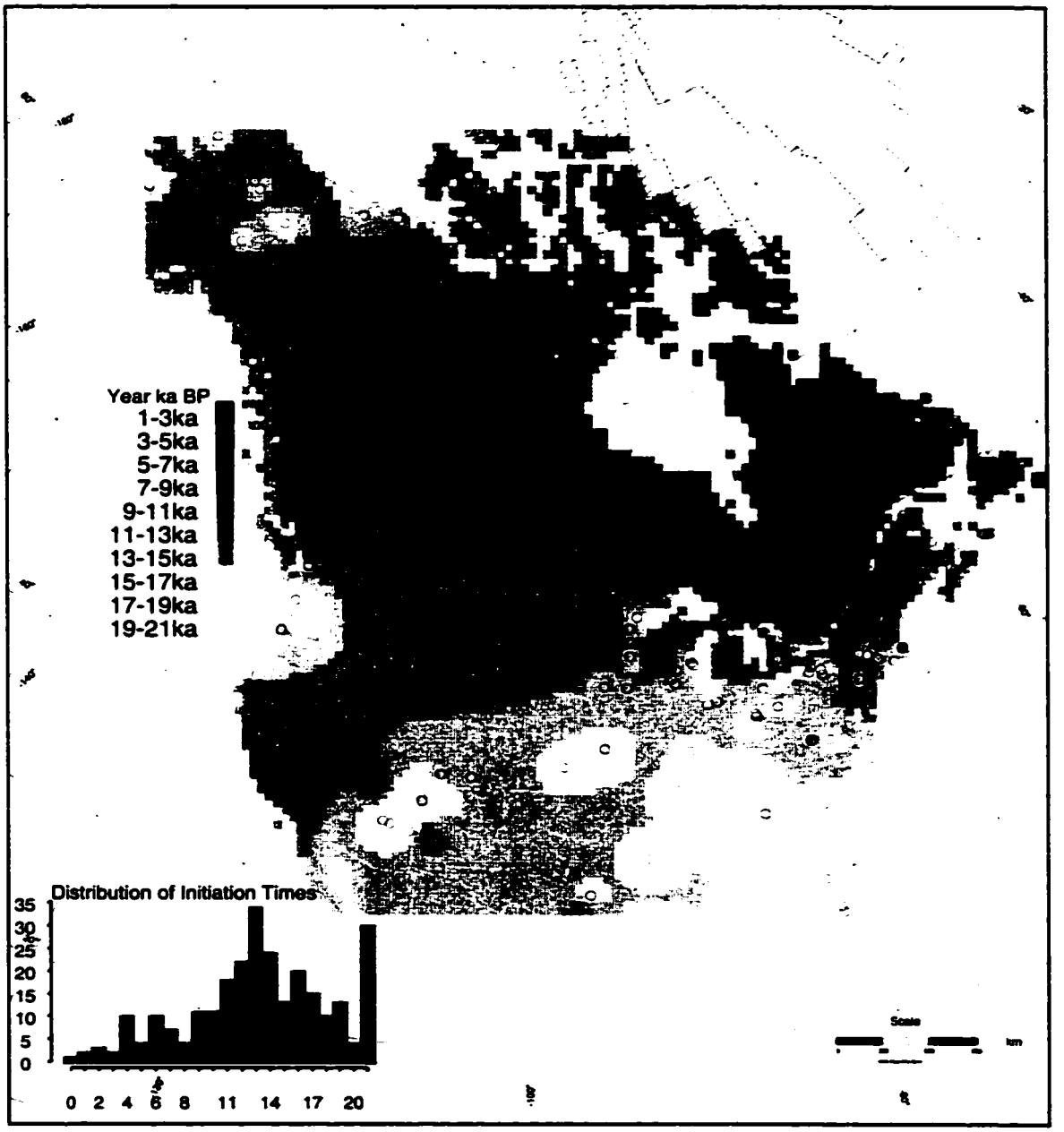


Figure 6.4: Isochrones of aquatic initiations in North America. ***Inlay:*** histogram of initiation times trough the last 21 ka with years on the x–axis and the number of sites at time t on the y–axis.

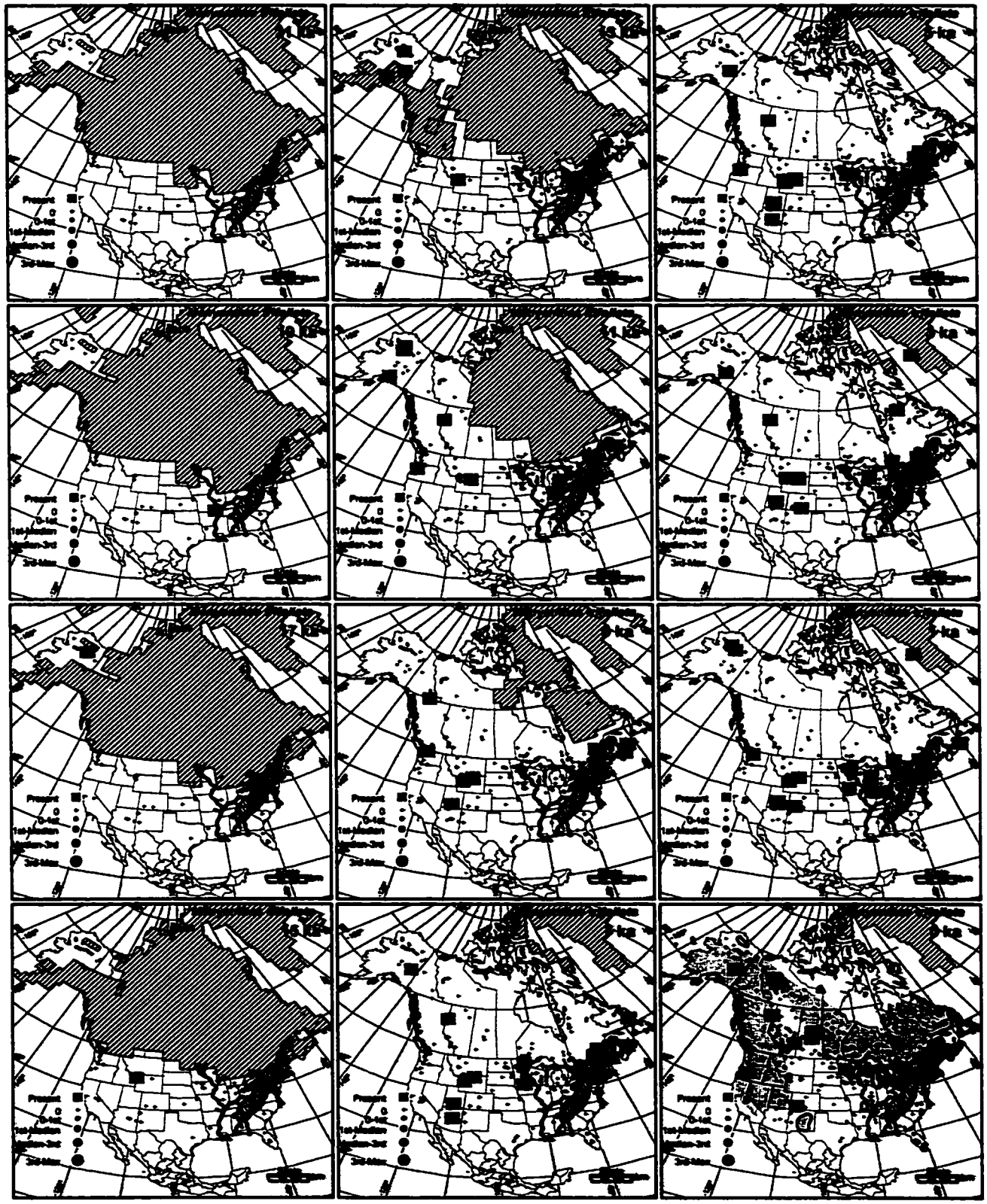


Figure 6.5: *Menyanthes trifoliata* with modern range boundaries modified from Hultén 1968. The dark lines in eastern North America are the Mississippi, Ohio, Alabama and Chattahoochee Rivers in the eastern US. Hatched region in the eastern US represents the Appalachian mountains (> 300 m). ***Legend***: Proportional circles represent the interquartile classes for the proportions of the individual aquatic taxon over all time periods.

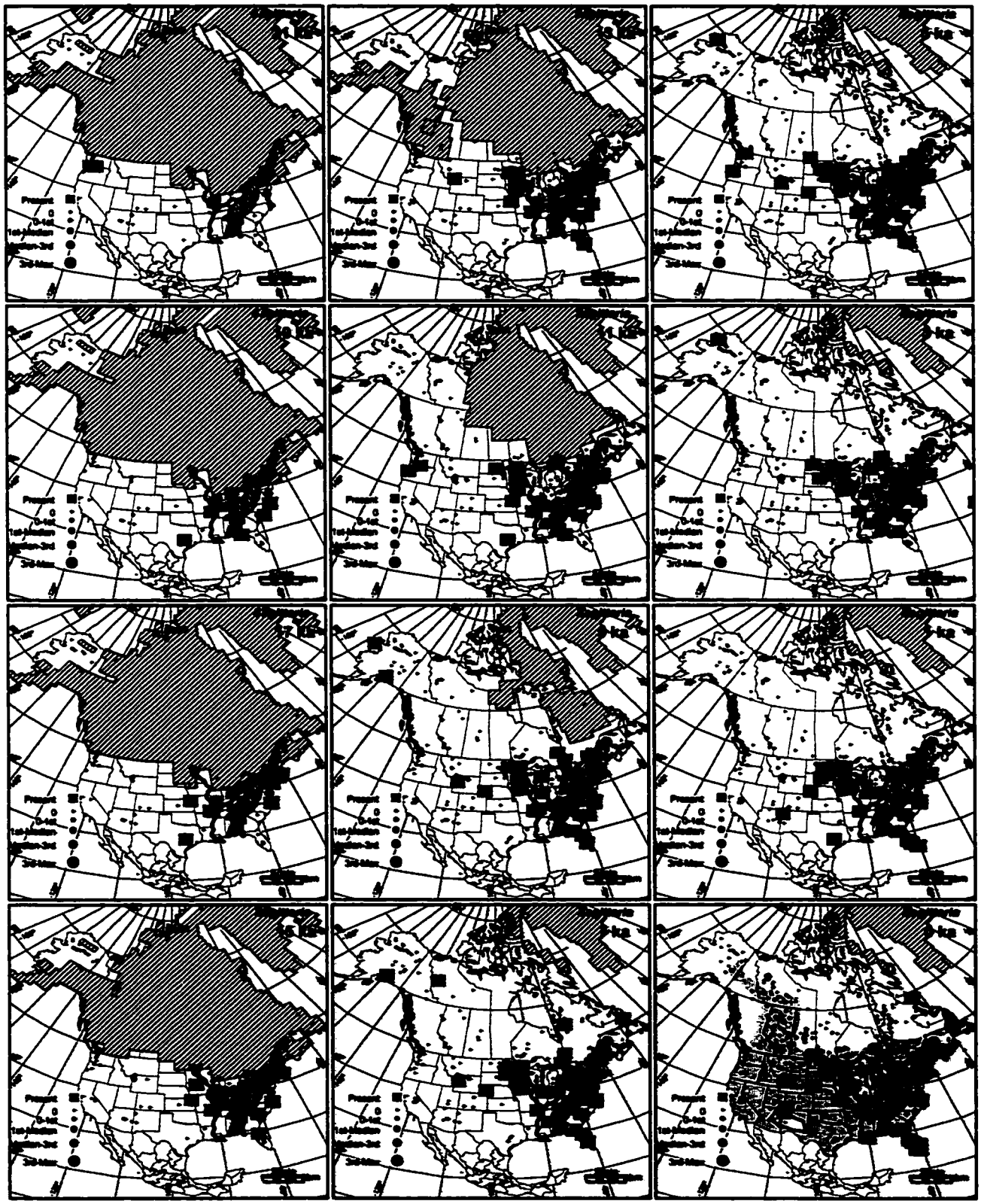


Figure 6.6: *Sagittaria* with modern range boundaries from Beal *et al.* 1982, Sculthorpe 1967 and Hultén 1968. See *Figure 6.5* for symbolization.

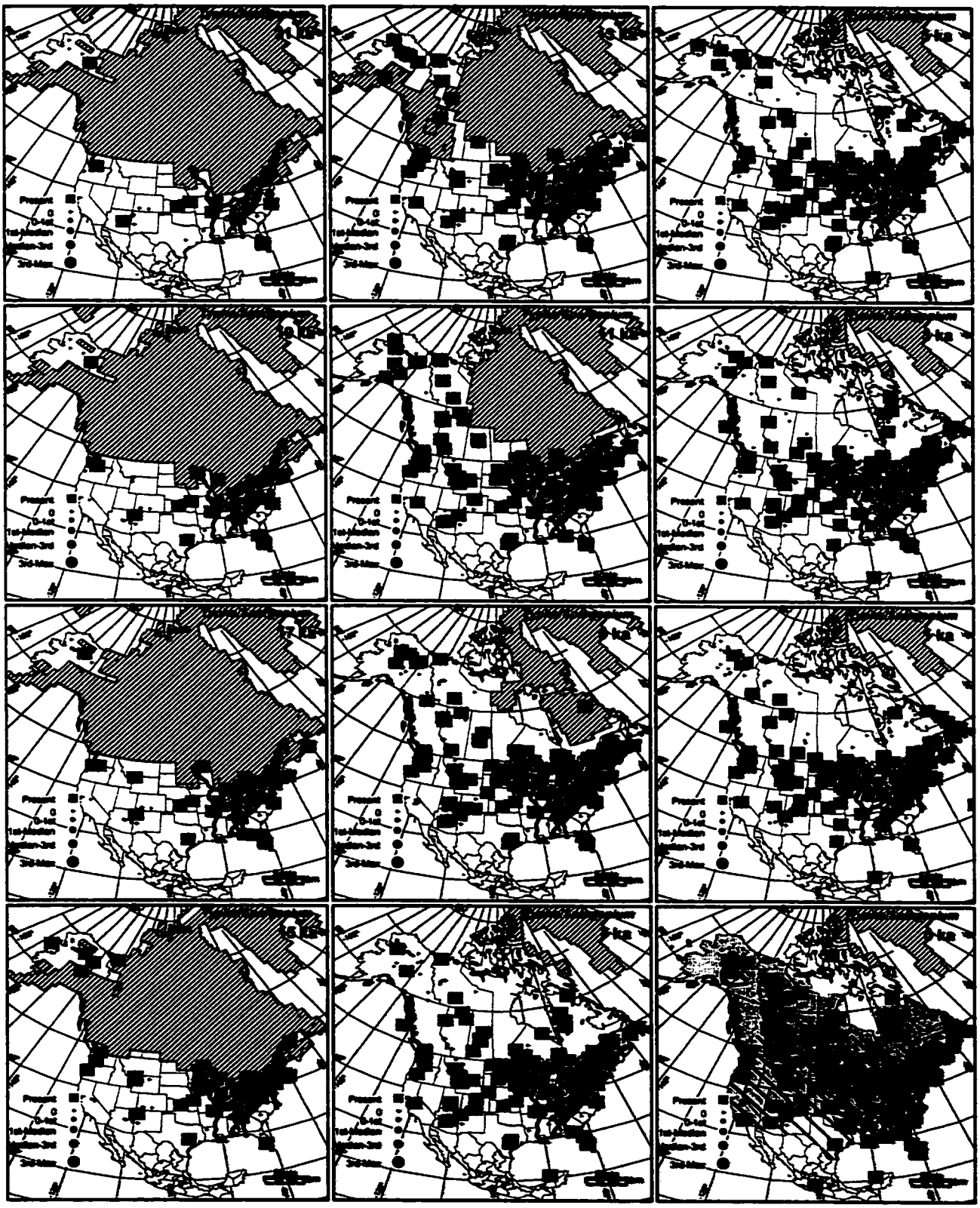


Figure 6.7: *Typha/Sparganium* with modern range boundaries for *Typha* (hatched pattern) from Grace & Harrison 1986, Porslid & Cody 1980, Lewis *et al.* 1983, Hotchkiss & Dozier 1949, and boundary for *Sparganium* (grey fill) from Hultén 1968 and Harms 1973.

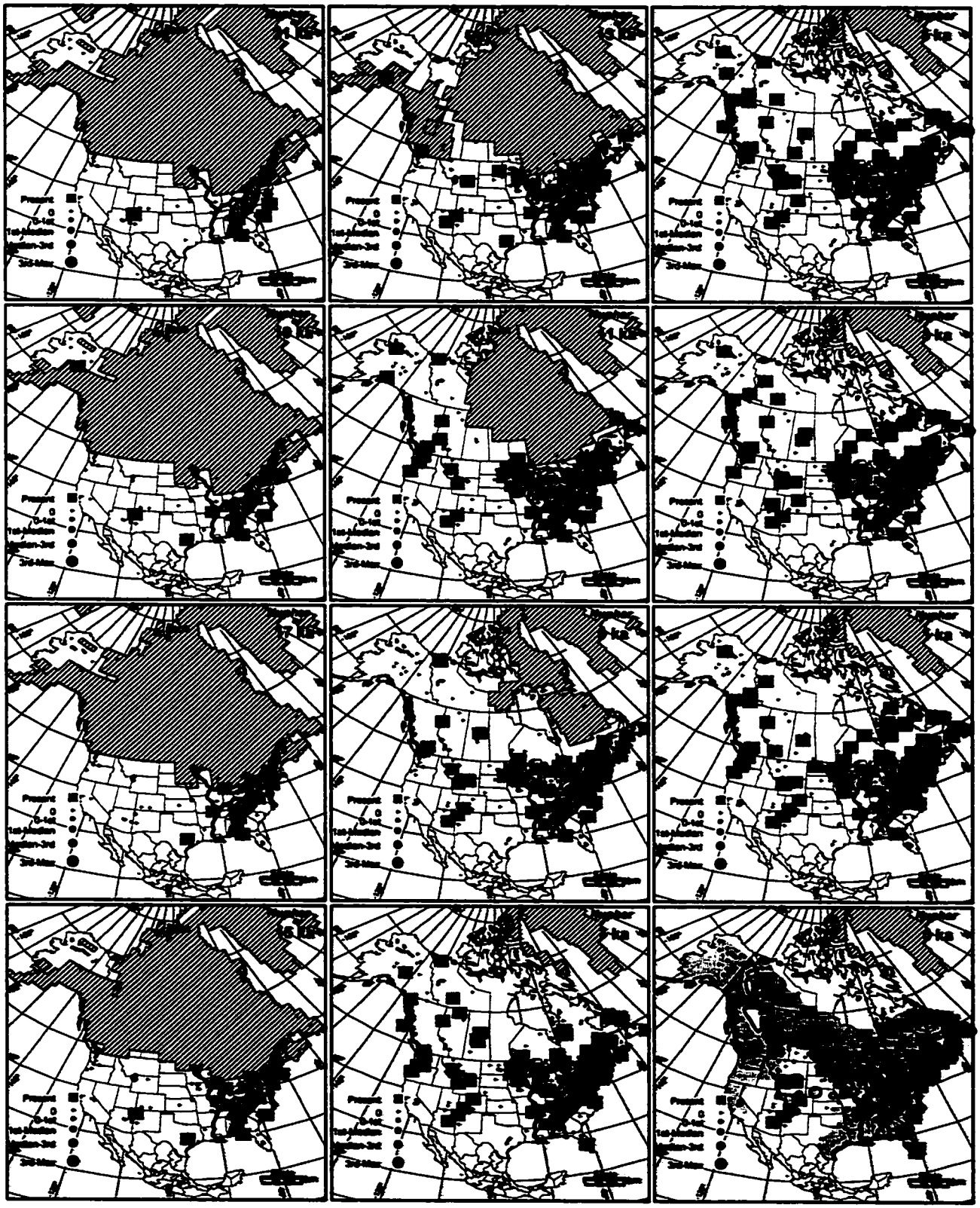


Figure 6.8: *Nuphar* with modern range boundary from Wiersema & Hellquist 1997.

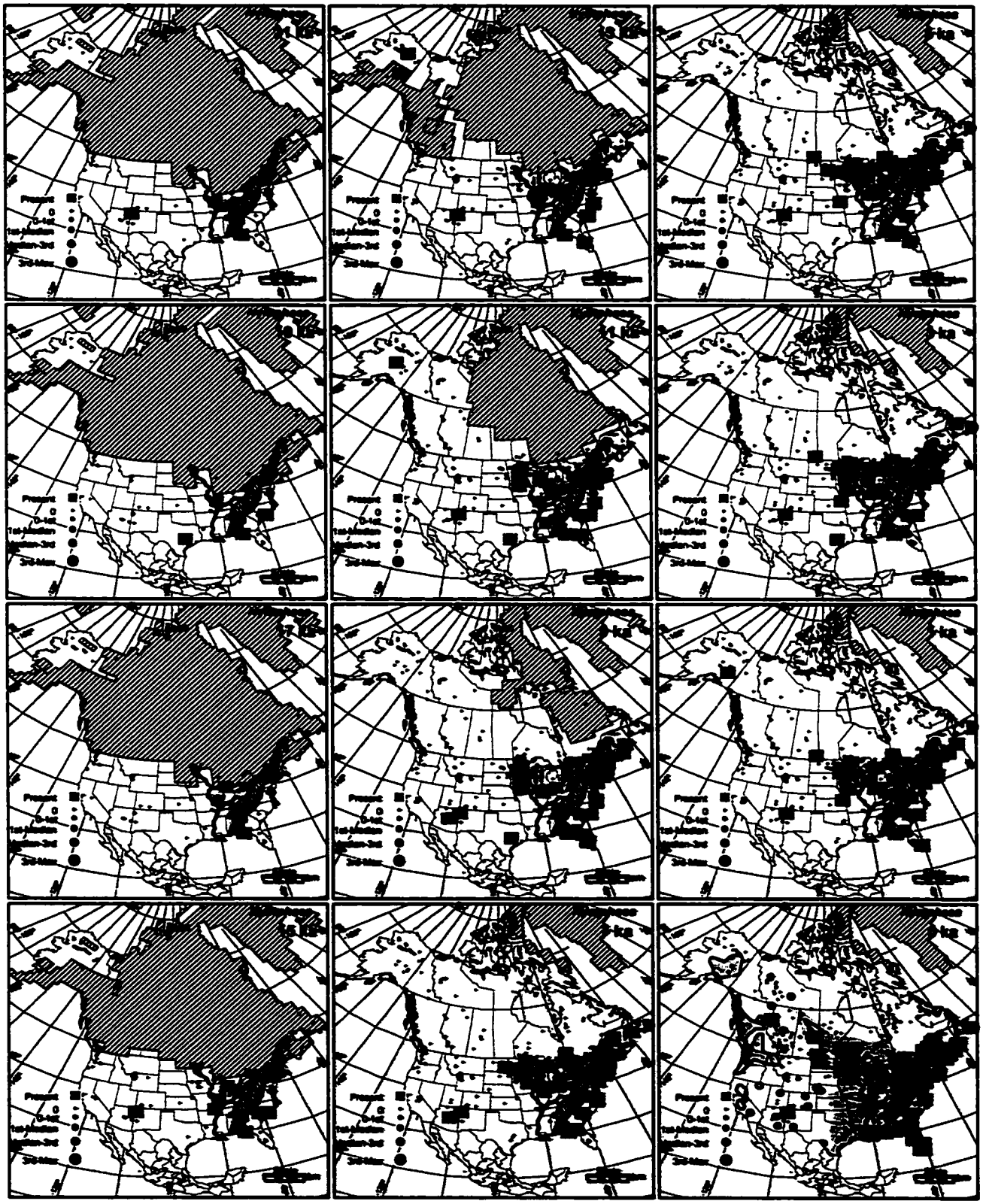


Figure 6.9: *Nymphaea* with modern range boundary from Wiersema & Hellquist 1997

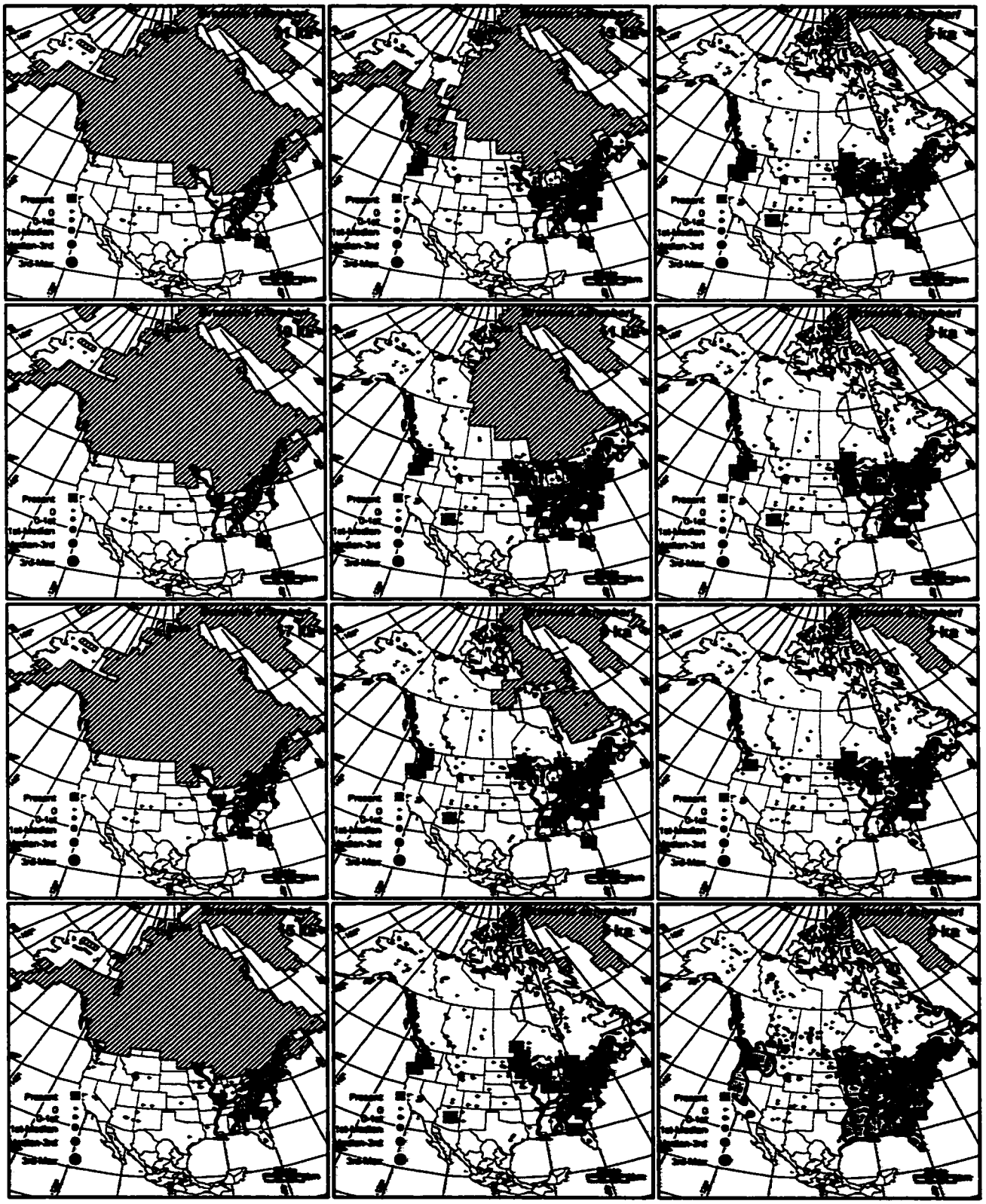


Figure 6.10: *Brasenia schreberi* with modern range boundary from Wiersema 1997.

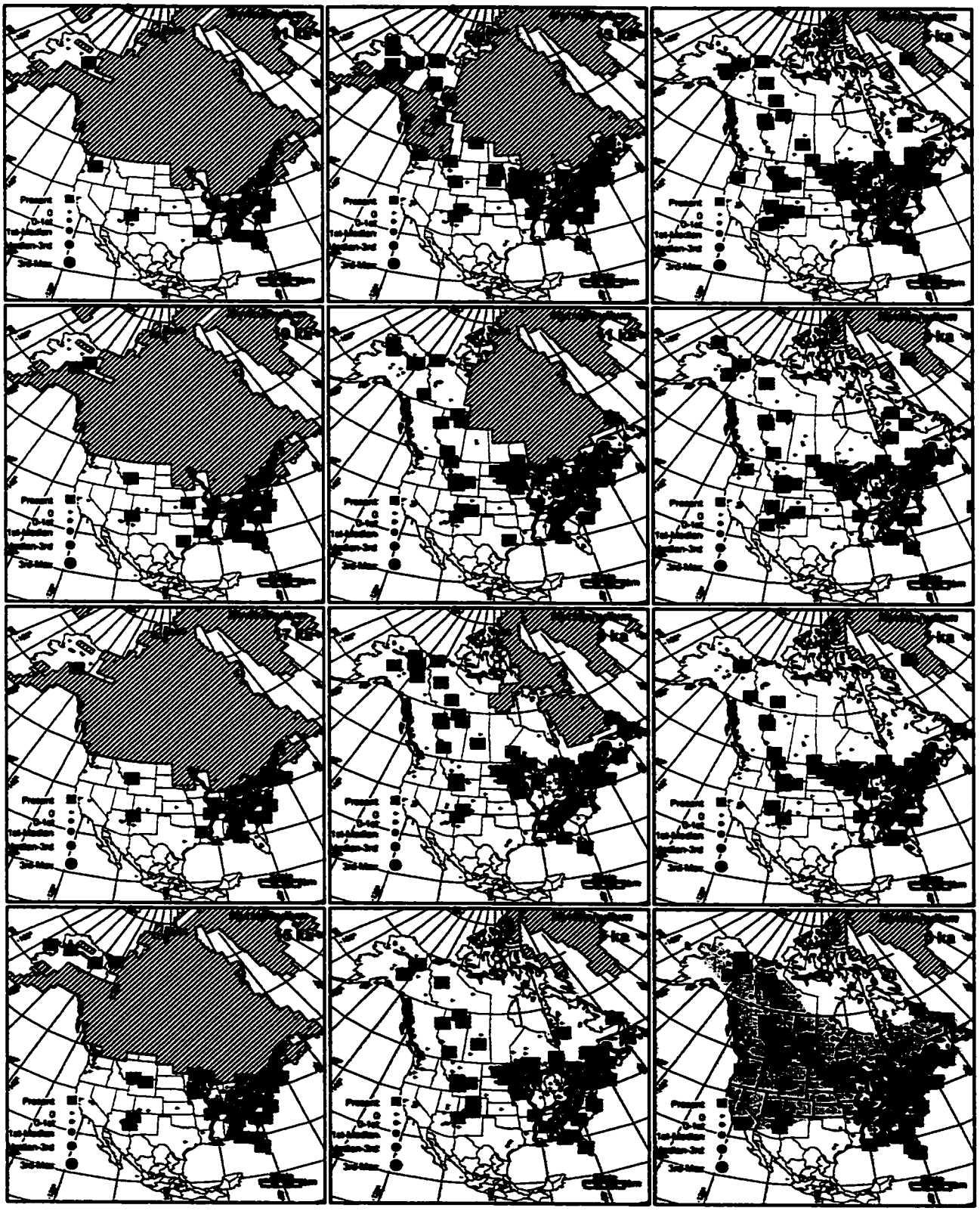


Figure 6.11: *Myriophyllum* with modern range boundary from Lewis *et al.* 1983, Porslid & Cody 1980, and Muenscher 1944.

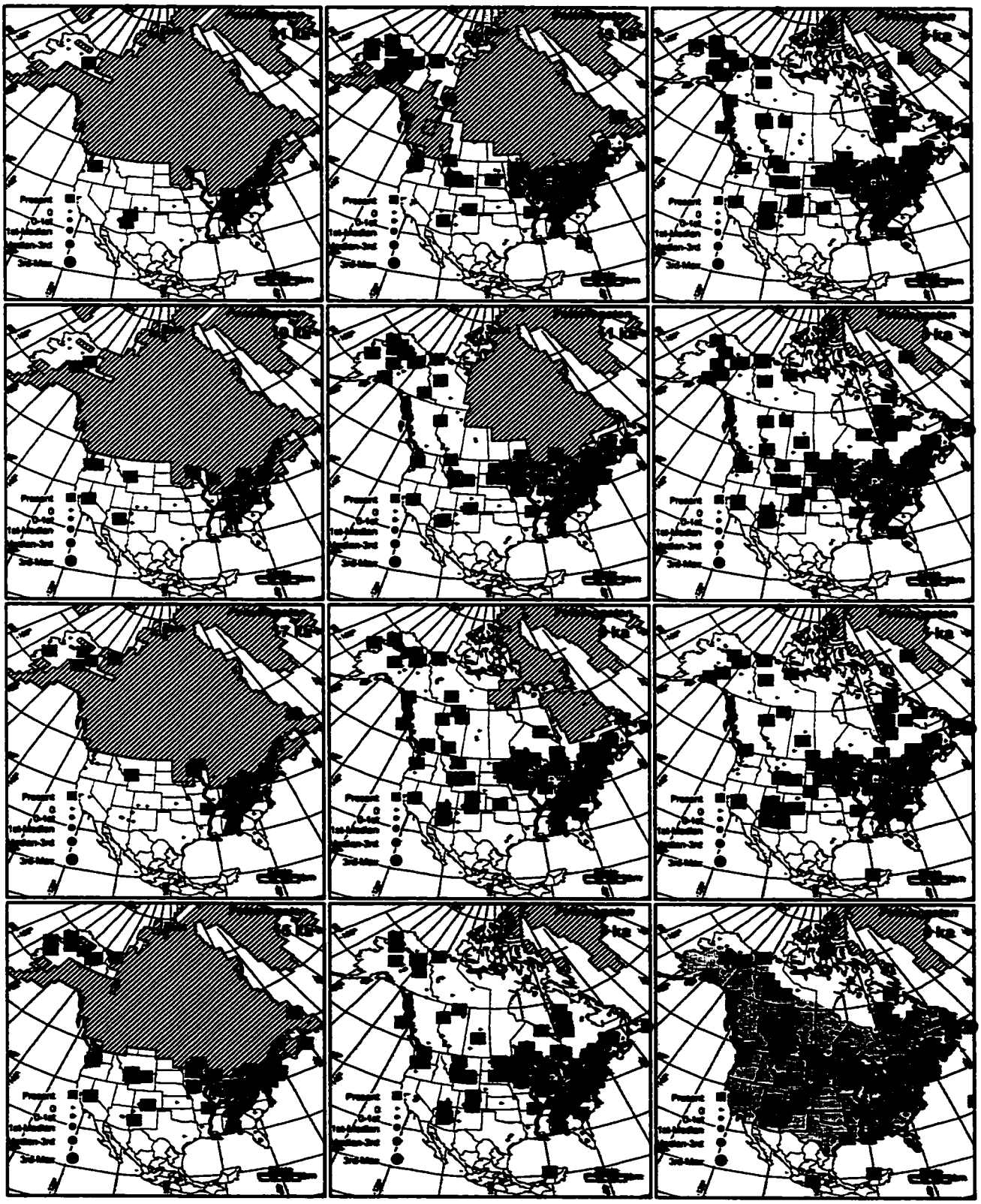


Figure 6.12: *Potamogeton* with modern range boundaries from Haynes 1974 and Lewis *et al.* 1983.

