Université d'Ottawa • University of Ottawa

PERMISSION DE REPRODUIRE ET DE DISTRIBUER LA THÈSE

PERMISSION TO REPRODUCE AND DISTRIBUTE THE THESIS

NOM DE L'AUTEUR / NAME OF AUTHOR:	André Ernest VIAU
ADRESSE POSTALE / MAILING ADDRESS:	55 - 400, rue Bédard Hull (Québec) J8Y 6A2
GRADE / DEGREE: Ph.D.(Geography)	**ANNÉE D'OBTENTION / YEAR GRANTED** 2003
TITRE DE LA THÈSE / TITLE OF THESIS: Millenial-scale Climate Variability in North America during the Past 14,000 Years	

3/1/03

DATE

(AUTEUR) SIGNATURE (AUTHOR)

Université d'Ottawa • University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

VIAU, André Ernest J.
AUTEUR DE LA THÈSE - AUTHOR OF THESIS

Ph.D. (Geography)
GRADE - DEGREE

Geography
FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Millennial-Scale Climate Variability in North America
During the Past 14,000 Years

Konrad Gajewski
DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

P. Johnson

D. Lagarec

B. Lauriol

P. Richard

J.-M. De Koninck, Ph.D.
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

Millennial-scale climate variability in North America during the past 14,000 years

André Ernest Viau

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

Examining Board: **Dr. Peter G. Johnson** (Professor of Geography, *University of Ottawa*), **Dr. Daniel Lagarec** (Professor of Geography and Chaiman of the Department of Geography, *University of Ottawa*), **Dr. Bernard Lauriol** (Professor of Geography, *University of Ottawa*), and **Dr. Konrad Gajewski** (Supervisor, Professor of Geography, *University of Ottawa*), External: **Dr. Pierre J.H. Richard** (Professor of Geography, *Université de Montréal*).

0-612-76506-7

MILLENNIAL-SCALE CLIMATE VARIABILITY IN NORTH AMERICA DURING THE PAST 14,000 YEARS

Abstract

Variations in the Earth's climate occur on many time and space scales. A recent focus of paleoclimate research is the so-called 1500-year North Atlantic quasi-periodic cycle, and has revolved around three main themes. First, what are the underlying causes and physical mechanisms governing these millennial-scale variations? Next, are they global or restricted to certain sensitive regions of the planet? Last, what is the magnitude of the temperature changes of these variations, and do they vary in time and space?

This dissertation explores millennial-scale climate variability in North America during the past 14,000 years using a dense network of fossil pollen data, which is used as proxy for climate variations. Three independent approaches are used to quantify these changes.

A mixture modelling analysis of radiocarbon dates on pollen transitions, a principal component analysis of pollen diagrams from all of North America, and a mean July temperature reconstruction based on the method of modern analogue (MAT) all reveal millennial-scale climate variability throughout North America during the past 14,000 years. The identified transitions generally correlate well with other proxy-climate records from the North Atlantic region. However, certain mismatches occurred particularly at 9, 6 and 4 ka BP. If we assume the dominant millennial-scale period is 1150-years, the records become more consistent. North American temperature variability was not unidirectional nor uniformly distributed in space, suggesting large-scale ocean-atmospheric reorganizations at the transitions. Correlation between the proxy-climate and cosmogenic nuclide records supports a variable solar output hypothesis as the fundamental cause for century to millennial-scale climate variability.

The mean July temperature of North America varied on the order of 0.2 to 0.4°C during the Holocene and 0.4° and 0.6°C during the deglaciation. Temperature was more variable during the late glacial, possibly due to the impact on the climate of massive meltwater pulses into the North Atlantic, further amplified through ocean dynamic processes. Recent global warming estimated as an increase in temperature of 0.4-0.6°C, is therefore greater than the estimated natural variation of the past 10,000 years, providing further evidence that recent increases in atmospheric CO_2 have played a major role in modern warming.

VARIATIONS CLIMATIQUES À L'ÉCHELLE MILLÉNAIRE EN AMÉRIQUE DU NORD DURANT LES DERNIERS 14,000 ANS.

Résumé

Les variations climatiques se manifestent à différentes échelles spatiales et temporelles. Une attention particulière, durant la dernière décennie, a été mise sur un cycle de 1,500-ans identifié dans les carottes marines et glaciaires de la région Nord-Atlantique. Le débat central, sur ces variations naturelles climatiques, touche trois principaux thèmes. Le premier thème concerne les causes et mécanismes physiques fondamentaux gouvernant ces variations. Le deuxième thème porte sur la répartition spatiale de cette variabilité. Enfin, il s'agit de l'amplitude de ces changements de température.

Cette dissertation explore les variations climatiques, à l'échelle millénaire, en Amérique du Nord durant les derniers 14,000 ans. Un vaste réseau de données polliniques est utilisé comme indicateur indirect pour la reconstruction climatique. Trois approches indépendantes ont été utilisées pour cette analyse, dont une analyse des distributions des datations de radiocarbone, une analyse en composantes principales des diagrammes polliniques, et une reconstruction des températures moyennes pour le mois de juillet basée sur la méthode des analogues (MAT).

Ces trois approches ont toutes montré une variabilité à l'échelle millénaire du climat de l'Amérique du Nord, durant les derniers 14,000 ans. Les transitions identifiées ont pu être corrélées avec d'autres données pour la région Nord-Atlantique. Toutefois, certaines discordances existent entre les différentes séquences, soit à 9, 6, et 4 milles ans passés. Une périodicité de 1150 ans pourrait réconcilier cette inconsistance. La variabilité spatiale des températures de l'Amérique du Nord n'était ni unidirectionnelle, ni uniformément distribuée dans l'espace. Ceci suggère qu'il y eut des réorganisations de la circulation océanique et atmosphérique à ces transitions. La corrélation entre les données paléoclimatiques et les données cosmogéniques soutient l'hypothèse que la cause fondamentale des variations climatiques aux échelles centenaire et millénaire est la variabilité solaire.

À l'échelle continentale de l'Amérique du Nord, la température moyenne pour le mois de juillet aurait varié de 0.2 à 0.4 °C durant l'Holocène, et de 0.4 à 0.6 °C durant la période de déglaciation. Les résultats montrent que l'amplitude de la variation de la température a augmenté lors de la dernière déglaciation. De fortes pulsations hydriques, affluant des eaux froides dans le Nord de l'Atlantique, amplifiées par des processus océaniques pourraient être à la source de cette variabilité. Le réchauffement planétaire moderne, associé à une augmentation des températures de l'ordre de 04° à 0.6°C, est donc plus considérable que les variations naturelles estimées pour les derniers 10,000 ans. Cette analyse souligne davantage que l'augmentation récente du CO_2 atmosphérique est la cause principale du réchauffement global durant le dernier siècle.

Avec la plus grande affection et un respect incontournable, cet ouvrage est dédié a papa et maman,

Ernest Viau et Fleurette Charbonneau en l'honneur de leur 45^{ième} anniversaire de mariage.

Félicitations, je vous aime pour toujours.

PREFACE

Three chapters of this dissertation were written for publication in peer-reviewed scientific journals, although they are expanded here and contain details and explanations that were not published or that are to be submitted. I am the primary author on all of the chapters although they involved collaborations.

CHAPTER 2 was published in *Geology* (Viau *et al.*, 2002). The solution to the mixture problem is difficult mathematically and there was no standard methodology to find a solution. Dr. Philippe Finès and Dr. D.E. Atkinson provided the necessary statistical help. Dr. M.C. Sawada provided technical support and participated in discussions.

CHAPTER 3 also contains contributions from Dr. Philippe Finès who provided valuable ideas on the implementation and testing of a complex analysis of pollen assemblages through time, space and taxonomic realms. Dr. D.E Atkinson's programming abilities made it possible to automate the testing of the robustness of the model.

CHAPTER 4 uses data generously provided by Dr. Thompson Webb III and Dr. John Williams, and Dr. P.J. Bartlein. Dr. M.C. Sawada provided the programs to automate the MAT computations. Dr. Philippe Finès's involvement in the later stages of this analysis again provided the comments on statistical methodology.

ACKNOWLEDGEMENTS

I would like to thank those whose support was fundamental to the success of this doctoral dissertation. Working in an environment that promotes professionalism can only lead to success. The Laboratory for Paleoclimatology and Climatology (LPC) of the University of Ottawa represents such an environment. Dr. Konrad Gajewski, my supervisor and founder of the LPC, has fostered several new researchers in paleoclimate and paleocological fields. His acute scientific knowledge and skills aided me in developing analytical and methodological skills needed to become a competent researcher. My thesis committee members: Dr. Daniel Lagarec for his numerous and precious discussions that helped me apply rigour and critique to my research, an indispensable skill for scientific advancement; Dr. Bernard Lauriol for his helpful discussions, teachings, references and support; Dr. Peter G. Johnson for his support and undeniable expertise in the Yukon territory; and Dr. Pierre J.H. Richard, my external examiner, for his unquestionable scientific achievements and knowledge. The Climate System History and Dynamics (CSHD) research grant network of which I am a active member. A special thanks to my colleagues and friends, Dr. Michael C. Sawada and Dr. David E. Atkinson, a pair of unmistakable competent researchers and my mentors in the past, present and future. Dr. Philippe Finès's high standards in statistical analyses through numerous discussions brought increased quality and a timely end to this dissertation. Finally, this research dissertation would not of been possible without the hundreds if not thousands of researchers that contributed to the North American Pollen Database (NAPD), a yet unequalled achievement in paleoclimatological research. A special mention of thanks goes to Dr. Thompson Webb III and Dr. John Williams of Brown University for their contributions to the modern pollen dataset, and Dr. Patrick Bartlein of the University of Oregon for his contribution to the modern climate dataset.

This research dissertation was financially supported by a Natural Sciences and Engineering Research Council (NSERC-PGS B), Fonds pour la Formation des Chercheurs et l'Aide à la Recherche (FCAR), and a University of Ottawa Excellence Scholarship (UOES), research grants.

Lastly, this thesis could not have been possible without the continuous love and support of my parents, Mr. And Mrs. Ernest Viau. A special moment of recognition goes to my late uncle Fernand Viau for his unforgettable encouragements. I thank my children, Marc, Shawn and Vincent as they were instrumental in my capacity to continue even in times of troubled waters. To Mr. James Becker, for his everlasting respect and friendship through the years, I say thank you my friend. Thank you also to the LPC members, Daniel Lefebvre, Marcel Valade and Hélène, Lynne and Lisle McCleery, Nicole and Richard Holtman, my numerous nieces and nephews, and the Viau and Charbonneau families, for their support.

André E. Viau, 2002.

MILLENNIAL-SCALE CLIMATE VARIABILITY IN NORTH AMERICA DURING THE PAST 14,000 YEARS

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

Chapter 1

MILLENNIAL-SCALE CLIMATE VARIABILITY: A REVIEW AND UNRESOLVED ISSUES

"From the time when written records began, and doubtless long before that, thoughtful men have had two preoccupations: the past and the future. It is this impelling sense of time and process that distinguishes such individuals from the 'practical man', absorbed as he is in the immediate present. Although we still have with us this divergence of attention and emphasis, the birth of modem science has been, in a curious way, the result of their fusion. For man has, at length, leamt that the key to past and future lies in the scrutiny of evidence that is before his eyes, here and now..."

Sears, 1964.

1. INTRODUCTION

The Quaternary period covers approximately the last 2 million years of our planet's history during which significant environmental changes occurred at many spatial and temporal scales. The Quaternary period was punctuated by glacial-interglacial climatic regimes. Massive continental ice-sheets, especially in the northern hemisphere, expanded and contracted at a pace of approximately 100,000 years at least for the last 800,000 years. Climatic controls on the Earth's surface at low frequency time-scales, such as glacial-interglacial variations (10^4-10^5), are likely due to orbital variations (i.e. precession, axial tilt and eccentricity changes) and plate-tectonic causes. However, there remain fundamental unresolved issues regarding higher frequency time-scales (10^2-10^3) of climate variability on the Earth's surface, which may result from interactions between external and internal forcings of the Earth's climate system (Peixoto and Oort, 1992).

The Earth's climate is a complex system in which nonlinear interactions occur between its components at all spatial and temporal scales. Paleoclimatologists have extensively studied climate variations at scales of 10^6 to 10^4 and meteorologists have investigated climate variability on the order of

10^0 to 10^1. However, there remains a gap in our understanding of the century to millennial scale climate variability. Studies have shown that the last glacial period was punctuated by several cool stadials and warm interstadials, such as Heinrich and Dansgaard-Oeschger events. These millennial-scale climate variations have been tentatively linked to changes in the North Atlantic thermohaline circulation (THC) sensitivity to massive fresh meltwater pulses into the North Atlantic as a possible causal factor. During the last deglaciation, millennial-scale climate variations between cool and warm periods continued to be evident, and include events such as the warm Bolling-Ållerød and cold Younger Dryas period.

During the Holocene, there is evidence from ice cores and other proxy-climate records for an abrupt cooling centered at 8,200 yr. BP. Other more subtle warm and cool intervals are also evident in the Holocene paleoclimate record. The Medieval Warm Period (900-1200 A.D.) and Little Ice Age (1450-1850 A.D.) climate oscillations are good examples of the most recent of these millennial-scale climate variations. Moreover, millennial-scale climate variability is evident throughout most of the Quaternary period. Although several hypotheses have been proposed, we are still unable to identify a suitable mechanism responsible for such climate variability. A need for more detailed research into these millennial-scale climate variations is therefore warranted.

Before any attempt can be made to explore mechanisms that govern millennial-scale climate variability, evidence from all components of the climate system must be gathered. Establishing the timing and spatial extent of these millennial-scale variations will lead to insight into which external and/or internal forcings may be responsible. The purpose of this dissertation is to provide a better understanding of the spatial extent of this millennial-scale climate variability. Results from this type of research can aid climate modellers by providing better understanding about atmosphere-cryosphere-ocean-biosphere interactions at millennial scales.

Our climate is facing a potential human-induced doubling of atmospheric CO_2 within a relatively short-time scale, which limits the degree to which past analogues can be used to predict our future climate state. However, understanding past climate history and dynamics at century to millennial-scale will provide essential tools to better estimate our climate system response to anthropogenic perturbation. Determining and understanding past natural climate variability is fundamental for research on recent anthropogenic forcing to assist in distinguishing natural climate variability from human-induced effect. This kind of research can only lead to improved climate models and allow more accurate estimates and better planning strategies on the impact of anthropogenic influence in the near future.

1.1. BACKGROUND LITERATURE

Recently, there has been considerable focus on millennial-scale climate change. This scale of variability seems to operate independently from glacial-interglacial climate regimes. The marine and ice core records offer convincing evidence that the Earth's climate has been punctuated by a pervasive quasi-periodic 1,500 ± 500 year climate variability between cool and warm conditions at least in the North Atlantic region and during the past one million years (Bond *et al.*, 1997, 2001; MacManus *et al.*, 1999; O'Brien *et al.*, 1995; Stuiver *et al.* 1997; Oppo *et al.* 1998; Raymo *et al.* 1998; and Bianchi and McCave, 1999). Millennial-scale climate variability was first observed in terrestrial records more than a century ago in European peat stratigraphies (In: Birks and Birks, 1980) and subsequently in botanical and alpine glacier records (e.g. Denton and Karlen, 1973, Wendland and Bryson, 1974). However, it remains to be seen if this millennial-scale climate variability was global or restricted to the North Atlantic region. As discussed below, the continental to global scale expression of millennial-scale climate variability is not well established.

The last 2.75 Ma have been marked by glacial-interglacial climate oscillations that have been recurring at approximately a 100,000 yr periodicity at least for the past 800,000 years (Berger *et al.*, 1993). These glacial-interglacial oscillations have, since the 1980's, been attributed to astronomical causes. According to this early twentieth century "Milankovitch theory", long-term climate change should unfold with clocklike precision (Kerr, 1999) assuming the lag of the climate system is shorter than the scale of this forcing. Superimposed on these orbitally forced changes are evidence of higher-frequency climate oscillations as shown in proxy-climate records derived from ice cores, marine, and terrestrial sediments. Although mainly observed in the North Atlantic and Greenland ice sheet records, other climate proxy records distributed around the world have been tentatively correlated to those identified in the North Atlantic. Based on this information, millennial-scale climate variability is a pervasive feature of the global climate system. Moreover, the evidence suggests synchronous and global millennial scale climate variability throughout most of the last million years. These higher frequency climate variations may result from complex nonlinear relationships between the climate system components and/or scale interactions between external and internal forcings, or of yet unknown external forcings.

It is this natural climate variability operating on all time-scales that poses the major difficulty in understanding the impact of recent anthropogenic forcing on the climate. Although mechanisms for such natural forcings have been advanced, such as the role of ocean-atmosphere-ice interactions (Broecker and Denton, 1989), orbital harmonics (Pestiaux *et al.*, 1988), nonlinear feedbacks within the climate system (Peixoto and Oort, 1992; Rind, 1999), volcanism (Zielinski *et al.*, 1996, Crowley, 2000) and solar output variability (Van Geel *et al.*, 1999, Crowley, 2000, Bond *et al.*, 2001), there is no clear indication as to which one or combination are responsible for millennial-scale climatic fluctuations.

Evidence for Holocene climate variations first came from the terrestrial records before radiocarbon dating methods were developed. The two main proxy-indicators for climate change then were alpine glacier advances and retreats and peat stratigraphy. In the case of peat accumulation, layers of humidified *Sphagnum* were taken as indicators of fast peat growth and therefore deposited during wetter climatic conditions (Birks and Birks, 1980). Layers of darker, more decomposed peat were indicative of drier, warmer conditions. Most of these studies were conducted in northern Europe, and a millennial-scale climatic sequence was established for the Holocene, subsequently called the Blytt-Sernander climatic scheme. However, only some of these climate transitions could be accurately dated by correlation with the known archaeological record.

Later, other terrestrial records provided evidence of Holocene millennial-scale climate variability. For example, Denton and Karlén (1973) found synchronous advances of mountain glaciers in North America and Europe. They suggested that these correspond to climate variations that were part of a regular millennial-scale pattern, which when projected backward coincided with climatic shifts of the preceding glaciations and when projected forward predicted a progressive warming over the next few centuries (Denton and Karlén, 1973). However, the synchronicity of these events was questioned by Grove (1979).

Holocene climate variability was also identified in pollen records (e.g. Bryson and Wendland, 1967, Wendland and Bryson, 1974; Bryant and Holloway, 1985 and Wright *et al.,* 1993). However, large-scale syntheses were difficult, partly because it was difficult to find corroborative evidence in the marine and Greenland ice core records. For example, measurement of oxygen isotope, methane concentrations, and snow accumulation in Greenland ice cores revealed no evidence of millennial-scale variations during the Holocene, except perhaps a brief cooling event centered around 8,200 years ago (Alley *et al.*, 1997), although Canadian ice core records showed these fluctuations (Fisher, 1983). However, O'Brien *et al.* (1995) later demonstrated, using measurements of soluble impurities in the new GISP2 ice core, that Holocene atmospheric circulation above the ice cap was punctuated by a series of millennial-scale climate shifts that appeared to correlate with Denton and Karlen's glacial advances. This prompted Bond *et al.* (1997) to analyze high-resolution marine sediments in two cores from opposite sides of the North Atlantic. They postulated that periodic ice-rafting events (IRD) in the North Atlantic region were not restricted to glacial periods but also occurred during the Holocene at a 1470 ±500 year rhythm. The authors then concluded that the cause of these millennial-scale variations would lie in ocean-atmosphere interactions.

Several recent marine and ice core studies have established a persistent millennial-scale climate periodicity, in some cases operating independently from glacial-interglacial modes. For example, Stuiver *et al.* (1997) found a strong 1470-year periodicity in ^{18}O from the GISP2 ice core (i.e. from Greenland) for the period 50,000-12,000 BP. In a deep-sea core obtained in the eastern North Atlantic, MacManus *et al.* (1999) detected a recurring millennial-scale pattern of sea-surface temperature change extending over the past 500,000 years. In another North Atlantic core, Oppo *et al.* (1998) also found persistent millennial-scale climatic oscillations covering the period of 500,000-340,000 years BP. Raymo *et al.* (1998) found that millennial-scale climate oscillations were operating well before 1 Ma ago in the North Atlantic. Bianchi and McCave (1999) examined the grain sizes of deep-sea sediments in cores extracted from the northeast Atlantic Ocean south of Ireland. This study showed evidence of millennial-scale warm/cool temperature variations at approximately every 1,500 years for the last 11,000 years. These authors postulated that these temperature variations were of similar magnitude to the Medieval Warm Period and Little Ice Age. In yet another study, Bond *et al.*, (2001) showed a strong correlation between North Atlantic IRD events and the $^{14}C_{atm}$ and ^{10}Be record, which are proxies for solar output variability, and postulated a variable solar output as an explanation of century to millennial-scale climate variability.

Millennial-scale climate variations between cool and warm conditions are also evident during the Holocene in other parts of the world. For example, millennial-scale climatic variations were found across northern Africa and southern Asia (Sirocko *et al.*, 1993), in western North American vegetation (Heusser, 1998), in the Sahel region (Street-Perrott and Perrott, 1990), in West Asia and North Africa (Gasse and Van Campo, 1994), in peat bogs in north-west Scotland (Anderson, 1998), in the midwest of North America (Jacobson and Grimm, 1986, Dorale *et al.*, 1992), in the Pacific northwest of North America (Whitlock and Bartlein, 1997), and in western Europe (Joussaume and Guiot, 1999; Labeyrie and Jouzel, 1999).

1.2 Sequence of events - linkages of millennial-scale variations from Glacial to Holocene

There is evidence of sudden and short-lived warm events that occurred many times during the last glacial period between 110,000 and 10,000 years ago in the North Atlantic (Dansgaard *et al.*, 1993; Mayewski *et al.*, 1997). Between 115,000 and 14,000 years ago, there are 24 of these, which have subsequently been named "Dansgaard-Oeschger events" (DO). Each warm interstadial is preceded or followed by a cold stadial and these "Dansgaard-Oeschger" events last approximately 1,500 years (Bond *et al.*, 1993, Bond and Lotti, 1995, Mayewski *et al.*, 1997). Subsequent work has identified many of these DO events in different places around the world. Nine of these were greater than 2,000 years in length and eight of these are evident in the Antarctic Vostok ice core record (Bender *et al.*, 1994). Evidence is also found in marine sediments where 17 DO events have been identified in the Santa Barbara basin in the Northeast Pacific (Behl and Kennett, 1996). Seventeen DO events have been identified in ocean cores extracted from the Bermuda Rise (Adkins *et al.*, 1997) and the Arabian Sea (Sirocko *et al.*, 1993). All these DO events show temporal synchronicity between the different records.

Ice-core and ocean data suggest that these millennial-scale warm interstadials (i.e. DO events) both began and ended abruptly. In general the abrupt transition in climate at the start was followed by a more gradual decline involving a series of smaller cooling events terminating in a larger one, which returned the climate to glacial conditions (Rasmussen *et al.*, 1997). Sakai and Peltier (1996) simulated with success the natural Dansgaard-Oeschger millennial-scale oscillations during glacial periods using a multi-basin model of the global thermohaline circulation and postulated that the North Atlantic thermohaline circulation must have played an important role in, for example, the Younger-Dryas climate reversal.

Heinrich events were first described as periods of intense ice-rafting debris (IRD) in glacial age marine sediments at roughly 10,000-year intervals in the eastern part of the North Atlantic (Heinrich, 1988). These are associated with extreme and short-lived cold events and seem to be restricted to glacial periods (Broecker *et al.*, 1992; Bond *et al.*, 1992,1993; Andrews, 1998). Evidence of a climate impact of Heinrich events also show up in the GRIP and GISP2 ice core record as a decrease in temperature (Mayewski *et al.*, 1997) as well as in European and North American pollen (Guiot *et al.*, 1993; Grimm *et* *al.*, 1993) and other marine sediments records (Bond *et al.*, 1992,1993; Bond and Lotti, 1995). There is also evidence of cooler temperature in other regions of the world, as far as the Arabian Sea and the Tibetan plateau (Thompson *et al.*, 1989; Schultz *et al.*, 1998). It is therefore probable that Heinrich events were sufficiently intense to have had a global impact on the climate. The timing of the onset, duration and termination of Heinrich events are still disputed because of their duration, which is no more than a few centuries (Mayewski *et al.*, 1997). However, since Heinrich events lasted only a few centuries at most, it is probable that the onset and termination cannot have been much longer than a few decades (Mayewski *et al.*, 1997). Although Heinrich events are considered to be glacial phenomena, there is debate about whether such events continued during deglaciation.

The last deglaciation was characterised by a warm interstadial, the Bolling-Ållerød, followed by a sudden cooling event named the Younger Dryas, which is considered the final phase of cold, glacial conditions preceding the abrupt warming at the beginning of the Holocene around 10,000 years ago. It has also been suggested that the Younger Dryas had the general features of a Heinrich event and is sometimes referred to as H0 (Bond *et al.*, 1997). Because of the problems associated with radiocarbon dating, the timing of the onset, duration and termination of the Younger Dryas have long been disputed (Anderson, 1997). There have been disagreements over the timing of the Younger Dryas as revealed by dendrochronology, varved lake sediments, ice cores and uranium/thorium dates on corals (Anderson, 1997).

The beginning and termination of the cold Younger Dryas event is estimated at 12,820 ± 260 cal/yr to 11,640 ±250 cal/yr from the GISP2 core (Alley *et al.*, 1993) and 12,700 ±100 to 11,550 ±90 cal/yr BP from the GRIP core (Johnsen *et al.*, 1992). Both ice cores, which have nearly annually resolved chronologies, agree that the Younger Dryas event lasted for approximately 1,300 years. While there are still some uncertainties in the precise timing of the Younger Dryas oscillation, most high-resolution ice and terrestrial sequences agree that the onset and termination of the event was abrupt (Dansgaard *et al.*, 1989; Broecker *et al.*, 1989; Lehman and Kegwin, 1992; Taylor *et al.*, 1997). A temperature decline is also evident in Western Europe, although it is more subtle (4-5°C), and in northeastern North America (Cwynar *et al.*, 1994). Outside the North Atlantic region, the impact of Younger Dryas cooling is not well established. Model simulations suggest that the associated change of surface air temperature is particularly large in the northern Atlantic and its adjacent regions, but relatively small in the rest of the world (Manabe and Stouffer, 1995).

Although the causes and consequences associated with this event are not fully understood, it has become increasingly clear that this climatic oscillation is associated with a major change in the thermohaline circulation (THC) of the North Atlantic which may be triggered by a complex combination of iceberg discharge and meltwater input (Dansgaard *et al.*, 1989; Broecker *et al.*, 1989). The abrupt onset and termination of the Younger Dryas cold phase raises fundamental questions about nonlinearities and feedback mechanisms within the ocean-ice-atmosphere system (Anderson, 1997).

The abrupt Younger-Dryas to Holocene transition marks the start of our present interglacial, defined as beginning approximately 10,000 years ago. During most of the Holocene, other climate events of several time scales have been identified. One event that is of particular interest is a sudden cooling at 8,200 years ago. This corresponds with the Cochrane-Cockburn ice-sheet re-advance (8000-8200 yr BP; Denton and Karlén, 1973) and vegetation change in Greenland (Fredskild, 1973). This widespread cool event presumably lasted at most a few hundred years and is clearly evident in the Greenland ice cores, where the temperature decrease is estimated to be half the magnitude of the Younger Dryas to Holocene difference (Alley *et al.*, 1997; Mayewski *et al.*, 1997). Given that the event lasted only a few hundred years, it is most probable that the onset and termination of this event must have been rapid, perhaps in a few decades or so. Sirocko *et al.*, (1993) suggested that a failure of the summer monsoon rains over North Africa and southern Asia occurred at this time, although the correlation may not be reliable. Cold and dry conditions also seem to have occurred in northeastern South America, eastern North America and northwestern Europe (Alley *et al.*, 1997).

1.3. RESEARCH QUESTIONS

Although marine and ice cores records offer convincing evidence for millennial-scale climate change at least for the North Atlantic region, little has been done to corroborate the timing, magnitude, and geographical extent within the terrestrial record. Much of the evidence is based on a few cores, which are assumed to be representative of large geographic regions. We do not yet know how widespread these millennial-scale events are, or what causes them. Uncovering the timing, magnitude, and geographical extent of terrestrial Holocene and late-glacial millennial-scale climate change will aid climate modellers better understand the interactions and feedbacks within the climate system such as the biosphere-atmosphere feedback.

THE MAIN OBJECTIVE OF THIS RESEARCH DISSERTATION IS TO DETERMINE WHETHER THERE IS EVIDENCE OF PERSISTENT MILLENNIAL-SCALE CLIMATE VARIABILITY IN THE NORTH AMERICAN TERRESTRIAL FOSSIL RECORD DURING THE HOLOCENE AND LATE GLACIAL.

The fundamental research questions are:

1. *Are there times of change evident in the terrestrial fossil record of North America that are synchronous in space, that would support a conclusion of millennial-scale climate change during the Holocene and Late Glacial?*
2. *What is the geographical extent of these changes in North America?*
3. *Are there differences in the magnitude of these events through the Late Glacial and Holocene?*

1.4 HYPOTHESES

Recent marine and ice cores studies suggest a pervasive 1,500 +- 500 years climate periodicity operating independently from glacial/interglacial modes, at least in the North Atlantic region. Although the correlations may not be reliable, other marine, ice, and terrestrial records from sites around the world have been tentatively correlated with this North Atlantic climate variability. Uncovering the presence/absence and timing of these millennial-scale climate variations on a continental-scale may bring insight into the cause of millennial-scale climate variability. Therefore the first hypothesis to be tested is:

H_1: NORTH AMERICAN TERRESTRIAL FOSSIL RECORDS SHOW MILLENNIAL-SCALE CLIMATE VARIATIONS AT THE PACE OF 1,500 ± 500 YEARS CONSISTENT WITH THE ICE AND MARINE RECORDS DURING THE LATE GLACIAL AND HOLOCENE.

Next, one previously-suggested hypothesis is that these millennial-scale climate variations were restricted to the North Atlantic and adjacent land masses (Manabe and Stouffer, 1995), while the alternative is that they were widespread, perhaps even global in extent (Alley *et al.*, 1997). Therefore the second hypothesis being tested is:

H_2: MILLENNIAL-SCALE CLIMATE CHANGES WERE RESTRICTED TO NORTHEASTERN NORTH AMERICA DURING THE LATE GLACIAL AND HOLOCENE.

Lastly, it has been suggested that temperature variations within millennial-scale climate variations are greater during glacial maxima, more subtle during interglacials, and greatest during periods of ice sheet growth and disintegration (Bond *et al*, 1997, Bianchi and McCave, 1999). The last hypothesis being tested is:

H_3: TEMPERATURE VARIABILITY WITHIN THESE MILLENNIAL-SCALE CLIMATE VARIATIONS WAS GREATER DURING THE TRANSITIONAL PERIOD OF THE LAST DEGLACIATION, AND SUBTLER DURING THE HOLOCENE IN NORTH AMERICA.

1.5 CONCEPTUAL FRAMEWORK

In order to address these questions, we must first determine the temporal and spatial scales of the study. Next we must choose the appropriate available data and understand the assumptions regarding their use. Finally, we must choose the methods and techniques that are suitable for this kind of analysis.

Determining the spatial and temporal scales and the data suitable for this research involves two questions. First, what temporal scales are necessary to study the climatic scale of interest? A general standard to study any particular temporal scale of interest is to analyse data with a resolution that is one scale below the frequency of interest and a period that is long enough to draw reasonable conclusions. In this research, the scale of interest is 1,500 year or the millennial scale (10^3). Therefore, to address the frequency of interest requires data with at least century-scale resolution (10^2). The second consideration is to determine an appropriate period or length of time; more than one cycle is required for any reasonable conclusion. Millennial-scale climate variability is relatively well understood for the last glacial period (i.e. warm interstadials and cool stadials) and well documented for the last deglaciation phase (i.e. warm Bolling-Ållerød and cool Younger Dryas). However, the linkages between these millennial-scale climate variations and the Holocene are not well established. In order assess these linkages; the period must be sufficiently long to include at least part of the documented millennial-scale climate record. In this research, the period of study involves the last 14,000 calendar years, which is adequate to resolve the assumption of continued millennial-scale climate variability between the last deglaciation phase and the Holocene. Thus, this period (10^4) is adequate for the study of millennial-scale climate variability and to study the linkages between glacial and interglacial millennial-scale climate variability.

Second, what spatial scale is required to answer the questions asked? For a climatological question, a continental scale, such as North America, is the minimum that can provide useful paleoclimatic information. It is imperative that the data chosen is geographically extensive enough to represent most regions of the continent.

The only available terrestrial data that is geographically extensive enough and temporally adequate for this kind of analysis is pollen data available from the North American Pollen Database (NAPD)(Grimm, 2000), which will be discussed later. In order to proceed with the analysis using any proxy climate data, we must first be aware of the assumptions underlying the use of the climate sensor to determine its suitability for the questions of interest and to consider while interpreting the results. The methodology under consideration is dependent on a number of assumptions, such as the relationship between pollen counts extracted from lake or peat sediment cores and their surrounding vegetation composition, as well as the relationship between vegetation and climate. However, Webb and Bryson (1972) have noted that it is the pollen-climate relationship that is critical irrespective of the relation of pollen and vegetation or vegetation to climate. Questions about vegetation response time to climate change and whether vegetation is in dynamic equilibrium with climate still remain under debate, although it is now understood that this question is scale-dependent (Prentice, 1986; Appendix A).

1.5.1 POLLEN-VEGETATION RELATIONSHIP

Records of past vegetation can be found in lake sediments, bogs and other sedimentary terrestrial environments, from which pollen grains can be extracted and identified. Since pollen grains are well mixed in the air, the sedimentary records consist of an integration of the pollen from the "catchments area". A pollen count from sediment is generally a statistically valid sample of the contemporary pollen in the air above an area (Watts, 1973) and therefore the pollen assemblages are related to the surrounding vegetation (Birks and Birks, 1980).

Studies have shown that there is a strong relation between pollen and vegetation at regional and continental scales (Davis and Webb, 1975). Comparisons between isopoll (i.e. contour line of equal percentages) and vegetation maps for eastern North America and the Northwestern boreal forest show that the pollen record represents major vegetation types in the area (Delcourt *et al.*, 1984, Delcourt and Delcourt, 1985, Anderson *et al.*, 1991).

Although there is concern about long distance transport, and over-and under-representation of certain species (Davis and Webb, 1975), pollen diagrams can be used to determine changes in vegetation through time if these considerations are accounted for. Spatial and temporal limits of resolution of past vegetation change obtained from pollen diagrams is still an important area of research (Gajewski, 1998). These spatial and temporal limitations result from several factors including sedimentation rates, bioturbation, sampling intervals, and errors associated with radiocarbon dating (Green, 1983, Gajewski, 1998).

1.5.2. VEGETATION-CLIMATE RELATIONSHIP

Vegetation has a strong relationship with modern climate conditions at least at the regional and continental scales. Vegetation boundaries have shown to be constrained by climatic conditions prevailing over certain regions. For example, the northern limit of the circumpolar Boreal forest coincide well with the mean winter position of the polar front (Bryson, 1966) and oak exists within the range of maritime tropical air over southeastern North America. This is true at a given time, such as the present, but whether vegetation changes lag or not climate change remains a major paleoecological question. Current methods for estimating past climatic patterns from pollen data require that the vegetation be in dynamic equilibrium with the climate (Webb, 1986). There is still debate over the extent to which vegetation changes lag behind climatic change, and whether or not vegetation is ever in equilibrium with climate (Webb, 1986). Vegetation does not respond in a simplistic way to changes in climate and there is no single answer to the equilibrium question. Because climate varies on all time scales, "judgement about equilibrium of vegetation to climate should be made separately for different time scales of climate change" (Webb, 1986). In order to achieve equilibrium between vegetation and climate changes at a particular time scale, the vegetation response time must be sufficiently fast compared to the time scale of climatic variation to which it is responding (Prentice, 1986, Webb, 1986).

The time required for vegetation to respond completely to climatic forcing at a time-scale of 10^4 is still unknown, but records of the vegetation response to climatic events of century to millennia scale duration provide evidence for relatively short response times (Webb, 1986). For example, Gajewski, (1987) shows the timing of individual taxa response associated with the Little Ice Age event (1450-1850 AD) along a transect of differing plant communities across eastern North America. An example of millennial-scale vegetation response during glacial times is the alternating pine-oak dominancy in a core extracted from Lake Tulane, Florida, that correlates well with the North Atlantic Heinrich events (Grimm *et al.*, 1993). This latter example illustrates well that climatic conditions favoured pine over oak or vice-versa without any of the two genera becoming locally extinct. In yet another lake core from northern Florida, there is evidence for the sudden cooling event of 8,200 BP. (Watts and Stuiver, 1980). In light of this evidence, this research assumes that vegetation was in equilibrium at the millennial-scale during the Holocene and Late Glacial. Appendix A offers a detailed discussion regarding the vegetation-climate equilibrium question.

1.6 STUDY AREA

The study area is North America, excluding Mexico where there are few data available. The current climate conditions are shown in Figure 1.1a, b. and the modern distribution of vegetation is shown in Figure 1.1c. There are a number of ecotones varying both latitudinally and longitudinally, which can be correlated with frontal zones between air masses (Bryson, 1966). The diversity of climates and topographical features of North America provides an excellent area to study synchronous climate change at the continental scale. For example, if a climate change is evident in the northeastern part of North America, then we can ask ourselves if it is also evident in other regions or across other ecotones of the continent. Note that neither the magnitude nor the direction of the change needs to be the same in all areas. Regional differences may give insight into general atmospheric circulation pattern anomalies of the past.

1.7 DATA

Published and unpublished radiocarbon-dated fossil pollen spectra from a dense network of sites from the NAPD (Grimm, 2000) will be used as a proxy for relative paleovegetation change through the Holocene and Late Glacial.

1.7.1 FOSSIL POLLEN DATA

The number of available sites in North America from the NAPD (Figure 1.2) that can be used for paleoecological/paleoclimatological analysis varies through time and in space (Table 1.1). There are 752 sites unevenly distributed throughout the continent. Some regions such as the Great Lakes region are well represented while other regions have fewer data (Figure 1.2). Appendix B offers a detailed description of the content of the NAPD.

1.7.2 Modern pollen dataset

A new modern pollen dataset was assembled and subsequently named WWKGVS2000. Modern pollen refers to pollen extracted from the uppermost 1-2 cm of sediments of a lake that were deposited in the past few years. This modern pollen dataset is derived from Arvizinis and Webb (1985), where Webb and Williams (pers. comm.) added a number of new modern pollen sites (WW2000). Then sites from the Canadian Pollen Database (KG2000) and Global Pollen Database (GPD, Grimm, 2000) were merged into a single dataset (WWKGVS2000). This gave a final modern dataset of 4590 sites, which covers most regions of North America (Figure 1.3). The complete procedure is documented in Sawada (2001).

1.7.3 Pollen sum

In general, the choice of a pollen sum should emphasize the question of interest and maximize the signal-to-noise ratio (Jackson, 1994). Including pollen taxa that are not necessary for the question of interest increases the noise and decreases the signal. For example, the exclusion of aquatic taxa is common if terrestrial vegetation is of primary interest. In this dissertation, I use different sets of pollen sums as will be discussed in detail below.

1.8 Methodological framework

This section presents a general overview of the methodology, methods, and techniques, which I used for testing the hypotheses. Each subsequent chapter contains a more detailed methods section.

Gajewski (1993) put the assumption that vegetation is in equilibrium with climate at certain time scales into perspective. He suggested that studying networks of sites at a regional to continental scale would reveal major trends in vegetation change, and hence broad-scale climate change. Rather than studying climate at individual sites, geographic arrays of sites are needed in order to reduce the signal-to-noise ratio of the response of vegetation to climate changes at specific time scales (Gajewski, 1993). With this kind of approach, widely separated sites may show different directions of change in vegetation and pollen assemblages, although the vegetation at those sites may be responding to the same change in climate (Gajewski, 1993, 1987).

1.8.1 H_1: NORTH AMERICAN TERRESTRIAL FOSSIL RECORDS SHOW MILLENNIAL-SCALE CLIMATE VARIABILITY AT THE PACE OF 1,500 ±500 YEARS CONSISTENT WITH THE ICE AND MARINE RECORDS DURING THE LATE GLACIAL AND HOLOCENE.

Palynologists have long interpreted vegetation change and climate variability based on biostratigraphical discontinuities in pollen diagrams. With the advent of radiocarbon dating, it has been possible to date these events and construct chronologies of past vegetation change. Wendland and Bryson (1974) assumed that climate changes were abrupt step functions that were reflected in discontinuities within sedimentary sequences that were preferentially dated by researchers. In this way, they identified globally synchronous discontinuities for the Holocene using 855 ^{14}C dates from lake cores, glaciers, sea levels and peats. They found relative peaks in the distribution of radiocarbon dates using the method of partial collectives (Johnson, 1966, Bryson, 1966, Bryson and Wendland, 1967) to identify the most significant times of climatic change (Figure 1.4). To test this first hypothesis, I use an approach that is conceptually similar to the method of partial collectives used by Wendland and Bryson (1974), but implemented using a new methodology to determine relative modes in the distribution of radiocarbon dates for North America.

An expanded set of 3076 ^{14}C dates is currently available which are extracted from published and unpublished biostratigraphical sequences for North America (Grimm, 2000). The distribution of ^{14}C dates is then analyzed using a nonlinear mixture modeling approach (Titterington et al., 1985, Everitt and Hand, 1981). Essentially, the method involves fitting multiple Gaussian curves to solve the multimodal distribution in order to identify the most significant modes or biostratigraphical discontinuities of the distribution of ^{14}C dates through time. The most significant modes found are then correlated with major millennial-scale climatic change from the ocean, ice and other terrestrial records. The main advantage of this approach is the fact that age-depth modeled chronologies, and taxonomic resolution within sedimentary sequences is not a factor (see Appendix B). The main disadvantage lies with dependence between the Gaussian curves fitted to the distribution that must sum to 1 in order to maintain a true probability distribution function (PDF).

1.8.2 H_2: Millennial-scale climate changes were restricted to Northeastern North America during the Late Glacial and Holocene.

A central question concerning millennial-scale climate change is to determine the geographical extent of these changes. Are they global phenomena or restricted to the North Atlantic and adjacent terrestrial regions? Most marine and ice core analyses that have identified these millennial-scale climate variations have been undertaken in the North Atlantic region. Although inferences and correlations have been made with other sites from around the globe, corroborating evidence remain sparsely distributed in space and cannot, at this time, confirm a linkage to the North Atlantic events. Objectively uncovering the geographical extent on a continental-scale will assist in understanding the spatial extent of these millennial-scale climate events.

Gajewski (1987) showed vegetation response to century-scale events such as the Little Ice Age using Principal Component Analysis (PCA) (Figure 1.5). By using a spatial array of sites associated with different plant communities, he was able to show the timing of plant community response to the Little Ice Age. Incorporation of both time and space dimensions reduced the signal-to-noise ratio at that scale. In other words, the pollen data of individual sites may not warrant an interpretation of a climate change following a subtle change in vegetation at that site, however, if several sites in a region associated with different plant communities show a synchronous change in vegetation, whether subtle or not, this would support an interpretation of climate control.

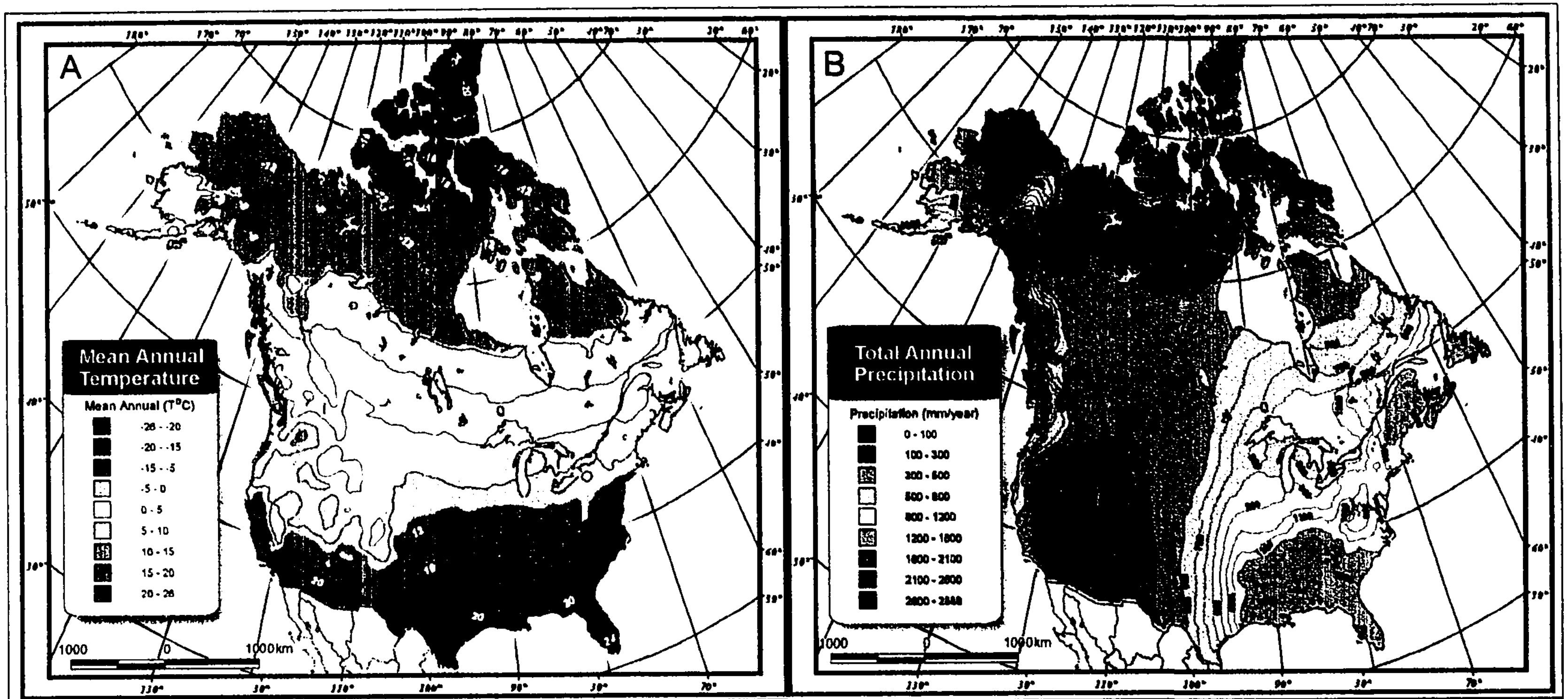
To approach this question, I used Principal Component Analysis (PCA) on the North American Pollen Database (NAPD). PCA has been widely applied to questions in geographical and ecological studies. In this study, a PCA was simultaneously implemented in *time, space and taxonomic axes* to test for spatial and temporal correlations between different regions of North America. PCA is a technique from the family of factor analysis. The main applications of factor analysis techniques are: (1) to reduce the number of variable using the correlations between the variables and (2) to detect structure in the relationship between variables. Therefore, factor analysis is applied as a data reduction or structure detection method. In essence, a PCA identifies variance common and unique in time and space and extracts this information as new variables called components which, when plotted, summarize the major spatio-temporal patterns that exist in the data (Legates, 1991).

1.8.3 H_3: TEMPERATURE VARIABILITY WITHIN THESE MILLENNIAL-SCALE CLIMATE VARIATIONS WAS GREATER DURING THE TRANSITIONAL PERIOD OF THE LAST DEGLACIATION, AND SUBTLER DURING THE HOLOCENE IN NORTH AMERICA.

The marine and ice core records have suggested that these millennial-scale climate events were strong during glacial maxima (i.e. Dansgaard-Oeschger-Heinrich events), more subtle during interglacials (Bond *et al.,* 1997), but strongest during periods of ice sheet growth and disintegration (i.e. Bolling-YD-Holocene-8, 200 event transitions). As stated above, the glacial period has been well studied, and there is evidence that vegetation has responded abruptly in glacial time at the millennial scale (Grimm *et al,* 1993, Guiot *et al.,* 1993).

Chapter 4 illustrates an approach to test this hypothesis. Here, I use the modern analogue technique (MAT) (Overpeck *et al.*, 1985) to reconstruct past temperatures from a dense network of fossil pollen records (Figure 1.2). Essentially, this method compares a fossil pollen spectrum to all modern pollen samples available for the geographic extent of the research area. For each fossil pollen spectrum at a given target time (e.g. 100 years), the best analog is the modern site (Figure 1.3) that possesses the least dissimilar modern pollen spectrum. From this, climate variables such as temperature and precipitation can be extracted and assigned to the past. Consequently, a mean continental July temperature anomaly curve is reconstructed for the Holocene and Late Glacial period for North America. Where opposing directions of change may exist regionally at this scale of interest, an average temperature anomaly curve for the entire continent should give insight on the magnitude of change during the Late Glacial and Holocene at the continental-scale. Maps of the mean July temperature anomalies were also constructed. This approach is the basis for quantifying our current global warming and is here applied at the continental scale of North America.

The results of these three studies are to illustrate the spatio-temporal evidence of millennial-scale climate variability in North America during the past 14,000 calendar years BP. The reader should bear in mind that the results and interpretations presented below emphasize the continental-scale expression of millennial-scale climate variability in North America.



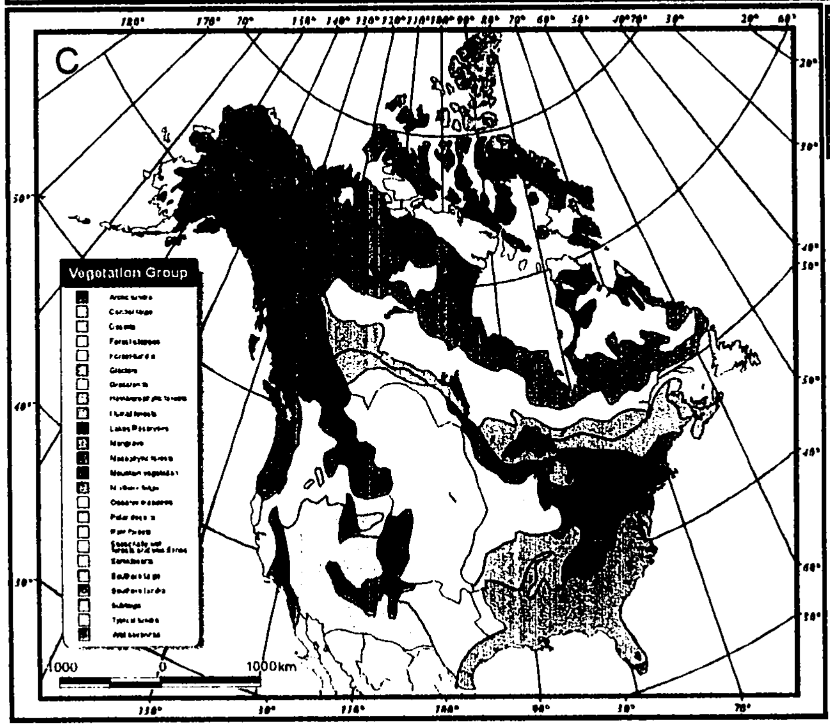


Figure 1.1: North American climate and vegetation
(Provided by M.C. Sawada)

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).

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

Albers Equal-Area Conic
Cen. Md.- 100W; Orig. - 55 N
Standard Parallel 1 - 29.5 N
Standard Parallel 2 - 80.0 N

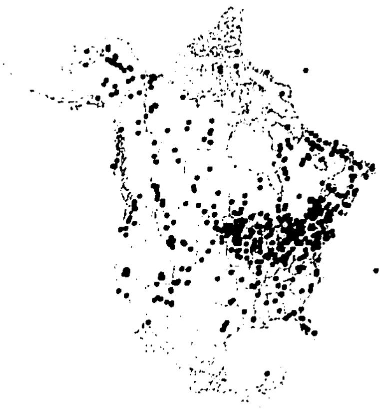


Figure 1.2: Fossil pollen site distribution in North America (source:NAPD, Grimm, 2000).

Time	# sites
0-1000	639
1000-2000	566
2000-3000	560
3000-4000	553
4000-5000	545
5000-6000	532
6000-7000	513
7000-8000	478
8000-9000	470
9000-10000	453
10000-11000	390
11000-12000	304
>12000	253

Table 1.1: Site density through time in North America (source:NAPD, Grimm, 2000).

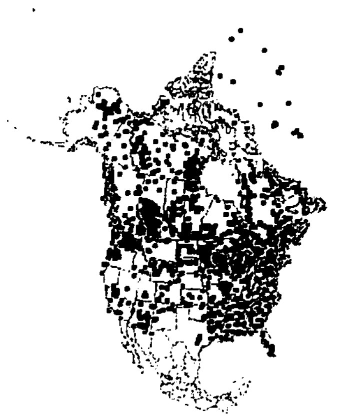


Figure 1.3: Modern pollen site distribution in North America (source: WWKGVS2000).

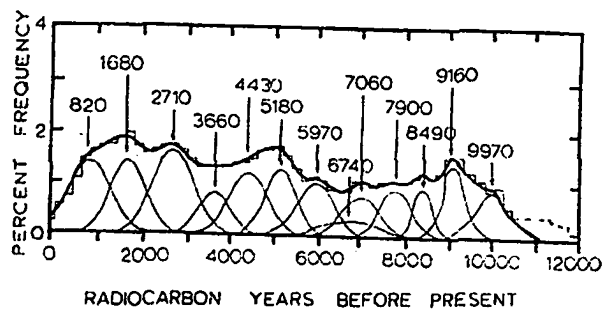


Figure 1.4: Partial collectives used to solve the multimodal distribution of the total sample (source: Wendland and Bryson, 1974).

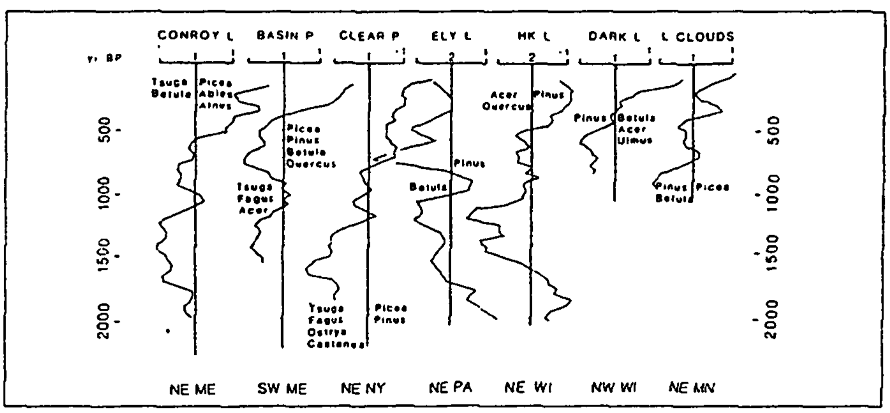


Figure 1.5: Principal components analysis of 7 fine-resolution pollen diagrams (source: Gajewski, 1987).

Chapter 2

WIDESPREAD EVIDENCE OF MILLENNIAL-SCALE CLIMATE VARIABILITY IN NORTH AMERICA DURING THE PAST 14 000 YEARS

Abstract

There is debate concerning the spatial extent and magnitude of a recently proposed 1500-yr climate oscillation during the Holocene. Existing evidence is largely restricted to the North Atlantic and adjacent landmasses and during the Holocene, the spatial extent, magnitude, and effects of these climate variations on the terrestrial environment have not been established. This chapter shows that millennial-scale climate variability caused changes in vegetation communities across all of North America during the past 14 000 calendar years (cal yr). Times of major transitions identified in pollen records occurred at 600, 1650, 2850, 4030, 6700, 8100, 10190, 12900, and 13800 cal yr B.P., consistent with ice and marine records. It is here suggested that North Atlantic millennial-scale climate variability is associated with rearrangements of the atmospheric circulation with far-reaching influences on the surface climate.

2. INTRODUCTION

Evidence from marine and ice-core records suggests a persistent North Atlantic millennial-scale climate oscillation operating for the past million years with a 1500 ± 500 yr quasi-periodic cycle, independent of the glacial-interglacial cycles (McManus *et al.*, 1999; Raymo *et al.*, 1998; Oppo *et al.*, 1998). Rapid environmental changes occurred during the Wisconsinan, most likely associated with changes in ocean-circulation patterns and large outflows of icebergs in the North Atlantic region (Dansgaard *et al.*, 1993; Bond *et al.*, 1992; Bond and Lotti, 1995; Grimm *et al.*, 1993; Guiot *et al.*, 1993).

Evidence for millennial-scale Holocene climate variability was first observed in the 1800s, when alternating climate periods were interpreted from peat bogs of northern Europe. This variability became established as the Blytt-Sernander zonation (Flint, 1971). Synchronous advances of North American and European alpine glaciers during the Holocene also suggested millennial-scale climate variations (Denton and Karlén, 1973) although this interpretation was questioned (Grove, 1979). More recent measurements from soluble impurities in Greenland ice (O'Brien *et al.*, 1995), and marine sediments (Bianchi and McCave, 1999; Bond *et al.*, 1997, 2001) show similar Holocene millennial-scale climate shifts in the North Atlantic region, which are also most likely due to reorganizations of the atmosphere-ocean system (Bianchi and McCave, 1999; Bond *et al.*, 1997, 2001). We here show that Holocene millennial-scale events are found in the terrestrial pollen record and are widespread throughout North America.

One approach to identify times of change was proposed by Wendland and Bryson (1974), who analyzed terrestrial records of botanical and archaeological change to search for climate transitions in the Holocene. They assumed that climate changes are abrupt step functions causing discontinuities within sedimentary pollen and archaeological sequences and further that paleoecologists preferentially date these biostratigraphical changes. They used the frequency distribution of 855 ^{14}C dates and the method of partial collectives to identify relative peaks in the multimodal distribution of these records (Johnson, 1966; Bryson, 1966), finding modes consistent with the Blytt-Sernander framework.

The method of partial collectives has some problems. The number of normal curves to be fitted is determined by the researcher on a trial and error basis, and the initial estimates (ie. means, standard deviations and percentage of total sample) must be determined *a priori* by the researcher to best represent the distribution. A consequence is that the suggested initial values may force or influence the solution. In this chapter we use the general approach of Wendland and Bryson (1974) on an updated database of radiocarbon dates from North American pollen diagrams. We apply a new method to solve some of the problems associated with the method of partial collectives.

2.1 METHODS AND DATA

A sedimentary sequence undergoing palynological analysis typically has radiocarbon dates obtained at the base and at significant discontinuities in the pollen assemblage series. Spatially widespread synchronous discontinuities should therefore correspond to major climate changes. These would be identified as modes within the multimodal frequency distribution of all radiocarbon dates with normally distributed variation.

2.1.1 DATA USED

We analyzed a set of 3076 ^{14}C dates obtained from the North American Pollen Database (Grimm, 2000; Appendix B) used to date sequences in over 700 pollen diagrams from across North America. We only analyzed ^{14}C dates from pollen diagrams because these are transitions in a proxy with similar lag times to climate changes.

Several versions of the data were used in order to test the robustness of the results. Basal, cultural impact (< 150 yr B.P.) and dates older than 12000 ^{14}C/yr B.P. were excluded, leaving 2372 ^{14}C dates.

Since the data are imprecise (that is, with a non null standard deviation for each date), we wanted to obtain results that were robust in presence of this imprecision. To do this, the dates were allowed to be perturbated by plus or minus one times their standard deviation uniformly distributed in that interval (Appendix C). Another way to take into account the imprecision of the data was to regroup dates in different classes of equal width, for example 0-200,200-400, and so on.

2.1.2 MIXTURE MODEL

Clustering of ^{14}C dates from the pollen sequences was solved using a nonlinear mixture modeling approach (Titterington *et al.*, 1985; Everitt and Hand, 1981). Normal curves were fitted to the multimodal distribution of all ^{14}C dates (n = 2372 observations). The model is based on expectation-maximization algorithms (Newton-Raphson algorithm). For n distributions the model contains $3n$ parameters: the expected value μ_i, the standard deviation σ_i, and the posterior probability of the distribution p_i, for $i = 1...n$, where p_i must sum to unity. This algorithm was implemented in SAS (Statistical Analysis Software) using macro-instructions (Appendix C). The model works in two steps. First, the FASTCLUS (cluster analysis procedure) produces initial values for the model parameters. Next, the NLP (nonlinear fitting procedure) procedure determined the optimal values of the parameters to maximize the log-likelihood ratio (NLL). An experimental dataset was developed in order to test the methodology, which consisted of six distinct modes with random variability around each mode. As expected, the log-likelihood curve flattens out at 6 to yield the optimal fit (Figure 2.1). Then, the mixture model was run on the original data where a typical execution of the algorithm included a loop on n (from 2 to 12). As expected, the log-likelihood ratio increases with n. At optimal solution, the log-likelihood (NLL) ratio curve flattens out, where adding new parameters does not greatly improve the fit, therefore yielding the optimal number of components or modes (Figure 2.2b). A standard deviation constraint was implemented to avoid distributions with very large standard deviations but was optimized to ensure that upper bounds were not too tight, which would lead to nonoptimal solutions. Clearly, there was a need to choose the right cutoff. Consequently, we ran several versions of the algorithm, in which the upper bound of the standard deviations varied between 500 and 1500 yr (Fig. 2.2c).

2.2 RESULTS

2.2.1 STATISTICAL ANALYSIS

The optimal solution found nine modes representing times in which many pollen diagrams across North America recorded vegetation changes (Fig. 2.1a). The dates chosen by the model are presented in Table 2.1.

We needed to ascertain that these events were really associated with transitions in pollen sequences, as the method is based on the assumption that palynologists tend to date transitions in their pollen sequences. We scoured the literature and examined the pollen diagrams associated with the identified modes and confirmed that dates typically corresponded to changes in the pollen assemblages. Another issue was the imprecision inherent in each ^{14}C date. These have a standard deviation, and we required robust results in the presence of this imprecision. Reanalysis with each date randomly perturbed by plus or minus one standard deviation (uniformly distributed in that interval) led to discrepancies caused only by extremely high standard deviations on a few dates. This approach, along with a variable bandwidth classification scheme, ensured consistency with the original analyses (Table 2.1).

Our null hypothesis is that randomly distributed dates throughout the study interval can account for the original distribution of dates. This was tested using a Monte Carlo approach to determine if the distribution of dates could be reproduced by chance alone. Results under randomization (100 times) show that the original distribution of dates cannot be reproduced by chance alone (99%). In addition, the original distribution of dates showed 7 significant modes but a mean of 3.8 under randomization (98%)(Figure 2.3).

To depict the spatial distribution of sites associated with each mode, all sites with a dated transition within ± 200 years of each mode were mapped (Figure 2.4). The difference in K functions (Rowlingson and Diggle, 1993) was then used to determine whether the mapped distribution of sites associated with each mode were a spatially random subset of the population of all ^{14}C sites in North America. Results show that only modes T8 and T9 departed (at distances > 100 km) and T1 and T4 marginally departed from spatially random subsamples (Figure 2.5).

Based on these analyses, we conclude that the original dated transitions are representative samples of transitions in North American pollen diagrams. Although many of these pollen transitions have been noted in regional summaries (see Bryant and Holloway, 1985), a continental-scale synthesis has not been previously attempted. The transitions identified here are synchronous with events found in ice cores and marine records further suggesting a large-scale climate cause (Table 2.2).

2.2.2 THE LATE GLACIAL

The pollen stratigraphic transition identified at 13 800 cal yr B.P. (T9) correlates with the Older Dryas-Allerød transition centered at ca. 13 500 cal yr B.P. (Anderson, 1997). This abrupt interstadial is associated with vegetation changes in most of Europe and North America (Anderson, 1997) and clearly evident in marine (Bond *et al.*, 1997; 2001) and Greenland ice cores (Dansgaard *et al.*, 1993). The 12 900 cal yr B.P. mode (T8) coincides with the transition between the Ållerød and Younger Dryas (Alley *et al.*, 1993; Johnsen *et al.*, 1992) when there is considerable evidence of a temperature decline in Western Europe (Anderson, 1997) and northeastern North America (Cwynar *et al.*, 1994). Previous syntheses and model simulations suggested that the Ållerød and Younger Dryas events were restricted to the North Atlantic (Manabe and Stouffer, 1995). Although the site distribution associated with these two modes shows vegetation transitions from all parts of North America (Fig. 2.4) our spatial analysis shows a tendency towards an eastern North American bias. Although our analysis does not provide an interpretation of the direction of the changes associated with each of these climate transitions (e.g., cooling, warming), it does indicate that the vegetation changed at these times across all of North America, and this is true of the other transitions presented here.

2.2.3 THE HOLOCENE

Transition 7 is dated at 10 190 cal yr B.P. and marks the beginning of the Boreal period (Table 2.2). One transition of particular interest is centered at 8100 cal yr B.P. (T6) and correlates with the 8200 yr B.P. cold event evident in the marine (Bond *et al.*, 1997, 2001) and Greenland GISP2 ice core (Alley *et al.*, 1997). In our analysis, this transition shows up as a distinct mode with little variation around the estimated mean (Table 2.1; Figure 2.2a). Although data from the ice core suggests a time span of only 400 yr with only half the magnitude of the Younger Dryas to Holocene transition (Alley *et al.*, 1997), our results suggest that this climate event induced changes in vegetation across North America (Fig. 2.4), supporting the hypothesis that this 8200 yr B.P. event was sufficient in magnitude to have had at least a hemispheric if not global impact (Alley *et al.*, 1997).

The middle Holocene mode (T5) is particularly broad even though constrained by the analysis. This climate transition has been identified globally (Steig, 1999) and in midwestern North America is a transition towards a warm and dry climate state (Winkler *et al.*, 1986). One reason for the large standard deviation is that analysts have historically tried to fit postglacial pollen diagrams into a three-part climate scheme and therefore dated the beginning and end of the "Hypsithermal Period" ignoring the variability within this period. In addition, the scale interaction of the millennial-scale events with the lower-frequency Milankovitch forcing may mean climate changes were more subdued during this time (Fisher, 1982; Gajewski, 1987). In general, the intervals between modal means in the middle Holocene are longer and less consistent between different proxies, some lasting more than 2000 yr (Table 2.2). During glacial times, some Dansgaard-Oeschger events (D/O) lasted anywhere from centuries to millennia (Dansgaard *et al.*, 1993), and the marine record also shows fewer ice-rafting debris (IRD) events in the middle Holocene (Bond *et al.*, 1997).

T4 is centered at 4030 cal yr B.P., consistent with the onset of Neoglaciation in the Northern Hemisphere (Denton and Karlén, 1973; Wendland and Bryson, 1974). This widespread transition is identified in pollen data from most parts of the world (Bryant and Holloway, 1985; Denton and Karlén, 1973; Wendland and Bryson, 1974; Harvey, 1979). It is associated with an IRD event in the North Atlantic (Table 2.2) and represents a cooling as a consequence of decreasing insolation in the Northern Hemisphere (Kutzbach, 1981). The next two modes at 2850 (T3) and 1650 cal yr B.P. (T2) are geographically widespread in North America (Figure 2.3) and found in different proxies (Table 2.2). These represent transitions towards cooler climatic conditions (Denton and Karlén, 1973; Bryant and Holloway, 1985; Harvey, 1979) at ca. 2800 yr B.P. followed by a climate amelioration around 1600 yr B.P. Evidence of vegetation change during these times has been recorded across North America (Bryant and Holloway, 1985; Harvey, 1979). T2 represents the beginning of a climate amelioration culminating in the maximum warming of the Medieval Warm Period 1000 yr ago (Bryant and Holloway, 1985).

Finally, T1, centered at 600 cal yr B.P., is interpreted as yet another reorganization of the atmospheric circulation culminating in the Little Ice Age, with maximum cooling 300 yr ago. The site distribution associated with this event suggests that this transition may have been widespread in North America, although the record becomes somewhat disturbed by human impact at this time (Figure 2.4). However, there is clear evidence of vegetation change in North America for this period (Gajewski, 1987; Nichols, 1967).

2.3 DISCUSSION AND CONCLUSIONS

This chapter presents a robust analysis demonstrating continental-scale synchronicity of the quasi-periodic 1500 yr cycle with marine and ice-core records. Furthermore, this result is based on hundreds of sites and thousands of radiocarbon dates, thus ensuring better dating of these events than previously possible. Because these climate changes are more extensive than previously identified, these results will be essential to climate modelers in determining the appropriate forcing.

The present findings also bring new insight into discussions of climate-vegetation interactions. During the Holocene, millennial-scale variations were sufficient to cause rapid changes in vegetation composition synchronously throughout North America. These vegetation changes are synchronous with transitions in marine and ice-core records, a discovery with significant implications for the long-running debate regarding whether vegetation was in equilibrium with climate changes of the late Quaternary. As such, these results clarify the issue of vegetation response to the rapid global warming predicted for this century. This large-scale synchronicity has implications for interpretation of past vegetation and for archaeological records as well. These results illustrate large-scale synchronicity of rapid changes between climate regimes in the Holocene and the late glacial, thereby suggesting impacts on societies through the past 10 000 yr.

In summary, the large-scale nature of these transitions and the fact they are found in different proxies confirms the hypothesis that Holocene and late-glacial climate variations of millennial-scale were abrupt transitions between climatic regimes as the atmosphere-ocean system reorganized in response to some forcing. Although several mechanisms for such natural forcing have been advanced, recent evidence points to a potential solar forcing (Bond *et al.*, 2001) associated with ocean-atmosphere feedbacks acting as global teleconnections agents. Because these transitions are identifiable across North America and presumably the world, the spatio-temporal evolution of the climate changes at this scale can be quantified. This exercise can lead to a better understanding of scale interactions within the climate system with implications for global warming scenarios due to human-induced CO_2 greenhouse gas forcing.

TABLE 2.1: Optimal modes identified in mixture model showing mean and standard deviations.

Components	Original	STD	Class	STD	Perturb	STD	Difference
T1	647	264	647	263	649	275	2
T2	1763	578	1688	506	1840	604	152
T3	2757	156	2679	302	2757	158	77
T4	3705	971	3764	835	3741	734	59
T5	5908	1000	5845	1000	5731	1000	178
T6	7259	66	7260	72	7278	105	19
T7	9003	978	9019	994	8973	974	47
T8	10749	527	10775	525	10643	482	132
T9	11771	83	11778	87	11596	243	182

Note: Dates are presented in radiocarbon years
Class represents modal values identified when original data are classified under different bin width (100 to 300).
Perturb represents modal values identified when original dates are randomly perturbed by ± 1 times their individual standard deviaton error.
Difference represents variations between modal values under different scnerarios.

TABLE 2.2: Millennial-scale climate transitions for the Holocene and Late Glacial identified in several studies.

Period *	Terrestrial						Marine **	GISP2††	Significant discontinuities	
	Botanical This study	Alpine glaciers†		Botanical § Major	Botanical § Minor	Regional Syntheses #			Date	Diff.
LIA maximum		330				300	300			
Pacific	(T1) *600*			*730*		*700*		*600*	*650*	
neo-Atlantic			Exp							
Scandic	(T2) 1650		Con	1550		1600	1400	1500	1550	900
sub-Atlantic	(T3) 2850	2800	Exp	2780		2600	2800	2750	2800	1250
sub-Boreal 1	(T4) 4030		Con		4000	4000	4200		4050	1250
sub-Boreal		5300	Exp	5900	5000	5500	5900	5550	5500	1450
Atlantic 2	(T5) 6700		Con		6850			6300	6600	1100
Atlantic 1	(T6) 8100	7800	Exp		8100	7500	8100	8100	8100	1500
Atlantic				9500			9400	9300	9400	1300
Boreal	(T7) 10190	10300		10450		10000	10300	10600	10300	900
pre-Boreal			Con	11500		11500	11100	11300	11300	1000
Younger Dryas	(T8) 12900	12800	Exp	12800		12800	12600		12800	1500
Allerod	(T9) 13800		Con						13800	1000
Late Glacial			Exp							

Note: Radiocarbon dates were calibrated to calendar years using INTCAL98 Calib 4.2 software (Stuiver and Reimer, 1993).
Significant discontinuities indicate synchronous dates present in at least 2 differemt records (eg. Marine, ice and one of terrestrial).

*Blytt-Sernander names assigned to these periods following Wendland and Bryson (1974) and used here for comparison use only.

†From Denton and Karlén (1973).

§From Wendland and Bryson (1974).

#From Bryant and Holloway (1985).

**From Bond et al. (1997, 2001).

††From O'Brien et al. (1995).

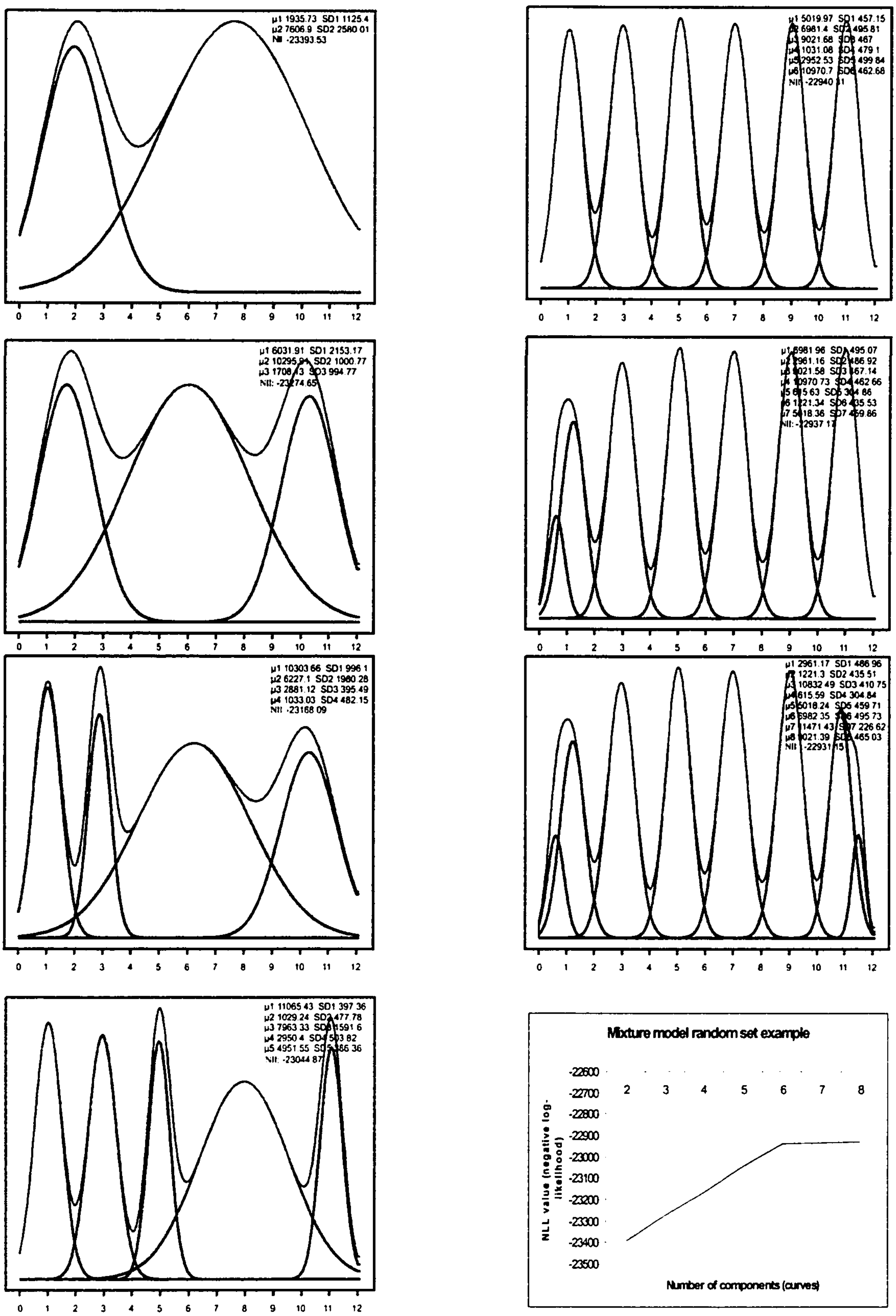


Figure 2.1: Mixture modeling results on random set of dates with 6 distinct modes.

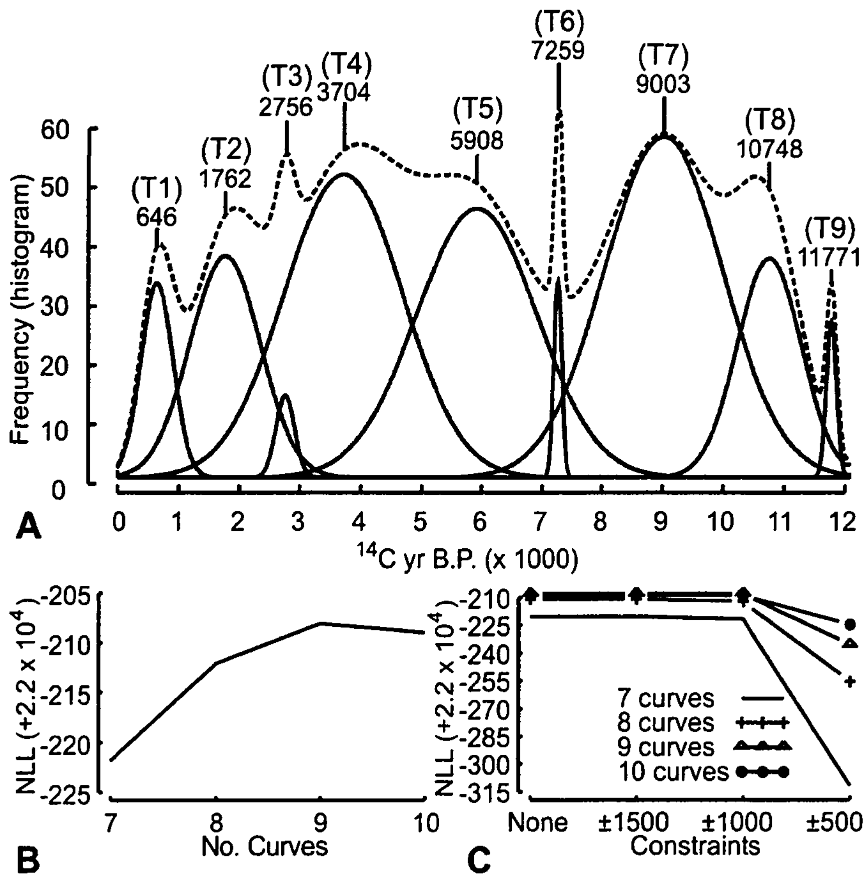


Figure 2.2: Model output and sensitivity results. A: Transitions identified; B: Maximum likelihood statistic; C: Standard deviation contraints (see text for explanation)

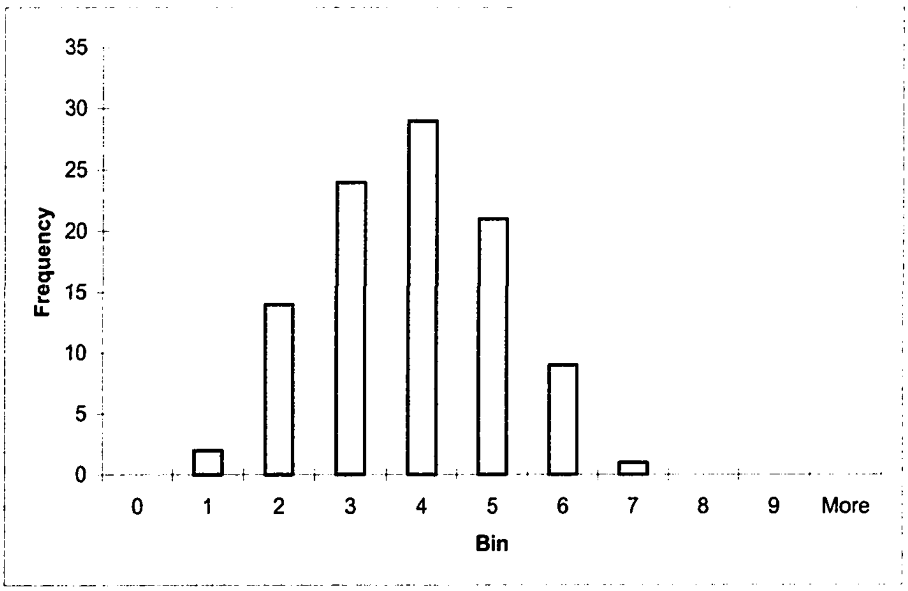


Figure 2.3: Significant peaks identified under randomization (100 times) (see text for explanation).

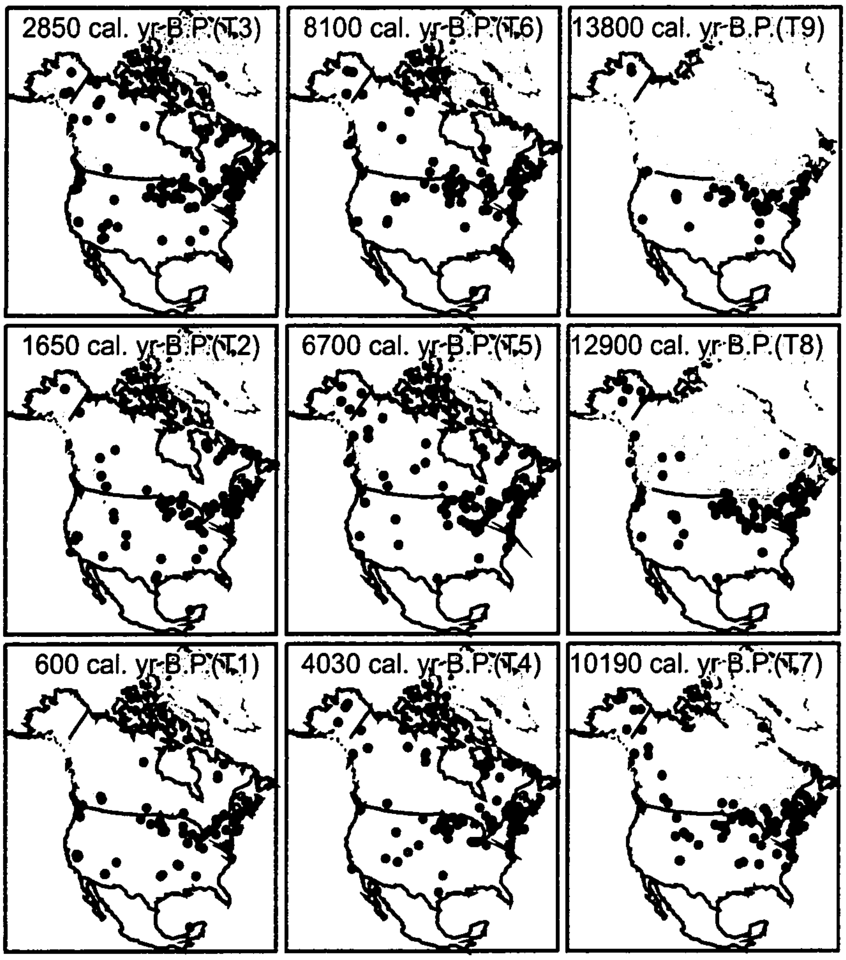


Figure 2.4: Site distributions for synchronous times of environmental change identified in this study. Site locations containing dated transitions (+/- 200 yr) identified in this study were extracted from North American Pollen Database (Grimm, 2000).
Ice-sheet contours were obtained from Dyke and Prest (1987).

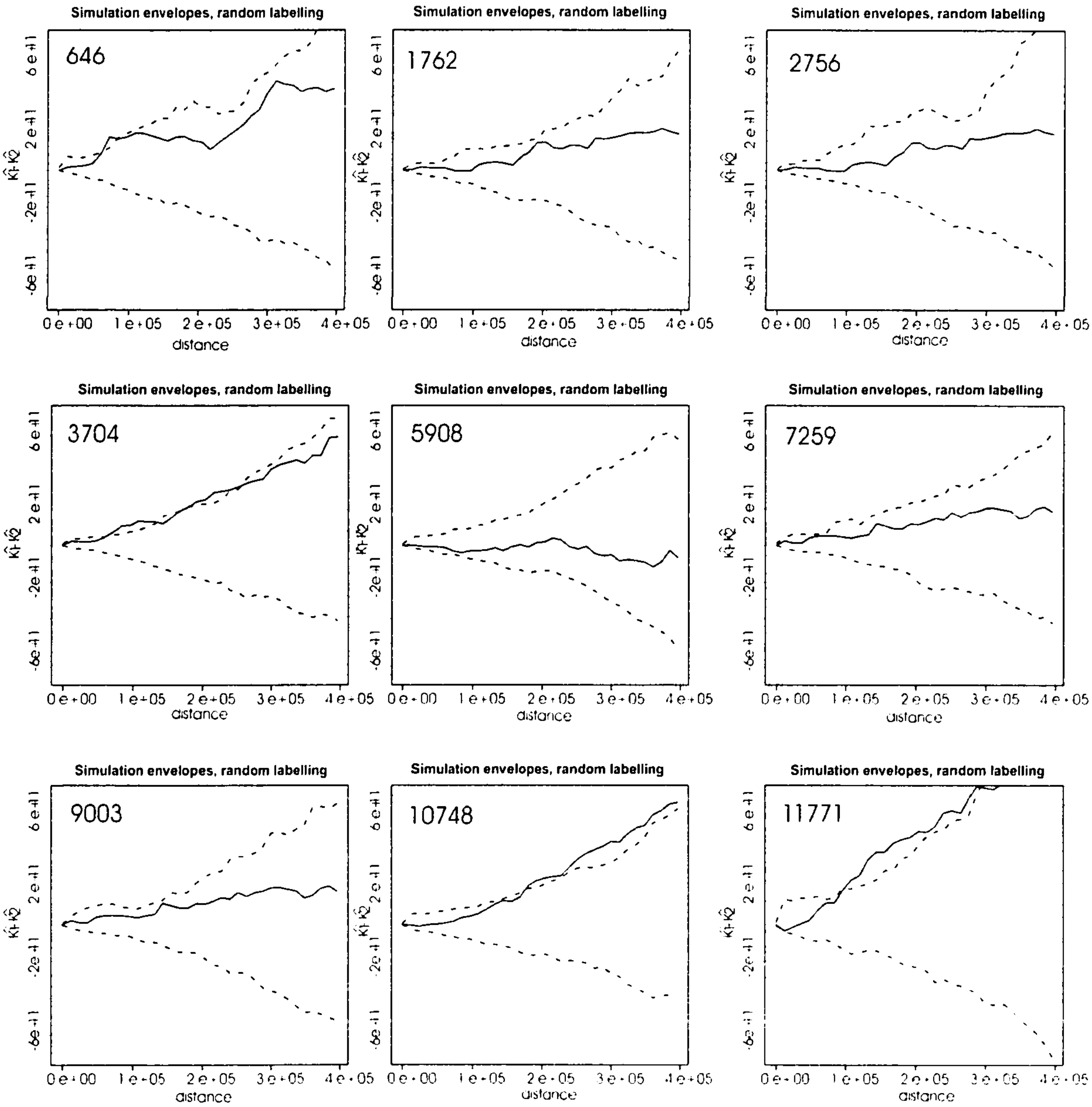


Figure 2.5: The difference in K functions (K_{1-2}) (Ripley 1976; Rowlingson and Diggle, 1993) was used to test if the spatial distributions in Figure 2.4 could be considered as random subsets of the heterogeneous spatial arrangement ^{14}C sites within North America. We computed K_1 from all n_1 points composing a given mode and K_2 from all n_2 sites with ^{14}C dates older than the lower age bound of the mode considered. We then determined the difference in K functions ($K_{1-2} = K_1$-K_2) up to a distance of 700 km. Monte Carlo upper and lower simulation envelopes of K_{1-2} were estimated in 100 repeated simulations using the fixed n_1+ n_2 locations but assigning randomly n_1 of these locations as modal events.

Chapter 3

POLLEN TIME SERIES AND MILLENNIAL-SCALE CLIMATE VARIABILITY OF NORTH AMERICA DURING THE PAST 14,000 YEARS

Abstract

Although pollen time series have been widely used to infer past vegetation and climate changes at regional and local scales, little research has focused on continental scales. In this chapter, millennial-scale climate variability in North America was investigated using Principal Component Analysis (PCA) of 125 pollen time series. Changes in vegetation, and thus climate, were quasi-periodic during the past 14,000 calendar years BP. The PCA revealed community-level changes in vegetation that occurred synchronously across all of North America. These results are consistent with times of major climate transitions identified in the Greenland ice sheet and North Atlantic marine paleoclimate proxy records. Millennial-scale climate variability was synchronous and widespread which is most likely associated with reorganizations of the ocean-atmospheric circulation patterns.

3. INTRODUCTION

Chapter 2 suggested common modal times of vegetation change across North America during the past 14,000 years. However, these results were based solely on the distribution of ^{14}C dates under the assumption that palynologists preferentially dated these transitions. Here we use the actual pollen assemblages to explore vegetation response to millennial-scale.

Millennial-scale climate variability has been suggested to be a pervasive feature of our climate system, at least in the North Atlantic region and during the past millions years (Bond *et al.*, 1997, 2001; MacManus *et al.*, 1999; O'Brien *et al.*, 1995; Stuiver *et al.* 1997; Oppo *et al.* 1998; Raymo *et al.* 1998; and Bianchi and McCave, 1999). A central unanswered question concerning millennial-scale climate variability is the geographical extent of these changes. One hypothesis is that millennial-scale climate variability is a North Atlantic regional thermohaline circulation phenomenon (Manabe and Stouffer, 1995). The alternative is that it is widespread, perhaps even global in extent (Alley *et al.*, 1997). Although several individual multi-proxy record sites from around the world have correlated their records to those from the North Atlantic, it remains unclear as to how widespread millennial-scale climate variability has been in the past or what are the underlying causes.

Records of past vegetation are found in lake sediments, bogs and other terrestrial environments, from which pollen grains can be extracted and identified. Since pollen grains are well mixed in the air, the fossil record consists of an integration of the pollen from the "catchment area". A pollen count from sediment is generally a statistically valid sample of the contemporary pollen in the air above an area (Watts, 1973) and therefore pollen assemblages are related to the surrounding vegetation (Birks and Birks, 1980). Studies have shown that there is a strong relationship between pollen and vegetation at regional and continental scales (Davis and Webb, 1975). Comparisons between isopoll and vegetation maps for eastern North America and the Northwestern boreal forest show that pollen assemblages are related to the major vegetation types in the area (Delcourt *et al.*, 1984; Delcourt and Delcourt, 1985, Anderson *et al.*, 1991).

Vegetation also has a strong relationship with modern climate conditions, at least at regional and continental scales. Vegetation boundaries have been shown to be constrained by climatic conditions prevailing over certain regions. For example, the northern limit of the circumpolar Boreal forest coincides well with the mean winter position of the polar front (Bryson, 1966) and oak exists within the range of maritime tropical air over southeastern North America.

Pollen diagrams can be used to determine changes in vegetation through time if considerations such as long distance transport and over-and under-representation of certain species (Davis and Webb, 1975) are accounted for. The time required for vegetation to respond completely to climatic forcing at a time-scale of 10^4 years is unknown, but records of the vegetation response to climatic variability of century to millennial scales provide evidence for sufficiently rapid response (e.g. Webb, 1986; Gajewski, 1987; Overpeck, 1987; Grimm *et al.*, 1993), suggesting that pollen diagrams can be interpreted as proxy-climate records.

One approach to identify times of climate change at century to millennial-scales using vegetation data was proposed by Gajewski (1987, 1993) who analyzed vegetation response to century-scale climatic events such as the Little Ice Age using Principal Component Analysis (PCA). By using a spatial array of sites associated with different plant communities, he was able to show the timing of plant community response to the Little Ice Age (1450-1850 A.D.). Incorporation of both time and space dimensions was essential to identify climate impacts. The use of PCA at individual sites may not warrant an interpretation of a climate change causing a subtle change in vegetation, however if several sites in a region associated with different plant communities show a synchronous change in vegetation, whether subtle or not, this would support an interpretation of climate control.

3.1 METHODS AND DATA

In this study, we use PCA for reasons proposed by Gajewski (1987), but at a different spatial and temporal scale. In addition, this analysis explicitly analyzes the time, space and taxonomic axes simultaneously. A PCA was implemented (Appendix D) to explore common times of change between the different pollen assemblages across North America and during the past 14,000 years BP.

3.1.1 DATA

Published and unpublished radiocarbon-dated pollen time series from a dense network of sites from the North American Pollen Database (NAPD) (Grimm, 2000) are used as proxy for relative paleovegetation and paleoclimatological changes through the Holocene and Late Glacial. A second dataset of high-resolution fossil pollen time series (Gajewski, 1987) is used to explore the past 1,000 years and validate the automated methodology used in this study.

3.1.1.1 TEMPORAL RESOLUTION

Age-depth models were extracted from the NAPD (Grimm, 2000) at each fossil pollen site. The pollen times series were then linearly interpolated at 250 yr intervals, yielding a total of 48 observations for the period of study (250-14,000 cal/yr BP) (Appendix E). This sampling interval was high enough to ensure satisfactory analysis at the millennial-scale. The higher resolution dataset typically has a temporal resolution of 40 years for the past 1,000 years (Gajewski, 1987).

3.1.1.2 SPATIAL RESOLUTION

We extracted only pollen sequences with continuous records for the past 12,000 radiocarbon-years. This selection criterion resulted in 125 sites. The selected sites are unevenly distributed throughout the continent, where some regions such as the Great Lakes region are well represented while other regions have fewer data (Figure 3.1; Table 3.1). The site distribution representing the 40-year resolution dataset is presented in Figure 3.2, which consists of an east-west transect in northeastern North America.

3.1.1.3 TAXONOMIC RESOLUTION

The analyses were done several times using different pollen sums to test different methodological assumptions and scenarios and attain robust results (Appendix F). The taxonomic resolution or pollen sums varied from high (47 taxa)(Table 3.2) to low (only trees and herbs) (Appendix F).

3.1.2 PRINCIPAL COMPONENT ANALYSIS

PCA has been widely applied to approach multivariate questions in geographical and ecological studies. A novel approach using PCA, as explained below, was implemented in SAS (Statistical Analysis Software) (Appendix D) in the time, space and taxonomic dimensions to explore the spatial and temporal correlations between different pollen time series across North America and to extract common times of change during the past 14,000 cal/yr BP.

PCA is a technique from the family of factor analysis (Legates, 1991). The main applications of factor analysis techniques are: (1) to reduce the number of variables by considering the correlation between the variables and (2) to detect structure in the relationship between variables. In essence, a PCA identifies variance common and unique in time and space and extracts this information as new variables called components, which summarize the major spatio-temporal patterns that exist in the data (Legates, 1991).

The rates and timing of pollen, hence vegetation and by inference climate change, were examined using a novel methodological framework using PCA. In order to simultaneously compare different pollen time series from all sites and to extract a common signal in both time and space, we merged the pollen proportions of all 125 sites into one axis as variables while keeping time independent on the second axis of the input matrix. This resulted in a two-dimensional representation of a 3-dimensional matrix where taxa and location would be simultaneously correlated with each other through time. Essentially, this approach enabled a comparison of the variation through time between each taxon at all sites extracting common temporal explained variance. This merging procedure resulted in a 2-dimensional matrix of 5875 variables (i.e. sites x taxa) by 48 time periods.

The inherent complexity of this matrix (i.e. 2-D representation of a 3-D matrix) resulted in too many columns in relation to the number of rows. This can lead to an "ill-conditioned matrix", where the inverse correlation matrix may contain negative values. Although this does not necessarily invalidate the analysis, it is desirable to avoid this potential instability. Bootstrapping was used to artificially increase the number of rows of the input data matrix to avoid this potential problem. Each taxon proportion through space and time were replicated, with some random variability, in the time domain until the number of rows (i.e. time) equaled the number of columns (i.e. taxa and space) in the input matrix. Considering the proportional nature of the data, the bootstrap was ensured to have no impact on the computation of the correlation matrix by remaining at .00001% of a centroid error radius of each taxon proportion value (Figure 3.3). In the case of zero proportion values, the bootstrap was constrained to retain zero as the proportion value. In addition, all columns or variables summing up to zero were excluded prior to the PCA analysis to avoid a singular matrix. Comparison between the analyses of the bootstrapped matrix versus the original matrix revealed no change in factor scores and loadings.

3.2 RESULTS

3.2.1 PCA RESULTS

Ten factors were retained to depict the spatio-temporal structures of vegetation change, hence climate change, at the continental-scale of North America during the past 14,000 years. The relationship between tree and herb proportions was used to determine the appropriate number of factors to retain. When the proportion of trees increase, the resulting herb proportions will decrease and vice-versa (i.e. trees = minus herbs) at all sites. This relationship among several others was tested in order to achieve robust results and is documented in Appendix F.

The ten factors summarize 66% of the total explained variance in the pollen time series distribution. In this study, factor scores show the time evolution of the common variance of all pollen assemblages while factor loadings depict spatial information on taxa and site contributions. For example, a particular factor score depicts common temporal variance between all variables, in this case taxa in space, while factor loadings will give insight on which taxa and sites contribute the most to the factor score.

All ten factor scores (i.e. solid line) were summed to illustrate the combined temporal evolution of the common variance explained in North American pollen time series during the past 14,000 years (Figure 3.4). Each individual factor scores (i.e. broken lines) is illustrated to show its iterative contribution to the total (Figure 3.4). Factor loadings are used to depict the spatial coherency between the pollen time series distribution. The numbers of sites that are significantly correlated with these factors are shown in Table 3.3. This serves to efficiently summarize the extent to which these factors reflect continental-scale variability at the scale of the factor scores. Significant factors are considered where loading is above the 9[th] percentile (positive) or below the 1[st] percentile (negative)(Table 3.3). The number of sites for which a particular taxon has a significant loading was summed as an index of the importance of the taxon for each particular component. For example, *Abies* or fir is correlated significantly at 26 sites with factor 1 (Table 3.3). This was done in order to evaluate the relative importance of each taxon to the overall principal component. Next, in order to assess the site contributions to the overall component, taxon loading above the 9[th] percentile or below the 1[st] percentile was again extracted at each individual site. This time the number of taxa with significant loadings were summed for every site (Table 3.4). For example, site 1 has 13 taxa that significantly contribute to factor 1 irrespective of the biological reason for this change or direction of change (i.e. increase or decrease in abundance)(Table 3.4).

In general, the first few factor score time series plots reveal long-term trends while higher-order factors reveal higher-frequency variability. Because principal components are additive, these higher-frequency variations are superimposed on the long-term trends (Figure 3.4). The first two factor scores depict long-term trends in vegetation and climate change that are likely related to deglaciation and astronomical forcings. Basically, herbs were not alone anymore as trees expanded at the continental-scale of North America. The significant taxa are those major components of the vegetation, for example *Picea*, *Pinus*, *Quercus*, *Acer*, *Fagus*, *Artemisia*, *Cyperaceae* and *Poaceae* (Table 3.3). Virtually all sites are correlated with these factor scores (Table 3.4). Factors 3 and 4 suggests higher temporal variability during the late phase of deglaciation while factors 5 and 6 illustrate the mid-Holocene warm period and subsequent neoglacial cooling trends, which are likely associated with secular changes in insolation in the northern hemisphere (Figure 3.4). Factors 3 through 6 remain significantly correlated with sites across all of North America (Table 3.4) and the major trees and herbs listed above continue to load significantly at numerous sites. Other taxa, for example *Ambrosia*, *Asteraceae*, *Chenopodiaceae*, *Cupressaceae*, *Ericacaeae*, *Fraxinus*, *Populus*, *Myrica*, *Ostrya/Carpinus*, *Tsuga*, *Ulmus*, and *Salix*, also show significant (1^{st} and 9^{th} percentile) loadings at several sites (Table 3.3). Scores of factor 7 show higher-frequency variability during the early Holocene and late glacial while factor scores 8, 9 and 10 depict higher-frequency explained variability during the mid-and late Holocene (Figure 3.4). Again many taxa are contributing significantly to factors 7, 8, 9 and 10 (Table 3.3) and across all of North America (Table 3.4). In summary, these results illustrate the well-documented long-term trends evident in individual pollen time series. However, as demonstrated by this analysis, superimposed on these long-term trends are higher-frequency variability common at most sites that may not be clearly evident at individual site locations. All pollen taxa and not just a few of the most abundant types caused this variability. The incorporation of time and space reveal these higher-frequency vegetation transitions. Because these are expressed throughout North America, these must be caused by climate changes.

A second set of high-resolution pollen time series was used to analyze in more detail the last 1,000 years and also to verify the methodology used in this analysis. Because there are fewer data and they have been extensively studied before (Gajewski, 1987), we wanted to be assured that the scale relations found are not just artifacts of the method. Again, the PCA separates by scale, where the first factor score illustrates a long-term trend associated with neoglacial cooling (Gajewski, 1987) (Figure 3.5). The second factor resolves the Little Ice Age and preceding Medieval Warm Period as previously demonstrated using PCA at individual sites along an east-west transect in northeastern North America (Gajewski, 1987) (Figure 3.5). The remaining factor scores illustrate even higher-frequency temporal explained variance although it was not evaluated for significance.

The majority of taxon contributed to some degree to the overall scores throughout North America (Table 3.3). Major tree and herb taxa, such as *Picea*, *Pinus*, *Quercus*, *Acer*, *Ambrosia*, *Artemisia*, and *Poaceae* (Appendix F1.5, F1.6), continually loaded significantly at many sites in all ten factors (Table 3.3). The number of sites having these major taxa in their respective pollen records may have biased some of these results. Conversely, these major taxa are found in most regions of North America and the fact that they contributed significantly to the overall factor scores suggests a common signal at the continental-scale. Although minor taxa, such as *Cephalanthus*, *Liliaceae*, *Magnoliaceae*, *Nyssa*, and *Rumex* among others (Appendix F1.7, F1.8) do not show up significantly in many sites does not preclude that they did not contribute significantly to the overall factor scores through significant contribution at a few sites (Table 3.3). These types have a limited geographical distribution and therefore are only important at regional scales. Virtually all sites contributed significantly to the first 10 components irrespective of their pollen composition or direction of change (i.e. increase or decrease in taxa abundance).

Although beyond the scope of this study, these results suggest community-level changes in pollen assemblages occurring synchronously across all of North America. The loadings between different taxa could further be assessed on local and regional scales for additive insight into the nature of these climate impacts. In summary, these results suggest common times of community-level vegetation changes across North America in response to climate change at low and high frequencies during the past 14,000 cal/yr BP.

3.2.2 EVENT CORRELATION

Quantitative spatio-temporal correlation of the 125 pollen records was used to identify times when vegetation changes occurred synchronously across North America (Figure 3.6 and Table 3.5). The results suggest that times of vegetation change in North America were not randomly distributed through time, but instead were centered on certain dates (Figure 3.6). Vegetation change was more substantial during the period of deglaciation before 10,000 years BP, as expressed by the higher-order factor scores, which translates to higher explained variances (Figure 3.4). Suggestions of subtler climate changes during the Holocene are supported by lower-order factor scores (i.e. lower explained variances) (Figure 3.4). Therefore, these results support the hypothesis that millennial-scale climate variability was more intense during the last deglaciation and subtler during the Holocene (Overpeck, 1987; Grimm and Jacobson, 1991; Huntley *et al.*, 1992; Bond *et al.*, 1997, 2001; Bianchi and McCave, 2001). Conversely, the potential for vegetation to respond will be greater during the degalciation than during the Holocene. During the late glacial period, the well-documented Bolling-Ållerød, Younger-Dryas, and Holocene transition are known to be abrupt climatic changes (Dansgaard *et al.*, 1989; Broecker *et al.*, 1989; Lehman and Kegwin, 1992; Taylor *et al.*, 1997; Alley *et al.*, 1993).

Correlations between different proxy records also reveal unequivocal times of climate change at the millennial-scale (Table 3.6). For example, this study identifies the Ållerød (T10), Younger Dryas (T9), and Boreal (T7) oscillations that correlate well with the marine, GISP2 and other terrestrial climate-proxy records (Table 3.6). This study has also identified the well-documented pre-Boreal transition (T8) as having an impact on the vegetation. This particular change was missed by the analysis of Chapter 2. Transitions at 8200 (T6) and 6800 (T5) cal/yr BP were identified here as well as in other proxy-climate records (Table 3.6). A climate change is revealed at *circa* 5300 years BP (T4) by the PCA, as well as in different proxy-climate records, although it was missed in Chapter 2 (Table 3.6). As stated previously, this is probably because palynologists used to preferentially date a three-part Holocene, neglecting mid-Holocene variability (*circa* 7000-4000 cal/yr BP). One transition that was not clearly identified in this study is *circa* 4000 cal/yr BP, which is associated with northern hemispheric neoglacial cooling. However, there are some indications in this analysis that support this transition as a progressively decreasing trend in the resulting factor score curve (Figure 3.6). The late Holocene transitions at circa 3000 (T3), 1300 (T2), and 600 (T1) cal/yr BP also correlate reasonably well with times of climate change in many proxy-climate records (Table 3.6).

3.3 DISCUSSION AND CONCLUSIONS

This chapter presents a robust analysis demonstrating continental-scale synchronicity of vegetation response to millennial-scale climate variability. The results of principal components analysis compare well with independent paleoclimate proxy records. Periods of rapid change in pollen diagrams identified in this study closely compare with the marine and ice core records of the North Atlantic. However, there remains some ambiguity in the timing of some of these transitions. For example, the onset of the Atlantic period circa 9400 years cal/yr BP is clearly evident in the GISP2 and marine record but not in the terrestrial record (Table 3.6). The 4000 cal/yr BP event identified in the marine and terrestrial record is absent in the GISP2 record while the *circa* 6700-year event is absent in the marine record (Table 3.6). These periods correspond to interactions of millennial-scale variability with lower-frequency northern hemispheric insolation (orbital parameters) changes during the past 10,000 years.

This raises questions about the origin of the hypothesized 1,500-year cycle. Were these transitions due to some millennial-scale climate forcing, for example long-period variations in solar output, or were they the result of the response of the atmosphere to northern hemispheric insolation change due to orbital configuration (eg. Milankovitch parameters), or a combination of both? The quasi-periodic nature of millennial-scale climate variability still remains enigmatic as it may result from interactions between orbital and century to millennial-scale climate forcings. It remains to be seen what causes millennial-scale climate variability. Moreover, the fact that vegetation changes synchronously across all of North America supports the hypothesis that millennial-scale climate variability is sufficiently large to cause widespread changes in surface climates.

Fossil pollen time series can be used to infer millennial-scale climate variations at the continental-scale and most likely at century-scale frequencies as described in the high-resolution analysis of the past 1,000 years (e.g. Gajewski, 1987). In order to expand our knowledge of Holocene climate variability using fossil pollen time series, the precision and accuracy of age models need to be improved. Dating control needs to be improved significantly if we acknowledge that fossil pollen sedimentary sequences have the ability to record century-to millennial-scale, and perhaps even decadal scale variability. An increased sampling interval may also be the key to improving the quality of the climate reconstructions using these data.

In summary, the large-scale nature of these millennial-scale climate oscillations and the fact they are found in different proxies confirms the hypothesis that Holocene and late glacial climate variations of millennial-scale were abrupt transitions between climatic regimes, as the atmosphere-ocean system reorganized in response to some forcing. There is still debate concerning the true nature of this so-called quasi-periodic cycle of 1,500 ± 500 years, or if indeed it is periodic. Although several mechanisms for such natural forcing have been advanced, such as such as the role of ocean-atmosphere-ice interactions

(Broecker and Denton, 1989), variable solar output (e.g. Perry and Hsu, 2000; Crowley, 2000; Bond *et al.*, 2001; Van Geel *et al.*, 1999), orbital harmonics (Pestiaux *et al.*, 1988), nonlinear feedbacks within the climate system (Peixoto and Oort, 1992; Rind, 1999), and volcanism (Zielinski *et al.*, 1996, Crowley, 2000), there is no clear indication as to which one or combination are responsible for millennial-scale climatic fluctuations. If the Earth is indeed warming from the cold times of the LIA, it becomes imperative that we understand these natural decadal to millennial-scale climate dynamics. An increase in atmospheric CO_2 concentrations will inevitably play a major role not only at seasonal, annual and inter-annual scales but most likely may overwhelm the natural climate variability at decadal to millennial-scales of our future climate evolution.

Table 3.1. Sites from NAPD contain containing continuous pollen time series from 0-12,000 radiocarbon years (125 sites).

Id	E#	SiteName	lat	long	Id	E#	SiteName	lat	long	Id	E#	SiteName	lat	long
1	969	Devon Island Glacier	75.4	-82.5	43	883	Cub Lake [Idaho]	44.1	-111.2	85	199	Pine Bog	41.3	-74.8
2	514	Kaiyak Lake	68.2	-161.4	44	859	Divide Lake	43.9	-110.2	86	690	No Bottom Pond	41.3	-70.2
3	521	Screaming Yellowlegs Pond	67.6	-151.4	45	862	Hedrick Pond	43.8	-110.6	87	585	Fresh Pond	41.2	-71.6
4	519	Redstone Lake	67.3	-152.6	46	464	Sherd Lake	44.3	-107.0	88	259	Tannersville Bog	41.0	-75.3
5	508	Ranger Lake	67.1	-153.7	47	288	Wolsfeld Lake	45.0	-93.6	89	192	Panther Run Pond	40.8	-77.4
6	506	Angal Lake	67.1	-153.9	48	473	Cottonwood Lake	44.8	-99.9	90	153	Longswamp	40.5	-75.7
7	515	Joe Lake	66.8	-157.2	49	189	Lake West Okoboji	43.3	-95.2	91	257	Szabo Pond	40.4	-74.5
8	520	Rebel Lake	67.4	-149.8	50	691	Clear Lake [Iowa]	43.2	-93.4	92	447	Gould Pond	45.0	-69.3
9	518	Ped Pond	67.2	-142.1	51	462	Roberts Creek	43.0	-91.5	93	174	Moulton Pond	44.6	-68.6
10	3500	Sleet Lake	69.3	-133.6	52	52	Colo Marsh	42.0	-93.3	94	240	Silver Lake [N.S.]	44.6	-63.6
11	3501	Reindeer Lake	69.1	-132.2	53	82	Edward Lake	44.4	-80.3	95	443	Ross Pond	43.9	-69.4
12	1551	Tuktoyaktuk 5	69.1	-133.5	54	1157	Kellners Lake	44.2	-87.8	96	1094	Swamp Lake	38.0	-119.8
13	1259	Hanging Lake	68.4	-138.4	55	92	Gass Lake	44.1	-87.7	97	695	Barrett Lake	37.6	-119.0
14	1549	Maria Lake	68.3	-133.5	56	1083	Wylde Lake	43.9	-80.4	98	882	Exchequer Meadow	37.0	-119.8
15	1552	Lateral Pond	66.0	-135.9	57	1155	Hanson Marsh	43.4	-89.7	99	1313	Head Lake	37.7	-105.5
16	522	Wien Lake	64.3	-151.3	58	72	Devils Lake [Wisconsin]	43.4	-89.7	100	880	Como Lake	37.6	-105.5
17	854	Compass Pond	50.0	-56.2	59	276	Vestaburg Bog	43.4	-84.9	101	868	Cheyenne Bottoms	38.5	-98.7
18	850	Davis Lake	46.6	-122.3	60	1056	Brophy Ditch	43.2	-81.7	102	64	Cupola Pond	36.8	-91.1
19	445	Carp Lake	45.9	-120.9	61	286	Wintergreen Lake	42.4	-85.4	103	441	Powers Fort Swale	36.6	-90.6
20	857	Battle Ground Lake	45.8	-122.5	62	87	Frains Lake	42.3	-83.6	104	57	Cranberry Glades	38.2	-80.3
21	849	Guardipee Lake	48.6	-112.7	63	1060	Lake Erie	41.9	-82.8	105	115	Jackson Pond	37.5	-85.7
22	1970	Forest Lake	46.5	-112.2	64	50	Clear Lake	41.7	-86.5	106	8	Anderson Pond	36.0	-85.5
23	94	Glenboro Lake Site	49.4	-99.3	65	205	Pretty Lake	41.6	-85.3	107	60	Crider's Pond	40.0	-77.6
24	1085	Irvin Lake	47.1	-93.6	66	925	Ladd Lake	41.4	-84.8	108	34	Buckle's Bog	39.6	-79.3
25	290	Wood Lake	45.3	-90.1	67	241	Silver Lake [Ohio]	40.4	-83.8	109	1148	Browns Pond	38.2	-79.6
26	143	Lily Lake [US:Minnesota]	45.1	-92.8	68	265	Torren's Bog	40.4	-82.5	110	390	Quarles Pond	38.0	-79.1
27	58	Crates Lake	49.2	-81.3	69	924	Fudger Lake	40.1	-83.5	111	100	Hack Pond	38.0	-79.0
28	261	Petit Lac Terrien	46.6	-70.6	70	467	Tonawa Lake	44.9	-77.2	112	220	Rockyhock Bay	36.2	-76.7
29	78	Lac Dufresne	45.9	-70.4	71	104	Heart Lake [New York]	44.2	-74.0	113	1334	San Agustin Plains	33.9	-108.3
30	271	Unknown Pond	45.6	-70.6	72	1077	Porqui Pond	44.2	-79.8	114	1336	San Agustin Plains	33.9	-108.3
31	266	Lac à la Tortue	45.5	-73.3	73	1058	Cranberry Lake	44.1	-78.1	115	215	Rayburn's Dome	32.5	-93.2
32	2339	Graham Lake	45.2	-77.4	74	68	Deer Lake Bog	44.0	-71.8	116	30	Boriack Bog	30.4	-97.1
33	486	Lac du Diable	48.9	-66.1	75	448	Poland Spring Pond	44.0	-70.4	117	197	Pigeon Marsh	34.7	-85.2
34	16	Basswood Road Lake	45.3	-67.3	76	275	Van Nostrand Lake	44.0	-79.4	118	281	White Pond	34.2	-80.8
35	145	Little Lake [New Bruns.]	45.1	-66.7	77	228	Sandogardy Pond	42.8	-71.7	119	37	Cahaba Pond	33.5	-86.5
36	856	Leading Tickles	49.5	-55.5	78	13	Balsam Lake	42.5	-74.6	120	95	Goshen Springs	31.7	-86.1
37	1078	Robinson's Pond	48.3	-58.8	79	7	Allenberg Bog	42.3	-78.9	121	38	Camel Lake	30.3	-85.0
38	1069	Joes Pond	48.2	-58.2	80	18	Belmont Bog	42.3	-77.9	122	461	Clear Pond [S.C]	33.8	-79.0
39	866	Slough Creek Pond	44.9	-110.4	81	105	Heart's Content Bog	42.2	-74.0	123	239	Sheelar Lake	29.5	-82.0
40	884	Gardiner's Hole	44.9	-110.7	82	77	Duck Pond	41.9	-70.0	124	268	Lake Tulane	27.6	-81.5
41	858	Cygnet Lake Fen	44.7	-110.6	83	247	Spring Lake	41.7	-76.4	125	147	Lake Annie	27.2	-81.4
42	865	Mariposa Lake	44.2	-110.2	84	222	Rogers Lake	41.4	-72.1					

Table 3.2. Pollen taxa used in PCA

v1	*Abies*	v26	*Ostrya/Carpinus*
v2	*Acer*	v27	*Picea undiff.*
v3	*Alnus undiff.*	v28	*Pinus undiff.*
v4	*Ambrosia type*	v29	*Plantaginaceae*
v5	*Apiaceae*	v30	*Poaceae*
v6	*Aquifoliaceae*	v31	*Polygonum*
v7	*Artemisia*	v32	*Polypodiaceae*
v8	*Asteraceae*	v33	*Populus*
v9	*Betula undiff.*	v34	*Pteridium*
v10	*Carya*	v35	*Quercus*
v11	*Castanea*	v36	*Ranunculaceae*
v12	*Cephalanthus*	v37	*Rubiaceae*
v13	*Chenopodiaceae/Amaranthaceae*	v38	*Rumex*
v14	*Corylus*	v39	*Salix*
v15	*Cupressaceae*	v40	*Sarcobatus*
v16	*Cyperaceae*	v41	*Saxifragaceae*
v17	*Ericaceae undiff.*	v42	*Shepherdia*
v18	*Fagus*	v43	*Sphagnum*
v19	*Fraxinus*	v44	*Taxodium*
v20	*Larix/Pseudotsuga*	v45	*Tilia*
v21	*Liliaceae*	v46	*Tsuga*
v22	*Lycopodium undiff.*	v47	*Ulmus*
v23	*Magnoliaceae*		
v24	*Myrica*		
v25	*Nyssa*		

Table 3.3: Number of sites where significant taxon loadings are above 9th and 1st percentile for each factor.

varid	taxa	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Sum
v1	*Abies*	26	21	26	34	21	16	22	16	20	18	220
v2	*Acer*	44	21	14	14	29	24	11	23	17	14	211
v3	*Alnus undiff.*	34	36	30	27	30	31	26	31	16	27	288
v4	*Ambrosia type*	9	18	23	24	39	33	44	43	42	26	301
v5	*Apiaceae*	4	6	9	13	19	19	16	13	17	22	138
v6	*Aquifoliaceae*	4	5	8	9	14	12	17	9	11	10	99
v7	*Artemisia*	29	15	36	31	21	26	13	13	25	18	227
v8	*Asteraceae*	28	25	22	27	23	26	22	16	28	18	235
v9	*Betula undiff.*	33	30	41	33	27	23	20	12	18	22	259
v10	*Carya*	34	24	9	10	19	9	14	13	18	17	167
v11	*Castanea*	10	11	6	7	16	12	11	18	12	17	120
v12	*Cephalanthus*	5	3	4	3	5	7	5	7	3	7	49
v13	*Chenopodiaceae/Amaranthaceae*	9	22	32	22	26	22	26	27	28	28	242
v14	*Corylus*	7	20	20	17	18	16	12	17	19	27	173
v15	*Cupressaceae*	12	27	25	33	27	25	20	26	26	24	245
v16	*Cyperaceae*	34	38	22	19	21	24	26	28	19	26	257
v17	*Ericaceae undiff.*	6	14	15	15	20	20	20	17	16	24	167
v18	*Fagus*	39	32	10	7	16	16	9	17	14	12	172
v19	*Fraxinus*	21	16	23	16	21	19	13	23	26	20	198
v20	*Larix/Pseudotsuga*	13	12	28	20	13	16	20	14	11	22	169
v21	*Liliaceae*	0	2	3	7	6	8	6	6	9	9	56
v22	*Lycopodium undiff.*	9	6	14	15	13	12	21	18	18	11	137
v23	*Magnoliaceae*	2	3	3	3	5	2	4	5	6	4	37
v24	*Myrica*	4	13	7	18	18	13	16	16	18	14	137
v25	*Nyssa*	8	11	6	6	13	8	8	11	6	8	85
v26	*Ostrya/Carpinus*	19	21	22	21	16	19	14	17	20	16	185
v27	*Picea undiff.*	84	16	20	19	11	6	14	6	9	7	192
v28	*Pinus undiff.*	42	53	45	29	15	22	15	10	11	11	253
v29	*Plantaginaceae*	1	2	4	8	15	14	17	25	18	13	117
v30	*Poaceae*	31	33	36	26	25	22	32	28	21	19	273
v31	*Polygonum*	1	6	7	12	13	17	17	10	18	16	117
v32	*Polypodiaceae*	10	15	17	26	19	19	22	19	17	28	192
v33	*Populus*	10	15	26	28	25	28	26	21	15	27	221
v34	*Pteridium*	1	13	15	15	20	13	20	17	19	17	150
v35	*Quercus*	45	37	18	23	17	13	18	12	24	8	215
v36	*Ranunculaceae*	2	2	8	16	14	15	25	13	16	21	132
v37	*Rubiaceae*	3	0	3	6	4	7	4	5	7	7	46
v38	*Rumex*	3	2	8	13	22	22	28	28	24	16	166
v39	*Salix*	31	12	32	31	18	22	21	20	22	18	227
v40	*Sarcobatus*	4	4	7	4	7	11	6	7	9	8	67
v41	*Saxifragaceae*	2	4	4	3	4	2	6	8	8	8	49
v42	*Shepherdia*	3	2	9	11	0	10	13	8	9	14	79
v43	*Sphagnum*	8	18	12	15	12	11	16	15	20	16	143
v44	*Taxodium*	2	3	1	1	4	2	1	5	3	2	24
v45	*Tilia*	11	11	9	12	8	9	11	18	24	22	135
v46	*Tsuga*	22	24	22	20	25	28	16	37	17	24	235
v47	*Ulmus*	14	39	17	20	20	20	22	15	22	22	211

Table 3.4. Site contributions to the overall component. Taxon loading above the 9th percentile or below the 1st percentile extracted at each individual site. The number of taxa with significant loadings are summed for every site.