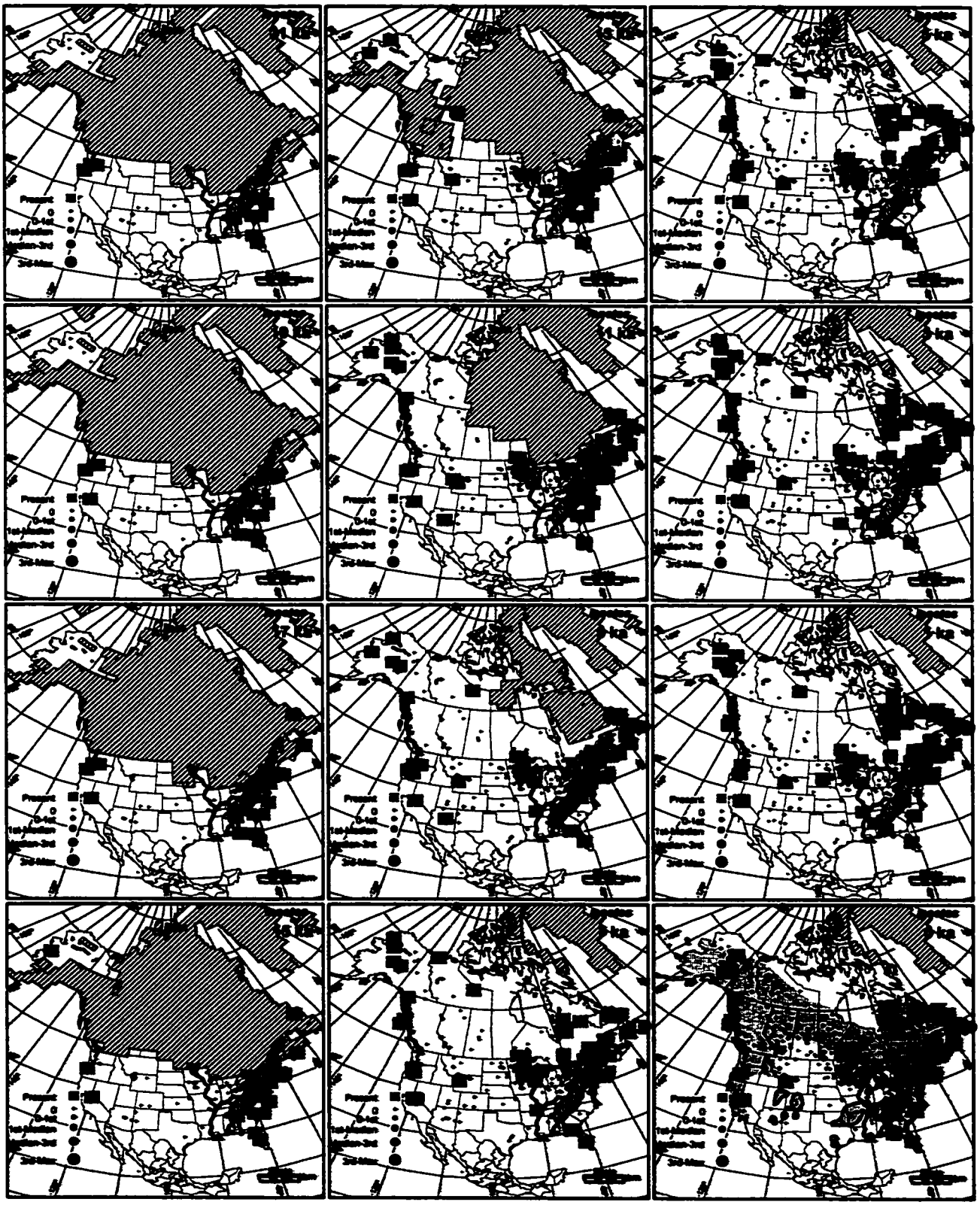


Figure 6.13: *Isoëtes* with modern range boundaries from Taylor *et al.* 1993.

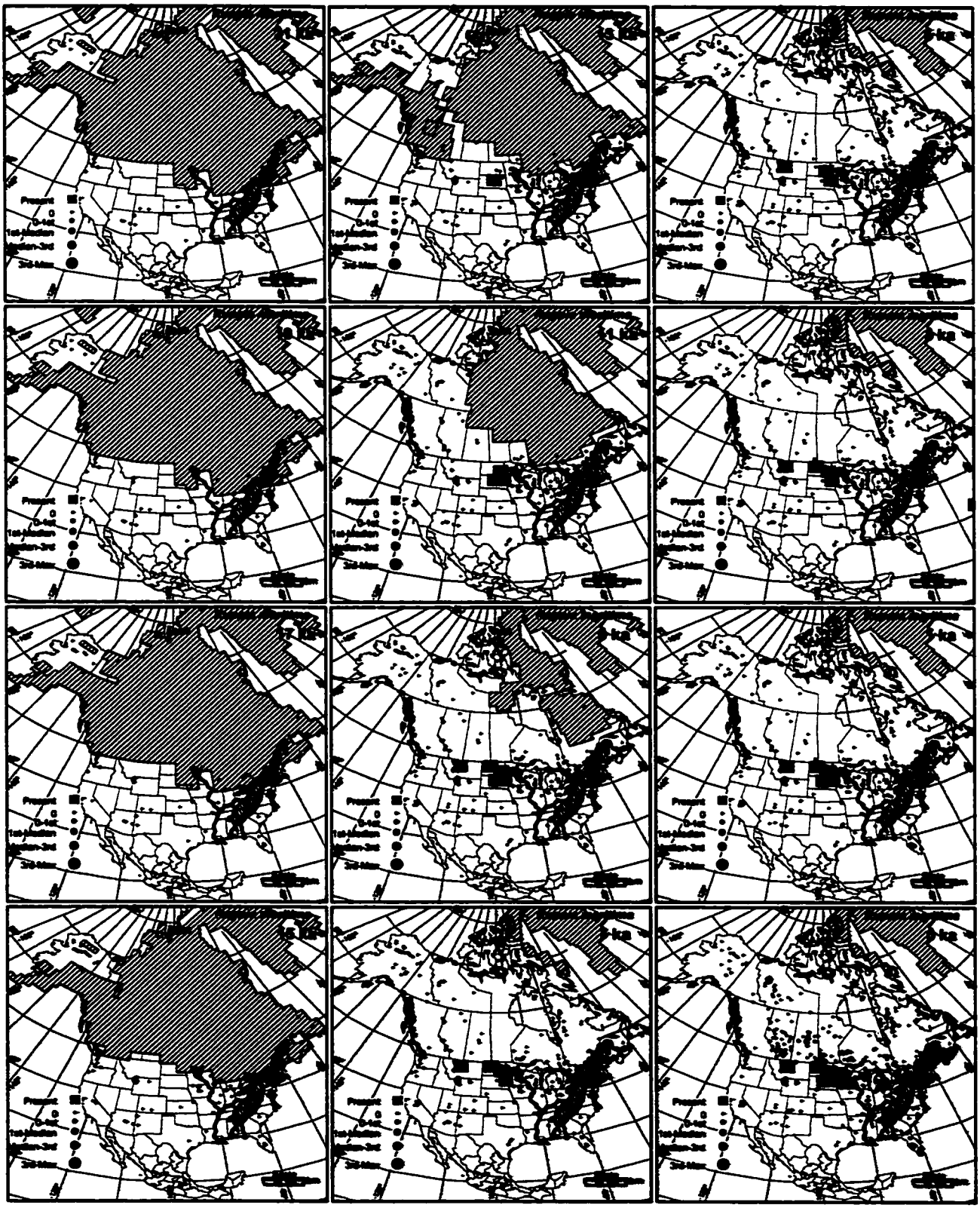


Figure 6.14: *Ruppia maritima.*

Chapter 7

7.1 Summary

The purpose of this dissertation was to improve our understanding of the late Quaternary climate of North America. The first chapters included a reconstruction of marine and terrestrial paleoclimates followed by a critical review of pollen source area and the modern analog technique (MAT). The latter chapters involved the exploration of wetland pollen and spores to assess their climatic signal at the continental scale and demonstrated their potential for paleoclimatic reconstructions while contributing to our understanding of the biogeographic distribution and migrational history of these taxa over the past 21,000 ka.

Paleoclimate model verification requires large-scale comparisons to fossil data. This dissertation took a first step towards a hemispheric synthesis of paleoclimates by undertaking a direct comparison of marine and terrestrial paleoclimates on the same spatial and temporal scales using the same quantitative paleoclimate reconstruction techniques. The quantitative climate reconstructions based on marine dinocysts and terrestrial pollen are consistent through the Holocene in northeastern North America. The combined and separate multivariate analyses of marine and terrestrial data from Hudson Bay, Labrador and the St. Lawrence Estuary differentiated tundra, boreal forest and deciduous forest assemblages in time and space demonstrating the large-scale climate signal present in both proxies. The marine pollen record reflected vegetation changes of the regional terrestrial environment and this raised a question regarding the spatial scale at which pollen samples reflect surrounding vegetation. The coherent pollen signal provided confidence in the derived marine chronological controls allowing us to directly compare quantitative climatic reconstructions of sea surface and terrestrial air temperatures from dinocysts and pollen respectively, utilizing the Modern Analog Technique (MAT). The climate reconstructions suggested that that the three geographic regions in northeastern North America had differing climate histories associated with their location with respect to deglaciation and air mass

boundaries. Prior to 6000 yr BP, cool temperatures reconstructed along the Labrador margins, both in the marine and terrestrial environments were in agreement with climate simulations that indicated the persistence of an anticyclone over the Québec-Labrador ice sheet. In both Labrador and northwestern Québec, a late Holocene cooling was evident at sites in the forest-tundra, but not in the boreal forest, suggesting movements in the mean position of the polar front over recent time. However, the high frequency climatic changes reconstructed for the St. Lawrence Estuary and Gulf, and a cooling reconstructed for the period prior to 8000 yr BP, were considered less reliable due to the larger values of the squared chord distance dissimilarity coefficients as fossil pollen became more different from modern pollen spectra.

Using fossil pollen sequences to reconstruct paleoclimate or vegetation at any spatial scale demands an understanding of how pollen is related to plant abundance. Most comparisons of pollen and plant abundance have been on small spatial scales (moss polsters) and only two regional-scale comparisons have been made using data from Wisconsin and Michigan. This thesis adds to previous studies by extending the pollen-plant abundance comparison into the Northern Mixed Hardwood - Boreal forests of North America.

If the relation between pollen proportions and plant proportions varies between different physiognomic regions, then inferences made from fossil pollen 'down-core' regarding surrounding vegetation abundance could only be made with regard to the dominant vegetation of the time. Chapter 3 shows that pollen-plant abundance relations between the Midwest and Québec are similar despite different regional vegetation types and suggests that the same inferences about pollen-plant relations can be made throughout a fossil core. Moreover, understanding the pollen source area of lake basins allows the determination of the spatial scale at which pollen spectra are influenced by surrounding vegetation. As such, the pollen source area of a given taxon determines the ecological/climatological scale of processes that can be studied and inferred 'down-core' from fossil pollen using basins of a particular size. Chapter 3 illustrates that pollen source area at the regional scale is still equivocal but the strong relations uncovered between pollen and plant abundance suggest that such lakes can be used to study paleoecologicical processes and landscape dynamics of the better dispersed taxa at a regional

scale. Furthermore, Chapter 3 presents the first comparisons of crown cover to plant abundance at a large regional scale in North America. Although crown cover should be more closely related to pollen productivity, Chapter 3 shows that crown cover and basal area record the same general aspects of plant abundance and are comparable in pollen-plant abundance studies. Moreover, the general pollen-plant abundance relations are robust and not strongly dependant on the plant abundance measures or plant abundance distance-weightings applied. These results and similar studies will eventually allow fossil pollen sequences to be calibrated in terms of plant abundance and allow more accurate interpretations of the spatial and physiognomic structure of paleovegetation at the regional scale.

Climatic reconstructions prior to ~8ka in the terrestrial and marine sequences in Chapter 2 were interpreted as less reliable due to greater values of the squared chord distance dissimilarity coefficient. Questions have arisen about the application of the MAT and how to identify non-analog fossil pollen assemblages and the critical limit of squared chord distance (SQD) that identify such non-analogs.

Chapter 4 identified how the limitations of the modern pollen or other proxy database as well as the assumptions of the MAT can affect the robustness of the climate reconstruction and the critical limit of squared chord distance that identify non-analog pollen assemblages. The choice of pollen set is a critical decision in the reconstruction of continental scale paleoclimates. Larger pollen sets better distinguish between different vegetation zones in North America because the rare types provide regional signals. Stochastic simulation was used to show that the critical values of squared chord distance must be adjusted for when utilizing larger pollen sets, and so the SQD limit of what may be considered a good analog increases with the number of pollen types employed in the reconstruction. Moreover, the critical limit of a good analog as measured by the SQD coefficient increases because larger pollen sets contain more zero observations for a greater number of continentally rare but regionally important taxa. The expected value of dissimilarity between pollen spectra increases with the number of types with zero values in the pollen set. The statistical analysis of the modern temperature and precipitation anomalies contributed to our

understanding of the degree to which spatial biases may be imposed on the MAT as a cause of the spatially varying density of the modern sampling network and identified areas that require denser sampling. This is reflected in the modern climate spaces that are sampled by the modern site network. The climatic reconstructions for the 6 ka are consistent with model simulations for a warmer continental interior; however, a moister signal in the west is equivocal. Under two different pollen sets, the spatial distributions of temperature and precipitation anomalies are different and this is important if the goal of the study is comparison with paleoclimate model simulations. The methodology presented in Chapter 4 can be applied to other proxy datasets in order to better understand the limitations of specific paleoclimate reconstructions.

The first step to utilizing new elements of proxy climate data requires that the quality of data be explored and biological-proxy-climate relations be demonstrated. Chapters 5 and 6 demonstrate that aquatic and wetland taxa exhibited changes in their distributions and ranges over the late Quaternary at the continental scale. Changes in range boundaries over the late Quaternary are evidence of a response to a large-scale forcing, namely climatic changes during this time. As such, these chapters represent the first step in our understanding of those underutilized parts of the pollen record.

The findings of chapters 5 and 6 are significant from a number of perspectives. Climate has been shown to play significant roles in the continental scale distribution of arboreal taxa; however, non-arboreal taxa such as *Sphagnum* and aquatic macrophytes have not been extensively studied in the fossil record at continental scales. Contemporary climate has been shown to play an important role in the regional distribution of peatlands and Chapter 5 provided evidence that this climatic signal is evident in the distribution of *Sphagnum* spores in North America. This study of past peatland distributions utilizing *Sphagnum* contributed to a better understanding of paleoclimate in North America and the role of peatlands in carbon storage (Appendix F).

Aquatic pollen are frequently recorded in palynological studies, however they have not been systematically studied at the continental scale except recently by Dieffenbacher-Krall and Jacobson (2001). As expected most of the aquatic taxa exhibited wide climatic ranges that were delineated strongly by temperature gradients at the continental scale. The aquatic taxa are not found in very warm and dry or in cold and dry environments. Moreover, it was demonstrated that aquatic pollen is not found outside the range boundaries of the aquatic plants at a continental scale and so aquatics can be useful paleoclimate indicators for regions or time periods where pollen transport of arboreal taxa may be problematic. The migration pathways were defined and it was demonstrated that aquatic taxa apparently migrated quickly following deglaciation. Particularly interesting was the distribution of aquatic taxa and their migration from the Midwestern states, Canada and Alaska where these taxa were found at full glacial. Future work will compare these migration rates to those of arboreal taxa as the rapidity of aquatic movements may be useful for understanding rapid plant migrations in general (Ptelka *et al.*, 1998; Clark *et al.*, 1998).

7.2 FUTURE RESEARCH

As expected, although this work resolved many questions it has raised many others. Pollen source area is not an unequivocal concept and there are many unresolved issues regarding the degree to which vegetation patchiness affects the ability of lakes to sense the spatial variation in vegetation pattern. Although much of the recent work has investigated this question at very small scales, vegetation patterns exist at many time and space scales. Future work at the landscape scale needs to explicitly incorporate spatial relations, for example using geostatistics, to better quantify pollen-vegetation relations.

Although the MAT is now routinely used for paleoclimate reconstructions, this work has shown that many questions remain regarding its application. Basic questions about the optimal pollen set to use, the method to identify the best analog, the criteria to identify a non-analog, and the data that is needed to reliably apply the method have not been fully resolved. Future work employing the general methodology identified in this study is needed to ensure reliable paleoclimate reconstructions.

The analysis presented here of the time-space distribution of wetland pollen and spores has shown a potential to deal with outstanding questions of paleoclimatology and biogeography. Future work will compare the migration patterns of these wetland taxa to those of arboreal taxa, to attempt to better understand the mechanisms of migration and factors affecting the speed of the postglacial colonization of land made available by the retreating ice-sheets. The potential to use lowland taxa for paleoclimate reconstruction was only suggested in this present dissertation. The next step is to combine these results with those of the global lake-level database, which can avoid some of the problems identified with the latter (Viau, 1999). These can then be used to further constrain the reconstructions derived using upland taxa.

The establishment of large databases of paleoclimate proxy data, specifically the NAPD (Grimm, 1999; Contributors to the NAPD, 1999), the availability of faster computers and new statistical techniques such as spatial statistics and resampling methods permit us to pose questions that were not possible only a few years ago. Application of the approaches presented in this thesis allows us to revisit old methods and results and improve our understanding of paleoclimatology and biogeography.

REFERENCES

Adkins, J.F., Boyle, E.A., Keigwin, L., and Cortijo, E. 1997. Variability of the North Atlantic thermohaline circulation during the last interglacial period. *Nature*, 390:154-156

Aerts, R., Verhoeven, J.T.A., Whigham, D.F., and Whigham, D.F. 1999. Plant-mediated controls on nutrient cycling in temperate fens and bogs. *Ecology*, 80(7):2170-2181

Andersen, S.T. 1973. The differential pollen productivity of trees and its significance for the interpretation of a pollen diagram from a forested region, in H.J.B. Birks and R.G. West (*eds.*), *Quaternary Plant Ecology*,109-115.

Anderson, P.M., Bartlein, P.J., Brubaker, L.B., Gajewski, K., and Ritchie, J.C. 1991. Vegetation-pollen-climate relationships for the arcto-boreal region of North America and Greenland. *Journal of Biogeography*, 18:565-582.

Anderson, P.M., Bartlein, P.J., Brubaker, L.B., Gajewski, K., and Ritchie, J.C. 1989. Modern analogs of late-Quaternary pollen spectra from the western interior of North America. *Journal of Biogeography*, 16: 573-596.

Anselin, L. 1995. Local indicators of spatial association – LISA. *Geographical Analysis*, 27(2):93-115.

Arigo, R., Howe, S.E., and Webb, T., III. 1986. Climatic calibration of pollen data: an example and annotated computing instructions, in Berglund, B.E., (eds.), *Handbook of Holocene Palaeoecology and Palaeohydrology*, 817-849.

Avizinis, J., and Webb, T., III. 1985. *The computer file of modern pollen and climatic data at Brown University*, Department of Geological Sciences, Brown University, Providence, R.I. 02912.

Bailey, T.C., and Gatrell, A.C. 1995. *Interactive spatial data analysis*. Longman Scientific and Technical, Essex, U.K.

Barrat-Segretain, M. -H. 1996. Germination and colonisation dynamics of *Nuphar lutea* (L.) Sm. in a former river channel. *Aquatic Botany*, 55(1), 31-38.

Bartlein, P.J., Prentice, I.C., and Webb, T., III. 1986. Climatic response surfaces from pollen data for some eastern North American taxa. *Journal of Biogeography*, 13:35-57.

Beal, E.O., Wooten, J.W., and Kaul, R.B. 1982. Review of the *Sagittaria engelmanniana* complex (Alismataceae) with environmental correlations. *Systematic Botany*, 7(4): 417-432.

Bedish, J.W. 1967. Cattail moisture requirements and their significance to marsh management. *American Midland Naturalist*, 78(2): 288-300.

Behl, R.J., and Kennett, J.P. 1996. Brief interstadial events in the Santa Barbara basin, NE Pacific, during the past 60 kyr. *Nature*, 379: 243-246

Bender, M., Sowers, T., Dickson, M-L., Orchardo, J., Grootes, P, Mayewski, P.A., and Meese, D.A. 1994. Climate correlations between Greenland and Antarctica during the past 100,000 years. *Nature*, 372:663-666

Bilodeau, G., de Vernal, A., and Hillaire-Marcel, C. 1990. Postglacial paleoceanography of Hudson Bay: stratigraphic, microfaunal, and palynological evidence. *Canadian Journal of Earth Science*, 27: 946-963.

Birks, H.H. 1980. Plant macrofossils in Quaternary lake sediments. *Arch. Hydriobiol. Beih.*, 15: 1-60.

Birks, H.J.B. 1995. Quantitative paleaeoenvironmental reconstructions, in Maddy, D., and Brew, J.S. (*eds.*), *Statistical Modelling of Quaternary Science Data*, Technical Guide 5, Quaternary Research Association, Cambridge,161-253.

Blackford, J. 2000. Palaeoclimatic records from peat bogs. *TREE*, 15(5): 193-198.

Boots, B.N. 1986. *Voroni (Thiessen) Polygons*. Geo Books, Norwich.

Bourgeois, J. 2000. *Modern and Holocene pollen assemblages from Arctic Ice Caps*. Ph.D. thesis, Department of Geography, University of Ottawa, Ottawa.

Bradshaw, R.H.W., and Webb, T., III. 1985. Relationships between contemporary pollen and vegetation data from Wisconsin and Michigan, USA. *Ecology*, 66(3): 721-737.

Brasier, M.D. 1994. *Microfossils*. Chapman and Hall, London.

Calcote, R. 1995. Pollen source area and pollen productivity evidence from forest hollows. *Journal of Ecology*, 83: 591-602.

Calleja, M., Rossignol-Strick, M., and Duzer, D. 1993. Atmospheric pollen content off West Africa. *Review of Palaeobotany and Palynology*, 79: 335-368.

Chambers, R.M., Meyerson, L.A., and Saltonstall, K. 1999. Expansion of *Phragmites australis* into tidal wetlands of North America. *Aquatic Botany*, 64(3-4): 261-273.

Cheddadi, R. 2000 *editor*. *European Pollen Database*. Data archived at IGBP PAGES/ World Data Center-A for Paleoclimatology, NOAA/NGDC Paleoclimatology Program. Boulder, Co, 2000. (www.ngdc.noaa.gov/paleo)

Clark, J.S., Fastie, C., Hurtt, G., Jackson, S.T., Johnson, C., King, G.A., Lewis, M., Lynch, J., Pacala, S., Prentice, C., Schupp, E.W., Webb, T., III., and Wyckoff, P. 1998. Reid's paradox of rapid plant migration: Dispersal theory and interpretation of paleoecological records. *Bioscience*, 48(1): 13-24

Clymo, R.S. 1984. The limits to peat bog growth. *Philosophical Transactions of the Royal Society of London B*, 303:605-654.

Clymo, R.S., and Hayward, P.M. 1982. The Ecology of *Sphagnum*, in *Bryophyte Ecology*, A.J.E. Smith (*ed.*). Chapman and Hall, New York, 229-289.

COHMAP members. 1988. Climate changes of the last 18,000 years: observations and model simulations. *Science*, 241: 1043-1052.

Contributors to the NAPD, 1999. *IGBP PAGES/World Data Center for Paleoclimatology*, NOAA/NGDC Paleoclimatology Program, Boulder, Colorado, USA. http://www.ngdc.noaa.gov/paleo/

Cook, C.D.K. 1985. Range extensions of aquatic vascular plant species. *Journal of Aquatic Plant Management*, 23: 1-6.

Cooper, D.W. 1968. The significance level in multiple tests made simultaneously. *Heredity*, 23: 614-617.

Cressie, N. 1995. *Statistics for spatial data.* John Wiley and Sons, New York.

Crow, G.E., and Hellquist, C.B. 1981a. Aquatic Vascular Plants of New England: Part 3. Alismataceae. *Station Bulletin*, 518:32. New Hampshire Agricultural Experiment Station, University of New Hampshire, Durham, New Hampshire.

Crow, G.E., and Hellquist, C.B. 1981b. Aquatic Vascular Plants of New England: Part 2. Typhaceae *and* Sparganiaceae. *Station Bulletin* 517:21. New Hampshire Agricultural Experiment Station, University of New Hampshire, Durham, New Hampshire.

Dansgaard, W., Johnsen, S.J., Clausen, H.B., Dahl-Jensen, D., Gudestrup, N.S., Hammer, C.U., Hvidberg, C.S., Steffensen, J.P., Svinbjornsdottir, A.E., Jouzel, J., and Bond, G. 1993. Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, 364:218-220.

Davis, M.B. 1976. Pleistocene biogeography of temperate deciduous forests. *Geoscience and Man*, 13: 13-26.

Davis, M.B. 1981. Quaternary history and stability of forest communities in Forest Succession, in Darrell C. West, Herman H. Shugart and Daniel B. Botkin (*eds.*). Springer Verlag, New York, 132-151.

Davis, M.B., Schwartz, M.W., and Woods, K. 1991. Detecting a species limit from pollen in sediments. *Journal of Biogeogrpahy*, 18:653-668.

Davis, M.B., Woods, K.D., Webb, S.L., and Futyma, R.P. 1986. Dispersal versus climate: Expansion of *Fagus* and *Tsuga* into the Upper Great Lakes region. *Vegetatio*, 67: 93-103.

Davis, R.B., and Webb, T., III. 1975. The contemporary distribution of pollen in Eastern North America: A comparison with the vegetation. *Quaternary Research*, 5: 395-434.

Davison, A. C., and Hinkley, D. V. 1998. *Bootstrapping methods and their application.* Cambridge, UK: Cambridge University Press.

de Vernal, A., Goyette, C., and Rodrigues, C.G. 1989. Contribution palynostratigraphique (dinokystes, pollen et spores) à la connaissance de la mer de Champlain: coupe de Saint-Cesaire, Québec. *Canadian Journal of Earth Science*, 26: 2450-2464.

de Vernal, A., Guiot, J., and Turon, J.L. 1993. Late and postglacial paleoenvironments of the Gulf of St. Lawrence: marine and terrestrial palynological evidence. *Géographie physique et Quarternaire*, 47 : 167-180.

de Vernal, A., Henry, A., and Bilodeau, G. 1986. Techniques de préparation et d'analyse en micropaléontologie. *Les Cahiers du GEOTOP 3*, Open File, Université du Québec a Montréal.

de Vernal, A., Hillaire-Marcel, C., and Bilodeau, G. 1996. Reduced meltwater outflow from the Laurentide ice margin during the Younger Dryas. *Nature* 381: 744-777.

de Vernal, A., Turon, J.-L., and Guiot, J. 1994. Dinoflagellate cyst distribution in high-latitude marine environments and quantitative reconstruction of sea-surface salinity, temperature, and seasonality. *Canadian Journal of Earth Science*, 31: 48-62.

Delcourt, P.A., Delcourt, H.R., and Webb, T., III. 1984. Atlas of mapped distributions of dominance and modern pollen percentages for important tree taxa of eastern North America. *American Association of stratigraphic Palynologists Foundation*. AASP Contributions Series No. 14: 131.

Denton, G.H., and Karlén, W., 1973: Holocene climatic variations - Their patterns and possible cause. *Quaternary Research*, 3:155-205.

Dieffenbacher-Krall, A.C., 1998. *Aquatic Plants in Quaternary Science*. Ph.D. Thesis (Unpublished) in Plant Science, University of Maine, December 1998.

Dieffenbacher-Krall, A.C., and Jacobson, G.L., Jr. 2001 (in press). Post-glacial changes in the geographic ranges of certain aquatic vascular plants in North America, in *Biology and Environment*: Proceedings of the Royal Irish Academy.

Dodge, J.D. 1994. Biogeography of marine armored dinoflagellates and dinocysts in the NE Atlantic and North Sea. *Review of Palaeobotany and Palynology*, 84: 169-180.

Dowsett, H.J., and Robinson, M.M. 1997. Application of the Modern Analog Technique (MAT) of sea surface temperature estimation to Middle Pliocene North Pacific planktonic foraminifer faunas. *Palaeontologia Electronica*, 1-1 [http://www-odp.tamu.edu/paleo/1998_1/dowsett/issue1.htm].

Duane, A., and Harland, R. 1990. Late Quaternary dinoflagellate cyst biostratigraphy for sediments of the Porcupine Basin, offshore Western Ireland. *Review of Palaeobotany and Palynology*, 63: 1-11.

Dugle, J.R., and Copps, T.P. 1972. Pollen characteristics of Manitoba Cattails. *Canadian Field Naturalist*, 86: 33-40.

Dyke, A., Dale, J.E., and McNeely, R.N. 1996. Marine mollusks as indicators of environmental change in glaciated North America and Greenland during the last 18000 years: *Géographie physique et Quaternaire*, 50: 125-184.

Edwards, M.E., Bigelow, N.H., Finney, B.P., and Eisner, W.R. 2000. Records of aquatic pollen and sediment properties as indicators of late-Quaternary Alaskan lake levels. *Journal of Paleolimnology*, 24: 55-68.

Engstrom, D.R., and Hansen, B.C.S. 1985. Postglacial vegetational change and soil development in southeastern Labrador as inferred from pollen and chemical stratigraphy. *Canadian Journal of Botany*, 63: 543-561.

Esselink, P., Helder, G.J.F., Aerts, B.A., and Gerdes, K. 1997. The impact of grubbing by Greylag Geese (*Anser anser*) on the vegetation dynamics of a tidal marsh. *Aquatic Botany*, 55(4): 261-279.

Faegri, K., and Iversen, J. 1975. *Textbook of pollen analysis*. 3rd edition. Blackwell, Oxford, 295pp.

Fagerlind, F. 1952. The real significance of pollen diagrams. *Botaniska Notiser*, 2: 186-224.

Fall, P.L. 1992. Spatial patterns of atmospheric pollen dispersal in the Colorado Rocky Mountains, USA. *Review of Palaeobotany and Palynology*, 74: 293-313.

FAUNMAP Working Group 1996. Spatial response of mammals to late Quaternary environmental fluctuations. *Science*, 272: 1601-1606.

Fedorova, I.T., and Volkova, Y.A. 1990. World Vegetation Cover Map. Global analog map of 1:80,000,000 scale in Russian polyconical projection. Paper manuscript on one sheet. Unpublished, in Kineman, J.J. *et al.*, (*eds*.) 2000. Global Ecosystems Database Project. 2000. Global Ecosystems Database Version II: Database, User's Guide, and Dataset Documentation. US Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, Boulder, Colorado. KGRD #35. Two CDROMs and publication on the World Wide Web.

Field, M., Huntley, M., and Muller, H. 1994. Eemian climate fluctuations observed in a European pollen record. *Nature*, 371:779.

Flora of North America Editorial Committee (eds.). 1993; 1997; 2000, *The Flora of North America North of Mexico. Vol. 2, 3, 22.* Oxford University Press.

Gajewksi, K., Viau, A., Sawada, M., Atkinson, D.E., and Wilson, S. 2001. *Sphagnum* peatland distribution in North America and Eurasia during the past 21,000 years. *Global Biogeochemical Cycles*, 15(2): 297.

Gajewski, K. 1987. Climatic impacts on the vegetation of eastern North America during the past 2000 years. *Vegetatio*, 68: 179-190.

Gajewski, K. 1991. Représentation pollinique actuelle à la limite des arbres au Nouveau-Québec. *Canadian Journal of Earth Science*, 28 : 643-648.

Gajewski, K. 1993. The role of paleoecology in the study of global climatic change. *Review of Palaeobotany and Palynology*, 79: 141-151.

Gajewski, K. 1995. Modern and Holocene pollen assemblages from some small arctic lakes on Somerset Island, NWT, Canada. *Quaternary Research*, 44: 228-236.

Gajewski, K., and Garralla, S. 1992. Holocene vegetation histories from three sites in the tundra of Northwestern Québec. *Arctic and Alpine Research*, 24: 329-336.

Gajewski, K., Payette, S., and Ritchie, J.C. 1993. Holocene vegetation history at the boreal-forest -- shrub-tundra transition in northwestern Québec. *Journal of Ecology*, 81: 433-443.

Gajewski, K., Vance, R., Sawada, M., Fung, I., Gignac, L.D., Halsey, L., John, J., Maisongrande, P., Mandell, P., Mudie, P., Richard, P.J.H., Sherrin, A., Soroko, J., and Vitt., D. 2000. The climate of

North America and adjacent ocean waters ca. 6ka. *Canadian Journal of Earth Sciences*. 37: 661-681.

Gerdol, R. 1995. The growth dynamics of *Sphagnum* based on field measurements in a temperate bog and on laboratory cultures. *Journal of Ecology*, 83: 431-437.

Getis, A., and Ord, J.K. 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3):189-206.

Gignac, L.D., and Vitt, D.H. 1990. Habitat limitations of *Sphagnum* along climatic, chemical and physical gradients in mires of western Canada. *The Bryologist*, 93: 7-22.

Gignac, L.D., and Vitt, D.H. 1994, Responses of northern peatlands to climate change: effects on bryophytes. *Journal of Hattori Botanical Laboratory*, 75: 119-132.

Gignac, L.D., Nicholson, B., and Bayley, S.E. 1998. The utilization of Bryophytes in bioclimatic modeling: present distribution of peatlands in the Mackenzie River Basin, Canada. *The Bryologist*, 101: 560-571.

Gignac, L.D., Vitt, D.H., and Bayley, S.E. 1991. Bryophyte response surfaces along ecological and climatic gradients. *Vegetatio*, 93:29-45.

Gleason, H.A., and Cronquist, A. 1963. *Manual of Vascular Plants of Northeastern United States and Adjacent Canada*. D. Van Nostrand Company, Toronto.

Golterman, H.L., Clymo, R.S., Best, E.P.H., and Lauga, J. 1988. Methods of exploration and analysis of the environment of aquatic vegetation. *Vegetation of Inland Waters*, (ed. by J.J. Symoens), Kluwer Academic Publishers, Dordrecht, Netherlands.

Goodchild, M.F. 1987. *Spatial Autocorrelation*. CATMOG 47.

Gorham, E. 1988. Canada's peatlands: their importance for the global carbon cycle and possible effects of "greenhouse" climatic warming. *Transactions of the Royal Society of Canada*, Series V. 3: 5-8.

Gorham, E. 1991. Northern peatlands: role in the carbon cycle and probable responses to climatic warming. *Ecological Applications*, 1: 182-195.

Goulden, M.L., Wofsy, S.C., Harden, J.W., Trumbore, S.E., Crill, P.M., Gower, S.T., Fries, T., Daube, B.C., Fan, S.M., Sutton, D.J., Bazzaz, A., and Munger, J.W. 1998. Sensitivity of boreal forest carbon balance to soil thaw. *Science*, 279:214-217

Grace, J.B., and Harrison, J.S. 1986. The biology of Canadian weeds. 73. *Typha latifolia* L., *Typha angustifolia* L. and *Typha x glauca* Godr. *Canadian Journal of Plant Science*, 66: 361-379.

Grace, J.B., and Wetzel, R.G. 1981. Habitat partitioning and competitive displacement in cattails (Typha): Experimental field studies. *American Naturalist*, 118: 463-474.

Grace, J.B., and Wetzel, R.G. 1998. Long-term dynamics of *Typha* populations. *Aquatic Botany*, 61: 137-146.

Griffin, K. 1977. Paleoecological aspects of the Red Lake Peatland, northern Minnesota. *Canadian Journal of Botany*, 55: 172-192.

Griffith, D.A. 1987. *Spatial Autocorrelation: A primer*. Association of American Geographers, Resource Publications in Geography.

Grimm, E. 1999. *editor*. North American Pollen Database and Global Pollen Database, Data archived at IGBP PAGES/World Data Center-A for Paleoclimatology, NOAA/NGDC Paleoclimatology Program, Boulder, Co., (www.ngdc.noaa.gov/paleo).

Grimm, E.C., and Jacobson, G.L. 1992. Fossil-pollen evidence for abrupt climate changes during the past 18000 years in eastern North America. *Climate Dynamics*, 6:179-184.

GRIP Members 1993. Climate instability during the last interglacial period recorded in the GRIP ice core. *Nature*, 364:203-207.

Groot, J. 1971. Distribution of pollen and spores in the oceans, in B.M. Funnell and W.R. Riedel (*eds.*), *The Micropalaeontology of Oceans*, 359-360. Cambridge University press.