Id	SiteName	Lat	Long	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
1	Allenberg Bog	42.3	-78.9	13	6	7	3	11	5	8	5	7	5
2	Anderson Pond	36.0	-85.5	11	12	11	5	14	8	6	10	4	6
3	Balsam Lake	42.5	-74.6	6	7	3	5	1	8	1	4	2	4
4	Basswood Road Lake	45.3	-67.3	9	7	11	5	6	3	9	7	1	9
5	Belmont Bog	42.3	-77.9	8	12	11	9	6	13	7	4	14	6
6	Boriack Bog	30.4	-97.1	3	3	7	2	7	8	4	4	7	10
7	Buckle's Bog	39.6	-79.3	9	10	3	4	7	6	2	3	0	0
8	Cahaba Pond	33.5	-86.5	6	9	7	6	6	12	12	9	6	7
9	Camel Lake	30.3	-85.0	8	5	6	10	9	7	7	7	7	5
10	Clear Lake	41.7	-86.5	9	2	5	4	7	2	8	2	1	7
11	Colo Marsh	42.0	-93.3	5	11	6	12	5	9	7	0	5	7
12	Cranberry Glades	38.2	-80.3	10	13	3	7	7	8	4	5	4	7
13	Crates Lake	49.2	-81.3	8	9	2	8	5	5	10	8	8	8
14	Crider's Pond	40.0	-77.6	9	7	11	5	11	11	13	8	5	10
15	Cupola Pond	36.8	-91.1	10	6	6	5	6	8	7	8	7	4
16	Deer Lake Bog	44.0	-71.8	10	6	12	10	10	7	9	8	6	5
17	Devils Lake	43.4	-89.7	5	11	10	3	4	7	4	10	5	6
18	Duck Pond	41.9	-70.0	9	3	6	7	7	15	8	12	13	12
19	Lac Dufresne	45.9	-70.4	2	10	11	7	11	13	12	13	10	7
20	Edward Lake	44.4	-80.3	6	6	4	11	8	9	3	9	11	4
21	Frains Lake	42.3	-83.6	9	3	7	12	8	6	9	6	8	8
22	Gass Lake	44.1	-87.7	3	10	7	6	4	7	7	5	6	6
23	Glenboro Lake Site	49.4	-99.3	3	6	4	5	5	2	3	6	6	2
24	Goshen Springs	31.7	-86.1	9	10	4	4	9	9	7	9	14	6
25	Hack Pond	38.0	-79.0	9	4	8	9	7	8	7	11	10	14
26	Heart Lake	44.2	-74.0	7	4	10	11	8	8	11	7	7	4
27	Heart's Content Bog	42.2	-74.0	5	7	5	4	7	5	5	5	9	6
28	Jackson Pond	37.5	-85.7	10	10	6	10	11	7	9	13	13	6
29	Lily Lake	45.1	-92.8	6	8	14	7	9	6	4	3	6	7
30	Little Lake	45.1	-66.7	6	9	8	10	11	7	8	6	4	8
31	Lake Annie	27.2	-81.4	3	15	5	2	7	5	5	4	3	6
32	Longswamp	40.5	-75.7	17	6	9	3	9	6	9	12	10	4
33	Moulton Pond	44.6	-68.6	10	7	6	3	10	7	5	10	8	4
34	Lake West Okoboji	43.3	-95.2	7	3	12	7	5	7	2	2	6	1
35	Panther Run Pond	40.8	-77.4	11	12	5	15	12	14	6	9	7	7
36	Pigeon Marsh	34.7	-85.2	10	6	5	4	4	5	4	4	6	2
37	Pine Bog	41.3	-74.8	6	4	1	2	3	1	2	5	7	2
38	Pretty Lake	41.6	-85.3	9	4	5	6	0	10	5	6	5	3
39	Rayburn's Dome	32.5	-93.2	9	10	5	4	7	6	4	4	1	0
40	Rockyhock Bay	36.2	-76.7	9	11	6	10	5	10	9	6	12	12
41	Rogers Lake	41.4	-72.1	9	7	6	6	7	4	2	6	4	8
42	Sandogardy Pond	42.8	-71.7	9	1	6	2	3	2	3	3	1	2

43	Sheelar Lake	29.5	-82.0	1	8	7	11	8	9	7	6	6	11
44	Silver Lake	44.6	-63.6	6	3	11	4	1	3	2	5	3	5
45	Silver Lake [Ohio]	40.4	-83.8	5	8	10	9	7	5	14	4	11	6
46	Spring Lake	41.7	-76.4	7	5	4	4	11	9	10	8	6	11
47	Szabo Pond	40.4	-74.5	14	10	7	9	9	6	5	5	7	6
48	Tannersville Bog	41.0	-75.3	8	9	6	5	11	8	7	9	6	7
49	Petit Lac Terrien	46.6	-70.6	4	7	6	6	7	9	9	14	15	6
50	Torren's Bog	40.4	-82.5	2	3	3	4	6	3	5	5	3	6
51	Lac à la Tortue	45.5	-73.3	9	7	5	5	6	4	7	7	6	10
52	Lake Tulane	27.6	-81.5	4	9	3	3	5	5	4	7	8	5
53	Unknown Pond	45.6	-70.6	7	6	7	7	2	7	7	5	8	5
54	Van Nostrand Lake	44.0	-79.4	6	6	5	8	8	8	3	5	3	9
55	Vestaburg Bog	43.4	-84.9	8	5	4	3	5	4	6	8	5	7
56	White Pond	34.2	-80.8	9	1	8	7	13	11	9	10	7	12
57	Wintergreen Lake	42.4	-85.4	5	2	10	7	3	3	7	5	2	9
58	Wolsfeld Lake	45.0	-93.6	8	7	13	17	4	6	11	4	5	7
59	Wood Lake	45.3	-90.1	8	6	10	10	5	6	9	9	12	12
60	Quarles Pond	38.0	-79.1	6	11	8	15	12	8	12	10	12	14
61	Powers Fort Swale	36.6	-90.6	10	9	6	10	7	10	16	5	5	5
62	Ross Pond	43.9	-69.4	7	5	5	7	8	7	14	10	5	6
63	Carp Lake	45.9	-120.9	7	4	0	7	6	3	6	8	7	8
64	Gould Pond	45.0	-69.3	7	5	5	4	11	7	7	10	6	6
65	Poland Spring Pond	44.0	-70.4	6	7	12	6	8	9	10	6	5	10
66	Clear Pond	33.8	-79.0	7	1	2	0	4	2	5	3	4	7
67	Roberts Creek	43.0	-91.5	3	16	13	11	11	6	8	14	8	13
68	Sherd Lake	44.3	-107.0	3	2	4	2	2	6	5	5	6	4
69	Tonawa Lake	44.9	-77.2	5	4	11	8	4	3	6	8	2	3
70	Cottonwood Lake	44.8	-99.9	6	2	11	9	7	10	8	8	10	6
71	Lac du Diable	48.9	-66.1	6	12	10	7	4	4	4	3	6	8
72	Angal Lake	67.1	-153.9	6	1	6	5	1	2	3	1	3	2
73	Ranger Lake	67.1	-153.7	6	3	2	4	1	2	3	1	5	4
74	Kaiyak Lake	68.2	-161.4	7	3	4	2	5	5	2	3	4	4
75	Joe Lake	66.8	-157.2	7	3	1	4	4	0	4	1	1	5
76	Ped Pond	67.2	-142.1	6	5	4	8	5	3	5	6	2	5
77	Redstone Lake	67.3	-152.6	6	3	3	4	3	0	0	3	6	2
78	Rebel Lake	67.4	-149.8	6	3	5	6	2	5	2	2	3	5
79	Screaming Yel.Pond	67.6	-151.4	4	4	3	5	4	3	5	4	6	5
80	Wien Lake	64.3	-151.3	3	0	4	4	1	4	2	3	2	5
81	Fresh Pond	41.2	-71.6	9	5	2	11	12	10	10	9	9	11
82	No Bottom Pond	41.3	-70.2	8	8	3	3	12	5	9	12	8	3
83	Clear Lake [Iowa]	43.2	-93.4	8	2	12	6	2	2	7	4	4	3
84	Barrett Lake	37.6	-119.0	1	6	3	6	7	9	6	8	10	10
85	Guardipee Lake	48.6	-112.7	2	3	3	8	7	12	9	11	13	11
86	Davis Lake	46.6	-122.3	5	3	4	6	7	5	9	6	4	2
87	Compass Pond	50.0	-56.2	6	1	6	9	2	4	6	7	8	7
88	Leading Tickles	49.5	-55.5	3	7	9	9	5	4	5	3	8	8

89	Battle Ground Lake	45.8	-122.5	6	10	8	6	6	5	8	7	4	8
90	Cygnet Lake Fen	44.7	-110.6	5	4	3	10	3	5	2	3	9	5
91	Divide Lake	43.9	-110.2	0	5	9	8	7	7	4	4	4	10
92	Hedrick Pond	43.8	-110.6	0	7	5	3	2	3	3	5	6	6
93	Mariposa Lake	44.2	-110.2	3	3	3	3	6	5	3	6	7	5
94	Slough Creek Pond	44.9	-110.4	2	5	5	6	5	4	4	5	2	5
95	Cheyenne Bottoms	38.5	-98.7	11	6	7	6	6	2	3	1	0	2
96	Como Lake	37.6	-105.5	3	5	4	4	10	9	8	4	6	7
97	Exchequer Meadow	37.0	-119.8	4	3	5	7	5	4	6	5	7	6
98	Cub Lake [Idaho]	44.1	-111.2	5	3	8	5	6	6	8	9	10	5
99	Gardiner's Hole	44.9	-110.7	5	5	9	9	8	7	8	6	3	4
100	Fudger Lake	40.1	-83.5	6	5	3	11	8	8	4	9	4	15
101	Ladd Lake	41.4	-84.8	6	8	7	9	4	3	3	11	5	10
102	Devon Island G.	75.4	-82.5	1	4	0	1	8	1	1	2	8	4
103	Brophy Ditch	43.2	-81.7	7	8	6	3	10	7	4	12	8	4
104	Cranberry Lake	44.1	-78.1	8	3	8	5	11	5	9	7	7	3
105	Lake Erie	41.9	-82.8	6	8	12	7	15	9	7	8	8	7
106	Joes Pond	48.2	-58.2	3	3	8	10	3	11	7	8	8	6
107	Porqui Pond	44.2	-79.8	3	12	8	7	8	5	5	9	10	9
108	Robinson's Pond	48.3	-58.8	2	9	8	5	8	7	9	9	6	7
109	Wylde Lake	43.9	-80.4	5	7	7	7	4	5	10	7	8	3
110	Irvin Lake	47.1	-93.6	4	9	9	7	7	7	7	6	8	11
111	Swamp Lake	38.0	-119.8	2	6	1	1	2	4	5	8	6	7
112	Browns Pond	38.2	-79.6	8	11	8	13	6	9	13	4	14	12
113	Hanson Marsh	43.4	-89.7	3	9	7	5	11	14	8	16	5	8
114	Kellners Lake	44.2	-87.8	5	7	5	3	4	4	4	1	7	4
115	Hanging Lake	68.4	-138.4	5	5	4	9	2	4	9	3	7	4
116	Head Lake	37.7	-105.5	2	5	7	7	6	5	8	6	13	8
117	San Agustin Plains	33.9	-108.3	3	6	3	3	7	7	3	6	7	5
118	San Agustin Plains	33.9	-108.3	7	7	0	2	0	0	0	1	1	1
119	Maria Lake	68.3	-133.5	5	4	5	6	2	5	3	2	2	2
120	Tuktoyaktuk 5	69.1	-133.5	4	2	5	1	4	2	9	4	7	2
121	Lateral Pond	66.0	-135.9	3	1	8	4	4	3	6	4	5	6
122	Forest Lake	46.5	-112.2	3	1	2	3	6	7	5	4	7	9
123	Graham Lake	45.2	-77.4	9	5	7	6	6	8	7	8	7	1
124	Sleet Lake	69.3	-133.6	5	4	2	2	7	6	2	2	4	5
125	Reindeer Lake	69.1	-132.2	5	6	6	10	3	6	4	3	5	6

Table 3.5. Millennial-scale climate transitions identified in this study. Calibrated age estimates were determined by averaging the estimates in the sum of all 10 factors and independent factor 10 estimates. The difference column represents the pacing between events.

Event	Sum of 10 factors (^{14}C)	Factor 10 estimates (^{14}C)	Average calibrated yr BP *	Pacing
T0	600	600	600	0
T1	1500	1250	1300	800
T2	3000	2750	3000	1700
T3	4750	4500	5300	2300
T4	6000	6000	6800	1500
T5	7500	7250	8200	1400
T6	9250	9000	10300	2100
T7	10250	10000	11700	1400
T8	10750	10750	12900	1200
T9	11500	11500	13500	600
Pacing				1450
STD				550

* Calibration set from Stuiver and Reimer, 1998.

TABLE 3.6: Millennial-scale climate transitions for the Holocene and Late Glacial identified in several studies.

Period *	Terrestrial							Marine **	GISP2[††]	Significant discontinuities	
	Botanical This study	Botanical[§§]	Alpine glaciers[†]		Botanical [§] Major	Minor	Regional Syntheses [#]			Date	Diff.
LIA maximum			330				300	300			
Pacific	(T1) 600	600			730		700		600	650	
neo-Atlantic				Exp							
	(T2) 1300							1400			
Scandic		1650		Con	1550		1600		1500	1500	900
sub-Atlantic	(T3) 3000	2850	2800	Exp	2780		2600	2800	2750	2800	1300
sub-Boreal 1		4030		Con		4000	4000	4200		4050	1250
sub-Boreal	(T4) 5300		5300	Exp	5900	5000	5500	5900	5550	5500	1450
Atlantic 2	(T5) 6800	6700		Con		6850			6300	6700	1200
Atlantic 1	(T6) 8200	8100	7800	Exp		8100	7500	8100	8100	8100	1400
Atlantic					9500			9400	9300	9400	1300
Boreal	(T7) 10300	10190	10300		10450		10000	10300	10600	10300	900
pre-Boreal	(T8) 11700			Con	11500		11500	11100	11300	11500	1200
Younger Dryas	(T9) 12900	12900	12800	Exp	12800		12800	12600		12800	1300
Allerod	(T10)13500	13800		Con						13650	850
Late Glacial				Exp							

Note: Radiocarbon dates were calibrated to calendar years using INTCAL98 Calib 4.2 software (Stuiver and Reimer, 1993).
* Blytt-Sernander names assigned to these periods following Wendland and Bryson (1974).
† From Denton and Karlén (1973).
§ From Wendland and Bryson (1974).
\# From Bryant and Holloway (1985).
** From Bond et al. (1997, 2001).
†† From O'Brien et al. (1995).
§§ From Viau et al. (2002).

Figure 3.1: Location of sites used in Principal Component Analysis (PCA).
125 sites identified with continuous sequences between 0 and 12,000 radiocarbon years.

Figure 3.2. Location of sites used in Principal Component Analysis (PCA) of the past 1,000 years BP. Sites contain continuous pollen sequences with an average 40-year resolution..

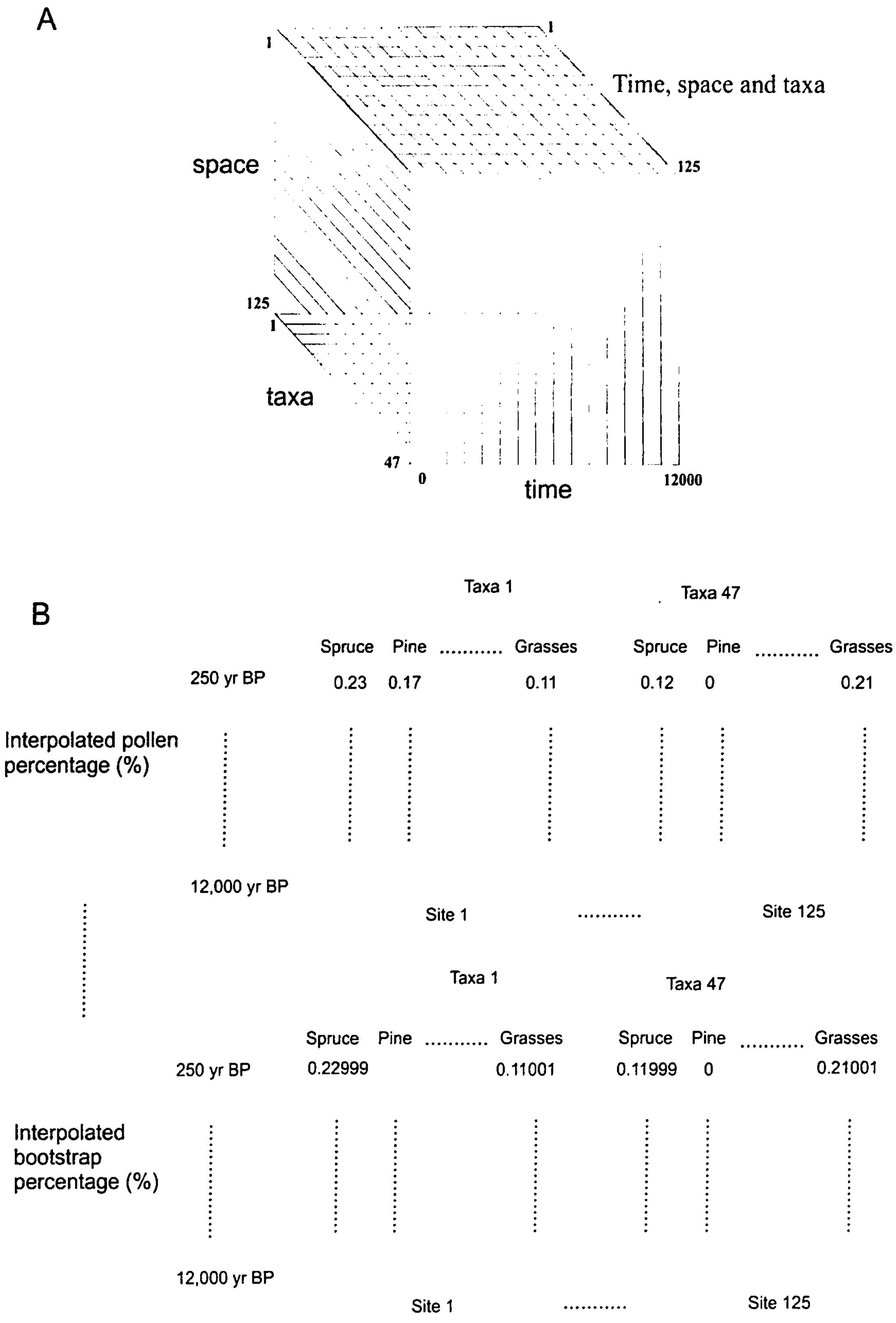


Figure 3.3: Methodology and input data. A: Representationof of the original data. B: Method used to collapse 3D data matrix in 2D for input tp PCA.

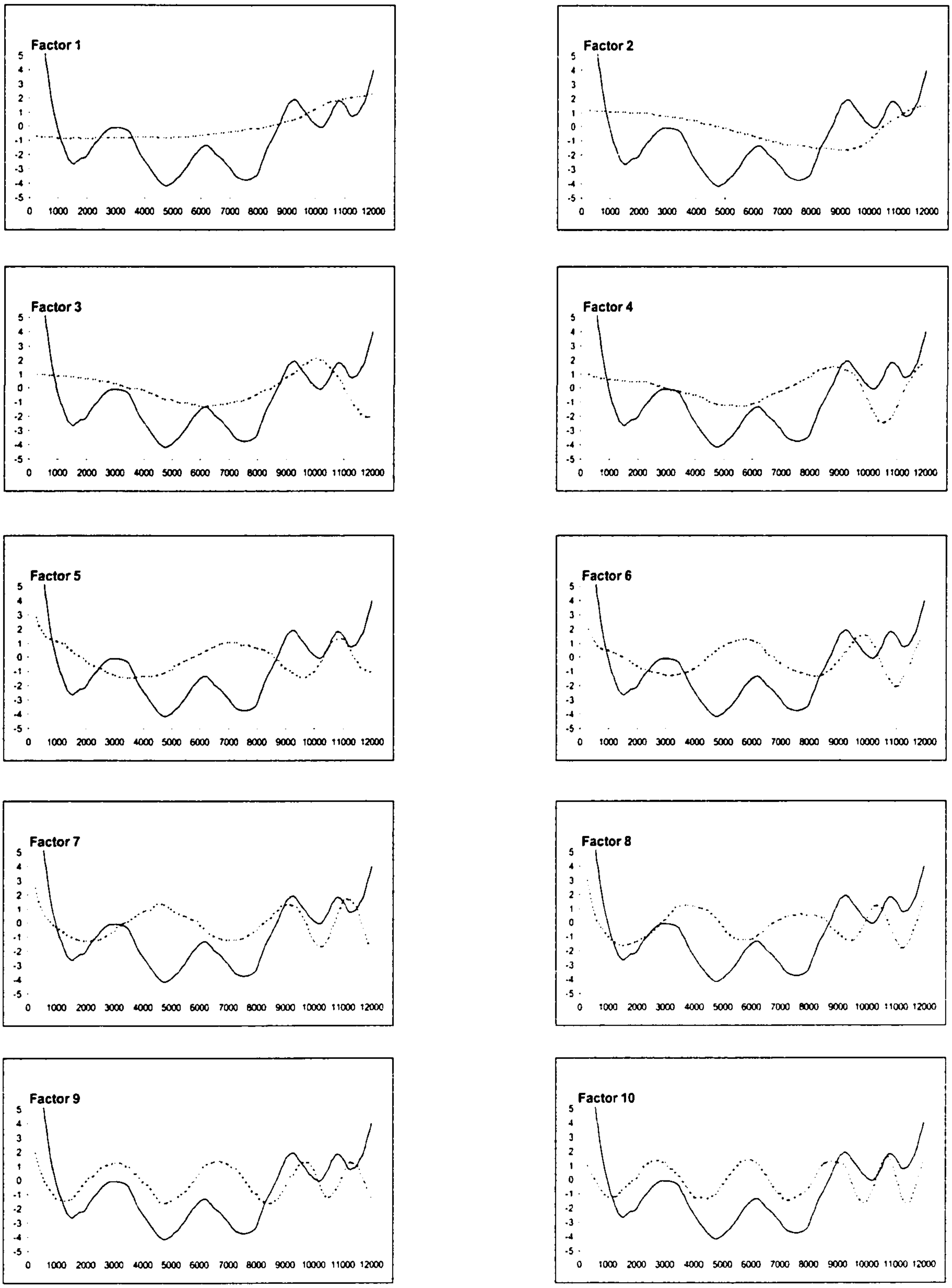


Figure 3.4:. Factor scores of all 125 pollen times series. Each individual factor scores (i.e. broken lines) is illustrated to show its iterative contribution to the total. All ten factor scores (i.e. solid line) were summed to illustrate the combined temporal evolution of the common variance.

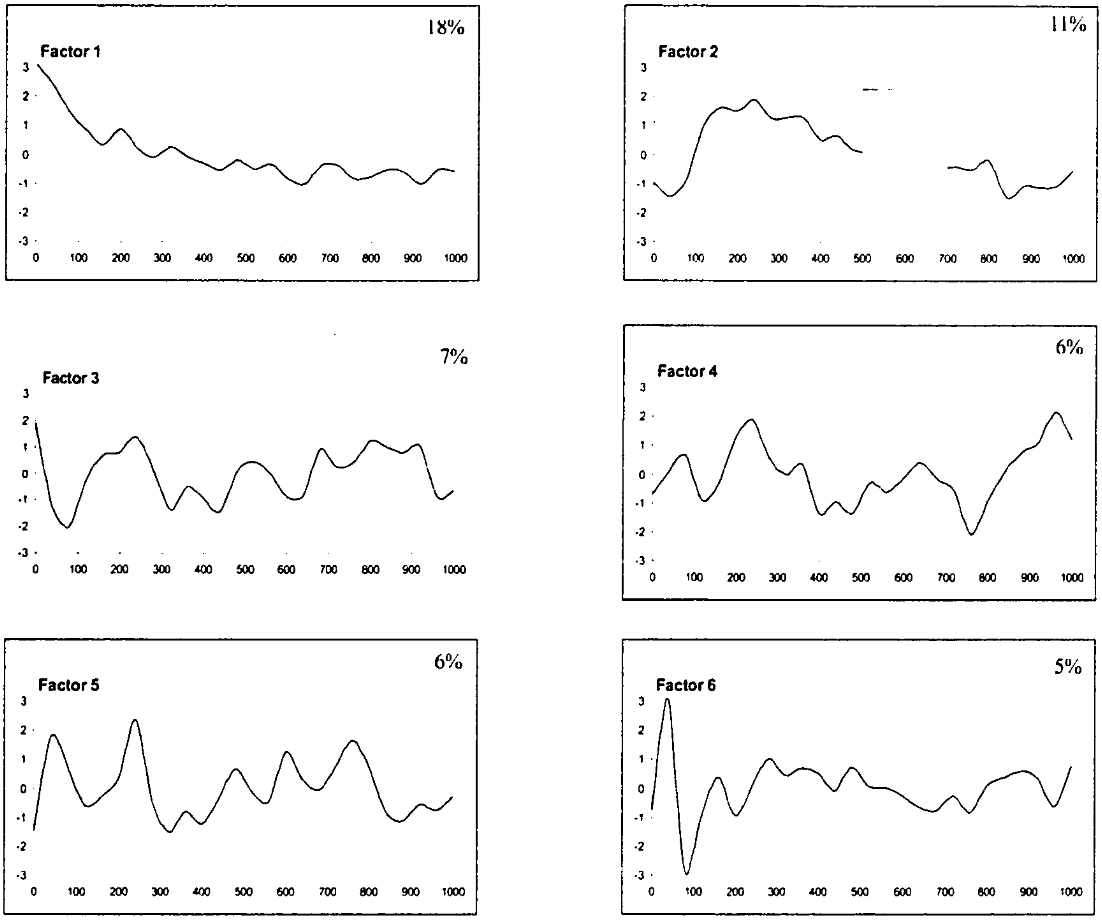


Figure 3.5: Factor scores of PCA on high-resolution pollen sites. x-axis = time (14C); y-axis = scores. Shadowed section in Factor 2 is an estimate of the transition towards the LIA.

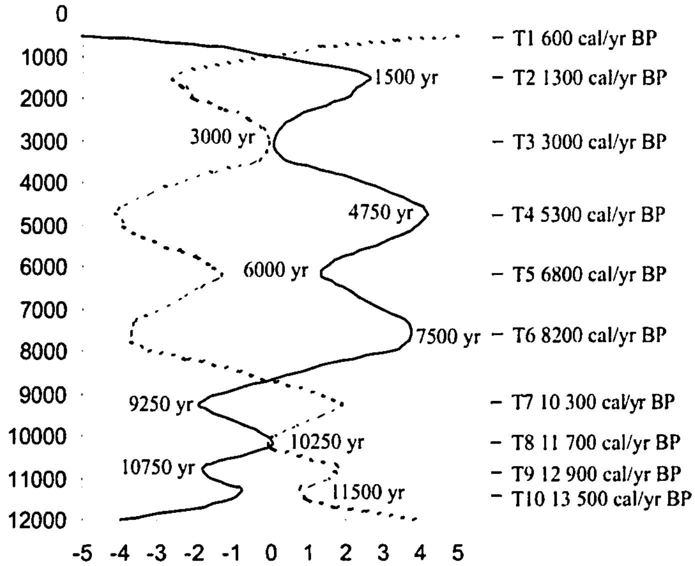


Figure 3.6:: Identified times when vegetation changes occurred synchronously across North America. Dates of transitions are expressed in radiocarbon years. x-axis = scores ; y-axis = time (14C).
Secondary Y-axis indicate calibrated dates on transitions as in Figure 3.5. Broken line represents opposite directions of change.

Chapter 4

MEAN-CONTINENTAL JULY TEMPERATURE VARIABILITY IN NORTH AMERICA DURING THE LATE GLACIAL AND HOLOCENE.

Abstract

This chapter presents a mean July temperature reconstruction for North America during the past 14 000 calendar years (cal yr) based on 752 pollen diagrams extracted from the North American Pollen Database (NAPD). Spectral analysis and band-pass filtering of the reconstructed mean-continental July temperature time series revealed that Holocene and late glacial climate variability in North America may have been periodic and with higher amplitudes during the late glacial. Times of century and millennial-scale temperature transitions closely correlate with ice, marine and terrestrial proxy-climate data and cosmogenic nuclide records recording variations in solar output. A circa 1150-year climate periodicity is shown to better reconcile all of the climate-proxy records. Holocene century to millennial-scale temperature variations in North America are most likely associated in part with solar output variability and internal feedback mechanisms such as the thermohaline circulation (THC) resulting in rearrangements of the ocean-atmospheric circulation patterns at these times of transition.

4. INTRODUCTION

Two independent quantitative methods are illustrated in the preceding chapters, which reveal that pollen sedimentary records can be used to study century to millennial-scale climate variability of the past 14,000 years at the continental-scale of North America. In this chapter we present a quantitative temperature reconstruction based on these pollen data. This will further quantify the spatial and temporal nature of Holocene and late glacial climate change in North America.

Unresolved issues concerning recent global warming due to enhanced CO_2 concentration in the atmosphere result from the difficulty in establishing the magnitude of surface temperature change that can be attributed to human impact versus natural climate variability (IPCC, 2001). Current debate concerning century to millennial-scale climate variability centers on the timing, magnitude and spatial extent of surface temperature changes. Although the processes of climate change are not fully understood, several quasi-periodic oscillations of past natural climate variability are revealed in the climate proxy records at century to millennial-scales (Jones *et al.*, 1999; Mann *et al.*, 1999; Esper *et al.*, 2002; Viau *et al.*, 2002; Bond *et al.*, 1997, 2001; MacManus *et al.*, 1999; O'Brien *et al.*, 1995; Stuiver *et al.* 1997; Oppo *et al.* 1998; Raymo *et al.* 1998; and Bianchi and McCave, 1999). Before any attempt can be made to explore mechanisms that govern century to millennial-scale climate variability, evidence from all climate components must be gathered. Establishing the timing, magnitude and spatial extent of these climate variations will lead to insight into which external and/or internal forcings may be responsible. Potential candidates for these hypothesized variations and abrupt transitions in temperatures have been proposed. These include internal forcings such as ocean-atmosphere-ice interactions (Broecker and Denton, 1989), and nonlinear feedbacks within the climate system (Peixoto and Oort, 1992; Rind, 1999). External forcings include variations in latitudinal and seasonal distribution of incoming solar radiation (Milankovitch orbital parameters), variable solar output (e.g. Perry and Hsu, 2000; Crowley, 2000; Bond *et al.*, 2001; Van Geel *et al.*, 1999), orbital harmonics (Pestiaux *et al.*, 1988), and volcanism (Zielinski *et al.*, 1996, Crowley, 2000).

The North Atlantic marine and Greenland ice core records have suggested a quasi-periodic 1500-year cycle to be a pervasive feature of climate variability during the past 1 m.y. (e.g. Bond *et al.*, 1997, 2001; Stuiver *et al*, 1997). However, there remains certain unresolved issues concerning the true nature of this 1500-year cycle as evidenced by comparison of several records of this variability (Viau *et al.*, 2002). Evidence from these multi-proxy studies suggests that the magnitude of temperature change at century to millennial-scale were strong during glacial maxima (i.e. Dansgaard-Oeschger-Heinrich events), more subtle during interglacials (Bond *et al.*, 1997), but strongest during ice sheet growth and disintegration (i.e. Bolling-Ållerød-YD-Holocene transitions). This hypothesis suggests that interactions between external and internal forcings maybe the key in understanding natural century to millennial-scale climate variability. Uncovering the magnitude and current direction of natural century to millennial-scale temperature variability on a continental-scale during the Holocene and late glacial can bring insight on these issues and also help in quantifying recent anthropogenic impact.

Records of vegetation change preserved in pollen time series have been used to document past climate variability on several time and spatial scales (Bartlein *et al.*, 1984; COHMAP members, 1988; Grimm *et al.*, 1993, Guiot *et al.*, 1993; Viau *et al.*, 2002; Gajewski *et al.*, 2000) as well as the magnitude and timing of higher-frequency climate variability such as the Little Ice Age (LIA) and preceding Medieval Warm Period (MWP) (Bernabo, 1981, Gajewski, 1987). Studies have shown that vegetation response is relatively short at century to millennial-scale (Webb, 1986) and there is evidence that vegetation has responded abruptly in glacial time at the millennial scale (Grimm *et al*, 1993, Guiot *et al.*, 1993). Thus, these records can be used to study this timescale of climate variability.

4.1 METHODS AND DATA

4.1.1 POLLEN DATA

Fossil pollen records from lake, bogs, and other terrestrial sediments are here used as proxy for late glacial and Holocene century to millennial-scale temperature changes at the continental-scale of North America. These temperature changes are estimated using the method of modern analogues (Modern Analogue Technique (MAT)) (Overpeck *et al.*, 1985). Continuous and discontinuous published and unpublished radiocarbon dated fossil pollen time series from a dense network of sites were extracted from the NAPD (Grimm, 2000) using Microsoft Access database queries.

The number of available sites in North America from the NAPD that can be used for paleoecological/paleoclimatological analysis varies through time and in space (Figure 4.1; Table 4.1). There are 752 sites unevenly distributed throughout the continent. The Great Lakes region is well represented while other regions have fewer data (Figure 4.1).

A new modern pollen dataset was assembled and subsequently named WWKGVS2000. This modern pollen dataset is derived from the modern data of Arvizinis and Webb (1985), where Webb and Williams (pers. comm.) added new modern pollen sites (WW2000). Then sites from the Canadian Pollen Database (Gajewski, unpublished) and Global Pollen Database (GPD, Grimm, 2000) were merged into a single modern pollen dataset (WWKGVS2000). This gave a final modern dataset of 4590 sites, which covers most regions of North America (Figure 4.2). The complete merging procedure is documented in Sawada (2001).

4.1.1.1 POLLEN SUM

In general, the choice of a pollen sum should emphasize the question of interest and maximize the signal to-noise ratio (Jackson, 1994). Including pollen taxa that are not necessary to the questions of interest increases the noise in the dataset and decreases the signal. Here, we use a set of 89 types (Table 4.2) in order to incorporate the major taxa from all regions of North America. Using an expanded pollen sum, including species and generic-level herbs and shrubs, may help in distinguishing regional climate differences and yield more accurate temperature estimates.

4.1.1.2 TEMPORAL RESOLUTION

All pollen time series were resampled at 100-cal/yr intervals by linear interpolation between the pollen spectra (Appendix G). This is done in order to tract subtle temperature changes and also to provide several realizations of climate maps between each transition identified in chapters 2 and 3. In this way, random variations can be smoothed to determine patterns common for the time periods.

4.1.2 MODERN ANALOGUE TECHNIQUE (MAT)

Past climates can be estimated from pollen spectra using the modern analogue technique with dissimilarity measures based on squared chord distance measurements (Overpeck *et al.*, 1985). The squared-chord measurement has been identified as the most robust coefficient to reduce the signal-to-noise ratio in pollen data (Overpeck *et al.*, 1985). Essentially, this method compares a fossil pollen spectrum to all modern pollen samples available for the geographic extent of the research area. For each fossil pollen spectrum at a given target time, the best analog is the modern site that possesses the least dissimilar modern pollen spectrum. From this, climate variables such as temperature and precipitation can be extracted to represent similar climatic conditions of the past.

All interpolated pollen spectra for a given time interval (up to 752 sites) were compared to the modern data (4590 sites), which provided a corresponding number of times series (up to 752 time series) of mean July temperatures for the past 14,000 cal/yr BP (Appendix G). Whereas opposing directions of temperature change may exist regionally at this scale, a mean continental July temperature reconstruction at 100 years interval should give insight on the magnitude of continental temperature variability during the Late Glacial and Holocene. Maps of the mean July temperature anomalies were also constructed for regional evaluation of climate patterns.

4.1.3 Mapping reconstructed temperature variations in North America

In order to depict spatial patterns of temperature variations in North America through time, maps at 100-year interval were constructed for the past 14,000 calendar years BP. Modern July temperatures were estimated using the distribution of sites for each time interval and a 30 by 30 km North American continental-scale temperature grid (P. Bartlein, pers. comm.). Using a nearest neighbor distance statistic, modern temperature values were interpolated to the location of each fossil site. Each site was also assigned a reconstructed temperature value based on the output of the MAT reconstruction for all time intervals. A local polynomial-smoothing algorithm was then applied to create continental grids for both modern and reconstructed temperatures for each time interval. The best search radius was prescribed individually for each time period based on the distribution of sites in order to minimize extrapolations. Using the modern and reconstructed grids, July temperature anomalies were calculated on a grid-to-grid basis resulting in 140 anomaly temperature grids for North America during the past 14,000 cal/yr BP.

4.2 RESULTS

4.2.1 RECONSTRUCTED TEMPERATURE VARIABILITY

The reconstructed mean July temperature anomaly curve for North America broadly resembles the isotopic ^{18}O curve of the GRIP and GISP2 ice cores. There is a warming of 4-5°C between 14,000 and 9,000 cal/yr BP during the deglaciation phase. Superimposed on this warming are smaller temperature fluctuations (Figure 4.3). Temperature changes are smaller during the past 10,000 years and reveal a three-part division of the Holocene at 9,000-6000, 6000-3000, and 3000-0 cal/yr BP (Figure 4.4). The rapid warming of the early Holocene, slightly warmer middle Holocene, and cooler late Holocene (i.e. neoglacial) is distinctively observable (Figure 4.4). Superimposed of these long-term temporal structures are subtle but yet appreciable oscillations between warming and cooling on millennial-scales (Figure 4.4). Each millennial-scale peak is numbered 1 through 12 and represents smaller temperature variations on the order of magnitude between 0.1 and 0.6°C (Figure 4.3 and 4.4).

4.2.2 SPECTRAL ANALYSIS AND BAND-PASS FILTERING

The reconstructed mean-continental July temperature anomaly time series was further examined using spectral analysis (Overpeck, 1986, 1987). Spectral analysis requires that the temperature time series be stationary. The time series was detrended using a 10^{th} order polynomial (adjusted r^2 of .99) (Figure 4.5) where these residuals are presented in Figure 4.6. The resulting curve retained the basic higher-frequency variations of the original data where each peak is labeled from 1 to 12 and associated troughs labeled alphabetically a to l (Figure 4.6). A small extrapolation can be observed at modern value, but this is not surprising given the order of the polynomial. Stationarity was achieved on the mean but there remains some non-stationarity in variance during the late glacial (Figure 4.6). However, the detrending procedure reproduces both the timing and magnitude of temperature change of the undetrended series at the millennial-scale. Once lower-frequencies are removed, maximum continental-scale temperature anomalies from the trends were on the order of approximately 0.4°C magnitude for the Holocene and 0.6°C for the late glacial.

A spectral analysis was performed on the detrended time series to identify possible periodicities. A significant narrow range of concentration of variance was identified in a frequency band bounded by 1-cycle/1100 years and 1-cycle/1200 years associated with a central periodicity of 1166 cal/yr (Figure 4.7). Based on these results, a band pass filter was applied to the series to retain only periods between 1100 and 1200 years. This provides a smoothed model of the residual times series that can be compared to other proxy-climate records. Visual inspection indicates that it reproduces the data well (Figure 4.8). The millennial-scale periodic component had higher amplitudes during the early Holocene and to a lesser extent during the late Holocene, with a more attenuated signal during the middle-Holocene. This concords with previous studies (e.g. Viau *et al.*, 2002; Grimm *et al.*, 1991; Overpeck, 1987) (Figure 4.8).

4.2.3 SPATIAL PATTERNS OF POSTGLACIAL CHANGES

Maps of the July temperature anomalies across North America, interpolated to 100-year interval are presented in Figures 4.9a-n. By interpolating to this fine interval, several maps can be used to deduce spatial information on several time scales; there are several replicates between each transition, so only the large-scale patterns are interpreted. The general characteristics of all 140 maps reveal coherent spatial patterns of opposing temperature anomalies (Figure 4.9a-n). At any transition, some areas warmed, while other cooled suggesting times of climate change during the past 14,000 years were not unidirectional in all places but resulted from atmospheric circulation changes. The long-term trends identified earlier (Figure 4.3; 4.4) are revealed in space. Spatial patterns of temperature anomaly rarely surpasses ± 1°C during the first 9,000 years BP (Figure 4.9a-i) while between 14,000 and 9,000 years BP, the spatial patterns of anomalies are on the order of ± 2-4°C (Figure 4.9j-n). The early Holocene period between 9,000 and 6,000 years BP is characterized by a generalized warming east of the Cordilleran, high northern latitudes and along the margins of the remnants of the Laurentide ice sheet (Figure 4.9g-i). The three-part Holocene describe above can be seen as a deeper southward penetration of warmer temperature anomalies during the middle Holocene in northern latitudes, northeastern North America and the southwest, especially during the period of 6,000 to 4,000 years BP (Figure 4.9e-g). Generalized continental-scale cooling in Eastern North America, high-northern latitudes, the Great Lakes and Midwest United States are characteristic of the neoglacial after 4,000 years BP (Figure 4.9a-d).

Superimposed on these long-term trends are fluctuations at a millennial-scale. Although influenced by the long-term deglaciation trend, the transition identified at circa 12,900 years BP shows relatively warmer anomalies in the southeast United States and in the Pacific northwest until approximately 12,500 years BP when a cooling can be identified in these regions (Figure 4.9m). During the middle Holocene, spatial patterns of temperature anomalies prior to the transition around 5,800 BP indicate an east-west band of warm anomalies that is clearly evident across Canada (Figure 4.9f). Following this transition, a generalized cooling is evident by the opening of an east-west band of negative anomaly leading to suggest a deeper penetration of polar air masses (Figure 4.9f). A gradual warming beginning around 1,500 years BP and reaching its maximum around 900 years ago (Esper *et al.*, 2002) is depicted by a deeper penetration of warmer temperature anomalies in northern Canada, northern Quebec and the Western Pacific regions (Figure 4.9a, b). A gradual cooling following 900 years BP marks the onset of the Little Ice Age (LIA) with maximum negative anomalies between 300 and 600 years ago (Figure 4.9a).

In summary, the spatial patterns of temperature anomalies through the past 14,000 years depict well the long-term and century to millennial-scale trends in the July temperature anomalies curve (Figure 4.3; 4.4). Regional temperature anomaly patterns are small on the order of ± 1°C during the first 9,000 years BP, whereas between 14,000 and 9,000 BP, the anomalies are on the order of ± 2-4°C. The reconstructed spatial patterns of temperature support the hypothesis that century to millennial-scale climate variability is associated with reorganization of ocean-atmospheric circulation patterns.

4.2.4 CORRELATION WITH OTHER PROXY RECORDS

Marine ice-rafting events (IRD) are indicative of cooler ice-filled surface waters advected deeper southward and eastward from Nordic and Labrador seas (Bond *et al.*, 1997, 2001). This has been linked to reduction or cessation of North Atlantic Deep Water Formation (NADW) and the Thermohaline circulation (THC) during glacial and interglacial climate regimes (Bond *et al.*, 1997; Broecker and Denton, 1989) and successfully modeled assuming an increase in fresh meltwater flux into the North Atlantic (Sakai and Peltier, 1994). IRD events are also indicative of basin-wide cooler North Atlantic climates. Recently, the marine IRD record has been correlated with the tree-ring derived terrestrial ^{14}C and ^{10}Be cosmogenic nuclide records which are proxies for solar variability and ocean forcing (Bond *et al*, 2001). Virtually all IRD events fall closely within a cooling period in this record providing support for widespread climate impact (Figure 4.8). Cross-spectral coherency of the marine and nuclide records at the 300-500 and 900 to 1100-year band (Bond *et al.*, 2001) correlates reasonably well with the 1150±125 year periodic component identified in this study.

4.3 DISCUSSION AND CONCLUSIONS

Results presented in this study reveal several important aspects of Holocene and late glacial climate variability. They demonstrate that pollen time series can be used to explore century to millennial-scale climate variability at the continental-scale of North America. A mean July temperature reconstruction for North America identified temperature variability at several scales. The detrended temperature time series revealed a periodic 1150 yr periodicity that can resolve the mismatches between the different climate-proxy records. As discussed in Chapter 3, ambiguity in the correlation of some of these transitions between the different climate proxy records revolves around certain key periods when lower-frequency forcings dominate, such as the abrupt warming around 9000 years BP, the middle Holocene maximum around 6000 years ago and neoglacial cooling circa 4000 years ago. One plausible explanation is that millennial-scale events would be attenuated around these preferred times due to large-scale interactions, thus not being easily identifiable in the different records. Virtually all significant dates identified in the different climate-proxy records are identified in the cool millennial-scale troughs in our record. These major millennial-scale climate reorganizations occurred at *circa* 600, 1600, 2800, 4000, 5200, 6300, 7400, 8500, 9600,10700,11800, and 12900 cal/yr BP, and these correlate reasonably well with the marine, terrestrial and ice records.

What causes these reorganizations between cool and warm modes of ocean-atmospheric operation? Bond *et al.*, (2001) showed a strong correlation between North Atlantic IRD events and the $^{14}C_{atm}$ and ^{10}Be records which are proxies for solar output variability, and postulated a variable solar output hypothesis to explain century to millennial-scale climate variability. The majority of millennial-scale marine IRD events fall within cooling periods in our record. A periodic component of 1150-year and associated century-scale cycles are reasonably within the range of the cross-spectral band of 300-500 years and 900-1100 years relationship identified between the marine and cosmogenic nuclide record. Therefore this study supports the solar-climate relationship hypothesis.

A relationship between solar variability and climate response has long been debated; resolution of this issue is inhibited because the magnitude of change of solar output is supposedly too weak to cause changes in atmospheric circulation patterns. Several processes in the climate system can amplify these solar output changes. Model simulations suggest that a change of less that 0.1 % in total irradiance could presumably affect surface climates through stratosphere/troposphere dynamic processes (Haigh, 1996, Shindell *et al.*, 1999). A weak solar output decrease on the order of 0.24% has been shown to affect surface climates using sun-like stars as proxies (Lean *et al*, 1992). Moreover, it is now recognised that solar output variability is much greater in the UV spectrum than the visible (SOHO experiment on solar wind; IPCC, 2001). These small variations in solar output would in turn affect stratospheric ozone production, therefore warming or cooling the stratosphere which in turn would, through dynamical processes, influence upper-troposphere wave patterns and thus surface climates.

Assuming that the relationship between our record, the marine IRD and the ^{14}C and ^{10}Be record exists, an externally driven variable solar output would not account for the asymmetry and increased magnitude of change observed during the Holocene and late glacial period. Interactions between orbital, solar variable output, and internal feedback mechanisms are probably needed to explain the surface observations.

Variations in cosmogenic nuclide production are related to solar winds and solar activity. Increases in the cosmogenic nuclides result from a lowered solar activity and diminished winds while a decrease in cosmogenic nuclides indicate an increase in solar activity and winds (Bond *et al.*, 2001). Therefore, IRD events, which are representative of North Atlantic basin-wide cooling events, are associated with an increase in nuclide production and therefore diminished solar irradiance. It is also well-documented that IRD events are related to changes in the thermohaline circulation (THC) through salinity and temperature processes (NADW formation), which have been linked to fresh meltwater fluxes into the North Atlantic from the Nordic seas (Broecker *et al.*, 1989, 1995, 1997). The interesting point here is that IRD events correlate with decreased solar activity. An ocean-atmospheric reoganization following an increased solar activity would be a more likely candidate to provide the trigger mechanism for the meltwater fluxes and IRD deposition into the North Atlantic and associated changes in THC at least during interglacials.

A possible scenario is as follows. An increase in solar activity leads to a reorganization of the ocean-atmospheric circulation through changes in stratospheric ozone production and dynamical processes in the upper-troposphere. Follows a warming of the northern high latitude atmosphere and a slight northern shift of the northern hemisphere subtropical jet (Shindell *et al.*, 1999). The ocean-atmospheric response to increased irradiance would lead to increased continental ice and sea-ice retreat, which would be further amplified through surface-albedo feedbacks. This, in association with a contraction of polar vortex, could presumably provide increased fresh meltwater influx, through continental runoff and iceberg discharge, into the North Atlantic through the Labrador and Nordic sea outlets. As gradually decreasing solar activity followed, deposits of IRD deep within warmer regions of the subpolar ocean circulation would result, in phase with decreasing solar irradiance. As solar irradiance decreases even more, this would lead to decreased heating of the stratosphere through ozone production processes. The cooling would be propagated downward from the stratosphere to the troposphere, causing a cooling of the northern high latitude atmosphere and a slight southward migration of the northern subtropical jet (Shinddell *et al.*, 1999). Cooler northern latitude ocean surface temperatures would further amplify an already less saline North Atlantic to a threshold by which NADW formation would decrease or cease providing the necessary trigger for a diminished or halted THC. This would in turn provide an additional amplification for a basin wide North Atlantic cooling. The ocean-atmospheric response to decreased solar activity along with the amplification feedbacks would lead to renewed expansion of continental ice and sea-ice extent in the northern high latitudes until solar irradiance increased once again. The sun may indeed be the external modulator for decadal to millennial-scale climate variability, but internal feedback mechanisms are required to explain a global variable asymmetric oscillating response.

The results presented here show climate variability on the order of a maximum of ± 0.4°C temperature magnitude during the Holocene and ± 0.6°C during the late glacial period. These continental-scale temperature variations on century to millennial-scales are comparable to those estimated for the last 1,000 years, for example the Litle Ice Age (LIA) and preceding Medieval Warm Period (MWP)(Mann *et al.*, 1999; Esper *et al.*, 2002). These are the most recent manifestations of century to millennial-scale temperature variations.

Although not present in our analysis, due to resolution and methodology, the recent global warming has been measured as a 0.4 to 0.6°C increase in global temperatures in the past 150 years, with most of the increase since the 1960's (IPCC, 2001). The abruptness and magnitude of this recent temperature increase is truly exceptional. Even being conservative and allowing a 0.4°C magnitude temperature change due to the natural climate variability, this would leave approximately 0.2 °C increase in hemispheric temperatures not observed during the past 10,000 years. Futhermore, this has occurred during the northern hemispheric neoglacial cooling.

If human influence through atmospheric CO_2 increases can account for a minimum of 0.2°C above the 0.4°C centennial to millennial-scale temperature change over the past 10,000 years in less than 150 years, the upper bound of predicted global warming scenarios for the next 100 years maybe a more realistic estimate. Skepticism over the fact that global warming is merely a natural recovery from the LIA therefore appears questionable, based on our analysis. The magnitude and rate of acceleration of warming over the past 150 years is unprecedented even when superimposing natural century to millennial scale modern warming. It appears more likely that human-induced concentrations of atmospheric CO_2 has and will continue to play a major role in our climate evolution of the near future. Therefore, an international cooperation at political, economic and research levels is required for planning response strategies for the next generations.

Figure 4.1: Fossil pollen site distribution in North America (source:NAPD, Grimm, 2000).

Time	# sites
0-1000	639
1000-2000	566
2000-3000	560
3000-4000	553
4000-5000	545
5000-6000	532
6000-7000	513
7000-8000	478
8000-9000	470
9000-10000	453
10000-11000	390
11000-12000	304
>12000	253

Table 4.1: Site density through time in North America (source: NAPD, Grimm, 2000). Periods based on radiocarbon intervals.

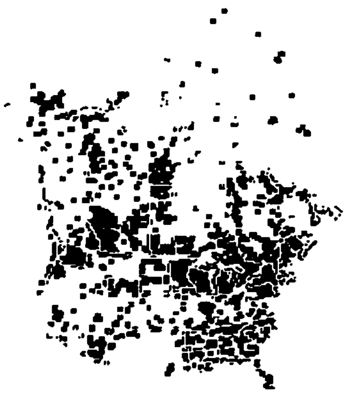


Figure 4.2: Modern pollen site distribution in North America (source: WWKGVS2000).

Table 4.2: Pollen sum used in North American climate reconstruction.

v1 *Abies*	v31 *Dryas*	v61 *Picea mariana*
v2 *Acer*	v32 *Ephedra*	v62 *Picea undiff.*
v3 *Alnuscrispa*	v33 *Ericaceae/Chamadaphne/Ledum*	v63 *Pinus strobus*
v4 *Alnusrugosa*	v34 *Ericaceae Empetrum*	v64 *Pinus undiff.*
v5 *Alnus undiff.*	v35 *Ericaceae Vaccinium*	v65 *Plantaginaceae*
v6 *Ambrosia type*	v36 *Ericaceae undiff.*	v66 *Platanus*
v7 *Amorpha*	v37 *Eriogonum*	v67 *Poaceae*
v8 *Apiaceae*	v38 *Euphorbiaceae*	v68 *Polygonum*
v9 *Aquifoliaceae*	v39 *Fabaceae*	v69 *Polypodiaceae*
v10 *Arceuthobium*	v40 *Fagus*	v70 *Populus*
v11 *Artemisia*	v41 *Fraxinus*	v71 *Potentilla*
v12 *Asteraceae*	v42 *Juglans*	v72 *Prosopis*
v13 *Betula undiff.*	v43 *Larix Pseudotsuga*	v73 *Pteridium*
v14 *Brassicaceae*	v44 *Larrea*	v74 *Quercus*
v15 *Cactaceae*	v45 *Liliaceae*	v75 *Ranunculaceae*
v16 *Campanulaceae*	v46 *Liquidambar*	v76 *Rhamnaceae*
v17 *Caprifoliaceae*	v47 *Lycopodium annotinum*	v77 *Rubiaceae*
v18 *Carya*	v48 *Lycopodium clavatum*	v78 *Rumex*
v19 *Caryophyllaceae*	v49 *Lycopodium complanatum*	v79 *Salix*
v20 *Castanea*	v50 *Lycopodium selago*	v80 *Sarcobatus*
v21 *Ceanothus*	v51 *Lycopodium undiff.*	v81 *Saxifragaceae*
v22 *Celtis*	v52 *Magnoliaceae*	v82 *Selaginella*
v23 *Cephalanthus*	v53 *Malvaceae*	v83 *Shepherdia*
v24 *Cercocarpus*	v54 *Myrica*	v84 *Sphagnum*
v25 *Chenopodiaceae/Amaranthaceae*	v55 *Nyssa*	v85 *Taxodium*
v26 *Chrysolepis*	v56 *Onagraceae*	v86 *Thalictrum*
v27 *Clethra*	v57 *Ostrya/Carpinus*	v87 *Tilia*
v28 *Corylus*	v58 *Oxyria*	v88 *Tsuga*
v29 *Cupressaceae*	v59 *Papaveraceae*	v89 *Ulmus*
v30 *Cyperaceae*	v60 *Picea glauca*	

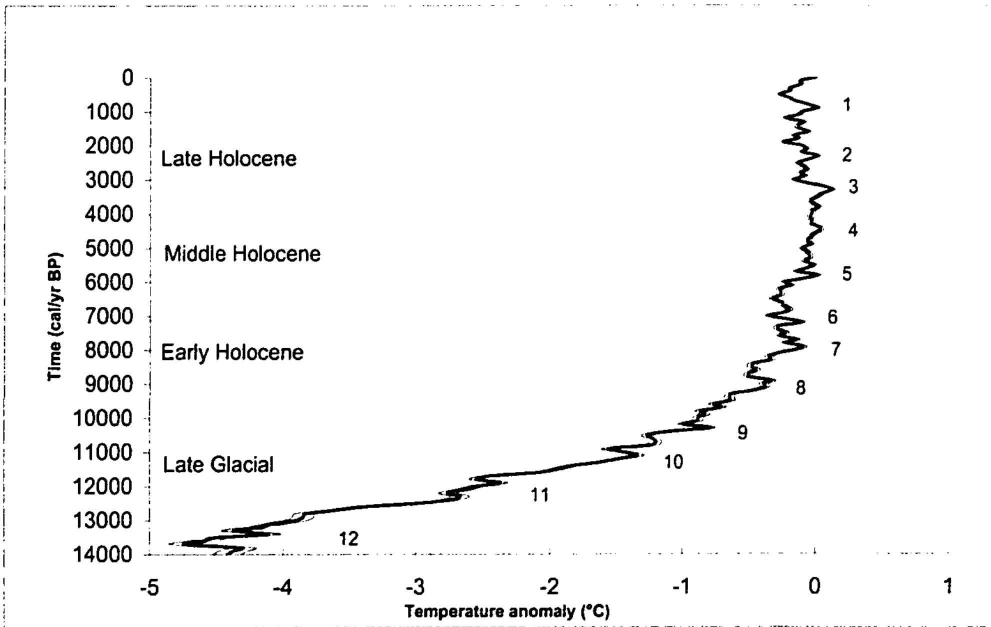


Figure 4.3: Mean North American July temperature anomaly reconstruction for the past 14,000 cal/yr BP. 95% confidence level corrected for trends. Numbers represent millennial-scale peaks.

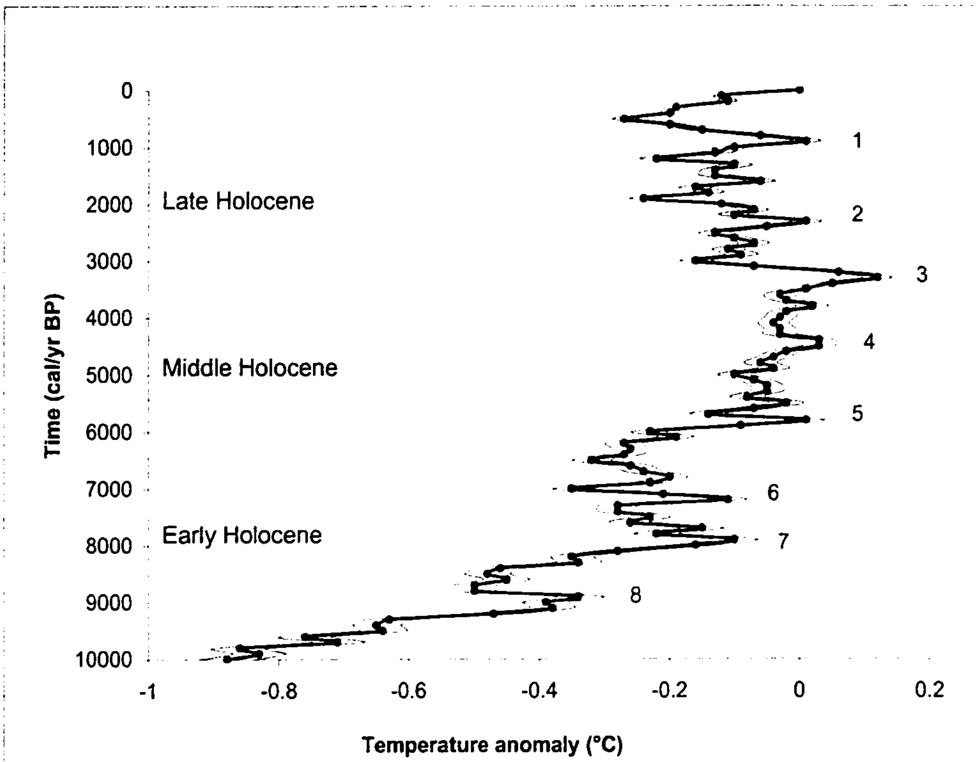


Figure 4.4: Mean North American July temperature anomaly reconstruction for the past 10,000 cal/yr BP. 95% confidence level corrected for trends. Numbers represent millennial-scale peaks.

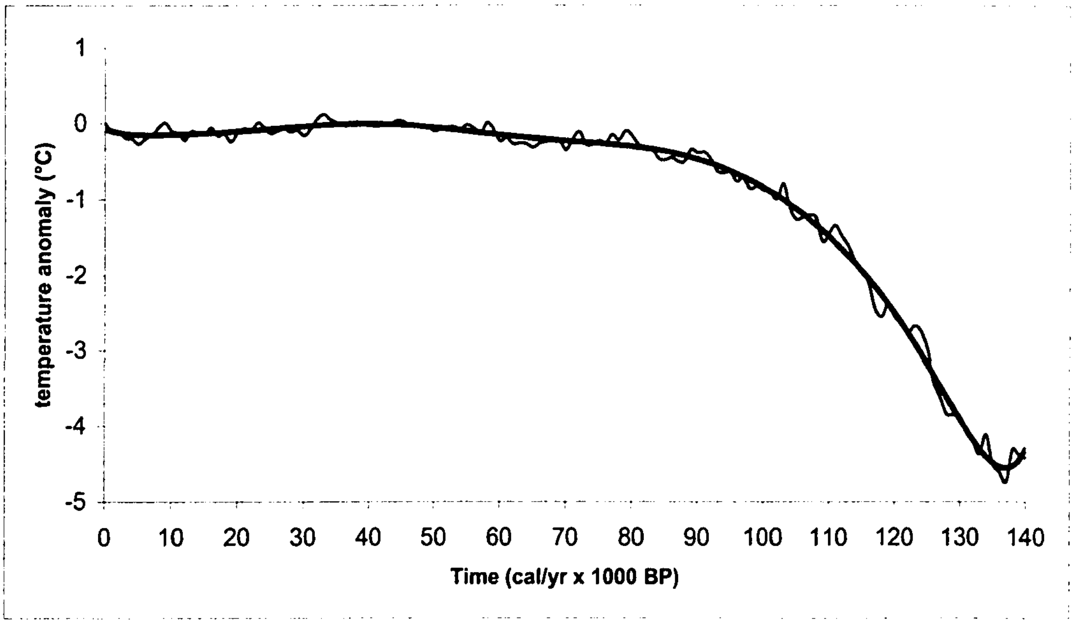


Figure 4.5: Detrending of the mean July temperature anomaly of North America using 10th order polynomial. Adjusted r^2 of .99.

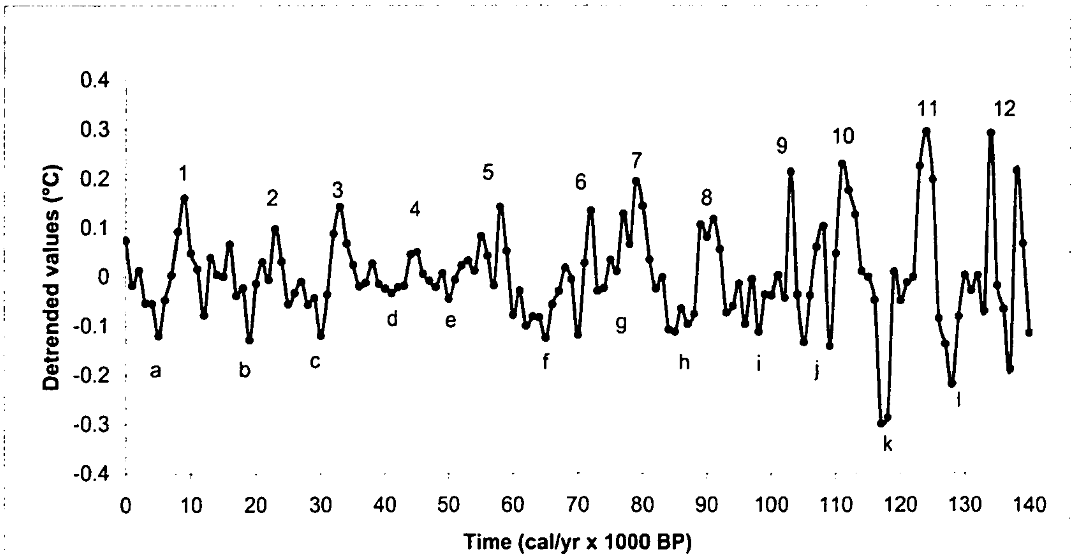


Figure 4.6: Detrended time series. Numbers associated with millennial-scale peaks as in Figure 4.3 and 4.4. Letters are millennial-scale troughs.

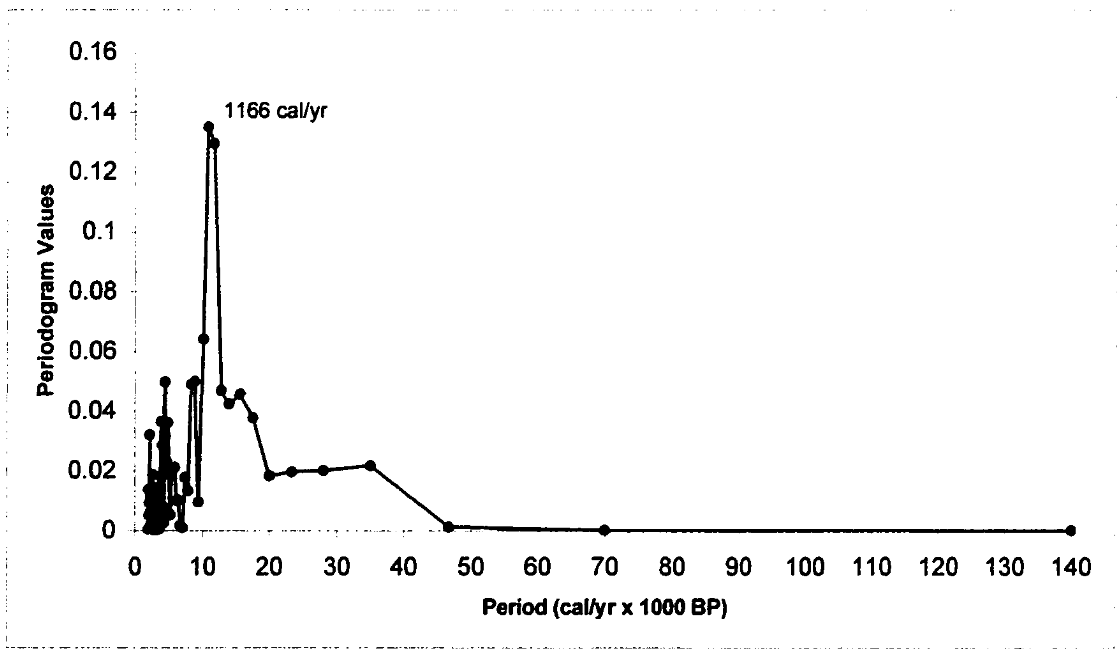


Figure 4.7: Spectral analysis of detrended mean July temperature curve of North America derived from pollen assemblages.

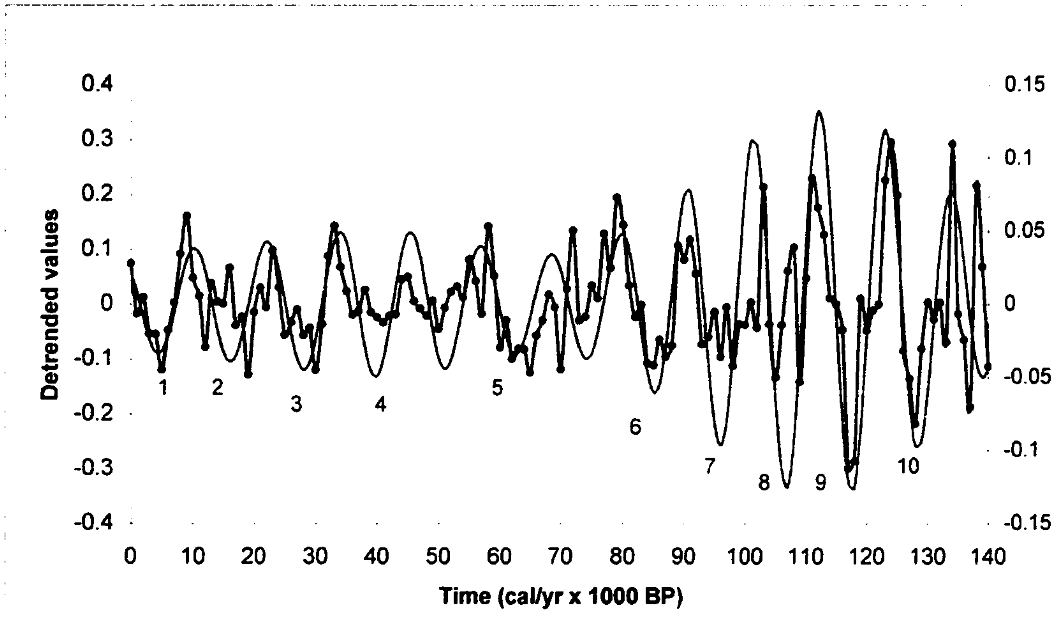


Figure 4.8: Bandpass filtering of actual mean July temperature curve of North America. Frequency range used is 0.0781 to 0.1094 or 1/1100 to 1/1200 cycles. Secondary *y-axis* indicates bandpass anomaly values. Numbers indicate marine IRD events identified in marine cores of the North Atlantic (Bond *et al.*, 1997, 2001).

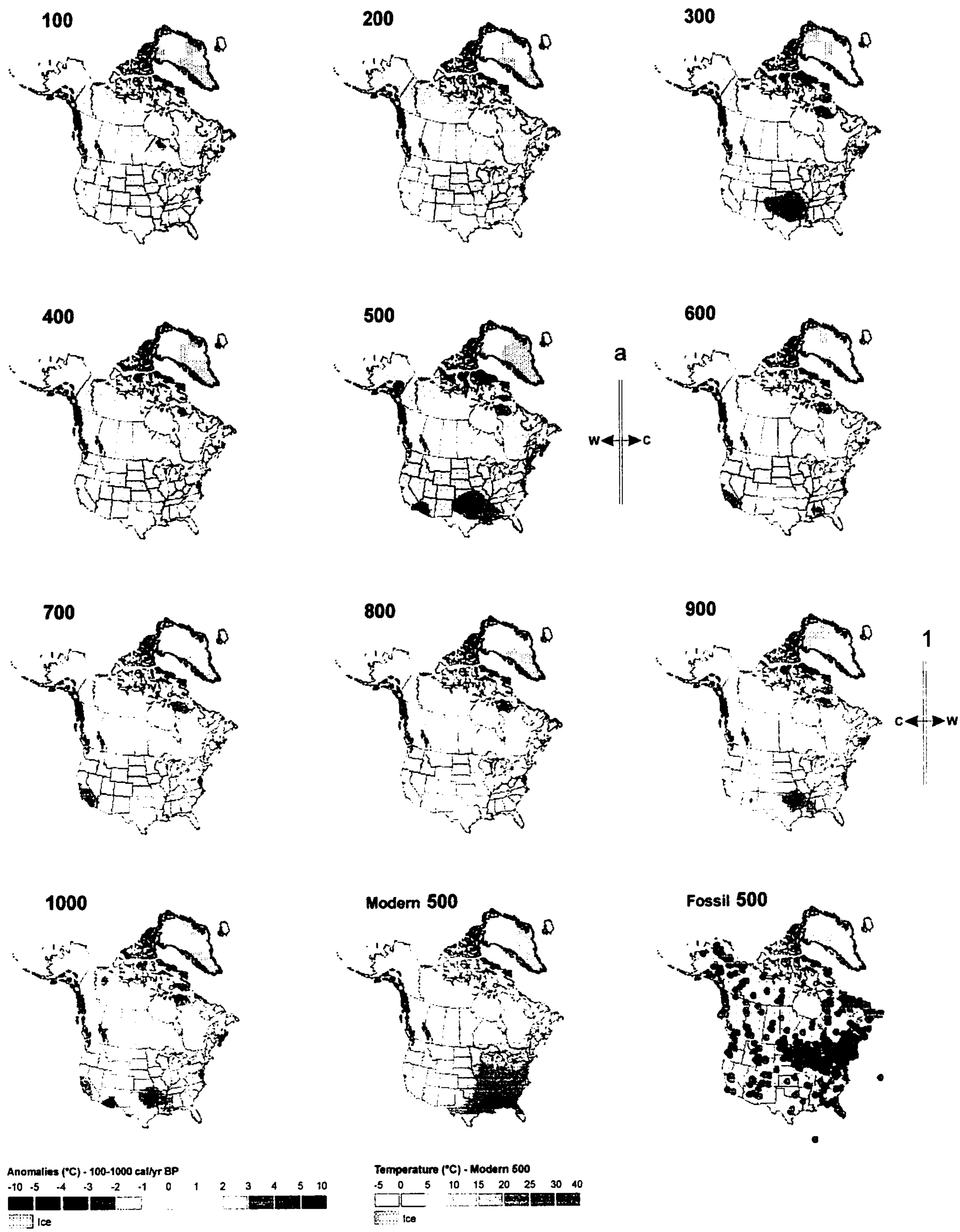


Figure 4.9 a: Temperature anomaly maps for North America for 100-1000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 500 cal/yr BP.

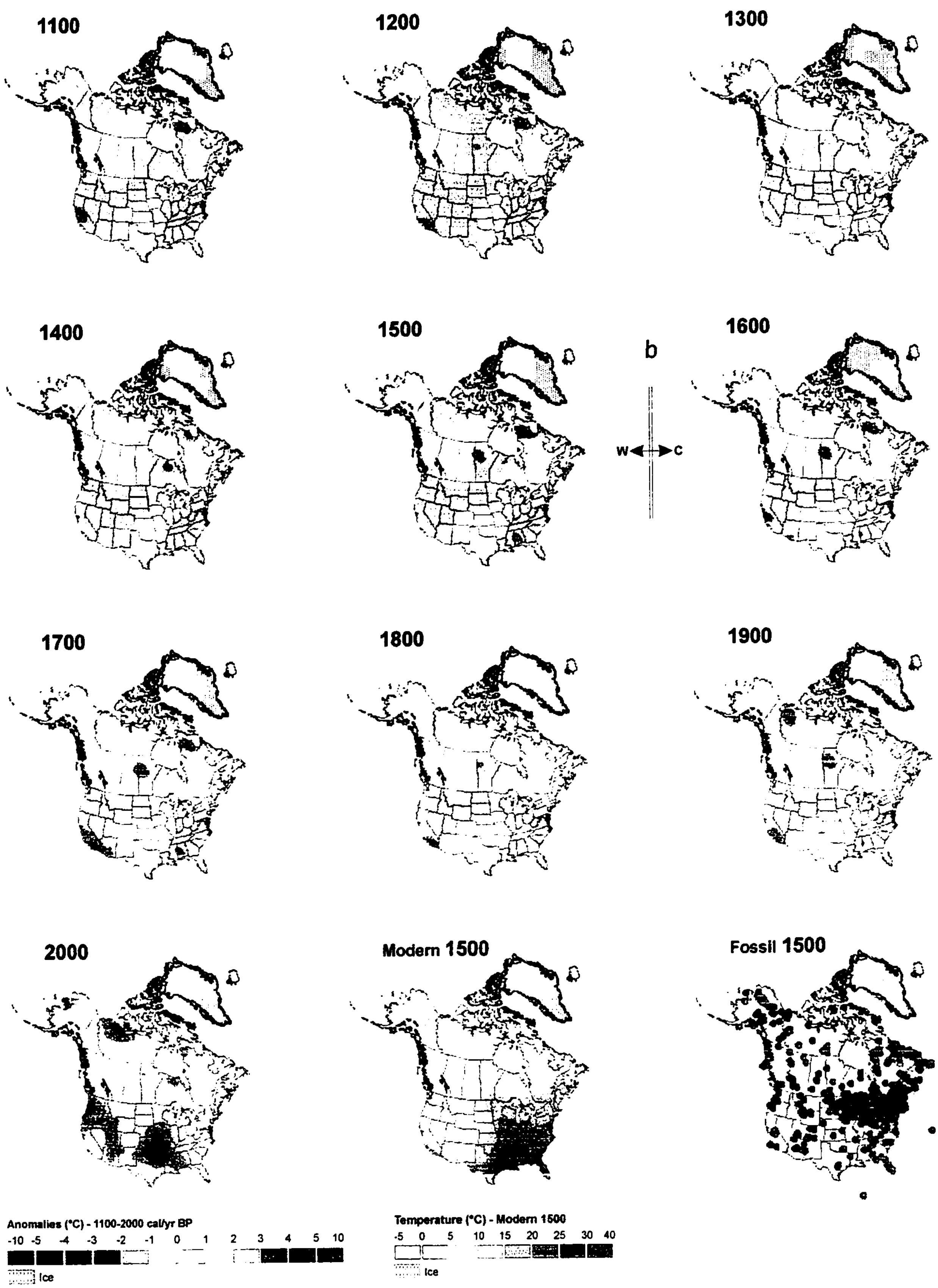


Figure 4.9 b: Temperature anomaly maps for North America for 1100-2000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 1500 cal/yr BP.

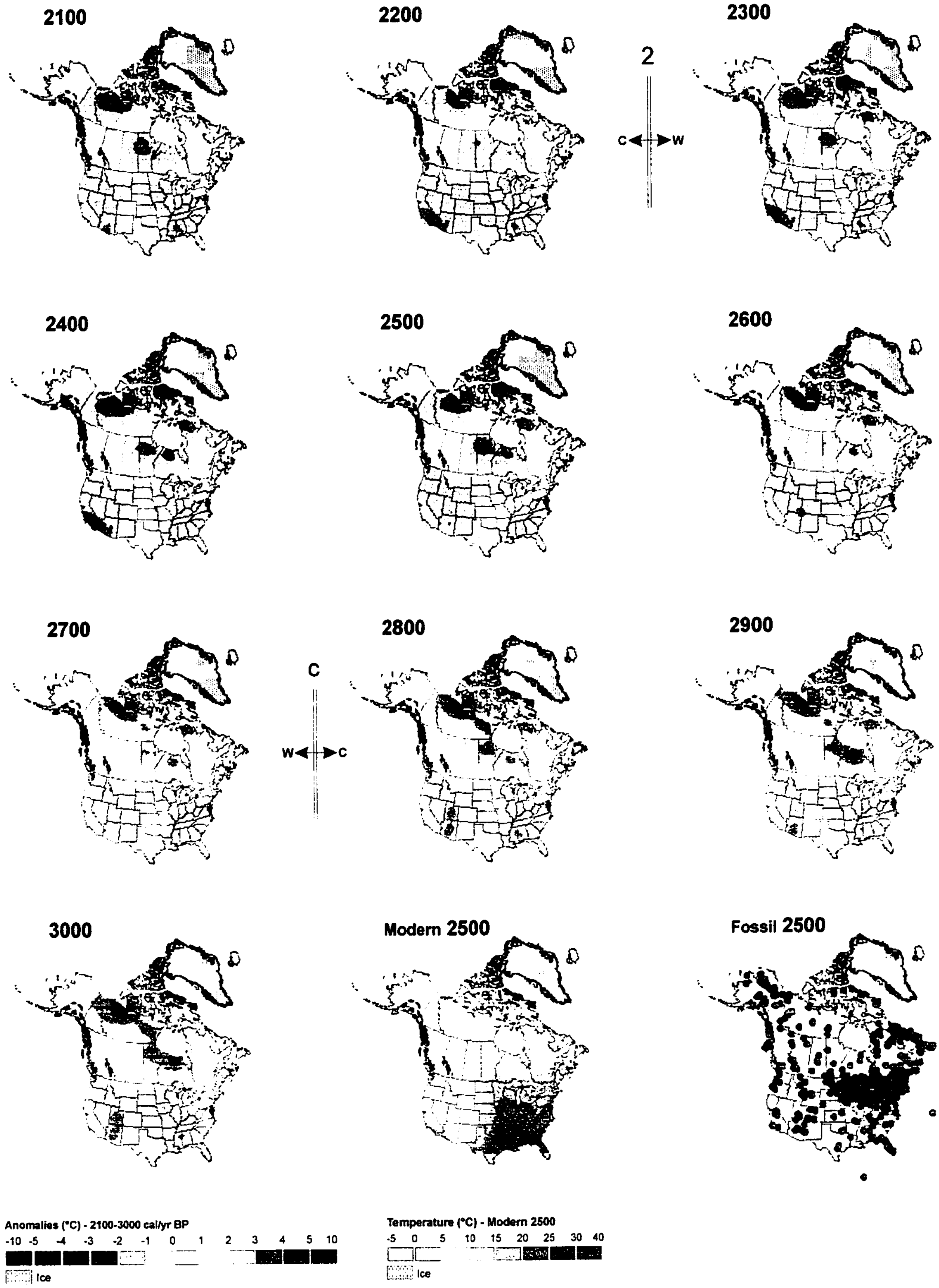


Figure 4.9 c: Temperature anomaly maps for North America for 2100-3000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 2500 cal/yr BP.

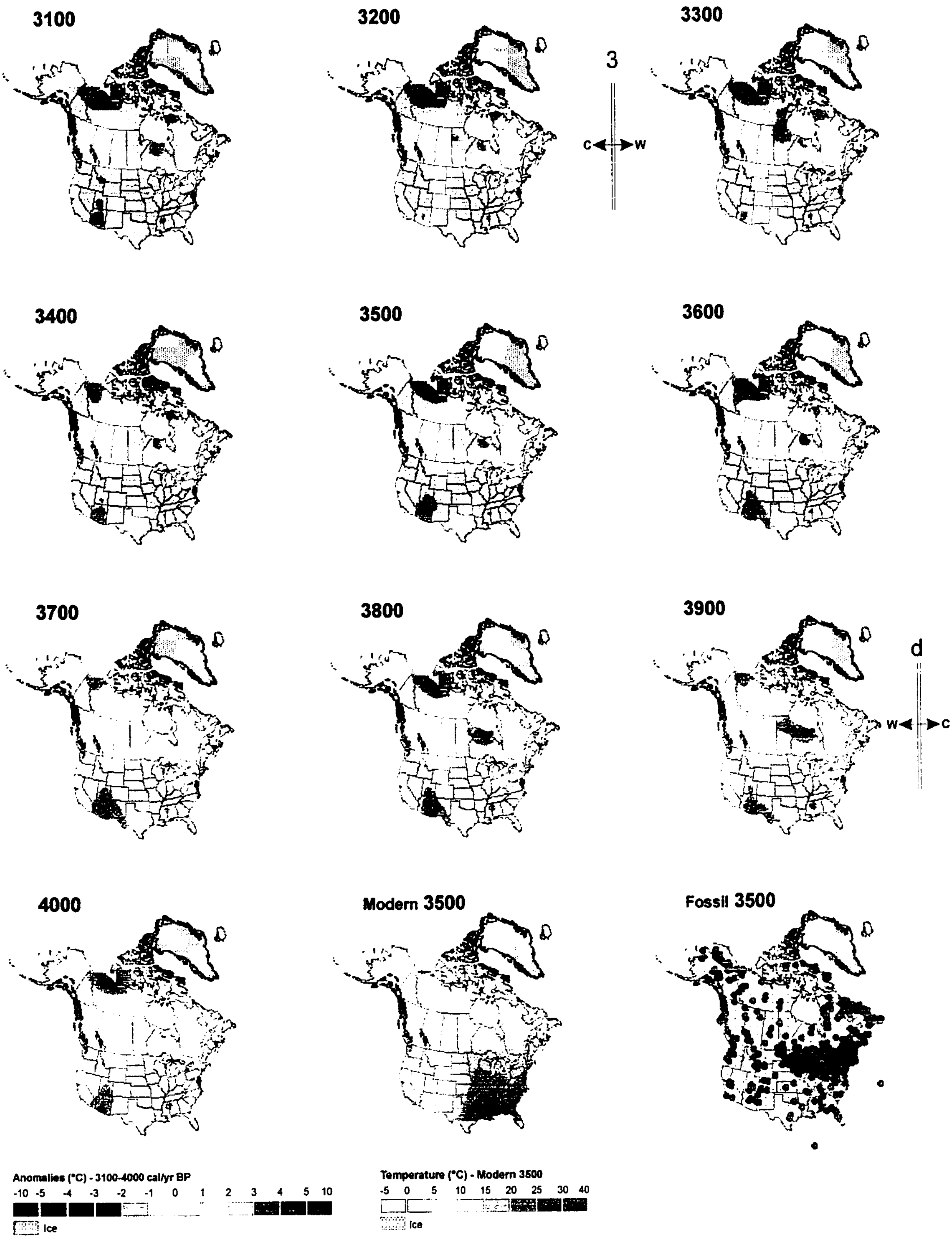


Figure 4.9 d: Temperature anomaly maps for North America for 3100-4000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 3500 cal/yr BP.

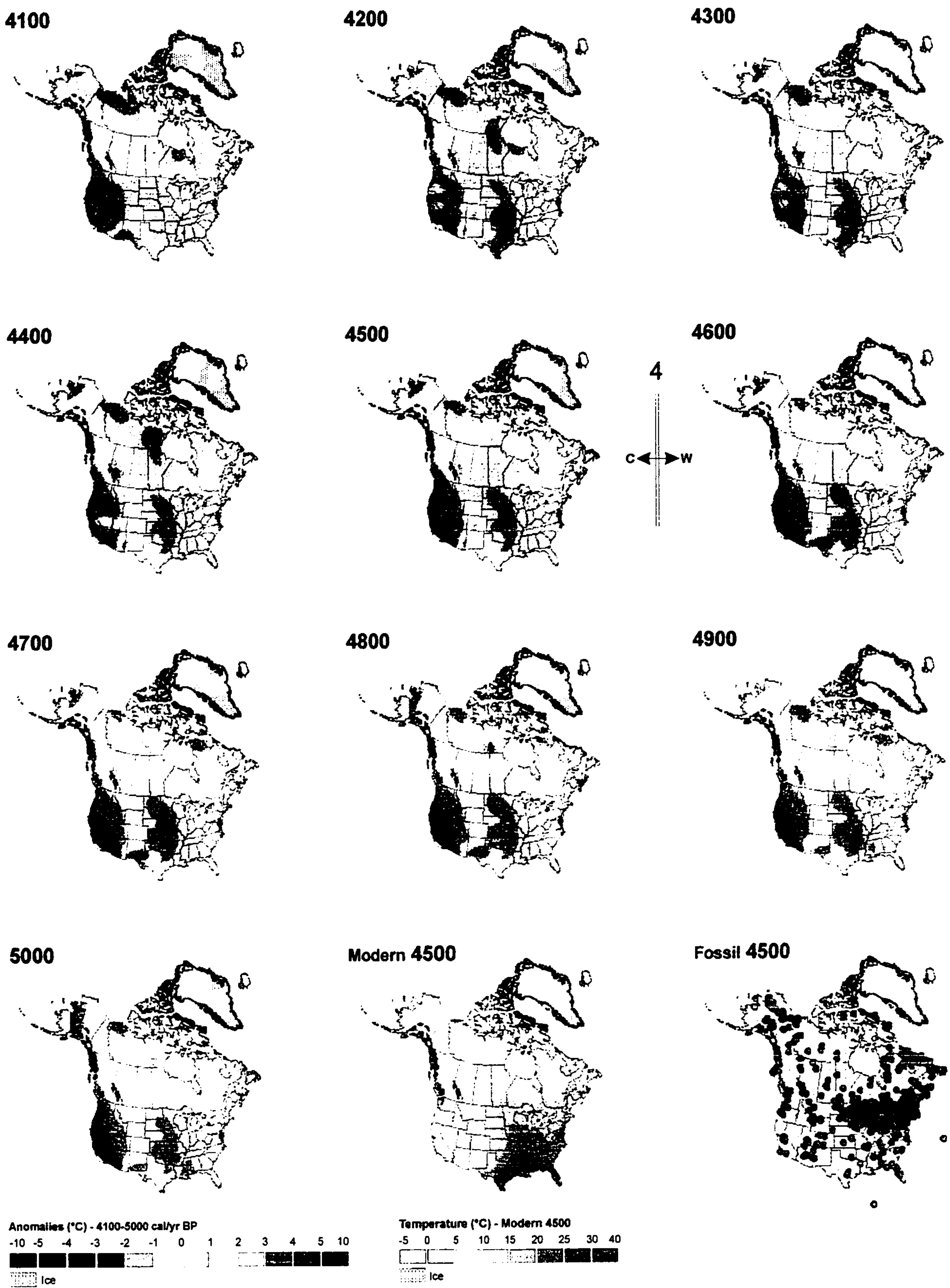


Figure 4.9 e: Temperature anomaly maps for North America for 4100-5000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 4500 cal/yr BP.

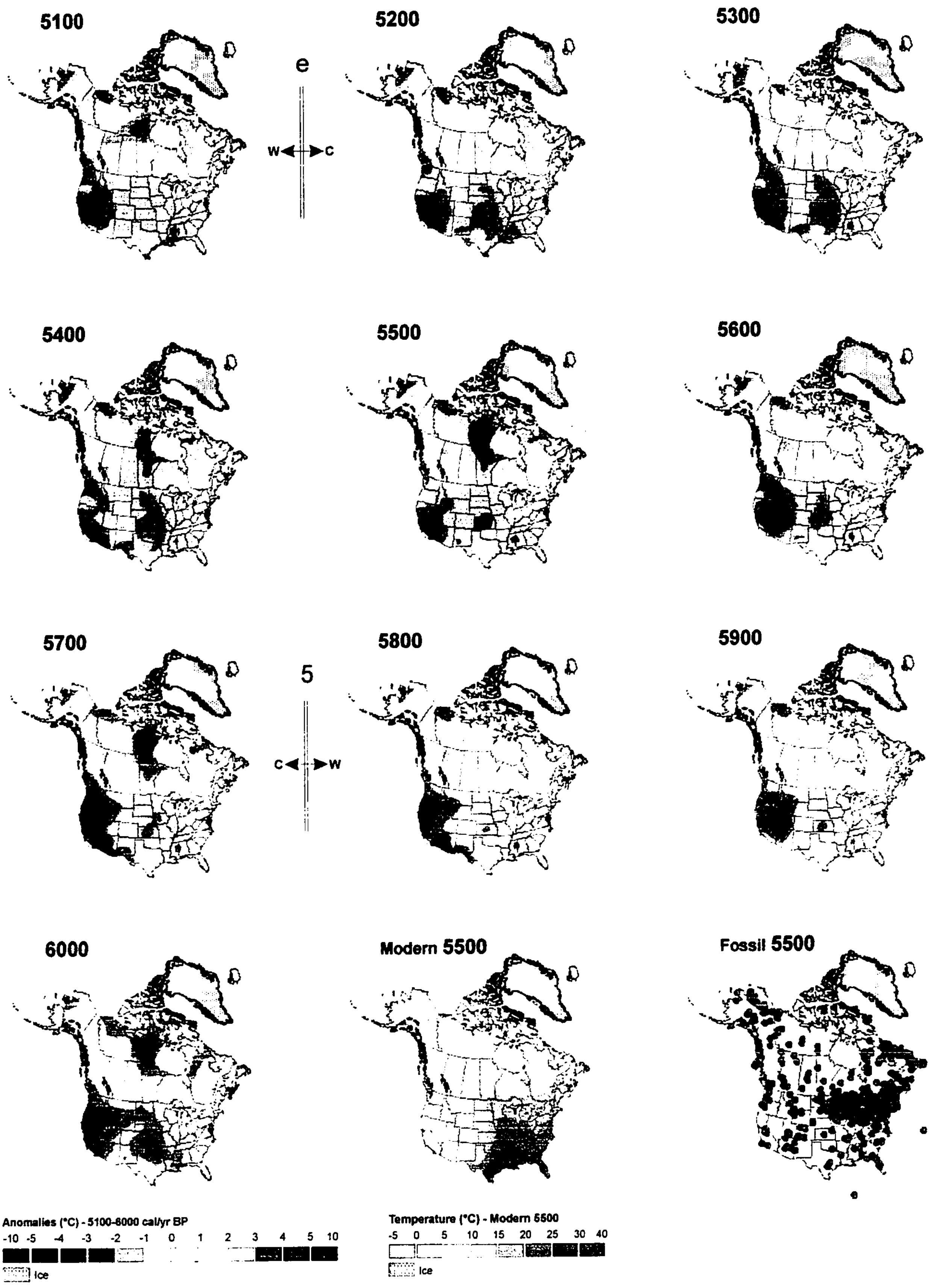


Figure 4.9 f: Temperature anomaly maps for North America for 5100-6000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 5500 cal/yr BP.

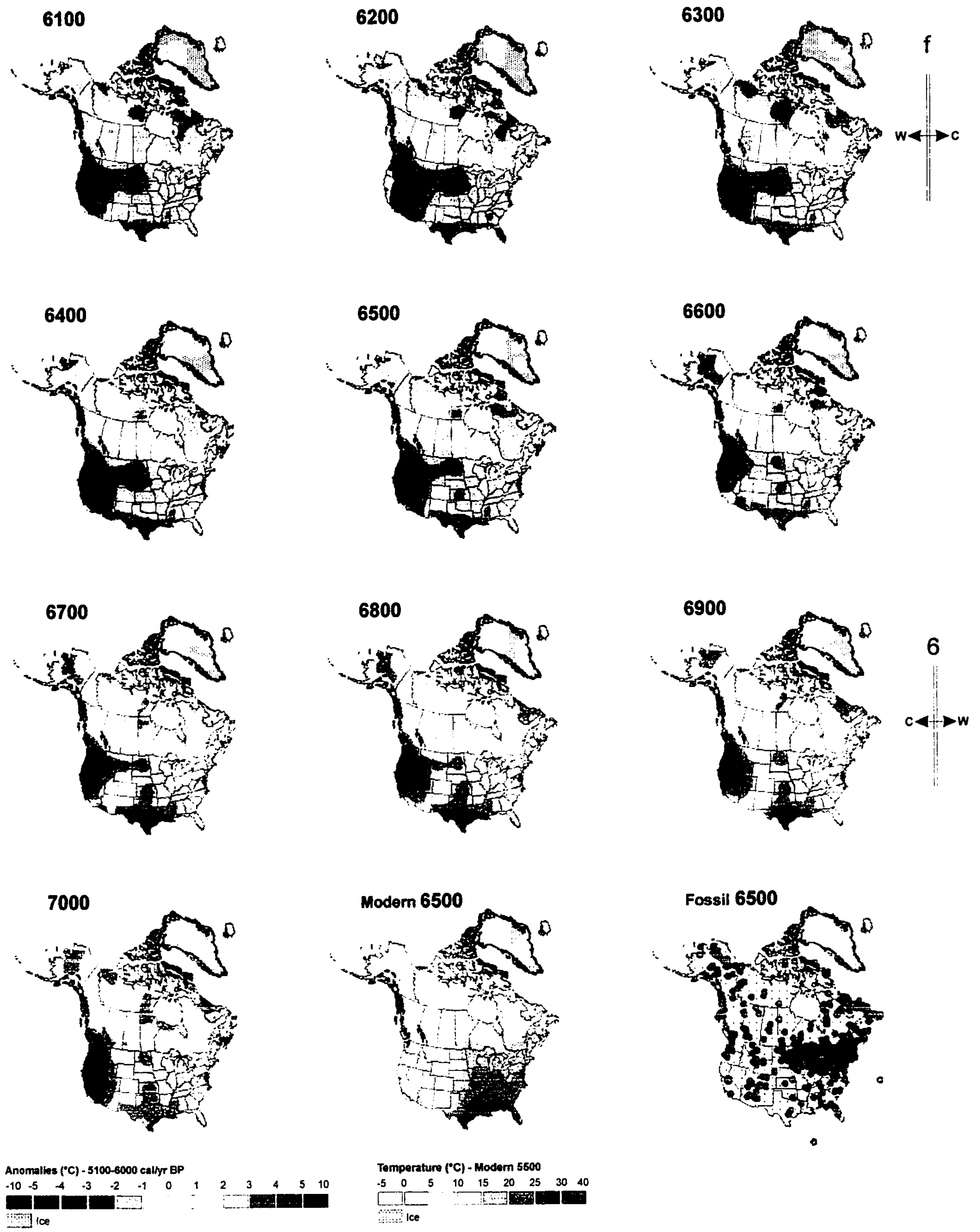


Figure 4.9 g: Temperature anomaly maps for North America for 6100-7000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 6500 cal/yr BP.

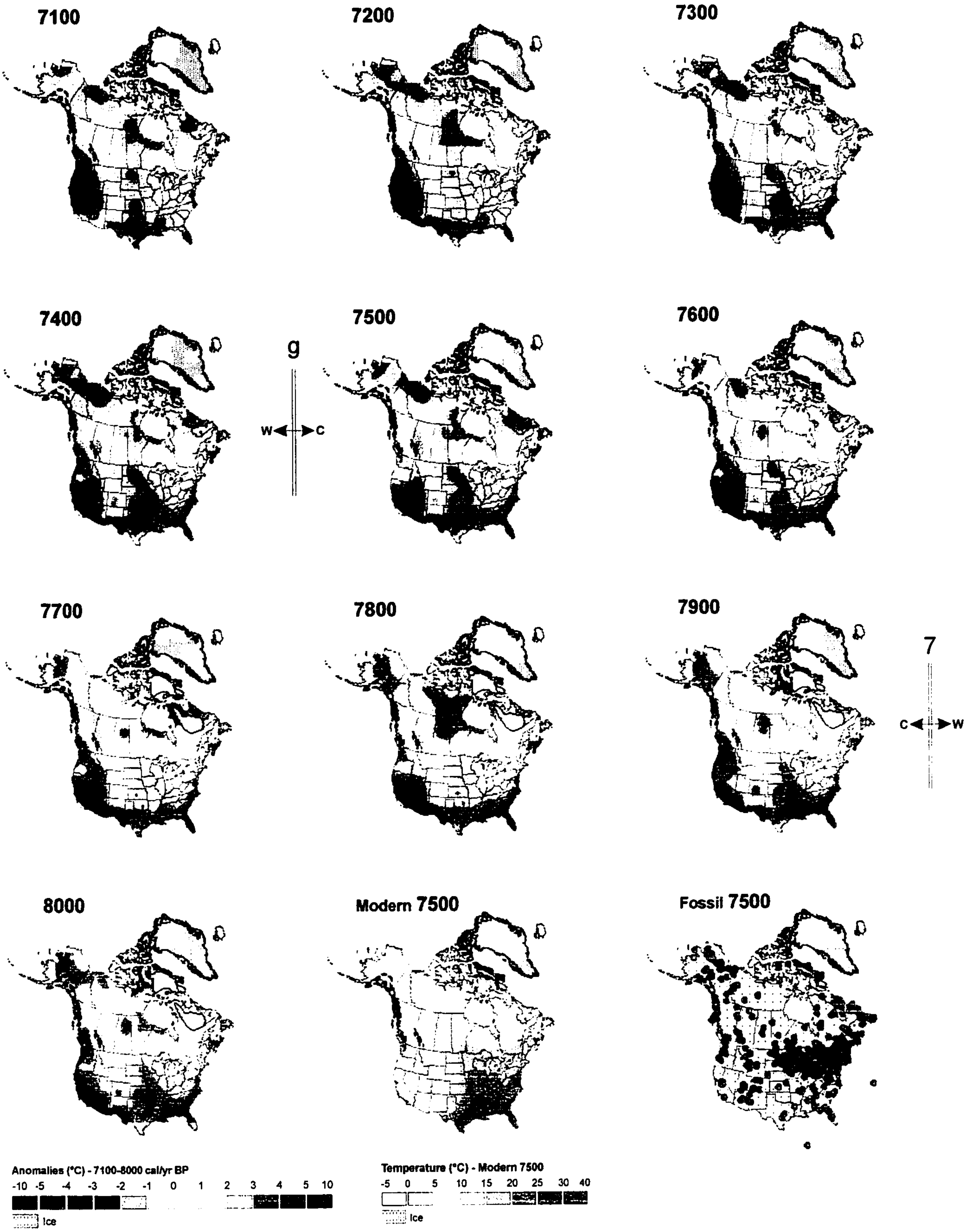


Figure 4.9 h: Temperature anomaly maps for North America for 7100-8000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 7500 cal/yr BP.

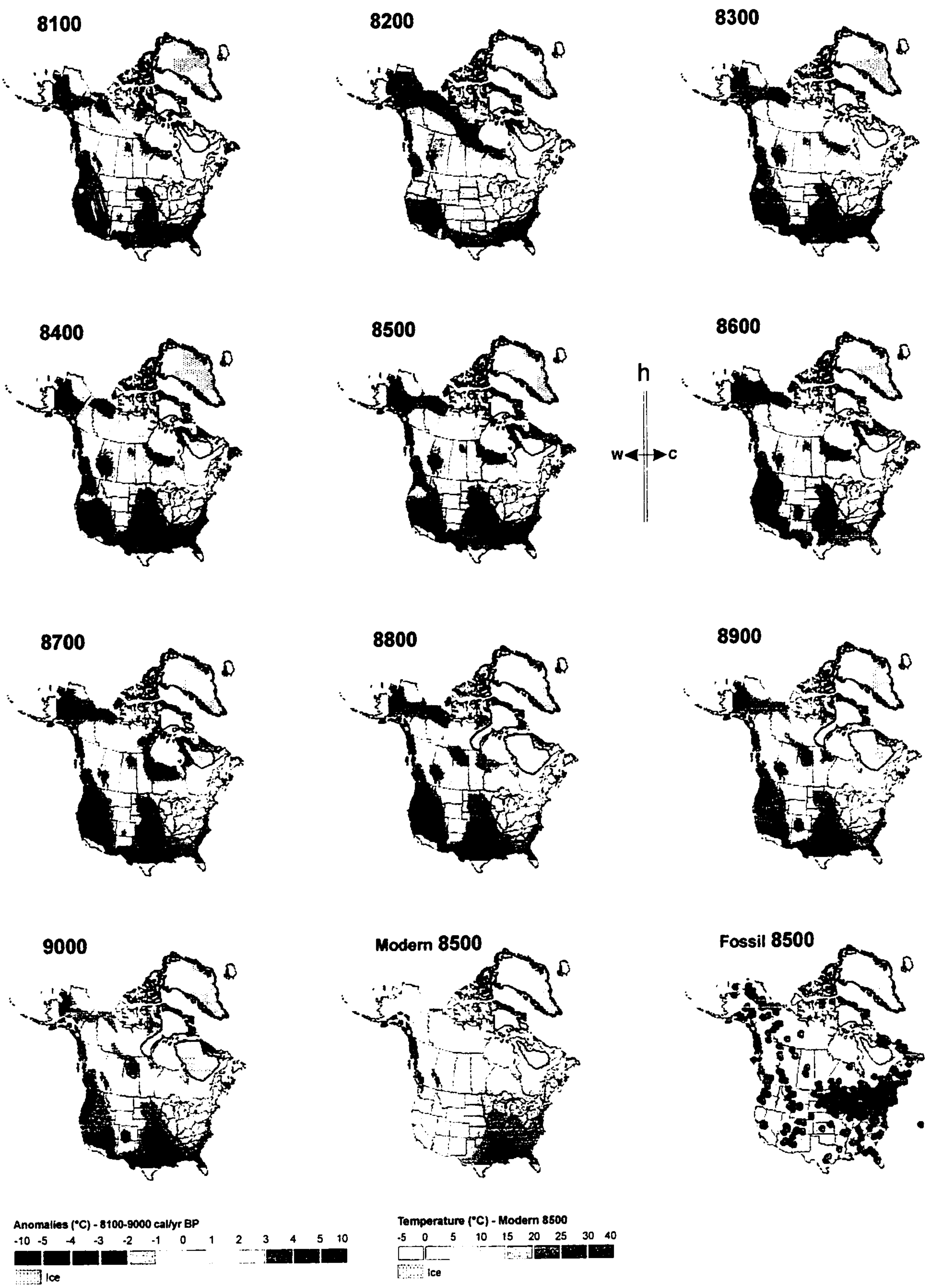


Figure 4.9 i: Temperature anomaly maps for North America for 8100-9000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 8500 cal/yr BP.

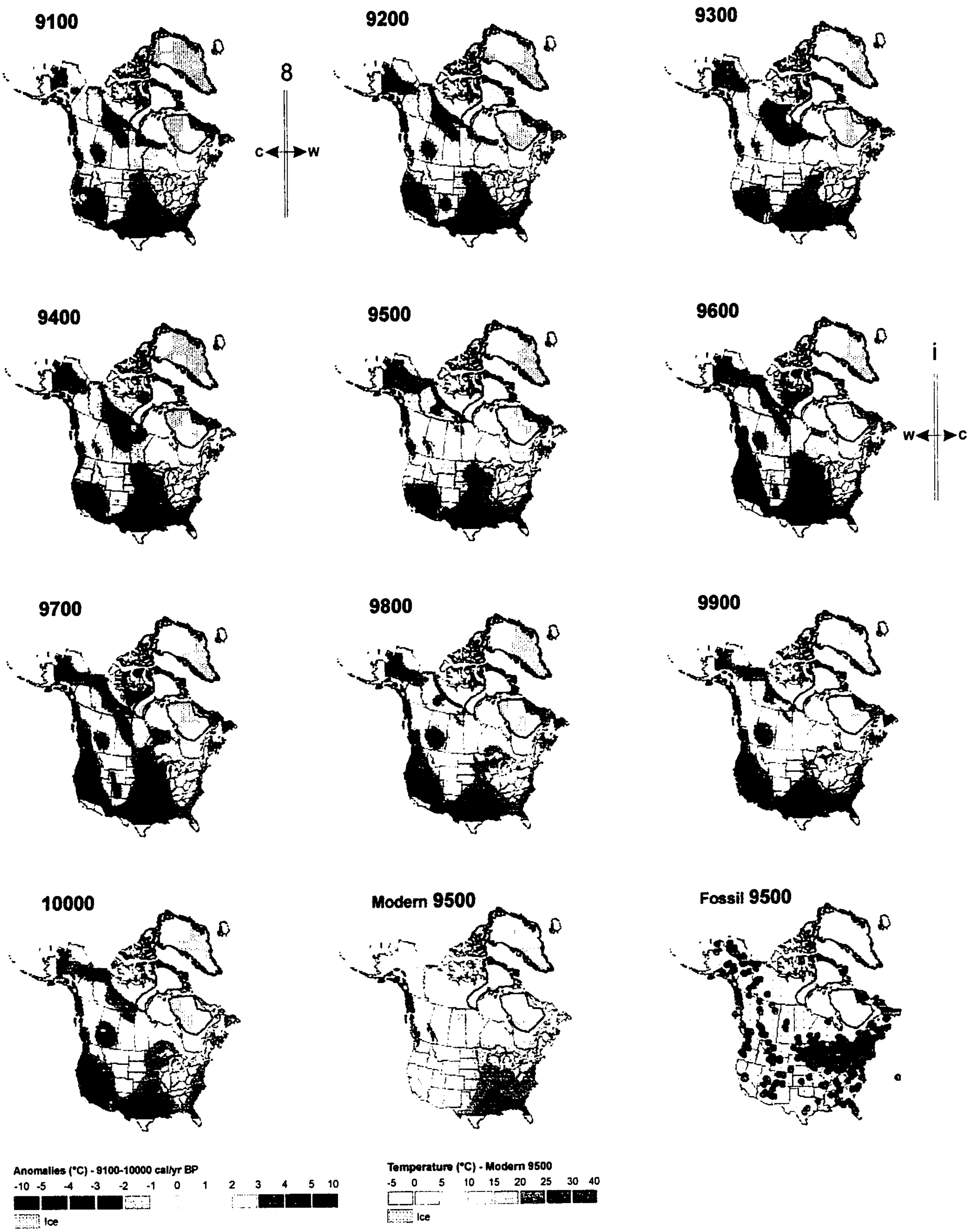


Figure 4.9 j: Temperature anomaly maps for North America for 9100-10000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 9500 cal/yr BP.

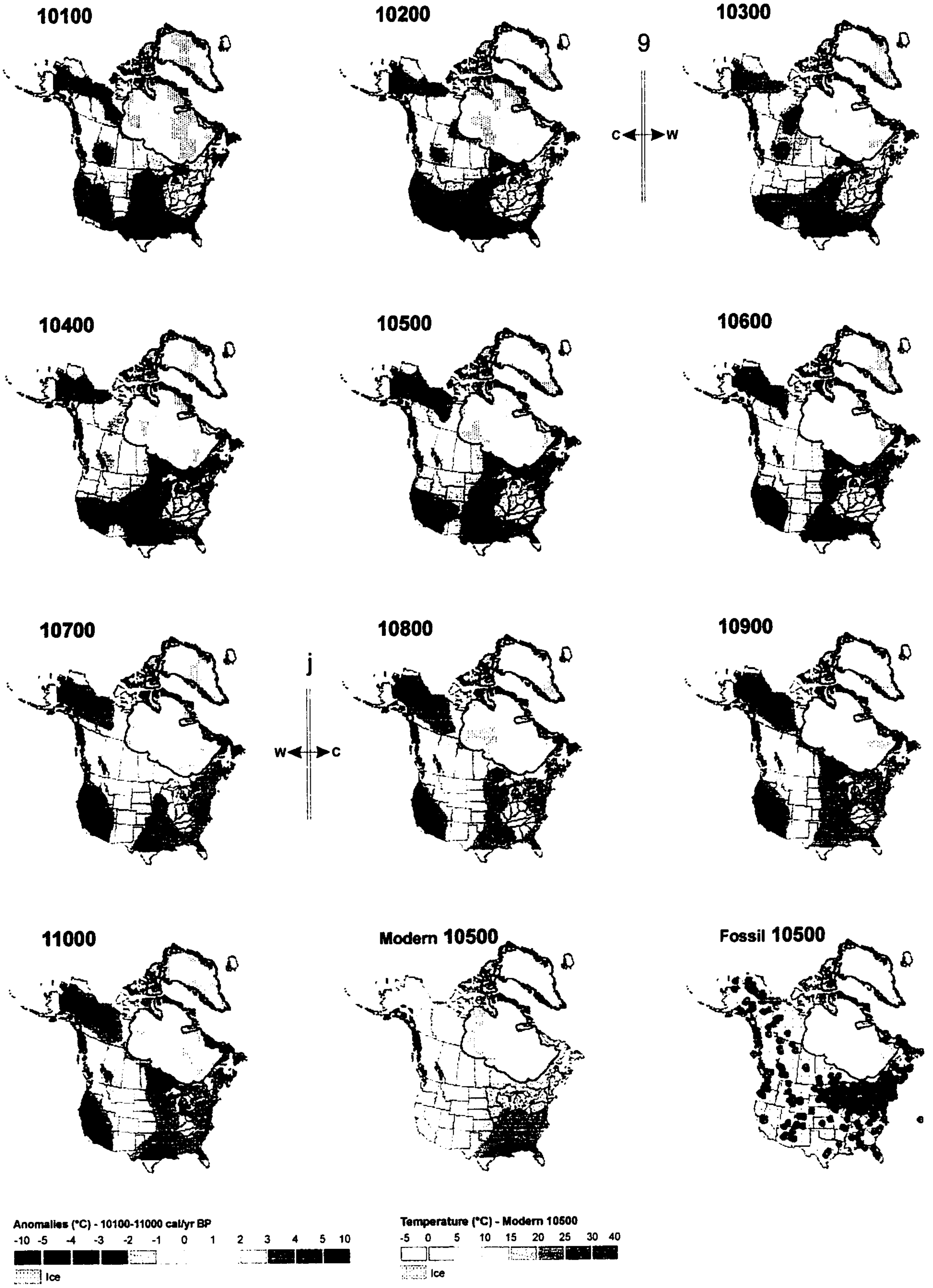


Figure 4.9 k: Temperature anomaly maps for North America for 10100-11000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 10500 cal/yr BP.

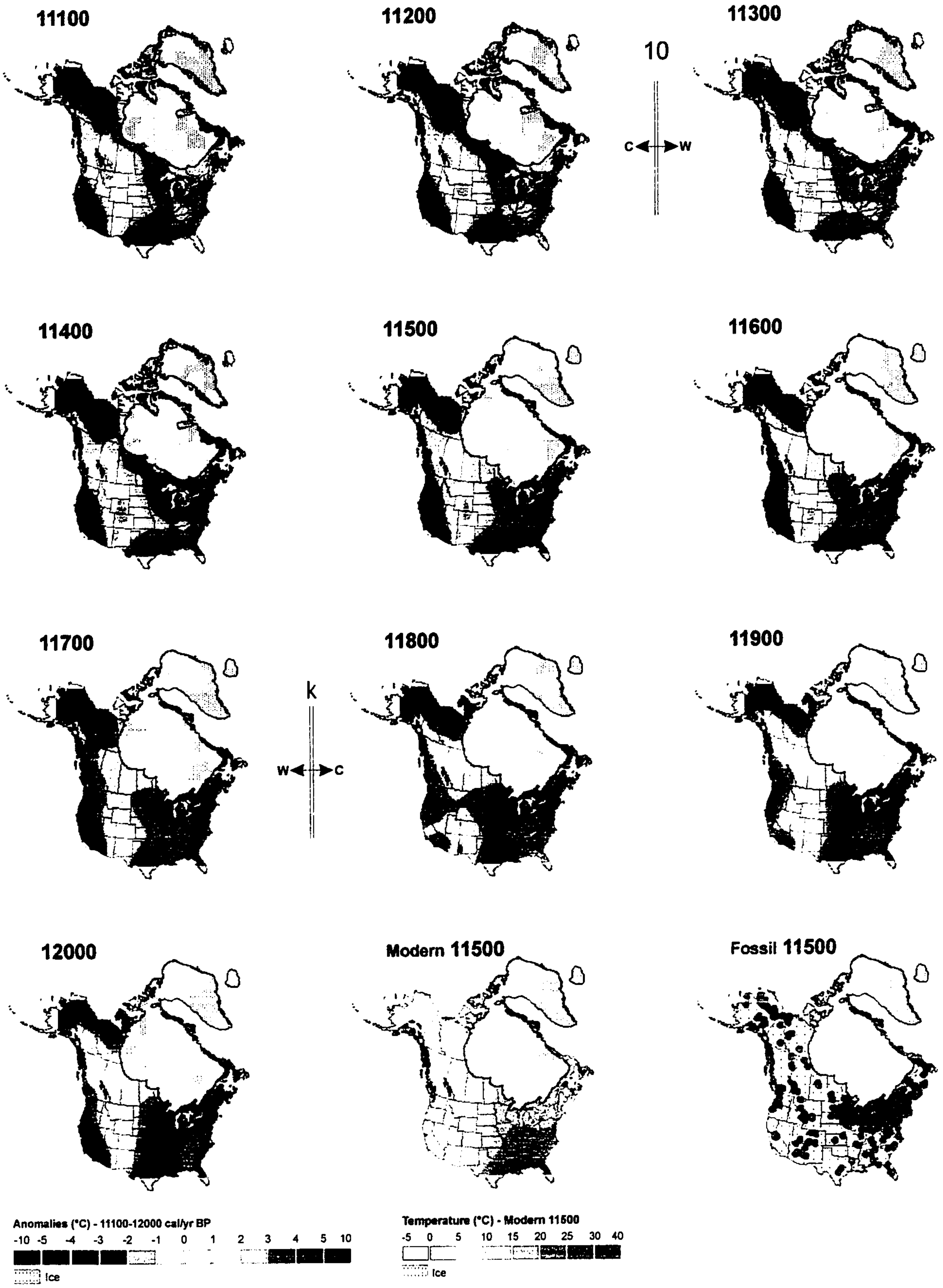


Figure 4.9 I: Temperature anomaly maps for North America for 11100-12000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 11500 cal/yr BP.

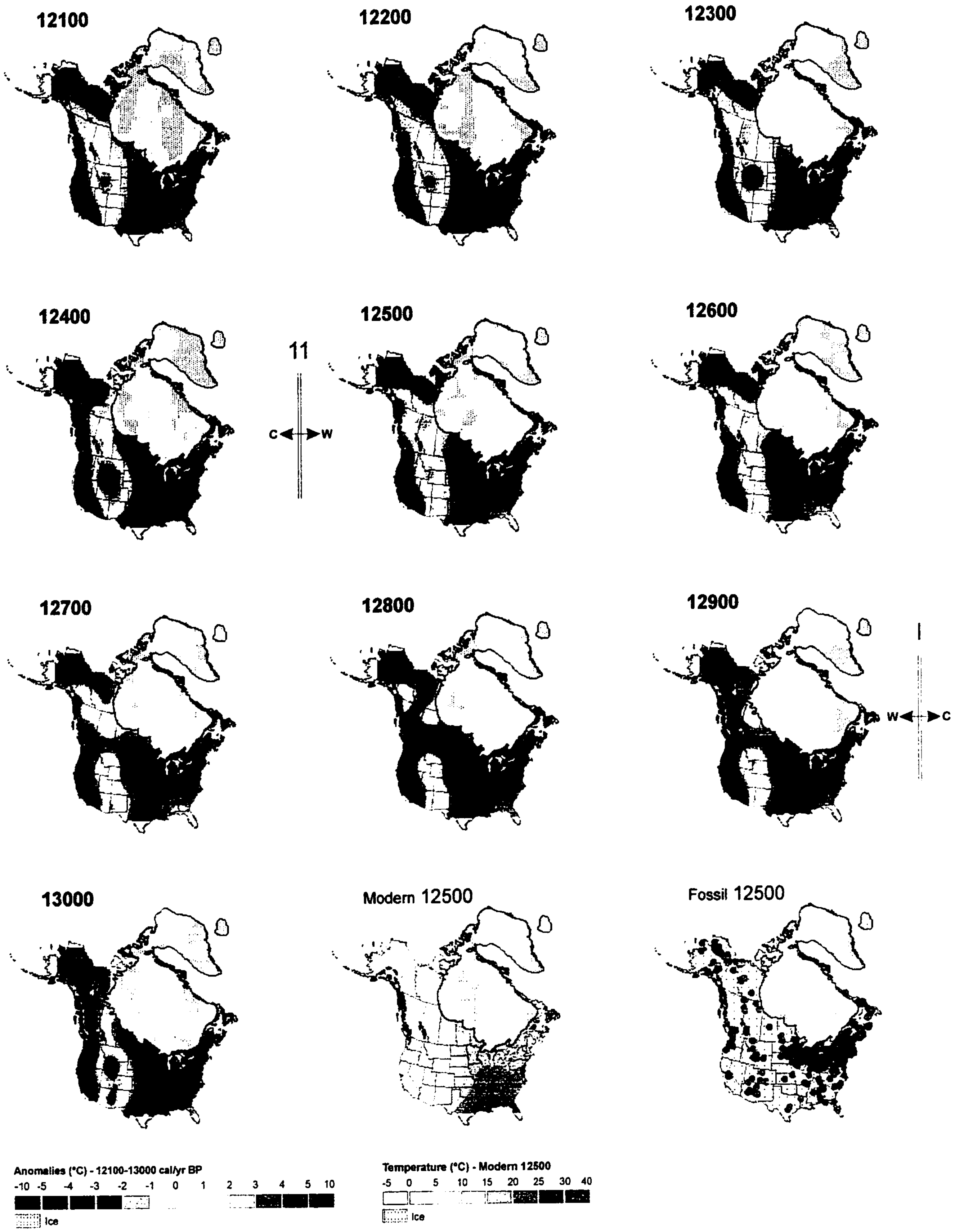


Figure 4.9 m: Temperature anomaly maps for North America for 12100-13000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 12500 cal/yr BP.

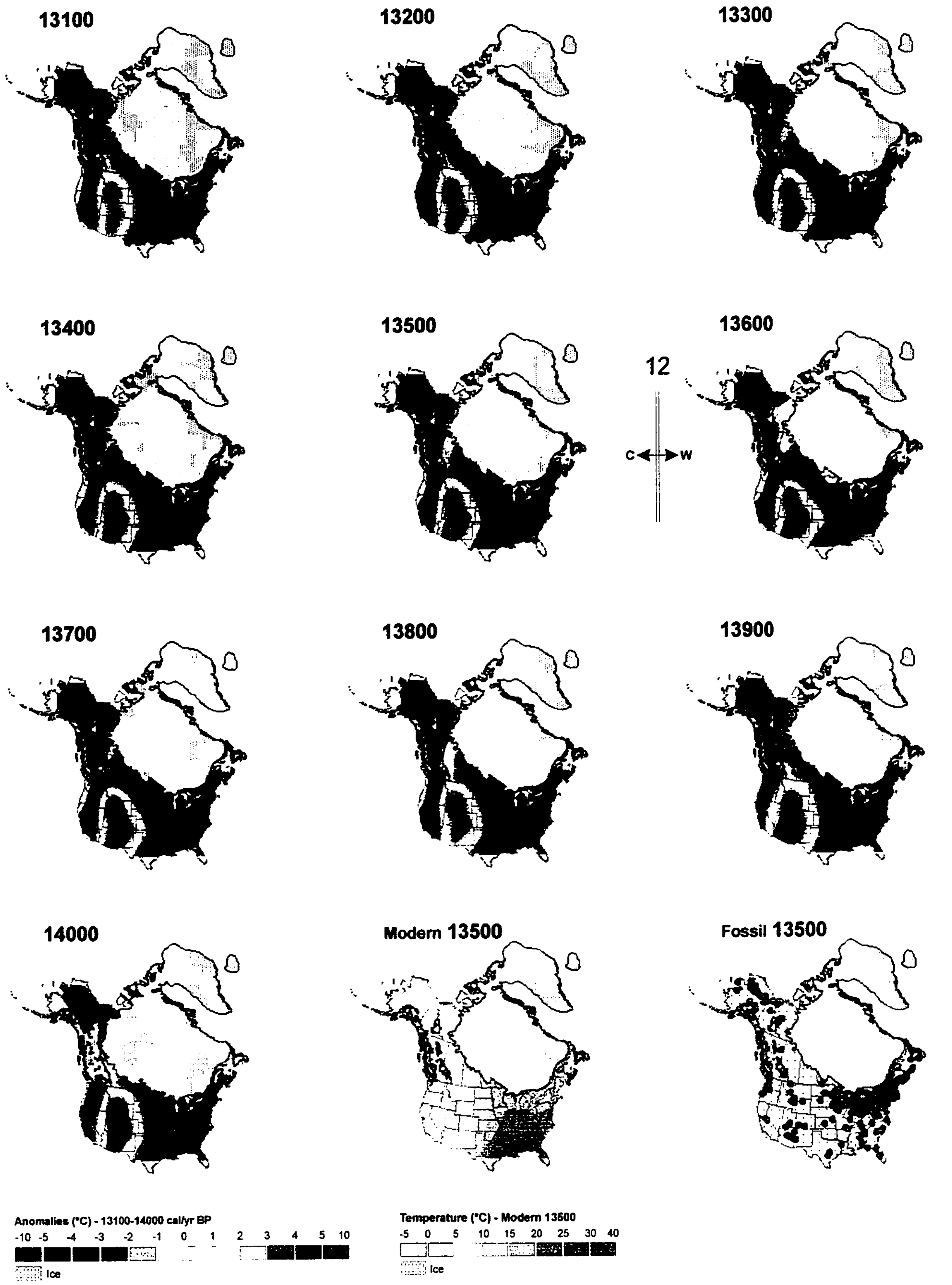


Figure 4.9 n: Temperature anomaly maps for North America for 13100-14000 cal/yr BP. Double vertical bars represent transitions where W=warming and C=cooling. Letters and numbers represent troughs and peaks identified in Figure 4.6. The last 2 figures show the modern-day climate that has been reconstructed using the site density at 13500 cal/yr BP.

Chapter 5

SUMMARY AND CONCLUSIONS

"Climate has influenced human activities. Two great developments emerged at about the time of major environmental change between 12 and 10 ka - the earliest appearance of agriculture in the Old World, and the cultural changes accompanying extinction of the Pleistocene mega fauna in the New World. Both developments may have been caused at least indirectly by the types of climatic change examined here. Now we may be faced with the reverse: human modification of climate and of related aspects of the physical environment. Application of a well-tested climate model is at present the only method for predicting these climatic and environmental changes and can help in planning a response to them"

COHMAP Members, 1988.