Groot, J.J., and Groot, C.R. 1971. Horizontal and vertical distribution of pollen and spores in quaternary sequences, in B.M. Funnell and W.R. Riedel (*eds.*), *The Micropalaeontology of Oceans*, 493-504. Cambridge University press.

Grootes, P.M., Stuiver, M., White, J.W.C., Johnsen, S., and Jouzel, J. 1993. Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores. *Nature*, 366:552-554

Guiot, J. 1990. Methodology of paleoclimatic reconstruction from pollen in France. *Palaeogeography, Palaeoclimatology, and Palaeoecology*, 80: 49-69.

Halsey, L., Vitt, D.H., and Gignac, L.D. 2000. *Sphagnum*-dominated peatlands in North America since the last glacial maximum: their occurrence and extent. *The Bryologist*, 103: 334-352.

Halsey, L.A., Vitt, D.H., and Bauer, I.E. 1998. Peatland initiation during the Holocene in continental western Canada. *Climatic Change*, 40:315-342.

Hammer, U.T., and Heseltine, J.M. 1988. Aquatic macrophytes in saline lakes of the Canadian prairies. *Hydrobiologia*, 158: 101-116.

Harms. 1973. Taxonomic studies of North American *Sparganium*. I. *S. hyperboreum* and *S.minimum*. *Canadian Journal of Botany*, 51: 1629-1641.

Hartleb, C.F., Madsen, J.D., and Boylen, C.W. 1993. Environmental factors affecting seed germination in *Myriophyllum spicatum* L. *Aquatic Botany*, 45(1): 15-26

Haynes, R.R. 1974. A revision of North American *Potamogeton subsection pusilli* (*Potamogeton*). *Rhodora*, 76: 564-649.

Heinselman, M. 1970. Landscape evolution, peatland types, and the environment in the Lake Agassiz peatlands natural area, Minnesota. *Ecological Monographs*, 40: 235-261.

Heslop-Harrison, Y. 1955. *Nuphar* Sm. *Journal of Ecology*, 43(1): 342-364.

Heusser, L.E., and Morley, J.J. 1990. Climatic change at the end of the last glaciation in Japan inferred from pollen in three cores from the northwest Pacific Ocean. *Quaternary Research*, 34: 101-110.

Hickman, M., and Schweger, C.E. 1993. Late glacial – early Holocene palaeosalinity in Alberta, Canada – climate implications. *Journal of Paleolimnology*, 8: 149-161.

Hofstetter, R. 1983. Wetlands in the United States, in A.J. Gore (*eds.*), *Mires: swamp, bog, fen and moor. Regional Studies, Ecosystems of the world 4B*, 201-244. Elsevier, Amsterdam.

Hooghiemstra, H., Stalling, H., Agwu, C.O.C., and Dupont, L.M. 1992. Vegetational and climatic changes at the northern fringe of the Sahara 250,000 years BP: evidence from 4 marine pollen records located between Portugal and the Canary Islands. *Review of Palaeobotany and Palynology*, 74: 1-53.

Hotchkiss, N., and Dozier, H.L. 1949. Taxonomy and distribution of N. American Cat-tails. *Am. Mid. Nat.*, 41: 237-254.

Hultén, E. 1968. *Flora of Alaska and Neighboring Territories: a manual of the vascular plants*, 1008. Stanford University Press, Stanford California.

Huntley, B. 1996. Quaternary palaeoecology and ecology. *Quaternary Science Reviews*, 15: 591-606.

Huntley, B., and Webb, T., III. 1989. Migration: species' response to climatic variations caused by changes in the earth's orbit. *Journal of Biogeography*, 16: 5-9.

Huntley, B., Bartlein, P.J., and Prentice, I.C. 1989. Climatic control of the distribution and abundance of beech (*Fagus* L.) in Europe and North America. *Journal of Biogeography*, 16:551-560.

Husband, B.C., and Hickman, M. 1984. Growth and biomass allocation of *Ruppia occidentalis* in three lakes, differing in salinity. *Canadian Journal of Botany*, 63: 2004-2014.

Imbrie, J.J., Hays, J., Martinson, D., McIntyre, A., Mix, A., Morley, J., Pisias, N., Prell, W., and Shackleton, N. 1984. in A.L. Berger et al. (*eds.*), *Milankovich and Climate, Part 1*, 269-305.

Jackson, S.T. 1990. Pollen source area and representation in small lakes of northeastern United States. *Review of Palaeobotany and Palynology*, 63: 53-76.

Jackson, S.T. 1994. Pollen and spores in Quaternary lake sediments as sensors of vegetation composition: theoretical models and empirical evidence, in A. Traverse (*ed.*), *Sedimentation of Organic Particles*, 253-286. Cambridge University Press.

Jackson, S.T., and Dunwiddie, P.W. 1992. Pollen dispersal and representation on an offshore island. *New Phytologist*, 122: 187-202.

Jackson, S.T., and Kearsley, J.B. 1998. Quantitative representation of local forest composition in forest-floor pollen assemblages. *Journal of Ecology*, 86: 474-490.

Jackson, S.T., and Lyford, M.E. 1999. Pollen dispersal models in Quaternary plant ecology: Assumptions, parameters, and prescriptions. *The Botanical Review*, 65(1): 39-75.

Jackson, S.T., and Smith, S.J. 1994. Pollen dispersal and representation on an isolated, forested plateau. *New Phytologist*, 128: 187-193.

Jackson, S.T., and Wong, A. 1994. Using forest patchiness to determine pollen source areas of closed-canopy pollen assemblages. *Journal of Ecology*, 82: 89-99.

Jackson, S.T., Overpeck, J.T., Webb, T., III., Keattch, S.E., and Anderson, K.H. 1997. Mapped plant-macrofossil and pollen records of late Quaternary vegetation change in eastern North America. *Quaternary Science Reviews*, 16: 1-70.

Jackson, S.T., Webb, T., III., Prentice, I.C., and Hansen, J.E. 1995. Exploration and calibration of pollen/vegetation relationships: a PC program for the extended R-value models. *Review of Palaeobotany and Palynology*, 84: 65-374.

Jacobson, G.L., and Grimm, E.C. 1986. A numerical analysis of Holocene forest and prairie vegetation in central minnesota. *Ecology*, 67(4):958-966.

Jacobson, G.L., Jr., and Bradshaw, R.H.W. 1981. The selection of sites for paleoecological studies. *Quaternary Research*, 16:80-96.

Jacobson, G.L., Jr., Webb, T., III, and Grimm, E.C. 1987. Patterns and rates of vegetation change during the deglaciation of eastern North America, in Ruddiman , W.F., and Wright, H.E., Jr., (*eds.*), North America and adjacent oceans during the last deglaciaion: Boulder, Colorado, Geological Society of America, *The Geology of North America*, v. K-3.

Janssen, C.R. 1973. Local and regional pollen deposition, in Birks, H.J.B., and West, R.G. (eds.), *Quaternary Plant Ecology*, 31-42. Blackwell, Oxford.

Janssens, J., Hansen, B., Glaser, P., and Whitlock, C. 1992. Development of a raised-bog complex, in H.E. Wright, Jr., B. Coffin, and N. Aaseng (*eds.*), *The patterned peatlands of Minnesota*. University of Minnesota Press, Minneapolis.

Jettè, H., and Richard, P.J.H. 1992. Contribution à l'histoire postglaciaire de la végétation en Gaspésie méridionale, Québec. *Géographie physique et Quaternaire*, 46 : 273-284.

Johnson, D., Kershaw, L., MacKinnon, A., and Pojar, J. 1995. *Plants of the Western Boreal Forest and Aspen Parkland*, p. 392. Lone Pine Publ. and Canadian Forest Service, Edmonton Alberta Canada.

Johnston, R.J. 1980. *Multivariate Statistical Analysis in Geography*, pp. 127-182. Longman, London, New York.

Jolly, D., Harrison, S.P., Damnati, B., and Bonnefille, R. 1998. Simulated Climate and Biomes of Africa During the Later Quaternary: Comparison with Pollen and Lake Status Data. *Quaternary Science Reviews*, 17: 629-657.

Jonasson, S., and Shaver, G.R. 1999. Within-stand nutrient cycling in arctic and boreal wetlands. *Ecology*, 80(7): 2139-2150.

Kapp, R. 1969. *Pollen and Spores*. Wm. C. Brown Co. U.S.A. 249.

Kermack, K.A., and Haldane, J.B.S. 1950. Organic correlation and allometry. *Biometrika*, 37: 30-41.

Korvena, E.V. 1971. Spores and pollen in Mediterranean bottom sediments, in Funnell, B.M., and Riedel, W.R. (*eds.*), *The Micropalaeontology of Oceans*, 361-371. Cambridge, Cambridge University press.

Kuc, M. 1973. Bryogeography of Expedition area, Axel Heiberg Island, N.W.T., Canada. *Bryophytorum Bibliotheca, Band 2.*

Kuhry, P., and Vitt, D. 1996. Fossil carbon/nitrogen ratios as a measure of peat decomposition. *Ecology*, 77: 271-275.

Kutzbach, J.E., and Ruddiman, W.F. 1993. Model description, external forcing, and surface boundary conditions, in Wright, H.E., Kutzbach, J.E., Webb, T., III., Ruddiman, W.F., Street-Perrott, F.A., and Bartlein, P.J. (*eds.*), *Global Climates Since the Last Glacial Maximum*,12-23. University of Minnesota Press, Minneapolis.

Kutzbach, J.E., and Webb, T., III. 1993. Conceptual basis for understanding late-Quaternary climates, in Wright, H.E., Kutzbach, J.E., Webb, T., III., Ruddiman, W.F., Street-Perrott, F.A., and Bartlein, P.J. (eds.), *Global Climates Since the Last Glacial Maximum*, 5-11. University of Minnesota Press, Minneapolis.

Labelle, C., and Richard, P.J.H. 1981. Végétation tardiglaciaire et postglaciaire au sud-est du Parc des Laurentides, Québec. *Géographie physique et Quaternaire*, 35 : 345-359.

Laird, K.R., Fritz, S.C., Grimm, E.C., and Mueller, P.G. 1996. Century-scale paleoclimatic reconstruction from Moon Lake, a closed-basin lake in the northern Great Plains. *Limnology and. Oceanography*, 41(5): 890-902.

Lam, N, 1983. Spatial interpolation methods: a review. *American Cartographer*, 10: 129,149.

Lamb, H.F. 1980. Late Quaternary vegetational history of southeastern Labrador. *Arctic and Alpine Research*, 12: 117-135.

Lamb, H.F. 1985. Palynological evidence for postglacial change in the position of tree limit in Labrador. *Ecological Monographs*, 55: 241-258.

Langanger, M., Jokl, S., and Musso, M. 2000. UV-reflectance in flowers of *Nymphaea alba* L. and *Nuphar lutea* (L.) Sm. (Nymphaeaceae). *Aquatic Botany*, 67: 13-21.

Leemans, R., and Cramer, W. 1991. *The IIASA database for mean monthly values of temperature, precipitation and cloudiness of a global terrestrial grid.* International Institute for Applied Systems Analysis (IIASA), RR-91-18, (ftp.pik-potsdam.de).

Legendre, P., and Fortin, M.J. 1989. Spatial pattern and ecological analysis. *Vegetatio*, 80:107-138.

Levac, E., and de Vernal, A. 1997. Postglacial changes of terrestrial and marine environments along the Labrador coast: Palynological evidence from cores 91-045-005 and 91-045-006, Cartwright Saddle. *Canadian Journal of Earth Science*, 34(10): 1358-1365.

Lewis, W.H., Vinay, P., and Zenger, V.E. 1983. *Airborne and Allergenic Pollen of North America.* The Johns Hopkins University Press, Baltimore, Maryland, USA, 254 pp.

Lézine, A.M. 1991. West African paleoclimates during the last Climatic Cycle Inferred from an Atlantic deep-sea pollen record. *Quaternary Research,* 35: 456-463.

Lézine, A.M., and Denèfle, M. 1997. Enhanced anticyclonic circulation in the eastern North Atlantic during cold intervals of the last deglaciation inferred from deep-sea pollen records. *Geology,* 25: 119-122.

Lynch, E.A. 1996. The ability of pollen from small lakes and ponds to sense fine-scale vegetation patterns in the central rocky Mountains, USA. *Review of Palaeobotany and Palynology,* 94: 197-210.

MacRae, R.A., Fensome, R.A., and Williams, G.L. 1996. Fossil dinoflagellate diversity, originations, and extinctions and their significance. *Canadian Journal of Botany,* 74: 1687-1694.

Marcoux, N., and Richard, P.J.H. 1995. Végétation et fluctuations climatiques postglaciaires sur la côte septentrionale gaspésienne, Québec. *Canadian Journal of Earth Science,* 32: 79-96.

Marret, F. 1994. Distribution of dinoflagellate cysts in recent marine sediments from the east Equatorial Atlantic (Gulf of Guinea). *Review of Palaeobotany and Palaynology,* 84: 1-22.

Martinson, D.G., Pisias, N.G., Hays, J.D., Imbrie, J., Moore, T.C., and Shackleton, N.J. 1987. Age dating and the orbital theory of ice ages: Development of a high-resolution 0 to 300,000-year chronostratigraphy. *Quaternary Research,* 27:1-29

Matsuoka, K. 1994. Holocene dinoflagellate cyst assemblages in shallow water sediments of the Tsushima Islands, west Japan. *Review of Palaeobotany and Palynology,* 84: 155-168.

McCaughey, T.L., and Stephenson, G.R. 2000. Time from flowering to seed viability in purple loosestrife (*Lythrum salicaria*). *Aquatic Botany,* 66: 57–68.

McGlone, M. 1996. When history matters: scale, time, climate and tree diversity. Global Ecology and Biogeography Letters, 5:309-314.

Miller, R.G. 1977. Developments in multiple comparisons. *Journal of the American Statistical Association,* 72(360): 779-788.

Mitchell, S.F., and Wass, R.T. 1996. Grazing by black swans (*Cygnus atratus* Latham), physical factors, and the growth and loss of aquatic vegetation in a shallow lake. *Aquatic Botany,* 55(3): 205-215.

Monroe, J.S., and Wicander, R. 1998. *Physical Geology: Exploring the earth,* 3rd edition. Wadsworth Publishing Company, Toronto.

Moore, P., and Bellamy, D. 1974. *Peatlands.* Springer-Verlag, New York.

Moran, P.A.P. 1950. Notes on continuous stochasic phenomena. *Biometrika,* 37:17-23.

Morton, J.F. 1975. Cattails (*Typha* spp.) – weed problem or potential crop? *Economic Botany,* 29: 7-29.

Muenscher, W.C. 1944. *Aquatic Plants of the United States,* 374. Comstock Publishing Company, Inc. Cornell University, Ithica New York.

Nilsson, S., and Praglowski, J. (*eds*). 1992. *Erdtman's Handbook of Palynology 2nd ed.* Munksgaard, Copenhagen, 580 p.

Ning, S., and Dupont, L.M. 1997. Vegetation and climatic history of Southwest Africa: a marine palynological record of the last 300,000 years. *Vegetation History and Archaeobotany*, 6: 117-131.

NRCAN, 2001. *Forest Regions of Canada*, adapted from J.S. Rowe, Forest Regions of Canada (Ottawa: Canadian Forestry Service, Dept. of Fisheries and Environment, 1977). Pacific Forestry Center, Canadian Forest Service. http://www.pfc.cfs.nrcan.gc.ca.

Oden, N.L. 1984. Assessing the significance of a Spatial Correlogram. *Geogrpahical Analysis*. 16(1):1-16.

Oechel, W.C., Hastings, S.J., Vourlitis, G., Jenkins, M., Riechers, G., and Grulke, N. 1993. Recent change of Arctic tundra ecosystems from a net carbon dioxide sink to a source. *Nature*, 361:520-523.

Ord, J.K., and Getis, A. 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27(4):287-305.

Overpeck, J.T., Webb, R.S., and Webb, T., III. 1992. Mapping eastern North America vegetation change over the past 18,000 years: no-analogs and the future. *Geology*, 20:1071-1074.

Overpeck, J.T., Webb, T., III., and Prentice, I.C. 1985. Quantitative interpretation of fossil pollen spectra: Dissimilarity Coefficients and the method of Modern Analogs. *Quaternary Research*, 23-1: 87-108.

Parsons, R.W., and Prentice, I.C. 1981. Statistical approaches to R-values and the pollen-vegetation relationship. *Review of Palaeobotany and Palynology*, 32: 127-152.

Parsons, R.W., Prentice, I.C., and Saarnisto, M. 1980. Statistical studies on pollen representation in Finnish lake sediments in relation to forest inventory data. *Ann. Bot. Fennici*, 17: 379-393.

Peltier, W. 1993. *Time dependent topography through glacial cycle*, IGBP PAGES/ World Data Center-A for Paleoclimatology Data Contribution Series #93-015, NOAA/NGDC Paleoclimatology Program, Boulder, Co. ftp.ngdc.noaa.gov/paleo/pmip/peltier_ice/

Peltier, W. 1994. Ice Age paleotopography. *Science*, 265: 195-201.

Penfound, T.Wm. 1952. An outline for ecological life histories of vascular hydrophytes. *Ecology*, 33-1: 123-128.

Penfound, T.Wm., and Earle, T.T. 1948. The biology of the Water Hyacinth. *Ecological Monographs*, 18-4: 447-472.

Peng, C., Guiot, J., and van Campo, E. 1998. Estimating changes in terrestrial vegetation and carbon storage: using palaeoecological data and models. *Quaternary Science Reviews*, 17: 719-736.

Peyron, O., Jolly, D., Bonnefille, R., Vincens, A., and Guiot, J., 2000. The climate of East Africa from pollen data, 6000 years ago. *Quaternary Research*, 54-1: 90-101.

Pip, E. 1979. Survey of the ecology of submerged aquatic macrophytes in central Canada. *Aquatic Botany*, 7: 339-357.

Pitelka, L.F., and the Plant Migration Workshop Group 1997. Plant migration and climate change. *American Scientist*, 85:464-474

Porsild, A.E., and Cody, W.J. 1980. *Vascular Plants of Continental Northwest Territories, Canada*, 667. National Museum of Natural Sciences, Canada.

Prentice, I.C. 1980. Multidimensional scaling as a research tool in Quaternary palynology: a review of theory and methods. *Review of Paleobotany and Palynology*, 31: 71-104.

Prentice, I.C. 1985. Pollen representation, source area, and basin size: toward a unified theory of pollen analysis. *Quaternary Research*, 23: 76-86.

Prentice, I.C. 1986. Forest-composition calibration of pollen data, in B.E. Berglund (ed.), *Handbook of Holocene Palaeoecology and Palaeohydrology*, pp. 799-816. John Wiley and Sons Ltd.

Prentice, I.C. 1988. *Records of vegetation in time and space: the principles of pollen analysis*, in B. Huntley and T. Webb III., (*eds.*) *Vegetation History, 17-42*. Kluwer Academic Publishers.

Prentice, I.C., and R.W. Parsons. 1983. Maximum likelihood linear calibration of pollen spectra in terms of forest composition. *Biometrics*, 39: 1051-1057.

Prentice, I.C., and Webb, T., III. 1986. Pollen percentages, tree abundance and the Fagerlind effect. *Journal of Quaternary Science*, 1-1: 35-43.

Prentice, I.C., Bartlein, P.J., and Webb, T., III. 1991. Vegetation and climate change in eastern North America since the last glacial maximum. *Ecology*, 72-6: 2038-2056.

Prentice, I.C., Berglund, B.E., and Olsson, T. 1987. Quantitative forest-composition sensing characteristics of pollen samples from Swedish lakes. *Boreas*, 16: 42-54.

Prentice, I.C., Harrison, S.P., Jolly, D., and Guiot, J. 1998. The Climate and Biomes of Europe AT 6000 yr BP: Comparison of Model Simulations and Pollen-based Reconstructions. *Quaternary Science Reviews*, 17: 657-668.

Punt, W. 1976. Sparganiaceae and Typhaceae, in W. Punt (*ed.*), *The Northwest European Pollen Flora, I.* Parts 1-7, pp. 75-88. Elsevier, Amsterdam.

Richard, P.J.H. 1977. Histoire post-wisconsinienne de la végétation du Québec méridional par l'analyse pollinique. *Service de la Recherche, Direction Générale des Forêts, Ministère des Terres et Forêts, Québec*, 1: 312 et 2: 142.

Richard, P.J.H. 1996. Le couvert végétal du Québec-Labrador il y a 6000 ans BP: Essai. *Géographie physique et Quaternaire*, 49 : 117-140.

Richard, P.J.H., and Poulin, P.H. 1976. Un diagramme pollinique au mont des Eboulements, région de Charlevoix, Québec. *Canadian Journal of Earth Science*, 13: 145-156.

Ridley, H.N. 1930. *The Dispersal of Plants Throughout the World.* L. Reeve and Company, Lloyds Bank Building, Ashford Kent, UK. 744 p.

Riggs, D.S., Guarnieri, J.A., and Addelman, S. 1978. Fitting straight lines when both variables are subject to error. *Life Sciences,* 22: 1305-1360.

Ritchie, J.C. 1986. Vegetation-climate equilibrium in relation to palaeoecology. *Vegetatio,* 67: 65-74.

Saucier, J.P., Bergeron, J.F., Grodin. P., and Robitaille, A. 1998. *Les régions écologiques du Québec méridional (3e version): un des éléments du système hiérarchique de classification écologique du territoire* mis au point par le ministère des Ressources naturelle du Québec. Supplément - L'Aubelle, février-mars 1998.

Sawada, M. 1999. ROOKCASE: An Excel 97/2000 Visual Basic (VB) add-in for exploring global and local spatial autocorrelation. *Bulletin of the Ecological Society of America,* 80-4:231-234.

Sawada, M., Gajewski, K., de Vernal, A., and Richard, P. 1999. Comparison of marine and terrestrial Holocene climatic reconstructions from northeastern North America. *The Holocene,* 9-3:267-277.

Sawada, M., Soroko, J., and Gajewski, K. (1997). Assessing the reliability of Paleoclimate reconstructions from the method of modern analogues. Annual meeting of the American Geophysical Union, San Francisco California, December 1997.

Sculthorpe, C.D. 1967. *The Biology of Aquatic Vascular Plants,* 610. Edward Arnold Ltd., London U.K.

Smith, S.G. 1967. Experimental and natural hybrids in North American *Typha* (Typhaceae). *American Midland Naturalist,* 78(2): 257-287.

Snyder, J. 1987. *Map projections – a working manual,* U.S. Geological Survey Professional Paper 1395, U.S. Government Printing Office, Washington D.C.

Sokal, R.R., and Oden, N.L. 1978. Spatial autocorrelation in biology 1. Methodology. *Biological Journal of the Linnean Society,* 10:199.

Sokal, R.R., Oden, N.L., and Thomson, B.A. 1998a. Local spatial autocorrelation in biological variables. *Biological Journal of the Linnean Society,* 65:41-62.

Sokal, R.R., Oden, N.L., and Thomson, B.A. 1998b. Local spatial autocorrelation in a biological model. *Geographical Analysis,* 30(4):331-354.

Søndergaard, M., Bruun, L., Lauridsen, T., Jeppesen, E., and Madsen, T.V. 1996. The impact of grazing waterfowl on submerged macrophytes: In situ experiments in a shallow eutrophic lake. *Aquatic Botany,* 53(1-2): 73-84.

S-PLUS 5 for UNIX Guide to Statistics. Data Analysis Products Division, Mathsoft, Seattle.

Stuckey, R.L. 1993. Phytogeographical outline of aquatic and wetland angiosperms in continental eastern North America. *Aquatic Botany,* 44: 259-301.

Stuiver, M., and Reimer, P. 1993. Extended 14C database and revised CALIB radiocarbon calibration program. *Radiocarbon,* 35: 215-230.

Stuiver, M., and Reimer, P. 2000. *Quaternary Isotope Laboratory HTML CALIB 4.2 manual*, http://depts.washington.edu/qil/calib/

Sugita, S. 1993. A model of pollen source area for an entire lake surface. *Quaternary Research*. 39: 239-244.

Sugita, S. 1994. Pollen representation of vegetation in Quaternary sediments: theory and method in patchy vegetation. *Journal of Ecology*. 82: 881-897.

Tauber, H. 1965. Differential pollen dispersion and the interpretation of pollen diagrams. *Dan. Geol. Unders., Raekke*. 2: 89.

Taylor, C.W., Luebke, N.T., Britton, D.M., Hickey, R.J., and Brunton, D.F. 1993. *Isoetaceae Reichenbach: Quillwort family in The Flora of North America North of Mexico. Vol. II*, 64-75. Flora of North America Editorial Committee (*ed.*), New York: Oxford University Press.

Taylor, K.C., Mayewski, P.A., Alley, R.B., Brook, E.J., Gow, A.J., Grootes, P.M., Meese, D.A., Saltzman, E.S., Severinghaus, J.P., Twickler, M.S., White, J.W.C., Whitlow, S., and Zielinski, G.A. 1997. The Holocene-Younger Dryas transition recorded at Summit Greenland. *Science*, 278:825-827

Thompson, R.S., Anderson, K.H., and Bartelin, P.J. 1999. Quantitative paleoclimatic reconstructions from late pleistocene plant macrofossils of the yucca mountain region. USGS Open-File Report 99-338.

Tiefelsdorf, M., and Boots, B. 1997. A note on the extremities of Local Moran's I_is and their impact on Global Moran's *I*. *Geographical Analysis*, 29(3):248-257.

Trebitz, A.S., Nichols, S.A., Carpenter, S.R., and Lathrop, R.C. 1993. Patterns of vegetation change in Lake Wingra following a *Myriophyllum spicatum* decline. *Aquatic Botany*, 46(3-4): 325-340.

Turner, A.M., and McCarty, J.P. 1998. Resource availability, breeding site selection, and reproductive success of red-winged blackbirds. *Oecologia*. 113: 140-146.

Turner, N.J. 1981. A gift for the taking: the untapped potential of some food plants of North American Native peoples. *Canadian Journal of Botany*. 59: 2331-2357.

Upton, G.J.G., and Fingleton, B. 1985. *Spatial data analysis by example: Volume I Point Pattern and Quantitative Data*, pp. 151-213. John Wiley & Sons, Toronto.

van Breemen, N. 1995. How *Sphagnum* bogs down other plants. *Trends in Ecology and Evolution*, 10: 270-275.

van der Putten, W.H. 1997. Die-back of *Phragmites australis* in European Wetlands: an overview of the European Research Programme on Reed Die-back and Progression (1993 -1994). *Aquatic Botany*. 59(3-4): 263-275.

Vance, R.E., and Mathewes, R.W. 1994. Deposition of modern pollen and plant macroremains in a hypersaline prairie lake basin. *Canadian Journal of Botany*. 72: 539-548.

Vardy, S., Warner, B., and Aravena, R. 1998. Holocene climate and the development of a subarctic peatland near Inuvik, Northwest Territories, Canada. *Climate Change*. 40: 285-313.

Versteegh, G.J.M., and Zonneveld, D.A.F. 1994. Determination of (palaeo-)ecological preferences of dinoflagellates by applying Detrended and Canonical Correspondence analysis to Late Pliocene dionoflagellate cyst assemblages of the south Italian Singa section. *Review of Palaeobotany and Palaynology*, 84: 181-199.

Vesper, S.J., and Stuckey, R.L. 1977. The return of aquatic vascular plants into the Great Lakes region after the Late-Wisconsin Glaciation. in R.C. Romans, *Geobotany*. Premium Press, New York.

Viau, A. 1999. *Lake Level Variations and Global Hydrological Change: A Spatio-temporal Analysis*. M.A. Thesis, University of Ottawa, 210 pp.

Vitt, D.H. 1990. Growth and production dynamics of boreal mosses over climatic, chemical and topographic gradients. *Botanical Journal of the Linnean Society*, 104:35-39.

Vitt, D.H. 1994. An overview of factors that influence the development of Canadian peatlands. *Memoirs of the Entomological Society of Canada*, 169: 7-20.

Vitt, D.H., Halsey, P., and Zoltai, S.C. 1994. The bog landforms of continental western Canada in relation to climate and permafrost pattern. *Arctic and Alpine Research*, 26:1-13.

Wall, D. 1971. The lateral and vertical distribution of dinoflagellates in Quaternary sediments, in B.M. Funnell and W.R. Riedel (*eds.*), *The Micropalaeontology of Oceans*, pp. 399-405. Cambridge University press, Cambridge.

Webb, T., III. 1974. Corresponding patterns of pollen and vegetation in lower Michigan: a comparison of quantitative data. *Ecology*, 55-1: 17-28.

Webb, T., III. 1986. Is vegetation in equilibrium with climate? How to interpret late-Quarternary pollen data. *Vegetatio*, 67: 75-91.

Webb, T., III. 1987. The appearance and disappearance of major vegetational assemblages: Long-term vegetational dynamics in eastern North America. *Vegetatio*, 69: 177-187.

Webb, T., III. 1993. Constructing the past from late-Quaternary pollen data: Temporal resolution and a zoom lens space-time perspective, in S.M. Kidwell and A.K. Behrensmeyer (eds.), *Taphonomic Approaches to Time Resolution in Fossil Assemblages*, Paleontological Society Short Courses in Paleontology, 6: 81-101.

Webb, T., III. 1997. The appearance and disappearance of major vegetational assemblages: Long-term vegetational dynamics in eastern North America. *Vegetatio*, 69:177-187.

Webb, T., III., Anderson, K.H., Bartlein, P.J., and Webb, R.S. 1998. Late Quaternary climate Change in Eastern North America: A Comparison of Pollen-derived Estimates with Climate Model Results. *Quaternary Science Reviews*, 17: 587-606.

Webb, T., III., Bartlein, P.J., Harrison, S.P., and Anderson, K.H. 1993. Vegetation, lake-levels, and climate in Eastern North America for the past 18,000 years in H.E.Wright, J.E., Kutzbach, T.Webb, III.,

W.F.Ruddiman, F.A.Street-Perrott, and P.J.Bartlein (*eds.*), *Global Climates Since the Last Glacial Maximum*, 415-467. University of Minnesota Press, Minneapolis.

Webb, T., III., Howe, S.E., Bradshaw, R.H.W., and Heide, K.M. 1981. Estimating plant abundance from pollen percentages: the use of regression analysis. *Review of Palaeobotany and Palynology*, 34: 269-300.

Webb, T., III., Laseski, R.A., and Bernabo, J.C. 1978. Sensing vegetational patterns with pollen data: choosing the data. *Ecology*, 59-6: 1151-1163.

Webb, T., III., Richard, P.J.H., and Mott, R.J. 1983. A mapped history of Holocene vegetation in Southern Quebec. National Museum of Natural Sciences, National Museums of Canada, *Syllogeus Series* 49: 273-336.

Webb, T., III., Yeracaris, G.Y., and Richard, P. 1978. Mapped patterns in sediment samples of modern pollen from southeastern Canada and Northeastern United States. *Géographie physique et quaternaire* 45-2: 163-176.

Weider, R., McCormick, A., and Lang, G. 1981. Vegetational analysis of Big Run Bog, a nonglaciated *Sphagnum* bog in West Virginia. *Castanea*, 46:16-29.

Weihe, P.E., and Neely, R.K. 1997. The effects of shading on competition between purple loosestrife and broad-leaved cattail. *Aquatic Botany*, 59(1-2): 127-138.

Wiersema, J.H. 1997. Cabombaceae A. Richard: Water-shield Family, in Flora of North America Editorial Committee (*ed.*), *The Flora of North America North of Mexico*, Vol. III, pp. 78-80. New York, Oxford University Press.

Wiersema, J.H., and Hellquist, C.B. 1997. Nymphaeaceae Salisbury: Water-lily family, in Flora of North America Editorial Committee (*ed.*), *The Flora of North America North of Mexico*, Vol. III, pp. 66-77. New York, Oxford University Press.

Williams, D.B. 1971a. The distribution of marine dinoflagellates in relation to the physical and chemical conditions, in B.M. Funnell and W.R. Riedel (*eds.*), *The Micropalaeontology of Oceans*, pp. 91-95. Cambridge University Press, Cambridge.

Williams, D.B. 1971b. The occurrence of dinoflagellates in marine sediments, in B.M. Funnell and W.R. Riedel (*eds.*), *The Micropalaeontology of Oceans*, pp. 231-243. Cambridge University Press, Cambridge.

Williams, J.W., Webb, T., III., Richard, P.H., and Newby, P. 2000. Late quaternary biomes of Canada and the eastern United States. *Journal of Biogeography*, 27: 585-607.

Williams, K.M., Andews, J.T., Weiner, N.J., and Mudie, P.J. 1995. Late Quaternary Paleoceanography of the Mid- to Outer Continental Shelf, East Greenland. *Arctic and Alpine Research*, 27: 352-363.

Winkler, M. 1988. Effect of climate on development of two Sphagnum bogs in south-central Wisconsin. *Ecology*, 69: 1032-1043.

Wright, H.E., Jr., Kutzbach, J.E., Webb, T., III., Ruddiman, W.F., Street-Perrott, F.A., and Bartlein, P.J., (*editors*). 1993. *Global Climates since the Last Glacial Maximum.* Minneapolis, University of Minnesota Press.

Zoltai, A., and Pollett, F. 1983. Wetlands in Canada: their classification, distribution, and use, in A.J. Gore (*ed.*), *Mires: swamp, bog, fen and moor, Regional Studies, Ecosystems of the world 4B*, pp. 245-268. Elsevier, Amsterdam.

Zoltai, S. 1995. Regional variations in peatland ecosystems of west-central Canada through time. *Gunneria*, 70: 35-42.

Zoltai, S., and Vitt, D. 1990. Holocene climate change and the distribution of peatlands in western interior Canada. *Quaternary Research*, 33, 231-240.

Appendix A

ASSEMBLY OF THE NEW MODERN POLLEN DATASET

FINAL MODERN DATASET SUMMARY

The final modern dataset WWKGVS2000 has 4590 modern pollen sites. How these were derived is described below. The datasources for the modern data were Avivinas and Webb (1985) which was reworked with new sites were added by Dr. Jack Williams (National Center for Ecological Analysis and Synthesis at the University of California, Santa Barbara) and Prof. T. Webb, III (Brown Univeristy, Department of Geological Sciences) WW2000 (WW2000). To this was added the Canadian Pollen Database (Prof. K. Gajewski, Department of Geography, University of Ottawa), here called KG2000 and sites from the Global Pollen Database (Grimm, 1999; Contributors to the GPD, 1999), called the GPD in this explanation.

A-1 SITE INFORMATION FIELDS IN WWKGVS2000

These are fields in the modern dataset that contain site information for each record.

Variable	DataType	Description
ID	[numeric]	KEY
ENUM	[numeric]	GPD enumber for GPD sites, -9999 for WW2000 and KG2000 sites
SITENUM	[numeric]	Original Site number from GPD (!= Enumber), Key Field from WW2000 and KG2000
SITECODE	[string]	Same as SITENUM but with first character specifying which dataset a record came from (W = WW2000, K = KG2000, G = GPD)
SITENAME	[string]	Site names from original datasets. Site name from WW2000 is [Name1]+[Name2] fields
LON2D	[numeric]	Longitude in Decimal Degrees to 2 decimal places
LAT2D	[numeric]	Latitude in Decimal Degrees to 2 decimal places
LONG	[numeric]	Longitude in full precision from original databases
LAT	[numeric]	Laititude in full precisions from orignal databases
ELEV	[numeric]	Elevation from original databases – missing elevations 200 of them were determined using geotopo 30 arc second elevation from the USGS (for sites where elevation was added see note at end of this document).
GTOPELV	[numeric]	Elevation extracted from GTOPO30 for all sites (see end of paper for details on extraction).
POLLSUM	[numeric]	pollen sum as calculated from [Abies]+...+[Ulmus]

A-1.1 CLIMATE DATA FIELDS

The following are also included as climate variables for each site and were extracted from P.J. Bartlein's modern gridded climatology for each modern pollen site for North America by Dr. Jack Williams.

Variable	Description
JanT	January Temperature
JulT	July Temperature
AnnPrecip	Total Annual Precipitation
GDD0	Growing Degree Days Greater than 0 °C
GDD5	Growing Degree Days Greater than 5 °C
Tcld	Temperature of Coldest Day
Twm	Temperature of Warmest Day
MIPT	?

Data extracted from Leeman's and Cramer (1991) for each site in the database by M.Sawada.

Variable	Description**
PJANUARY	Precipitation January (mm)
PFEBRUARY	
PMARCH	
PAPRIL	
PMAY	
PJUNE	
PJULY	
PAUGUST	
PSEPTEMBER	
POCTOBER	
PNOVEMBER	
PDECEMBER	
JANUARY	Mean January Temperature (°C)
FEBRUARY	
MARCH	
APRIL	
MAY	
JUNE	
JULY	
AUGUST	
SEPTEMBER	
OCTOBER	
NOVEMBER	
DECEMBER	

A-2 BETWEEN DATASET DUPLICATION REMOVAL - GEOGRAPHIC FILTER

Because of differences in the precision of which geographic coordinates were entered or maintained in the three databases we chose to use decimal degrees rounded to two decimal places in order to determine which

** *P denotes precipitation fields, e.g.. PJanuary is January precipitation. The rest are mean monthly temperatures.*

sites in the three databases were located at the same coordinates. Rounding to decimal places requires we accept that identical sites are defined by those as little as ~111 m and as much as 1113 m apart. These errors consider strictly the direction in y as the errors in x depend on the latitude and will be smaller than these numbers the further north samples are located, e.g., at 45°N the same difference would produce errors of ~787 m maximum.

The following rules were applied to compare the samples.

1. Prefer WW2000 samples over GPD samples

In the GPD 1879 samples are duplicated by at least one site in WW2000;
these 1879 samples were excluded from the GPD in favor of Those in WW2000.

2. Prefer KG2000 over GPD and WW2000.
a. Remove GPD samples that replicate KG samples. (126 samples)
b. Remove WW2000 samples that replicate KG samples. (8 samples)

total: 301 KG2000
1017 GPD
3311 WW2000

4629 SITES with no geographic duplicates based on 0.01 distance threashold (~1400 m).

NOTE: To ensure that these sites are duplicates we could have looked at the elevation fields of two geographic duplicate sites and determined the difference in elevation. If the elevation difference was zero then these sites are likely the same. Maybe 10% of the samples had discrepancies in elevation with the same geographic coordinates. Considering the resolution of our climate data we did not believe that further analysis of these small discrepancies on a few of the sites would be worth while at this time.

A-2.1 DATASET VARIABLES

The combined dataset is composed of the following variables:

Variable	Variable	Variable	Variable
ID	Ceanothus	Magnoliaceae	Shepherdia
ENUM	Celtis	Malvaceae	Sphagnum
SITENUM	Cephalanthus	Myrica	Taxodium
SITECODE	Cercocarpus	Nyssa	Thalictrum
SITENAME	Chenopodiaceae.Amaranthaceae	Onagraceae	Tilia
LON2D	Chrysolepis	Ostrya.Carpinus	Tsuga
LAT2D	Clethra	Oxyria	PJANUARY
LONG	Corylus	Papaveraceae	PFEBRUARY
LAT	Cupressaceae	Picea.undiff	PMARCH
ELEV	Cyperaceae	Picea.glauca	PAPRIL
POLLSUM	Dryas	Picea.mariana	PMAY
Abies	Ephedra	Pinus.undif	PJUNE
Acer	Ericaceae.undif	Pinus.strobus	PJULY
Alnus.undiff	Ericaceae.Chamadaphne.Ledum	Plantaginaceae	PAUGUST
Alnus.crispa	Ericaceae.Empetrum	Platanus	PSEPTEMBER

Alnus.rugosa	Ericaceae.Vaccinium	Poaceae	POCTOBER
Ambrosia	Eriogonum	Polygonum	PNOVEMBER
Amorpha	Euphorbiaceae	Polypodiaceae	PDECEMBER
Apiaceae	Fabaceae	Populus	JANUARY
Aquifoliaceae	Fagus	Potentilla	FEBRUARY
Arceuthobium	Fraxinus	Prosopis	MARCH
Asteraceae	Juglans	Pteridium	APRIL
Artemisia	Larix.Pseudotsuga	Quercus	MAY
Betula.undif	Larrea	Ranunculaceae	JUNE
Brassicaceae	Liliaceae	Rhamnaceae	JULY
Cactaceae	Liquidambar	Rubiaceae	AUGUST
Campanulaceae	Lycopodium.undiff	Rumex	SEPTEMBER
Caprifoliaceae	Lycopodium.annotinum	Salix	OCTOBER
Carya	Lycopodium.clavatum	Sarcobatus	NOVEMBER
Caryophyllaceae	Lycopodium.complanatum	Saxifragaceae	DECEMBER
Castanea	Lycopodium.selago	Selaginella	

A-3 Within Dataset Geographic Duplication - Identical Long/Lats

There are 1218 sites that have long/lats that duplicate one or more of the 4629 sites. This would leave 3411 sites if removed. However, we retain all duplicates through the next filter. That is, only those sites that are geographic duplicates and not spectral duplicates will be retained as acceptable duplicates.

For example, suppose we have two sites A and B. The longitudes and latitudes of these two sites are exactly the same. However, the spectra of A and B are different. Thus we would retain both A and B in the final dataset. On the other hand, if we had two sites C and D with identical longitudes and latitudes, and sites C and D have zero squared chord distance between them then this implies that they are at the same location AND have the same relative counts and so their spectra are identical. Thus, we would only keep one of these in the final modern dataset, which one we keep does not matter.

We will retain duplicates for the final analog reconstruction because each record in a set of duplicated sites is an random sample of the pollen at that particular site. Because multiple samples come from a particular lake, that set of samples should be strongly correlated. If a fossil spectrum is similar to one of these sites it should find its next best analog at a nearest neighbor geographic duplicate. However, by the same argument, we could justify taking the average of the duplicate spectra since this should approach the expected value of the true proportions of each individual taxon if we have more duplicates – getting to the true pollen spectra of the lake.

A-4 Within Dataset Duplication Removal – Geographic and Spectral Filter

From the between dataset geographic duplicates we acquired 4629 modern sites. There are 1494 duplicate longitude/latitude points. Of these 1494 duplicates, there are 276 sites in the modern database that together have 1218 geographic duplicates (the sites have zero distance between them). However, of these duplicate sites there is only 1 spectral duplicate - Round Lake 1710 and 3705 site ID.

```
ID 1710 W493 RDLK0 -86.63 41.23 216
ID 3705 W3521 ROUNDLK -86.63 216
```

Two programs were written to determine whether we had spectral and geographic duplicates in the dataset.

A-4.1 ELIM.DUPLICATES.2.F

```
function(inModern, t1, tn, long, lat)
{
stime <- proc.time()
duplist <- vector("numeric")
stime <- proc.time()
data1 <- as.matrix(inModern[, t1:tn])
dimnames(data1) <- NULL
data2 <- data1
data1 <- t(data1)
for(i in 1:length(data2[, 1])) {
if(length(which(duplist == i)) == 0) {
x <- colSums((data1^0.5 - data2[i,  ]^0.5)^2)
dups <- which(x == 0)[which(x == 0) != i]
if(length(dups) > 0) {
dists <- apply(inModern[dups, c(long, lat)], 1, euclidean.distance.f, unlist(inModern[i,
c(long, lat)]))
distZero <- which(dists == 0)
dups1 <- dups[distZero]
if(length(dups1) > 0) {
  duplist <- c(duplist, dups1)#
  write(paste(c(i, dups1)), "d:\\rdupsIdenticalLongLatSpectra.txt", ncolumns =
length(dups1) + 1, append = T)
  write(paste(i, inModern[i, 4], inModern[i, 5], inModern[i, 6], inModern[i, 7],
inModern[i, 10], inModern[
    dups1, 4], inModern[dups1, 5], inModern[dups1, 6], inModern[dups1, 10]),
"d:\\rdupssiteinfo.txt", ncolumns
     = length(dups1) * 10, append = T)
}
}
}
write(proc.time() - stime, "d:\\time.txt", append = T)
}
duplist
}
```

A-4.2 WITHIN DATASET SPECTRAL DUPLICATION

There are some possible discrepancies in the geographic coordinates entered for some sites. In particular, there are a number of sites that are likely identical sites but have their coordinates entered incorrectly in either WW2000 or GPD. These sties have identical pollen spectra. Here Dr. Jack Williams' sites were retained over the GPD sites. ID refers to the ID of the 4629 sites. Here we can see by the pollen set that the sites are duplicated.

RowNames	ID	ENUM	SITENUM	SITECODE	SITENAME	LON2D	LAT2D	LONG	LAT	ELEV	POLLSUM
307	307.00	35.00	35.00	G35	Bugbee Bog	-72.15	44.37	-72.15	44.37	398	13100.00
2476	2476.00	-9999.00	1664.00	W1664	BUGBEE	-72.00	44.83	-72.00	44.83	398	13100.00
327	327.00	129.00	129.00	G129	Lac Louis	-79.12	47.29	-79.12	47.29	300	232.00
2502	2502.00	-9999.00	1693.00	W1693	LACLOUIS	-79.12	47.28	-79.12	47.28	300	232.00
330	330.00	136.00	136.00	G136	Lake B	-106.08	53.80	-106.08	53.80	553	229.00
3638	3638.00	-9999.00	3426.00	W3426	LAKEB	-106.07	53.80	-106.07	53.80	553	229.00
350	350.00	240.00	239.00	G239	Silver Lake [Nova Scotia]	-63.64	44.56	-63.64	44.56	69	96.00

2539	2539.00	-9999.00	1732.00	W1732	SILVERNS	-63.65	44.57	-63.65	44.57	69	96.00
584	584.00	977.00	893.00	G893	Haeusser Farm Pond	-91.70	44.20	-91.70	44.20	-9999	270.00
1555	1555.00	-9999.00	326.00	W326	3	-91.70	44.17	-91.70	44.17	213	270.00
586	586.00	979.00	895.00	G895	Strum	-91.39	44.55	-91.39	44.55	271	214.00
1557	1557.00	-9999.00	328.00	W328	5	-91.38	44.55	-91.38	44.55	268	214.00
587	587.00	980.00	896.00	G896	Dead	-92.05	44.58	-92.05	44.58	-9999	292.00
1558	1558.00	-9999.00	329.00	W329	6A	-92.03	44.57	-92.03	44.57	210	292.00
591	591.00	985.00	901.00	G901	Menomin	-91.93	44.88	-91.93	44.88	248	304.00
1563	1563.00	-9999.00	334.00	W334	10	-91.92	44.88	-91.92	44.88	248	304.00
592	592.00	986.00	902.00	G902	Tainter	-91.89	44.94	-91.89	44.94	266	318.00
1564	1564.00	-9999.00	335.00	W335	11	-91.80	44.82	-91.80	44.82	268	318.00
599	599.00	993.00	909.00	G909	Bass [WI1]	-91.49	45.42	-91.49	45.42	397	337.00
1571	1571.00	-9999.00	342.00	W342	18	-91.50	45.43	-91.50	45.43	397	337.00
606	606.00	1000.00	916.00	G916	Little St. Germain	-89.50	45.90	-89.50	45.90	-9999	350.00
1576	1576.00	-9999.00	349.00	W349	22	-89.47	45.90	-89.47	45.90	496	350.00
626	626.00	1022.00	938.00	G938	White Clay	-88.40	44.80	-88.40	44.80	246	327.00
1602	1602.00	-9999.00	377.00	W377	44	-88.40	44.82	-88.40	44.82	250	327.00
636	636.00	1032.00	948.00	G948	Stratton	-89.18	44.29	-89.18	44.29	270	380.00
1611	1611.00	-9999.00	387.00	W387	53	-89.18	44.30	-89.18	44.30	274	380.00
637	637.00	1034.00	950.00	G950	Tuttle	-89.29	43.93	-89.29	43.93	248	447.00
1613	1613.00	-9999.00	389.00	W389	55	-89.28	43.93	-89.28	43.93	247	447.00
639	639.00	1036.00	72.00	G72	Devils Lake [Wisconsin]	-89.73	43.42	-89.73	43.42	294	379.00
1615	1615.00	-9999.00	391.00	W391	57	-89.72	43.42	-89.72	43.42	294	379.00
642	642.00	1039.00	955.00	G955	Montfort	-90.43	42.97	-90.43	42.97	-9999	334.00
1618	1618.00	-9999.00	394.00	W394	60	-90.45	42.95	-90.45	42.95	329	334.00
643	643.00	1040.00	956.00	G956	Cox Hollow	-90.11	43.01	-90.11	43.01	292	332.00
1619	1619.00	-9999.00	395.00	W395	61	-90.10	43.00	-90.10	43.00	299	332.00
644	644.00	1041.00	957.00	G957	Raemisch Farm Pond	-90.00	43.00	-90.00	43.00	-9999	301.00
1620	1620.00	-9999.00	396.00	W396	62	-90.05	42.97	-90.05	42.97	372	301.00
645	645.00	1042.00	958.00	G958	Stewart	-89.80	43.00	-89.80	43.00	-9999	296.00
1621	1621.00	-9999.00	397.00	W397	63	-89.75	43.02	-89.75	43.02	363	296.00
651	651.00	1048.00	964.00	G964	Gibbs	-89.18	42.78	-89.18	42.78	267	200.00
1626	1626.00	-9999.00	403.00	W403	69	-89.20	42.78	-89.20	42.78	268	200.00
652	652.00	1050.00	966.00	G966	Ripley	-88.99	43.00	-88.99	43.00	255	317.00
1628	1628.00	-9999.00	405.00	W405	71	-89.00	43.00	-89.00	43.00	262	317.00
670	670.00	1073.00	989.00	G989	Lac Patricia	-64.68	56.67	-64.68	56.67	538	335.00
2526	2526.00	-9999.00	1718.00	W1718	PATRICIA	-64.70	56.67	-64.70	56.67	540	335.00
946	946.00	1632.00	1415.00	G1415	Station 3 [Lichti-Federovich & Ritchie	-65.00	63.05	-65.00	63.05	61	996.00

					1968]						
1455	1455.00	- 9999.00	166.00	W166	LR3	-95.00	63.50	-95.00	63.50	61	996.00
969	969.00	1723.00	1467.00	G1467	Station 94 [Lichti-Federovich & Ritchie 1968]	-107.05	54.67	- 107.05	54.67	472	997.00
1495	1495.00	- 9999.00	257.00	W257	LR94	-107.50	54.67	- 107.50	54.67	472	997.00
974	974.00	1751.00	1494.00	G1494	Station 122 [Lichti-Federovich & Ritchie 1968]	-98.67	55.05	-98.67	55.05	233	993.00
1521	1521.00	- 9999.00	285.00	W285	LR122	-98.67	55.50	-98.67	55.50	233	993.00
975	975.00	1752.00	1495.00	G1495	Station 123 [Lichti-Federovich & Ritchie 1968]	-96.05	54.67	-96.05	54.67	233	994.00
1522	1522.00	- 9999.00	286.00	W286	LR123	-96.50	54.67	-96.50	54.67	233	994.00
995	995.00	1972.00	1693.00	G1693	Site 2 [Heusser C.J. 1978]	-124.08	44.50	- 124.08	44.50	46	100.00
3427	3427.00	- 9999.00	3104.00	W3104	ORESU 6	-124.03	44.50	- 124.03	44.50	46	100.00
996	996.00	1984.00	1705.00	G1705	Site 14 [Heusser C.J. 1978]	-123.50	44.35	- 123.50	44.35	305	100.00
3439	3439.00	- 9999.00	3116.00	W3116	ORESU 18	-123.50	44.52	- 123.50	44.52	305	100.00
997	997.00	2042.00	1759.00	G1759	Site 15 [Mack R.N. V.M. Bryant Jr. and W. Pell. 1978]	-116.25	47.72	- 116.25	47.72	1554	850.00
2914	2914.00	- 9999.00	2558.00	W2558	IDASU 18	-116.08	47.72	- 116.08	47.72	1554	850.00
998	998.00	2068.00	1783.00	G1783	Site 40 [Mack R.N. V.M. Bryant Jr. and W. Pell. 1978]	-117.27	46.27	- 117.27	46.27	732	420.00
3307	3307.00	- 9999.00	2968.00	W2968	WASU184	-117.27	46.68	- 117.27	46.68	732	420.00
1003	1003.00	2117.00	1832.00	G1832	Upper Satsop Lakes 1	-123.52	47.41	- 123.52	47.41	945	1000.00
3193	3193.00	- 9999.00	2854.00	W2854	WASU 1	-123.35	47.41	- 123.35	47.41	945	1000.00
1062	1062.00	2358.00	843.00	G843	Beckman Lake	-93.18	45.43	-93.18	45.43	278	447.00
2644	2644.00	- 9999.00	1853.00	W1853	BECKMMN	-93.18	45.42	-93.18	45.42	-999	447.00
1063	1063.00	2359.00	844.00	G844	Cedar Bog Lake	-93.20	45.41	-93.20	45.41	275	469.00
2645	2645.00	- 9999.00	1854.00	W1854	CEBOGMN	-93.20	45.42	-93.20	45.42	-999	469.00
1095	1095.00	2399.00	1162.00	G1162	Spiritwood Lake	-98.59	47.07	-98.59	47.07	437	212.00
1367	1367.00	- 9999.00	51.00	W51	NDMCA01	-98.58	47.10	-98.58	47.10	444	212.00
1096	1096.00	2400.00	350.00	G350	Glatsch Lake	-92.62	47.28	-92.62	47.28	426	349.00
1442	1442.00	- 9999.00	137.00	W137	GLATCH	-92.43	47.47	-92.43	47.47	425	349.00
1097	1097.00	2401.00	89.00	G89	Hayes Lake	-93.75	49.58	-93.75	49.58	391	490.00
1726	1726.00	- 9999.00	511.00	W511	ONTMCA8	-93.73	49.65	-93.73	49.65	410	490.00
1108	1108.00	2843.00	2219.00	G2219	Site 6 [Swain unpublished]	-88.05	42.93	-88.05	42.93	259	290.00
1894	1894.00	- 9999.00	905.00	W905	C100PTYS	-88.50	42.93	-88.50	42.93	259	290.00
1710	1710.00	- 9999.00	493.00	W493	RDLK0	-86.63	41.23	-86.63	41.23	216	468.00
3705	3705.00	- 9999.00	3521.00	W3521	ROUNDLK	-86.63	41.23	-86.63	41.23	216	468.00

3812	3812.00	-9999.00	4102.00	W4102	ARCH12	-111.00	32.00	-32.00	728	42.00
								111.00		
3839	3839.00	-9999.00	4129.00	W4129	ARCH17	-116.00	34.00	-34.00	610	42.00
								116.00		

The following sites were excluded from the final dataset.

307	307.00	35.00	35.00	G35	Bugbee Bog
327	327.00	129.00	129.00	G129	Lac Louis
330	330.00	136.00	136.00	G136	Lake B
350	350.00	240.00	239.00	G239	Silver Lake [Nova Scotia]
584	584.00	977.00	893.00	G893	Haeusser Farm Pond
586	586.00	979.00	895.00	G895	Strum
587	587.00	980.00	896.00	G896	Dead
591	591.00	985.00	901.00	G901	Menomin
592	592.00	986.00	902.00	G902	Tainter
599	599.00	993.00	909.00	G909	Bass [WI1]
606	606.00	1000.00	916.00	G916	Little St. Germain
626	626.00	1022.00	938.00	G938	White Clay
636	636.00	1032.00	948.00	G948	Stratton
637	637.00	1034.00	950.00	G950	Tuttle
639	639.00	1036.00	72.00	G72	Devils Lake [Wisconsin]
642	642.00	1039.00	955.00	G955	Montfort
643	643.00	1040.00	956.00	G956	Cox Hollow
644	644.00	1041.00	957.00	G957	Raemisch Farm Pond
645	645.00	1042.00	958.00	G958	Stewart
651	651.00	1048.00	964.00	G964	Gibbs
652	652.00	1050.00	966.00	G966	Ripley
670	670.00	1073.00	989.00	G989	Lac Patricia
946	946.00	1632.00	1415.00	G1415	Station 3 [Lichti-Federovich & Ritchie 1968]
969	969.00	1723.00	1467.00	G1467	Station 94 [Lichti-Federovich & Ritchie 1968]
974	974.00	1751.00	1494.00	G1494	Station 122 [Lichti-Federovich & Ritchie 1968]
975	975.00	1752.00	1495.00	G1495	Station 123 [Lichti-Federovich & Ritchie 1968]
995	995.00	1972.00	1693.00	G1693	Site 2 [Heusser C.J. 1978]
996	996.00	1984.00	1705.00	G1705	Site 14 [Heusser C.J. 1978]
997	997.00	2042.00	1759.00	G1759	Site 15 [Mack R.N. V.M. Bryant Jr. and W. Pell. 1978]
998	998.00	2068.00	1783.00	G1783	Site 40 [Mack R.N. V.M. Bryant Jr. and W. Pell. 1978]
1003	1003.00	2117.00	1832.00	G1832	Upper Satsop Lakes 1
1062	1062.00	2358.00	843.00	G843	Beckman Lake
1063	1063.00	2359.00	844.00	G844	Cedar Bog Lake
1095	1095.00	2399.00	1162.00	G1162	Spiritwood Lake
1096	1096.00	2400.00	350.00	G350	Glatsch Lake
1097	1097.00	2401.00	89.00	G89	Hayes Lake
1108	1108.00	2843.00	2219.00	G2219	Site 6 [Swain unpublished]
1710	1710.00	-9999.00	493.00	W493	RDLK0
3839	3839.00	-9999.00	4129.00	W4129	ARCH17

These spectral duplicates were determined using the following code:

A-4.3 NEWDUPS.SPECTRA_ELIM.DUPLICATES.3.F

```
function(inModern, t1, tn, long, lat)
{
stime <- proc.time()
duplist <- vector("numeric")
stime <- proc.time()
data1 <- as.matrix(inModern[, t1:tn])
dimnames(data1) <- NULL
data2 <- data1
data1 <- t(data1)
for(i in 1:length(data2[, 1])) {
if(length(which(duplist == i)) == 0) {
x <- colSums((data1^0.5 - data2[i,  ]^0.5)^2)
dups <- which(x == 0)[which(x == 0) != i]
if(length(dups) > 0) {
write(paste(c(i, dups)), "c:\\rdupsIdenticalLongLatSpectra.txt", ncolumns = length(dups)
+ 1, append = T)
write(paste(i, inModern[i, 4], inModern[i, 5], inModern[i, 6], inModern[i, 7],
inModern[i, 10], inModern[dups, 4], inModern[dups, 5], inModern[dups,
   6], inModern[dups, 10]), "c:\\rdupssiteinfo.txt", ncolumns = length(dups) * 10, append
= T)
duplist <- c(duplist, dups)#
}
}
write(proc.time() - stime, "c:\\time.txt", append = T)
}
duplist
}
```

The modern.percent was then subsetted with this newdups.spectra vector creating a newdups.spectra.set table which was then run trough the following program to get all duplicates.

A-4.4 DUPLICATES_DUPS.NEW.F

```
function()
{
out <- vector("numeric")
for(i in 1:length(newdups.spectra.set[, 1])) {
d <- which(as.vector(unlist(colSums(((t(as.matrix(modern.percent[, 12:100]))^0.5) -
(unlist(newdups.spectra.set[i, 12:100])^0.5))^2))) == 0)
out <- c(out, d)
}
out
}
```

This is the table above

A-5 ELEVATION FILLING/EXTRACTION FOR MISSING DATA

1. GTOPO 30arc second sets for North America are held in the following tiles:

W180N90	W140N90
W100N90	W060N90
W140N40	W100N40

These are in Band Interleaved by Line (BIL) format and are 16bit motorola byte order.

2. These were brought into IDRISI32 and converted to IDRISI *rst* files
3. –9999 is used in the files to indicate missing data and/or oceans.
4. The –9999 is reclassed as 0 because we need to merge the 6 sections (each 58 megs) into a single mosaic in ArcView. ArcView Image Analyst recongnizes ERDAS IMG files. ERDAS IMG files cannot hold negative numbers.
5. 140N40 tile has minimum values of elevation –79 m and greater. As such in order to make this into an ERDAS file we reclassed all values less than 0 as 0 in this image.
6. The ERDAS files were created and added as themes to an ArcView View.
7. These were exported as TIFF files.
8. The TIFF files were imported back into arcview and converted to GRID themes (decreases file size and access time).
9. Then the 6 tiles were merged, to do this I wrote the following Avenue program:

```
theView = av.getactivoc
thethemes = theview.getthemes
theGrid={}
for each t in theView.GetThemes
  if (t.Is(GTHEME)) then
    ' get grid and display source name
    theGrid.add(t.GetGrid)
  end
end
filenam="c:\temp\geotopo\nwgrd1"'saved one of the tiles
in this grid
filenam.asstring
athename=Grid.MakeSrcName(filenam)
newGrid=grid.make(athename)
newgrid=newGrid.Merge (theGrid)
thegtheme=gtheme.make(newgrid)
theview.addtheme(thegtheme)
```

10. Once merged, the new layer was saved and a text file was added to the project containing the ID, LONG, LAT and ELEVATION fields of the modern database.
11. The "Summarize by Zones" was used with the unique ID's on the Merged tiles to get the elevations for all of the Modern Data Set points.
12. Those with missing elevations were extracted and added to the original elevation field of the modern dataset. One point has no topo elevation in East Greeenland. AS well some GTOPO elevations are zero because points fall on coastal islands not covered by GTOPO cells.

    (300 300 -9999 1024 K1024 SOMMERSO -14.67 81.58 -14.67 81.58 25 NA)

Those sites with missing elevations are (with the missing elevations added here):

ID	ENUM	SITENUM	SITE	CODE					SITENAME
LON2D	LAT2D		LONG	LAT	ELEV				
120	120	-9999	29	K29	ALPEN SIDING	-113.00	54.45	-113.00000	54.45000	688
121	121	-9999	30	K30	REINDEER L	-102.58	56.67	-102.58000	56.67000	337
123	123	-9999	32	K32	SEWELL LAKE	-99.25	49.58	-99.25000	49.58000	375
183	183	-9999	92	K92	GMS15	-114.17	53.55	-114.17000	53.55000	753
355	355	256	255	G255	Swan Lake	-102.51	41.72	-102.50861	41.72139	1160
438	438	637	585	G585	GMM047	-128.17	65.75	-128.16667	65.75000	92
439	439	641	589	G589	GMM055	-122.23	62.33	-122.23333	62.33333	633
440	440	649	597	G597	GMM064	-115.97	62.83	-115.96667	62.83333	156
441	441	654	602	G602	GMM075	-114.67	55.25	-114.66667	55.25000	578
442	442	683	631	G631	GMM121	-111.08	49.92	-111.08333	49.91667	835
452	452	730	675	G675	Bad Medicine Lake	-95.38	47.13	-95.38333	47.13333	488
454	454	735	680	G680	Cottonwood Lake [Minnesota]	-96.02	45.76	-96.01667	45.76111	335
455	455	737	682	G682	Dalton Lake	-95.92	46.17	-95.91528	46.17389	411
457	457	739	684	G684	Faith Pond	-95.10	47.28	-95.10000	47.28333	462
459	459	741	686	G686	Fox Pond	-95.56	47.17	-95.55833	47.16500	510
461	461	746	691	G691	Horse Pond	-95.87	47.19	-95.87222	47.19028	402
468	468	758	702	G702	Reichow	-95.78	47.19	-95.77500	47.18778	467
469	469	759	703	G703	Rholl Lake	-96.17	47.60	-96.16667	47.60000	365
475	475	765	709	G709	Stone Lake	-93.52	45.40	-93.52222	45.40000	305
476	476	766	710	G710	Toenjes Lake	-93.98	45.41	-93.98139	45.40972	293
477	477	767	711	G711	Twin Lakes	-93.17	45.85	-93.16778	45.84889	305
488	488	780	722	G722	Pasqua Lake	-104.00	50.78	-104.00000	50.78000	481
583	583	976	892	G892	Merrick State Park	-91.75	44.15	-91.74806	44.15167	207
588	588	981	897	G897	Silver Birch	-92.03	44.60	-92.02722	44.59806	210
589	589	982	898	G898	Thompson [WI]	-91.99	44.63	-91.98667	44.63444	243
596	596	990	906	G906	Friendly Pond	-92.00	45.20	-92.00000	45.20000	335
597	597	991	907	G907	Island	-91.40	45.30	-91.40000	45.30000	333
600	600	994	910	G910	Park	-91.10	45.50	-91.10000	45.50000	361
604	604	998	914	G914	Weber	-89.70	46.00	-89.70000	46.00000	494
605	605	999	915	G915	Helmet	-89.70	46.00	-89.70000	46.00000	494
609	609	1005	921	G921	North [WI1]	-88.20	45.90	-88.20000	45.90000	403
612	612	1008	924	G924	Hilbert	-88.40	45.70	-88.40000	45.70000	473
620	620	1016	932	G932	Mueller	-88.91	45.16	-88.91417	45.16000	438
627	627	1023	939	G939	Kroening	-88.87	44.80	-88.87111	44.80472	289
628	628	1024	940	G940	Bog Pond	-88.90	44.80	-88.90000	44.80000	287
629	629	1025	941	G941	Pine [WI2]	-88.66	44.69	-88.65778	44.68611	245
630	630	1026	942	G942	Twin	-88.99	44.57	-88.98944	44.56722	267
631	631	1027	943	G943	North [WI2]	-89.10	44.60	-89.10000	44.60000	305
632	632	1028	944	G944	School Section	-88.93	44.55	-88.92944	44.54972	256
635	635	1031	947	G947	Silver	-89.37	44.35	-89.37083	44.34917	341
641	641	1038	954	G954	Farm Pond	-90.00	43.10	-90.00000	43.10000	294
646	646	1043	959	G959	Weedy Farm Pond	-89.70	43.10	-89.70000	43.10000	274
648	648	1045	961	G961	Harriet	-89.45	42.93	-89.44833	42.92639	303
666	666	1068	984	G984	James Bay	-80.20	50.22	-80.20000	50.22000	214
707	707	1166	1070	G1070	Wells Mastodon Site	-86.43	41.08	-86.43333	41.08333	236
731	731	1261	1145	G1145	Lamb Lake	-92.11	48.17	-92.10722	48.16528	390
734	734	1268	1151	G1151	Shelby Pond	-81.64	31.08	-81.63694	31.08083	6
756	756	1338	1216	G1216	Dogfish Lake	-92.18	48.18	-92.18333	48.18333	427
758	758	1340	1218	G1218	Burntside Lake	-91.99	47.92	-91.99472	47.92417	421