5.1 SUMMARY

The purpose of this dissertation was to improve our understanding of millennial-scale climate variability on the continental-scale of North America. The first chapter summarized the evidence pertaining to millennial-scale climate variability from the literature. This revealed that the evidence for an hypothesized global natural millennial-scale climate variability was not as compelling as implied and raised several unexplored and unresolved issues. Among these issues were questions pertaining to the timing, spatial extent, and magnitude of these millennial-scale climate variations. Archives of climate variability preserved in pollen time series offer the opportunity to investigate large-scale response to climate variation on several time scales. Using a dense network of pollen time series, in Chapters 2, 3, and 4, we explore millennial-scale climate variability on the continental-scale of North America during the past 14,000 years using different approaches in order to achieve robust results, and thus contribute to our current understanding of these natural climate fluctuations.

Paleoclimate information can be extracted from many environments. This is important because it increases our capability to understand not only different physical components of our climate system but also broad-scale patterns of our planet's climate history. Paleoclimate model verification requires large-scale comparisons to fossil data to determine possible physical mechanisms that govern our climate system response to perturbations and improve our predictive capabilities for potential climate changes in the future. One major uncertainty that arises in paleoclimate research is the lack of spatial corroborating evidence. For example, the Greenland ice core records offer a rich temporal record of past climate variations on several time-scales, however it remains a single point in space until it can be corroborated with records found in other regions. This lack of the spatial dimension has plagued large-scale paleoclimate research for many generations. The last few decades have seen assembled, through tedious and dedicated effort, large databases of paleoclimate information. These have been utilized for paleoclimate model verification at the astronomical-scale (Milankovitch scale). Due to increasing computing power, it is now possible to synthesize the voluminous and rich information that these paleoclimate archives contain at different spatial and temporal scales.

As seen in Chapter 1, the literature offers convincing evidence for millennial-scale climate fluctuations during the past million years. Although terrestrial paleoclimatologists have documented millennial-scale climate variability for more than a century, new results from high-resolution records have caused a renewed focus on century to millennial-scale climate variability. However, the lack of large-scale terrestrial syntheses at these scales has left a biased record, where most of the evidence comes from ice and ocean cores extracted from the North Atlantic region. This has lead to two divergent schools of thought. First, millennial-scale climate variability is a regional phenomena confined to the North Atlantic and adjacent landmasses. The alternative hypothesis is that millennial-scale climate variability is a global phenomenon, however the evidence has only been found in few records scattered around the world. Both hypotheses can be supported by the limited data available. Perhaps the North Atlantic is more sensitive to millennial-scale climate variability but a global response of the climate system is also a viable explanation given our current understanding of climate dynamics. It is only by exploring the timing, spatial extent, and magnitude of change that we can bring insight to this debate.

A rigorous and objective analysis of all proxy-climate records is needed to distinguish among these hypotheses. This dissertation makes a first step towards understanding the hemispheric response to millennial-scale climate variability by exploring the spatial and temporal extent, and the magnitude of climate changes on a continental scale, which is the minimum required from a paleoclimatological perspective. If North American climate variations were found to be synchronous to North Atlantic climate variability, this would provide support for the suggestion of hemispheric ocean-atmospheric reorganizations in response to some internal or external forcing. However, this would not preclude the idea that the North Atlantic region is more sensitive to these perturbations. It may in time turn out that millennial-scale climate changes are a global phenomenon where certain key regions are more sensitive than others. This is not as ambiguous as it seems, where global warming or El Nino phenomena have been found to have global impacts, although the magnitude and direction of change varies spatially.

Chapter 2 uses radiocarbon dates on pollen time series extracted from the North American Pollen Database (NAPD). It is solely based on an established assumption that palynologists tend to date transitions rather than randomly selected depths partly because of the cost associated with radiocarbon dating. This has the effect of minimizing errors that could incur by the analysis of the complete database, such as age-depth model chronologies and different taxonomic and sampling resolution. The only errors remaining are the method of dating that was utilized (i.e. AMS versus more conventional methods and material dated) and so-called characteristic ^{14}C plateaus. The presumption that preferentially dated transitions should yield common modal times of change was tested on the continental-scale of North America. Using an improved mixture modeling methodology, it was found that common transitions were not confined to the North Atlantic region but widespread throughout most regions of North America, and that these correlate reasonably well in timing with the North Atlantic (i.e. GISP2 and marine) and other terrestrial records (i.e. botanical, archeological and continental ice records).

Chapter 3 further investigates these transitions using a different methodological approach. Given that the mixture model analysis suggested widespread synchronous transitions in the pollen data, it would be expected that these times of change would be evident in the actual pollen time series. Conversely, errors arising from different taxonomic resolution, problems in deriving age-depth chronologies, and sampling resolution could accumulate and mask the signal of coherent changes in these pollen records. A Principal Component Analysis (PCA) in the time, space, and taxonomic dimensions was implemented to search for commonalities between the different regions of North America. Millennial-scale transitions were widespread and synchronous throughout North America at the vegetation community-level, and these closely correlated with the identified modes of transitions found using the mixture modeling approach of the previous analysis and the other climate-proxy records. Although the majority of transitions identified in this analysis closely correlated with the previous analyses, certain mismatches were noted to be preferentially distributed around certain time periods. Following a close examination of the different proxy-climate records, mismatches that were found related to scale interaction with lower-frequency forcings. Around 9000, 6000, and 4000 years BP, identified times of millennial-scale transitions were less coherent between the various records. These periods correspond with the abrupt early Holocene warming at 9000 yr BP, mid-Holocene thermal maximum around 6000 yr BP, and northern hemispheric neoglacial cooling approximately 4000 yr BP. In addition, analysis of the terrestrial, marine, and ice core records revealed more transitions than previously implied, with a mean pacing of 1,200 years between events. This raised a question regarding the true nature of the hypothesized quasi-periodic 1500-yr cycle. This study also raised the possibility that sampling aliasing may have caused the mismatch between records based on sample resolution.

Chapter 4 is a first attempt to quantify the magnitude of these millennial-scale climate changes at the continental-scale of North America during the past 14,000 years. This was accomplished using yet another different investigative tool, the method of modern analog (MAT). A newly updated modern pollen dataset (4590 sites) and over 700 fossil pollen sites were used to reconstruct a mean-continental July temperature anomaly curve for the past 14,000 years at 100-yr intervals. The use of this large dataset enabled the reconstruction of climate variations at several timescales. In addition, a quasi-periodic variation with a period of approximately 1150-yr can explain the mismatches in the different climate-proxy records. Potential causes for these millennial-scale temperature excursions were also explored through correlation with cosmogenic records, a proxy for solar variability, which has been proposed as a potential forcing mechanism for century to millennial-scale climate variability. The temperature change associated with millennial-scale variations was shown to be more variable during the late glacial and to a lesser extent in early Holocene. The middle-Holocene was the least variable period. Analysis of the spatial patterns of climate change revealed that temperature variations at this scale were not unidirectional. This supports the hypothesis that these climate changes caused large-scale ocean-atmospheric circulation changes. Maximum continental-scale temperature changes were found to be on the order of approximately 0.4°C during the last 10,000 years, which is comparable to reconstructions of high-resolution northern hemispheric temperature variations of the past 1,000 years. This has serious implications for our current understanding of global warming, estimated as a rise in global temperatures of 0.4 to 0.6°C in the past 150 years. These results support the hypothesis that modern global warming cannot be explained simply by natural climate variability over the past 10,000 years but that human-induced atmospheric CO_2 has and will continue to play a major role in our climate evolution in the near future.

5.2 CONCLUSIONS

The fundamental hypotheses of this research dissertation are:

H_1: NORTH AMERICAN TERRESTRIAL FOSSIL RECORDS SHOW MILLENNIAL-SCALE CLIMATE VARIATIONS AT THE PACE OF 1,500 ± 500 YEARS CONSISTENT WITH THE ICE AND MARINE RECORDS DURING THE LATE GLACIAL AND HOLOCENE.

Conclusions of this dissertation:

a) Millennial-scale climate variability is evident in the fossil pollen records of North America during the past 14,000 years.

b) The marine, ice and other terrestrial records show consistent times of change in most, but not all cases.

c) Mismatches occur at certain times when larger-scale interactions are operating, such as during the abrupt early Holocene warming at 9,000-yr BP, the mid-Holocene thermal maximum around 6,000-yr BP and the Northern Hemispheric neoglacial cooling approximately 4,000 years ago.

d) Likely reasons for these mismatches include sampling aliasing and/or large-scale interactions attenuating the different records at these preferred times, making millennial-scale climate variations more difficult to identify.

e) An 1150 ± 125-yr quasi-periodic component is more consistent with the new reconstruction.

Based on these conclusions, the first hypothesis is verified but with reservations as to the pacing. An 1150 ± 125-yr quasi-periodic component does fall within the 1500-yr ± 500 ranges, however this dissertation has further refined this estimate.

H_2: MILLENNIAL-SCALE CLIMATE CHANGES WERE RESTRICTED TO NORTHEASTERN NORTH AMERICA DURING THE LATE GLACIAL AND HOLOCENE.

Conclusions of this dissertation:

f) Widespread synchronous millennial-scale climate variability is found throughout North America during the past 14,000 years based on three independent methods.

g) Modes of preferentially dated vegetation transitions in the fossil pollen records are found throughout North America.

h) Community-level changes in vegetation are revealed at synchronous times in different regions of North America during the past 14,000 years.

i) Temperature variations are widespread but not unidirectional throughout North America

j) The magnitude of temperature variability is not uniformly distributed throughout North America

Based on these conclusions, this hypothesis is rejected and the alternative hypothesis is accepted. Millennial-scale climate variability was widespread throughout North America although the North Atlantic region may be more sensitive to temperature variability. Thus, the climate variability at this scale is not just due to an isolated North Atlantic phenomenon, but to large-scale changes in the global atmosphere-ocean circulation.

H_3: TEMPERATURE VARIABILITY WITHIN THESE MILLENNIAL-SCALE CLIMATE VARIATIONS WAS GREATER DURING THE TRANSITIONAL PERIOD OF THE LAST DEGLACIATION, AND SUBTLER DURING THE HOLOCENE IN NORTH AMERICA.

Conclusions of this dissertation:

k) Temperature variability is greater during the late glacial and early Holocene and more subdued during the middle Holocene.

l) Mean-continental July temperature variability is on the order of 0.2 to 0.4°C during the Holocene and 0.4 to 0.6 °C during the late glacial

m) Vegetation and climate changes and transition timing are more readily identifiable and consistent between various proxy-climate records during the late glacial than during the Holocene

Based on these conclusions, this hypothesis cannot be rejected.

These results have enabled conclusions to be made about causes in the climate system dynamics during the late glacial and Holocene.

n) Correlation between marine, ice, and terrestrial records and cosmogenic nuclide records (i.e. ^{10}Be and $^{14}C_{atm}$) support variability in solar output as a fundamental cause of climate variations in frequency bands of approximately 300-500 years and 900 to 1100 yr.

o) Internal mechanisms, including North Atlantic Ocean overturning (THC), surface-albedo feedback mechanisms, and ocean-atmospheric interactions are most likely required to explain the asymmetry and abruptness seen in the records during the Holocene. Cryospheric feedbacks are likely to amplify climate variability at these scales during the late glacial and into the last glacial period.

p) Recent global warming exceeds natural century to millennial-scale continental to hemispheric temperature variations of the past 10,000 years and therefore human-induced CO_2 contribution is likely to be an important contributor to modern warming.

This dissertation supports the contention that century to millennial-scale northern hemispheric climate variability was widespread and synchronous during the past 14,000 years due to large-scale ocean-atmospheric circulation changes. The quasi-periodic nature of the paleoclimate record resulted from a weak periodic solar forcing in frequency bands of 300-500 and 900-1100 years and amplified through powerful internal feedback mechanisms, as witnessed by the abruptness and asymmetry of the paleoclimatic record. Recent warming cannot be attributed solely to century to millennial-scale climate variability of the past 10,000 years; therefore atmospheric CO_2 concentration has and will continue to play a major role in climate changes in the future.

The Earth's climate system is complex where there are many sources of instabilities. This poses a major problem in identifying internal mechanisms responsible for natural century to millennial-scale climate variability (for more information, see Appendix H). It is imperative that research continues to better understand the relationships between the various components of the climate system. This will increase our understanding of the climate sytem as human influence have undoubtedly become a major player in our climate system.

5.3 FUTURE RESEARCH IN CLIMATE CHANGE

Our understanding of century to millennial-scale climate variability is still limited. The first problem is that the spatial extent of these climate oscillations is not well established. Although some attempts have been made to correlate interhemispheric synchroneity (Lowell *et al.*, 1995, Bluiner *et al.*, 1998; Steig *et al.*, 1998), poor geographical coverage and lack of high-resolution data have plagued this analysis. A second problem arises from proxy dating and proxy-climate reconstructions. It has been possible to correlate abrupt climate events between terrestrial and ocean records (Lowell *et al.*, 1995; Kukla *et al.*, 1997; Viau *et al.*, 2002). However, is this merely an exercise in "wiggle matching"? Errors associated with current dating methods are greater than the resolution of the transitions we are attempting to correlate. Another problem stems from age-depth models in sediment cores, which are typically linearly interpolated between dates. Age-depth models are frequently based on previously derived models, such as the SPECMAP curve, which is orbitally tuned. This makes the determination of leads and lags between the climate system components impossible.

As I have suggested above, we are still limited in explaining century to millennial-scale climate variability, although the results presented here have shown that the hemispheric extent of these must be part of the explanation. This makes prediction of future climate transitions difficult. Models (GCMs) have not yet been successful in simulating global climate transitions that have been identified in proxy-climate data. Models are needed to improve our understanding of the underlying relationships between the different components of the climate system. For example, it appears possible to simulate Dansgaard-Oeschger (DO) events by small increases in meltwater fluxes into the North Atlantic causing diminished NADW formation (Sakai and Peltier, 1996). Models can be regarded as tools to help understand responses and processes of the climate system to perturbations, but more work is needed to improve them.

Our understanding of climate change has evolved substantially in the past decades. We have seen evidence that the climate can "flip" from one mode to another in a few decades and lasting centuries to millennia in both glacial and interglacial regimes. However, our climate system has internal instabilities resulting from inherent nonlinearities between the different components of the climate system. These non-linear relationships are key elements in these rather sudden climate transitions. An increased geographical coverage and detailed high-resolution paleoclimatic record, at least for the Quaternary period, is necessary to establish spatial extent and synchroniety of these sub-orbital events and to determine the causative factors. This can lead us to appreciate how these factors play a role in climate today and how they may play in the future. Furthermore, unless the natural variability of climate is fully understood, it will be extremely difficult to recognize any human-induced effects on climate that may become apparent in the future.

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: 154-56.

Alley, R.B., Mayewski, P.A., Sowers, T., Stuiver, M., Taylor, K.C., and Clark, P.U., 1997, Holocene climatic instability: A prominent widespread event 8200 yr ago: Geology, 25: 483-86.

Alley, R.B., Meese, D.A., Shuman, C.A., Gow, A.J., Taylor, K.C., Grootes, P.M., White, J.W.C., Ram, M., Waddington, E.D., Mayewski, P.A., and Zielinski, G.A., 1993, Abrupt increase in Greenland snow accumulation at the end of the Younger Dryas event: Nature, 362: 527-528.

Anderson, D.E. 1998: A reconstruction of Holocene climatic changes from peat bogs in north-west Scotland. Boreas 27: 208-224.

Anderson, D.E. 1997: Younger Dryas research and its implications for understanding abrupt climatic change. Progress in Physical Geography 21(2): 230-249.

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

Andrews, J.T. 1998: Abrupt changes (Heinrich events) in late Quaternary North Atlantic marine environments. Journal of Quaternary Science 13: 3-16.

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

Bartlein, P.J., Webb, T, III, and Fleri, E. 1984. Holocene climate changes in the northern Midwest: pollen derived estimates. Quaternary Research 22: 361-374.

Behl, R.J. and Kennett, J.P. 1996: Brief interstadial events in the Santa Barbara basin, NE Pacific, during the past 60kyr. Nature 379: 243-46.

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

Berger, W., Bickert, T., Jansen, E., Wefer, G. 1993: The central mystery of the Quaternary ice age. Oceanus 36: 53-56.

Bernabo, J.C. 1981. Quantitative estimates of temperature over the last 2700 years in Michigan based on pollen data. Quaternary Research 15: 143-159.

Bianchi, G.G., and McCave, I.N., 1999: Holocene periodicity in North Atlantic climate and deep-ocean flow south of Iceland. Nature 397: 515-517

Birks, H.J.B., and Birks, H.H., 1980: Quaternary paleoecology. University of Cambridge. 288 pp.

Blunier, T., Chappellaz, J., Scwander, J., Dallenback, A., Stauffer, B., Stocker, T.F., Raynaud, D., Jouzel, J., Clausen, H..B., Hammer, C.U. and Johnsen, S.J. 1998: Asynchrony of Antarctic and Greenland climate change during the last glacial period. Nature 394: 739-43.

Bond, G., Heinrich, H., Broecher, W., Labeyrie, L., McManus, J., Andrews, J., Huon, S., Jantschik, R., Clasen, S., Simet, C., Tedesco, K., Klas, M., Bonani, G. and Ivy, S. 1992: Evidence for massive discharges of icebergs into the North Atlantic Ocean during the last glacial period. Nature 360: 245-49.

Bond, G., Showers, W., Cheseby, M., Lotti, R., Almasi, P., deMenocal, P., Priore, P., Cullen, H., Hajdas, I. and Bonani, G. 1997: A pervasive millinnial-scale cycle in North Atlantic Holocene and glacial climates. Science 278: 1257-65.

Bond, G., Kromer, B., Beer, J., Muscheler, R., Evans, N.E., Showers, W., Hoffman, S., Lotti-Bond, R., Hajdas, I. and Bonani, G. 2001, Persistent solar influence on North Atlantic climate during the Holocene: Science, 294: 2130-2136.

Bond, G.C. and Lotti, R. 1995: Iceberg discharges into the North Atlantic on millennial time scales during the last deglaciation. Science 267: 1005-10.

Bond, G.C., Broecker, W., Johnsen, S., McManus, J., Labeyrie, L., Jouzel, J. and Bonani, G. 1993: Correlations between climate records from North Atlantic sediments and Greenland ice. Nature 365: 143-47.

Broecker, W.S., Bond. G., Klas, M., Clark, E. and McManus, J. 1992: Origin of the North Atlantic's Heinrich events. Climate Dynamics 6: 265-73.

Broecker, W.S., and Denton, G.H. 1989: The role of ocean-atmosphere reorganizations in glacial cycles. Geochimica et Cosmochimica Acta 53: 2465-2501.

Broecker, W.S., Kennett, J.P., Flower, B.P., Teller, J.T. Trumbore, S., Bonani, G., Wolfli, W. 1989: Routing of meltwater from the Laurentide Ice Sheet during the Younger Dryas cold episode. Nature 341: 318-321.

Broecker, W.S. 1995: The glacial world according to Wally. New York: Eldigio Press.

Broecker, W.S. 1997: Thermohaline circulation, the Achilles heel of our climate system: will manmade CO2 upset the current balance? Science 278: 1582-88.

Bryant, V.M., and Holloway, R.G., editors, 1985, Pollen records of Late-Quaternary North American sediment, Dallas,Texas, American Association of stratigraphic palynologists foundation, Library of Congress Catalog Number, 85-71610, 426 p.

Bryson, R.A. 1966. Air masses, streamlines, and the Boreal Forest. Geographical Bulletin 8(3): 228-269.

Bryson, R.A., and Wendland, W.M. 1967. Tentative climatic patterns for some late-Glacial and Post-Glacial episodes in central North America. In: Life, Land and Water, W.J. Mayer-Oakes, Ed., pp. 271-298, University of Manitoba Press, Winnipeg, Canada.

COHMAP Members, 1988. Climatic changes of the last 18,000 years: Observations and model simulations. Science 241: 1043-1052.

Crowley, 2000. Causes of climate change over the past 1,000 years, Science 289: 270-277.

Cwynar, L.C., Levesque, A.J., Mayle, F.E., and Walker, I. 1994: Wisconsinan late-glacial environmental change in New Brunswick: a regional synthesis. Journal of Quaternary Science 9: 161-64.

Dansgaard, W., Johnson, S.J., Clausen, H.B., Dahl-Jensen, D., Gundenstrup, N.S., Hammer, C.U., Hvidberg, C.S., Steffensen, J.P. and Sveinbjornsdottier, A.E. 1993: Evidence for general instability of past climate from a 250-kyr ice-core record. Nature 364: 218-20.

Dansgaard, W., White, J.W.C. and Johnsen, S.J. 1989: The abrupt termination of the Younger Dryas climate event. Nature 339: 532-34.

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

Delcourt, H.R., and Delcourt, P.A. 1985. Quaternary Palynology and Vegetational History of the Southeastern United States. In: Pollen records of Late-Quaternary North American sediments. Bryant, V.M. Jr, Holloway, R.G. (eds.) Austin, Texas: American Association of Stratigraphic Palynologists Foundation.

Delcourt, P.A., Delcourt, H.R., Webb III, T., 1984. Atlas of mapped distributions of dominance and modem pollen percentages for important tree taxa of eastern North America. AASP Contribution Series 14. Dallas, Texas: American Association of Stratigraphic Palynologists Foundation.

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

Dorale, J.A., Gonzalez, L.A., Reagan, M.K., Pickett, D.A., Murrell, M.T. and Baker, R.G. 1992: A high-resolution record of Holocene climate change in speleothem calcite from Cold Water Cave, northeast Iowa. Science 258: 1626-30.

Dyke. A.S., and Prest, V.K., 1987. Late Wisconsinan and Holocene history of the Laurentide ice sheet. Géographie Physique e Quaternaire, 41: 237-264.

Esper, J., Cook, E.R. and Schweingruber, F.H. 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. Science 295: 2250-2253.

Everitt, B.S. and D.J. Hand (1981). Finite Mixture Distributions. Chapman and Hall, London.143 pp

Fisher, D.A. 1983. Carbon-14 production compared to oxygen isotope records from Camp Century, Greenland and Devon Island, Canada. Climatic Change 4: 419-426.

Fisher, D.A., 1982, Carbon-14 production compared to oxygen isotope records from Camp Century, Greenland, and Devon Island, Canada: Climate Change, 4: 419-426.

Flint, R.F., 1971, Glacial and Quaternary Geology: New York, Wiley, 892 p.

Fredskild, B., 1973: Studies in vegetational history of Greenland. Meddelser om Gronland, 198(4): 4-245.

Gajewski, K. 1993. The role of paleoecology in the study of global climatic change. Review of Paleobotany and Palynology 79: 228-236.

Gajewski, K. 1998. Climate changes of the past millennium reconstructed from high resolution pollen records. In: MacIver, D.C., and Meyer, R.E. (eds.) Proceedings of the Workshop on Decoding Canada's Past: Climate variations and biodiversity change during the last millennium, Atmospheric Environments Services, Canada, pp-69-85.

Gajewski, K., 1987, Climatic impacts on the vegetation of eastern North America for the past 2,000 years: Vegetatio, 68: 179-190.

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., Sherin, 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.

Gasse, F., and Van Campo, E. 1994: Abrupt post-glacial climate in West Asia and North Africa monsoon domains. Earth and Planetary Science Letters 126: 435-456.

Green, D., 1983. The ecological interpretation of fine resolution pollen records. New Phytologist 94: 459-477.

Grimm, E.C., Jacobson, G.L., Watts, W.A., Hansen, B.C.S. and Maasch, K. 1993: A 50 000 year record of climate oscillations from Florida and its temporal correlation with the Heinrich events. Science 261: 198-200.

Grimm, E.C., 2000, ed., North American Pollen Database and Global Pollen Database, Data archived at IGBP PAGES/World Data Center-A for Paleoclimatology: Boulder, Colorado, National Oceanic and Atmospheric Administration/National Geophysical Data Center Paleoclimatology Program, http://www.ngdc.noaa.gov/paleo/, July 1999.

Grimm, E.C., and Jacobson, G.L Jr. 1991. Fossil-pollen evidence for abrupt climate changes during the past 18,000 years in eastern North America. Climate Dynamics 6: 179-184.

Grove, J.M., 1979, The glacial history of the Holocene: Progress in Physical Geography,3: 1-54.

Guiot, J.L., de Beaulieu, J.L., Chedaddi, R., David, F., Ponel, P. and Reille, M.. 1993: The climate in western Europe during the last glacial/interglacial cycle derived from pollen and inscet remains. Palaeogeography, Palaeoclimatology, Palaeoecology 103: 73-93.

Haigh, J.D., 1996: The impact of solar variability on climate. Science, 272: 981-984.

Harvey, L.D.D., 1979, Solar variability as a contributing factor to Holocene climate change: Progress in Physical Geography, 3: 487-530.

Heinrich, H. 1988: Origin and consequences of cyclic ice rafting in the northeast Atlantic Ocean during the past 130 000 years. Quaternary Research 29: 142-52.

Heusser, L. 1998: Direct correlation of millennial-scale changes in western North American vegetation and climate with changes in the California Current system over the past 60 kyr. Paleoceanography 13: 252-262.

Huntley, B. 1992. Rates of change in the European palynological record of the last 13,000 years and their climatic interpretation. Climate Dymanics 6: 185-191.

IPCC Climate Change: 1995: The Science of Climate Change. Houghton, J.T, Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A., and Maskell, K. (Eds.). Cambridge University Press. 572 pp.

IPCC Climate Change: 2001: The Scientific Base. Houghton, J.T, Ding, Y.,Griggs, D.J., Noguer, M., van Der Linden, P.J., Dai, X., Maskell, K., and Johnson, C.A. (Eds.). Cambridge University Press. 572 pp.

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, 253-386, Cambridge Unvivesity Press.

Jacobson, G.L., Webb, T.III., Grimm. E.C. 1986: Patterns and rates of vegetation change during the deglaciation of eastern North America. In the Geology of North America, Geological Society of America K-3, 277-288.

Johnsen, S.J., Clausen, H.B., Dansgaard, W., Gundenstrup, N.S., Hammer, C.U. and Tauber, H. 1992: The Eem stable isotope record along the GRIP ice core and its interpretation. Quaternary Research 43: 117-24.

Johnson, D., 1966, Non-linear parameter estimation for the partial collective model of airmass analysis: Department of Meterorology, Madison, University of Wisconsin, Technical Report 24, Nonr 1202(07): 63-73.

Jones, P.D., New, M., Parker, D.E., Martin, S. and Rigor, I.G. 1999. Surface air temperature and its changes over the past 150 years. Reviews of Geophysics 37: 173-199.

Joussaume, S. And Guiot, J. 1999: Reconstruire les chauds et froids de l=Europe. La Recherche 321: 54-59.

Kerr, R.A. 1999: Glacial periods don't keep time. Nature 396: 274-275.

Kukla, G., McManus, J., Roussear, D.-D. and Chuine, I. 1997: How long and how stable was the last interglacial? Quaternary Science Reviews 16: 605-12.

Kutzbach, J.E., 1981, Monsoon climate of the early Holocene: Climate experiment with the Earth's orbital parameters for 9000 years ago: Science 214: 59-61.

Labeyrie, L., and Jouzel, J. 1999: Les soubresauts millenaires du climat. La Recherche 321: 60-62.

Lean, J., Skumanich, A., and White, O. 1992. Estimating the sun's radiative output during the Maunder minimum. Geophysical Research Letters, 19(15): 1591-1594.

Legates, D. 1991. The effect of domain shape on Principal Component Analysis, International Journal of Climatology 11: 135-146.

Lehman, S.J. and Keigwin, L.D. 1992: Sudden changes in North Atlantic circulation during the last deglaciation. Nature 356: 757-762.

Lowell, T.V., Heusser, C.J., Andersen, B.G., Moreno, P.I., Hauser, A., Heusser, L.E., Schluchter, C., Marchant, D.R. and Denton, G.H. 1995: Interhemispheric correlation of late Pleisocene glacial events. Science 269: 1541-49.

MacManus, J.F., Oppo, D.W., and Cullen, J.L., 1999: A 0.5 million-year record of millennial-scale climate variability in the North Atlantic. Science 283: 971-974.

Manabe, S. And Stouffer, R.J. 1995: Simulation of abrupt climate change induced by freshwater input to the North Atlantic. Nature 378: 165-167.

Mann, M.E., Bradley, R.S. and Hughes, M.K. 1999. Northern Hemisphere temperatures duing the past millennium: Inferences, uncertainties, and limitations. Geophysical Research Letters 26: 759-762.

Mayewski, P.A., Meeker, L.D., Twickler, M.S., Whitlow, S., Yang, Q., Lyons, W.B. and Prentice, M. 1997: Major features and forcing of high-latitude Northern Hemisphere atmospheric circulation using a 110 000 year-long glaciochemical series. Journal of Geophysical Research 102: 26 345-65.

Nichols, H., 1967, Pollen diagrams from sub-arctic central Canada: Science 155: 1665-1668.

O'Brien, S.R., Mayewski, A., Meeker, L.D., Meese, D.A., Twickler, M.S. and Whitlow, S.I. 1995: Complexity of Holocene climate as reconstructed from a Greenland ice core. Science 270: 1962-64.

Oppo, D.W., McManus, J.E., and Cullen, J.L., 1998, Abrupt climate events 500,000 to 340,000 years ago: Evidence from subpolar North Atlantic sediments: Science 279: 1335-1338.

Overpeck, J.T., Webb III, T., Prentice, I.C. 1985. Quantiative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. Quaternary Research 23: 87-1008.

Overpeck, J.T. 1987. Pollen time series and Holocene climate variability of the Midwest United States. In: W.H Berger and L.D. Labeyrie (eds.), Abrupt Climatic Change: 137-143.

Overpeck, J.T. 1986. Time series analysis of Holocene pollen data: Paleoclimatological and Paleoecological Applications. Ph.D. Thesis, Brown University.

Peixoto, J. P., and Oort, A.H. 1992: Physics of Climate. American Institute of Physics. New York, 520 pp.

Perry, C.A., and Hsu, K.J. 2000. Geophysical, archeological, and historical evidence support a solar-output model for climate change. Proceeding to the National Academy of Science (PNAS), 97:12433-12438.

Pestiaux, P. Van Der Mersh, I., Berger, A., and Duplessy, J.C., 1988: Paleoclimatic variability at frequencies ranging from 1 cycle per 10000 uears to 1 cycle per 1000 years: evidence for nonlinear behavior of the climate system. Climatic Change 12: 9-37.

Prentice. I.C. 1986. Vegetation responses to past climatic variation. Vegetatio 67: 131-141.

Rasmussen, T.L., van Weering, T.C.E. and Labeyrie, L. 1997: Climatic instability, ice sheets and ocean dynamics at high northern latitudes during the last glacial period (58-10 KA BP). Quaternary Science Reviews 16: 73-80.

Raymo, M.E., Ganley, K., Carter, S., Oppo, D.W., and MacManus, J., 1998: Millennial-scale climate instability during the early Pleistocene epoch. Nature 392: 699-702.

Rind, D., 1999: Complexity and Climate. Science 284: 105-107.

Ripley, B.D. 1976. The second-order analysis of stationary point processes. Journal of Applied Probability, 13:255-266.

Rowlingson, B. and Diggle, P., 1993, Splancs: spatial point pattern analysis code in S-Plus: Computers and Geosciences, 19: 627-655.

Sakai, K. And Peltier, W.R. 1996: Dansgaard-Oeschger ocsillations in a coupled atmosphere-ocean climate model. Americam Meteorological Society 10: 949-970.

Sawada, M.C. 2001. Late Quaternary Paleoclimates and Biogeography of North America. Unpublished Ph.D. thesis, University of Ottawa, Ottawa, Ontario, Canada.

Schultz, H., von Rad, U. and Erlenkeuser, H. 1998: Correlation between Arabian Sea and Greenland climate oscillations of the past 110 000 years. Nature 393: 54-57.

Sears, P.B. 1964. The goals of paleoclimatological reconstructions. In: The reconstruction of past environments, J.H. Hester and J. Shoenwetter (eds.) 4-5, Fort Burgwin Research Center Publication No. 3, Fort Burgwin, New Mexico, USA.

Shindell, D., D. Rind, N. Balachandran, J. Lean, and P. Lonergan, 1999: Solar cycle variability, ozone, and climate. Science, 284: 305-308.

Sirocko, F., Sarnthein, M., Erlenkeuser, H., Lange, H., Arnold, M. and Duplessy, J.-C. 1993: Century-scale events in monsoonal climate over the past 24 000 years. Nature 364: 322-24.

Steig, E.J., Brook, E.J., White, J.W.C., Sucher, C.M.,Bender, M.L., Lehman, S.J., Morse, L.D., Waddington, E.D., Clow, G.D. 1998: Synchronous climate changes in Antartica and the North Atlantic. Science 282: 92-95.

Steig, E.J., 1999, Mid-Holocene climate change: Science 286: 1485-1487.

Street-Perrott, F.A., Perrott R.A 1990: Abrupt climate fluctuations in the tropics: the influence of Atlantic ocean circulation. Nature 343: 607-612.

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

Stuiver, M., Braziunas, T.F., and Grootes, P.M., 1997: Is there evidence for solar forcing of climate in the GISP2 oxygen isotope record? Quaternary Research 48: 259-266.

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: 825-27.

Thompson, L.G., Mosley-Thompson, E., Davis, M.E., Bolzan, J.F., Dai, J., Yao, T., Gundenstrup, N., Wu, X., Lein, L. and Xie, Z. 1989: Holocene-late Pleistocene climatic ice core records from Qinghai - Tibetan Plateau. Science 246: 474-77.

Titterington, D.M., A.F.M. Smith and U.E. Makov (1985). Statistical Analysis of Finite Mixture Distributions, Wiley, New York. 243 pp.

Van Geel, B., Raspopov, O.M., Renssen, H., van der Plitch, J., Dergachev, V.A., and Meijer, H.A.J., 1999: The role of solar forcing upon climate change. Quaternary Science Reviews 18: 331-338.

Viau, A.E, Gajewski, K., Fines, P., Atkinson, D.E., and Sawada, M.C. 2002. Widespread evidence of 1500 yr climate variability in North America during the past 14 000 yr. Geology 30(5): 455-458.

Watts, W.A. 1973. Rates of change and stability in vegetation in the perspective of long periods of time. In:Birks, H.J.B., West, R.G. eds., Quaternary Plant Ecology. Oxford: Blackwell Scientific.

Watts, W.A., and M. Stuiver. 1980. Late Wisconsin climate of northern Florida and the origin of species-rich deciduous forest. Science 210:325-327.

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

Webb III, T, and Bryson, R.A. 1972. Late and Postglacial climatic change in the northern Midwest, USA. Quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis. Quat. Res. 2: 70-115.

Wendland, W.M., and Bryson, R.A., 1974, Dating climatic episodes of the Holocene: Quaternary Research: 4: 9-24.

Whitlock, C., Bartlein, P.J. 1997: Vegetation and climate change in northwestern America during the past 125 kyr. Nature 388: 57-61.

Winkler, M.G., Swain, A.M., and Kutzbach, J.E., 1986, Middle Holocene dry period in the northern Midwestern United States: Lake levels and pollen stratigraphy: Quaternary Research 25: 235-250.

Wright, H.E. Jr., Kutzbach, J.E., Webb, T.III., Ruddiman, W.F., Street-Perrott, F.A., Bartlein, P.J. 1993: Global climates since the last glacial maximum.. The University of Minnesota Press. 569 pp.

Zielinski, G.A., Mayewski, P.A., Meeker, L.D., Withlow, S., and Twickler, M.S., 1996: A 110,000 yr record of explosive volcanism from the GISP2 (Greenland) ice core. Climate Change 1996: 109-118.

APPENDIX A - VEGETATION-CLIMATE RELATIONSHIP

There are two schools of thought regarding vegetation-climate relationship. According to one hypothesis, "the dynamic vegetation response", migrational lags caused delays of thousands of years in the postglacial dynamics of forest compositions. The alternative hypothesis, "the complex past climatic changes", explains these changes as dynamic equilibrium responses to changes in seasonality and climatic pattern anomalies (Webb, 1986, Prentice, 1986).

Given that the time scale of interest for this research is millennial and the space scale is regional to continental, there are two major questions that pertain to whether vegetation was in dynamic equilibrium with climate at these scales during the Late Glacial and Holocene period. The first concerns how plant species migrated so quickly during the Holocene in response to deglaciation of the North American continent (Pitelka *et al.*, 1997; Clark *et al.*, 1998, Sawada, 2001). The second question is whether vegetation was in equilibrium with climate during the large-scale migrations of the Late Quaternary and Holocene (Webb, 1986, 1987).

How plant migrated so rapidly during the Late Quaternary is still not fully understood (Clark *et al.*, 1998). Simulations incorporating long-distance stochastic events indicate that local populations may establish at a great distance from the main population and then expanded rapidly under favorable conditions (Clark *et al.*, 1998). There is some data to support this (Ritchie and MacDonald, 1986, Bennett, 1985).

There have been many studies that have focused on migrations of tree species during the Late Quaternary and Holocene. Migration maps (isochrone maps) were first used by Davis (1976, 1981) to illustrate the changes in the range boundaries of a given taxa through time. By determining the presence/absence of a taxa above a threshold in the pollen record, migrations maps were constructed for the major tree species. This has led to the interpretation that trees were not in dynamic equilibrium with climate and that significant lags occurred between vegetation and climate. However, the presence or absence of a particular species was based on some 'rational limit' or rise in pollen percentages in the record (Davis *et al.* 1991). For example, if beech (*Fagus* spp.) pollen was found to be greater than 1.5%

then it would be considered present in local vegetation. This approach has a number of problems such as variations in the proportions due to the presence/absence of major pollen types, under-and over-representation of species, dispersal rates and production, or simply that the tree species may have been present in very low numbers, therefore not recorded in the pollen record (Ritchie and MacDonald 1986; Bennett, 1985, Gajewski *et al.*, 1993). Alternatives to migration maps are isopoll maps (Webb, 1986) - contoured lines of equal pollen percentages. These maps demonstrate the spatial expansion or contraction of different plant taxa throughout the Late Quaternary. In addition to showing the migration fronts of individual taxa, they also illustrate the abundance of pollen within their ranges. These isopoll maps lead to the interpretation of the changing patterns of tree taxa through time and depict how new and different plant community associations have evolved (Webb, 1987).

Both isopoll and migration maps show that taxa migrated at rapid rates in response to changing climates during the Late glacial and Holocene in North America (Huntley and Webb, 1989). Rates of migration of plant taxa varied and are estimated to range from 300 to 1000 meters per year (Brubaker, 1986; Huntley and Webb, 1989). Climate has played a role in the movements of most trees at some time during the mid- to late-Holocene (Davis, 1976). What is in question is how the various taxa migrated at speeds exceeding those which would be predicted by contemporary knowledge of tree species life-history traits. Important traits that affect migration rates are seed production, dispersal ability and seedling survival (Brubaker, 1986). For example, early succession species such as alder (*Alnus* sp.), birch (*Betula* sp.) and poplar (*Populus* sp.) produce seeds that are more easily dispersed by the wind, compared to late successional species such as beech (*Fagus sp.*), oak (*Quercus* sp.) and fir (*Abies* sp.) (Brubaker, 1986). Therefore, their migration rates should exceed those of late successional species. However, the problem is that many plant taxa with varying life-history traits moved at the same rate during the Holocene (Brubaker, 1986, Webb, 1982).

Modeling exercises have shown that the migration rates were far to high to be produced from life-history traits alone (Clark *et al.*, 1998). Tree species migrations have been viewed as originating in a diffusive manner along a frontal boundary area where propagules are normally distributed away from the front. However, using standard diffusion equations, Clark *et al.* (1998) illustrate that this approach cannot explain the rapid rates of Holocene tree migrations. Preliminary models indicate that long-distance stochastic establishment, or the establishment of local outlying population events may explain the observed rapid migration (Clark *et al.*, 1998). Under this scenario, a climate change providing favorable conditions for the growth and expansion of these outlying populations would result in an increased rate of migration of the newly established plants. The question of how tree migrations occurred so rapidly is still unanswered but model simulations are beginning to question old theories of migration. Further analysis of fossil pollen networks and modeling exercises may provide evidence for this 'outlying' population hypothesis (Bennett, 1985). For example, spruce moved most rapidly northward between ~12 ka and 9.5 ka in western Canada (Ritchie and MacDonald, 1986). In fact, in the region from central Alberta to the Mackenzie River Delta spruce pollen increased almost simultaneously between 10ka and 9.5 ka over an 11 site transect along the ice-free corridor between the Western Cordillera and the wasting Laurentide ice sheet (Ritchie and MacDonald, 1986). In this case, the renewed flowing Mackenzie River could have rapidly dispersed propagules of spruce. The rapid migration of spruce over a 2000 km latitudinal range in less than 1000 years has implications for questions involving the response rate of vegetation to climate change.

The second major question is whether plant taxa were in equilibrium with climate during their rapid migrations. The debate centers around the fact that interpretation of mapped pollen data has lead to two separate, yet equally plausible interpretations of vegetation change during the Late Quaternary: '*Complex dynamic vegetation responses*' and '*Complex past climatic changes*' (Webb, 1986).

First, the complex dynamic vegetation response is based on endogenous factors. Life-history traits such as seed production, dispersal mechanisms, propagule viability, fecundity, shade tolerance, growth rate, and fire susceptibility are endogenous variables. For example, the dispersal mechanism of seeds may have important impacts in determining the rate of taxa on migration resulting in complex vegetation responses observed in mapped patterns for different taxa. Migration maps suggest that *Fagus grandifolia* lagged climate by approximately 1,000 years in reaching the western shore of Lake Michigan (Davis *et al.*, 1986). The authors suggested that since *Tsuga canadensis* (wind dispersed) was found on the western shore of Lake Michigan 1,000 years prior to *Fagus* and because both *Fagus* and *Tsuga* share a similar climatic niche today, then *Fagus* lagged behind climate by 1,000 years because of its dispersal mechanism. Thus, when compared to modern assemblages, there were some plant communities whose composition was not reflective of their climatic situation. This is known as the no-analogue problem (Webb, 1986).

However, accepting the dynamic vegetation response hypothesis requires the assumption that *Fagus* and *Tsuga* occupy the same climatic space. In other words, the assumption is that modern association of the two trees represents the full range of the shared climatic space. If this is true then *Tsuga* did occur on the western shores of Lake Michigan 1,000 years before *Fagus* even if both taxa occupy the same modern ecological space. Therefore it can be concluded that climatic conditions were favorable for *Fagus* west of Lake Michigan 1,000 years before its migration into the area. The delay of *Fagus* was due to its dispersal mechanism and by the topographical barrier of Lake Michigan. Hence, *Fagus* lagged behind climate by 1,000 years west of Lake Michigan.

Although convincing, this logical sequence is not without problems. First, there is no independent climate evidence with which to compare the vegetation changes. Secondly, *Fagus* may have been present in very low numbers west of Lake Michigan and thus not appeared in the pollen record (Bennett, 1985; Ritchie and MacDonald, 1986). In addition, it is not climate that is being compared to the vegetation changes but rather it is modern *Fagus-Tsuga* climate relations that are used to analyze the pollen record. These modern relations assume that all possible combinations of temperature and precipitation, which are acceptable for *Fagus* and *Tsuga*, are represented within the range of their modern association.

Alternatively, the complex climate changes school of thought can also be used to interpret the same pollen data. This hypothesis is based on the fact that the complex vegetation response argument may be assuming an oversimplified view of how climate changes, hence, how different taxa respond to those changes. The idea behind this hypothesis is that exogenous (external) forcing may be complex. For example, temperature and precipitation vary independently from each other both seasonally and spatially and thus cause different responses in different taxa (Webb, 1986). According to this hypothesis, *Fagus* migration in North America may not have been delayed by its dispersal mechanism but rather the climate conditions west of Lake Michigan delayed its expansion into the area for 1,000 years. In another example from Europe, *Fagus* migrated northward at the same time that *Ostyra-Carpinus* showed a southward migration after the mid-Holocene climate optimum (Huntley and Webb, 1989). Because of the southward migration of the northern range boundary of *Ostyra-Carpinus* it had been concluded that *Fagus* moved north at a time of climate cooling and, therefore, the movement of *Fagus* northward could not be due to climate (Huntley and Webb, 1989). However, at 10ka the northern hemisphere experienced maximum insolation in summer, which meant increased seasonality. Since then maximum insolation in the northern hemisphere occurs in winter thereby decreasing seasonality (Huntley and Webb, 1989). This lead to warmer winters with more snow, which can potentially protect *Fagus* propagules and seeds (Huntley and Webb, 1989). Thus, seasonality may have been the factor of climate that led to *Fagus* migrating north, rather than cooler mean annual temperatures.

It is clear that species found together today were not necessarily found together in the past and that species found together today migrated at different rates and at different times during the Holocene. However does this mean that past vegetation was not in equilibrium with climate? The answer to this question largely depends on which hypothesis - complex dynamic vegetation or complex past climates - one accepts. However, it is likely that both hypotheses were important on various time and spatial scales.

Until recently there has been no definitive way to select among these alternative hypotheses since the presence of a species in outlying marginal areas cannot necessarily be determined by pollen percentages alone. What is required is either independent vegetation data or new ways to synthesize pollen spectra. For example, if a species existed regionally but was not strongly evident in the pollen record, it would provide evidence that vegetation-climate lags were not due to complex dynamic vegetation responses but the result of complex climates of the past. Independent climate evidence comes from General Circulation Models applied to past time periods (Vettoretti et *al.* 1997; Kutzbach and Webb 1993). However, GCM models applied to past climates are still preliminary and have yet reproduced broad regional-scale past climates evident in the proxy record reconstructions (Gajewski *et al.* 2000; Jackson *et al.* 1997; Viau, 1999; Viau and Gajewski, 2001). Independent large-scale vegetation macrofossil reconstructions are able to provide evidence of the presence of vegetation taxa, which may not be evident in pollen records. For example, Jackson *et al.* (1997) compared macrofossil distributions in eastern North America with isopoll maps at 3 ka intervals to test specific hypotheses, including the "no analogue" question through the Holocene. They suggested that the large-scale no-analogue assemblages around 12 ka west of Appalachia for the association of *Picea*, *Cyperaceae* and *Artemesia*, where *Abies*, *Larix* and *Betula* should have been found in the region with spruce was not due to migration lags. They showed that macrofossils of *Abies*, *Larix* and *Betula* occurred over most of the no-analogue range (Jackson *et al.*, 1997). This independent macrofossil evidence suggests that the forest composition was dominated by different genera. Since the taxa were present, migration lags can be ruled out as a cause of the no-analogue assemblage.

As Jackson *et al.* (1997) state "unique combinations of climatic variables led to unique combinations of plant species on the landscape". Climates different from today would have selected certain taxa to be more successful, based on their life-history traits, leading to different competitive relations under the large-scale environmental conditions present at the time. This finding supports the argument that plant taxa responded individually to past climate changes. However, the pollen assemblages were not due to vegetation assemblages unrepresentative of their climatic situation. Rather, the vegetation assemblages were likely in equilibrium with the climates of their time and the non-analogues are due to climates without modern analogues. Much of the evidence previously interpreted as resulting from desequilibrium conditions can instead be interpreted as resulting from the individualistic response of plant taxa to the different regional patterns of temperature and precipitation of the past (Webb, 1986).

*Acknowledgment: Initial unpublished material provided by M. C. Sawada. Modified and revised by A. E. Viau.

REFERENCES

Bennett, K.D. 1985. The spread of Fagus grandifolia across eastern North America during the last 18,000 years. *Journal of Biogeography*, 12:147-164.

Brubaker, L.B., 1986. Responses of tree populations to climatic change. *Vegetatio* 67: 119-130.

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., and Wyckoff, P. 1998. Reid's paradox of rapid plant migration: Dispersal theory and interpretation of paleoecological records. *Bioscience*, 48(1): 13-24

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*, 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., adn Futyma, R.P. 1986. Dispersal versus cliamte: Expansion of Fagus and Tsuga into the Upper Great Lakes region. *Vegetatio* 67:93-103.

Gajewski, K., Payette, S., and Ritchie, J.C. 1993. Holocene vegetation history at the Boreal Forest--shrub-tundra transition in northwestern Québec. *Arctic and Alpine Research* 24, 329-336.

Gajewski, K., Vance, R., Sawada, M., Fung, I., Gignac, L.D., Halsey, L., John, J., Maisongrande, P., Mandell, P., Mudie, P., Richard, P., Sherrin, P., Soroko, J., Vitt, D. 2000. The climate of North America and adjacent ocean waters *ca.* 6ka. *Canadian Journal of Earth Sciences.* 37: 661-681.

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

Jackson, S.T., Overpeck, J.T., Webb, T., Keattch, S.E., 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.

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

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

Prentice. I.C. 1986. Vegetation responses to past climatic variation. *Vegetatio* 67, 131-141.

Ritchie, J.C. and MacDonald, G.M. 1986. The patterns of post-glacial spread of white spruce. *Journal of Biogeography*, 13:527-540.

Sawada, M.C. 2001. *Late Quaternary Paleoclimates and Biogeography of North America*. Unpublished Ph.D. thesis, University of Ottawa, Ottawa, Ontario, Canada.

Vettoretti, G., W.R. Peltier, and N.A. McFarlane, Simulations of Mid-Holocene climate using an atmospheric general circulation model, *J. Clim.*, 11, 2607-2627, 1998.

Viau, A.E. 1999. *Lake level fluctuations and Global hydrological change: a spatio-temporal analysis*. Unpublished M.A. thesis. University of Ottawa, Ontario, Canada, 210 pp.

Viau, A.E. and Gajewski, K., 2001. Holocene variations in the global hydrological cycle quantified by objective gridding of lake level databases. *Journal of Geophysical Research-Atmosphere* 106(D23): 31703-31716

Webb, T. 1982. *Is Vegetation in Equilibrium with Climate?* Unpublished Manuscript, Quaternary Research Center, Seattle, Washington.

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

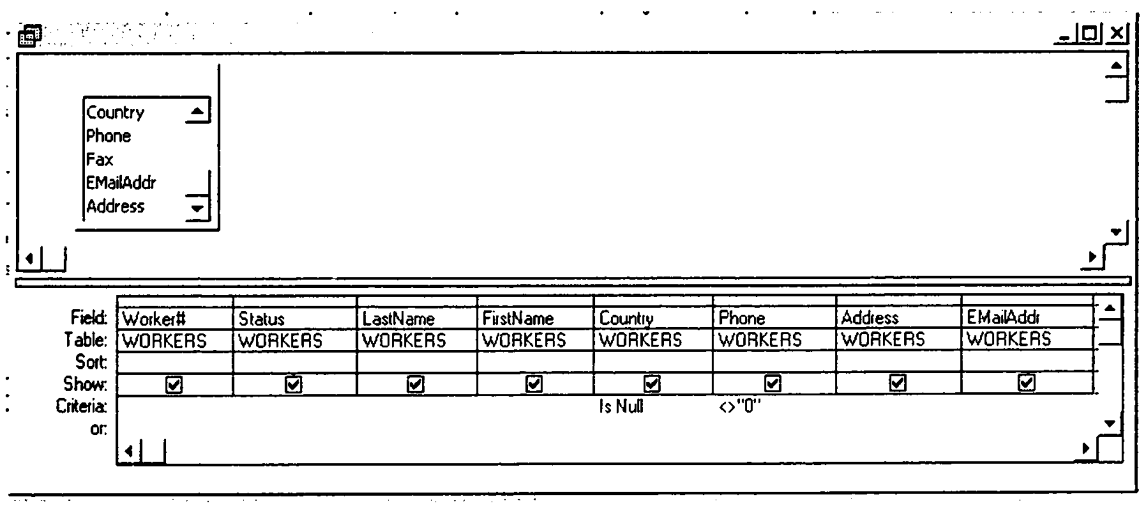
Webb, T. 1987. The appearance and disappearance of major vegetational assemblages: Long-term vegetational dynamics in eastern North America. *Vegetatio* 69:177-187

APPENDIX B: DESCRIPTION OF CONTENT OF NAPD

Table WORKERS

There are 747 individual workers whom contributed to entries in the NAPD, however,it is understood that this index could represent other worker than principal investigators.

a) 303 are listed as Active, 23 as Inactive, 7 deceased, and 414 with Unknown status.
b) 326 workers have a residence listed with 180 workers residing in the USA, 80 in Canada, and 66 workers from other countries of the world. 421 workers are listed with unknown residence (ie. null field).
c) 459 workers have no contact information while 248 have some contact information (ie. phone, address or e-mail).



Example of query for WORKERS table.

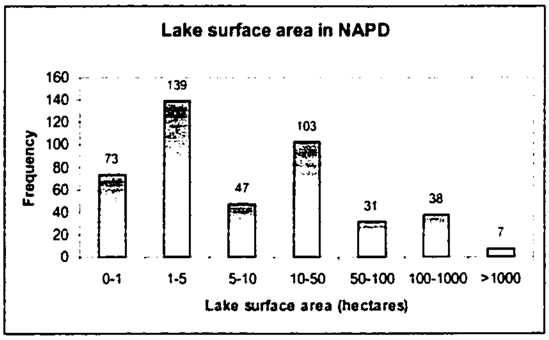
Table SITELOC

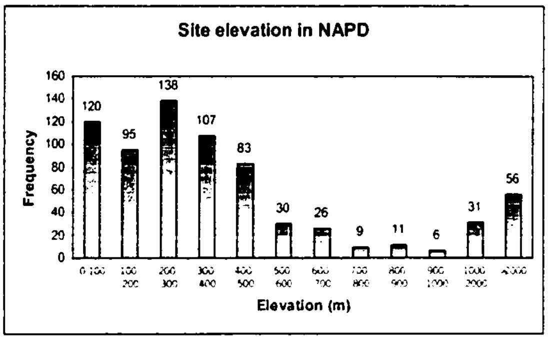
This table contains information about sites entered in the database. There are 752 sites in the NAPD.

a) 446 sites are in the USA, 300 in Canada, 4 in Mexico, 1 in Greenland, and 1 in Bermuda;
b) 438 sites contains surface area (ie. in hectares), 312 sites have no information on the surface area of the lake, leaving 2 sites with ambiguous information (ie.0 hectares, 1 hectare =100x100m, 1 hectare/100=km2).

c) 40 sites are missing elevation information, 712 have the elevations entered.
d) 477 sites or 63% are located between 40 and 50° latitude. 490 or 65% of sites lie in between –70 and –100° longitude West. Therefore 51% or 387 sites are restricted within the regional quadrant of the Great Lakes (latitudes 40 to 50° and –70 to –100° longitude).

Figure 1: Fossil pollen site distribution in North America (source:NAPD, Grimm, 2000)

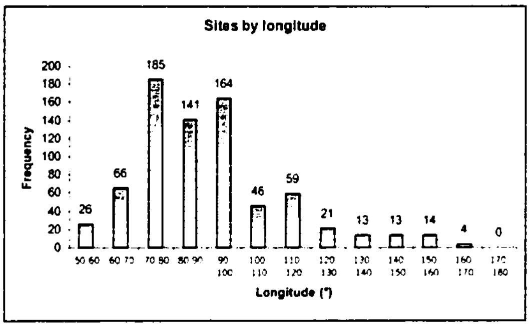


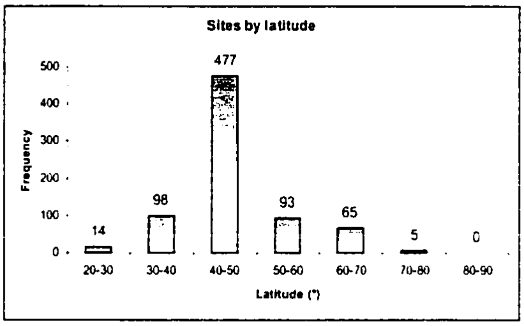


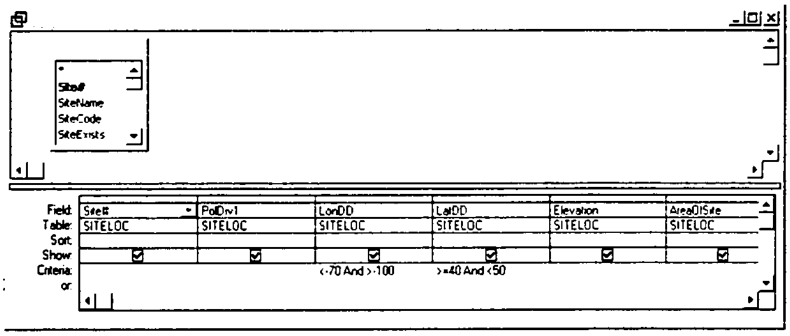
	Elevation (m)	Hectares	Km2
MIN	0	0.01	0.0001
MAX	3740	95125	951.3
AVG	526.6	313.5	3.1
STD	688.3	4615	46.2
MEDIAN	302	5.7	0.1
MODE	305	2	0.02

				1										
			2	17	2			4	1		1			
			1	10	7	1				1				
				1	8	1	6			4	1			
				1	4	2	3			12	20	13	1	
				2	6		4		2	5	4	13	8	
				1	9	4	3	8	12	11	51	17	2	
				4	1	26	2	107	41	57	58	12		
				9		2	5	8	18	23	12		1	
				2	2	19	2	4	5	10	3			
					2			5	8	6	9			
									2					

Figure 2: Arbitrary(15x15) grid with sum of sites.







Spatial statistics – distances based on nearest neighbor. Original Plate Carré (lon/lat) were re-projected into Lambert Oblique Azimuthal Equal Area projection (Datum NAD27) in meters for computation.

	MIN	1ST QUARTILE	MEDIAN	MEAN	3RD QUARTILE	MAX
Meters	37.5	8771.0	25647.3	52428.4	58513.0	1105940.3
Kilometers	0.04	8.77	25.65	52.43	58.51	1105.94

Table SITEDESC

This table stores textual descriptions such as surrounding vegetation and site descriptions.

a) There are 595 sites that contains site description information (out of 752).
b) 546 sites with site description
c) 519 sites with physiographic description
d) 446 sites that contain information about the vegetation surrounding
e) 583 sites have vegetation formation information. 205 different vegetation formation types are entered in the NAPD.
f) 0 sites contain an IGBPtype region code.

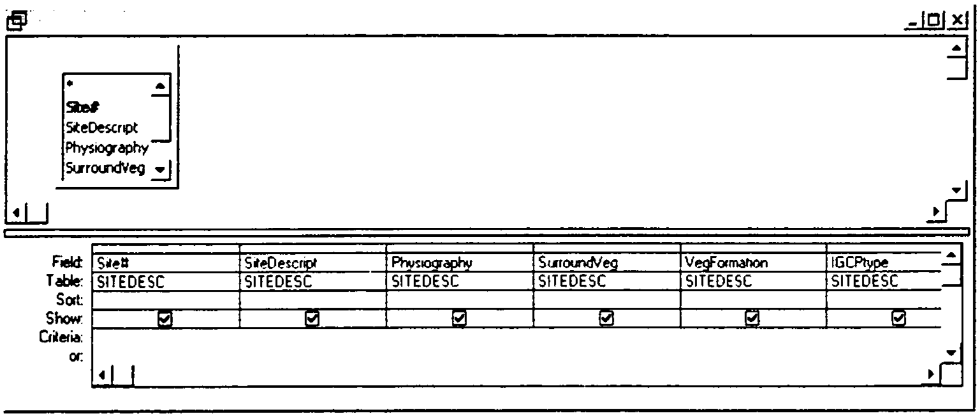


Table SITEINFO

This table stores cross-reference information between sites and publications assigned with them.

a) There is presently no information stored in this table.

Table ENTITY

This table is the core table of the NAPD and stores information about the description of an entity. An entity is a core, a section, or a surface sample. There can also be multiple entities for a single site.

a) There are 909 entities stored in the NAPD associated with 752 individual sites which means that some sites have more than one entity associated with them.
b) However, each entity has an individual sigle (ie. Unique code for the entity).
c) 739 entities are cores of which 538 core entities store modern sample information and 201 entities have fossil information only (ie. starts at 2000 BP).
d) 73 entities are sections where 14 entities are sections from a core and 59 section from other sources. 31 sections have modern information while 42 do not.
e) 97 entities are neither a core nor a section where 85 entities are modern surface sample and 12 are not.
f) Therefore, 654 entities contain modern information leaving 255 entities with no modern samples.

	Cores	Sections	Other	Total
Entities	739	73	97	909
Modern	538	31	85	654
Fossil	201	42	12	255
Total	739	73	97	909

g) Each entity has a unique identifier for its description. This unique identifier is related to table DESCR for a full textual description.
h) There are 43 distinct identifier.