760	760	1342	1220	G1220	Harding Lake	-146.91	64.44	-146.90750	64.44444	219
761	761	1343	1221	G1221	Tangle Lakes	-146.06	63.03	-146.06333	63.02944	868
762	762	1344	1222	G1222	Point Moronzof Peat Section	-149.48	61.12	-149.48000	61.12000	699
927	927	1548	1337	G1337	Hidden Lake	-144.66	63.94	-144.65833	63.94000	388
1021	1021	2301	2010	G2010	Watchtower Basin	-117.08	52.78	-117.08333	52.78333	2096
1024	1024	2305	2014	G2014	Site 1 [Jensen 1998]	-80.35	26.36	-80.35306	26.35889	4
1025	1025	2306	2015	G2015	Site 2 [Jensen 1998]	-80.37	26.36	-80.36611	26.35917	4
1026	1026	2307	2016	G2016	Site 3 [Jensen 1998]	-80.35	26.35	-80.34583	26.35444	4
1027	1027	2308	2017	G2017	Site 4 [Jensen 1998]	-80.34	26.35	-80.34000	26.35139	4
1028	1028	2309	2018	G2018	Site 5 [Jensen 1998]	-80.31	26.35	-80.30778	26.35278	4
1029	1029	2310	2019	G2019	Site 6 [Jensen 1998]	-80.33	26.26	-80.33306	26.26139	3
1030	1030	2311	2019	G2019	Site 6 [Jensen 1998]	-80.33	26.26	-80.33306	26.26139	3
1031	1031	2312	2020	G2020	Site 7 [Jensen 1998]	-80.32	26.28	-80.32444	26.27500	3
1032	1032	2313	2021	G2021	Site 8 [Jensen 1998]	-80.31	26.31	-80.31472	26.30694	3
1033	1033	2314	2022	G2022	Site 9 [Jensen 1998]	-80.31	26.31	-80.31361	26.31361	4
1034	1034	2315	2023	G2023	Site 10 [Jensen 1998]	-80.31	26.34	-80.30639	26.33611	4
1035	1035	2316	2023	G2023	Site 10 [Jensen 1998]	-80.31	26.34	-80.30639	26.33611	4
1036	1036	2317	2024	G2024	Site 11 [Jensen 1998]	-80.30	26.34	-80.30361	26.34306	4
1037	1037	2318	2025	G2025	Site 12 [Jensen 1998]	-80.32	26.35	-80.32000	26.34806	4
1038	1038	2319	2026	G2026	Site 13 [Jensen 1998]	-80.34	26.34	-80.33833	26.34167	4
1039	1039	2320	2026	G2026	Site 13 [Jensen 1998]	-80.34	26.34	-80.33833	26.34167	4
1040	1040	2321	2027	G2027	Site 14 [Jensen 1998]	-80.38	26.29	-80.37917	26.28944	3
1041	1041	2322	2028	G2028	Site 15 [Jensen 1998]	-80.39	26.30	-80.38639	26.30222	3
1042	1042	2323	2029	G2029	Site 16 [Jensen 1998]	-80.39	26.31	-80.38528	26.31250	3
1043	1043	2324	2030	G2030	Site 17 [Jensen 1998]	-80.37	26.33	-80.36889	26.32500	4
1044	1044	2325	2031	G2031	Site 18 [Jensen 1998]	-80.34	26.34	-80.33694	26.34056	4
1045	1045	2326	2032	G2032	Site 19 [Jensen 1998]	-80.37	26.33	-80.36583	26.33306	4
1046	1046	2327	2033	G2033	Site 20 [Jensen 1998]	-80.36	26.31	-80.35778	26.30778	4
1047	1047	2328	2034	G2034	Site 21 [Jensen 1998]	-80.36	26.28	-80.35667	26.27639	3
1048	1048	2329	2035	G2035	Site 22 [Jensen 1998]	-80.36	26.30	-80.35583	26.29639	3
1049	1049	2330	2036	G2036	Site 23 [Jensen 1998]	-80.37	26.31	-80.36861	26.30611	3
1050	1050	2331	2037	G2037	Site 24 [Jensen 1998]	-80.39	26.25	-80.38889	26.25056	3
1051	1051	2332	2038	G2038	Site 25 [Jensen 1998]	-80.37	26.31	-80.37278	26.31278	4
1052	1052	2333	2039	G2039	Site 26 [Jensen 1998]	-80.38	26.33	-80.37861	26.32611	4
1053	1053	2334	2040	G2040	Site 27 [Jensen 1998]	-80.38	26.33	-80.37778	26.33361	4
1054	1054	2335	2041	G2041	Site 28 [Jensen 1998]	-80.38	26.36	-80.37778	26.35889	4
1055	1055	2336	2042	G2042	Site 29 [Jensen 1998]	-80.30	26.33	-80.30417	26.33111	4
1056	1056	2337	2043	G2043	Site 30 [Jensen 1998]	-80.33	26.32	-80.32639	26.32056	4
1057	1057	2338	2044	G2044	Site 31 [Jensen 1998]	-80.34	26.31	-80.34278	26.31111	4
1060	1060	2347	2053	G2053	Site F1	-80.31	26.28	-80.30667	26.28167	3
1061	1061	2348	2054	G2054	Site U3	-80.41	26.29	-80.41139	26.28750	3
1064	1064	2365	2066	G2066	USGS Gage 15	-80.38	26.39	-80.38083	26.38722	4
1065	1065	2366	2067	G2067	USGS Gage 34	-80.49	25.87	-80.48500	25.86917	2
1066	1066	2367	2068	G2068	USGS Gage 65	-80.72	25.81	-80.72000	25.81444	1
1067	1067	2368	2069	G2069	USGS Gage 63	-80.53	26.19	-80.53111	26.18861	4
1068	1068	2369	2070	G2070	USGS Gage 76	-80.48	25.01	-80.48278	25.00750	0
1069	1069	2370	2071	G2071	USGS Gage 62	-80.75	26.17	-80.75139	26.17444	4
1070	1070	2371	2072	G2072	WCA 3A Site 3-4-1	-80.47	26.17	-80.46944	26.16694	3
1071	1071	2372	2073	G2073	WCA 3A Site 3-4-2	-80.52	26.19	-80.52306	26.19000	4
1072	1072	2373	2074	G2074	WCA 3A Site 3-4-3	-80.61	26.17	-80.60861	26.17028	4
1073	1073	2374	2075	G2075	WCA 3A Site 3-4-4	-80.47	26.04	-80.46944	26.03583	3
1074	1074	2375	2076	G2076	Loxahatchee Site 1	-80.34	26.52	-80.33889	26.51667	4
1075	1075	2376	2077	G2077	Loop Road Site 2	-80.80	25.32	-80.79861	25.31944	1
1076	1076	2377	2078	G2078	Loxahatchee Site 3	-80.34	26.52	-80.33889	26.52278	4
1077	1077	2378	2079	G2079	Loxahatchee Site 4	-80.34	26.52	-80.34194	26.52278	4
1078	1078	2379	2080	G2080	Loop Road Site 3	-80.80	25.28	-80.79528	25.27861	1
1079	1079	2380	2081	G2081	Loop Road Site 1	-80.80	25.32	-80.79528	25.31944	1
1080	1080	2381	2082	G2082	Taylor Creek Site 1A	-80.64	25.20	-80.64333	25.19778	1
1081	1081	2382	2083	G2083	Taylor Slough Site 2	-80.64	25.21	-80.64389	25.20556	1
1082	1082	2383	2084	G2084	Taylor Slough Site 1A	-80.74	25.21	-80.73639	25.21306	1
1083	1083	2384	2085	G2085	Rotenberger Site 1	-80.85	26.36	-80.84917	26.36333	4
1084	1084	2385	2086	G2086	Taylor Slough Site 7	-80.65	25.29	-80.64639	25.28722	1
1085	1085	2386	2087	G2087	CA315 (AAS)	-80.67	25.97	-80.66889	25.97417	3
1086	1086	2387	2088	G2088	EPA Site 7	-80.42	26.48	-80.41556	26.48111	4
1087	1087	2388	2089	G2089	Mud Creek Site 1	-80.60	25.22	-80.60472	25.21972	1
1088	1088	2389	2090	G2090	Taylor Creek Site 2	-80.64	25.21	-80.64417	25.20556	1
1138	1138	3190	2504	G2504	Menominee Site 1 [Webb unpublished]	-88.51	44.87	-88.51000	44.87000	249
1139	1139	3191	2505	G2505	Menominee Site 2 [Webb unpublished]	-88.52	44.88	-88.52000	44.88000	243
1140	1140	3192	2506	G2506	Menominee Site 3 [Webb unpublished]	-88.85	44.98	-88.85000	44.98000	327
1141	1141	3193	2507	G2507	Menominee Site 4 [Webb unpublished]	-88.49	44.90	-88.49000	44.90000	258
1142	1142	3194	2508	G2508	Menominee Site 5 [Webb unpublished]	-88.55	44.89	-88.55000	44.89000	256
1143	1143	3195	2509	G2509	Menominee Site 6 [Webb unpublished]	-88.60	44.89	-88.60000	44.89000	259
1144	1144	3196	2510	G2510	Menominee Site 7 [Webb unpublished]	-88.71	45.07	-88.71000	45.07000	335
1145	1145	3197	2505	G2505	Menominee Site 2 [Webb unpublished]	-88.52	44.88	-88.52000	44.88000	243
1146	1146	3198	2511	G2511	Menominee Site 8 [Webb unpublished]	-88.70	45.01	-88.70000	45.01000	302
1147	1147	3199	2512	G2512	Menominee Site 9 [Webb unpublished]	-88.66	44.88	-88.66000	44.88000	266
1148	1148	3200	2513	G2513	Menominee Site 10 [Webb unpublished]	-88.63	45.08	-88.63000	45.08000	320
1149	1149	3201	2514	G2514	Menominee Site 11 [Webb unpublished]	-88.57	44.98	-88.57000	44.98000	289
1150	1150	3202	2515	G2515	Menominee Site 12 [Webb unpublished]	-88.75	45.02	-88.75000	45.02000	320
1151	1151	3203	2516	G2516	Menominee Site 13 [Webb unpublished]	-88.53	44.94	-88.53000	44.94000	256
1152	1152	3204	2517	G2517	Menominee Site 14 [Webb unpublished]	-88.58	44.91	-88.58000	44.91000	259
1153	1153	3205	2509	G2509	Menominee Site 6 [Webb unpublished]	-88.60	44.89	-88.60000	44.89000	259
1154	1154	3206	2509	G2509	Menominee Site 6 [Webb unpublished]	-88.60	44.89	-88.60000	44.89000	259
1155	1155	3207	2518	G2518	Menominee Site 15 [Webb unpublished]	-88.89	45.01	-88.89000	45.01000	375
1156	1156	3208	2519	G2519	Menominee Site 16 [Webb unpublished]	-88.63	44.98	-88.63000	44.98000	272
1157	1157	3209	2520	G2520	Menominee Site 17 [Webb unpublished]	-88.66	44.95	-88.66000	44.95000	285
1158	1158	3210	2521	G2521	Menominee Site 18 [Webb unpublished]	-88.62	45.11	-88.62000	45.11000	336
1159	1159	3211	2522	G2522	Menominee Site 19 [Webb unpublished]	-88.58	45.00	-88.58000	45.00000	295

1160	1160	3212	2523	G2523	Menominee Site 20 [Webb unpublished]	-88.57	44.89	-88.57000	44.89000	258
1161	1161	3213	2504	G2504	Menominee Site 1 [Webb unpublished]	-88.51	44.87	-88.51000	44.87000	249
1162	1162	3214	2524	G2524	Menominee Site 21 [Webb unpublished]	-88.51	44.89	-88.51000	44.89000	259
1163	1163	3215	2522	G2522	Menominee Site 19 [Webb unpublished]	-88.58	45.00	-88.58000	45.00000	295
1164	1164	3216	2525	G2525	Menominee Site 22 [Webb unpublished]	-88.70	45.09	-88.70000	45.09000	350
1165	1165	3217	2526	G2526	Menominee Site 23 [Webb unpublished]	-88.72	44.87	-88.72000	44.87000	286
1166	1166	3218	2512	G2512	Menominee Site 9 [Webb unpublished]	-88.66	44.88	-88.66000	44.88000	266
1167	1167	3219	2512	G2512	Menominee Site 9 [Webb unpublished]	-88.66	44.88	-88.66000	44.88000	266
1168	1168	3220	2527	G2527	Menominee Site 24 [Webb unpublished]	-88.55	45.00	-88.55000	45.00000	271
1169	1169	3221	2528	G2528	Menominee Site 25 [Webb unpublished]	-88.52	45.28	-88.52000	45.27667	380
1170	1170	3222	2529	G2529	Menominee Site 26 [Webb unpublished]	-88.83	45.02	-88.83000	45.02000	376
1171	1171	3223	2530	G2530	Menominee Site 27 [Webb unpublished]	-88.62	45.05	-88.62000	45.05000	308
1172	1172	3224	2531	G2531	Menominee Site 28 [Webb unpublished]	-88.82	45.03	-88.82000	45.03000	352
1173	1173	3225	2532	G2532	Menominee Site 29 [Webb unpublished]	-88.63	44.89	-88.63000	44.89000	260
1174	1174	3226	2533	G2533	Menominee Site 30 [Webb unpublished]	-88.52	44.90	-88.52000	44.90000	254
1175	1175	3227	2534	G2534	Menominee Site 31 [Webb unpublished]	-88.89	44.98	-88.89000	44.98000	346
1176	1176	3228	2535	G2535	Menominee Site 32 [Webb unpublished]	-88.50	44.99	-88.50000	44.99000	273
1177	1177	3229	2513	G2513	Menominee Site 10 [Webb unpublished]	-88.63	45.08	-88.63000	45.08000	320
1178	1178	3230	2536	G2536	Menominee Site 33 [Webb unpublished]	-88.61	45.10	-88.61000	45.10000	320
1179	1179	3231	2537	G2537	Menominee Site 34 [Webb unpublished]	-88.54	44.93	-88.54000	44.93000	257
1180	1180	3232	2521	G2521	Menominee Site 18 [Webb unpublished]	-88.62	45.11	-88.62000	45.11000	336
1181	1181	3233	2534	G2534	Menominee Site 31 [Webb unpublished]	-88.89	44.98	-88.89000	44.98000	346
1182	1182	3234	2534	G2534	Menominee Site 31 [Webb unpublished]	-88.89	44.98	-88.89000	44.98000	346
1183	1183	3235	2538	G2538	Menominee Site 35 [Webb unpublished]	-88.84	44.98	-88.84000	44.98000	320
1184	1184	3236	2529	G2529	Menominee Site 26 [Webb unpublished]	-88.83	45.02	-88.83000	45.02000	376
1185	1185	3237	2509	G2509	Menominee Site 6 [Webb unpublished]	-88.60	44.89	-88.60000	44.89000	259
1186	1186	3238	2522	G2522	Menominee Site 19 [Webb unpublished]	-88.58	45.00	-88.58000	45.00000	295
1187	1187	3239	2512	G2512	Menominee Site 9 [Webb unpublished]	-88.66	44.88	-88.66000	44.88000	266
1188	1188	3240	2539	G2539	Menominee Site 36 [Webb unpublished]	-88.77	45.01	-88.77000	45.01000	334
1189	1189	3241	2527	G2527	Menominee Site 24 [Webb unpublished]	-88.55	45.00	-88.55000	45.00000	271
1190	1190	3242	2540	G2540	Menominee Site 37 [Webb unpublished]	-88.50	44.87	-88.50000	44.87000	258
1191	1191	3243	2541	G2541	Menominee Site 38 [Webb unpublished]	-88.52	45.09	-88.52000	45.09000	289
1192	1192	3244	2537	G2537	Menominee Site 34 [Webb unpublished]	-88.54	44.93	-88.54000	44.93000	257
1193	1193	3245	2542	G2542	Menominee Site 39 [Webb unpublished]	-88.96	45.06	-88.96000	45.06000	395
1194	1194	3246	2543	G2543	Menominee Site 40 [Webb unpublished]	-88.71	44.87	-88.71000	44.87000	283
1195	1195	3247	2531	G2531	Menominee Site 28 [Webb unpublished]	-88.82	45.03	-88.82000	45.03000	352
1196	1196	3248	2544	G2544	Menominee Site 41 [Webb unpublished]	-88.54	44.89	-88.54000	44.89000	255
1197	1197	3249	2545	G2545	Menominee Site 42 [Webb unpublished]	-88.52	45.11	-88.52000	45.11000	313
1198	1198	3250	2546	G2546	Menominee Site 43 [Webb unpublished]	-88.51	45.09	-88.51000	45.09000	275
1199	1199	3251	2547	G2547	Menominee Site 44 [Webb unpublished]	-88.63	44.92	-88.63000	44.92000	272
1200	1200	3252	2514	G2514	Menominee Site 11 [Webb unpublished]	-88.57	44.98	-88.57000	44.98000	289
1201	1201	3253	2517	G2517	Menominee Site 14 [Webb unpublished]	-88.58	44.91	-88.58000	44.91000	259
1202	1202	3254	2548	G2548	Menominee Site 45 [Webb unpublished]	-88.57	45.05	-88.57000	45.05000	280
1203	1203	3255	2525	G2525	Menominee Site 22 [Webb unpublished]	-88.70	45.09	-88.70000	45.09000	350
1204	1204	3256	2549	G2549	Menominee Site 46 [Webb unpublished]	-88.64	44.98	-88.64000	44.98000	274
1205	1205	3257	2550	G2550	Menominee Site 47 [Webb unpublished]	-88.80	45.00	-88.80000	45.00000	344
1206	1206	3258	2551	G2551	Menominee Site 48 [Webb unpublished]	-88.62	44.89	-88.62000	44.89000	259
1207	1207	3259	2542	G2542	Menominee Site 39 [Webb unpublished]	-88.96	45.06	-88.96000	45.06000	395
1208	1208	3260	2552	G2552	Menominee Site 49 [Webb unpublished]	-88.52	44.95	-88.52000	44.95000	258
1209	1209	3261	2553	G2553	Menominee Site 50 [Webb unpublished]	-88.84	45.08	-88.84000	45.08000	381
1210	1210	3262	2554	G2554	Menominee Site 51 [Webb unpublished]	-88.88	45.11	-88.88000	45.11000	390
1211	1211	3263	2555	G2555	Menominee Site 52 [Webb unpublished]	-88.92	45.05	-88.92000	45.05000	411
1212	1212	3264	2556	G2556	Menominee Site 53 [Webb unpublished]	-88.93	44.98	-88.93000	44.98000	350
1213	1213	3265	2557	G2557	Menominee Site 54 [Webb unpublished]	-88.78	45.09	-88.78000	45.09000	381
1214	1214	3266	2558	G2558	Menominee Site 55 [Webb unpublished]	-88.50	45.10	-88.50000	45.10000	275
1215	1215	3267	2559	G2559	Menominee Site 56 [Webb unpublished]	-88.56	44.89	-88.56000	44.89000	258
1216	1216	3268	2560	G2560	Menominee Site 57 [Webb unpublished]	-88.75	45.06	-88.75000	45.06000	350
1217	1217	3269	2561	G2561	Menominee Site 58 [Webb unpublished]	-88.52	45.04	-88.52000	45.04000	269
1218	1218	3270	2562	G2562	Menominee Site 59 [Webb unpublished]	-88.60	44.95	-88.60000	44.95000	271
1219	1219	3271	2563	G2563	Menominee Site 60 [Webb unpublished]	-88.55	45.01	-88.55000	45.01000	272
1220	1220	3272	2564	G2564	Menominee Site 61 [Webb unpublished]	-88.75	45.07	-88.75000	45.07000	350
1221	1221	3273	2565	G2565	Menominee Site 62 [Webb unpublished]	-88.89	45.04	-88.89000	45.04000	365
1222	1222	3274	2566	G2566	Menominee Site 63 [Webb unpublished]	-88.51	45.03	-88.51000	45.03000	272
1223	1223	3275	2505	G2505	Menominee Site 2 [Webb unpublished]	-88.52	44.88	-88.52000	44.88000	243
1224	1224	3276	2505	G2505	Menominee Site 2 [Webb unpublished]	-88.52	44.88	-88.52000	44.88000	243
1317	1317	3500	2740	G2740	Sleet Lake	-133.58	69.28	-133.58333	69.28333	30
1318	1318	3501	2741	G2741	Reindeer Lake	-132.17	69.12	-132.16667	69.11667	9
1867	1867	-9999	878	W878	MCA-NS01	-91.32	49.47	-91.32000	49.47000	457
1868	1868	-9999	879	W879	MCA-NS02	-90.78	50.12	-90.78000	50.12000	409
1869	1869	-9999	880	W880	MCA-NS03	-90.67	50.23	-90.67000	50.23000	458
1870	1870	-9999	881	W881	MCA-NS04	-90.48	50.42	-90.48000	50.42000	408
1881	1881	-9999	892	W892	MCA-NS15	-85.15	49.72	-85.15000	49.72000	273
1883	1883	-9999	894	W894	MCA-NS17	-81.72	49.25	-81.72000	49.25000	237
2639	2639	-9999	1848	W1848	CRYSPMO	-99.17	46.90	-99.17000	46.90000	531
2640	2640	-9999	1849	W1849	LAMBTMN	-93.90	45.28	-93.90000	45.28000	304
2641	2641	-9999	1850	W1850	CEDARMN	-93.88	45.27	-93.88000	45.27000	294
2642	2642	-9999	1851	W1851	SUGARMN	-94.02	45.35	-94.02000	45.35000	303
2643	2643	-9999	1852	W1852	WYOMIMN	-93.05	45.33	-93.05000	45.33000	276
2644	2644	-9999	1853	W1853	BECKMN	-93.18	45.42	-93.18000	45.42000	283
2645	2645	-9999	1854	W1854	CEDOGMN	-93.20	45.42	-93.20000	45.42000	280
2646	2646	-9999	1855	W1855	WAMPSMN	-90.42	48.05	-90.42000	48.05000	516
2647	2647	-9999	1856	W1856	ROADMN	-90.48	48.05	-90.48000	48.05000	533
2648	2648	-9999	1857	W1857	SPEDMN	-90.43	48.05	-90.43000	48.05000	529
2649	2649	-9999	1858	W1858	CLWTRMN	-90.32	48.08	-90.32000	48.08000	556
2650	2650	-9999	1859	W1859	SQINTMN	-90.47	48.05	-90.47000	48.05000	533

A-6 SUMMARY

The final modern dataset WWKGVS2000 has 4590 modern pollen sites. There are still areas in the southwestern United States where data is sparse, this is also true of the Arctic regions of North America. There are no sites in Mexico which may limit the ability of analog reconstructions. There are also data lacking in the western Cordilleran regions of North America.

Appendix B

CALIBRATING THE DATES IN THE NAPD AGEBASIS TABLE FOR PURPOSES OF DERIVING NEW AGE-DEPTH CHRONOLOGIES FOR THE NAPD.

Some of the dates in the NAPD Age Basis table do not require re-calibration to calendar years using the INTCAL98 Calibration Curve. Therefore, these must be excluded or manually changed and the rest of the dates, e.g., 14C dates, sent for re-calibration. Recalibration was undertaken using programs in the S+ language ("#" indicates comments in the code).

B-1 SITES EXCLUDED FROM DATE CALIBRATION

The following were excluded from the recalibration from ^{14}C to Calendar Years:

```
# Which should not be excluded from date conversion
calibT_is.na(calibT==F)

# rcode.exclude.new
[1] "ANL" "DEN" "CAL" "CCA" "AMR" "ESH" "TSD"

# Get rcode column
rcode_napdab[,6]

# Match rcodes in column with those that must be excluded
calibT_is.na(match(rcode,rcode.exclude.new))

# Make column of exclusion inclusion and add to original table
napdab_cbind(napdab,calibT)

# Make an index field so that converted numbers can be put back again.
napdab_cbind(napdab,ID=1:length(napdab[,1]))

# Get those records that are TRUE for date conversion
tocalibratex_napdab[which(napdab[,7]==T),c(5,8)

# Calibrate the dates by getting their mean
tocalibratex.calibrated.mean_apply(as.matrix(tocalibratex[,1]),1,calcal.new.f,intcal98[,4
],((intcal98[,1]-1950)*-1),handle=T,mean)
tocalibratex.calibrated.median_apply(as.matrix(tocalibratex[,1]),1,calcal.new.f,intcal98[
,4],((intcal98[,1]-1950)*-1),handle=T,median)
tocalibratex.calibrated.max_apply(as.matrix(tocalibratex[,1]),1,calcal.new.f,intcal98[,4]
,((intcal98[,1]-1950)*-1),handle=T,max)
tocalibratex.calibrated.min_apply(as.matrix(tocalibratex[,1]),1,calcal.new.f,intcal98[,4]
,((intcal98[,1]-1950)*-1),handle=T,min)
```

```
# Add the calibrated to the uncalibrated list
tocalibratex_cbind(tocalibratex,i98mean=tocalibratex.calibrated.mean,i98medain=tocalibrat
ex.calibrated.median,i98max=tocalibratex.calibrated.max,i98min=tocalibratex.calibrated.mi
n)

# Copy age column over in napdab
napdab_cbind(napdab,ICAL98Mean=napdab[,5],ICAL98Median=napdab[,5],ICAL98Max=napdab[,5],IC
AL98Min=napdab[,5])

# Put the calibrated dates in the napdab table
napdab[tocalibratex[,2],9:12]_tocalibratex[,3:6]

# Change the date of tsuga decline to 5300 calibrated years
napdab[which(napdab[,6]=="TSD"),9:12]_c(5300,5300,5300,5300)
```

Name	Description
E.	Entity Number
Chron.	Chronology
Sample.	Sample Number
DepthCM	Depth (cm)
Age	Age c14 Years BP
RCode	Sample Type
CalibT	Logical: TRUE means the value was calibrated, False = Sample not calibrated using INTCAL98
ID	Primary KEY
ICAL98Mean	Calibrated INTCAL98 using mean of n possible ages
ICAL98Median	Calibrated INTCAL98 using median of n possible ages
ICAL98Max	Calibrated INTCAL98 using maximum of n possible ages
ICAL98Min	Calibrated INTCAL98 using minimum of n possible ages

B-2 Creating the Calibrated Datasets from INTCAL 1998: Remodeling the NAPD

The remodelling of the NAPD requires taking each of the dates and depths from the AGEBASIS table and creating linear equations from one date to another. For example, for E# 2, we have 3 dates at 3 depths,

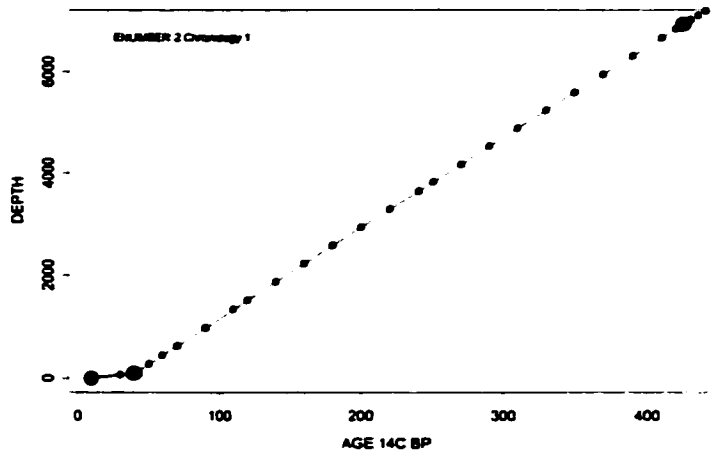


Figure: Graph showing linearly interpolated (red) dates from depth based on original radiocarbon dates (Blue dots)

This is done for each of the E#'s and for each Chronology for each E# which range between 1-4 possible chronologies. A few sites have no chronology at all!

B-2.1 Remodelling the AGEBASIS Table

The remodelling involves the two tables AGEBASIS and AGEDEPTH. The AGEBASIS table contains the following fields:

E.	Chron.	Sample.	DepthCM	Age	RCode
2	1	1	10	0	TOP
2	1	2	40	100	ESH
2	1	3	425	6090	C14
3	1	1	0	2000	TOP
3	1	2	50	4055	C14
3	1	3	160	5945	C14
3	1	4	230	5985	C14
3	2	1	50	4055	C14
3	2	2	195	5962	CAV
4	1	1	527	30	TOP

In this table not all Ages are to be calibrated, and these include "ANL","DEN","CAL","CCA","AMR","ESH","TSD". Specifically, TSD is the *Tsuga* Decline and is set to 5300 ka with AgeUp 5100 and AgeLo 5500, the others are all substituted as they are. Also is excluded, CCA should be excluded as well. "TOP" the *Ambrosia* rise is to be calibrated as well as the European Settlement Horizon. The AGEDEPTH Table contains the following fields:

E.	Chron.	Sample.	DepthCM	AgeBP

2	1	1	10	0
2	1	2	30	66
2	1	3	50	255
2	1	4	60	411
2	1	5	70	566
2	1	6	90	877
2	1	7	110	1189

B-2.2 CALIBRATING THE AGEDEPTH TABLE

A number of sites in AGEDEPTH have only one observed level in the core. Therefore, these were not modeled and but their dates were calibrated.

Site 211 – Chron 2 – all NA in AGEDEPTH
Site 329 – Chron 1 – all NA in AGEDEPTH except for Depths 197-290.

A further number of sites have no radiocarbon chronologies (that is modeled ages) for part of their AGEDEPTH sequences but do have a middle set of modeled numbers, specifically,

Site 370 – Chron 1 – NA is only from depths 20 to 60.

All others are given chronologies except:

Site 86 – Has no data in AGEBASIS so its radiocarbon chronology is simply transferred and corrected ages.

Another site has 4 different depths 143 to 160 but has only one date associated with all three, namely

Site 2295 – Has only one entry in the AGEBASIS table but has 4 associated depths very close together in the AGEDEPTH table all with the same 14C date so these were directly converted to calandar years – see below.

```
#make the datasets of calibrated modelled age-depths
        model.mean_remodel.f(ageBasisAgeCol = 9)
        model.median_remodel.f(ageBasisAgeCol = 10)
        model.max_remodel.f(ageBasisAgeCol = 11)
        model.min_remodel.f(ageBasisAgeCol = 12)

        #Add the calibrated dates to the napdad.calibrated matrix, also add the vectors
        that tell whether a date was extrapolated within the model (xxEx - 1 =
        extrapolated within age-depth model)
        napdad.calibrated_cbind(napdad,calMean=model.mean[,5],calMedian=model.median[,5],c
        alMin=model.min[,5],calMax=model.max[,5],
        calMeanEx=model.mean[,6],calMedianEx=model.median[,6],calMinEx=model.min[,6],calMa
        xEx=model.max[,6])

        #Calibrate the single date for site 2295
        napdad.calibrated[which(napdad.calibrated[,1]==2295),6:9]_c(1286.54,1286.54,1286.5
        4,1286.54)

        #Calibrate the dates for site 86 - these are directly calibrated - only the mean
        dates were returned for all sets.
        toadd.to_apply(as.matrix(napdad.calibrated[which(napdad.calibrated[,1]==86),5]),1,
        calcal.new.f,intcal98[,4],((intcal98[,1]-1950)*-1),handle=T,mean)
        napdad.calibrated[which(napdad.calibrated[,1]==86),6:9]_toadd.to
```

napdad.calibrated - calibrated dataset contains the following fields:

Name	Description
E.	Entity Number
Chron.	Chronology
Sample.	Sample Number
DepthCM	Depth (cm)
AgeBP	Age 14C
calMean	Calendar Years Mean Age of n Possible
calMedian	Calendar Years Median Age of n Possible
calMin	Calendar Years Minimum Age of n Possible
calMax	Calendar Years Maximum Age of n Possible
calMeanEx	Logical: False = not extrapolated beyond range of Calibrated Dates in AGEBASIS age-depth model
calMedianEx	Logical: False = not extrapolated beyond range of Calibrated Dates in AGEBASIS age-depth model
calMinEx	Logical: False = not extrapolated beyond range of Calibrated Dates in AGEBASIS age-depth model
calMaxEx	Logical: False = not extrapolated beyond range of Calibrated Dates in AGEBASIS age-depth model

B-2.3 CALIBRATING THE THE NAPD 14C TABLE

If any date is negative then the calibrated date is equal to the C14 date.

E.	Sample.	AgeBP	AgeSDUp	AgeSDLo
2	1	6090	85	85
3	1	4055	120	120
3	2	5945	120	120
3	3	5985	140	140
4	1	5720	100	100

The following dates in napdc14 do not have any associated AgeUp or AgeDown
which.na(napdc14[,5])
[1] 178 943 977 978 1056 1121 1238 1241 1242 1244 1397 1398 1400 1696 1836 1837 1969 1970 2048 2049 2352 2356 2357 2358 2716 2780 2811 2836 2929 2930 2931 2933 2934 2989 3053 3062

CALIBRATION PROCEDURE

```
#Calibration of napdc14 table
napdc14.mean_apply(as.matrix(napdc14[,3]),1,calcal.new.f,intcal98[,4],((intcal98[,
1]-1950)*-1),handle=T,mean)
```

```
napdc14.median_apply(as.matrix(napdc14[,3]),1,calcal.new.f,intcal98[,4],((intcal98
[,1]-1950)*-1),handle=T,median)
napdc14.max_apply(as.matrix(napdc14[,3]),1,calcal.new.f,intcal98[,4],((intcal98[,1
]-1950)*-1),handle=T,max)
napdc14.min_apply(as.matrix(napdc14[,3]),1,calcal.new.f,intcal98[,4],((intcal98[,1
]-1950)*-1),handle=T,min)

napdc14.mean.upsd_apply(as.matrix(napdc14[,3]+napdc14[,4]),1,calcal.new.f,intcal98
[,4],((intcal98[,1]-1950)*-1),handle=T,mean)
napdc14.mean.losd_apply(as.matrix(napdc14[,3]-
napdc14[,5]),1,calcal.new.f,intcal98[,4],((intcal98[,1]-1950)*-1),handle=T,mean)
napdc14.median.upsd_apply(as.matrix(napdc14[,3]+napdc14[,4]),1,calcal.new.f,intcal98[,4],((int
cal98[,1]-1950)*-1),handle=T,median)
napdc14.median.losd_apply(as.matrix(napdc14[,3]-
napdc14[,5]),1,calcal.new.f,intcal98[,4],((intcal98[,1]-1950)*-1),handle=T,median)
napdc14.max.upsd_apply(as.matrix(napdc14[,3]+napdc14[,4]),1,calcal.new.f,intcal98[
,4],((intcal98[,1]-1950)*-1),handle=T,max)
napdc14.max.losd_apply(as.matrix(napdc14[,3]-
napdc14[,5]),1,calcal.new.f,intcal98[,4],((intcal98[,1]-1950)*-1),handle=T,max)
napdc14.min.upsd_apply(as.matrix(napdc14[,3]+napdc14[,4]),1,calcal.new.f,intcal98[
,4],((intcal98[,1]-1950)*-1),handle=T,min)
napdc14.min.losd_apply(as.matrix(napdc14[,3]-
napdc14[,5]),1,calcal.new.f,intcal98[,4],((intcal98[,1]-1950)*-1),handle=T,min)

napdc14.calibrated_cbind(napdc14,napdc14.mean, napdc14.mean.upsd,
napdc14.mean.losd , napdc14.median, napdc14.median.upsd, napdc14.median.losd,
napdc14.max, napdc14.max.upsd, napdc14.max.losd, napdc14.min, napdc14.min.upsd,
napdc14.min.losd)
```

napdc14.calibrated – final calibrated dataset

Calibration of these requires converting the AgeBP to Calendar Years in ka. I've added the following fields to the ^{14}C Table:

Ageka	Calibrated INTCAL98 – the mean of n possible dates
AgekaSDUp	Calibrated with 14C date + SDUp – the mean of n possible dates
AgekaSDLo	Calibrated with 14C date – SDLo – the mean of n possible dates
mnka	Calibrated based on minimum of n possible calendar years returned
mnkaSDup	Minimum of the n possible calendar years returned from AgekaSDUp + SDUp
mnkaSDLo	Minimum of the n possible calendar years returned from AgekaSDUp - SDUp
mxka	Calibrated based on the maximum of the n possible calendar years returned
mxkaSDUp	Maximum of the n possible calendar years returned from AgekaSDUp + SDUp
mxkaSDLo	Maximum of the n possible calendar years returned from AgekaSDLo – SDLo
mdka	Median of the n possible calendar years
mdkaSDUp	Median of the n possible calendar years returned from AgekaSDUp + SDUp
mdkaSDLo	Median of the n possible calendar years returned from AgekaSDLo – SDLo

```
> names.f(napdc14.calibrated)

     Position                    name
[1,] "1"      "E."
[2,] "2"      "Sample."
[3,] "3"      "AgeBP"
[4,] "4"      "AgeSDUp"
[5,] "5"      "AgeSDLo"
[6,] "6"      "napdc14.mean"
[7,] "7"      "napdc14.mean.upsd"
[8,] "8"      "napdc14.mean.losd"
```

```
 [9,] "9"      "napdc14.median"
[10,] "10"     "napdc14.median.upsd"
[11,] "11"     "napdc14.median.losd"
[12,] "12"     "napdc14.max"
[13,] "13"     "napdc14.max.upsd"
[14,] "14"     "napdc14.max.losd"
[15,] "15"     "napdc14.min"
[16,] "16"     "napdc14.min.upsd"
[17,] "17"     "napdc14.min.losd"
```

B-2.4 CALIBRATING THE NAPD AGEBOUND TABLE

The NAPD AgeBound Table contains the following:

E.	Chron.	Top	Bottom
2	1	0	6000
3	2	3390	6030
4	1	30	9890
5	1	0	10000
6	1	2250	9970
7	1	150	12000

Calibration Procedure

```
napdagebound.mean.top_apply(as.matrix(napdagebound[,3]),1,calcal.new.f,intcal98[,4
],((intcal98[,1]-1950)*-1),handle=T,mean)
napdagebound.median.top_apply(as.matrix(napdagebound[,3]),1,calcal.new.f,intcal98[
,4],((intcal98[,1]-1950)*-1),handle=T,median)
napdagebound.max.top_apply(as.matrix(napdagebound[,3]),1,calcal.new.f,intcal98[,4]
,((intcal98[,1]-1950)*-1),handle=T,max)
napdagebound.min.top_apply(as.matrix(napdagebound[,3]),1,calcal.new.f,intcal98[,4]
,((intcal98[,1]-1950)*-1),handle=T,min)
napdagebound.mean.bottom_apply(as.matrix(napdagebound[,4]),1,calcal.new.f,intcal98
[,4],((intcal98[,1]-1950)*-1),handle=T,mean)
napdagebound.median.bottom_apply(as.matrix(napdagebound[,4]),1,calcal.new.f,intcal
98[,4],((intcal98[,1]-1950)*-1),handle=T,median)
napdagebound.max.bottom_apply(as.matrix(napdagebound[,4]),1,calcal.new.f,intcal98[
,4],((intcal98[,1]-1950)*-1),handle=T,max)
napdagebound.min.bottom_apply(as.matrix(napdagebound[,4]),1,calcal.new.f,intcal98[
,4],((intcal98[,1]-1950)*-1),handle=T,min)

napdagebound.calibrated_cbind(napdagebound,napdagebound.mean.top,
napdagebound.median.top, napdagebound.max.top, napdagebound.min.top,
napdagebound.mean.bottom, napdagebound.median.bottom, napdagebound.max.bottom,
napdagebound.min.bottom )

The following fields were added to the AgeBound Table

napdagebound.calibrated - is final calibrated dataset
```

Topka	Calibrated INTCAL98 - the mean of n possible dates
Bottomka	Calibrated INTCAL98 - the mean of n possible dates
mdTopka	Median of the n possible calendar years
mdBottomka	Median of the n possible calendar years

mnTopka	Calibrated based on minimum of n possible calendar years returned
mnBottomka	Calibrated based on minimum of n possible calendar years returned
mxTopka	Calibrated based on the maximum of the n possible calendar years returned
mxBottomka	Calibrated based on the maximum of the n possible calendar years returned

B-3 THE REMODELING PROGRAM CODE

remodel.f

```
function(ageBasis = napdab, ageBasisEnum = 1, ageBasisAgeMod = 2, ageBasisDepthCol
= 4, ageBasisAgeCol = 9, ageDepth = napdad, ageDepthEnum = 1, ageDepthAgeMod = 2,
ageDepthDepths
 = 4, ageDepthC14 = 5)
{
#tables must contain the calibrated dates before using theis program
#ageBasis is the table with the depths and original radiocarbon dates used in
making an age-depth model which is the AGEBASIS table
#ageDepth is the table with the depths for all samples in each core that need to
be modeled which is the P_AGEDEPT table
#
#
#> napdab[1:5,]   is ageBasis
#  E. Chron. Sample. DepthCM  Age RCode calibT ID INTCAL98
#1  2      1       1      10    0   TOP  FALSE  1    0.000
#2  2      1       2      40  100   ESH  FALSE  2  100.000
#3  2      1       3     425 6090   C14   TRUE  3 6924.817
#4  3      1       1       0 2000   TOP  FALSE  4 2000.000
#5  3      1       2      50 4055   C14   TRUE  5 4526.597
#ageBasisEnum - the enumber column of the ageBasis table here called napdab
#ageBasisAgeMod - the column containing the age model used e.g., age model 1 or 2
or 3 or 4 for each also called the Chronology which is preferred
#ageBasisDepthCol - the column which contains the depths of the dated levels in
the core
#ageBasisAgeCol - the column which contains the ages 14C or Calibrated for each
level in the core in ageBasisDepthCol
#
#
#> napdad[1:5,]   is ageDepth
#  E. Chron. Sample. DepthCM AgeBP
#1  2      1       1      10     0
#2  2      1       2      30    66
#3  2      1       3      50   255
#4  2      1       4      60   411
#5  2      1       5      70   566
#ageDepthEnum - the column with the enumber for the ageDepth table or napdad
dataset here.  Each Enumber corresponds to a unique site.
#ageDepthAgeMod - the column containing the age modelled chronology (e.g., all
core depths counted and all estimated ages for each core depth using the data in
ageBasis)
#ageDepthDepths - the column containing the depths in each core.
#ageDepthC14 -
#This is
#> names.f(napdab)   is ageBasis
#[[1]]:
#[1] "E."       "Chron."   "Sample."  "DepthCM"  "Age"      "RCode"    "calibT"
"ID"       "INTCAL98"
#> names.f(napdad)   is ageDepth
#[[1]]:
#[1] "E."      "Chron."  "Sample." "DepthCM" "AgeBP"
enumout <- vector("numeric")
agemodelout <- vector("numeric")
depthout <- vector("numeric")
```

```
        calageout <- vector("numeric")
        c14age <- vector("numeric")
        rangeout <- vector("logical")
        enum <- unique(ageDepth[, ageDepthEnum])#ENUMBER
        totalsites <- length(enum)
        for(i in 1:totalsites) {
        #age model 1 or 4
        for(k in 1:4) {
        #From the table ageDepth (napdad), for a given age Chronology (ageDepthAgeMod) and
        a given site (ageDepthEnum) get the locations of all
        #for that Enumber and chronology in the large ageDepth table and store this in
        currentdepths variable
         currentdepths <- which(ageDepth[, ageDepthEnum] == enum[i] & ageDepth[,
        ageDepthAgeMod] == k)#Get the age depth for current model
        #write(c(any(currentdepths), enum[i]), "c:\\sdf.txt", append = T)
        #From teh table ageBasis (containing the actual depth and c14 or calibrated dates)
        for the same chronology and enumber as in currentdepths
        #get the depth and c14 ages to create a linear interpolated model to estimate ages
        at currentdepths.
         currentdepthage <- which(ageBasis[, ageBasisEnum] == enum[i] & ageBasis[,
        ageBasisAgeMod] == k)#Get the age depth for current model
        #See if there are any depths to be estimated because there are some age models in
        ageBasis that have no corresponding depth profiles within
        #the ageDepth table
         if(any(currentdepths)) {
        #
        #put in depthage a two column vector containing the actual depths and 14C or
        calibrated ages (Depth, Date)
         depthage <- ageBasis[currentdepthage, c(ageBasisDepthCol, ageBasisAgeCol)]#get
        cols 4 and 9 depth and calibrated age intcal98
        #take the depths from depthage and store in depth
         depth <- depthage[, 1]#
        #take the ages for the depths and put them in ageCalBP
         ageCalBP <- depthage[, 2]#get detphs to be calibrated from master depth file
        #get the depth data from ageDepth (napdad) in a vector for which ages will be
        estimated by linear interpolation from depth and ageCalBP
         depthforinterpolation <- ageDepth[currentdepths, ageDepthDepths]#
        #get the c14 ages from the original ageBasis age model table and put them in c14
         c14 <- ageDepth[currentdepths, ageDepthC14]#
        #make vector showing which of the target depths lead to extrapolated dates because
        they fall outside of the depth range at which there are actual sampled dates.
         maximumDepthInAgeBasis <- max(depth)
         minimumDepthInAgeBasis <- min(depth)
         mineVector <- depthforinterpolation < minimumDepthInAgeBasis |
        depthforinterpolation > maximumDepthInAgeBasis#x>=max(y)|x<=min(y)
        #Ensure that there are at least two ages and depths in the age basis model table
        so that a linear equation can be created between the two points.
         testlen <- length(depth)
         if(any(testlen) & (testlen >= 2)) {
        #do the linear interpolation of x = depth, y = ageCalBP (from ageBasis) to the
        xout = depthforinterpolation (from ageDepth) with rule = 3 for linear
        extrapolation
        #to ensure that all depths in ageDepth are given a modelled age. May be useful
        here to put in a separate column showing which are modelled outside the data
        #range and which are modelled within the data range in ageBasis.
        #outputs a list with x and y where x = depthforinterpolation vector and y = the
        interpolated ages at those depths
           agedepthmodel <- approx(depth, ageCalBP, xout = depthforinterpolation, rule =
        3)
         }
         else if(length(depthforinterpolation) == testlen) {
           agedepthmodel <- list(x = depthforinterpolation, y = ageCalBP)
}
else {
        #if there are not enough dates and depths in the ageBasis table then simply put
        NA's for all the depthforinterpolation depths and no estimated ages.
```

```
                                        agedepthmodel <- list(x = depthforinterpolation, y
        = rep(NA, length(depthforinterpolation)))
}
                                enumout <- c(enumout, rep(enum[i],
        length(depthforinterpolation)))# the current enumber
                                agemodelout <- c(agemodelout, rep(k,
        length(depthforinterpolation)))#the current agemodel
                                depthout <- c(depthout, agedepthmodel$x)#the depths at
        which ages were interpolated from the agemodel
                                calageout <- c(calageout, agedepthmodel$y)#the ages that
        were interpolated or modelled at the depthforinterpolation depths
c14age <- c(c14age, c14)# the original c14dates in the file
rangeout <- c(rangeout, mineVector)
}
}
}
        #return a large matrix with all the information
                cbind(enum = enumout, agebasis = agemodelout, depth = depthout, c14yrBP =
        c14age, intcal98 = calageout, extrapolated = rangeout)
        }
```

Appendix C

RADIOCARBON TO CALENDAR YEAR CONVERSION

C-1 DESCRIPTION

Using INTCAL98 (Stuiver and Reimer 1998) as input, this function linearly interpolates a given 14C year to its calendar year equivalent. The user can choose to report all intercepts, or some summary of the intercepts. This function requires a DLL called "linint.dll" to be loaded before use – *see notes*.

USAGE

calcal.new.f(targetdate, radiocarbon, calandar, handle = F, how = mean)

REQUIRED ARGUMENTS

targetdate	***Numeric***. 14C date for which the calendar equivalent is desired.
radiocarbon	***Vector Numeric***. A vector of radiocarbon years.
calandar	***Vector Numeric***. A vector of calendar years corresponding to those in the radiocarbon vector.

The radiocarbon and calendar arguments must have the same length. The targetdate argument cannot be larger than the largest 14C year in radiocarbon.

OPTIONAL ARGUMENTS

handle	***Logical***. Should multiple intercepts be handled. Default is False in which case all intercepts will be returned to the user.
how	***Function***. If handle=T, then how should multiple intercepts be handled. Any S summary function can be named here such as mean (default), min, max, median. The function will return a vector with the "how" of the multiple intercepts.

VALUE

Vector with 1 component, which is the calibrated age(s) in Cal yr BP.

NOTE

This function requires the Dynamic Linked Library (DLL) called <linint.dll> to be loaded before use. Place linint.dll in your working directory and use dll.load("linint.dll","linint","cdecl"). This will be unloaded when the session is closed.

DETAILS

The function has been tested against the CALIB 4.0 output and provides results within +/- 1 yr. To get results in cal yrBP you need to subtract 1950 from the AD.BC column of INTCAL98 before sending it to this function. If you give negative radiocarbon dates to this function they are forced to 0 cal yrBP.

AUTHOR

M.Sawada
Laboratory of Paleoclimatology and Climatology
Department of Geography
University of Ottawa
Ottawa, ON KIN 6N5
msawada@aix1.uottawa.ca

REFERENCES

Stuiver, M., and Reimer, P.J., 1993. Radiocarbon, 35:215-230.
Stuiver, M., Reimer, P.J., Bard, E., Beck, J.W., Burr, G.S., Hugen, K.A., Kromer, B., McCormac, F.G., v.d. Plicht, J., and Spurk, M. (1998). Radiocarbon 40:1041-1083.

SEE ALSO

approx, mean, min, max, median

EXAMPLES

```
#Calibrate a single date assuming radiocarbon years are column 4 and calendar
years are column 1 of the intcal98 matrix
        calcal.new.f(2200,intcal98[,4],(intcal98[,1]-1950)*-1)
        [1] 2299.359 2267.931 2178.339 2170.935 2156.548

#Calibrate single date but return the mean
        calcal.new.f(2200,intcal98[,4],(intcal98[,1]-1950)*-1,handle=T,how=mean)
        [1] 2214.622

#Calibrate single date but return the maximum
        calcal.new.f(2200,intcal98[,4],(intcal98[,1]-1950)*-1,handle=T,how=max)
        [1] 2299.359

#Calibrate a vector of dates (returns a list)

        apply(as.matrix(c(2200,3200)),1,calcal.new.f,intcal98[,4], (intcal98[,1]-
        1950)*-1),handle=F,how=mean)
        [[1]]:
        [1] 2299.359 2267.931 2178.339 2170.935 2156.548

        [[2]]:
        [1] 3439.681 3430.075 3400.333
```

```
#Calibrate a vector of dates and return the mean

        apply(as.matrix(c(2200,3200)),1,calcal.new.f,intcal98[,4], (intcal98[,1]-
        1950)*-1),handle=T,how=mean)
        [1] 2214.622 3423.363
```

C-1.1 CODE

```
> calcal.new.f
function(targetdate, radiocarbon, calandar, handle = F, how = mean)
{
lr <- length(radiocarbon)
lc <- length(calandar)
if(lr != lc)
stop(message = "Length of radiocarbon and calandar must be the same")
if(lr < 2)
stop(message = "You need at least two radiocarbon years and two calandar years for
interpolation (two xy pairs)")
l <- rep(0, length(radiocarbon))
if(targetdate < 0) {
#if radio is less than zero then calibrated date equald the radiocarbon date
ot <- targetdate
}
else if((targetdate > 20265.1) & (targetdate <= 27120)) {
#if radio is between 20265.1 and 27120 14C yrs. use Bard (1990) for calibration
ot <- ((30470 - 22955)/(27120 - 19262)) * (targetdate - 19262) + 22955
}
else if((targetdate > 27120) & (targetdate <= 50000)) {
#elseif radio is between 27120 to 50000 use Mazaud(1991) for calibration
ot <- (((50000 - targetdate)/(50000 - 27120)) * 3350) + targetdate
}
else if(targetdate > 50000) {
#elseif radio is greater than 50000 use Mazaud(1991) for calibration
ot <- targetdate
}
else {
if(targetdate > max(radiocarbon))
 stop(message = paste(targetdate, ":Radiocarbon Date must less than or equal to",
max(radiocarbon), ", the maximum 14C date in argument <radiocarbon>!"))
alldates <- .C("linint",
x = as.double(l),
as.double(targetdate),
as.double(radiocarbon),
as.double(calandar),
length(radiocarbon))$x
ot <- unique(alldates[alldates != -99999])
}
if(handle) {
how(ot)
}
else {
ot
}
}
```

C-1.2 C++ CODE FOR THE "LININT.DLL" CALLED BY THE S+ FUNCTION CALCAL.NEW.F

```
/*
#####################################################################################
###################
```

```
#@#32-bit DLL for determining linear interpolations from Radiocarbon years to
Calandar Years using S+
#@#(c) M.Sawada 2000
#@#InputSee Main Function Below for Descriptions
#@#OutputInterpolated y-values.
#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#@#
@#@#@#@#@#@#@#@#@#@#@#@#@#

  working function is double returns results of y to thing
*/
double
inter(double *tofill, double *testval,double *xvec, double *yvec, long *length)
{
//initialize variables for slope m and intercept b method
long i;//counter
//for each row and row+1 in the xvec and yvec's
for (i=0; i<*length; i++){

//check to see if interpolated value will be within the domain of the segment
being tested including endpoints
if(((xvec[i]>=*testval) & (*testval >= xvec[i+1])) | ((xvec[i] <= *testval) &
(*testval <= xvec[i+1])))
{

//make sure x1 and x2 are not the same point or will cause division by zero in
interpolation e.g., xvec[i]-xvec[i+1]==0
if((xvec[i] != xvec[i+1])){

//linearly interpolate the y-value
tofill[i]=yvec[i]+(yvec[i+1] - yvec[i])/(xvec[i+1]-xvec[i])*(*testval-xvec[i]);
}
else{
//if x1 and x2 are the same then the interpolated value is just the y1 or y2
tofill[i]=yvec[i];
}//endif


}//endif
else{

//if testval is outside domain of current test segment then put a -99999.
tofill[i]=-99999;

}//endif

}//endfor
//return the interpolated value array back to *outarray
return *tofill;
}//endinter

 /*
 Main function is void
 outarrayis an array with zeros of equal length as xarray that is filled with
interpolated values
 testvalueis the single testvalue sent to this function for determining the
interpolated date(s)
 xarrayis the x-values - a vector of radiocarbon years
 yarrayis the y-values - a vector of the same length as xarray with corresponding
calandar years
 n is the length of xarray
*/
void linint(double *outarray,double *testvalue,double *xarray,double *yarray, long
*n)
{
```

```
inter(outarray,testvalue,xarray,yarray,n); //send to interpolation function and
provide results back to user

}
```

Appendix D - ROOKCASE

SAWADA, M. 1999. ROOKCASE: An Excel 97/2000 Visual Basic (VB) add-in for exploring global and local spatial autocorrelation. Bulletin of the Ecological Society of America, 80(4):231-234.

DEPARTMENTS

Technological Tools

Note: Dr. David Inouye is the editor of the **Technological Tools** section. Anyone wishing to contribute articles or reviews to this section should contact him at the Department of Zoology, University of Maryland, College Park, MD 20742, E-mail: di5@umail.umd.edu.

ROOKCASE: AN EXCEL 97/2000 VISUAL BASIC (VB) ADD-IN FOR EXPLORING GLOBAL AND LOCAL SPATIAL AUTOCORRELATION

Introduction

Since its first comprehensive review (Cliff and Ord 1973), the quantitative assessment of spatial autocorrelation (SA) has acquired a fundamental role in ecological data analysis, spawning numerous reviews (e.g., Sokal and Oden 1978*a*,*b*, Legendre and Fortin 1989, Dutilleul and Legendre 1993, Legendre 1993, Rossi 1996) and applications (e.g., Sokal and Menozzi 1982, Legendre and Troussellier 1988, Fortin et al. 1989, Graumlich and Davis 1993, Sokal et al. 1998*a*). Measures of SA deal simultaneously with both location and attribute information for a mapped variable Z. SA exists when there is a significant similarity/dissimilarity between the values of a variable Z at all pairs of adjacent locations i, j (Upton and Fingleton 1985). Such a pattern deviates significantly from a map where each value Z_{ij} is assigned randomly across locations. Common measures of global SA are Moran's I (Moran 1950), a product-moment like coefficient, and Geary's C (Geary 1954), a squared-difference coefficient akin to the semi-variogram. The known distributional behavior of these statistics allows hypotheses of spatial influence to be tested. Current developments in quantitative geography have also given rise to Local Indicators of Spatial Association (LISA) such as local Moran's I (I_i) and local Geary's C (C_i) (Anselin 1995, Sokal et al. 1998*a*,*b*), as well as Local Spatial Autocorrelation (LSA) statistics such as the Getis-Ord statistics, G_i and G_i^* (Getis and Ord 1992, Ord and Getis 1995). These local spatial autocorrelation statistics are only now finding applications in the ecological (Sokal et al. 1998*a*,*b*) and environmental sciences (Derksen et al. 1998, Wulder and Boots 1998). LSA/LISA can delineate statistically significant locations of nonstationarity or "hot spots" in mapped patterns where global SA is absent (Ord and Getis 1995, Sokal et al. 1998*b*). However, in the presence of global SA, the distributional properties of LISA/LSA statistics are less certain (Tiefelsdorf and Boots 1997, Sokal et al. 1998*b*), and significance testing is problematic. Nevertheless, LISA (I_i and C_i) are proportional to I and C and can be used to determine the relative contribution of specific locations to the global SA, a property not shared by LSA (Anselin 1995). Consequently, LISA can be used in the exploratory spatial data analysis (ESDA) of patterns exhibiting spatial autocorrelation. The use of LSA/LISA in exploratory spatial data analysis, along with further simulation and theoretical studies, will begin to address their potential for testing specific location-based hypotheses.

Considering the potential of LSA/LISA and the importance of spatial autocorrelation, freely available easy-to-use programs and source code, suitable for research and teaching, are not widely available. Accordingly, I have put forth a user-friendly visual Excel 97/2000 Add-In (Table 1) with source code to facilitate the use of these statistics in the classroom and at an applied level.

Development environment

Excel is easy to use and provides a small learning curve for ROOKCASE, an advantage in the classroom because time is spent learning spatial autocorrelation rather than software. Furthermore, the ease of data input/output, numerous built-in functions, graphing capabilities, and wide availability of Excel make it an ideal platform for ROOKCASE. Finally, because VB for Applications (VBA) is part of Excel, the Add-In has no dependent files and all function code can be used in other VB applications.

Table 1. ROOKCASE user forms, statistics calculated, and details of operation not presented in text. **Performance** is based on average time on an Intel P233 MHz computer using a lag distance where all locations have four neighbors. Arguments are omitted in function calls because they are documented in the function code.

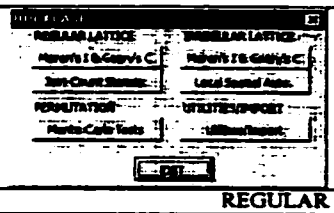
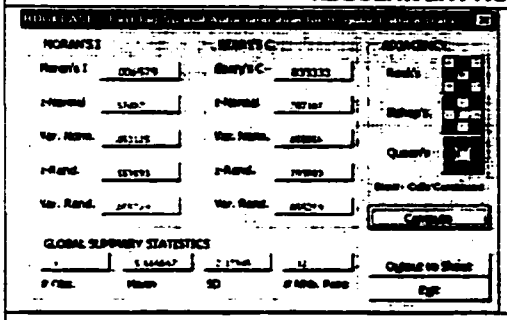
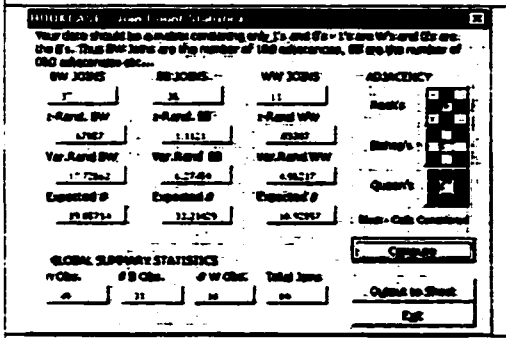
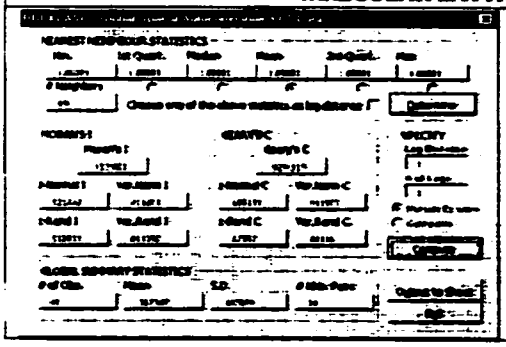
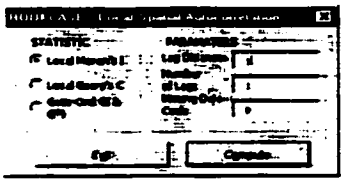
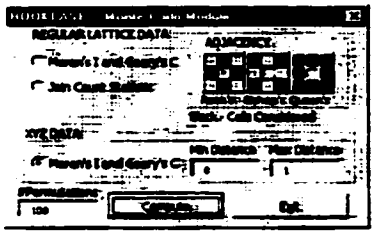
User Form	Operation and Description
	**Main Interface** REGULAR LATTICE - ***DATA FORMAT A*** IRREGULAR LATTICE - ***DATA FORMAT B*** PERMUTATION - ***DATA FORMAT A or B*** UTILITIES/IMPORT - Create random $n \times m$ array; convert DATA FORMAT A to B; Import IDRISI ASCII images
REGULAR LATTICE ***DATA FORMAT A***	
	**Purpose:** First-lag SA according to specified ADJACENCY **Functions Called:** jcRooks; jcBishops; jcQueens **Performance:** 100 x 100 matrix in 10 seconds using Rook's Adjacency
	**Purpose:** Join-Count SA for nominal data according to ADJACENCY **Input:** Data as 1's and 0's. **Function Called:** jcRooks; jcBishops; jcQueens **Performance:** Same as above.
IRREGULAR LATTICE ***DATA FORMAT B***	
	**Purpose:** Compute *I* and *C* and correlograms. **Input:** 1) User can determine "Lag Distance" from NEAREST NEIGHBOUR STATISTICS by first checking Determine and then clicking on the checkbox beside "Choose one... as lag distance" and then a radio button *or*, 2) Enter "Lag Distance" manually **Output:** 1) "# of Lags", if > 1 outputs correlogram data to a new worksheet. Use Excel to graph the resultant correlogram values. **Functions Called:** ICxyz. **Performance:** 400 *XYZ* Pairs in 2 Seconds 1 Lag; 13 seconds for 10 lags.

Table 1(continued). ROOKCASE user forms, statistics calculated, and details of operation not presented in text.

User Form	Operation and Description
	**Purpose:** Calculates a LISA/LSA for each *XYZ* location **Input:** "Missing Data Code" if local variance is undefined - when *XYZ* has no neighbors. If the code is not zero then the global statistic cannot be correctly reported by this module. **Output:** 1) A LISA/LSA, z-variate(s), and number of neighbors for each location. 2) For LISA a global statistic is output with the sum of neighbors, sum of LISA and constant of proportionality. 3) G_i and G_i^* statistics are output as z-variates and in raw form. 4) All lags are cumulative **Function Called:** LocalI; LocalC, LocalG **Performance:** 400 *XYZ* pairs in 5 seconds 1 lag, 10 seconds for 10 lags.
MONTE-CARLO MODULE ***DATA FORMAT A*** or ***B***	
	**Purpose:** Creates n SA coefficients for n permutations of the original data. **Input:** 1) "Min Distance" and "Max Distance" define a particular lag-class (e.g., 50-100 units) **Output:** 1) One statistic for each permutation; 2) The global statistic and a Monte-Carlo p-value where $p = 1+\#\{pSA \geq SA\}/n+1$ if the Joint-Count $\geq E_i$, if $I \geq E_i$ or $C \geq E_i$ and $p = 1+\#\{pSA \leq SA\}/n+1$ if $I \leq E_i$ or if $C \leq E_i$, pSA equals the number of permuted SA coefficients and SA the observed SA value and E_i the expected value of the statistic. **Functions Called:** Any appropriate global SA function **Performance:** 400 *XYZ* data 100 permutations in 2 min., 100 x 100 matrix, 100 permutations in 2 min.

Calculations/equations

The equations for global Moran's I, Geary's C, join-counts and their moments under the normal and randomization hypotheses are found in Griffith (1987) and Sokal and Oden (1978*b*). The formulae for the Getis-Ord statistics, G_i and G_i^*, and their moments under randomization are from Getis and Ord (1992) and Ord and Getis (1995). The equations for Local Moran's I (I_i) and Local Geary's C (C_i) are from Anselin (1995), with moments under total randomization from Sokal et al. (1998*b*). Spatial weights, w_{ij}, in ROOKCASE are symmetrical with $w_{ij} = 1$ if locations i and j are adjacent and are 0 otherwise. However, the program source code can be easily modified to include other adjacency definitions such as row-standardized weights or asymmetrical weights. Adjacency is defined as Rook's (four cardinal neighbors), Bishop's (four diagonal neighbors), or Queen's (four cardinal plus four diagonal neighbors) in regular lattice data or by distance in irregularly spaced *X,Y,Z* data. Program output has been tested whenever possible against published data sets.

Program structure

Microsoft Excel 97/2000 VBA is used for the visual interface and Excel objects are employed for data input and output via six user input forms (Table 1). Each form passes arrays and constants to one of 11 functions that calculate I, C, Join-Counts, I_i, C_i, or G_i and G_i^*. Function code is explicitly documented with respect to passed variables and output so that users may easily modify/use ROOKCASE functions.

Program retrieval

The program is freely available for download at <http://www.uottawa.ca/academic/arts/lpcweb/> and by clicking on "Freeware/Data/LPC Reports." Alternatively, e-mail the author, who will forward the program to you as an attachment.

Installation

(1) Place ROOKCASE.xla in a *directory*. (2) In Excel choose Tools/Add-Ins ... /Browse ... . (3) Navigate to the *directory*. (4) Choose ROOKCASE.xla. (5) Ensure that the checkbox beside RookCase in the Add-Ins Available listbox is checked. A floating button will appear in Excel from which the program is executed. The program code is accessed using the Alt + F11 key sequence and clicking on ROOKCASE.xla in the VBA editor's Project Explorer.

Operation

First, beginning in cell A1 and with no text headings, the user enters data as either a n x m raster of Z-values (called ***DATA FORMAT A***), or as

X,Y,Z in columns A, B, and C, respectively (called ***DATA FORMAT B***), where *X,Y* are in some Euclidean coordinate system (e.g., not latitude and longitude). Second, the user chooses a module appropriate for the data (Table 1). In those modules using ***DATA FORMAT B*** (Table 1), a "Lag Distance" must be specified and must be in the *XY* coordinate units, e.g., meters, and specifies that points separated by this distance or less are used to calculate the spatial autocorrelation statistic. If the "# of Lags" is greater than one, the program creates a correlogram based on mutually exclusive or cumulative distance classes. For example, a lag distance of one unit with three lags produces three SA coefficients for distance classes (adjacencies) from > 0 to ≤ 1, >1 to ≤ 2, >2 to ≤ 3 under the mutually exclusive model, whereas the cumulative model provides coefficients for the classes >0 to ≤ 1, >0 to ≤ 2, >0 to ≤ 3. Program output is labeled and self-explanatory; however, some important abbreviations include: "Var. Norm." and "*z*-Normal," which are the variance and standard *z*-variate based on the normality assumption while "*z*-Rand" and "Var. Rand" are based on the total randomization hypothesis. The "Mean" and sample standard deviations "SD" are reported for *Z* observations. Output Fields labeled "# Nhb. Pairs" report the number of neighbor pairs at the current adjacency definition. Finally, the button "Output to Sheet" directs form fields to a new workbook. All outputs, both user-specified or program driven, are automatically labeled.

Acknowledgments

This note is a contribution to the Climate System History and Dynamics (CSHD) Program, funded by the Natural Sciences and Engineering Research Council (NSERC) and the Atmospheric Environment Service (AES).

Literature cited

Anselin, L. 1995. Local indicators of spatial association—LISA. Geographical Analysis **27**(2):93–115.

Cliff, A. D., and J. K. Ord. 1973. Spatial autocorrelation. Pion, London, UK.

Derksen, C., M. Wulder, E. LeDrew, and B. Goodison. 1998. Associations between spatially autocorrelated patterns of SSM/I-derived prairie snow cover and atmospheric circulation. Hydrological Processes **12**:2307–2316.

Dutilleul, P., and P. Legendre. 1993. Spatial heterogeneity against heteroscedasticity: an ecological paradigm versus a statistical concept. Oikos **66**(1):149–171.

Fortin, M.-J., P. Drapeau, and P. Legendre. 1989. Spatial autocorrelation and sampling design in plant ecology. Vegetatio **83**:209–222.

Geary, R. 1954. The contiguity ratio and statistical mapping. Incorporated Statistician **5**:115–145.

Getis, A., and J. K. Ord. 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis **24**(3):189–206.

Graumlich, L. J., and M. B. Davis. 1993. Holocene variation in spatial scales of vegetation pattern in the upper Great Lakes. Ecology **74**:826–839.

Griffith, D. A. 1987. Spatial autocorrelation: a primer. Association of American Geographers, Resource Publications in Geography, Commercial Printing, Philadelphia, Pennsylvania, USA.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology **74**:1659–1673.

Legendre, P., and M.-J. Fortin. 1989. Spatial pattern and ecological analysis. Vegetatio **80**:107–138.

Legendre, P., and M. Troussellier. 1988. Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. Limnology and Oceanography **33**(5): 1055–1067.

Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. Biometrika **37**:17–23.

Ord, J. K., and A. Getis. 1995. Local spatial autocorrelation statistics: distributional issues and an application. Geographical Analysis **27**(4):287–305.

Rossi, J.-P. 1996. Statistical tool for soil biology. XI. Autocorrelogram and Mantel test. European Journal of Soil Biology **32**(4):195–203.

Sokal, R. R., N. L. Oden, and B. A. Thomson. 1998*a*. Local spatial autocorrelation in biological variables. Biological Journal of the Linnean Society **65**:41–62.

Sokal, R. R., N. L. Oden, and B. A. Thomson. 1998*b*. Local spatial autocorrelation in a biological model. Geographical Analysis **30**(4):331–354.

Sokal, R. R., and P. Menozzi. 1982. Spatial autocorrelations of HLA frequencies in Europe support demic diffusion of early farmers. American Naturalist **119**:1–17.

Sokal, R. R., and N. L. Oden. 1978*a*. Spatial autocorrelation in biology 1. Methodology. Biological Journal of the Linnean Society **10**:199–228.

Sokal, R. R., and N. L. Oden. 1978*b*. Spatial autocorrelation in biology 2. Some biological implications and four applications of evolutionary and ecological interest. Biological Journal of the Linnean Society **10**:229–249.

Tiefelsdorf, M., and B. Boots. 1997. A note on the extremities of Local Moran's I_is and their impact on Global Moran's *I*. Geographical Analysis **29**(3):248–257.

Upton, G. J. G., and B. Fingleton. 1985. Spatial data analysis by example. Volume 1. Point pattern and quantitative data. John Wiley and Sons, Toronto, Ontario, Canada.

Wulder, M., and B. Boots. 1998. Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. International Journal of Remote Sensing **19**(11):2223–2231.

M. Sawada
Laboratory for Paleoclimatology and Climatology
Department of Geography
University of Ottawa
Ottawa, Ontario, Canada K1N 6N5
E-mail: msawada@aix1.uottawa.ca

Appendix E

MODERN ANALOG CALCULATOR

E-1 DESCRIPTION

Creates a modern analog object using the Modern Analog Technique (MAT) employing squared chord distance.

USAGE

calc.dist.f (inFossil, inModern, LongMod = 6, LatMod = 7, t1Mod = 2, tnMod = 5, LongFossil = 6, LatFossil = 6, t1Fossil = 2, tnFossil = 5, TopNAnalogs = 1, sitenames = seq(1, length(inFossil[, 1]), 1), is.Modern.VS.Modern = F, dist.calc = "spherical")

REQUIRED ARGUMENTS

inFossil	***Matrix Numeric***. Contains the fossil sites for which you want the best modern analog. These must be proportions of percentages.
inModern	***Matrix Numeric***. Contains the modern sites for which the best modern analog will be found. These must be proportions or percentages.
LongMod	***Numeric***. The column number of the **inModern** table that contains the Longitude of the site and/or the X-Coordinate (if X,Y is used then dist.calc must be set to "euclidean")
LatMod	***Numeric***. The column number of the **inModern** table that contains the Latitude of the site and/or the X-Coordinate (if X, Y is used then dist.calc must be set to "euclidean")
t1Mod	***Numeric***. The column number of the **inModern** table that contains the first taxon of the pollen sum.
tnMod	***Numeric***. The column number of the **inModern** table that contains the last taxon of the pollen sum. Note that the pollen sum must be continuous from **t1Mod...tnMod**
LongFossil	***Numeric***. The column number of the **inFossil** table that contains the Longitude of the site and/or the X-Coordinate (if X,Y is used then dist.calc must be set to "euclidean")
LatFossil	***Numeric***. The column number of the **inFossil** table that contains the Latitude of the site and/or the X-Coordinate (if X, Y is used then dist.calc must be set to "euclidean"). There cannot be a mix of X,Y and Long, Lat between **inModern** and **inFossil**, they must be the same units (meters, projected or DD).
t1Fossil	***Numeric***. The column number of the **inFossil** table that contains the first taxon of the pollen sum.
tnFossil	***Numeric***. The column number of the **inFossil** table that contains the last taxon of the pollen sum. Note that the pollen sum must be continuous from **t1Mod...tnMod**

OPTIONAL ARGUMENTS

TopNAnalogs	***Numeric***. the number of best analogs to return to the user. If is.Modern.VS.Modern = T then this must be equal to 2.
sitenames	***Numeric/Text***. The site names for the fossil sites.
is.Modern.VS.Modern	***Logical***. Whether or not the analysis is for Modern vs. Modern or a regular reconstruction.
dist.calc	***Character***. Two choices of either "euclidean" or "spherical" to determine how distances and directions to the best modern analog are calculated. NOTE: if "spherical" is selected data must be Longitude/Latitude and only distance to best analog is returned. If choice is "euclidean" then both inModern and inFossil matrices must have X,Y coordinates in meters or feet.

VALUE

A list with the following components:

x	***Vector Numeric***: the x coordinates of the fossil sites
y	***Vector Numeric***: the y coordinates of the fossil sites
xmodanalog	***Matrix:*** the x coordinates of the n best modern analog sites for each fossil site where each row in the matrix corresponds to the nth fossil sites best analog.
ymodanalog	***Matrix***: the y coordinates of the n best modern analog sites for each fossil site where each row in the matrix corresponds to the nth fossil sites best analog.
sqdist	***Matrix*** the squared chord distances of the best modern analogs for each fossil site where each row in the matrix corresponds to the nth fossil sites squared chord distance.
position	***Matrix***: The row position in the inModern matrix of the nth best analog for a given site. The matrix is also by row so that column n row n corresponds to the position of the nth best modern analog row in inModern.
statistics	***Matrix***: The statistics for the squared chord distances for the n best modern analogs. Statistics are by column corresponding to each row in inFossil.
topndistances	***Matrix***: The distances from the fossil site to the best n modern analogs in row order like above.
topndiststats	***Matrix***: The statistics for the squared chord distances for the n best modern analogs. Statistics are by column corresponding to each row in inFossil.
cartesiandir	***Matrix***: The directions from the fossil site to the best n modern analogs in row order like above.
componentx	***Matrix***: The x-component of the vector from the fossil site to the best n modern analogs in row order like above.
componenty	***Matrix***: The y-component of the vector from the fossil site to the best n modern analogs in row order like above.

NOTE

The order of taxon in inModern and inFossil is very important, they columns must match such that column n in each table contains the same taxon. inModern should also contain columns containing the climate parameters for each modern site.

DETAILS

Use the object$position[,n] for subsetting the inModern for the actual reconstructions of climates or other variables within the inModern file.