Descriptor	Description	Total entities
ARCH	archaeological	1
CEST	estuarine	1
COAS	coastal	3
CSLT	salt marsh	1
FLUV	fluvial	6
HOLL	small hollow	1
LACU	lacustrine	190
LCIR	cirque lake	5
LDRA	drained lake	5
LDUN	interdunal lake	2

LEXP	explosion crater	1
LFLU	fluvial origin, natural open water	13
LGLA	glacial origin, natural open water	109
LGMC	glacial meltwater channel	1
LGSC	glacial scour lake	32
LKET	kettle lake	138
LLSL	landslide dammed lake	1
LMOR	morainally dammed lake	12
LNAT	natural open water	64
LPLY	playa	5
LPRO	pro-glacial lake	1
LRES	reservoir	1
LSOL	solution hollow, natural open water	25
LTHK	thermokarst lake	1
LUNK	origin unknown, natural open water	3
LVOL	volcanic origin lake basin	3
LWND	wind origin lake basin	1
MPAC	packrat (Neotoma) midden	8
OTHR	other	2
TBOG	bog	44
TCAV	cave sediment	3
TCUT	stream-cut exposure	7
TERR	terrestrial	18
TFEN	fen	20
TFLO	floating mire	3
TLUS	loess	4
TMIR	mire	98
TMOS	moss polster	23
TMSH	marsh	17
TRAI	raised bog	7
TSOI	soil	3
TSWA	swamp	7
UNKN	Unknown	19
Total	43	909

i) Presented below are summary counts for first level hierarchy descriptions.

Descriptor	Description	Total entities
ARCH	archaeological	1
COAS	coastal	5
FLUV	fluvial	6
HOLL	small hollow	1
LACU	lacustrine	613
MPAC	packrat (Neotoma) midden	8

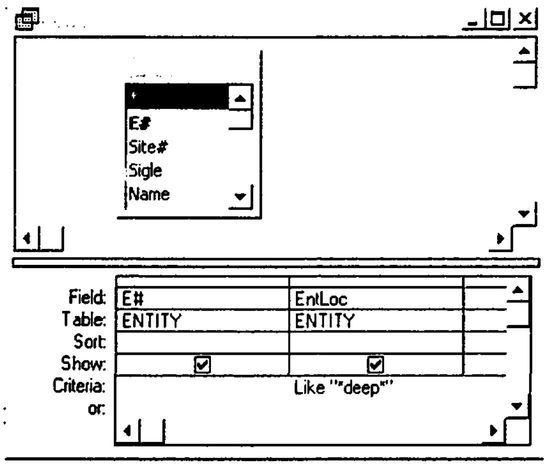
OTHR	other	2
TERR	terrestrial	75
TMIR	mire	179
UNKN	Unknown	19
Total	10	909

j) Out of 753 cores or section, 17 have lamination and 736 do not.

E.	Site.	SiteName	LonDD	LatDD	Elevation	AreaOfSite	HasAnLam
12	12	Lac à l'Ange	-70.68	47.48	648	0.5	Y
15	15	Basin Pond	-70.05	44.47	124	14	Y
54	54	Conroy Lake	-67.88	46.28	140	10	Y
59	59	Crawford Lake	-79.83	43.5	102		Y
66	66	Dark Lake	-91.47	45.27	334	5.3	Y
86	86	Found Lake	-78.5	45.5	488		Y
89	89	Hayes Lake	-93.75	49.58	391		Y
107	107	Hell's Kitchen Lake	-89.7	46.18	488	3	Y
146	146	Little Pine Lake	-91.48	45.28	330	4	Y
224	223	Ruby Lake	-91.45	45.28	335	6.9	Y
857	792	Battle Ground Lake	-122.49	45.8	155	13.5	Y
907	841	Ely Lake	-75.83	41.76	384	14	Y
908	842	Clear Pond [New York]	-74.02	43.75	513	10	Y
1141	1053	Molas Lake	-107.68	37.74	3200	16	Y
1142	1054	Molas Pass Bog	-107.69	37.73	3220	1	Y
1155	1063	Hanson Marsh	-89.67	43.43	381	0	Y
1277	1160	Moon Lake	-98.15	46.87	444	35	Y

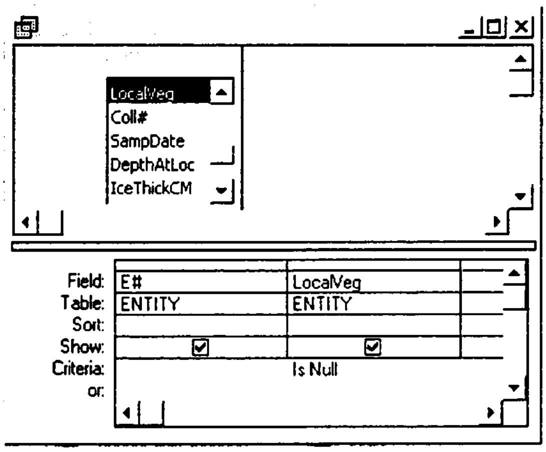
k) The location where the core was taken in the lake includes 239 different location entries in the NAPD. In addition, only 433 entities have entries concerning the location leaving 476 entities with unknown location of sample for a total of 909 entities. 84 entities were cored at the deepest part of the lake. 176 entities were collected at the center of the lake. 11 entities were sample in the middle of the lake, 3 at maximum water depth, 4 at greatest water depth. Therefore it is assumed that 278 entities were sampled near or in the deepest part of the lake. 155 entities of these 433 entities were collected at other locations that can be ambiguous in interpretation (ie. 20m from shore; wettest part of fen etc…).

Location	deepest part of lake	Other	Unknown	Total
Entities	278	155	476	909



l) The ENTITY table also stores information concerning local vegetation at the sites. 552 of 909 entities have stored information on local vegetation. 406 of 552 entities are stored as "open water" suggesting lakes. 146 entities have some information on local vegetation, however this information is very general (ie. cultivated field, sphagnum, ericaceae, sedge, peatland....)

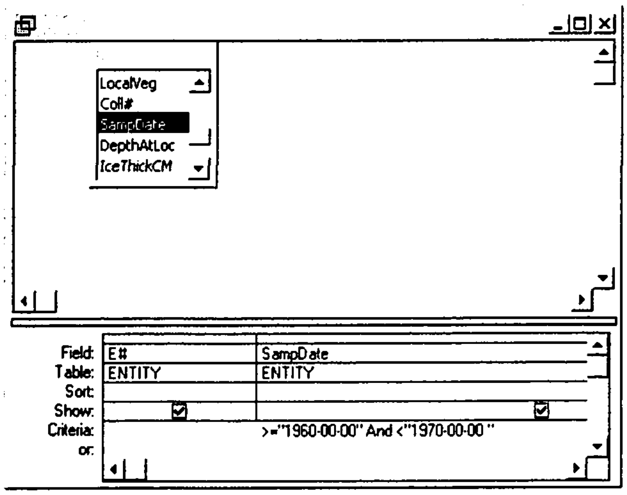
Local vegetation	Open water	Other	Unknown	Total
Entities	406	146	357	909



m) This table stores information on the dates that the sample were collected. However, there seems to be many unknown dates, 400 entities with a entry of "0000-00-00". As far as the rest of the dates it suggests that most of the entities in the NAPD were collected from the 60's through the 90's.

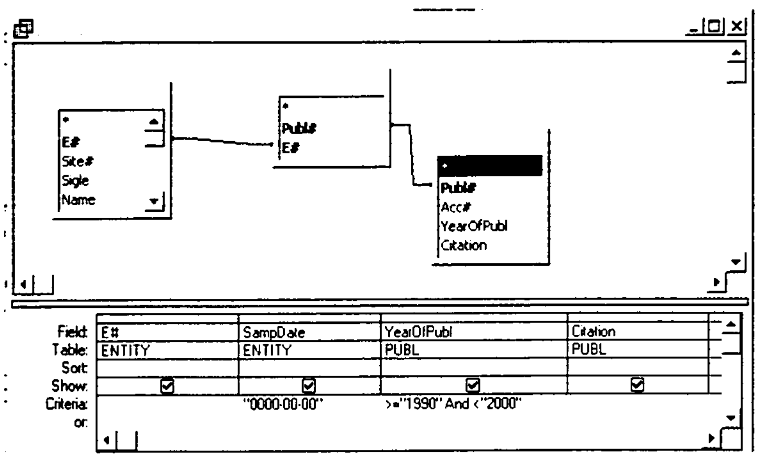
n) It may be worthwhile to cross-check the database with known publications as to when the samples were collected.

Dates	<1960	1960-1970	1970-1980	1980-1990	>1990	0000-00-00	Total
Entities	18	96	211	153	31	400	909

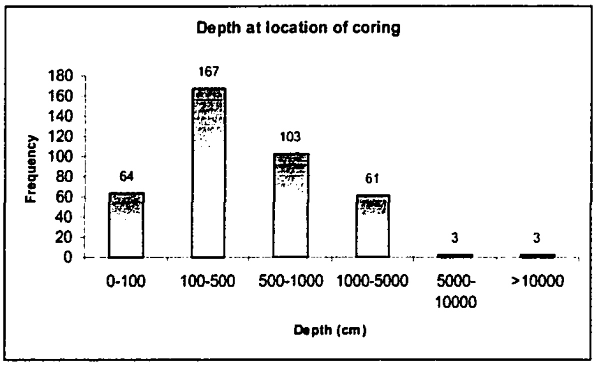


Year of publ.	Unknown	"0000"	<1960	1960-1970	1970-1980	1980-1990	>1990	Total
Entities	90	3	6	51	92	119	74	435

Although there are only 400 entities with sample dates of "0000-00-00", this table gives a good approximation of the year of publication for these entities. As we can notice there are 35 entities that have multiple year of publications overlapping decades.



o) Where cores and surface samples are concerned is information on the water depth at the coring location. There are 401of 909 entities with depth information.



MIN	4
MAX	22300
AVG	738.59
STD	1713.8
MEDIAN	400
MODE	100

p) Perhaps of less concern, except Hiller, for example tends to disturb the sediments, is the sampling device used in the coring. 208 entities have no stored information. 313 entities used a modified or regular Livingston piston corer while 308 entities used other coring devices such as the Hiller sampler, plastic tube with piston and others.

q) Core diameter information is also stored for cores only. Out of 477 entities, 325 have a 5 cm diameter, 68 have a diameter of less than 5 cm, and 85 entities have a diameter greater than 5 cm. Two entities have a core diameter greater than 100 cm with a maximum of 400 cm.

Table P_ENTITY

This table stores additional information on entities as it is a subtype table of ENTITY.

a) Where does the data come from is a fundamental question in any statistical analysis. It can give some insight into how the data was collected.

b) Each 909 entity has data source information. 686 entities source come from the "originator" of the analysis. 193 come from the COHMAP members (Webbs database) and 12 from publications. 18 come from individuals or other sources.
c) Of concern is how the counts were entered in the NAPD.

Dataform	Raw counts	Raw percent.	Digitized counts	Digitized percent.	Null	Total
Entities	729	2	10	13	155	909

d) 906 entities come from published work leaving only 3 entities with unpublished data.

Table GEOCHRON

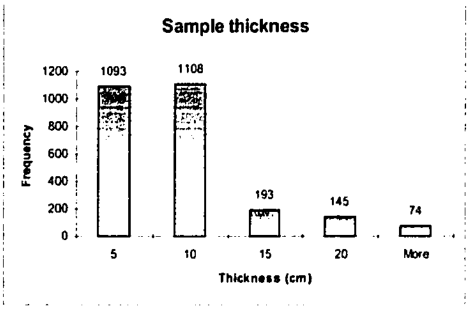
This table stores geochronological information.

a) 674 entities have geochronological information.

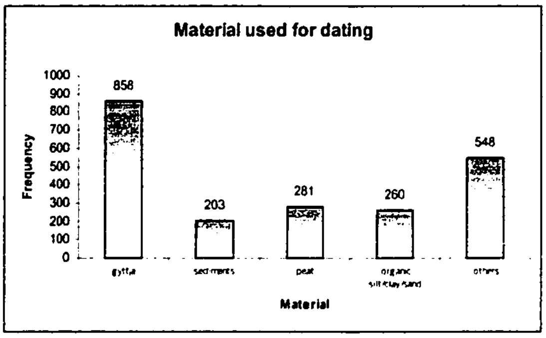
Method	AAR	C14	ESR	FT	KAR	210P	si32	Tl	USERIES	Total
Entities	0	660	0	0	0	12	1	1	0	674

AAA- amino acid racemization
C14 – radiocarbon
ESR - electron spin resonance
FT - fission track
KAR - potassium/argon
210P - lead-210
Si32 – Silicate isotope
TL – thermoluminescence
Useries – uranium series

b) There are 672 entities that contain information on sample thickness.



b) There are 384 different materials reported in the NAPD for dating purposes. However, the main materials are gyttja, organic sediments and peat. Others include wood, needles, leaves, clay, shells, ostracods, mollusc, etc....



Note: these values may contain errors due to overlaps. However, the errors are not significant enough to make any differences.

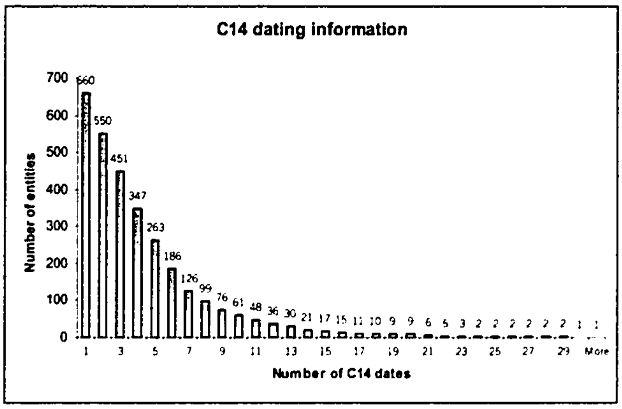
Table AAR

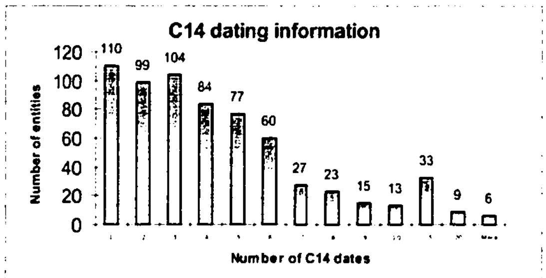
No Amino Acid Racemization dating information is presently entered in the NAPD.

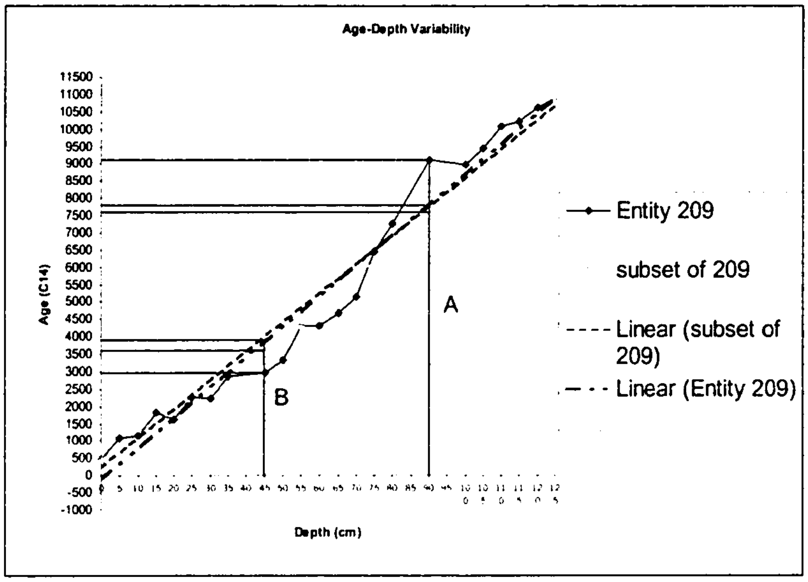
Table C14

There are 660 entities with stored C14 dating information.

a) These two frequency histograms reveal quality dating control information of entities. These dates are used to derive an age-depth model to estimate ages at depths not sampled. Unless a constant linear sedimentation rate is assumed; less dates used will result in a less reliable estimate (see Figure below). Moreover, given that sedimentation rates are non linear, the model imposed will suffer from temporal resolution and therefore loose some variability in the sedimentation rates. Therefore the estimate of the age may be off by several centuries or smoothed through







In this example, entity 209 has 23 C14 dates on which to rely for age/depth estimate. However, using a subset of only 3 dates, we can see that in both A and B the difference in age estimate can be substantial (eg. > 1000 yrs) .

A.

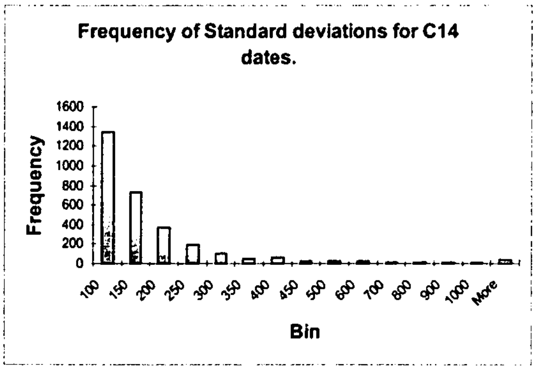
B.

These figures show the distribution of sites that have 5 or more C14 dates in A. and 10 or more dates in B.

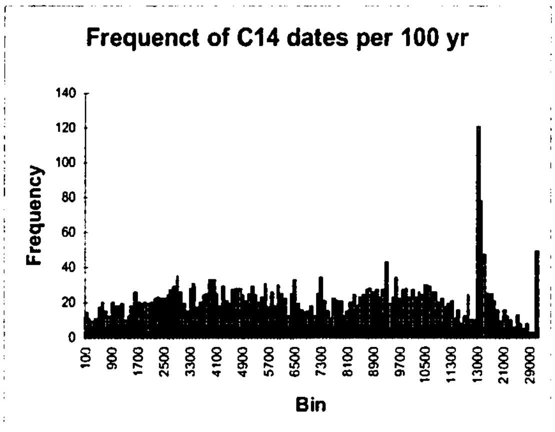
c) Except for dates greater than 25000 years, the C14 dates entered in the database are considered good.
d) The method used to make the radiocarbon determination.

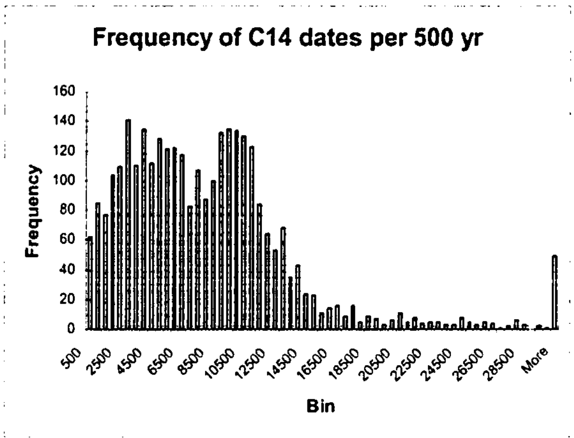
Method	Gas decay count	Accelerator	Decay count liquid scintillation	decay count unknown	No information	Total
Entities	1563	269	659	331	254	3076

e) There are 3076 C14 dates associated with the 660 entities.



	Left SD	Right SD
Min	0	0
Max	14480	14480
Avg	179.52221	179.68345
Std	438.9776	440.39285
Mode	80	80
Median	110	110





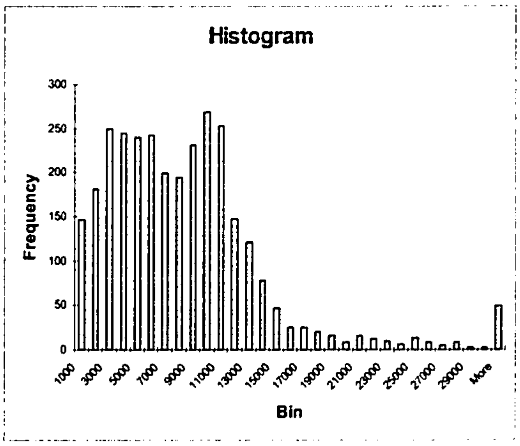


Table ESR

There are presently no entries in this table.

Table FT

There are presently no entries in this table.

Table KAR

There are presently no entries in this table.

Table PB210

There are 12 entities that contain Pb-210 information. These sites are clustered between 44 and 48° latitude North and 84 to 94° longitude West. The minimum date is 1792 and maximum date of 1983. However the standard error for 1792 is 332 years. A total of 138 dates are included in this table.

Table SI32

Only one entity has stored silicon-32 dating information. Entity 969 is located on Devon Island ice cap. The dates span from present to 4600 years BP with a total of 8 dates with standard deviation ranging from 50 to 550 years BP.

Table TL

Only one entity has stored thermoluminescence dating information. Entity 1067 is of January Cave deposits in British Columbia, Canada. There are 4 dates recorded ranging from 1034 to 54132 years BP. The older dates have errors limits of almost 6000 years.

Table USERIES

There are presently no entries recorded in this table.

Table SYNEVENT

There are presently no entries recorded in this table.

Table EVENT

There are presently no entries recorded in this table.

Table ALSEGS

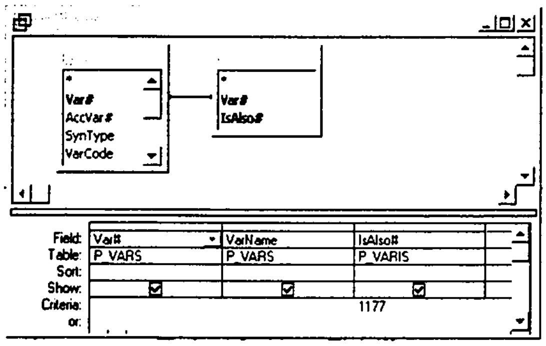
Although there are 8 entities with records in this table, there is no information recorded but the number of the segments which are all only one segment.

Table P_ANLDPT

This tables stores the information on annual laminations for each of an entity's segments. The same 8 entities stored in table ALSEGS contain information on depth, thickness and lamination counts for each entities's samples.

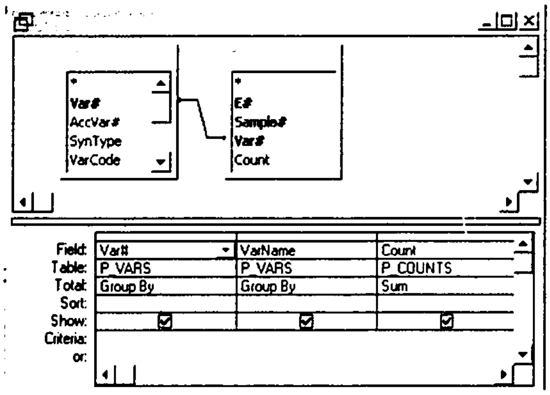
Table P_VARS

This table stores the names, codes and hierarchy information for each pollen variable entered in the NAPD.
There are 3673 individual pollen variable names.



This simple query extracts hierarchical information for a given variable. One simply enters the varcode of a pollen type in the IsAlso# field to retrieve all types underneath this variable.

Counts	variables
1	187
50	646
100	135
200	99
500	145
1000	81
5000	136
10000	45
25000	68
50000	23
75000	13
100000	5
More	61
Total	1644



Of 3673 identifiable pollen variables, only 1644 variables have actual counts entered in the NAPD. Of these 1644 variables, some pollen types are very poorly represented. See table above.

Table P-SAMPLE

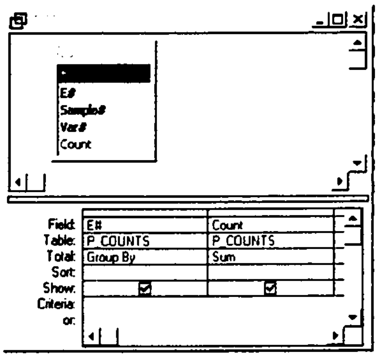
This table store information on depth, thickness of samples for each entity. There 891 entity with stored information. In addition, there is information on the analyst who did the counting and at what date the analysis was completed. Sample analysed are extracted at intervals ranging in majority from 5 to 10 cm. Thickness of sample ranges from as little as 0.5 to 10 cm.

Table P-COUNTS

This table store the pollen counts for each entity at every sample depth (ie. stored in P_SAMPLE table).
There are 887 entities with stored pollen counts. As previously shown (see table P_VARS) some variables are more represented than other.

Counts	Entities
200	3
500	38
1000	24
2000	20
5000	64
10000	98
25000	202
50000	100
100000	46
More	292

Note: The counts > 25000 are probably pre mil derived from percentages, not real counts



This query sums all pollen count for each entity to give insight into how many pollen grains are counted with some errors.

Table LITHOLOGY

There are 287 entities with stored lithological information like transition from dark to light gyttja.

Table LOI

This table store Loss-on-ignition information for a entity. There are 115 entities with stored LOI information.

Table PUBL

This table store publication information. Refer to Table WORKERS.

Table PUBLENT

This table is a cross-reference for publication information from entities. For example one can extract which entity a certain publication refers to.

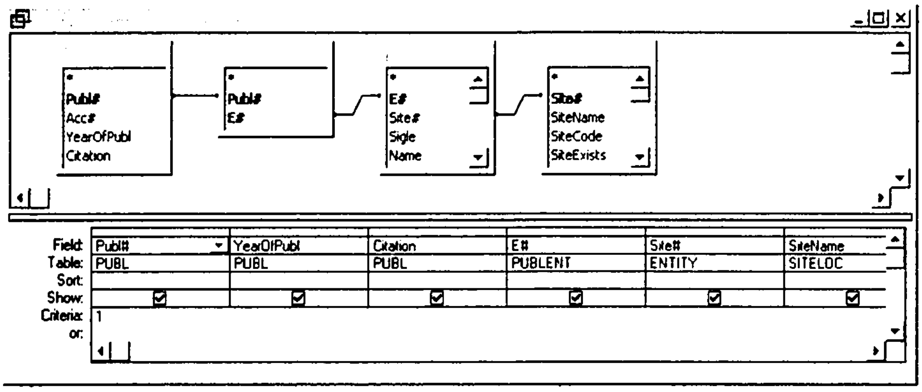


Table AUTHORS

This table cross-references authors with their publications. For more information, see Table Workers.

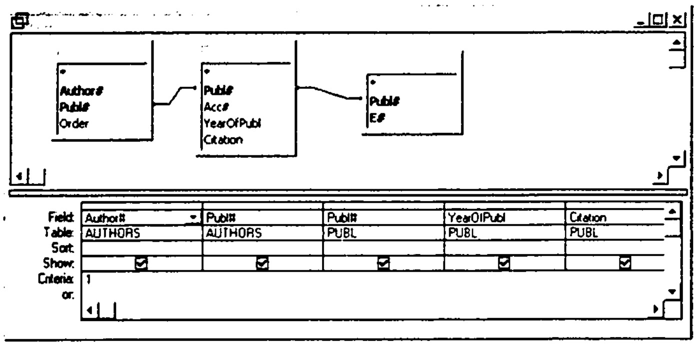


Table POLDIV1

Although there are 227 possible country codes entered in this table, only 5 codes relate to the NAPD. Can- Canada; Grl – Greenland; Mex – Mexico, USA – United States and BMU – Bermuda. All 909 entities have coded POLDIV1.

Table POLDIV2

This table stores information relating to 2nd oder polital subdivisions such as the canadian provinces, or states of the USA or even region of other countries.

Table POLDIV3

This table stores information on 3rd political subdivisions such as specific regions or counties.

Table SYNTYPE

Stores synonym codes and rationale for table p_vars.

Table INFOTYPE

Look-up table that stores full names for 3 character codes in the NAPD.

ICode	InfoType
AAR	amino acid racemization
ALG	Algae
ARC	archaeology
C14	radiocarbon
CES	cesium-134/cesium-137
CHA	Charcoal
CHI	chironomids
CLA	cladocerans
COL	coleopterans
DEN	dendrochronology
DIA	Diatoms
ESR	electron spin resonance
FIS	fission track
FOR	foraminiferans
GEO	geochemistry
GRA	granulometry
INS	other insects
ISO	other stable isotopes
KAR	Potassium/argon
LED	lead-210
LOI	loss on ignition
LUM	thermoluminescence
MAC	macrofossils
MAG	palaeomagnetism
MOL	Molluscs
OST	Ostracods
OTH	Other
OXY	stable oxygen isotopes
POL	Pollen
RAD	radon-226
RHI	rhizopods

TEF	Tephra
URA	u-series

Table DESCR

This is a fundamental table which stores information on site-description.

Descriptor	Description	Entities	Percentage	Hierarchy 1	Hierarchy 2	Hierarchy 3
AAIR	air pollen sampler	0	0.00		0.00	
AERI	Aerial	0	0.00	0.00		
ARCH	archaeological	1	0.11	0.11		
CEST	Estuarine	1	0.11		0.11	
COAS	Coastal	3	0.33	0.55		
CSLT	salt marsh	1	0.11		0.11	
FLUV	Fluvial	6	0.66	0.66		
HOLL	small hollow	1	0.11	0.11		
HVER	vernal pool	0	0.00		0.00	
LACU	Lacustrine	190	20.90	67.44		
LART	artificial open water	0	0.00		0.11	
LCIR	cirque lake	5	0.55			
LDEF	Deflation basin lake	0	0.00			
LDRA	drained lake	5	0.55		0.55	
LDUN	interdunal lake	2	0.22			
LEXP	Explosion crater	1	0.11			0.11
LFLU	fluvial origin, natural open water	13	1.43			1.43
LGLA	glacial origin, natural open water	109	11.99			32.67
LGMC	glacial meltwater channel	1	0.11			
LGSC	glacial scour lake	32	3.52			
LKET	kettle lake	138	15.18			
LLSL	Landslide dammed lake	1	0.11			
LMOR	morainally dammed lake	12	1.32			
LNAT	natural open water	64	7.04		44.33	
LOTH	other, artificial open water	0	0.00			0.00
LPER	periglacial origin, natural open water	0	0.00			0.11
LPLY	Playa	5	0.55		0.55	
LPRO	pro-glacial lake	1	0.11		0.11	
LRES	Reservoir	1	0.11			0.11
LSOL	solution hollow, natural open water	25	2.75			2.75
LSTK	stock pond	0	0.00			0.00
LTEC	tectonic lake	0	0.00			0.00
LTHK	thermokarst lake	1	0.11			
LUNK	origin unknown, natural open water	3	0.33			0.33

LVOL	volcanic origin lake basin	3	0.33		0.44	
LWND	wind origin lake basin	1	0.11			
MARI	Marine	0	0.00	0.00		
MIDD	animal midden	0	0.00	0.88		
MPAC	packrat (Neotoma) midden	8	0.88		0.88	
OTHR	Other	2	0.22	0.22		
PTLK	pollen trap (lacustrine)	0	0.00			
PTRP	pollen trap (terrestrial)	0	0.00	0.00	0.00	
TBLA	blanket bog	0	0.00			
TBOG	Bog	44	4.84			5.61
TCAV	cave sediment	3	0.33		0.33	
TCDU	cave sediment (animal dung)	0	0.00			
TCUT	stream-cut exposure	7	0.77			
TDET	organic detritus	0	0.00		0.00	
TERR	terrestrial	18	1.98	27.94		
TFEN	Fen	20	2.20			2.53
TFLO	floating mire	3	0.33			
TFLP	floodplain mire	0	0.00			
TLUS	Loess	4	0.44		0.44	
TMIR	Mire	98	10.78		19.69	
TMOR	mor humus	0	0.00			0.00
TMOS	moss polster	23	2.53		2.53	
TMSH	Marsh	17	1.87		1.87	
TOWT	open-water transition mire	0	0.00			
TRAI	raised bog	7	0.77			
TSOI	Soil	3	0.33		0.33	
TSOL	soligenous (mineraltrophic) mire	0	0.00			
TSWA	Swamp	7	0.77			0.77
TVAL	valley mire	0	0.00			0.00
UNKN	Unknown	19	2.09	2.09		

Table RATIONAL

This table stores the descriptions of the dating rationales, Used in conjucntion with AGEBASIS table.

RCode	Rationale
AMR	Ambrosia rise
ANL	annual laminations
ATL	beginning of the Atlantic
BOT	bottom of core
C13	Corrected for isotopic fractionation with C13

C14	uncorrected radiocarbon date
CAJ	adjusted for ancient carbon
CAL	C14 date calibrated to calendar years
CAV	Averaged radiocarbon date
CCA	calibrated radiocarbon date
CGT	radiocarbon greater than date
COM	Corrected for compaction
DEN	C14 date with dendrochronological correction
DGL	date of deglaciation
ESH	European settlement horizon
EXT	extrapolated date
GUE	Guess
HOL	beginning of the Holocene
INT	interpolated date
LST	Laarcher See tephra
OTH	Other
PB2	PB210 date
POL	pollen stratigraphic date
RCY	radiocarbon years calibrated from calendar years
SED	Sediment stratigraphic date
TEF	tephra date
TOP	top of core
TSD	Tsuga decline
USE	U-Series date

Table IGCPTYPE

This table stores the names and codes for the IGCP type regions. Refer to the IGCP catalogue of maps that indicate the type region boundaries.

Table P_GROUP

This research table stores the group to which each pollen variable belongs to except for pollen variables that have syntype code in p_VARS table. There is 3489 of 3673 variables that are assigned a group.

GroupId	GroupName	Variables	Variable with counts	Total counts
ALGA	Algae	27	17	710022.0
AQUA	Aquatics	148	95	1414459.8
BRYO	bryophytes	25	11	325003.0
DINO	dinoflagellates	21	2	3037.0
FUNG	Fungus	9	6	114398.0

INUN	indeterminables and unknowns	98	29	372619.3
NOPO	nonpollen	133	87	4450261401.2
PREQ	pre-Quaternary types	43	39	16831.5
RHIZ	rhizopods	31	16	3766.0
SPOR	taxonomically undifferentiated spores	38	16	100410.5
TRSH	trees and shrubs	1396	544	14595332.7
UPHE	upland herbs	1225	639	3638062.3
VACR	vascular cryptogams (pteriodphytes)	295	143	640539.6

Table GROUPS

This table stores codes and names for different pollen groups. It is useful to decode P_GROUP table.

GroupId	GroupCode	GroupName
ALGA	L	algae
AQUA	Q	aquatics
BRYO	M	bryophytes
DINO	D	dinoflagellates
FUNG	U	fungus
INUN	X	indeterminables and unknowns
NOPO	-	nonpollen
PREQ	Z	pre-Quaternary types
RHIZ	R	rhizopods
SPOR	S	taxonomically undifferentiated spores
TRSH	A	trees and shrubs
UPHE	B	upland herbs
VACR	F	vascular cryptogams (pteriodphytes)

Table CHRON

This research table registers each chronology developed for an entity. Not all entities have a developed chronology. Of 909 entities, 772 have 1 or more developed chronologies.

a) There are 1028 registered chronologies for the 772 entities. Therefore, 772 entities have at least one chronology, 191 entities have at least 2 chronologies, 60 entities have at least 3 chronologies and 5 entities have 4 chronologies.
b) 768 of 772 have a determined default chronology (most accepted) (ie. default="y") while 260 have not. Therefore, 4 entities have no default chronologies (E#, 137,385,873,1333)

E.	Site.	SiteName	PolDiv1	Default	Chron.
137	137	Lake of the Clouds [Minnesota]	USA	N	1
137	137	Lake of the Clouds [Minnesota]	USA	N	2
137	137	Lake of the Clouds [Minnesota]	USA	N	3
385	149	Lake Louise	USA	N	1
873	808	Tuskegee Pond	USA	N	1
1333	1213	Antelope Playa	USA	N	1

c) 931 of 1028 model are linear interpolations, 96 are other models such as polynomial regression and spline fittings leaving one null model.
d) The default chronology or preferred one has 693 linear interpolation models and 74 other models such as cubic splines and polynomial regression fits and 1 null model.
e) 189 preferred chronologies were prepared by COHMAP, 488 by members of NAPD-1, 91 by others.

Table AGEBASIS

This research table stores the data used to derive the age-depth curve for each chronology created for an entity. (see p_agedepth).

a) The chronologies are not all based on C14 dates but derived from other sources of dating information (23 rationales, see table RATIONAL above)

E.	Chron.	Sample.	DepthCM	Thickness	Age	AgeUp	AgeLo	RCode
2	1	1	10		0	30	-30	TOP
2	1	2	40		100	150	50	ESH
2	1	3	425		6090	6175	6005	C14
3	1	1	0	0	2000	2500	1500	TOP
3	1	2	50	20	4055	4175	3935	C14
3	1	3	160	20	5945	6065	5825	C14
3	1	4	230	20	5985	6125	5845	C14
3	2	1	50	20	4055	4295	3815	C14
3	2	2	195	90	5962	6146	5778	CAV

Table P_AGEDPT

This research table stores fitted age estimates for the sample depths for each chronology specified for an entity. Used in conjunction with table CHRON to determine the favored chronology. Applying the age-depth model to the data in the AGEBASIS table has generated these estimates.

E.	Chron.	Sample.	AgeBP

2	1	1	0
2	1	2	66
2	1	3	255
2	1	4	411
2	1	5	566
2	1	6	877
2	1	7	1189
2	1	8	1344
2	1	9	1655
2	1	10	1967
2	1	11	2278
2	1	12	2589
2	1	13	2900
2	1	14	3211
2	1	15	3367
2	1	16	3678
2	1	17	3989
2	1	18	4300
2	1	19	4611
2	1	20	4923
2	1	21	5234
2	1	22	5545
2	1	23	5856
2	1	24	6012
2	1	25	6090
2	1	26	6167
2	1	27	6245
2	1	28	6323
2	1	29	6478
3	1	1	2000
3	1	2	2410
3	1	3	2821
3	1	4	3232
3	1	5	3643
3	1	6	4055
3	1	7	4226
3	1	8	4398
3	1	9	4570
3	1	10	4742
3	1	11	4914
3	1	12	5085
3	1	13	5257
3	1	14	5429
3	1	15	5601

3	1	16	5773
3	1	17	5945
3	1	18	5950
3	1	19	5954
3	1	20	5967
3	1	21	5973
3	1	22	5979
3	1	23	5985
3	1	24	5990
3	1	25	6000
3	1	26	6002
3	2	1	3397
3	2	2	3529
3	2	3	3660
3	2	4	3792
3	2	5	3923
3	2	6	4055
3	2	7	4187
3	2	8	4318
3	2	9	4450
3	2	10	4581
3	2	11	4713
3	2	12	4844
3	2	13	4976
3	2	14	5107
3	2	15	5239
3	2	16	5370
3	2	17	5502
3	2	18	5633
3	2	19	5725
3	2	20	6028
3	2	21	6159
3	2	22	6291
3	2	23	6422
3	2	24	6554
3	2	25	6685
3	2	26	6817

As we can observe, there are multiple chronologies based on tables AGEBASIS and CHRON.

Table AGEBOUND

This research table stores a best estimate for top and bottom ages for an entity. This table provides a simple mechanism for evaluating extrapolated dates in P_AGEDPT table.

There 751 entities with stored best top/bottom estimates. Therefore, this table serves primarily for chronological control.

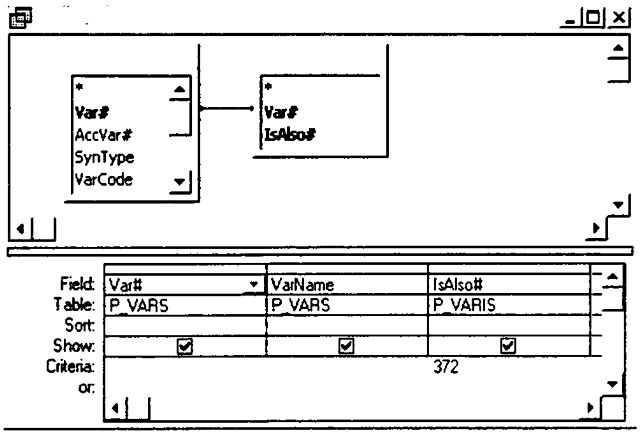
Table P_VTRANS

This research table stores variable names in alternate or preferred style to that stored in P_VARS table.

There are pesntly no stored information in this table.

Table P_VARIS

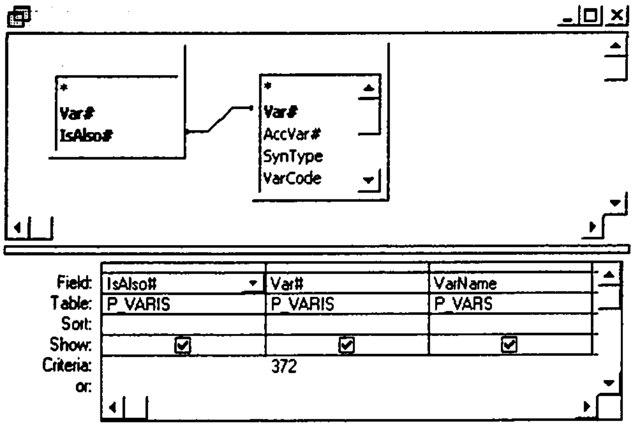
This view table stores the taxonomic hierarchy of variables on the NAPD. For example, this table can be used to query at any level in the hierarchy (eg. *Juglans* include *J. cinera, nigra* ......)



Var.	VarName	IsAlso.
144	Juglans cinerea	372
145	Juglans nigra	372
146	Juglans undiff.	372
372	Juglans	372

801	Juglans cinerea-type	372
802	Juglans nigra-type	372
1441	Juglans cf. J. cinerea	372
1442	Juglans cf. J. nigra	372
2664	Juglans regia	372
3135	Juglans australis	372

The above example query extracts all hierarchical variables under *Juglans.*

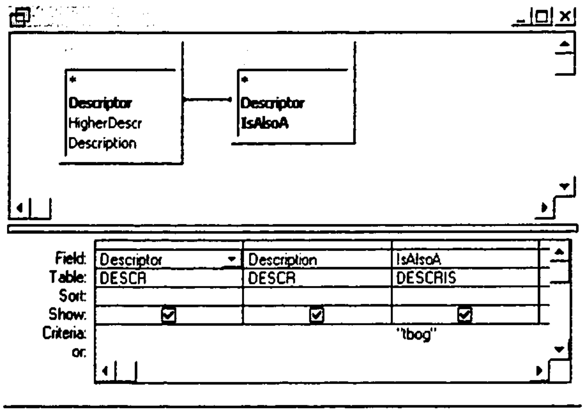


IsAlso.	Var.	VarName
372	372	Juglans
1710	372	Juglandaceae
2132	372	Magnoliopsida
2133	372	Magnoliophyta

The above example extract upward in the hierarchy, in this case for *Juglans.*

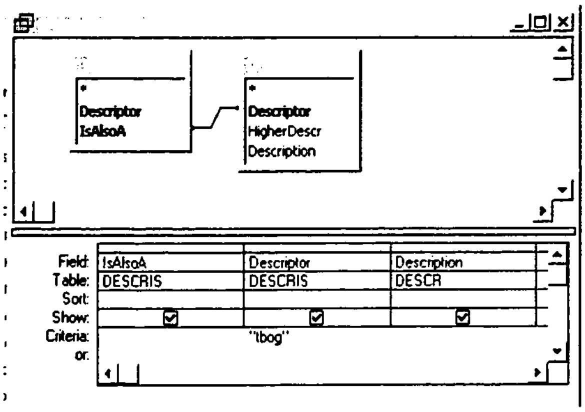
Table DESCRIS

This view table stores the hierarchy of modern entity site descriptions. Used in conjunction with table DESC. This view table is identical in terms of query to that of P-VARIS table.



Descriptor	Description	IsAlsoA
TBLA	blanket bog	TBOG
TBOG	bog	TBOG
TRAI	raised bog	TBOG

The above example extract downward in the hierarchy for "BOGS"



Descriptor	IsAlsoA	Description
TBOG	TBOG	bog
TBOG	TERR	terrestrial
TBOG	TMIR	mire

The above example extract upward hierarchical information for "BOGS".

APPENDIX C – SAS PROGRAMS FOR MIXTURE MODEL

This appendix provides the source SAS (Statistical Analysis Software) code.

Perturbation.sas

```
/* This program creates random perturbation dataset to test null hypothesis on
/* original distribution /*
/* Developped by A. E. Viau and P. Fines

options nodate nolabel ps=200 ls=100;
data pertur;
set fines.dates;
* Ajout d'une perturbation uniforme;
if sdpos ne . then do;
maxim=min(sdpos,200+(200/12000)*date);
datepert=date - maxim + ranuni(234559)*2*maxim;
end;
else datepert=date;
datepert=round(datepert,1);
ecart=datepert-date;
ecartabs=abs(datepert-date);
run;
data sel;
set fines.dates;
if sdpos > 200+(200/12000)*date then put date sdpos;
run;

proc plot data=pertur;
plot ecart*date;
plot ecart*maxim;
plot ecart*sdpos;
plot ecartabs*date;
plot ecartabs*maxim;
plot ecartabs*sdpos;
run;

data _null_;
file "c:\users\viau\viaup800\fines\pertucgu1.csv";

set pertur(rename=(datepert=date) keep=datepert) end=last;
        put date "," ;
run;
```

MixtureModel.sas

```
/* This sas application executes a mixture model on C14 dates extracted from the NAPD
/* It uses the NLP (non-linear program) algortim (ie. expectation maximizing or EM approach) to
/* solve the distribution of c14 dates. Basically it fits normal distributions to the dataset in order to determine and maximize
/*optimal parameterization (ie. optimal number of distribution; std constraints) and solution (ie. modes) */
/* Developped by: D.E Atkinson, A.E. Viau, P.Fines and SAS support network 1999-2000 */


data sel;
 set fines.dates;
  run;
data toto;
 date=0;
 output;
 date=12000;
 output;
run;
data new;
 set sel toto;
run;
************** ATTENTION: FERMER LA FENETRE DES GRAPHIQUES
            POUR QUE LE KILL PUISSE S'EXECUTER;

************* AJUSTER LES OPTIONS DU GPLOT POUR BWIDE >200, ETC;
************* VERIFIER SI PLOT+CHART UTILISE LES PERCENT;
************* NORMALISER LES COURBES NORMALES SELON LA VALEUR DE RUN;
************* Somme des courbes du FIT sur un seul histogramme;
************* Analyses avant et APRES FASTCLUS;
************ Bornes minimales sur les SD;
************ Superposition des asymptotes avec Labels;


options nomtrace nomprint;

goptions reset=all target=win device=win
     noprompt norotate noborder gunit=pct cback=white;

/******* Clean out the graphics libraries ************/
  *proc catalog cat=fines.gseg kill; *run;
  *proc catalog cat=work.gseg kill; *run;
/*******************************/

%global mu1 mu2 mu3 mu4 mu5 mu6 mu7 mu8 mu9 mu10 mu11 mu12 mu13 mu14 mu15
     sigma1 sigma2 sigma3 sigma4 sigma5 sigma6 sigma7 sigma8 sigma9 sigma10
     sigma11 sigma12 sigma13 sigma14 sigma15 bwide bmax run bornemax fclus;

options missing=0;

/*********************************************/
/******** MAIN BEGINS ********************/
/*********************************************/

%macro MAIN(file,nomvar,debut=&curvst,fin=&curvend);

  /*%do bwide=300 %to 300 %by 100; */

   %let loopend=%eval(&fin-&debut+1);

   %do bmax=12000 %to 12000 %by 12000;
     /** initialize data set "Assess" **/
     data assess; sum=-100000; fits=0;run;
```

```
%DO run=&debut %TO &fin;
                                        %let indice=%eval(&indice+1);
                                        /* INDICE is a unique identifying number for each plot */

  /*** An annotate data set to generate footnote information ***/
  /** FOOTNT1 is for normal+freq plot, FOOTNT2 is for scree plot **/
  data footnt1;
    length function color $8. text $30.;
    function="label"; color="black"; position='6'; style='swiss';
        xsys='5'; ysys='5'; hsys='5'; size=2.5;
    x=75; y=97; text="bin width: &bwide"; output;
    x=75; y=94; text="SD constrained to: &bmax years"; output;
    x=75; y=91; text="fitted curves: &run"; output;
    x=2 ; y=40; size=3; text="% frequency"; angle=90; output;
    x=40; y=5 ; size=3; text="years BP"; angle=0; output;
  run;
  data footnt2;
    length function color $8. text $30.;
    function="label"; color="black"; position='6'; style='swiss';
        xsys='5'; ysys='5'; hsys='5'; size=2.5;
    x=75; y=97; text="bin width: &bwide"; output;
    x=75; y=94; text="SD constrained to: &bmax years"; output;
  run;

  %mixinit(data=&file,comp=&run,var=&nomvar);

                                              /*proc means data=capture; var loglike; output sum=sum out=nllout; run;
                                              data nllout; set nllout; keep sum; run;*/

                                       /*** This directs output to a text file for S processing ****/
                                              data muout ; set inest; if _n_=1;
                                                     keep
                                                     %do xx=1 %to &run;
                                                            mu&xx sigma&xx p&xx
                                                     %end;
                                                     ;
                                              run;
                                              proc transpose data=muout out=mut; run;
                                              data muus sigmas ps; set mut;
                                                            retain getp getmu getsig;
                                                            if _n_=1 then do;
                                                                          getp = 1;
                                                                          getmu=0;
                                                                          getsig=0;
                                                                          end; put "first getp" getp;
                                                            if getp=1 then output ps;
                                                            if getmu=1 then output muus;
                                                            if getsig=1 then output sigmas;

                                                            if getsig=1 then do; getsig=0; getp=1; goto skipit; end;
                                                            if getmu=1 then do; getmu=0; getsig=1; goto skipit; end;
                                                            if getp=1 then do; getmu=1; getp=0; goto skipit; end;
                                                            drop getp getmu getsig;
                        skipit:                      put "GETP" getp "GETMU" getmu "GETSIG" getsig;
                                              run;

                                              data muus; set muus; drop _name_ col1; mu=col1; cnt=_n_; run;
                                              data sigmas; set sigmas; drop _name_ col1; sigma=col1; cnt=_n_; run;
                                              data ps; set ps; drop _name_ col1; p=col1; cnt=_n_; run;
                                              data finstats; merge muus sigmas ps; by cnt; drop cnt; run;

                                              /*********** rem out after first run */
                                       /*     data _null_;
                                                     file
"c:\users\fines\realfastclusonly\class300\meanclus&run.&bwide.&bmax.&dsname..csv";
```

```
                              set finstats end=last;
                                        put mu "," sigma "," p;
                              run; */
                    /******** End output redirect *********/

proc nlp data=&file inest=inest out=capture outvar=estcap maxit=1000 maxfu=1000; /*tech: Non teste: a essayer*/
   %mixlike(comp=&run,var=&nomvar,bornemax=&bmax);
run;
                         /******** rem out after first run */
                    /*   data testcap; set estcap (keep=_RHS_); if _n_=1; run;

                         data _null_;
                                   file
"c:\users\fines\classfastclusonly\class300\meanclus&run.&bwide.&bmax.&dsname..csv" MOD;
                                   set testcap; put "-9999 ," _RHS_;
                         run; */

                         proc means data=capture; var loglike; output sum=sum out=nllout; run;
                         data nllout; set nllout; keep sum; run;
                    /*** This directs output to a text file for S processing ****/
                         data muout ; set capture; if _n_=1;
                                   keep
                                   %do xx=1 %to &run;
                                             mu&xx sigma&xx
                                   %end;
                                   ;
                         run;

                         proc transpose data=muout out=mut; run;

                         data muus sigmas; set mut;
                                   if mod(_n_,2)=0 then output sigmas;
                                             else output muus;
                         run;


                         data pout ; set capture; if _n_=1;
                                   keep
                                   %do xx=1 %to &run;
                                             p&xx
                                   %end;
                                   ;
                         run;
                         proc transpose data=pout out=pmut; run;

                         data ps; set pmut;
                                   output ps;
                         run;

                         data muus; set muus; drop _name_ col1; mu=col1; cnt=_n_; run;
                         data sigmas; set sigmas; drop _name_ col1; sigma=col1; cnt=_n_; run;
                         data ps; set ps; drop _name_ col1; p=col1; cnt=_n_; run;
                         data finstats; merge muus sigmas ps ; by cnt; drop cnt; run;

                         data _null_;
                                   file
"c:\users\viau\viaup800\fines\test\meannlp&run.&bwide.&bmax.&dsname..csv";
                                   set finstats end=last;
                                             put mu "," sigma "," p ;

                                   if last then do;
                                             set nllout;
                                             put "-9999 ," sum;
                                   end;
                         run;
```

```
                    /******** End output redirect *********/
/******** These lines build up a table of fit attempts (2 curves, 3 curves, etc)
    and a display of their goodnes of fit
**/
proc means data=capture; var loglike; output out=capsum sum=sum; run;
data capsum; set capsum; fits=&run; run;

                    %if &fclus=YES %then %do;
                              %nrquote(data assess; set assess pseudo; run;);
                         %end;
                    %else %do;
                              %nrquote(data assess; set assess capsum(keep=sum fits); run;);
                         %end;
/******** Dump parameters MU and SIGMA to Macro variable
     space for the plotting program to grab.
******/

%do out = 1 %to &run; %put BUILD DS CAPTURE &out &loopend;
  data _null_;
                         %if &fclus=YES %then %do;
                                   %nrquote(set inest;);
                              %end;

                         %else %do;
                                   %nrquote(set capture;);
                              %end;
    call symput('mu'||left(put(&out,4.)),int(100*mu&out));
    call symput('sigma'||left(put(&out,4.)),int(100*sigma&out));
    call symput('par'||left(put(&out,4.)),int(100*p&out));
  run;
%end; /** out **/

/*** GENERATE NORMAL CURVE FITS TO PLOT AGAINST HISTOGRAM ***/
  data normcurv;
    %do out=1 %to &run;
      %put NORMAL CURVE GENERATE;
      %ncurv(&out,&&mu&out,&&sigma&out,&&par&out)
    %end; /** out **/
  run;

  %let start=&debut;
  %PLOTTER(m=&run, st=&start)

%END; /* run */

/***** Sort through the results table to find the best occurrence **/
data assess2; set assess; if fits=0 then return; output; run;
proc sort data=assess2; by fits; run;
data _null_; set assess2; if _n_=1 then call symput('bestfit',fits); run;

/**** Plot the best occurrence (include option to plot all occurrences) **/
  axis10
    value= (font=swiss height=3 pct color=black )
    minor=none
    label= (f=swiss color=black height=4 pct '# Fits')
    ;

/** specify axis and tick labels to preserve proper spacing, but color them white
   so they dont appear. Make sure the heights are okay. Note that the ORDER
   option specified here must match that used for final display
   (eg ORDER=(0 to 25 by 2) will not work, because it will stop at 24, which
   will not align properly with an axis specified as 0 to 25 by 5).
**/

  axis11
```

```
            label=(a=90 f=swiss height=4 pct color=black
                                                                   %if &fclus=YES %then %do;
                                                                            'Pseudo-F statistic'
                                                                   %end;
                                                                   %else %do;
                                                                            'NLL totaled'
                                                                   %end;
                                                             )
            value=(font=swiss height=3 pct color=black)
            ;

          proc gplot data=assess2 gout=fines.gseg anno=footnt2;
            plot sum*fits / vaxis=axis11 haxis=axis10;
            run;
          quit;

    %END;  /* of bmax loop */

  /* %END;*/ /*of bwide loop */
 %mend MAIN;
 /**********************************************/
 /******** MAIN ENDS ***********************/
 /**********************************************/


 /**********************************************/
 /******** MIXINIT BEGINS  ****************/
 /**********************************************/

 %macro mixinit(  /* Purpose: get initial values for NLP */
   data=&syslast, /* data set to be clustered */
   comp=3,       /* number of mixture components (clusters) */
   var=?,        /* list of variables */
   equal=0,      /* 0 for unequal within-cluster variances
              1 for equal within-cluster variances */
   outest=inest  /* name of data set to contain initial values */
   );

   proc fastclus data=&data replace=part out=out maxc=&comp outstat=fcgrab;
     var &var;
   run;

   /* Grab the pseudo-F statistic from the fast cluster results; use this stat
                        to generate the scree plot to assess goodness of fit of the curves */
           data pseudo; set fcgrab; if _type_="PSEUDO_F"; sum=over_all; fits=&comp; keep fits sum; run;

   proc means noprint data=out;
     var &var;
     output out=stat;
     class cluster;
   run;
   data tran; set stat(drop=_type_);
     length _name_ char $8;
     retain _type_ 'PARMS';
     retain nobs;
     if cluster=. then do;
       select (_stat_);
         when ('N') nobs=&var;
         %if &equal %then %do;
           when ('STD') do; _name_='sigma'; output; end;
         %end;
         otherwise;
       end;
       delete;
    end;
```

```
      char=left(put(cluster,8.));
      select (_stat_);
        when ('N') do; _name_='p'||char; &var=&var/nobs; end;
        when ('MEAN') _name_='mu'||char;
        %if ^&equal %then %do;
          when ('STD') _name_='sigma'||char;
        %end;
        otherwise delete;
      end;
      output;
    run;
    proc transpose data=tran out=inest(type=est);
      var &var;
      by _type_;
    run;
    proc print data=inest; run;
%mend mixinit;
/*********************************************/
/******** MIXINIT ENDS    ****************/
/*********************************************/


/*********************************************/
/******** MIXLIKE BEGINS ****************/
/*********************************************/

%macro mixlike(  /* Purpose: compute log likelihood for NLP */
  comp=3,       /* number of mixture components (clusters) */
  var=?,        /* list of variables */
  equal=0,      /* 0 for unequal within-cluster variances
                1 for equal within-cluster variances */
  bornemax=500
  );
  max loglike;
  %local i;
  parms
    %do i=1 %to &comp;
      p&i
      mu&i
      %if ^&equal %then sigma&i;
    %end;
    %if &equal %then sigma;
    ;
  bounds
    %do i=1 %to &comp;
      %if ^&equal %then &eps < sigma&i < &bornemax,;
      &eps < p&i
      %if &i<&comp %then ,;
    %end;
    %if &equal %then , &eps < sigma;
    ;
  lincon
    %do i=1 %to &comp;
      p&i
      %if &i<&comp %then +;
    %end;
    =1;
  %local si;
  %if &equal %then %let si=;
  %do i=1 %to &comp;
    %if ^&equal %then %let si=&i;
    nll&i=.5*(((mu&i-&var)/sigma&si)**2+&log2pi)+log(sigma&si);
          %end;
/*  loglikex=int(log(
    %do i=1 %to &comp;
```

```
      p&i * exp(-nll&i)
      %if &i<&comp %then +;
    %end;
    )*1000);
*/
  loglike=log(
    %do i=1 %to &comp;
      p&i * exp(-nll&i)
      %if &i<&comp %then +;
    %end;
    );

/*          %let nllsum = %eval(&nllsum + loglikex); */
            /* put "THIS IS THE LOGLIKE VALUE IN MIXLIKE" loglike; */
%mend mixlike;
/*********************************************/
/******** MIXLIKE ENDS    ****************/
/*********************************************/



/*********************************************/
/******** NCURV BEGINS    ****************/
/*********************************************/
%macro ncurv(which, /* to uniquely identify the variable */
        mn,   /* the mu value for the distribution */
        sd,   /* the sigma value for the distribution */
        par)  /* parameter p */
;
 /** FIRST - FUNCTION to generate a normal curve - used for plotting **/
 %let e=2.71828;
 %let range=%eval(4*&sd);
 %if &range=0 %then
                    %let range=7500;
 %if &sd=0 %then
                    %let sd=7500;
 %let start=%eval(&mn-&range);
 %let end=%eval(&mn+&range);

 %let step=%eval(&range*2/50);
          %put &range &start &end &step;
 %if &step=0 %then &step=1;

 /** fix the values for the dates range */
 %let start=0; %let end=1200000; %let step=1000;

 %do i = &start %to &end %by &step;
           nval&which=(&e**(-1*(((&i-&mn)**2)/(2*&sd**2))))/ (&sd*((2*3.14159)**.5));
           nval&which=nval&which*&par;
   count&which=&i;
   output;
 %end;
%mend ncurv;
/*********************************************/
/******** NCURV ENDS     ****************/
/*********************************************/



/*********************************************/
/******** PLOTTER BEGINS  ****************/
/*********************************************/

%Macro PLOTTER(rn=&run, st=&start);
```

```
data normcurv;
     set normcurv;
   %do xx=1 %to &m;
     count&xx = count&xx /100;
   %end;
run;

  data
   %do xx=1 %to &m;
     n&xx (rename=(count&xx=count) keep=count&xx nval&xx)
   %end;
     ;
   set normcurv;
   %do xx=1 %to &m;
     nval&xx=nval&xx*10000;
     if nval&xx or count&xx then output n&xx;
   %end;
   run;

data normtest;
  merge
    %do xx=1 %to &m;
      n&xx
    %end;
    ; by count;
run;
/** rescale the normals curves to 1 for entire plot **/
data normplot; set normtest(obs=1210);
 /* %do xx=1 %to &m;
    nval&xx = %eval(nval&xx *(100/&m)); /* /4 /&m */ /* Dividing by &m scales all normals

        plots to sum to 100; otherwise EACh

        normal plot sums to 100*/

/*%end; */
    nval=(nval1
    %do xx=2 %to &m;
      +nval&xx
    %end;
    );
  run;
/**************************/
                                   /** Capture NORMCURV to text file for plotting in S **/

                                   data _null_; set normplot; file
"c:\users\viau\viaup800\fines\test\normnlp&run.&bwide.&bmax.&dsname..csv";
                                          put
                                                 %do xx=1 %to &m;
                                                       nval&xx ","
                                                 %end;
                                                 nval
                                                       ;
                                          run;
 /** Use normplot data range to specify the axis scaling */
 proc means data=normplot noprint; var nval; output out=normstat max=max; run;
 data _null_; set normstat; call symput('normmax',max); run;
 data _null_;
    mx=int(symget('normmax')*100);
    if mx<40 then top=1.5; /* Modifier les valeurs */
    else if mx <90 then top=3;
    else if mx <140 then do top=5; topg=1.5; end;**** 5* 3.5;
    else if mx <190 then top=7;
    else if mx < 240 then top=8.5;
    else if mx < 290 then top=10;
```

```
    else top=12;
    ybyval=.5;
            ybyvalg=.5;
    call symput('ytop',top);
    call symput('ybyval',ybyval);
    call symput('ytopg',topg);
    call symput('ybyvalg',ybyvalg);
  run;
                                                    /* Force left axis scaling for the dates dataset */
  axis1  %if &normcont=YES %then %str(order=(0 to &ytopg by &ybyvalg)); /*&ytop by &ybyval)); */
       style=0
       label=none /* (a=90 f=swiss height=2 color='white' 'JUNK') */
       color=white
       value=none
       minor=none
       major=none /* (font=swiss height=2 pct color= white )*/
       origin=(0.5 in,0.5 in);

  axis3  order=(&binstart to &binend by &binlabl)
       value=(font=swiss height=2 pct)
       major=(n=5)
       minor=none
       label=none /* (f=swiss height=4 'Years BP') */
       length=5 in
       origin=(0.5 in, 0.5in)
       offset=(2 pct, 0.75 pct); /* need to offset because the bars
                   take up extra space at each end */

 %let loopend=%eval(&rn-&st+1);

  symbol1 i=j c=black width=2;
  symbol2 i=j c=grey width=1;

 proc gplot data=normplot /* normcurv */ gout=gseg ;
   plot
   %do cnter=1 %to &rn;
     %nrquote(nval&cnter *count=2 )
   %end;
   nval*count=1  / name="fits&indice" vaxis=axis1 haxis=axis3 overlay;
 run; quit;
* 2e passage, pour conserver;

/*
proc gplot data=normcurv gout=fines.gseg ;
   plot
   %do cnter=1 %to &rn;
     %nrquote(nval&cnter *count&cnter )
   %end;
   / name="fits&indice" vaxis=axis1 haxis=axis3 overlay;
 run; quit;
*/

axis5  order=(&binstart to &binend by &bwide)  color=white
     value= none /* (font=swiss height=.5 pct color=black ) */
     minor=none
     label=none /* (f=swiss color=black white height=3 pct 'Value') */
     length=5 in
     origin=(.5 in,.5 in );
     ;

/** specify axis and tick labels to preserve proper spacing, but color them white
     so they dont appear. Make sure the heights are okay. Note that the ORDER
     option specified here must match that used for final display
     (eg ORDER=(0 to 25 by 2) will not work, because it will stop at 24, which
     will not align properly with an axis specified as 0 to 25 by 5).
```

```
 **/
 axis6  order=(0 to &ytop by &ybyval)  /* &ytop by &ybyval) */
       label=none /* (a=90 f=swiss height=4  color=black 'Frequency') */
      value=(font=swiss height=2 color=black)
      origin=(.5 in, .5 in) /*(.9 in,.79 in ) */;

 pattern1 value=solid color=grayDD;
 pattern2 value=solid color=grayBB;
 pattern3 value=solid color=grayDD;

 proc gchart data=fines.dates gout=gseg anno=footnt1;
      vbar  date  /  type=percent  name="freq&indice" nolegend width=3
                     raxis=axis6 maxis=axis5
                     midpoints=&binstart to &binend by &bwide
                     ;
   run; quit;
  /* proc chart data=fines.dates; vbar date/type=percent midpoints=150 to 12150 by 300;
   run;*/
 *2e fois, pour conserver;
 /*
   proc gchart data=fines.dates gout=fines.gseg;
      vbar  date  /  type=percent  name="freq&indice" nolegend width=3
                     raxis=axis6 maxis=axis5
                     midpoints=&binstart to &binend by &bwide
                     ;
   run; quit;
 */
   proc greplay igout=gseg gout=fines.gseg tc=sashelp.templt template=whole nofs;
     treplay 1:freq&indice 1:fits&indice;
   run; quit;

 %MEND plotter;

************* DEBUT *****************;
%let log2pi=1.83787706640934;
%let eps=1e-4;
%let rawfile=c14datesdec8testsort;  /* don't type the extension .csv - SAS automatically adds it*/
%let curvst=4;  /** User sets these  **/
%let curvend=15; /* et non 15 */
%let normcont=YES; /** user set to NO to allow auto scaling */
%let binstart=0;
%let binend=12000;
%let binlabl=2000;  /* this controls the appearance of the axis */
%let indice=0;
%let bwide=000;
%let fclus=NO;  /* Set to NO to use NLP results, YES to use FASTCLUS results */
%let dsname=A;

%let nllsum=0;

libname fines "c:\users\viau\viaup800\fines\";
/*
 data fines.dates;
   infile "c:\users\viau\viaup800\fines\c14dec82000.csv" dsd;
   input date SDpos SDneg;
   if _n_<100 then list;
   if _n_ > 1 and date > -100 and date < 12001 then output;
run;
*/
/*data fines.dates;
   infile "c:\users\fines\&rawfile..csv" dsd;
   input date;
   if _n_<100 then list;
   output;
run; */
```

```
%MAIN(fines.dates,date,debut=&curvst,fin=&curvend)

/** test area
   proc means data=n1; var nval1; output sum=sum; run;

   data n1 (rename=(count1=count) keep=count1 nval1)
        n2 (rename=(count2=count) keep=count2 nval2)
        n3 (rename=(count3=count) keep=count3 nval3)
        n4 (rename=(count4=count) keep=count4 nval4)
        ;
    set normcurv;
    nval1=nval1*10000; nval2=nval2*10000; nval3=nval3*10000; nval4=nval4*10000;
    if nval1 or count1 then output n1;
    if nval2 or count2 then output n2;
    if nval3 or count3 then output n3;
    if nval4 or count4 then output n4;
    run;
  data normtest;
    merge n1 n2 n3 n4; by count; run;

  *remember to deal with scaling issues (eg all curves must total to 1);
  data normplot; set normtest(obs=121); nval=(nval1+nval2+nval3+nval4);  run;

symbol1 i=j c=black width=2;
symbol2 i=j c=grey width=1;
proc gplot data=normplot;
  plot nval*count=1 nval1*count=2 nval2*count=2 nval3*count=2 nval4*count=2 /overlay;
  run; quit;

proc gplot data=normplot;
  plot  ;
  run; quit;

%let m=7;
  data pseudo; set fcgrab; if _type_="PSEUDO_F"; curve=&m; keep curve over_all; run;

  proc sort data=capture out=captest; by date; run;
  proc gplot data=captest; plot loglike*date nll1*date nll2*date nll3*date
nll4*date nll5*date nll6*date/overlay; run;

**/
```

APPENDIX D – SAS PROGRAMS FOR PCA

PCAmainmerge.sas

```
/* This program is the main application that perform a Principal Component Analysis on NAPD pollen dataset */
/* in a multidimensional approach ; time, space and taxonomy                                                */
/* This module is used to set parameterization for PCA run.
/* Developped by: D.E. Atkinson and A.E. Viau (2001) */

DATA _null_; file print; put 'Rendu ici !!!';file log; put 'Rendu ici !!!';run;


options sasautos="c:\users\VIAU\PhD PCA analysis\auto\" missing=M fullstimer;

libname oldmaps v6 "c:\program files\sas institute\sas\v8\maps\OldWorldMap\";
libname plots "c:\users\viau\PhD PCA analysis\sasplots\";

%global startyr endyr minobs factend MAX MIN DSTODO FINALFACTOR droplist
        maxx minx maxy miny keeplist totlstns etotal laste bootkeep bootmerg
        drop0 drop1 drop2 drop3 drop4 drop5 drop6 drop7 drop8 drop9
        drop10 drop11 drop12 drop13 drop14 drop15 drop16 drop17 drop18 drop19
        drop20 drop21 drop22 drop23 drop24 drop25 drop26 drop27 drop28 drop29
        drop30 drop31 drop32 drop33 drop34 drop35 drop36 drop37 drop38 drop39
        drop40 drop41 drop42 drop43 drop44 drop45 drop46 drop47 drop48 drop49
        drop50 drop51 drop52 drop53 drop54 drop55 drop56 drop57 drop58 drop59
        drop60 drop61 drop62 drop63 drop64 drop65 drop66 drop67 drop68 drop69
        drop70 drop71 drop72 drop73 drop74 drop75 drop76 drop77 drop78 drop79
        drop80 drop81 drop82 drop83 drop84 drop85 drop86 drop87 drop88 drop89
        drop90 drop91 drop92 drop93 drop94 drop95 drop96 drop97 drop98 drop99
        keep0 keep1 keep2 keep3 keep4 keep5 keep6 keep7 keep8 keep9
        keep10 keep11 keep12 keep13 keep14 keep15 keep16 keep17 keep18 keep19
        keep20 keep21 keep22 keep23 keep24 keep25 keep26 keep27 keep28 keep29
        keep30 keep31 keep32 keep33 keep34 keep35 keep36 keep37 keep38 keep39
        keep40 keep41 keep42 keep43 keep44 keep45 keep46 keep47 keep48 keep49
        keep50 keep51 keep52 keep53 keep54 keep55 keep56 keep57 keep58 keep59
        keep60 keep61 keep62 keep63 keep64 keep65 keep66 keep67 keep68 keep69
        keep70 keep71 keep72 keep73 keep74 keep75 keep76 keep77 keep78 keep79
        keep80 keep81 keep82 keep83 keep84 keep85 keep86 keep87 keep88 keep89
        keep90 keep91 keep92 keep93 keep94 keep95 keep96 keep97 keep98 keep99
        kep0 kep1 kep2 kep3 kep4 kep5 kep6 kep7 kep8 kep9
        kep10 kep11 kep12 kep13 kep14 kep15 kep16 kep17 kep18 kep19
        kep20 kep21 kep22 kep23 kep24 kep25 kep26 kep27 kep28 kep29
        kep30 kep31 kep32 kep33 kep34 kep35 kep36 kep37 kep38 kep39
        kep40 kep41 kep42 kep43 kep44 kep45 kep46 kep47 kep48 kep49
        kep50 kep51 kep52 kep53 kep54 kep55 kep56 kep57 kep58 kep59
        kep60 kep61 kep62 kep63 kep64 kep65 kep66 kep67 kep68 kep69
        kep70 kep71 kep72 kep73 kep74 kep75 kep76 kep77 kep78 kep79
        kep80 kep81 kep82 kep83 kep84 kep85 kep86 kep87 kep88 kep89
        kep90 kep91 kep92 kep93 kep94 kep95 kep96 kep97 kep98 kep99;


%let d2r=atan(1)/45;


/*_____________ PRE-CLEAN ____________________*/

proc catalog cat=plots.plots kill ; run; quit;
proc datasets lib=work kill; run; quit;
```

```
/*____________ USER OPTIONS ____________________*/

/* Set date range (years BP) */
   %let startyr=0;
   %let endyr =12000;

          /* Set minimum number of obs reqd before proceeding
                    e.g. if minobs=20 then the date range must possess
                    at least 20 obs NOTE that this must be set in conjuction
                    with the date range, eg if minobs=20 but the date range
                    is only 1000-3000 years BP then NO data series will be
                    selected because they can never have 20 obs.*/
  %let minobs=48;

  /* Set variable start and end point, e.g. v1=1 and v2=89 is the entire set,
                    v1=10 and v2=35 are variables 10 through 35 */
          %let v1=1;
          %let v2=1;

          /* Set factor number endpoint; eg factend=5 means five factors will be
                    processed FOR THE FIRST RUN ONLY */
          %let factend=20;

          /* Set which major factor group gets processed; eg DSTODO=2 will
                    extract the scores for factor 2 from each site */
          %let DSTODO=1;
          /* If this is the second factor analysis remember to manually
                    skip the first part of the call, that is:
                                   %bigwrit
                                   %reader
                                   %mainds
                                   %plot1
                    */


          /* Set how many factor score series are retained for each site
                    eg if FINALFACTOR=5 then the first five series for each
                    site will be extracted. */
          %let FINALFACTOR=20;

          /* Remember to export the Data file WORK.SCREE1
                    using File/Export from the main menu in order to compute
                    spatial patterns in eigen structure and explained variance
                    patterns. */


          /* Type the path of the input data files. There will be a data file and
                    a location file. They should take the following format:
                    Data file: Comma Delimited with the following variable sequence:
                              e time var1-varN, where
                                        E is a unique site identifier
                                        TIME is an observation time stamp
                                        var1 - varN are the data series from one to N variables

                    Location file: Comma Delimited with the following variable sequence:
                              stnid long lat, where
                                        stnid is a unique site identifier (correponds to E in the Datafile)
                                        long is longitude (in Decimal Degrees)
                                        lat is latitude (in Decimal Degrees)
                    */

          %let DATAPATH=c:\users\viau\PhD PCA Analysis\RawData\;
```

```
%let RAWDATA=pcarealallapr19reduce7picea.csv;
%let LOCDATA=pcarealallapr19loc.csv;                    /*PCAPollenc14250yrint2test2b.csv;*/

%let fileout=SCN;
        /*
                FIL= output to a postscript file
                SCN= output to the screen
                PRT= output directly to the WINPRTG printer
        */
%let page=1; /* Set number of pages for output  - user must calculate
                                One page for every three components in the final
                                Loading Map/Component Score set plots
                                i.e. 1-3 components, 1 page
                                        4-6 components, 2 pages, etc.
                        */

%let maxe=130; /* Set maximum number of E sites */

/* These parameters control CUTPLOT */
        %let factor=COL1; /* Pick the factor to process FOR LOADINGS or loadplot, where COL1=Factor1, etc.
     need to chanfe f=2 for proper labelling see just before macro cmapmerg*/
        %let fthresh=0.4 ; /* Select a threshold for loading values (this is + and - ) */
        %let varsel=2; /* which variable to select */


/* This is the number of variables a given DROPLIST is to contain*/
        %let dropint=2000;

option nomprint;
/*************** Start running *************/
%bigmergl /*LIA run*/                          /*%bigmerg*/

/* This section adds the bootstrap pre-processing data
to make sure the PCA is stable (obs>variables) */

%let etotal=7; /* This is the number of sites (and not the last e-number),
                                                        determined using ACESS */

%macro pert;
  %do k=1 %to &laste;
/*          %let GO=0;
            data _null_; set ecodes; if &k=curre then call symput('GO',1); run;
            %if &GO=1 %then %goto skipv;*/

                    /* IF current e number is not present then skip past the V do loop
                            GOTO skipV */
                              %do var=1 %to &v2;
                              e&k.v&var=e&k.v&var-.01 + ranuni(123129)*.02;
                              if e&k.v&var le 0.0 then e&k.v& var = abs(e&k.v&var);
                              %end;
/* %SKIPV: %let junk=1; */
  %end;
%mend;
        proc transpose data=newviau out=vtt(keep=_name_);run;

        proc means data=newviau; output out=nvmean mean=; run;
        proc transpose data=nvmean(drop= _TYPE_ _FREQ_) out=nvmeant; run;


data dropDS; set nvmeant end=last;
                        length dlist $ 32500;
                        if _n_=1 then do;
                                dlist='';
                                tally=0;
                        end;
```

```
                              drop COL1 dlist totdropl tally;
                              retain dlist tally;
                              if COL1=0 then do;
                                        dlist=trim(left(dlist))||" "||trim(left(_NAME_));
                                        tally=tally+1;
                                        output;
                              end;
                              if last then do;
                                        totdropl=int(tally/&dropint);
                                        call symput('totdropl',trim(left(totdropl)));
                              end;
run;
options mprint;
%macro DSdrops;
          %do count=0 %to &totdropl;
                    data drop&count;
                              length dlist&count $ 32500;
                              retain dlist&count;
                              set dropDS (
                                                  firstobs=
                                                            %eval(&count*&dropint+1)
                                                  obs=
                                                            %eval(&dropint + &count*&dropint)
                                                  ) end=last;

                              if _n_= 1 /*%eval(&count*&dropint+1)*/ then do;
                                                  dlist&count='';
                              end;
                              dlist&count=trim(left(dlist&count))||" "||trim(left(_NAME_));
                              if last then call symput('drop'||trim(left(&count)),trim(left(dlist&count)));
                    run;
          %end;
%mend;
%dsdrops


/* Earlier form of droplist creation */
/*
          data _null_; set nvmeant end=last;
                              if _n_=1 then tally=0;
                              drop tally;
                              length dlist $ 32500;
                              if _n_=1 then dlist='';
                              retain dlist TALLY;
                              if COL1=0 then do;
                                        dlist=trim(left(dlist))||" "||trim(left(_NAME_));
                                        tally=tally+1;
                              end;
                              if last then do;
                                        call symput('dlist',trim(left(dlist)));
                                        put ">>>>>>>>>>>>>>>>>>>>>>>> NUMBER OF DROPS: " tally;
                              end;
          run;
*/
          proc transpose data=vtt out=vttt(drop=_NAME_ ); run;
%macro dropVTT;
          %do count=0 %to &totdropl;
                    data vttt; set vttt(drop=&&drop&count); run;
          %end;
%mend dropVTT;
%dropVTT
          proc transpose data=vttt out=vtt; run;
          /* Moved the following DATA step out from under the VTT above */
                                        data ecodes; set vtt end=last ;
                                                            keep currc;
```

```
                                test=substr(_name_,2,4);
                                testlen=length(test); put test;
                                        do lookat=1 to testlen;
                                                lookchar=substr(test,lookat,1);
                                                if lookchar not in ('0','1','2','3','4','5','6','7','8','9') then
goto out1;
                                        end;
                                out1: lastelen=lookat-1;
                                curre=substr(_name_,2,lastelen); put curre;
                                call symput('curre',substr(_name_,2,lastelen));


                        if last then do;
                                test=substr(_name_,2,4);
                                testlen=length(test);
                                        do lookat=1 to testlen;
                                                lookchar=substr(test,lookat,1);
                                                if lookchar not in ('0','1','2','3','4','5','6','7','8','9') then
goto out2;
                                        end;
                                out2: lastelen=lookat-1;
                                call symput('laste',substr(_name_,2,lastelen));
                        end;
                run; %put &laste;

%put &dlist;
/** Apply all the droplists to create the final data set **/

/*&totdropl;*/
%macro dropimp;
        %do count=0 %to &totdropl;
                data newviau; set newviau(drop=&&drop&count); run;
        %end;
%mend dropimp;
%dropimp

        /* The next 4 data steps determine row and column counts for the orignal data set
                and then divide them to get the number of bootstrap permutations required */
        data nvcount(keep= rcount);
                set newviau end=last;
                if last then do;
                        rcount=_n_;
                        output;
                end;
        run;

        data vttcount(drop=_name_); set vtt end=last; if last then do;
                        ccount=_n_;output;
                        end;
                run;
        data counts; merge nvcount vttcount; run;
        data _null_; set counts; perms=int(ccount/rcount); call symput('permcnt',trim(left(perms)));run;

        data toto;
           set newviau;
                output;
                do numpert=1 to &permcnt;
                 %pert;
                 output;
           end;
        run;

/* This and DSKEEP build up a series of KEEPlists for BOOTSTRAP */
data _NULL_; set vtt end=last;
                        if _n_=1 then do;
```

```
                    tallyB=0;
                end;
                drop tallyB;
                retain tallyB;
                tallyb=tallyB+1; put tallyb;
    if last then do;
                    totkeepl=int(tallyB/&dropint);
                    call symput('totkeepl',trim(left(totkeepl)));
    end;
run;

%macro DSkeep;
    %do count=0 %to &totkeepl;
        data keep&count;
                length klist&count $ 30000;
                retain klist&count;
                set vtt (
                        firstobs=
                            %eval(&count*&dropint+1)
                        obs=
                            %eval(&dropint + &count*&dropint)
                        ) end=last;

                if _n_= 1 /*%eval(&count*&dropint+1)*/ then do;
                        klist&count='';
                end;
                klist&count=trim(left(klist&count))||" "||trim(left(_NAME_));
                if last then call symput('keep'||trim(left(&count)),trim(left(klist&count)));
        run;
    %end;
%mend;
%dskeep



/* Original implemetation of the BOOTSTRAP keep list */
/*
    data _null_;
        length bootkeep $ 32500;
        set vtt end=last;
        retain bootkeep;
        if _n_=1 then bootkeep='';
        bootkeep=trim(left(bootkeep))||" "||trim(left(_NAME_));
        if last then call symput('bootkeep',left(trim(bootkeep)));
    run;
%put &bootkeep;
*/

/** Apply all the keeplists to create the final data set **/

%macro keepimp;
    %do count=0 %to &totkeepl;
        %if &count=0 %then %let bootmerg='';
        data boot&count; set toto(keep=&&keep&count); run;
        %let bootmerg=&bootmerg boot&count;
    %end;
    data bootstra; merge &bootmerg; run;
%mend keepimp;
%keepimp %put &bootmerg;
/* JUST IN CASE IT NEEDS SORTING */
/*
proc transpose data=bootstra out=boottran; run;
proc sort data=bootsort out=bootsort; by _NAME_; run;
proc transpose data=bootsort out=bootstra(drop=_NAME_); run;
```

```
*/
/* THIS IS THE OLD KEEPLIST FOR BOOTSTRA
        data bootstra; set toto(keep=&bootkeep);run;
*/
/*
        data bootstra;
          set toto(drop=e1v3 e1v12 e1v14 e1v17 e2v17 e7v3 numpert);
        run;
*/
/* END Bootstrap pre-prep section*/
data plots.bootstra; set bootstra; run;
data plots.newviau; set newviau; run;

/*Call PCA module*/

%pca2mergabcd(bootstra,noprint) /* or all */

/*******************************************************/
/*******************************************************/
/*    BEGIN OF FACTOR PREP SECTION          */
/*******************************************************/
/*******************************************************/


/** This first section verifies the Communality estimates PROC FACTOR
provides - Factor loadings are squared and summed. These results
match what SAS PROC FACTOR generates. **/

data pull; set pca1;
        if _TYPE_="PATTERN";
        drop _TYPE_ _NAME_;
run;

/* This squares the matrix */
proc iml;
        use pull;
         read all var _num_ into aaa;
                        asq=aaa##2;
                        create pull2 from asq;
                        append from asq;
run; quit;

proc summary data=pull2; output out=summpull
        sum=;
        var _all_;
        run;

/* This second section performs a communality estimate by Factor, and
not by variable, and then takes the percentage contribution of each
variable to the overall Factor.
Ok, so as it turns out, these are Eigenvalues! And ours calculated
here match SAS PROC factor. */

proc transpose data=pull out=pullt; run;

/* This squares the matrix */
proc iml;
        use pullt;
         read all var _num_ into aaa;
                        asq=aaa##2;
                        create pullt2 from asq;
                        append from asq;
run; quit;

proc summary data=pullt2; output out=sumpullt(drop=_TYPE_ _FREQ_)
```

```
            sum=;
            var _all_;
            run;

/* Now merge the E (that is site) codes back onto the data and
 process by site code*/

data pullsite; merge  ecodes pullt2; run;

data pullsite; set pullsite; curre=curre*1; run;

proc sort data=pullsite; by curre; run;

proc transpose data=pullsite out=pullabs(keep=_NAME_); run;
data _null_; set pullabs end=last;
            retain lablist;
            length lablist $ 32500;
            if _n_=1 then lablist="";
            lablist=trim(left(lablist))||" "||trim(left(_NAME_));
            if last then call symput('lablist',trim(left(lablist)));
run;

/* This gives a raw proportion value for each site by factor */
proc summary data=pullsite;
            output out=sumsite(drop=_FREQ_ _TYPE_) sum=;
            by curre;
            var &lablist;
run;

/* This gives a proportion that each site contributes to a given factor */
data sitetemp; set sumsite(keep=curre); run;
data sumsite; set sumsite(drop=curre); run;

proc iml;
          use sumsite;
           read all var _num_ into aaa;
                                        use sumpullt;
                                        read all var _num_ into eigenval;
                                        totcols=ncol(eigenval);
                                        totrows=nrow(aaa);
                                        result=j(nrow(aaa),ncol(aaa),0);
                                        /* Two ways to get a proportion by factor: first is a raw
                                                proportion, which will result in smaller values with
                                                increasing number of sites, the second scales the
                                                proportions by number of sites to help with contour plotting */
                                        do x=1 to nrow(aaa);
                                                result[x,]=aaa[x,]/eigenval;
/*                                              result[x,]=nrow(aaa)*(aaa[x,]/eigenval); */
                                        end;

                                        create sitecont from result;
                                        append from result;
run; quit;

data sitecont; merge sitetemp sitecont; run;

/*******************************************************/
/*******************************************************/
/*      END OF FACTOR PREP SECTION            */
/*******************************************************/
/*******************************************************/


%cutplot
```

```
/* Remake data set LOADPLOT using sitecont and the ecodes lat/long (and SUM) */

data elocs; drop stnid; set subloc; length curre $ 7; curre=left(substr(stnid,2)); run;
data sitecont; set sitecont; curre=left(curre); run;
proc sort data=elocs out=elocs; by curre; run;
proc sort data=sitecont out=sitecont; by curre; run;


data loadplot;
merge elocs
    sitecont(rename=( &factor=sum) keep=curre &factor in=abcd);
by curre;
if abcd;
run;

            /* Do the plotting and mapping*/

%basemap

%let wichfile=LOADPLOT; /* VARPLOT or LOADPLOT*/
%macro mainloop;
            %if &wichfile=VARPLOT %then %let fstop=10;
            %if &wichfile=LOADPLOT %then %let fstop=1;

%do f=1 %to &fstop;
            %pcamstn
            %mkgridm
            %if &wichfile=VARPLOT %then %do;
                        %contour(fact&f) /*   FACTOR WHEN VAR < 100*/
            %end;
            %if &wichfile=LOADPLOT %then %do;
                        %contour(sum)
            %end;

            %if &wichfile=LOADPLOT %then %do;
                        %let f=1;                                    /* change for desired factor output on LOADPLOTS*/
            %end;
%cmapmerg
%end;
            data _null_; set plotfile;
                        file "c:\users\viau\phd pca analysis\temp\plotfile.csv" ;

                        /*if _n_=1 then put "stnid,fulstnid,Factor1,Factor2,Factor3,Factor4,Factor5,varid,percent,long,lat";
                        put stnid "," fulstnid "," Factor1 "," Factor2 "," Factor3 "," Factor4 "," Factor5 "," varid "," percent "," long "," lat;
                        */
                        if _n_=1 then put "stnid,fulstnid,Fact1,Fact2,Fact3,Fact4,Fact5,Fact6,Fact7,Fact8,Fact9,Fact10,varid,percent,long,lat";
                        put stnid "," fulstnid "," Fact1 "," Fact2 "," Fact3 "," Fact4 "," Fact5 "," Fact6 "," Fact7 "," Fact8 "," Fact9 "," Fact10 ","
varid "," percent "," long "," lat;
run;
            /*%bigwrit*/
/* THE FOLLOWING IS HARD WIRED - WILL NEED TO BE CHANGED */
            /*%do x=1 %to 5;
                        data _null_; set lg&x;
                                    file "c:\users\viau\phd pca analysis\temp\PCAbysiteForFactor&x..csv" ;
                                    if _n_=1 then put "time,e2f&x,e3f&x,e4f&x,e5f&x,e6f&x,e7f&x";
                                    put time"," e2f&x "," e3f&x "," e4f&x "," e5f&x "," e6f&x "," e7f&x;
                        run;
            %end;*/

%mend mainloop;

%mainloop
```

Bigmergl.Sas

```
/* This module prepares the dataset for PCA analysis */
/* It is called by Pcamainmerg.sas main program        */

%macro bigmergl;

/* A  version that merges E and Var (create a unique identifier
            for all vars, eg v1e1 v2e1 v1e2 v2e2 etc) and then run a single PCA by time
            JAN 2002 */
%let dslist=;
/*

*/

%do e=1 %to 125; /*&laste; */
/*
 %let GO=0;
 data _null_; set ecodes; if &e=curre then call symput('GO',1); run;
 %if &GO=1 %then %goto skipv;
*/
          data v&e;
              infile "&DATAPATH&RAWDATA" dsd LRECL=2000; /* Watch for truncation message in LOG file */
              if _n_=1 then input;
                              input e @;
                              drop e;
                              if e="&e" then input time e&e.v&v1 - e&e.v&v2 ;
                              /* put time e&e.v&v1 - e&e.v&v2 ; */
                              if e="&e" then do;
                                                    %let dslist=&dslist v&e;
                                                    output;
                                        end;
          run;
/*        %skipv: %let junk=1;*/
%end;

          data viaux; merge &dslist; by time; run;

          Data basetime;
   DO time = &startyr TO &endyr by 250;
     output;
   end;
  run;

          data viaux2; merge basetime(in=a) viaux(in=b); by time; if a and b;  run;
          proc means data=viaux2 n; output out=viaux2n; run;
                    proc transpose data=viaux2n out=viaux2t(keep= _NAME_ COL1); run;
          data vreduce; set viaux2t; if _n_>2 and COL1>&minobs;run;


          data _NULL_; set vreduce (firstobs=2) end=last ;
                                        if _n_=1 then do;
                                                  tallyC=0;
                                        end;
                                        drop tallyC;
                                        retain tallyC;
                                        tallyC=tallyC+1; put tallyC;
                    if last then do;
                                                  totkeep3=int(tallyC/&dropint);
                                                  call symput('totkeep3',trim(left(totkeep3)));
                    end;
          run;
```

```
        %do count=0 %to &totkeep3;
                data kep3&count;
                        length klist3&count $ 30000;
                        retain klist3&count;
                        set VREDUCE (
                                        firstobs=
                                                %eval(&count*&dropint+1+1)
                                        obs=
                                                %eval(&dropint + &count*&dropint +1 )
                                        ) end=last;

                        if _n_= 1 /*%eval(&count*&dropint+1)*/ then do;
                                        klist3&count='';
                        end;
                        klist3&count=trim(left(klist3&count))||" "||trim(left(_NAME_));
                        if last then call symput('kep3'||trim(left(&count)),trim(left(klist3&count)));
                run;
        %end;

        %do count=0 %to &totkeep3;
                %if &count=0 %then %let KEP3MERG=;
                data KEP3&count; set Viaux2(keep=&&kep3&count); run;
                %let KEP3MERG=&KEP3MERG KEP3&count;
        %end;
        data NEWVIAU; merge &KEP3MERG; run;

/*
        data _null_; length keeplist $ 32500;
                set vreduce end=last; retain keeplist;
                if _n_=1 then keeplist="";
                if _n_>1 then keeplist=trim(keeplist)||" "||trim(_NAME_);
                        call symput('keeplist',keeplist);
                run;

%put &keeplist;
data newviau; set viaux2(keep= &keeplist); run;
*/
/*proc transpose data=viaux2(keep= &keeplist) out=viauxtest; run;*/


                /* %put &keeplist;*/

%mend bigmergl;
```

Pcamergabcd.sas

```
/*********************************************************************************
 This Macro runs the PCA, creating output with factor loadings for mapping and scores for plots
          for the variable-site merged PCA.
*********************************************************************************/
%macro PCA2mergabcd(file,OPT);

          proc factor data=&file/*(keep= &keeplist)*/ /*viaux2 d'habitude */
   /*  ( drop=time e1v3 e1v12 e1v14 e1v17 e2v17 e7v3  obs=20) */
     outstat=pca1 &OPT m=p score mineigen=0.0001
                    sing=1e-12
     ; run;

          proc score data=newviau /*&file*//*(keep= &keeplist)*/
          /*( drop=time e1v3 e1v12 e1v14 e1v17 e2v17 e7v3 )*/
      score=pca1 out=pca1scor; run;

  data score2;
   set pca1;
           length _NAME_ $9.;
   if _TYPE_="PATTERN" then output;
   if _TYPE_="EIGENVAL" then do;
      _NAME_="EIGENVAL"; output;
      end;
   *format _all_ 6.3 -r;
  run;

  proc transpose data=score2 out=score3(rename=(_name_=fulstnid)); run;
          data score3; set score3; drop leng ct junk hold target;
                    /* Extract e and var numbers  */
                              leng=length(fulstnid);
                              retain hold target;
                              hold=0;
                              do ct=1 to leng;
                                        if substr(fulstnid,ct,1)="v" then goto skipout;
                                        hold=hold+1;
                              end;
                    skipout: junk=1;
                              do ct=1 to leng;
                                        if substr(fulstnid,ct,1)="v" then target=ct;
                              end;
                              stnid=substr(fulstnid,1,hold);
                              varid=substr(fulstnid,target);
                    run;

  proc univariate data=score3 noprint; output out=toteigen sum=sum; var EIGENVAL; run;
          proc univariate data=score3 noprint; output out=totlstns n=n; var EIGENVAL; run;
          data _null_;
      set totlstns;
      call symput('totlstns',n);
  run;


          data _null_;
      set toteigen;
      call symput('toteigen',sum);
  run;

            data perc;
      set score3;
      percent=round(eigenval/&toteigen*100,1);
                              /*percent=round(eigenval,1);*/
```

```
                                drop eigenval;
run;

proc sort data=perc; by stnid; run;
data subloc; infile "&DATAPATH&LOCDATA" dsd;
                                        length stnid $ 5;
        input stnid long lat; run;
proc sort data=subloc; by stnid; run;

         data subloc; set subloc; stnid="e"||trim(left(stnid)); run;

data percent; length stnid $ 5; set perc; run;
         proc sort data=percent out=percent; by stnid; run;

data plotfile;
                                merge percent(in=a) subloc(keep=lat long stnid);
         by stnid; if a;
run;

         data time; set basetime; ob=_n_;run;
         data scores; set pca1scor; ob=_n_; run;
         /*data scores1; set scores;
         if mod(ob,5)=1;
         run;*/
         /*data scores2; set scores;
         if mod(ob,5)=2;
         run;
         %do f=1 %to 20; /*&factend;*/
                  /* Factor score plots here */
                  /*data scoreplot(drop=ob); merge time scores(keep=Factor&f); run;*/
                  data scoreplot(drop=ob); merge time scores(keep=Fact&f); run;
                  %plot2(Fact) /* OR factor if the number of variables < 100 */
                           %if &f=1 %then %str(data outscore; set scoreplot; run;);
                           %if &f>1 %then %str(data outscore; merge outscore scoreplot; by time; run;);
         /* data scoreplot(drop=ob); merge time scores2(keep=Fact&f); run;
                  %plot2(Fact) /* OR factor if the number of variables < 100 */
                  /*       %if &f=1 %then %str(data outscore; set scoreplot; run;); */
                  /*       %if &f>1 %then %str(data outscore; merge outscore scoreplot; by time; run;); */
         %end;


         data _null_; set outscore;
                  file "c:\users\viau\phd pca analysis\finalcomparison\OUTSCOREKGLIATEST.csv" lrecl=1000;
         /*if _n_=1 then put
"time,Fact1,Fact2,Fact3,Fact4,Fact5,Fact6,Fact7,Fact8,Fact9,Fact10,Fact11,Fact12,Fact13,Fact14,Fact15,Fact16,Fact17,Fact18,Fact19,Fact20";
                  put time "," factor1 "," factor2 "," factor3 "," factor4 "," factor5"," factor6 "," factor7 "," factor8 "," factor9"," factor10 ","
factor11 "," factor12 "," factor13
"," factor14 "," factor15 "," factor16 "," factor17"," factor18 "," factor19 "," factor20;*/
                  if _n_=1 then put
"time,Fact1,Fact2,Fact3,Fact4,Fact5,Fact6,Fact7,Fact8,Fact9,Fact10,Fact11,Fact12,Fact13,Fact14,Fact15,Fact16,Fact17,Fact18,Fact19,Fact20";
                  put time "," fact1 "," fact2 "," fact3 "," fact4 "," fact5 "," fact6 "," fact7 "," fact8 "," fact9 "," fact10
      "," fact11 "," fact12 "," fact13 "," fact14 "," fact15 "," fact16 "," fact17 "," fact18 "," fact19 "," fact20;
                  /* POSSIBLY INCREASE NUMBER OF FACTORS */
         run;
         /* Recombine everything here */
                  /* PROC GREPLAY etc. */
%mend PCA2merg;
```

Cutplot.sas

```
%macro cutplot;
/* This macro works with the results of the site/var merge
                    - prepare a count of the loadings exceeding a certain threshold
                            value at each site
                    - plot by factor loadings by variable
*/

            /*  Use PLOTFILE for this work */

            /* Count loadings by threshold */
                    /* User (or something) pick which factor and threshold value
                                    set threshold into the main call */
/*
                    Data LoadThres;
                            if _n_=1 then count=0;
                            set plotfile(keep=Fact&factor stnid lat long); */ /*FACTOR WHEN VARIABLE < 100*/
/*                          if Fact&factor > &fthresh OR Fact&factor < (-1*&fthresh) then count=1;*/ /*FACTOR WHEN VARIABLE
<100 */
/*                          else count=0;
                    run;

                    proc univariate data=loadthres noprint;
                            by stnid;
                            var count;
                            output out=xx sum=sum;
                    run;

                    proc sql;
                            create table loadplot as
                                    select distinct loadthres.stnid, xx.stnid, xx.sum, loadthres.lat,
                                                    loadthres.long from work.xx, work.loadthres
                                                    where loadthres.stnid=xx.stnid;
                    quit;
*/
                    data varplot;
                            set plotfile;
                            if varid="v&varsel";
                            keep lat long varid stnid fact1-fact5; /*faCTOR WHEN VARIABLE <100 */
                    run;

%mend cutplot;
```

Plot2.sas

```
/* This macro creates plots for factor scores */

%macro plot2(which);
symbol i=j value=none;
axis1 order=(-5 to 5 by 1);
                    proc gplot data=scoreplot gout=plots.plots;
                        plot &which&f *time / name="Score&f"
                                                                 overlay
                                                                 lv=2
                             vref=(-1 0 1)
                                                                 vaxis=axis1
                                                                 /*
                             vaxis=axis2
                             haxis=axis1
                                                                 */
                                                                 /*
                             %DO x = 1 %to %eval(&HLINECNT);
                               &&hline&x
                             %END;
                                                                 */
                             ;
                        run; quit;
%mend plot2;
```

Basemap.sas

```
/*****************************************************************************************
Macro BASEMAP extracts necessary basemaps from the MAPS library
*****************************************************************************************/
%macro basemap;
 %let pi=3.14159;

   data CANUS;
       set maps.canada4(where=(density < 6))
          maps.states (where=(state ^= 15 and state ^= 72 and density < 5));
       myid=SEGMENT*PROVINCE;
       colors=1;
       if state then province=state;
       xsys='2'; ysys='2';
   run;
   data GRNarea;
     set Oldmaps.worldmap(where=(id = 405));
     province='98';
     myid=SEGMENT*PROVINCE;
     colors=2;
   run;
   data basemap; set CANUS GRNarea; run;

 /* determine map extents so the contour grid can be specified */
 proc univariate data=basemap noprint;
    var x y; output out=mapxn max=maxx maxy min=minx miny;
 run;

 data mapLxn;
    set mapxn;
    maxx=maxx*(360/(2*&pi)); maxy=maxy*(360/(2*&pi)); minx=minx*(360/(2*&pi)); miny=miny*(360/(2*&pi));
    call symput('maxx',maxx);
    call symput('maxy',maxy);
    call symput('minx',minx);
    call symput('miny',miny);
 run;
%mend basemap;
```

Pcamnstn.sas

```
/* This prepares the locations for gridding in the PCA analysis performed
                                        on site/variable merged data. */

%macro PCAmstn;
    data origdot ;
        length function color style $8 text $20;
                                        %if &wichfile=VARPLOT %then %do;
                set varplot end=last;
                                        %end;
                                        %if &wichfile=LOADPLOT %then %do;
         set loadplot end=last;
                                        %end;
         xsys='2'; ysys='2'; hsys='3'; when='a'; function='label'; style='special';
         text='J'; size=2; position='+'; color='black'; province='-2';
         x=long*&d2r; y=lat*&d2r;
         output ;

         ysys='2'; hsys='3'; /*y=y-.003;*/ x=x-.008;
         function='label'; style='SWISSB'; size=2;
         position='6';
         color='black';
                                        %if &wichfile=VARPLOT %then %do;
                                                text=left(trim(int(fact&f*100))); /* FACTOR WHEN VARIABLE <100*/
                                        %end;
                                        %if &wichfile=LOADPLOT %then %do;
                                                text=left(trim(int(sum*100)));
                                        %end;
         province='-1';
         output;

        drop pid long lat fact1 fact2 fact3 fact4; /*factor when varible <100*/
        /** reverse vertical axis for display and store some values for grid setup **/
    run;
%mend PCAmstn;
```

Mkgridm.sas

```
/* This macro preps the grid for contouring in the PCA on
the site/variable merged data */

%macro mkgridm;

  data gridfeed;
                    %if &wichfile=VARPLOT %then %do;
                          set varplot ;
                          keep lat long fact&f; /* factor when varible <100*/
                                       %end;
                                       %if &wichfile=LOADPLOT %then %do;
                     set loadplot ;
                          keep lat long sum;
                                       %end;
  run;

  data _null_;
   set maplxn;
   xgap=(maxx-minx)/-100;
   ygap=(maxy-miny)/-100;    put xgap ygap;
   call symput('xgap',xgap);
   call symput('ygap',ygap);
   gxstart=maxx;
   gxend=minx-xgap;
   gystart=maxy +ygap;
   gyend=miny;
   call symput('gxstart',gxstart);
   call symput('gxend',gxend);
   call symput('gystart',gystart);
   call symput('gyend',gyend);
  run;
      %PUT GXSTART: &GXSTART;
      %PUT GXEND: &GXEND;
      %PUT GYSTART: &GYSTART;
      %PUT GYEND: &GYEND;
%put &wichfile;
          %if &wichfile=VARPLOT %then %do;
                    PROC G3GRID data=gridfeed out=pcagrid&f;
                         grid long*lat=fact&f / /* factor when variable <100*/

                                       spline smooth=.0001
                                       axis1=&gxstart to &gxend by &xgap
                                       axis2=&gystart to &gyend by &ygap;
                    run;
          %end;
          %if &wichfile=LOADPLOT %then %do;
                    PROC G3GRID data=gridfeed out=pcagrid&f;
                         grid long*lat=sum /

                                       spline smooth=.0001
                                       axis1=&gxstart to &gxend by &xgap
                                       axis2=&gystart to &gyend by &ygap;
                    run;
          %end;
%mend mkgridm;
```

Countour.sas

```
/* This macro performs contouring for mapping purposes */

%MACRO contour(var);
/* %let f=1; for testing */

PROC IML;

    rows=100;
    cols=100;
    grid=j(rows,cols,0);
    gridlat=j(rows,cols,0);
    gridlong=j(rows,cols,0);

    start getgrid(grid,gridlat,gridlong,rows,cols);
        use pcagrid&f;
        read all var {&var} into a;
        grid=shape(a,rows);
        read all var {lat} into b;
        gridlat=shape(b,rows);
        read all var {long} into c;
        gridlong=shape(c,rows);
    finish;
run getgrid(grid,gridlat,gridlong,rows,cols);

start MAINBODY(matrix,matlat,matlong,rows,cols);

    LongMin=&gxstart;
    LongMax=&gxend;
    LatMin=&gyend;
    LatMax=&gystart;

    gridlat=matlat;
    gridlong=matlong;
    grid=matrix;
    found1 = 0;
    found2 = 0;

    xinc=&xgap*-1;
    yinc=&ygap*-1;

    X1 = 0;
    X2 = 0;
    Y1 = 0;
    Y2 = 0;
    s1 = 0;
    s2 = 0;
    s3 = 0;
    s4 = 0;
    c1 = 0;
    c2 = 0;
    c3 = 0;
    c4 = 0;
    reverse=0;
    db=0;
    secondRn=0;

                if "&var"="sum" then do;
        cStart = 0;
        cEnd = .05;
        cStep =  .005;
                end;
```

```
                    if "&var"="fact&F" then do; /* FACTOR WHEN VARIABLE <100 */
        cStart = -2;
        cEnd = 2;
        cStep =  .2;
                    end;



    Do contour = cStart To cEnd by cStep;
      Do i = 1 to (rows-1);
        Do j = 1 to (cols-1);
          ULv = grid[i, j];
          URv = grid[i, j + 1];
          LRv = grid[i + 1, j + 1];
          LLv = grid[i + 1, j];
    /*
      'skip check - if all vertices are less than or greater than
      'the contour being checked then skip the checking process
    */
      If contour > ULv & contour > URv & contour > LRv & contour > LLv Then GoTo skipit;
      If contour < ULv & contour < URv & contour < LRv & contour < LLv Then GoTo skipit;
    /*
      'double line check - if diagonal vertices bracket the contour
      'then there are double contour lines.  This situation must
      'be dealt with individually
    */
      If ((ULv < contour) & (LRv < contour) & (URv > contour) & (LLv > contour)) |
        ((ULv > contour) & (LRv > contour) & (URv < contour) & (LLv < contour)) Then do;
          db = 1;
          reverse = 0;
          secondRn = 1;
          End;
      If ((ULv > contour) & (LRv > contour) & (URv < contour) & (LLv < contour)) |
        ((ULv < contour) & (LRv < contour) & (URv > contour) & (LLv > contour)) Then do;
          db = 1;
          reverse = 1;
          secondRn = 1;
          End;

      xpos = gridlong[i,j];
      ypos = gridlat[i,j];

reptRun:
      Do side = 1 To 4;
        If reverse = 1 & db = 1 Then side = 3;

        If found1 = 0 Then do;
          If side = 1 & found1 = 0 Then do;
            If ((ULv < contour) & (URv > contour)) Then do;
              found1 = 1;
              s1 = 1;
              X1 = xpos - (((contour - ULv) / (URv - ULv)) * xinc);
              Y1 = ypos;
            End;
            If ((ULv > contour) & (URv < contour)) Then do;
              found1 = 1;
              s1 = 1;
              X1 = xpos - (((ULv - contour) / (ULv - URv)) * xinc);
              Y1 = ypos;
            End;
          End;

          If side = 2 & found1 = 0 Then do;
            If ((URv < contour) & (LRv > contour)) Then do;
```

```
        found1 = 1;
        s2 = 1;
        X1 = xpos - xinc;
        Y1 = ypos - (((contour - URv) / (LRv - URv)) * yinc);
      End;
      If ((URv > contour) & (LRv < contour)) Then do;
        found1 = 1;
        s2 = 1;
        X1 = xpos - xinc;
        Y1 = ypos - (((URv - contour) / (URv - LRv)) * yinc);
      End;
    End;

    If side = 3 & found1 = 0 Then do;
      If ((LLv < contour) & (LRv > contour)) Then do;
        found1 = 1;
        s3 = 1;
        X1 = xpos - (((contour - LLv) / (LRv - LLv)) * xinc);
        Y1 = ypos - yinc;
      End;
      If ((LLv > contour) & (LRv < contour)) Then do;
        found1 = 1;
        s3 = 1;
        X1 = xpos - (((LLv - contour) / (LLv - LRv)) * xinc);
        Y1 = ypos - yinc;
      End;
    End;

    If side = 4 & found1 = 0 Then do;
      If ((ULv < contour) & (LLv > contour)) Then do;
        found1 = 1;
        s4 = 1;
        X1 = xpos;
        Y1 = ypos - (((contour - ULv) / (LLv - ULv)) * yinc);
      End;
      If ((ULv > contour) & (LLv < contour)) Then do;
        found1 = 1;
        s4 = 1;
        X1 = xpos;
        Y1 = ypos - (((ULv - contour) / (ULv - LLv)) * yinc);
      End;
    End ;

    If contour = ULv & found1 = 0 Then do;
      found1 = 1;
      c1 = 1;
      X1 = xpos;
      Y1 = ypos;
      End;
    If contour = URv & found1 = 0 Then do;
      found1 = 1;
      c2 = 1;
      X1 = xpos - xinc;
      Y1 = ypos;
      End;
    If contour = LLv & found1 = 0 Then do;
      found1 = 1;
      c3 = 1;
      X1 = xpos;
      Y1 = ypos - yinc;
      End;
    If contour = LRv & found1 = 0 Then do;
      found1 = 1;
      c4 = 1;
      X1 = xpos - xinc;
```

```
        Y1 = ypos - yinc;
        End;

    End; /* FOUND1 = 0 */

    If found1 = 1 Then do;
      If contour = ULv & c1 = 0 & found2 = 0 Then do;
        found2 = 1;
        X2 = xpos;
        Y2 = ypos;
        End;
      If contour = URv & c2 = 0 & found2 = 0 Then do;
        found2 = 1;
        X2 = xpos - xinc;
        Y2 = ypos;
        End;
      If contour = LLv & c3 = 0 & found2 = 0 Then do;
        found2 = 1;
        X2 = xpos;
        Y2 = ypos - yinc;
        End;
      If contour = LRv & c4 = 0 & found2 = 0 Then do;
        found2 = 1;
        X2 = xpos - xinc;
        Y2 = ypos - yinc;
        End;
      If side = 1 & s1 = 0 & found2 = 0 Then do;
        If ((ULv < contour) & (URv > contour)) Then do;
          found2 = 1;
          X2 = xpos - (((contour - ULv) / (URv - ULv)) * xinc);
          Y2 = ypos;
        End;
        If ((ULv > contour) & (URv < contour)) Then do;
          found2 = 1;
          X2 = xpos - (((ULv - contour) / (ULv - URv)) * xinc);
          Y2 = ypos;
        End;
      End;

      If side = 2 & s2 = 0 & found2 = 0 Then do;
        If ((URv < contour) & (LRv > contour)) Then do;
          found2 = 1;
          X2 = xpos - xinc;
          Y2 = ypos - (((contour - URv) / (LRv - URv)) * yinc);
        End;
        If ((URv > contour) & (LRv < contour)) Then do;
          found2 = 1;
          X2 = xpos - xinc;
          Y2 = ypos - (((URv - contour) / (URv - LRv)) * yinc);
        End;
      End;

      If side = 3 & s3 = 0 & found2 = 0 Then do;
        If ((LLv < contour) & (LRv > contour)) Then do;
          found2 = 1;
          X2 = xpos - (((contour - LLv) / (LRv - LLv)) * xinc);
          Y2 = ypos - yinc;
        End;
        If ((LLv > contour) & (LRv < contour)) Then do;
          found2 = 1;
          X2 = xpos - (((LLv - contour) / (LLv - LRv)) * xinc);
          Y2 = ypos - yinc;
        End;
      End;
```

```
            If side = 4 & s4 = 0 & found2 = 0 Then do;
              If ((ULv < contour) & (LLv > contour)) Then do;
                found2 = 1;
                X2 = xpos;
                Y2 = ypos - (((contour - ULv) / (LLv - ULv)) * yinc);
              End;
              If ((ULv > contour) & (LLv < contour)) Then do;
                found2 = 1;
                X2 = xpos;
                Y2 = ypos - (((ULv - contour) / (ULv - LLv)) * yinc);
              End;
            End;

          End; /* FOUND1=1 */
          db = 0;
      End; /* SIDE */

      If (X1 = 0 & Y1 = 0) | (X2 = 0 & Y2 = 0) Then GoTo skipPlot;
/*        'if this if is true then the condition is a contour value equals
        'a vertex and the current grid cell under investigation does not
        'have any other occurrences of the contour value in it, either
        'in other verticies or on any sides. The simple remedy is to just
        'skip the plotting of that line. A more complex solution is to set
        'more check variables. Note that the solution implemented here
        'will only work if the origin (0,0) is not in the plotting space,
        'that is, and therefore not a valid point for plotting.
*/

    xout=xout//x1;
    yout=yout//y1;
    fn=fn//'move';
    cont=cont//contour;

    xout=xout//x2;
    yout=yout//y2;
    fn=fn//'draw';
    cont=cont//contour;

    If secondRn = 1 Then do;
      reverse = 0;
      db = 0;
      secondRn = 0;
      found1 = 0;
      found2 = 0;
      GoTo reptRun;
      End;
skipPlot:
    found1 = 0;
    found2 = 0;
    X1 = 0;
    X2 = 0;
    Y1 = 0;
    Y2 = 0;
    s1 = 0;
    s2 = 0;
    s3 = 0;
    s4 = 0;
    c1 = 0;
    c2 = 0;
    c3 = 0;
    c4 = 0;
skipit:
    End; /* j */
  End; /* i */
End; /* contour */
```

```
/*
latrange=latmax-latmin;
longrang=longmax-longmin;

xout=(xout/100)*longrang+longmin;
yout=latmax-(yout/100)*latrange;
*/

contout=fuzz(cont);

done=xout||yout||contout;

create annoraw1 from done[colname={x y contour}];
append from done;

create annoraw2 from fn[colname={function}];
append from fn;

finish;

run getgrid(grid,gridlat,gridlong,rows,cols);
run mainbody(grid,gridlat,gridlong,rows,cols);

QUIT;

data annoraw;
      merge annoraw1 annoraw2;
run;


/*** FINAL DISPLAY ****/
data plot;
      set annoraw;
            color='black'; hsys='3'; xsys='2'; ysys='2'; size=0.1; line=1;
            when='a'; province='-3';
            x=&d2r*x;
            y=&d2r*y;
      if contour=0 then do;
            size=0.3; /* 2 */
            end;
      else do;
            size=.3;
            end;
run;

data pcaplot;
      length function $8;
      set  plot  origdot;
run;
%mend contour;
```

Cmapmerg.sas

```
/* This macro set parameters for mapping purposes */

%macro cmapmerg;
   goptions reset=all dev=win target=win hsize= 3 in vsize=3.5 in noprompt;
   /* PROC SORT DATA=origdot ; BY x y ; RUN; */

   data full;
     set origdot basemap plot;
   run;

   PROC GPROJECT data=full out=setfulx
                    polelong=100 paralel1=35 paralel2=80;
       id province;
   run;

          data dumb;
               set setfulx;
               if _n_=1;
          run;


     proc gproject data=setfulx out=setfull
           latmax=0.5
           latmin=-0.6
           longmax=0.45
           longmin=-0.6
           project=none;
           id province;
     run;

          data setfull;
                     set dumb setfull;
          run;

/* Reset Frame SIZE to .4 for screen testing, was 4. Nov 2001 */
     data frame;
          length function style color $ 8;
          retain xsys ysys '2' hsys '3' color 'black' line 1 style '1' size .4 when 'a' province '-90';
          function='move'; x=-0.6; y=-0.6; output;
          function='draw'; x=-0.6; y=0.5; output;
          function='move'; x=0.45; y=-0.6; output;
          function='draw'; x=0.45; y=0.5; output;
          function='move'; x=-0.6; y=0.5; output;
          function='draw'; x=0.45; y=0.5; output;
          function='move'; x=-.6; y=-0.6; output;
          function='draw'; x=.45; y=-0.6; output;
     run;
/* +++++++++++++++++++++++++++ */
data maintitl;
     length stmoval edmoval $9 text   $20 function style $8 ;
     stdate=&startyr; eddate=&endyr; fact="&f";

                    retain       function 'label' style 'SWISS' size 2.5 xsys ysys '5' hsys '3' color 'black'
                                                  when 'a' position '>' province '-66';
          text='Start: '||trim(left(stdate))||' ka BP'; y=89; x=70; output;
          text='End: '||trim(left(eddate))||' ka BP'; y=86; x=70; output;
                                                  if "&f" = 1 then text="Component One";
          else if "&f" = 2 then text="Component Two";
          else if "&f" = 3 then text="Component Three";
          else if "&f" = 4 then text="Component Four";
                                                  else if "&f" = 5 then text="Component Five";
                                                  else if "&f" = 6 then text="Component Six";
```

```
                                        else if "&f" = 7 then text="Component Seven";
                                        else if "&f" = 8 then text="Component Eight";
                                        else if "&f" = 9 then text="Component Nine";
        else text="Component Ten";
            y=83; x=70; output;
    drop stdate eddate fact stmonth stmoval edmonth edmoval;
run;
proc sort data=percent out=percent; by descending percent; run;
data _null_; set origdot end=last; if last then call symput('vcnt',left(_n_/2));run;

data plotper;
    length fact $2 style color $8 text $20 function $8 ;
    set percent end=last;
    fact=trim(left(percent));
    numstn=trim(left(&totlstns));
    retain              function 'label' style 'SWISS' size 2 xsys ysys '5' hsys '3' color 'black'
                                        when 'a' position '>' province '-66';
                        if &f=_n_ then do;
        text='Explained var: '||fact||'%'; y=79; x=70; output;

                            text="Stations used: &vcnt"; y=77; x=70; output;

        text='Contour interval: 20'; y=74; x=70; output;
        text='heavy line = 0'; y=71; x=70; output;
                                        text="Process variable: &varsel"; y=68; x=70; output;
                        end;
    drop fact1 fact2 fact3 MnT fact4 fact5 fact6 fact7 fact8 fact9 fact10 pid fact; /*factor when variable <100*/

    /* Figure label - only for final printing */
/*
    if ("&letters"='YES' AND "&factor" = 'factor1') then do;
      function='move'; color='white'; x=8; y=85; output;
      function='bar'; x=16; y=93; output;
      function='label'; x=10; y=98; text='A'; xsys='3'; ysys='3';
         position='9'; style='swissb'; color='black'; size=6; output;
    end;
    if ("&letters"='YES' AND "&factor" = 'factor2') then do;
      function='move'; color='white'; x=8; y=85; output;
      function='bar'; x=16; y=93; output;
      function='label'; x=10; y=98; text='C'; xsys='3'; ysys='3';
         position='9'; style='swissb'; color='black'; size=6; output;
    end;
    if ("&letters"='YES' AND "&factor" = 'factor3') then do;
      function='move'; color='white'; x=8; y=85; output;
      function='bar'; x=16; y=93; output;
      function='label'; x=10; y=98; text='E'; xsys='3'; ysys='3';
         position='9'; style='swissb'; color='black'; size=6; output;
    end;
*/
    /* box around the title section to mask the contour lines */
                        if last then do;
      function='move'; color='WHITE'; x=69; y=68; output;
      function='bar'; x=93; y=91; output;
                        end;
run;

    data setfull; set maintitl plotper setfull frame ; run;

    data land points;
   set setfull;
   if province > 0 then output land;
   if province < 0 then output points;
  run;

  pattern v=e;
```

```
goptions dev=win target=win;
   proc GMAP map=land data=basemap annotate=points gout=plots.plots;
         id province;
         choro colors /  nolegend
                     coutline=black
                     name="plot9&f"
                     ;
   run; quit;
%mend cmapmerg;
```

Bigwrit.sas

```
/* This macro runs a PCA on individual sites */

%macro bigwrit;

data viau1;
      infile "&DATAPATH&RAWDATA" dsd LRECL=2000; /* Watch for truncation message in LOG file */
      if _n_=1 then input;
      input e time var&v1.-var&v2.;
      output;
run;

data viau1; set viau1; if e='.' then return; output; run;

proc means data=viau1; output out=viauxn; run;
proc transpose data=viauxn out=viauxn; run;
data _null_; set viauxn; if _name_="e" then do; call symput('MAX',left(trim(col3))); call symput('MIN',left(trim(col2))); end; run;

proc sort data=viau1 out=viau1; by e time; run;
proc factor data=viau1 noprint outstat=pca1 m=p score mineigen=0; by e; run;

proc score data=viau1 score=pca1 out=pca1scor; by e; run;

/* GENERATE DATA TO PRODUCE SCREE PLOTS */
data scree1(drop=_TYPE_ _NAME_ rename=(time=eigen1 var1=eigen2 var2=eigen3 var3=eigen4 var4=eigen5
var5=eigen6 var6=eigen7 var7=eigen8 var8=eigen9 var9=eigen10 var10=eigen11));
set pca1; if _TYPE_='EIGENVAL'; run;
            PROC EXPORT DATA= WORK.scree1
        OUTFILE= "C:\users\VIAU\PhD PCA analysis\ScreeData.xls"
        DBMS=EXCEL2000 REPLACE;
            RUN;
/*  end scree plot generate */

     %do fact=1 %to &factend.; /* Set end point to how many wanted/available */
          %let merg&fact = ;
          %do e = &MIN %to &MAX;
               %let merg&fact =&&merg&fact f&fact.e&e;
               data f&fact.e&e(rename=(factor&fact=e&e.f&fact)) ;
                    set pca1scor;
                    if e=&e;
                    keep time factor&fact ;
               run;
          %end;
          data lg&fact; merge &&merg&fact; by time; run;
     %end;

%mend bigwrit;
```

APPENDIX E – "R" PROGRAM FOR INTERPOLATION

Interpolate.fossil.f

*** This R function interpolates linearly between dates in the age-depth model of the fossil pollen spectra Developed by M.C Sawada (LPC) 2000. ***

```
function(fossilDataFrame, colFirstTaxa = 10, colLastTaxa = 98, colCurrDates = 8, colEnum = 1, colSiteNum = 2,
colSiteName = 3, colLong = 4, colLat = 5, target.dates = targets)
{
  holdnames <- as.vector(unlist(names.f(fossilDataFrame)))  #
  #get enumbers and the date column plus the columns with the proportions/percentages of each taxon.
  datainterp1 <- as.matrix(fossilDataFrame[, c(colEnum, colCurrDates, colFirstTaxa:colLastTaxa)]) #
  #make no dimnames to speed up calculations
  dimnames(datainterp1) <- NULL  #
  #get list of enumbers for each site
  enumbers <- as.numeric(unique(datainterp1[, 1]))  #
  newcolend <- dim(datainterp1)[2]
  outtable <- vector("numeric")  #
  tstart <- proc.time()[1]
  #ttime <- proc.time()[1] - currtime
  #      currtime <- proc.time()[1]

  for(i in 1:length(enumbers)) {
    datainterp <- datainterp1[which(datainterp1[, 1] == enumbers[i]), ]  #
  #  write(c("prick", "dick"), "c:\\asdf.txt", ncolumns = 2, append = T)
    if(length(dim(datainterp)) == 2) {
      interpolatedPrcts <- data.frame(apply(datainterp[, 3:newcolend], 2, approx.multi.f, datainterp[, 2],
target.dates)) #
      inter <- cbind(ENUM = rep(fossilDataFrame[which(fossilDataFrame[, colEnum] == enumbers[i])[1],
colEnum], length(target.dates)), SITENUM = rep(
        fossilDataFrame[which(fossilDataFrame[, colEnum] == enumbers[i])[1], colSiteNum],
length(target.dates)), NAME = rep(fossilDataFrame[which(
        fossilDataFrame[, colEnum] == enumbers[i])[1], colSiteName], length(target.dates)), LONGITUDE =
rep(fossilDataFrame[which(fossilDataFrame[, colEnum
        ] == enumbers[i])[1], colLong], length(target.dates)), LATITUDE =
rep(fossilDataFrame[which(fossilDataFrame[, colEnum] == enumbers[i])[1], colLat],
        length(target.dates)), AGECALBP = target.dates, interpolatedPrcts)
      #write.table(inter, "c:\\xnt.txt", sep = ",")
      outtable <- rbind(outtable, inter[which(is.na(inter[, 8]) == F), ])  #
    }
    write(proc.time()[1] - tstart, "c:\\timeinterp1.txt", append = T)  #
    tstart <- proc.time()[1]
  }
  write(paste("final", proc.time()[1] - tstart), "c:\\timeinterp1.txt", append = T)  #
  outtable
}
```

APPENDIX F - SUPPLEMENTAL DATA AND INFORMATION - PCA ANALYSIS

With such a large analysis, it was necessary to ensure that numerical problems were not influencing the analysis. The principal component analysis was performed several times using different number of pollen sums. The resulting factor scores time plots using a high pollen sum (i.e. 89 and 47 types) (Table F1.1 and F1.2) yielded interesting but suspicious results (Figure F1.1). The resulting factor score plots are very smooth and the number of peaks and troughs seem to be doubling as the factor components increase (Figure F1.1). We proceeded to test other relationships to determine if indeed this was a true signal or resulted from methodological aliases or artifacts. Since our temporal and spatial axes were fixed (i.e. 125 sites and 48 time periods), we redid our PCA on 9 different combinations of subsets of the 47-type sum (Table F2). For example, trees and herbs were divided into major and minor taxa groups (Table F1.3 and F1.4) based on pollen proportions in the database. This selection criterion resulted into a subdivision of the 47-type pollen sum into a 24-type group of major trees and herbs, which represent higher pollen proportions, and a 23-type taxa group of minor trees and herbs, which consist of taxa with lower proportions. These two sub-dataset are represented by the letters C and D (Table F2). Further subdivision of these two major and minor groups (C and D) (Table F2) were also performed to analyze independently major tree, major herb, minor tree, and minor herb taxa, which are represented by letters (F,G and H,I) (Table F2; Table F1.5,1.6,1.8,1.9). Next, all tree taxa proportions were summed, as were all herb taxa proportions at each site, to explore independently tree and herb responses through space and time (A and B; Table F2). Finally, a principal component analysis was performed on each individual taxon (i.e. 47 types) through space and time to investigate individual taxon behavior. This resulted in 56 independent runs of the PCA model.