DEPENDANT FUNCTIONS

recursive.unique.2.f, great.circle.distance.f, euclidean.distance.f, euclidean.direction.f, euclidean.compx.f, euclidean.compy.f

AUTHOR

M.Sawada
Laboratory of Paleoclimatology and Climatology
Department of Geography
University of Ottawa
Ottawa, ON KIN 6N5
msawada@aix1.uottawa.ca

E-1.1 CODE

```
function(inFossil, inModern, LongMod = 6, LatMod = 7, t1Mod = 2, tnMod = 5, LongFossil =
6, LatFossil = 6, t1Fossil = 2, tnFossil = 5, TopNAnalogs = 1, sitenames = seq(1, length(
inFossil[, 1]), 1), is.Modern.VS.Modern = F, dist.calc = "spherical")
{
currtime <- proc.time()
nColsMatrix <- length(inFossil[, 1])
StatsDist <- matrix(NA, nrow = 4, ncol = nColsMatrix)
xmodanalog <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
ymodanalog <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
LocMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
PosMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
DistMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
DirMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
CompxMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
CompyMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
StatsSqdist <- matrix(NA, nrow = 4, ncol = nColsMatrix)
colNames <- vector("character")
StatsSqdistLabels <- c("Max", "Min", "Mean", "SD")
if(is.Modern.VS.Modern) {
bestAnalog <- 1
if(TopNAnalogs <= 1) {
stop("TopNAnalogs must be > 1: For modern vs. modern reconstruction you must choose to
return the 2nd-nth best analog")
}
}
else {
bestAnalog <- 0
}
data1 <- as.matrix(inModern[, t1Mod:tnMod])
data3 <- as.matrix(inFossil[, t1Fossil:tnFossil])
dimnames(data1) <- NULL
dimnames(data3) <- NULL
data1 <- t(data1)
for(i in 1:length(inFossil[, 1])) {
data2 <- data3[i,  ]
sqdistVec <- (data2^0.5 - data1^0.5)^2
x <- colSums(sqdistVec)
write(c("difference.f", as.character(proc.time() - currtime)), "c:\\time.txt", ncolumns =
2, append = T)
StatsSqdist[, i] <- c(max(x), min(x), mean(x), var(x)^0.5)
areUnique <- length(x[duplicated(x)])
if(areUnique > 0) {
x <- recursive.unique.2.f(x)
}
y <- rank(x)
zorder <- order(x[which(y > bestAnalog & y <= TopNAnalogs)])
x <- sort(x[which(y > bestAnalog & y <= TopNAnalogs)])
topn <- which(y > bestAnalog & y <= TopNAnalogs)[zorder]
if(dist.calc == "spherical") {
```

```
distances <- apply(inModern[topn, c(LongMod, LatMod)], 1, great.circle.distance.f,
inFossil[i, c(LongFossil, LatFossil)])
directions <- rep(NA, TopNAnalogs)
compx <- rep(NA, TopNAnalogs)
compy <- rep(NA, TopNAnalogs)
}
else if(dist.calc == "euclidean") {
distances <- apply(inModern[topn, c(LongMod, LatMod)], 1, euclidean.distance.f,
inFossil[i, c(LongFossil, LatFossil)])
directions <- apply(inModern[topn, c(LongMod, LatMod)], 1, euclidean.direction.f,
inFossil[i, c(LongFossil, LatFossil)])
compx <- apply(inModern[topn, c(LongMod, LatMod)], 1, euclidean.compx.f, inFossil[i,
c(LongFossil, LatFossil)])
compy <- apply(inModern[topn, c(LongMod, LatMod)], 1, euclidean.compy.f, inFossil[i,
c(LongFossil, LatFossil)])
}
StatsDist[, i] <- c(max(distances), min(distances), mean(distances), var(distances)^0.5)
xmodanalog[, i] <- unlist(inModern[, LatMod][topn])
ymodanalog[, i] <- unlist(inModern[, LongMod][topn])
LocMinRow[, i] <- x
PosMinRow[, i] <- topn
DistMinRow[, i] <- distances
DirMinRow[, i] <- directions
CompxMinRow[, i] <- compx
CompyMinRow[, i] <- compy
}
colNames <- sitenames
dimnames(LocMinRow) <- list(NULL, colNames)
dimnames(PosMinRow) <- list(NULL, colNames)
dimnames(xmodanalog) <- list(NULL, colNames)
dimnames(ymodanalog) <- list(NULL, colNames)
dimnames(StatsSqdist) <- list(StatsSqdistLabels, colNames)
dimnames(DistMinRow) <- list(NULL, colNames)
dimnames(DirMinRow) <- list(NULL, colNames)
dimnames(CompxMinRow) <- list(NULL, colNames)
dimnames(CompyMinRow) <- list(NULL, colNames)
dimnames(StatsDist) <- list(StatsSqdistLabels, colNames)
write(c("Final Time", as.character(proc.time() - currtime)), "c:\\time.txt", ncolumns =
2, append = T)
list(x = inFossil[, LongFossil], y = inFossil[, LatFossil], xmodanalog = xmodanalog,
ymodanalog = ymodanalog, sqdist = LocMinRow, position = PosMinRow, statistics =
StatsSqdist, topndistances = DistMinRow, topndiststats = StatsDist, cartesiandir =
DirMinRow, componentx = CompxMinRow, componenty = CompyMinRow)
}
```

Appendix F – Gajewski et al. 2001.

Gajewski, K., Viau, A., **Sawada, M.**, Atkinson, D., and, Wilson, S. 2001. ***Sphagnum* peatland distribution in North America and Eurasia during the past 21,000 years.** *Global Biogeochemical Cycles,* 15(2): 297 (2000GB001286)

GLOBAL BIOGEOCHEMICAL CYCLES, VOL. 0, NO. 0, PAGES XXX–XXX, MONTH 2001

Sphagnum peatland distribution in North America and Eurasia during the past 21,000 years

K. Gajewski, A. Viau, M. Sawada, D. Atkinson, and S. Wilson

Laboratory of Paleoclimatology and Climatology, Department of Geography, University of Ottawa, Ottawa, Ontario, Canada

Abstract. The distribution and abundance of *Sphagnum* spores in North America and Eurasia are mapped for the past 21 ka. The present-day distribution of abundant *Sphagnum* spores corresponds closely to areas with peatland development, with maximum *Sphagnum* abundance between 630 and 1300 mm annual precipitation and between −2° and 6°C mean annual air temperature. During the Wisconsin glaciation, there were apparently not large areas of peatland in North America, except in Alaska. High *Sphagnum* spore percentages were found in eastern North America during deglaciation. Major peatland development occurred in boreal North America after 9 ka and there was a southward movement of high *Sphagnum* spore abundance after 5 ka in the western Great Lakes region. Major peatland development began after 9 ka in Europe and Asia. On the basis of maps of the area supporting peatlands, carbon accumulation in peatlands is estimated to be low prior to 11 ka, increased slightly between 11 and 5 ka, and greatly increased during the past 5 ka.

1. Introduction

Since the last glacial maximum (LGM), there have been many biogeographic changes on the Earth's surface. The migration of the major tree species has been analyzed in some detail, especially in North America and Europe where there are extensive networks of pollen data [e.g., *Webb et al.*, 1998; *Prentice et al.*, 1998]. These studies emphasize the development of upland vegetation, but less is known about the development of wetlands and peatlands. Understanding the development of the lowland vegetation is important because peatlands are important components of the carbon cycle [e.g., *Gorham*, 1991]. *Sphagnum* is a characteristic taxon of peat bogs, and the remains of this moss make up much of the biomass of a bog. In the boreal zone, bogs and fens store large amounts of carbon in peat deposits that have accumulated up to 5–6 m or more during the postglacial.

A number of attempts have been made to reconstruct variations in carbon storage since the last glacial maximum. However, estimates have varied considerably [*Crowley*, 1995], and this is due in part to assumptions about the distribution and nature of the vegetation at full glacial. These studies typically reconstruct ecosystem classes defined by some criterion. Carbon estimates are then assigned to the classes to compute changes in storage (reviewed by *Peng et al.* [1998]). However, boreal peatlands are a major source or sink of carbon. Peatlands have not been rigorously included in previous reconstructions of biomes [e.g., *Peng et al.*, 1994, 1998; *Prentice and Fung*, 1990], probably due to lack of quantitative data on peatland distribution through time. Acknowledging this uncertainty, *Adams et al.* [1990] assumed there were few peatlands at full glacial because of the cooler and drier climate, a conclusion questioned by *Crowley* [1995].

In this paper, we investigate one aspect of the biogeographic changes through the glacial-interglacial transition, the distribution of *Sphagnum* peatlands. We present maps of the spatial patterns of *Sphagnum* spores for several time intervals to indicate the distribution of peatland limits though time. Although the emphasis of this study is North America, we present some results from the entire Northern Hemisphere.

Peatlands are ecosystems that sequester carbon through peat accumulation, owing to greater rates of biomass production than decomposition [*Kuhry and Vitt*, 1996]. Because oxygen supply is low in these water-saturated organic soils, decomposition is very slow [*Moore and Bellamy*, 1974]. Boreal and subarctic peatlands store an estimated 455 Pg of carbon, which is approximately one third of the global soil carbon pool, and remove an estimated 0.076 Pg of carbon from the atmosphere annually [*Gorham*, 1991].

Peatlands can be classified as either bogs or fens. Bogs are ombrotrophic systems that receive water and nutrients solely from precipitation and dry fall and are therefore poor in solutes. Fens are minerotrophic systems, which are fed by groundwater and are therefore usually richer in solutes, notably Ca^{2+} [*Gignac and Vitt*, 1990; *van Breemen*, 1995]. Fens may be alkaline or acidic and range from mesotrophic to oligotrophic, whereas bogs are acidic and oligotrophic [*Vitt et al.*, 1994]. These peatlands can be further subdivided on the basis of vegetation cover and the presence of local peat landforms [*Vitt et al.*, 1994; *Halsey et al.*, 1997].

Peatland development is strongly linked to climate, with bogs (and fens) occurring in areas where precipitation exceeds evapotranspiration and where total annual precipitation is typically greater than 500 mm [*Gignac and Vitt*, 1994]. Climate also plays an important role in the regional distribution of peatland types [*Moore and Bellamy*, 1974; *Gignac and Vitt*, 1990, 1994; *Gignac et al.*, 1998]. The extent of peatlands also depends on topography, and flat terrain provides an ideal setting for vast expanses of peatlands [*Zoltai and Pollett*, 1983].

Sphagnum is by far the dominant component of bogs and poor fen peat. No other group of mosses is as ecologically dominant on a worldwide basis [*Andrus*, 1986]. Calculations suggest that there is more carbon locked up in *Sphagnum* (dead and living) than is fixed by all terrestrial vegetation in one year [*Clymo and Hayward*, 1982]. The extensive carpet-like growth and slow rate of decay of *Sphagnum* are two of the primary reasons for its large volume in bogs. The low pH and low dissolved solute concentration of bogs are conducive to *Sphagnum* growth [*Gignac and Vitt*, 1990]. However, the general restriction of *Sphagnum* to these sites is only partly determined by the environment, with *Sphagnum* playing a considerable role in creating these conditions [*Andrus*, 1986].

Paper number 2000GB001286.
0886-6236/01/2000GB001286$12.00

The success of *Sphagnum* is due to its ability to create a habitat in which few other plants can flourishAn environment with high acidity, low nutrient availability, anoxia and low temperatures [*van Breemen*, 1995].

2. Methods

Spores of *Sphagnum* are preserved along with pollen grains in lake and bog sediments. They are easily identified and routinely counted along with other pollen grains. Pollen grains and spores have similar dispersal and preservation properties and data on their presence in sediments are available in public databases. *Sphagnum* data were extracted from sites available in the North American Pollen Database (NAPD) [*Grimm*, 2000b], European Pollen Database (EPD) [*Cheddadi*, 2000], and Global Pollen Database (GPD) [*Grimm*, 2000a] furnished by the National Oceanic and Atmospheric Administration (NOAA) Paleoclimatology Data Center. All available data for North America, Europe, and Siberia were used to make the maps.

Sphagnum percentages were computed within a sum based on all upland plants pollen and Bryophyte spores. Radiocarbon chronologies were obtained from the database. The radiocarbon equivalents of target calendar years spanning 0–21 ka by 2 ka intervals were determined using CALIB version 4.2 [*Stuiver and Reimer*, 1993; M. Stuiver and P. Reimer, Quaternary Isotope Laboratory HTML CALIB 4.2 manual, available at http://depts.washington.edu/qil/calif]. *Sphagnum* percentages were then linearly interpolated to these radiocarbon equivalents to produce our maps.

Maps were then drawn of the abundance of *Sphagnum* spores at 2000-year intervals. Geodetic coordinates of sample sites were projected using the forward spherical solution for the Albers Equal-Area Conic projection [*Snyder*, 1987], and *Sphagnum* percentages were interpolated using a standard inverse-distance squared algorithm [*Lam*, 1983] with a grid resolution of 50 km^2. In order to limit any further extrapolation within and outside the sampling space, we used a search radius of 750 km in North America and 250 km in Europe. Grid cells with *Sphagnum* proportions greater than or equal to 0.5% were selected if they fell within modern-day continental boundaries and outside of ice sheet boundaries for each time slice. The extent of the ice sheets was taken from *Peltier* [1993, 1994]. However, the modern sea level and coastline were retained for simplicity.

Response surfaces [*Bartlein et al.*, 1986; *Gignac et al.*, 1991] were computed to determine the modern climatic range in which *Sphagnum* occurs and is most abundant in North America. Mean annual temperature and total annual precipitation were extracted from *Leemans and Cramer* [1991] for the locations of 3005 core tops (age <100 yr B.P.) and surface sediment samples in the GPD. These variables were chosen because they are easily available and can be compared to known patterns of the modern climate. Spore percentages at sites with identical climate values were averaged since they would add no information to the response pattern. The response of *Sphagnum* percentages to the climate variables was estimated using kernel density interpolation [*Cressie*, 1995]. The quartic kernel function used Euclidean distance within a 100 × 100 unit grid defined by climate axes scaled to unity. Scaling was necessary because of the large differences in the ranges of the two environmental variables [*Gignac et al.*, 1991]. The scaled axes provided an equivalent resolution of 0.46°C by 19.7 mm, and the bandwidth was set at 3 times these values. This bandwidth and grid resolution limit extrapolation to 59.1 mm and 1.38°C. The resultant surface was smoothed using a 5 × 5 inverse-distance squared center-weighted moving average filter.

Local overrepresentation of *Sphagnum* percentages in bogs might affect the estimated climatic optimum of the response surface, but we found that the position of the optimum changed relatively little when surfaces were derived using only samples from lakes. The position of the abundance center is robust, but the absolute percentages of *Sphagnum* are elevated when bogs are included, which is of little concern since our purpose was to estimate the climatic range of the genus. Only the surface based on all samples is shown in this paper. In order to assess the degree to which our response surface represents the possible climates in North America where *Sphagnum* can occur, we compared plots of North American climate space with the subset of climates sampled by our sites used in the response surface. Finally, sampling intensity surfaces were constructed using all nonzero data in order to determine the extent to which sampling biases in climate space where *Sphagnum* spores occur could have affected the estimated *Sphagnum* response optimum.

The pollen databases are made up of contributions from many people over a long period of time, and we thus wanted to assess the reliability of our maps. Palynologists collect cores from both lakes and bogs. Although in the past, paleoecologists frequently cored bogs for pollen studies, the preference today is to use lake sediments. Spores of *Sphagnum* are present in bog sediments and vary from a few percent, even in *Sphagnum* peat, to significant amounts [e.g., *Gajewski*, 1987; *Hu and Davis*, 1995]. If we were to use bog samples only, the resultant maps would depend on the distribution of bogs that have been studied, which is not an extensive sample set. Bog stratigraphies are dominated by the local presence of spores in the wetland, but it is difficult to interpret the regional distribution of peatlands from these studies. However, spores are also present in lake sediments. These occasional spores do not necessarily indicate the presence of *Sphagnum* in the immediate area of the deposition site (local pollen and spores) because spores can be transported varying distances by the wind. However, as has been demonstrated with pollen maps [e.g., *Anderson et al.*, 1991], *Sphagnum* spores are more abundant in the region where the plants are found, with trace amounts found away from the source. Thus the distribution of *Sphagnum* spores from lake sediments can be used to indicate regions where *Sphagnum* peatlands are present, in the same way that the distribution of pollen of upland taxa can be used to indicate the distribution of trees. We plot the modern distribution of *Sphagnum* spore percentages to illustrate the clear relation between spores and peatland extent.

One potential problem became apparent in performing queries on these databases. Because *Sphagnum* is semi-aquatic, it is not certain that it is routinely recorded or subsequently entered into the database thus allowing some discrepancies in the distribution of this taxon. That is, while a positive value indicates that *Sphagnum* was deposited at that time within the sediment core, a zero value may not indicate absence of deposition but only that it was not entered into the database. Of course, a zero means only that *Sphagnum* spores were not recorded for the site and does not necessarily indicate the absence of the plant in the region. We therefore queried the database to determine the extent of this problem. We first identified all sites where a *Sphagnum* spore was counted at least once during the entire core sequence. For these sites, we then assumed that a zero at one level implied that no *Sphagnum* spores were observed for that time period (Figure 1). However, within a core where *Sphagnum* counts were zero for all levels, it is also possible that no spores were deposited at the site throughout the entire period of record, as might occur in the southwest United States. We therefore queried the database to determine those sites where *Sphagnum* was absent from the record but where the author has been known to enter *Sphagnum* at other sites. If an author counted and entered *Sphagnum* spores for other sites in the database, it is likely that one of their cores that have zero *Sphagnum* spores throughout actually signifies the absence of the spore at that site in question. The remaining sites are more

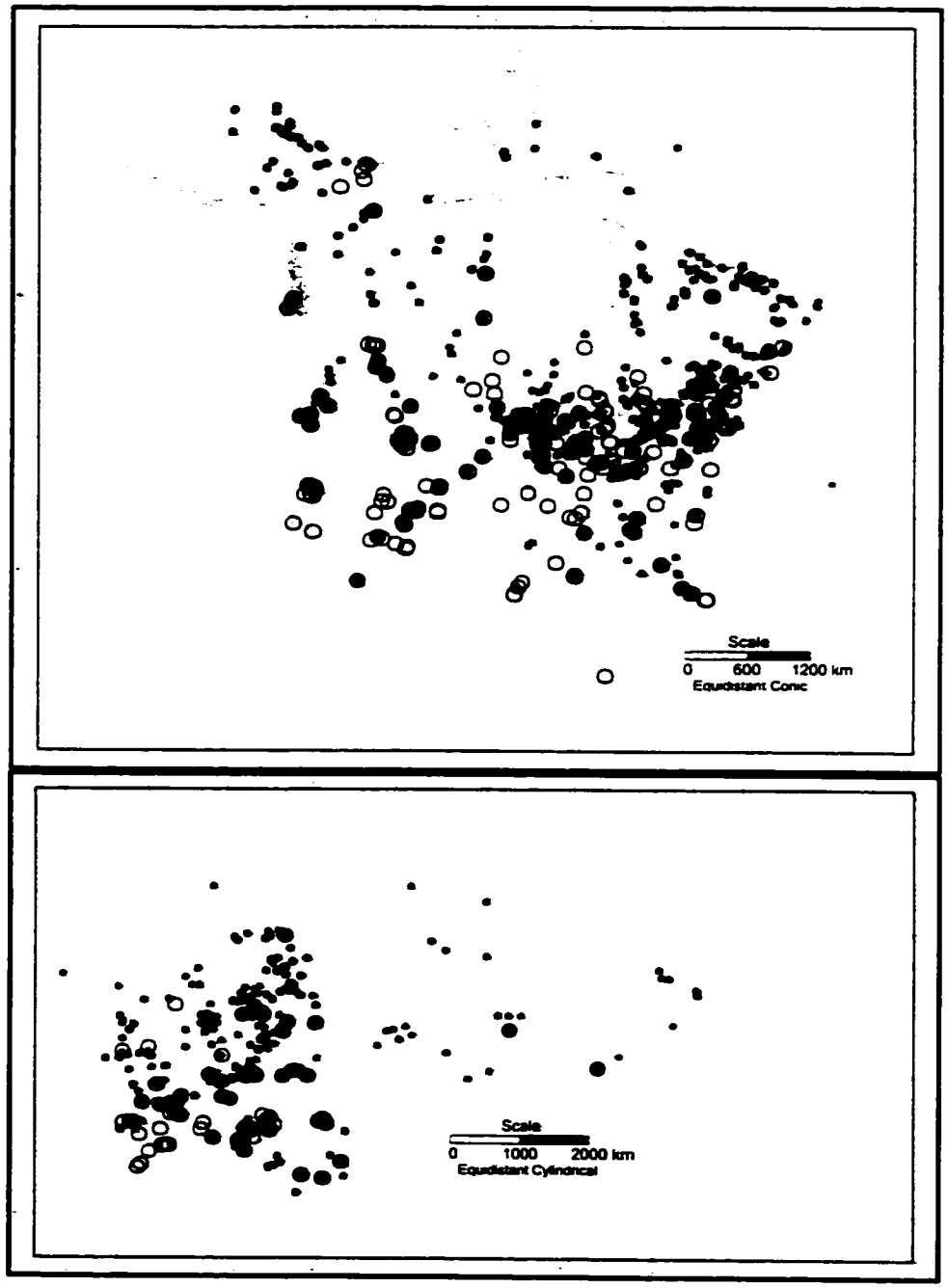


Figure 1. Location of core sites from the North American Pollen Database (NAPD) and European Pollen Database (EPD). Grey circles are sites where *Sphagnum* was recorded at least once during the period of the core. Open circles are sites where no *Sphagnum* spores were recorded. Open circles with a black dot are sites where *Sphagnum* is absent throughout the core but the author had recorded *Sphagnum* in at least one of their entries in the NAPD.

ambiguous, and a zero count may be interpreted as a lack of *Sphagnum* spores or simply that the data were not entered. Although we used all sites for the mapping, these categories can be used in the interpretation of the reliability of a particular zero *Sphagnum* value (Figure 1).

Estimates of area of peatland and total carbon (TC) stored in peatlands were made for every 2000-year time slice between 21 and 1 ka. To estimate the area of peatland extent in North America and Europe, the number of grid cells within the 0.5% *Sphagnum* isopoll was counted and multiplied by the grid resolution. Grid cells within the Canadian Arctic Archipelago were excluded from area calculations since most of these spores are probably transported from the boreal region to the south and percentages are inflated due to local underrepresentation of the tundra pollen [*Bourgeois*, 2000] (see below). The Arctic is too cold and dry to support extensive paludification, although small local peatlands have been recorded.

There are too few data from Asia to compute the area of peatlands for this continent directly. To arrive at a first approximation of the global peatland area for the different time periods, we assumed the total proportion of the world's peatland area in Asia has remained constant through time. We therefore computed the area of Asian peatlands as 0.765 times the sum of North American and European peatland area and added this to the areas we computed from our maps to arrive at the global total.

The carbon content of peatlands for time period t was estimated using $TC(t) = A_t \times D_t \times \rho \times C$, where A_t is the total area of peatland at time t (m^2), D_t is the peat thickness (m) at time t, ρ is the bulk density of the peat (g m^{-3}), and C is the carbon content of the peat (%). The land within the 0.5% isopoll consists of both upland and peatland vegetation. We used ranges of the modern ratio of peatland area to total area (0.1–0.2 for North America and 0.1–0.3 for Europe) [*Kivinen and Pakarinen*, 1981; *Vitt et al.*, 2000] to scale the total area within the 0.5% isopoll to A_t. Average modern peat thickness was estimated as 2.3 m in North America and 1.1 m in Europe [*Gorham*, 1991]. Values for mean bulk density (112×10^3 g m^{-3}) and carbon content (51.7%) were also taken from *Gorham* [1991] and were kept constant for all time periods.

Estimates of peat accumulation rates were needed to generate peat thickness values, D_t, for all time periods before the modern. We used an exponential curve ($be^{-(bt)}$) to model the accumulation rate through time (m yr^{-1}), where t = year (ka). This model reflects the low rates of accumulation in the full glacial and early Holocene and the increase in accumulation in the mid-Holocene [*Vitt et al.*, 2000]. The accumulation rate constant b was set to the modern value of 0.005 m yr^{-1} [*Gorham*, 1991]. The constant c, was set at 0.0002 for North America and 0.0004 for Europe in order to obtain carbon values that are negligible at full glacial and 455 Pg at present [*Gorham*, 1991]. Modifying the accumulation rate curve tested the sensitivity of total carbon estimates to the accumulation rate. Altering the rates did not significantly change the shape of the resulting carbon estimates unless a very high accumulation was introduced in the early Holocene, which is unreasonable.

There are very few data from which peatland extent in Asia may be accurately estimated. For this reason, carbon accumulation for Asia was roughly estimated using the ratio of its modern carbon total to the carbon total for North America and Europe (0.765).

3. Results

Crum [1984] reports *Sphagnum* plants from nearly every state of the United States and all of Canada. Although the northern range limit of many species coincides with tree line, several species are found in the low arctic, and others are found on Baffin Island. From the rest of the Arctic Islands, *Crum* [1986] reports only a few species from eastern Devon, southeast Ellesmere and southern Victoria Islands, although *Kuc* [1973] suggests a more widespread distribution.

The major area of peatland distribution in North America is in the boreal zone of Canada and Alaska (Figure 2) [*Hofstetter*, 1983; *Zoltai and Pollett*, 1983] and most sites across Alaska, Canada, and northeastern and northcentral United States have positive values of *Sphagnum* spores. The southern, modern-day distribution of positive values of *Sphagnum* spores corresponds closely to the distribution of peatlands (Figure 2) but not to the distribution of *Sphagnum* plants. The zone with maximum *Sphagnum* percentages is found in a band from north of the Great Lakes through to the Gaspé Peninsula. Although there are some sites in western Canada (e.g., Manitoba and Keewatin) with no *Sphagnum* spores reported, these come from older studies, and spores, although present, were probably not entered into the database. Not only are the percentages elevated in the area with maximum peatland extent, but also nearly all sites contain *Sphagnum* spores. In contrast, both the values of the *Sphagnum* spore percentages and the number of sites reporting spores decrease in the southeastern United States. Smaller bogs are located along the Atlantic Seaboard to Florida and at altitude in the Appalachians [*Hofstetter*, 1983; *Weider et al.*, 1981], and the scattered distribution of positive *Sphagnum* spore percentages reflects this less extensive presence of peatlands. Relatively high values of *Sphagnum* are also found in some arctic island samples. This is due to the low local pollen productivity that inflates the percentages of the *Sphagnum*, as there are not presently extensive peatlands in these regions. Much of this is presumably transported from the boreal peatlands to the south [*Bourgeois*, 2000].

The center of *Sphagnum* abundance ranges from ~630 to 1300 mm annual precipitation and between −2° and 6°C mean annual temperature (Figure 3a). The sites responsible for the *Sphagnum* percentage maximum are mainly from northern Ontario, meridional Québec, the Maritime Provinces, and the New England states (Figure 3c). Our *Sphagnum* response surface would suggest that peatlands can be found in areas of low annual temperature and precipitation in the Canadian Arctic Archipelago, but as discussed above, these percentages are inflated. Small clusters of positive values in areas with temperatures greater than 12°C and precipitation around 1200 mm (Figure 3a) represent samples from the southeastern United States (Figure 3c). In this region, the less extensive bogs are nevertheless recorded. Another cluster of high *Sphagnum* values is found at ~3°C and above 1440 mm, which represents a few sample sites in Labrador and Newfoundland (Figures 3a and 3c), areas with extensive bog development but few sample points. Although *Sphagnum* can occur along a wide range of precipitation and temperature, it is most abundant where precipitation is in excess of 500 mm yr^{-1} in cool to moderately cool temperatures [*Halsey et al.*, 1998].

The reliability of the response surface depends on how well our samples represent the possible climates in North America where *Sphagnum* can occur and whether limited sampling at those sites with nonzero *Sphagnum* percentages bias the response surface optimum. The convex hull defined on *Sphagnum* sample sites in climate space overlaps 55% of the total area of the convex hull defined by North American climate space (Figure 3d). The 45% not accounted for represents eastern coastal Alaska to Oregon, which is a geographically restricted area. Within the *Sphagnum* convex hull (Figure 3d), few samples occur where mean annual temperatures exceed 11°C and precipitation is less than 1000 mm y^{-1} which delimit climates of the American southwest. The aridity of this area prevents widespread bog development, and all of our samples have no recorded *Sphagnum* spores. We then used a

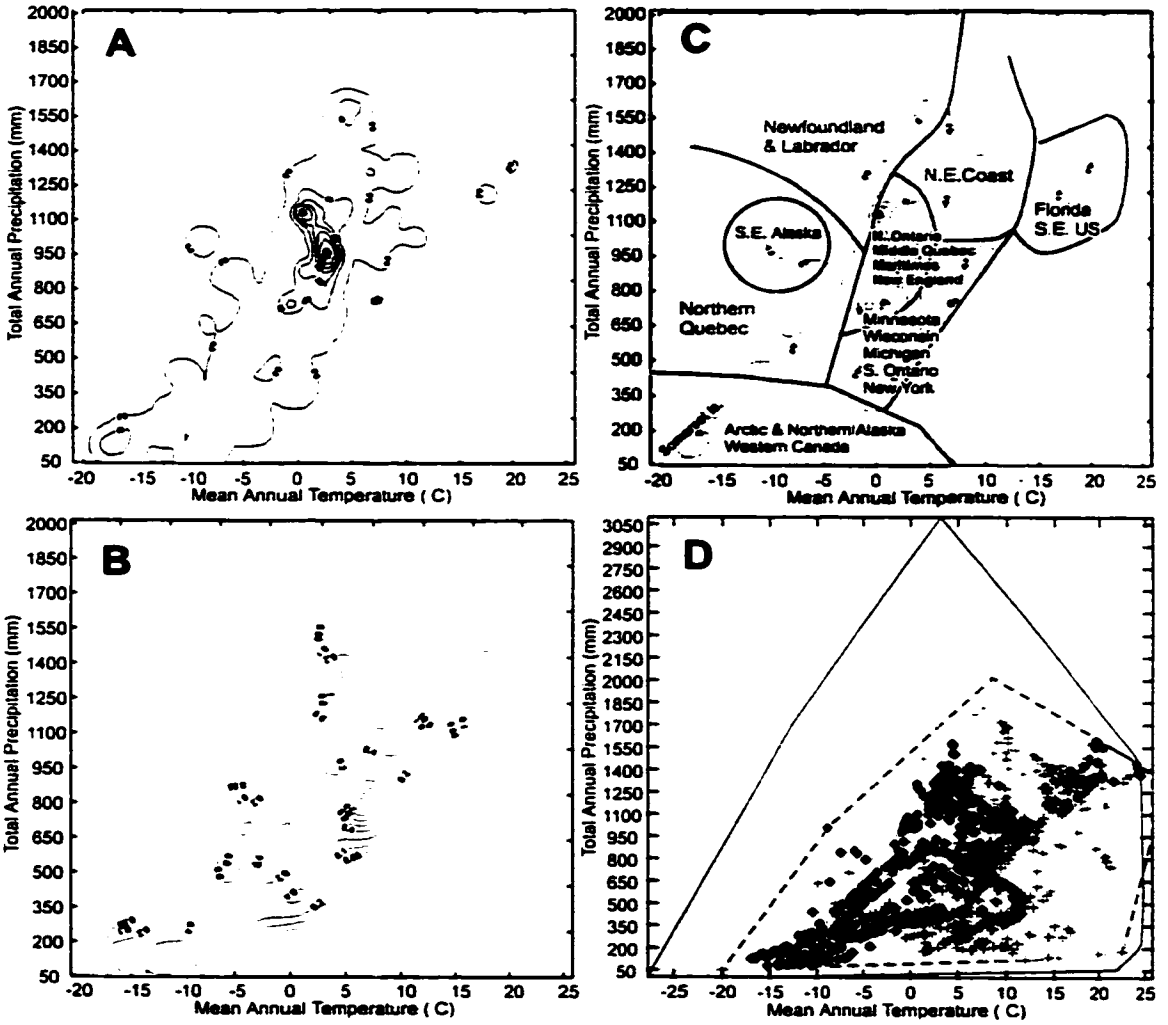


Figure 3. (a) Response surface of modern *Sphagnum* spore percentages plotted as a function of total annual precipitation and mean annual temperature. (b) Sampling intensity surface of sites with nonzero *Sphagnum* percentages. Contours represent number of sites per 0.46°C by 19.66 mm. (c) Geographic partitioning of *Sphagnum* spore distribution along climatic axes. (d) Degree of overlap between *Sphagnum* spore sampling and climate space. Solid outer line is the convex hull containing North American climate; the dotted line is the convex hull for *Sphagnum* sampling sites in North America. Small dots are individual grid points from *Leemans and Cramer* [1991] between 25° – 83° N and 49.5° – 168° W. Black diamonds are sediment sample locations with positive *Sphagnum* percentages, and small crosses are sites with zero *Sphagnum*.

Great Britain. By 11 ka, many sites record *Sphagnum* in eastern Europe, with occasional reports in western Europe. By 9 ka, there is extensive *Sphagnum* bog development across Europe, except for sites in the Mediterranean region. This pattern continues until the present.

There are few data available in the Global Pollen Database for Russia and Siberia, although much of central Siberia has extensive peatland development, in a band from around 50° to 65°N, with major peat basins in this zone from 60° to 90°E [*Kivinen and Pakarinen*, 1981; *Neustadt*, 1984]. The available cores do not extend to the full glacial. The few pollen diagrams available (Figure 6) suggest little peatland until around 10 ka and most peatland development after 8 ka [*Neustadt*, 1984].

Values for total area of *Sphagnum*-dominated peatlands and carbon stored in *Sphagnum* peatland areas remain low through the Early and Middle Holocene and rise rapidly in the Late Holocene (Figure 7). The Late Holocene increase is due to both

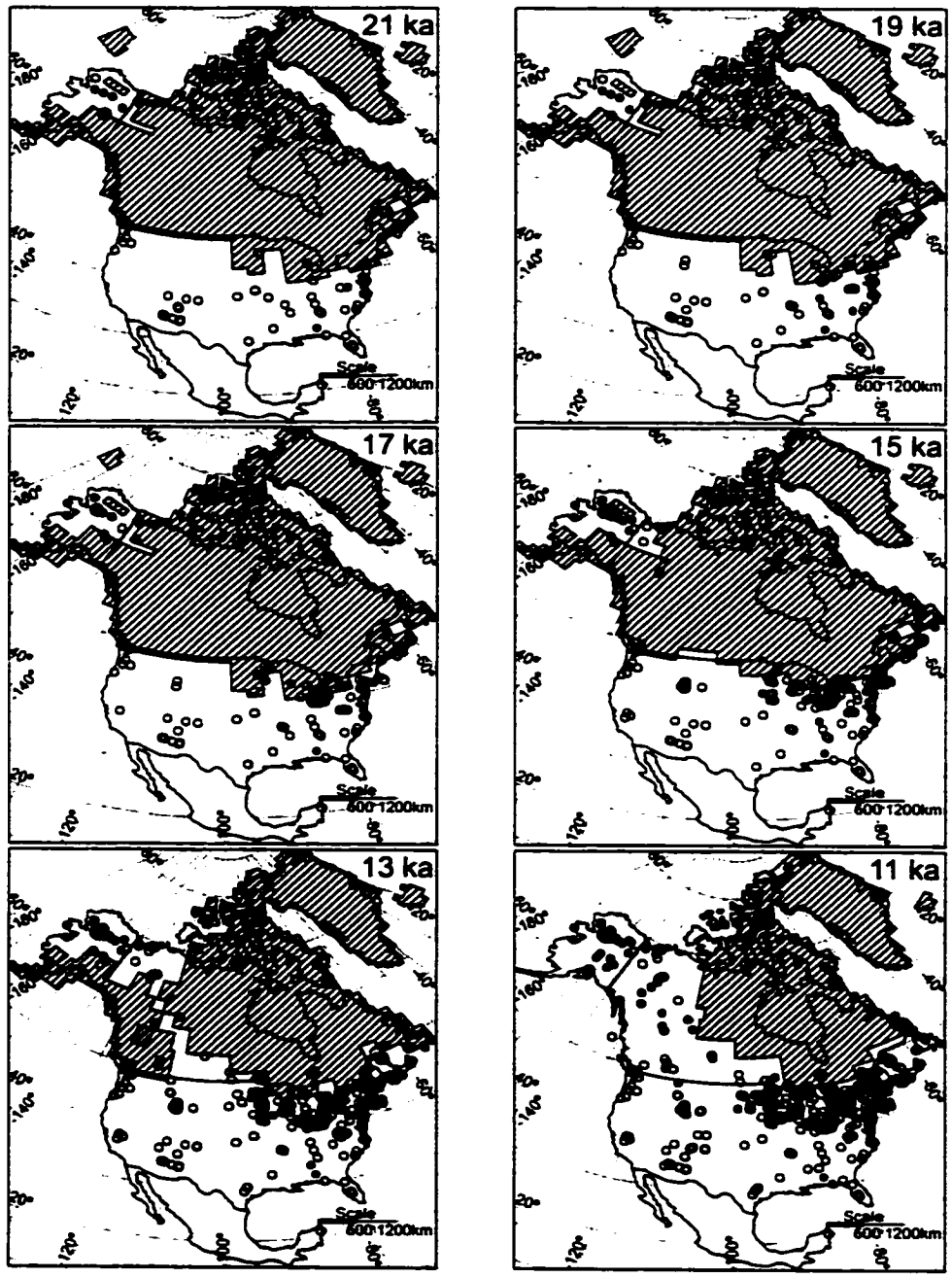


Figure 4. Map of *Sphagnum* distribution for the past 21,000 ka at 2000-year intervals for North America. Open circles represent zero *Sphagnum* percentages for that time period, small closed circles are positive values less than 5%, and large closed circles are values greater than 5%. Area within the 0.5% contour is shaded. Ice sheet location is hatched. Projection central meridian is 100°W (North America), 0°W (Europe); origin latitude 50°N; standard lines 35° and 80°N.

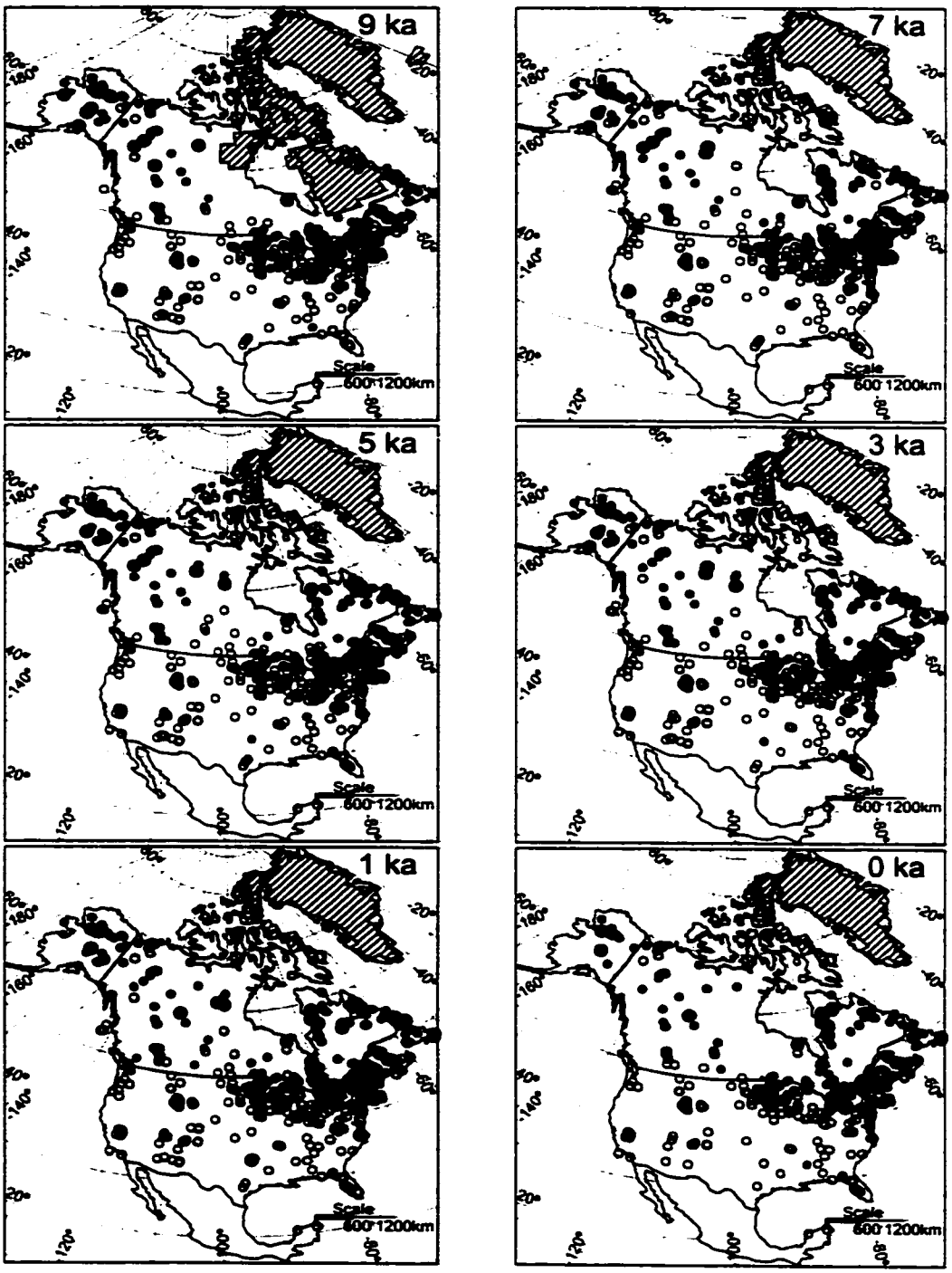


Figure 4. (continued)

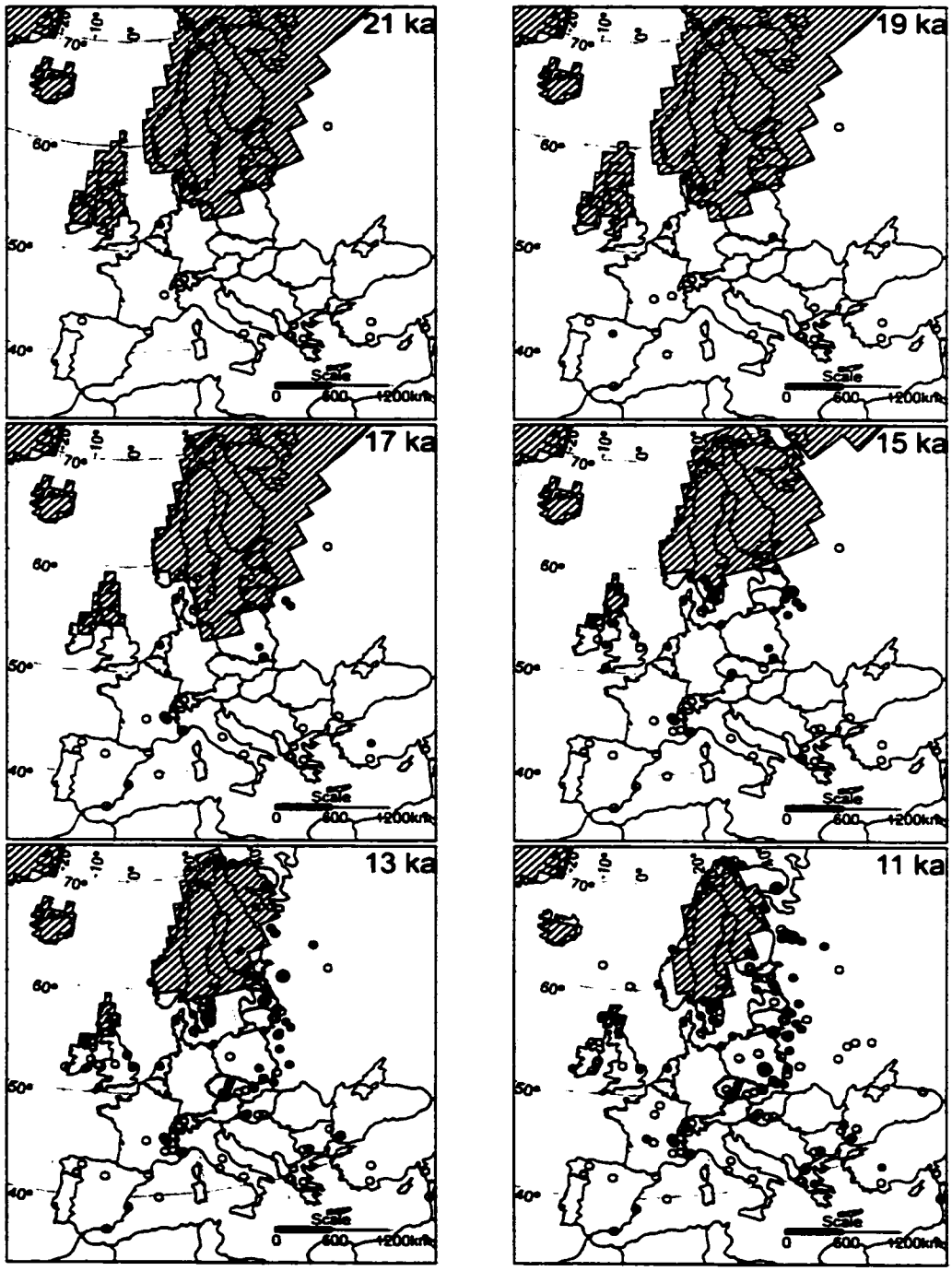


Figure 5. Map of *Sphagnum* distribution for the past 21,000 ka at 2000-year intervals for Europe. Symbols are the same as in Figure 4.

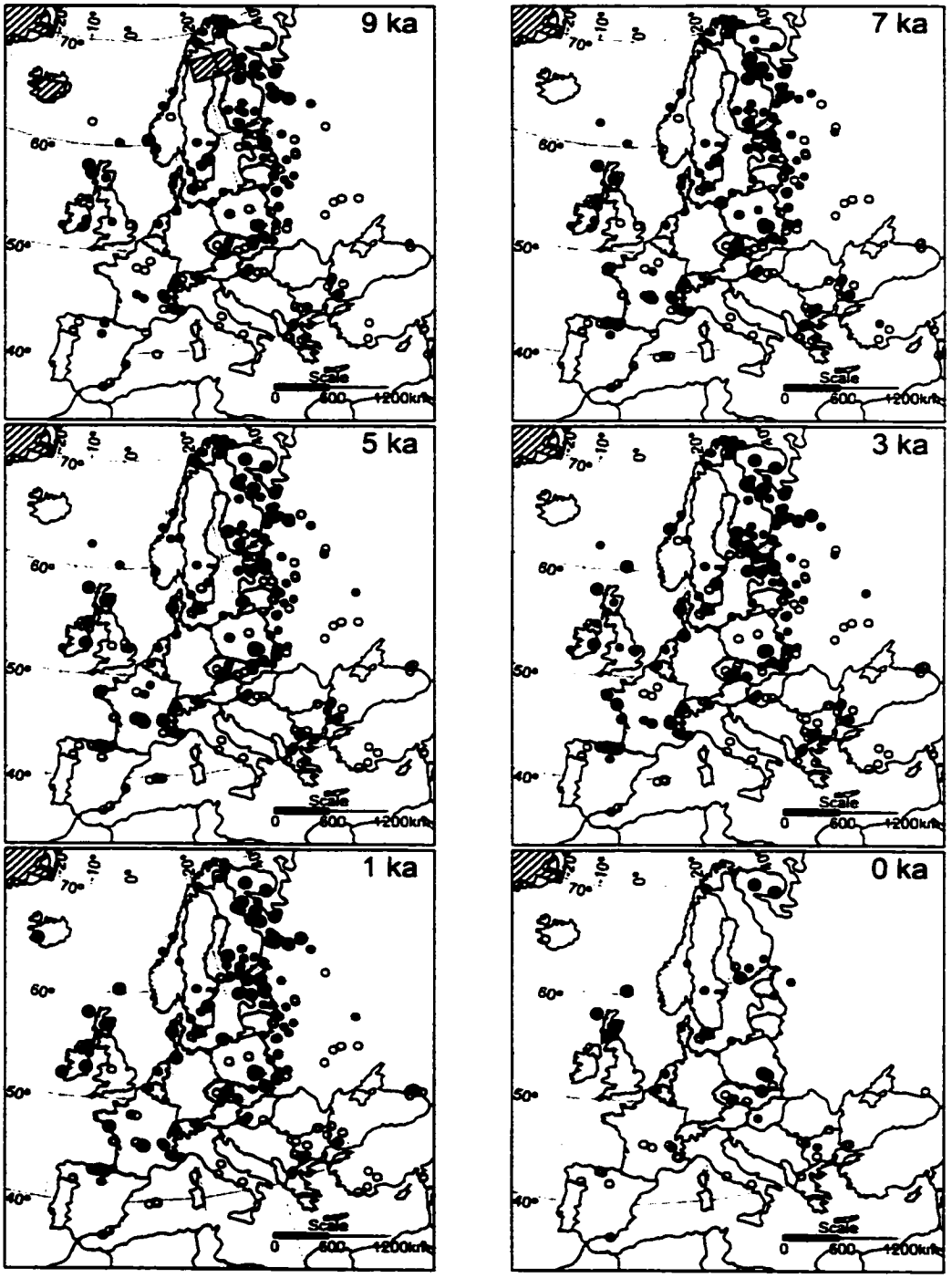


Figure 5. (continued)

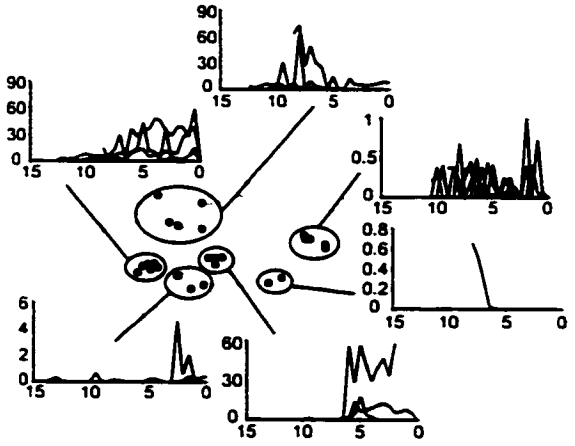


Figure 6. *Sphagnum* spore distribution through time in Siberia. The *x* axis of the inlaid graphs is time (years), and the *y* axis is *Sphagnum* spore percentage.

an increase in peatland area and to increasing mean peatland depth. Uncertainty in the ratio of peatland area to total area within the 0.5% isopoll was represented as an error bar on the results of our calculations. Calculations were constrained such that total peatland carbon matched the modern value of 455 Pg [*Gorham*, 1991].

4. Discussion

The response surfaces and maps of the modern distribution show that *Sphagnum* spores in sediments can be used to illustrate regions of peatland presence that are consistent with their modern geographic and climatic ranges (Figures 1 and 2). Accounting for long-distance transport, the area with abundant spores at most sites corresponds to the boreal region with extensive peatlands. Eastern United States, with smaller and less extensive peatlands, had smaller amounts of spores at some but not all sites and this reflects the smaller amounts of *Sphagnum* plants in less extensive bogs and moist areas. In the arid southwest, there are no *Sphagnum* spores reported in the modern data. The southern limit of the 0.5% isopoll corresponds to the southern limit of *Sphagnum* peatlands (Figure 3). The northern limit of *Sphagnum* peatlands on the landscape is not well captured by the 0.5% isopoll due to efficient long-distance transport of spores and low local production of pollen, which inflates the *Sphagnum* percentages. This effect is evident in all of the maps. Our response surface concords with previous work based on plant distribution [*Gignac et al.*, 1998; *Halsey et al.*, 1998].

The maps presented here illustrate the extent of peatlands during the past 21 ka. *Sphagnum* peatland extent was restricted during full glacial in North America and Europe. Although spruce and other boreal species were found far to the south of the Laurentide ice sheet, this vegetation was not a boreal forest in the sense that we consider it today. Today, much of the boreal zone is underlain by peatlands, in both North America and Eurasia [*Moore and Bellamy*, 1974; *Gorham*, 1991; *Neustadt*, 1984], but this has not always been the case. Models simulating conditions at full glacial should account for this difference in carbon storage. With the exception of Alaska, there is little

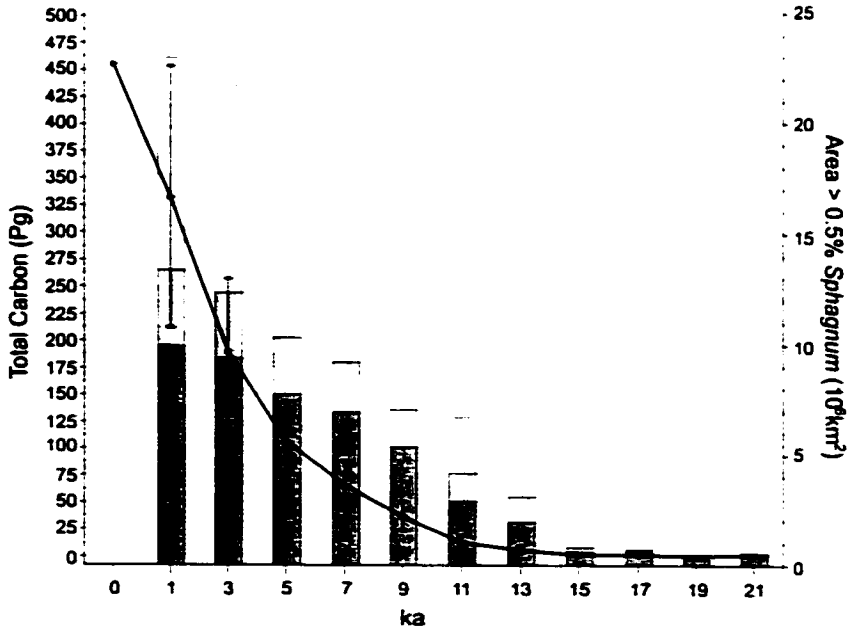


Figure 7. Changes in land area supporting *Sphagnum* peatland (bars) since the last glacial maximum. Light shading is North America, medium shading is Europe and dark shading is Asia. Carbon content in *Sphagnum* peatlands (lines); see text for explanation of calculation. Modern value from *Gorham* [1991].

evidence of extensive *Sphagnum* at this time. Little peat was accumulating in Europe. The interpretation of *Neustadt* [1984] only extends to the late glacial, and our results from Siberia cannot ascertain the extent of peatlands in Asia.

Sphagnum peatlands developed shortly after deglaciation (15–11 ka) in eastern North America. The area of peatland initiation and development moved eastward from Alaska and westward from the St. Lawrence region. This spread was related partly to the availability of land after deglaciation, but initial establishment at any site also depended on climate changes. Several studies have illustrated the importance of climate changes in initiating peat formation or in causing a transition from marsh to fen or fen to bog [*Vitt*, 1994; *Zoltai*, 1995; *Zoltai and Vitt*, 1990; *Winkler*, 1988; *Janssens et al.*, 1992; *Kuhry et al.*, 1993; *Vardy et al.*, 1998].

In this paper we interpret the closed *Sphagnum* contour as the area having *Sphagnum* peatlands present, that is, bogs and poor fens. Rich fens are dominated by other mosses, and their spores are not routinely identified in pollen preparations. However, the distinction between different wetland types in boreal zones depends on the relative influence of groundwater and rainwater (minerotrophic versus ombrotrophic) to the growing surface, and this is influenced by the wetland location in relation to the regional topography. The presence or absence of peat accumulation in peatlands is under the influence of the macroclimate, and this is what can be interpreted in this study. There may be a lag between initial peat formation and the establishment of *Sphagnum* in a particular peatland, and this time is longer in continental peatlands than in coastal ones [*Vitt et al.*, 1994]. Our estimates of peatland extent are thus conservative. However, if poor fens were accumulating peat during the full glacial, there should have been some trace of this left in the form of peat deposits, but there is little evidence of this accumulation, and our estimates should not have excessive error.

These reconstructions agree broadly with previous local and regional studies. *Halsey et al.* [1998] show calibrated radiocarbon dates from 90 locations across continental western Canada that indicate peat formation began ~8000–9000 years BP in nucleation zones along the upper elevations of the Montane region of Alberta and in northern Alberta. From 6 to 8 ka, peat formation expanded eastward into Manitoba, and after 6 ka the trend of southeasterly peatland expansion continued. Our results put this into a large-scale context and illustrate that the peatland development in this region is part of the larger-scale movement from west to east from the full glacial onward. In the western Great Lakes region, the Red Lake Peatlands [*Heinselman*, 1970; *Griffin*, 1977; *Janssens et al.*, 1992] and smaller bogs [*Winkler*, 1988] developed later, as the climate became cooler and moister in the middle to Late Holocene. The maps of *Halsey et al.* [2000], based on presence or absence of spores at a selected subset of North American sites, also agree broadly with our results.

Huntley and Birks [1983] presented maps of *Sphagnum* spore percentages at 9000, 4000, and 2000 ^{14}C years BP for Europe. Since they only mapped isopolls greater than 10%, they were recording the local presence of abundant *Sphagnum* or sample sites that happened to be bogs or peatlands. Our area of peatland extent is therefore much greater than these earlier results. In the British Isles, characteristic peat basal dates range between 8500 and 7500 years BP in Scotland to 6000 years BP in Wales and Ireland [*Taylor*, 1983]. There was accelerated peat development in blanket bogs after 2800 years BP. In Sweden, paludification occurred extensively in the Early and Middle Holocene and had almost ceased by 2000 years BP [*Sjors*, 1983].

Interestingly, there is some suggestion from the literature that *Sphagnum* reproduces essentially vegetatively and that capsules are rarely observed [*Andrus*, 1986; *Clymo and Duckett*, 1986; *Cronberg*, 1991]. The presence of numerous spores in lake sediments occasionally transported far from the source [*Bourgeois*, 2000] suggests that considerable numbers of spores are being produced and that these are well transported in the atmosphere.

This work provides at least two kinds of information. First, since the distribution of *Sphagnum* as well as peatlands is related to climate (Figure 3) [*Moore and Bellamy*, 1974; *Vitt et al.*, 1994], these maps indicate that portion of the Northern Hemisphere that supports *Sphagnum* growth and how this region has changed through time. Because of the transport of *Sphagnum* spores beyond the immediate source, the limits indicated here are more extensive than the actual limit of the plant. However, at the scale and resolution of global climate models (GCMs), these maps could be useful for paleoclimate verification.

Studies of CO_2 variations through time need to account for major sources and sinks of carbon, and there are great uncertainties in these estimates. Previous work has been forced to ignore or estimate carbon storage in peatlands due to lack of data on peatland extent [*Crowley*, 1995]. Our work is a contribution to the goal of better understanding carbon dynamics by presenting estimates of *Sphagnum* peatland extent through time. When rough estimates of the carbon stored in peatlands are made, the results indicate three periods. Before ~11 ka, there was little carbon stored in *Sphagnum* peatlands. In the Early Holocene, carbon gradually accumulated as the area of land supporting peatlands increased. In the Late Holocene, the carbon stored in these ecosystems increased more rapidly as the depth of the peatlands increased.

These maps only indicate where there was bog and/or poor fen development but not other wetland classes, including rich fen. In some regions, there is a delay between initiation of wetlands and the transition to *Sphagnum* bog, and peat can be accumulating during this period. For more precise estimates of carbon storage in peatlands, regional-scale models of peat growth [e.g., *Gignac et al.*, 1998] could be combined with reconstructions or simulations of past climates and tested against these maps.

The reduced abundance of *Sphagnum* peatlands through much of the past 21 ka illustrates significant ecosystem-level changes in boreal and temperate biomes. The hydrological balance and carbon storage were significantly different from the present, as illustrated by a relative lack of peatlands through much of the past 21 ka. These maps indicate changes in the major biomes through time, and the importance of using functional-type bioclimatic models [e.g., *Prentice et al.*, 1998; *Williams et al.*, 1998] in preference to static biomes [e.g., *Prentice and Fung*, 1990; *Adams et al.*, 1990] when reconstructing vegetation changes through time.

5. Summary

Maps based on *Sphagnum* spores in lake and bog sediments can be used to reconstruct changes in peatland extent in boreal and temperate biomes. The present-day distribution of abundant *Sphagnum* spores corresponds closely to areas with peatland development. In North America, maximum peatland development is in the boreal zone and, to a lesser extent, southward along the eastern coastal region. In Europe, the most extensive peatland development is found to the north of the Alps, and in Asia, central Siberia has extensive peatland development. Today maximum *Sphagnum* spore abundance is found between 630 and 1300 mm annual precipitation and between −2° and 6°C mean annual temperature in North America. During the Wisconsin glaciation, there were apparently not large amounts of carbon stored in peatlands. Peatlands were growing in Alaska and began developing in eastern North America during deglaciation. Major peatland development occurred in North America after 9 ka, and there was southward movement of high *Sphagnum* spore abundance after 5 ka in the western Great Lakes region. The major period of peatland development began after 9 ka in Europe and Asia.

Acknowledgments. This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Science Grant and by the Climate System History and Dynamics Project (CSHD) jointly funded by NSERC and the Atmospheric Environment Service. Further support came from NSERC Postgraduate Scholarships to M. Sawada and A. Viau, and an NSERC Postdoctoral Fellowship to S. Wilson. Data from the North American Pollen Database (NAPD), European Pollen Database (EPD), and Global Pollen database (GPD) were obtained from the U.S. National Oceanographic and Atmospheric Administration (NOAA) Paleoclimatology Data Center, and we acknowledge the use of these data, as well as the contributors to this database.

References

Adams, J., H. Faure, L. Faure-Denard, J. McGlade, and F. Woodward, *Increases in terrestrial carbon storage from the last glacial maximum to present*, Nature, *348*, 711–714, 1990.

Anderson, P. M., P. J. Bartlein, L. B. Brubaker, K. Gajewski, and J. C. Ritchie, *Vegetation-pollen-climate relationships for the arcto-boreal region of North America and Greenland*, J. Biogeogr., *18*, 565–582, 1991.

Andrus, R., *Some aspects of Sphagnum ecology*, Can. J. Botany, *64*, 416–426, 1986.

Bartlein, P. J., I. C. Prentice, and T. Webb, III *Climatic response surfaces from pollen data for some eastern North American taxa*, J. Biogeogr., *13*, 35–57, 1986.

Bourgeois, J., *Modern and holocene pollen assemblages from arctic ice caps*, Ph.D. thesis, Dep. of Geogr., Univ. of Ottawa, 2000.

Cheddadi, R., José M. (Ed.), *European Pollen Database*, www.ngdc.noaa.gov/paleo, World Data Cent. A for Paleoclimatology, , 2000.

Clymo, R. S., and J. G. Duckett, *Regeneration of Sphagnum*, New Phytol., *102*, 589–614, 1986.

Clymo, R. S., and P. M. Hayward, *The ecology of sphagnum*, in *Bryophyte Ecology*, edited by A. J. E. Smith., José M. pp. 229–289, Chapman and Hall, 1982.

Cressie, N., *Statistics for Spatial Data*, , John Wiley, 1995.

Cronberg, N., *Reproductive biology of Sphagnum*, Lindbergia, *17*, 69–82, 1991.

Crowley, T., *Ice age terrestrial carbon changes revisited*, Global Biogeochem. Cycles, *9*, 377–390, 1995.

Crum, H., *Sphagnopsida: Sphagnaceae, North American Flora Ser II*, Part 11, New York Bot. Garden, 1984.

Crum, H., *Illustrated moss flora of Arctic North America and Greenland, 2. Sphagnaceae*, Medd. Gronl. Biosci., *18*, 1–61, 1986.

Gajewski, K., *Environmental history of Caribou Bog, Penobscot Co.*, Maine, Nat. Can., *114*, 133–140, 1987.

Gignac, L. D., and D. H. Vitt, *Habitat limitations of Sphagnum along climatic, chemical and physical gradients in mires of western Canada*, Bryologist, *93*, 7–22, 1990.

Gignac, L. D., and D. H. Vitt, *Responses of northern peatlands to climate change Effects on bryophytes*, J. Hattori Bot. Lab., *75*, 119–132, 1994.

Gignac, L. D., D. H. Vitt, , and S. E. Bayley, *Bryophyte response surfaces along ecological and climatic gradients*, Vegetatio, *93*, 29–45, 1991.

Gignac, L., D. B. Nicholson, and S. E. Bayley, *The utilization of Bryophytes in bioclimatic modeling: Present distribution of peatlands in the Mackenzie River Basin, Canada*, Bryologist, *101*, 560–571, 1998.

Gorham, E., *Northern peatlands: Role in the carbon cycle and probable responses to climatic warming*, Ecol. Appl., *1*, 182–195, 1991.

Griffin, K., *Paleoecological aspects of the Red Lake Peatland, northern Minnesota*, Can. J. Bot., *55*, 172–192, 1977.

Grimm, E., José M. (Ed.), *Global Pollen Database*, www.ngdc.noaa.gov/paleo, World Data Cent. A for Paleoclimatology, 2000a.

Grimm, E. (Ed.), *North American Pollen Database*, www.ngdc.noaa.gov/paleo, World Data Cent. A for Paleoclimatology, 2000b.

Halsey, L., D. H. Vitt, and S. Zoltai, *Climatic and physiographic controls on wetland type and distribution in Manitoba, Canada*, Wetlands, *17*, 243–262, 1997.

Halsey, L., D. H. Vitt, and I. E. Bauer, *Peatland Initiation during the Holocene in continental western Canada*, Clim. Change, *40*, 315–342, 1998.

Halsey, L., D. H. Vitt, and L. D. Gignac, *Sphagnum-dominated peatlands in North America since the last glacial maximumTheir occurrence and extent*, Bryologist, *103*, 334–352, 2000.

Heinselman, M., *Landscape evolution, peatland types, and the environment in the Lake Agassiz peatlands natural area, Minnesota*, Ecol. Monogr., *40*, 235–261, 1970.

Hofstetter, R., *Wetlands in the United States*, in *MiresSwamp, Bog, Fen and Moor, Regional Studies, Ecosystems of the World 4B*, edited by A. J. Gore., pp. 201–244,. Elsevier Sci., 1983.

Hu, F. S., and R. B. Davis, *Postglacial development of a Maine bog and paleoenvironmental implications*, Can. J. Bot., *73*, 638–649, 1995.

Huntley, B., and H. J. B. Birks, *An Atlas of Past and Present Pollen Maps for Europe: 0–13 000 Years Ago*, Cambridge Univ. Press, 1983.

Janssens, J., B. Hansen, P. Glaser, and C. Whitlock, *Development of a raised-bog complex*, in *The Patterned Peatlands of Minnesota*, edited by H.E. Wright, Jr., B. Coffin, and N. Aaseng., Univ. of Minn. Press, 1992.

Kivinen, E., and P. Pakarinen, *Geographical distribution of peat resources and major peatland complex types in the world*, Ann. Acad. Sci. Fenn. Ser. A3, *132*, 1–28, 1981.

Kuc, M., *Bryogeography of Expedition area, Axel Heiberg Island, N.W.T. Canada*, Bryophytorum Bibl., *2*, 120 pp., 1973.

Kuhry, P., , and D. Vitt, *Fossil carbon/nitrogen ratios as a measure of peat decomposition*, Ecology, *77*, 271–275, 1996.

Kuhry, P., B. J. Nicholson, L. D. Gignac, D. Vitt, and S. E. Bayley, *Development of Sphagnum-dominated peatlands in boreal continental Canada*, Can. J. Bot., , *71*, 10–22, 1993.

Lam, N., *Spatial interpolation methodsA review*, Am. Cartographer, *10*, 129–149, 1983.

Leemans, R., and W. Cramer, *The IIASA database for mean monthly values of temperature, precipitation and cloudiness of a global terrestrial grid, RR-91-18*, Int. Inst. for Appl. Syst. Anal. (IIASA), 1991. (Available at ftp.pik-potsdam.de)

Moore, P., and D. Bellamy, *Peatlands*, Springer-Verlag, 1974.

Neustadt, M., *Holocene peatland development, in Late Quaternary Environments of the Soviet Union*, edited by A. Velichko, pp. 201–206, Univ. of Minn. Press, 1984.

Peltier, W., *Time Dependent Topography Through Glacial Cycle, ftp.ngdc.noaa.gov/paleo/pmip/peltier_ice/, Data Contrib. Ser. 93-015*, World Data Cent. A for Paleoclimatology, 1993.

Peltier, W., *Ice Age paleotopography*, Science, *265*, 195–201, 1994.

Peng, C., J. Guiot, E. van Campo, and R. Cheddadi, *The vegetation carbon storage variation in Europe since 6000 BP: Reconstruction from pollen*, J. Biogeogr., *21*, 19–31, 1994.

Peng, C., J. Guiot, and E. van Campo, *Estimating changes in terrestrial vegetation and carbon storage. Using palaeoecological data and models*, Quat. Sci. Rev., *17*, 719–736, 1998.

Prentice, I., S. Harrison, D. Jolly, , and J. Guiot, *The climate and biomes of Europe at 6000 yr B.P.: Comparison of model simulations and pollen-based reconstructions*, Quat. Sci. Rev., *17*, 559–668, 1998.

Prentice, K., and I. Fung, *The sensitivity of terrestrial carbon storage to climate change*, Nature, *346*, 48–51, 1990.

Sjors, H., *Mires of Sweden*, in *Mires Swamp, Bog, Fen and Moor, Regional Studies, Ecosystems of the World 4B*, edited by A. J. Gore, pp. 69–94, Elsevier New york, 1983.

Snyder, J., *Map Projections: A Working Manual, U.S. Geol. Surv. Prof. Pap 1395*, 1987.

Stuiver, M., and P. Reimer, *Extended ^{14}C database and revised CALIB radiocarbon calibration program*, Radiocarbon, *35*, 215–230, 1993.

Taylor, J. A., *Peatlands of Great Britain and Ireland*, in *Mires: Swamp, Bog, Fen and Moor, Regional Studies, Ecosystems of the World 4B*, edited by A. J. Gore, pp. 1–46, . Elsevier, New York, 1983.

van Breemen, N., *How Sphagnum bogs down other plants*, Trends Ecol. Evol., *10*, 270–275, 1995.

Vardy, S., B. Warner, and R. Aravena, *Holocene climate and the development of a subarctic peatland near Inuvik, Northwest Territories, Canada*, Clim. Change, *40*, 285–313, 1998.

Vitt, D. H., *An overview of factors that influence the development of Canadian peatlands*, Mem. Entomological Soc. Can., *169*, 7–20, 1994.

Vitt, D. H., L. Halsey, and S. C. Zoltai, *The bog landforms of continental western Canada in relation to climate and permafrost patterns*, Arct. Alp. Res., *26*, 1–13, 1994.

Vitt, D. H., L. Halsey, I. Bauer, and C. Campbell, *Spatial and temporal trends in carbon storage of peatlands of continental western Canada through the Holocene*, Can. J. Earth Sci., *37*, 683–693, 2000.

Webb, T., III, K. Anderson, P. Bartlein, and R. Webb, *Late Quaternary climate change in eastern North AmericaA comparison of pollen derived estimates with climate model results*, Quat. Sci. Rev., *17*, 587–606, 1998.

Weider, R., A. McCormick, and G. Lang, *Vegetational analysis of Big Run Bog, a nonglaciated Sphagnum bog in West Virginia*, Castanea, *46*, 16–29, 1981.

Williams, J., R. Summers, and T. Webb, III*Applying plant functional types to construct biome maps from eastern North America pollen data: Comparisons with model results*, Quat. Sci. Rev., *17*, 607–628, 1998.

Winkler, M., *Effect of climate on development of two Sphagnum bogs in south-central Wisconsin*, Ecology, 69, 1032–1043, 1988.

Zoltai, A., and F. Pollett, *Wetlands in Canada: Their classification, distribution, and use*, in *Mires: Swamp, Bog, Fen and Moor, Regional Studies, Ecosystems of the World 4B*, edited by A. J. Gore, pp. 245–268, Elsevier, New York, 1983.

Zoltai, S., *Regional variations in peatland ecosystems of west-central Canada through time*, Gunneria, 70, 35–42, 1995.

Zoltai, S., and D. Vitt, *Holocene climate change and the distribution of peatlands in western interior Canada*, Quat. Res., 33, 231–240, 1990.

D. Atkinson, K. Gajewski, M. Sawada, S. Wilson, and A. Viau, Laboratory of Paleoclimatology and Climatology, Department of Geography, University of Ottawa, 165 Waller Street, Ottawa, Ontario, K1N 6N5 Canada. (gajewski@aix1.uottawa.ca)

(Received April 11, 2000; revised August 15, 2000; accepted August 24, 2000.)