Our null hypothesis was that these factor scores time series plots resulted from the model used or were the result of chance alone. To test this, we constructed an artificial random dataset where pollen percentages at each sites were varied according to the distribution of the original data (i.e. gamma distribution). This analysis revealed no similar temporal evolution of factor scores compared with the original analyses (Figure F3). We therefore conclude that the factor scores of the original data were neither due to chance alone nor the result of the model used. Conversely, this meant that the data contained a significant amount of spatio-temporal autocorrelation. The temporal factor components revealed in this analysis are therefore representative of a true signal embedded in the pollen data.

In general, a decrease in the pollen sum (i.e. using the different susbsets) resulted in more variable factor scores; nevertheless at least the first 10 score plots resembled those obtained using all 47 types (Figure F1.1). For example, the analysis including only major tree and herb taxa (Table F2 (C)) revealed strikingly similar temporal evolution in factor scores (Figure F1.2). Similar but more variable factor score time series plots are obtained when using only minor tree and herb taxa pollen sums (Table F2 (D); Figure F1.3). Independent use of major tree and herb taxa pollen sums revealed slightly more variable temporal score plots (Table F2 (F,G), Figures F1.4 and F1.5). The temporal variability in factor scores also increases in the case of independent use of major trees, major herbs, minor trees and minor herbs subsets (Table F2 (H,I); Figures F1.6 and F1.7). Independent PCA performed on the sum of all tree taxa as well as the sum of all herbs revealed once more strikingly similar temporal factor score plots (Table F2 (A,B); Figures F1.8 and F1.9). Finally, results from the PCA performed on each individual taxon through space revealed time series score plots coherent with the overall scores obtained from the 47-type pollen sum, although the order and temporal variability of the factor scores may differ (Figure F2.1-F2.47).

These results raised the question on which factors to retain. In order to address this issue, we proceeded to test the relationships between the different subgroups (Table F2). There was a strong inverse relationship between trees and herbs factor scores where it is described as A = -B or trees = minus herbs (Table F2). By using this relationship, we retained 10 factors based on a number of spatial analyses (Figure F4 and F5; Table F3, Table F4). Although factor 11 has a high spatial autocorrelation (Figure F4; Table F3), it was not retained based on two criteria. First, the spatial coherency map depicting the relationship between herbs and trees shows that most regions begin to be affected by opposing signs where the relationship no longer is true (tree = herbs or A=B; Figure F5). Next, the relationship begins to breakdown during the middle Holocene as shown in the scores time series plot of Factor 11 in Figure F4. As seen in Chapter 2, the middle Holocene seems to be less variable as far as millennial-scale climate variability is concerned, compared to the early and late Holocene (e.g. Grimm *et al.*, 1992; Huntley *et al.*, 1992; Viau *et al.*, 2002). The middle Holocene was perhaps characterized by even more subtle changes in climate, therefore tree and herb tolerance and resilience may have been greater than the climate variability during this period.

The summary of the total explained variance for 10 factors for all 9 subgroup PCA runs is shown in Table F5. Major tree and herb taxa sums (C, F, G) reveal less overall variability while minor tree and herb taxa pollen sums (D, H, I) are more variable (Table F5). The original combination of minor and major taxa pollen sum (E) reveals a total explained variance that is intermediate (66%). This is enough to suggest that although minor taxa pollen sums are more variable, they contributed in part to reinforce the major type signal (Table F5). A detailed summary of explained variance by individual taxa indicates that some taxon may be important for some factor scores, for example *Picea* for factor 1, but overall all taxa contribute to some degree to the overall temporal evolution of the factor scores (Table F6).

The loading contributions of each taxon to factor scores and site contribution are illustrated in Chapter 3, which reveal significant overall contributions across North America. Therefore, this suggests that each taxon responded individually as can be seen in Figure F4. However, these analyses reveal common and unique times of variation amongst them through space and time. This raised the question on the interpretation of the resulting temporal factor scores and loadings. What do they represent? The temporal evolution of the factor score time series plots are here interpreted as trends representing different temporal scale and magnitude of vegetation changes common in time, space and taxonomy. For example, factor 1 scores can be clearly interpreted as a long-term trend associated with deglaciation, where herbs were not alone anymore as trees expanded at the continental-scale of North America (Figure F1.1). Other trends could represent seasonality changes due to orbital configuration or simply humidity and temperature changes associated with reorganization of atmospheric circulation patterns (Figure F1.1). The magnitude of change is representative of the order in factor scores where the most intense changes are first recorded in the early scores such as orbital and insolation changes affecting vegetation composition while more subtle changes are revealed at lower factor score plots, such as millennial-scale climate variability.

REFERENCES

Grimm, E.C., and Jacobson, G.L Jr. 1992. Fossil-pollen evidence for abrupt climate changes during the past 18,000 years in eastern North America. *Climate Dynamics* 6: 179-184.

Huntley, B. 1992. Rates of change in the European palynological record of the last 13,000 years and their climatic interpretation. *Climate Dymanics* 6: 185-191.

Viau, A.E, Gajewski, K., Fines, P., Atkinson, D.E., and Sawada, M.C. 2002. Widespread evidence of 1500 yr climate variability in North America during the past 14 000 yr. Geology 30(5): 455-458.

POLLEN SUMS USED IN PCA

Table F1.1 89 types pollen sum used in PCA

v1	*Abies*	v31	*Dryas*	v61	*Picea mariana*
v2	*Acer*	v32	*Ephedra*	v62	*Picea undiff.*
v3	*Alnus crispa*	v33	*Ericaceae/Chamead./Ledum*	v63	*Pinus strobus*
v4	*Alnus rugosa*	v34	*Ericaceae Empetrum*	v64	*Pinus undiff.*
v5	*Alnus undiff.*	v35	*Ericaceae Vaccinium*	v65	*Plantaginaceae*
v6	*Ambrosia type*	v36	*Ericaceae undiff.*	v66	*Platanus*
v7	*Amorpha*	v37	*Eriogonum*	v67	*Poaceae*
v8	*Apiaceae*	v38	*Euphorbiaceae*	v68	*Polygonum*
v9	*Aquifoliaceae*	v39	*Fabaceae*	v69	*Polypodiaceae*
v10	*Arceuthobium*	v40	*Fagus*	v70	*Populus*
v11	*Artemisia*	v41	*Fraxinus*	v71	*Potentilla*
v12	*Asteraceae*	v42	*Juglans*	v72	*Prosopis*
v13	*Betula undiff.*	v43	*Larix Pseudotsuga*	v73	*Pteridium*
v14	*Brassicaceae*	v44	*Larrea*	v74	*Quercus*
v15	*Cactaceae*	v45	*Liliaceae*	v75	*Ranunculaceae*
v16	*Campanulaceae*	v46	*Liquidambar*	v76	*Rhamnaceae*
v17	*Caprifoliaceae*	v47	*Lycopodium annotinum*	v77	*Rubiaceae*
v18	*Carya*	v48	*Lycopodium clavatum*	v78	*Rumex*
v19	*Caryophyllaceae*	v49	*Lycopodium complanatum*	v79	*Salix*
v20	*Castanea*	v50	*Lycopodium selago*	v80	*Sarcobatus*
v21	*Ceanothus*	v51	*Lycopodium undiff.*	v81	*Saxifragaceae*
v22	*Celtis*	v52	*Magnoliaceae*	v82	*Selaginella*
v23	*Cephalanthus*	v53	*Malvaceae*	v83	*Shepherdia*
v24	*Cercocarpus*	v54	*Myrica*	v84	*Sphagnum*
v25	*Chenopodiaceae/Amaranthaceae*	v55	*Nyssa*	v85	*Taxodium*
v26	*Chrysolepis*	v56	*Onagraceae*	v86	*Thalictrum*
v27	*Clethra*	v57	*Ostrya/Carpinus*	v87	*Tilia*
v28	*Corylus*	v58	*Oxyria*	v88	*Tsuga*
v29	*Cupressaceae*	v59	*Papaveraceae*	v89	*Ulmus*
v30	*Cyperaceae*	v60	*Picea glauca*		

Table F1.2. 47-types pollen sum (E)

v1	*Abies*	v26	*Ostrya/Carpinus*
v2	*Acer*	v27	*Picea undiff.*
v3	*Alnus undiff.*	v28	*Pinus undiff.*
v4	*Ambrosia type*	v29	*Plantaginaceae*
v5	*Apiaceae*	v30	*Poaceae*
v6	*Aquifoliaceae*	v31	*Polygonum*
v7	*Artemisia*	v32	*Polypodiaceae*
v8	*Asteraceae*	v33	*Populus*
v9	*Betula undiff.*	v34	*Pteridium*
v10	*Carya*	v35	*Quercus*
v11	*Castanea*	v36	*Ranunculaceae*
v12	*Cephalanthus*	v37	*Rubiaceae*
v13	*Chenopodiaceae/Amaranthaceae*	v38	*Rumex*
v14	*Corylus*	v39	*Salix*
v15	*Cupressaceae*	v40	*Sarcobatus*
v16	*Cyperaceae*	v41	*Saxifragaceae*
v17	*Ericaceae undiff.*	v42	*Shepherdia*
v18	*Fagus*	v43	*Sphagnum*
v19	*Fraxinus*	v44	*Taxodium*
v20	*Larix/Pseudotsuga*	v45	*Tilia*
v21	*Liliaceae*	v46	*Tsuga*
v22	*Lycopodium undiff.*	v47	*Ulmus*
v23	*Magnoliaceae*		
v24	*Myrica*		
v25	*Nyssa*		

Table F1.3. 24-type pollen sum (C)

v1	*Abies*
v2	*Acer*
v3	*Alnus undiff.*
v4	*Ambrosia type*
v5	*Artemisia*
v6	*Asteraceae*
v7	*Betula undiff.*
v8	*Carya*
v9	*Castanea*
v10	*Chenopodiaceae/Amaranthaceae*
v11	*Cupressaceae*
v12	*Cyperaceae*
v13	*Fagus*
v14	*Fraxinus*
v15	*Ostrya/Carpinus*
v16	*Picea undiff.*
v17	*Pinus undiff.*
v18	*Poaceae*
v19	*Populus*
v20	*Quercus*
v21	*Salix*
v22	*Tilia*
v23	*Tsuga*
v24	*Ulmus*

Table F1.4. 23-type pollen sum (D)

v1	*Apiaceae*
v2	*Aquifoliaceae*
v3	*Cephalanthus*
v4	*Corylus*
v5	*Ericaceae undiff.*
v6	*Larix/Pseudotsuga*
v7	*Liliaceae*
v8	*Lycopodium undiff.*
v9	*Magnoliaceae*
v10	*Myrica*
v11	*Nyssa*
v12	*Plantaginaceae*
v13	*Polygonum*
v14	*Polypodiaceae*
v15	*Pteridium*
v16	*Ranunculaceae*
v17	*Rubiaceae*
v18	*Rumex*
v19	*Sarcobatus*
v20	*Saxifragaceae*
v21	*Shepherdia*
v22	*Sphagnum*
v23	*Taxodium*

Table F1.5. 18-type pollen sum (F)

v1	*Abies*
v2	*Acer*
v3	*Alnus undiff.*
v4	*Betula undiff.*
v5	*Carya*
v6	*Castanea*
v7	*Cupressaceae*
v8	*Fagus*
v9	*Fraxinus*
v10	*Ostrya/Carpinus*
v11	*Picea undiff.*
v12	*Pinus undiff.*
v13	*Populus*
v14	*Quercus*
v15	*Salix*
v16	*Tilia*
v17	*Tsuga*
v18	*Ulmus*

Table F1.7. 11-type pollen sum (H)

v1	*Aquifoliaceae*
v2	*Cephalanthus*
v3	*Corylus*
v4	*Ericaceae undiff.*
v5	*Larix/Pseudotsuga*
v6	*Magnoliaceae*
v7	*Myrica*
v8	*Nyssa*
v9	*Sarcobatus*
v10	*Shepherdia*
v11	*Taxodium*

Table F1.6. 6-types pollen sum (G)

v1	*Ambrosia type*
v2	*Artemisia*
v3	*Asteraceae*
v4	*Chenopodiaceae/Amaranthaceae*
v5	*Cyperaceae*
v6	*Poaceae*

Table F1.8. 12-type pollen sum (I)

v1	*Apiaceae*
v2	*Liliaceae*
v3	*Lycopodium undiff.*
v4	*Plantaginaceae*
v5	*Polygonum*
v6	*Polypodiaceae*
v7	*Pteridium*
v8	*Ranunculaceae*
v9	*Rubiaceae*
v10	*Rumex*
v11	*Saxifragaceae*
v12	*Sphagnum*

Table F1.9 Pollen sum - 17 types (Gajewski,1987)

v1	*Abies*
v2	*Acer*
v3	*Alnus*
v4	*Artemisia*
v5	*Betula*
v6	*Castanea*
v7	*Cupressineae*
v8	*Fagus*
v9	*Fraxinus*
v10	*Gramineae*
v11	*Ostrya*
v12	*Picea*
v13	*Pinus*
v14	*Populus*
v15	*Quercus*
v16	*Tsuga*
v17	*Ulmus*

Table F2. Subdivision of 47-type pollen sum. Each letter corresponds to the pollen sum of Table F1.2 to F1.8

	Major taxa		Minor Taxa		All taxa	
Trees	F	18	H	11	A	29
Herbs	G	6	I	12	B	18
All taxa	C	24	D	23	E	47

Note: Values represent number of taxa included in subgroups.

Table F3: Spatial autocorrelation of the inverse relationship between herbs and trees (A=-B).

Factor	positive	negative	opposite	sum	ratio (%)
F1	87	32	6	125	95.2
F2	81	40	4	125	96.8
F3	73	47	5	125	96
F4	63	54	8	125	93.6
F5	64	57	4	125	96.8
F6	52	63	10	125	92
F7	62	57	6	125	95.2
F8	62	45	18	125	85.6
F9	52	58	15	125	88
F10	61	56	8	125	93.6
F11	52	48	25	125	80
F12	21	14	90	125	28
F14	27	20	78	125	37.6

Positive values represent number of sites having herbs and trees with positive direction (+,+). A=-B

Negative values represent number of sites having herbs and trees with negative directions (-,-). B=-A

Opposite values represent number of sites having herbs and trees with opposing direction (-,+ ; +,-). A=B or B=A

Table F4: Color-coded scheme illustrating spatial coherency by factor loadings. Same colors in both tree and herb loadings at sites indicate that the relationship between herbs and trees is true at those sites. Opposing colors between trees and herbs indicates that the relationship no longer holds at those sites.

				Factor 1		Factor 2		Factor 3		Factor 4		Factor 5		Factor 6	
ID	lat	long	sort	Trees	herbs	Trees	herbs	Trees	herbs	Trees	herbs	Trees	herbs	Trees	herbs
969	75.4	82.5	0	-0.54	-0.5	0.46	0.46	-0.04	-0.01	-0.12	-0.1	-0.06	-0	0.1	0.1
514	68.2	161	1	0.23	0.17	0.7	0.73	-0.09	-0.06	-0.14	-0.1	-0.08	0.08	0.54	0.55
521	67.6	151	1	0.55	0.53	-0.1	-0.08	-0.37	-0.39	0.513	0.48	-0.24	-0.3	-0.03	0.03
519	67.3	153	1	0.76	0.76	0.23	0.21	0.22	0.22	0.238	0.24	-0.01	0.07	0.47	0.46
508	67.1	154	1	0.36	0.35	-0.37	-0.38	-0.04	-0.08	0.296	0.32	0.15	0.03	-0.12	-0.2
506	67.1	154	1	0.74	0.74	-0.03	-0.02	-0.41	-0.42	0.094	0.1	0.02	0.08	0.29	0.3
515	66.8	157	1	0.73	0.74	0.35	0.34	-0.04	-0.03	0.127	0.09	-0.13	-0.1	0.22	0.23
520	67.4	150	2	0.52	0.49	-0.42	-0.42	0.07	0.06	0.434	0.44	-0.29	-0.3	0.14	0.25
518	67.2	142	2	0.92	0.93	0.27	0.24	0.13	0.14	-0.01	-0.1	-0.17	-0.19	-0.08	-0
3500	69.3	134	3	-0.14	-0.1	0.58	0.58	-0.49	-0.46	0.199	0.18	0.14	0.12	-0.05	-0.1
3501	69.1	132	3	0.89	0.89	0.14	0.11	0.18	0.19	0.021	0.02	0.12	0,14	0.11	0.05
1551	69.1	133	3	0.47	0.47	0.22	0.24	-0.69	-0.68	0.07	0.07	-0.03	-0.05	0.14	0.09
1259	68.4	138	3	0.25	0.25	-0.12	-0.13	-0.29	-0.31	-0.35	-0.3	0.1	0.17	0.16	0.11
1549	68.3	133	3	0.85	0.86	0.27	0.23	0.35	0.36	0.186	0.18	-0.11	-0.12	0.1	0.11
1552	66	136	3	0.69	0.7	0.2	0.17	0.44	0.44	0.193	0.19	0.08	0.09	0.27	0.22
522	64.3	151	4	0.58	0.58	0.41	0.42	-0.32	-0.3	0.067	0.04	-0.15	-0.2	-0.24	-0.2
854	50	56.2	5	0.95	0.96	0.25	0.21	0.15	0.14	0.017	-0	-0.02	-0.01	0.09	0.07
850	46.6	122	6	0.09	0.1	0.38	0.39	-0.35	-0.34	0.495	0.52	0.41	0.32	-0.04	-0.2
445	45.9	121	6	0.75	0.74	-0.48	-0.49	-0.17	-0.19	0.213	0.21	-0.08	-0.08	0.17	0.18
857	45.8	122	6	0.09	0.08	-0.69	-0.7	0.46	0.42	0.302	0.35	0.12	0.04	-0.07	-0.1
849	48.6	113	7	-0.24	-0.2	0.51	0.52	0.27	0.3	0.145	0.1	-0.3	-0.26	0.23	0.25
1970	46.5	112	7	0.66	0.63	0.13	0.13	-0.04	-0.01	-0.18	-0.2	-0.07	-0.09	-0.19	-0.1
94	49.4	99.3	8	-0.34	-0.4	-0.79	-0.78	0.17	0.13	-0.02	0.01	-0.02	0.03	0.2	0.22
1085	47.1	93.6	8	0.91	0.91	-0.21	-0.24	0.2	0.17	0.163	0.15	-0.03	-0.06	0.01	0.02
290	45.3	90.1	8	0.84	0.85	0.24	0.21	0.26	0.26	0.287	0.26	-0.03	-0.02	0.13	0.13
143	45.1	92.8	8	-0.56	-0.5	-0.08	-0.1	0.72	0.73	-0.03	-0	0.16	0.13	-0.22	-0.2
58	49.2	81.3	9	0.95	0.95	0.03	0.01	0	-0	-0.08	-0.1	-0.1	-0.13	-0.17	-0.1
261	46.6	70.6	10	0.81	0.82	0.27	0.24	0.14	0.14	0.01	-0	-0.23	-0.26	-0.22	-0.2
78	45.9	70.4	10	0.67	0.71	0.42	0.38	0.19	0.19	0.161	0.18	0.3	0.29	0.2	0.06
271	45.6	70.6	10	0.84	0.86	0.27	0.23	0	-0.02	-0.14	-0.1	0.09	0.12	0.18	0.12
266	45.5	73.3	10	0.89	0.93	0.09	0.04	-0.12	-0.09	-0.21	-0.1	0.08	0.03	-0.07	-0.1
2339	45.2	77.4	10	0.86	0.85	-0.04	-0.08	-0.3	-0.32	-0.3	-0.3	0.11	0.12	-0.07	-0.1
486	48.9	66.1	11	0.87	0.88	0.34	0.31	0.28	0.3	0.1	0.06	-0.14	-0.13	0.07	0.09
16	45.3	67.3	11	0.81	0.81	-0.15	-0.16	-0.35	-0.35	-0.19	-0.2	0.25	0.27	0.12	0.04
145	45.1	66.7	11	0.94	0.94	0.05	0.02	0.05	0.04	-0.11	-0.1	0.13	0.15	0.1	0.06
856	49.5	55.5	12	0.97	0.97	0.15	0.1	0.02	-0.01	-0.12	-0.2	-0.02	-0	-0.02	-0
1078	48.3	58.8	12	0.77	0.79	0.55	0.51	0.16	0.18	0.111	0.06	-0.05	-0.08	-0.04	-0
1069	48.2	58.2	12	0.78	0.79	0.38	0.34	-0.14	-0.16	-0.2	-0.2	0.13	0.11	-0.18	-0.2

866	44.9	110	13	0.8	0.81	0.4	0.37	0.03	0.04	-0.24	-0.3	-0.08	-0.1	-0.24	-0.2
884	44.9	111	13	0.78	0.79	0.49	0.46	0.09	0.09	0.01	-0	-0.02	-0.05	-0.1	-0.1
858	44.7	111	13	0.95	0.95	0.06	0.02	0.04	0.04	-0.2	-0.2	-0.08	-0.09	-0.06	-0
865	44.2	110	13	0.92	0.92	0.13	0.09	0.27	0.27	-0.11	-0.1	-0.13	-0.13	-0.09	-0.1
883	44.1	111	13	0.83	0.82	-0.42	-0.43	-0.22	-0.25	0.138	0.13	-0.09	-0.12	0	0.03
859	43.9	110	13	0.32	0.35	0.37	0.33	0.7	0.7	0.216	0.23	0.08	0.12	0.26	0.22
862	43.8	111	13	0.1	0.12	0.77	0.76	-0.17	-0.12	-0.11	-0.1	0.16	0.14	-0.08	-0.2
464	44.3	107	14	0.66	0.65	-0.15	-0.19	-0.14	-0.16	-0.25	-0.2	0.42	0.46	0.08	-0
288	45	93.6	15	-0.74	-0.7	0.23	0.22	0.49	0.51	-0.13	-0.1	-0.04	-0.06	-0.18	-0.2
473	44.8	99.9	15	-0.91	-0.8	-0.07	-0.01	-0.28	-0.41	0.076	0.14	-0.08	-0.15	-0.13	-0.1
189	43.3	95.2	15	-0.82	-0.8	0	0.01	0.5	0.52	0.132	0.13	-0.02	-0.06	-0.15	-0.1
691	43.2	93.4	15	-0.79	-0.8	0.07	0.04	0.52	0.55	-0.14	-0.1	-0.02	-0.03	-0.2	-0.2
462	43	91.5	15	-0.17	-0.1	0.52	0.48	0.08	0.09	-0.36	-0.4	-0.03	0.16	0.54	0.58
52	42	93.3	15	-0.81	-0.8	0.29	0.28	0.35	0.38	-0.17	-0.2	0.05	0.04	-0.21	-0.2
82	44.4	80.3	16	0.76	0.78	0.09	0.06	-0.16	-0.14	-0.33	-0.3	0.17	0.13	-0.22	-0.2
1157	44.2	87.8	16	-0.19	0.11	0.36	0.34	0.35	0.36	-0.35	-0.1	0.16	0.26	0.22	0.19
92	44.1	87.7	16	0.55	0.63	0.44	0.46	0.54	0.43	-0.05	0.01	0.19	0.16	0.1	0.08
1083	43.9	80.4	16	0.69	0.7	0.21	0.16	0.26	0.27	0.065	0.02	-0.33	-0.35	-0.16	-0.1
1155	43.4	89.7	16	-0.38	-0.4	0.66	0.7	0.01	0.06	0.066	0.02	-0.22	-0.13	0.22	0.27
72	43.4	89.7	16	-0.33	-0.2	0.54	0.57	0.36	0.35	0.004	-0	-0.16	-0.15	-0.04	0.02
276	43.4	84.9	16	0.39	0.63	0.04	0.05	0.66	0.5	-0.08	0.01	0.01	0.03	0.02	0.13
1056	43.2	81.7	16	-0.55	-0.5	0.63	0.64	-0.23	-0.19	0.123	0.11	0.14	0.1	-0.15	-0.2
286	42.4	85.4	16	0.79	0.85	-0.13	-0.11	0.31	0.2	0.158	0.09	-0.16	-0.1	0.18	0.22
87	42.3	83.6	16	-0.18	0.47	0.48	0.4	0.66	0.38	0.11	0.31	0.35	0.42	0.2	0.28
1060	41.9	82.8	16	-0.09	-0	0.24	0.18	0.5	0.5	-0.54	-0.5	0.07	0.3	0.54	0.52
50	41.7	86.5	16	0.21	0.5	-0.33	-0.37	0.65	0.51	-0.06	-0.1	-0.11	-0.02	0.25	0.25
205	41.6	85.3	16	0.71	0.87	0.17	0.09	0.32	0.23	0.204	0.08	-0.19	-0.23	-0.3	-0.1
925	41.4	84.8	16	0.21	0.38	0.53	0.43	0.15	0.05	-0.08	-0	-0.05	-0.01	0.06	0.1
241	40.4	83.8	16	0.44	0.62	0.65	0.53	0.46	0.39	-0.16	-0.1	-0.19	-0.2	0.06	-0
265	40.4	82.5	16	-0.57	-0.2	0.15	-0.07	0.33	0.35	0.23	0.53	0.03	0.07	-0.04	-0.4
924	40.1	83.5	16	-0.01	0.21	0.62	0.66	0.14	0.03	0.149	0.15	-0.17	-0.35	-0.46	-0.3
467	44.9	77.2	17	0.88	0.88	0.21	0.16	0.18	0.2	0.061	0.01	-0.24	-0.24	-0.06	0
104	44.2	74	17	0.9	0.92	0.29	0.24	0.23	0.22	-0.04	-0.1	-0.04	-0.04	-0.01	-0
1077	44.2	79.8	17	0.9	0.9	0.02	-0.03	0.12	0.12	-0.07	-0.1	0.12	0.09	-0.11	-0.1
1058	44.1	78.1	17	0.5	0.54	0.27	0.23	-0.05	-0.07	0.076	0.24	0.62	0.6	0.06	-0.1
68	44	71.8	17	0.68	0.67	0.32	0.29	0.1	0.1	-0.24	-0.3	-0.26	-0.32	-0.46	-0.4
448	44	70.4	17	0.91	0.91	0.05	-0	0.05	0.03	-0.06	-0.1	0.03	0.08	0.07	0.06
275	44	79.4	17	0.65	0.67	0.57	0.57	0.1	0.07	0.257	0.18	-0.2	-0.24	0.06	0.07
228	42.8	71.7	17	0.94	0.94	0.22	0.19	-0.12	-0.12	-0.1	-0.1	-0.09	-0.09	-0.11	-0.1
13	42.5	74.6	17	0.38	0.41	0.8	0.79	-0.16	-0.13	-0.01	-0.1	-0.1	-0.09	-0.03	-0
7	42.3	78.9	17	0.24	0.4	0.53	0.55	-0.39	-0.46	0.335	0.31	0.36	0.22	-0.22	-0.2
18	42.3	77.9	17	-0.07	-0.1	-0.36	-0.4	0.71	0.67	-0.45	-0.4	0.14	0.23	-0.01	-0
105	42.2	74	17	-0.58	-0.6	0.69	0.7	-0.16	-0.11	-0.09	-0.1	0.12	0.13	-0.07	-0.1
77	41.9	70	17	0.42	0.48	0.39	0.38	0.22	0.16	0.382	0.45	0.6	0.55	0.13	-0
247	41.7	76.4	17	-0.49	-0.3	0.46	0.57	0.04	-0.11	0.173	0.23	0.24	0.2	0.03	0.03
222	41.4	72.1	17	0.67	0.7	0.46	0.46	0.09	0.07	0.244	0.25	-0.03	-0.1	-0.15	-0.1

199	41.3	74.8	17	-0.46		0.19		0.01		0.157		-0.06		-0.14	
690	41.3	70.2	17	0.69	0.71	0.15	0.05	0.28	0.28	0.156	0.19	0.43	0.35	-0.25	-0.3
585	41.2	71.6	17	0.11	0.15	0.21	0.2	0.16	0.13	0.581	0.67	0.62	0.52	-0	-0.2
259	41	75.3	17	-0.49	-0.4	0.4	0.46	-0.06	-0.17	0.266	0.29	0.3	0.26	-0.02	-0
192	40.8	77.4	17	0.85	0.83	-0.44	-0.48	0.15	0.11	-0.02	-0	-0.07	-0.06	0.01	0.02
153	40.5	75.7	17	-0.7	-0.7	0.3	0.34	0.37	0.37	-0.05	-0.1	0.03	0.13	0.21	0.21
257	40.4	74.5	17	-0.12	-0.1	0.63	0.53	0.26	0.34	0.264	0.31	0.26	0.17	-0.2	-0.3
447	45	69.3	18	0.81	0.83	0.03	0.01	-0.17	-0.17	-0.31	-0.3	0.1	0.08	-0.21	-0.2
174	44.6	68.6	18	0.91	0.91	0.15	0.13	-0.2	-0.2	-0.11	-0.1	0.03	0.01	-0.13	-0.1
240	44.6	63.6	18	0.87	0.86	-0.18	-0.2	-0.26	-0.29	-0.21	-0.2	0.04	0.08	0.09	0.07
443	43.9	69.4	18	0.78	0.77	0.13	0.08	-0.09	-0.13	-0.17	-0.2	0.35	0.34	-0.06	-0.1
1094	38	120	19	-0.01	-0	-0.71	-0.69	0.03	0.07	0.133	0.11	0.08	0.12	0.04	0
695	37.6	119	19	0.25	0.28	-0.66	-0.69	0.45	0.4	0.062	0.07	0.24	0.14	-0.34	-0.4
882	37	120	19	0.9	0.91	-0.1	-0.12	-0.23	-0.23	0.118	0.11	0.1	0.1	0.15	0.11
1313	37.7	106	20	0.55	0.53	-0.63	-0.67	0.27	0.23	-0.06	-0	0.1	0.09	-0.02	-0
880	37.6	106	20	0.55	0.62	0.71	0.65	0.12	0.18	-0.05	-0.1	-0.13	-0.12	-0.09	-0.1
868	38.5	98.7	21	-0.9	-0.9	-0.21	-0.16	-0.26	-0.28	0.132	0.21	0.05	0.05	0.18	0.15
64	36.8	91.1	21	0.41	0.4	-0.65	-0.68	0.31	0.29	-0.19	-0.2	0.2	0.18	-0.1	-0.2
441	36.6	90.6	21	0.66	0.63	-0.53	-0.6	0.31	0.19	-0.02	0.04	0.11	0.01	-0.26	-0.3
57	38.2	80.3	22	0.22	0.22	-0.72	-0.75	0.3	0.22	-0.16	-0.1	-0.07	-0.13	-0.17	-0.2
115	37.5	85.7	22	0.7	0.79	0.25	0.15	-0.34	-0.29	0.414	0.38	-0.01	-0.05	0.02	0.04
8	36	85.5	22	0.33	0.3	-0.55	-0.53	-0.6	-0.64	0.107	0.14	0.23	0.24	0.26	0.2
60	40	77.6	23	-0.65	-0.6	0.02	0.05	0.29	0.32	0.244	0.25	0.01	0.1	0.42	0.42
34	39.6	79.3	23	0.82	0.83	0.37	0.35	-0.21	-0.2	0.122	0.13	0.27	0.24	-0.03	-0.1
1148	38.2	79.6	23	0.24	0.21	-0.54	-0.53	-0.12	-0.13	0.48	0.49	-0.16	-0.14	0.25	0.3
390	38	79.1	23	-0.5	-0.4	-0.02	-0.03	0.1	0.1	0.635	0.64	-0.04	-0.17	-0.22	-0.2
100	38	79	23	0.34	0.32	-0.68	-0.7	-0.06	-0.12	0.238	0.23	0.09	0.04	-0	-0
220	36.2	76.7	23	-0.4	-0.2	0.37	0.58	-0.24	-0.25	0.368	0.23	-0.28	-0.28	-0.09	0
1334	33.9	108	24	-0.66	-0.7	0.61	0.64	-0.08	-0.04	-0.15	-0.2	-0.15	-0.07	0.13	0.14
1336	33.9	108	24	-0.62	-0.6	0.74	0.77	-0.24	-0.2	0.052	0.04	0.02	0.02	-0.05	-0.1
215	32.5	93.2	25	0.8	0.81	-0.54	-0.53	0.16	0.09	-0.08	-0.1	-0.1	-0.1	0.09	0.14
30	30.4	97.1	25	-0.75	-0.8	0.23	0.23	0.03	0.05	-0.27	-0.3	0.15	0.16	-0.01	-0.1
197	34.7	85.2	26	0.87	0.86	-0.2	-0.23	-0.11	-0.14	-0.18	-0.2	0.23	0.24	-0.06	-0.1
281	34.2	80.8	26	-0.51	-0.5	0.01	0.03	0.56	0.55	0.131	0.12	0.3	0.3	-0.21	-0.3
37	33.5	86.5	26	-0.03	0.2	-0.67	-0.73	0.09	-0.14	0.116	0.17	0.48	0.27	-0.01	0.12
95	31.7	86.1	26	0.33	-0	-0.85	-0.9	0.17	0.17	0.171	0.15	-0.09	-0.12	0.05	0.08
38	30.3	85	26	-0.18	0.14	0.08	-0.38	-0.35	-0.22	-0.5	-0.5	0.34	0.37	-0.01	-0.1
461	33.8	79	27	0.75	0.76	-0.39	-0.34	-0.29	-0.34	-0.08	-0.1	0.05	0.07	0.12	0.12
239	29.5	82	28	0.43	0.46	-0.2	-0.23	0.23	0.23	0.473	0.44	-0.41	-0.38	0.06	0.17
268	27.6	81.5	28	0.61	0.6	-0.49	-0.54	0.55	0.52	0.063	0.08	0.12	0.09	-0.08	-0.1
147	27.2	81.4	28	0.77	0.75	-0.29	-0.32	-0.11	-0.12	-0.09	-0.1	-0.35	-0.34	-0.01	0.09

				Factor 7		Factor 8		Factor 9		Factor 10		Factor 11		Factor 12	
ID	lat	long	sort	Trees	herbs	Trees	herbs	Trees	herbs	Trees	herbs	Trees	herbs	Trees	herbs
969	75.4	82.5	0	0.06	0.01	-0.2	-0.2	0.16	0.14	0.02	0.03	0.27	0.1	-0.11	0.19
514	68.2	161	1	0.22	0.19	-0.11	-0.1	-0.01	-0	0.05	0	-0	0.02	0.012	0.01
521	67.6	151	1	0.12	0.12	0.03	0.02	0.04	0.07	0.13	0.16	0.09	0	-0.2	0.2
519	67.3	153	1	-0.06	-0.1	-0.01	-0	-0.03	-0	-0.03	-0	-0.1	-0.09	-0.06	-0.01
508	67.1	154	1	-0.3	-0.3	0.12	0.19	0.13	0.11	-0.01	-0	-0.1	-0.2	-0.16	-0.11
506	67.1	154	1	0.19	0.14	-0.32	-0.3	-0.11	-0	0.02	0.06	-0	-0.08	-0.09	0.02
515	66.8	157	1	0.03	-0	-0.11	-0.1	-0.14	-0.1	-0.07	-0	-0	-0.23	-0.34	0.14
520	67.4	150	2	0.06	0.09	0.2	0.27	0.23	0.22	0.11	0.07	-0.1	0.09	0.13	-0.04
518	67.2	142	2	-0	0.01	0.05	0.04	-0.03	-0	0.06	0.08	0.02	0.03	-0.02	0.05
3500	69.3	134	3	0.41	0.43	0.2	0.2	-0.05	-0.1	0.1	0.1	-0.1	-0.15	-0.03	-0.09
3501	69.1	132	3	-0.23	-0.2	0.03	0.08	0.05	0.03	-0	-0	-0	0.02	0.098	-0.07
1551	69.1	133	3	-0.15	-0.2	0.31	0.23	-0.13	-0.2	-0.18	-0.2	0.06	0.03	-0.05	0.12
1259	68.4	138	3	-0.23	-0.2	0.17	0.2	0.32	0.21	0.29	0.19	0.28	0.42	0.081	0.18
1549	68.3	133	3	-0.09	-0.1	-0.01	0.02	0.02	-0	0.01	0	0.03	0.01	-0.04	0.04
1552	66	136	3	-0.33	-0.3	0.01	0.04	0.05	0.03	-0.15	-0.2	-0	-0.02	-0	0.02
522	64.3	151	4	0.01	-0	-0.18	-0.2	0.16	0.15	0.01	0.01	0.07	0.14	0.182	-0.28
854	50	56.2	5	-0.07	-0.1	0.01	-0	-0.03	-0	-0.01	0	0.03	0.01	-0.03	0.03
850	46.6	122	6	-0.07	-0.1	-0.21	-0.2	-0.08	-0.1	0.12	0.16	0.12	-0.03	-0.01	0
445	45.9	121	6	0.05	0.03	-0.06	-0.1	0.03	0.01	-0.06	-0.1	0.06	0.06	0.027	-0.08
857	45.8	122	6	-0.12	-0.1	0.02	0.04	-0.07	-0.1	-0.03	-0.1	-0	0.03	0.21	-0.15
849	48.6	113	7	-0.09	-0.1	0.14	-0	-0.39	-0.5	-0.02	-0	0.24	0.22	0.252	-0.15
1970	46.5	112	7	-0.05	-0	0.26	0.35	0.27	0.25	0.27	0.24	-0.1	0.01	-0.02	-0.27
94	49.4	99.3	8	0.05	0.05	0.08	0.08	-0.2	-0.2	0.18	0.19	-0	-0.09	-0.12	0.14
1085	47.1	93.6	8	-0.02	-0	0.09	0.11	0.02	0.02	0.14	0.16	0.04	0.03	-0.06	0.04
290	45.3	90.1	8	-0.05	-0	-0.03	0.03	0.04	0.1	0.09	0.08	-0.1	-0.05	-0.02	0.02
143	45.1	92.8	8	-0.04	-0	0.03	0.03	0.12	0.11	0.08	0.07	0.1	0.13	0.016	0.07
58	49.2	81.3	9	0.15	0.16	-0.03	-0	-0.02	0.01	0.02	0.04	-0	-0.03	-0.02	-0.02
261	46.6	70.6	10	0.12	0.13	-0.04	-0.1	-0.07	-0	0.09	0.14	0.04	-0.03	-0.16	0.13
78	45.9	70.4	10	-0.21	-0.2	0.09	0.13	0.15	0.1	-0.23	-0.2	-0	-0.06	-0.11	0.11
271	45.6	70.6	10	-0.09	-0.1	-0.07	-0	0.13	0.11	-0.12	-0.1	-0.2	-0.08	0.066	-0.13
266	45.5	73.3	10	0.11	0.07	-0	-0	-0.05	-0.1	0.02	0.03	0.07	0.1	0.072	-0
2339	45.2	77.4	10	0.02	0.03	0.02	0.03	0.03	0.03	-0.12	-0.1	-0.1	-0.08	0.035	-0.12
486	48.9	66.1	11	-0.08	-0.1	0.02	-0	-0.05	-0.1	-0.01	0	0.04	0.03	-0	0.05
16	45.3	67.3	11	-0.04	-0.1	0.12	0.1	0.02	-0	-0.12	-0.2	0.04	0.11	0.128	-0.02
145	45.1	66.7	11	-0.11	-0.1	0.04	0.06	0.01	-0	-0.09	-0.1	-0.1	-0.06	-0.04	0.01
856	49.5	55.5	12	0.01	0.03	0	-0	-0.03	-0	-0	-0	-0	-0	0.02	-0.02
1078	48.3	58.8	12	0.05	0.06	0.03	0.01	-0.04	-0	-0.01	0.01	-0	-0.04	-0.03	-0.01
1069	48.2	58.2	12	0.02	0.01	0	-0	-0.09	-0.1	0.08	0.06	0.03	0.03	0.044	0
866	44.9	110	13	0.11	0.12	0.01	0	0.13	0.12	-0.11	-0.1	0.04	0.07	0.041	-0
884	44.9	111	13	-0.1	-0.1	-0.01	0.02	0.09	0.08	-0.06	-0.1	-0.1	0.04	0.067	-0.05
858	44.7	111	13	0.06	0.07	0.05	0.01	-0.02	-0	-0.14	-0.1	0.03	0.05	0.044	0.02
865	44.2	110	13	0.05	0.06	0.06	0.05	0	-0	-0.09	-0.1	-0	-0.02	-0	0

883	44.1	111	13	-0.01	-0	-0.01	0.01	0.16	0.15	-0.02	-0	-0	0.07	0.077	-0.09
859	43.9	110	13	-0.01	-0.1	-0.21	-0.2	-0.1	-0.1	-0.06	-0	0.09	-0.08	-0.1	0.09
862	43.8	111	13	-0.24	-0.2	0.01	0.01	0.14	0.07	-0.23	-0.2	0.06	0.04	0.023	-0.02
464	44.3	107	14	-0.26	-0.3	-0.08	-0.1	-0	-0	0.11	0.11	0.05	0.05	0.001	0.03
288	45	93.6	15	-0.12	-0.1	-0.07	-0.1	-0.06	-0	0	0.02	-0	-0.07	-0.02	0.01
473	44.8	99.9	15	0.15	0.15	0	0	0.03	0.02	-0.08	-0.1	-0	-0.05	-0.1	0.09
189	43.3	95.2	15	-0.01	0.03	0.02	0.05	0.11	0.11	-0.03	-0	0.01	0.01	0.017	0.02
691	43.2	93.4	15	-0.03	-0	0.04	0.08	0.06	0.09	0.05	0.05	-0.1	-0.04	0.012	-0.03
462	43	91.5	15	0.28	0.29	0.19	0.23	-0	-0	0.2	0.15	-0.1	-0.01	-0.03	0.04
52	42	93.3	15	-0.05	-0	0.12	0.12	0.04	0.02	0.01	0.02	0.04	-0.03	-0.06	0.06
82	44.4	80.3	16	0.17	0.19	-0.06	-0	0.12	0.13	-0.25	-0.2	-0.2	-0.15	0.008	-0.13
1157	44.2	87.8	16	0.08	-0	-0.31	-0.2	0.31	0.42	-0.03	-0.1	0.13	-0.11	-0.16	0.25
92	44.1	87.7	16	-0.05	-0	-0.02	-0	0.07	0.09	-0.02	-0	-0	0	-0	-0.08
1083	43.9	80.4	16	0.13	0.13	0.08	-0	-0.17	-0.2	0.26	0.3	0.18	0.2	0.038	0.14
1155	43.4	89.7	16	0.33	0.32	0.18	0.21	-0.02	-0	0.22	0.16	-0.3	-0.13	0.081	-0.23
72	43.4	89.7	16	0.41	0.48	0.33	0.31	0.02	-0	0.17	0.12	-0	0.12	0.026	-0.01
276	43.4	84.9	16	-0.31	-0.3	0.18	0.18	0.3	0.16	0.07	0.02	0.05	0.11	0.068	-0.08
1056	43.2	81.7	16	-0.14	-0.2	-0.12	-0.1	0.21	0.24	0.23	0.21	0.12	0.16	0.029	-0.03
286	42.4	85.4	16	0.05	0.04	-0.18	-0.2	-0.05	-0.1	0.15	0.13	-0.1	-0.06	0.092	-0.05
87	42.3	83.6	16	-0.19	-0	0.15	0.21	0.02	0.07	-0.14	-0	-0.1	-0.07	-0.02	-0.14
1060	41.9	82.8	16	0.16	0.12	-0.07	-0	0.11	0.11	0.12	0.09	-0	-0.02	-0.07	0.02
50	41.7	86.5	16	0.22	0.19	-0.15	-0.1	-0.15	-0.1	-0.03	-0	-0.2	-0.28	-0.31	0.09
205	41.6	85.3	16	0.15	0.24	-0.04	0.03	0.09	0	0.21	0.09	-0.1	-0.1	-0.09	0.07
925	41.4	84.8	16	0.08	0.16	0.25	0.33	-0.06	-0	-0.14	-0.1	-0.3	-0.3	-0.07	0.07
241	40.4	83.8	16	-0.01	-0.1	-0.19	-0.2	-0.04	-0	-0.04	-0.1	0.08	0.15	0.093	0.01
265	40.4	82.5	16	-0	-0.1	0.1	-0.1	-0.26	-0.2	-0.12	-0.2	-0.1	0.18	0.138	-0.05
924	40.1	83.5	16	-0.3	-0.2	-0.32	-0.1	-0.05	0.19	0.16	0.19	-0.2	-0.17	-0.04	-0.07
467	44.9	77.2	17	0.08	0.1	-0.02	-0.1	-0.13	-0.1	0.12	0.15	0.05	0.07	0.079	-0.05
104	44.2	74	17	0.07	0.06	0.01	0	-0	0	-0.04	-0	-0	-0.08	-0.12	0.05
1077	44.2	79.8	17	0.16	0.17	0.03	-0	-0.01	-0	0.02	-0	0.08	0.17	0.172	-0.03
1058	44.1	78.1	17	0.36	0.33	0.17	0.12	-0.04	-0.1	0.15	0.11	0.1	0.15	0.05	0.07
68	44	71.8	17	0.01	0.03	-0.08	-0.1	0.02	0.12	0.07	0.09	-0	-0.09	-0.14	0.12
448	44	70.4	17	-0.15	-0.2	0.13	0.11	-0.04	-0.1	0.13	0.1	0.11	0.15	0.13	-0.01
275	44	79.4	17	-0.23	-0.3	0.01	0.01	-0.01	-0	-0.04	-0.1	0.11	0.06	0.013	0.04
228	42.8	71.7	17	0.13	0.13	0.03	0	-0.01	-0	0	0.01	0.02	0.04	0.021	-0.03
13	42.5	74.6	17	-0.28	-0.3	0	0.01	0.01	0.03	-0	-0	-0.1	-0.01	-0.01	0.01
7	42.3	78.9	17	-0.13	-0	0	0.04	-0.02	0.04	0.28	0.25	-0.1	-0.12	-0.03	-0.03
18	42.3	77.9	17	0.06	0.04	-0.21	-0.2	-0.11	-0	0.19	0.23	0.13	0.07	-0.02	0.06
105	42.2	74	17	-0.05	-0.1	-0.12	-0.1	0	0.02	0.16	0.17	0.07	0.06	0.038	-0.07
77	41.9	70	17	0.05	0.01	0.06	0.08	0.11	0.1	-0	-0	-0	-0.08	-0.11	0.11
247	41.7	76.4	17	0.33	0.34	0.09	-0.1	-0.11	-0.3	-0.17	-0.2	0.34	0.25	-0.05	0.06
222	41.4	72.1	17	0.12	0.09	0	-0.1	0.02	-0	-0.13	-0.1	0.11	0.11	0.184	-0.17
199	41.3	74.8	17	-0.24		-0.43		0.02		0.04		-0.3		0.217	
690	41.3	70.2	17	0.24	0.25	0.14	0.1	-0.07	-0.1	0.01	0.01	0.05	0.07	0.071	0.04
585	41.2	71.6	17	0.26	0.26	0.03	-0	-0	-0	-0.06	-0	0.12	0.14	0.043	0.09
259	41	75.3	17	0.19	0.24	0.09	0.12	0.24	0.27	-0.08	-0.1	-0.1	0.05	-0.12	0.27

192	40.8	77.4	17	0.06	0.06	-0.05	-0.1	0.05	0.07	-0.08	-0.1	0.08	0.05	0.003	0.06
153	40.5	75.7	17	0.19	0.15	-0.05	-0.1	0.18	0.1	0.03	-0	0.1	0.19	0.201	-0.21
257	40.4	74.5	17	-0.07	-0.1	-0.16	-0.2	-0.48	-0.4	-0.01	0.03	-0.1	-0.26	0.05	0.01
447	45	69.3	18	0.19	0.2	-0.06	-0	0.06	0.07	-0.1	-0.1	-0.1	-0.01	0.086	-0.07
174	44.6	68.6	18	0.11	0.11	0.1	0.05	-0.11	-0.2	-0.01	0	0.05	0.06	0.091	-0.07
240	44.6	63.6	18	-0.06	-0.1	0.08	0.05	-0.16	-0.2	0.16	0.15	0.06	0.07	0.034	0
443	43.9	69.4	18	0.14	0.17	-0.09	-0	0.21	0.24	-0.09	-0.1	-0.1	-0.06	-0.07	0.05
1094	38	120	19	0.03	0.05	-0.31	-0.3	0.18	0.13	-0.07	-0.1	0.07	0.18	-0	0.08
695	37.6	119	19	-0.1	-0.1	-0.04	-0.1	-0.1	-0.1	-0.07	0.02	0.07	-0.01	-0.04	-0.02
882	37	120	19	-0	-0	-0.15	-0.1	-0.07	-0	0.18	0.19	0.04	0.01	0.001	-0.02
1313	37.7	106	20	-0.03	-0	0.06	0.05	-0.28	-0.3	-0.14	-0.1	-0.1	-0.09	0.068	-0.04
880	37.6	106	20	-0.1	-0.1	0	-0	-0.01	-0	-0.1	-0	0.18	0.06	-0.14	0.24
868	38.5	98.7	21	-0.07	-0.1	0.06	0.04	-0.01	-0	0.03	0.01	0.01	0.03	0.01	0.06
64	36.8	91.1	21	-0.2	-0.2	0.12	0.11	-0.28	-0.2	0.14	0.17	0.04	-0.08	-0.16	0.16
441	36.6	90.6	21	0.13	0.09	0.1	0.2	0.04	0.08	0.04	0.07	-0	-0.07	-0.11	0.07
57	38.2	80.3	22	0.13	0.15	0.23	0.22	-0.05	-0.1	-0.21	-0.2	0.08	0	-0.05	0.11
115	37.5	85.7	22	-0.29	-0.2	0.11	0.1	-0.02	0.01	0.02	0.05	-0.1	-0.09	-0.08	0.01
8	36	85.5	22	-0.14	-0.2	-0.05	-0	-0.15	-0.1	0.14	0.16	0.03	-0.05	-0.11	0.06
60	40	77.6	23	-0.25	-0.2	0.18	0.22	0.01	-0	0.1	0.05	-0.1	0.09	0.281	-0.29
34	39.6	79.3	23	0.02	0	-0.05	-0	-0.07	-0	0.09	0.11	-0.1	-0.09	-0	-0.03
1148	38.2	79.6	23	0.28	0.25	-0.24	-0.2	0.2	0.2	-0.15	-0.2	-0.1	0.09	0.222	-0.18
390	38	79.1	23	0.28	0.29	0.05	0.12	0.07	0.1	-0.03	-0	-0.1	-0.21	-0.01	-0.07
100	38	79	23	-0.1	-0.1	0.26	0.33	0.13	0.11	0.17	0.13	-0.1	-0.01	-0.18	0.19
220	36.2	76.7	23	0.01	-0.1	0.21	0.2	0.03	-0.1	-0.4	-0.4	0.32	0.24	-0.14	0.29
1334	33.9	108	24	0.12	0.11	0.11	0.06	-0.07	-0.1	-0.24	-0.2	-0	-0.03	-0.01	0.07
1336	33.9	108	24	-0.01	-0	0.01	-0	-0.02	-0	0.01	0.01	-0	-0.03	-0.03	0.03
215	32.5	93.2	25	-0.02	-0	0.03	0.05	0.01	-0	-0.07	-0.1	0.02	0.04	0.021	-0.01
30	30.4	97.1	25	-0.24	-0.2	0.31	0.31	-0.03	-0.1	0.13	0.12	0.03	-0.06	-0.2	0.1
197	34.7	85.2	26	0.21	0.19	-0.08	-0.1	0.01	0.03	0.09	0.08	0.06	0.04	0.016	0.03
281	34.2	80.8	26	0.2	0.17	0.25	0.25	0.13	0.01	0.06	0.02	-0	0.21	0.155	-0.13
37	33.5	86.5	26	0.25	0.13	-0.22	-0.3	-0.04	0.08	0.07	0.07	-0	-0.04	-0.06	-0.05
95	31.7	86.1	26	0.01	0.03	0.05	0.07	0.13	0.12	-0.15	-0.2	0.04	0.01	-0.02	0.1
38	30.3	85	26	0.12	0.19	-0.16	-0.1	-0.15	-0.1	-0.16	-0.2	-0.1	0.05	0.193	-0.14
461	33.8	79	27	-0.08	-0.1	0.2	0.14	-0.06	-0.1	0.1	0.03	0.1	0.2	0.175	-0.04
239	29.5	82	28	0.21	0.18	-0.19	-0.2	0.19	0.14	0.13	0.16	0.21	0.22	0.024	0.13
268	27.6	81.5	28	0.05	0.06	-0.02	-0	-0.02	-0	-0.13	-0.1	-0.1	-0.04	0.075	-0.1
147	27.2	81.4	28	0.1	0.12	0.2	0.21	0.08	0.06	-0.11	-0.1	-0.1	0.01	0.002	0.01

Table F5. Summary of total explained variance for all 9 PCA runs on the subgroups and the original 47-type pollen sums.

Category	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Factor10	Sum
47 types (E)	25	11	7	5	4	4	3	3	2	2	66
24 types (C)	30	13	8	5	4	4	3	3	2	2	74
23 types (D)	16	9	7	5	4	4	4	3	3	3	58
18 types (F)	31	14	8	5	4	4	3	2	2	2	75
6 types (G)	30	11	8	6	5	4	3	3	3	2	75
11 types (H)	19	9	7	5	4	4	3	3	3	3	60
12 types (I)	14	9	7	5	5	4	4	4	3	3	58
All trees (A)	42	17	9	5	4	3	3	2	2	2	89
All herbs (B)	42	17	8	5	4	4	3	2	2	2	89
LIA*	18	11	7	6	6	5	5	4	4	4	70

Note: Values expressed in percentages.

* Little Ice Age analysis at decadal resolution

Table F6. Summary of explained variance for each individual taxon PCA.

Varid	Taxa	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
v1	*Abies*	37	15	8	6	5	3	3	2	2	2
v2	*Acer*	34	11	6	6	5	4	3	3	2	2
v3	*Alnus undiff.*	29	15	8	6	5	5	4	3	2	2
v4	*Ambrosia type*	43	13	7	5	4	3	3	3	2	2
v5	*Apiaceae*	13	9	8	7	6	5	4	4	4	3
v6	*Aquifoliaceae*	18	11	9	6	6	6	5	4	4	4
v7	*Artemisia*	41	11	7	5	4	3	3	2	2	2
v8	*Asteraceae*	31	12	7	6	5	4	3	3	3	2
v9	*Betula undiff.*	36	14	10	6	4	3	3	2	2	2
v10	*Carya*	34	12	7	6	4	4	3	3	2	2
v11	*Castanea*	26	9	8	6	5	5	4	4	3	3
v12	*Cephalanthus*	21	12	10	9	7	6	5	4	4	4
v13	*Chenopodiaceae/Amaranthaceae*	17	13	9	7	5	5	4	3	3	3
v14	*Corylus*	23	12	8	5	4	4	4	3	3	3
v15	*Cupressaceae*	23	13	8	7	4	4	3	3	3	3
v16	*Cyperaceae*	38	14	7	5	4	3	3	3	2	2
v17	*Ericaceae undiff.*	19	12	6	6	6	5	5	4	4	3
v18	*Fagus*	47	11	6	4	4	4	3	3	2	2
v19	*Fraxinus*	26	10	8	5	5	4	4	3	3	3
v20	*Larix/Pseudotsuga*	26	12	8	7	4	4	3	3	3	2
v21	*Liliaceae*	14	13	9	7	6	6	5	4	4	4
v22	*Lycopodium undiff.*	24	9	7	6	5	4	4	4	3	3
v23	*Magnoliaceae*	25	14	12	10	8	8	6	5	4	3
v24	*Myrica*	19	10	9	7	6	6	4	4	4	3
v25	*Nyssa*	20	14	9	6	6	5	5	4	3	3
v26	*Ostrya/Carpinus*	29	14	8	6	5	4	3	3	3	2
v27	*Picea undiff.*	63	11	7	4	3	2	1	1	1	1
v28	*Pinus undiff.*	34	24	12	5	3	3	2	2	2	1
v29	*Plantaginaceae*	32	9	7	5	5	5	4	4	3	3
v30	*Poaceae*	33	15	9	8	5	3	3	3	3	2
v31	*Polygonum*	14	11	9	7	7	6	5	5	4	3
v32	*Polypodiaceae*	23	12	9	6	5	5	4	3	3	3
v33	*Populus*	25	12	9	6	5	4	4	3	3	3
v34	*Pteridium*	17	10	8	7	5	5	4	4	4	3
v35	*Quercus*	36	19	9	5	4	3	3	2	2	2
v36	*Ranunculaceae*	13	8	8	6	6	5	5	5	4	4
v37	*Rubiaceae*	17	10	9	8	6	6	5	5	4	4
v38	*Rumex*	32	13	11	7	4	3	3	3	2	2
v39	*Salix*	36	10	7	5	4	4	3	3	2	2
v40	*Sarcobatus*	15	10	8	7	6	5	5	4	4	4
v41	*Saxifragaceae*	19	11	10	7	7	6	5	5	4	4
v42	*Shepherdia*	22	12	9	8	6	5	5	4	4	3

v43	*Sphagnum*	20	15	7	6	5	4	4	4	3	3
v44	*Taxodium*	36	19	14	11	8	5	4	3	1	0
v45	*Tilia*	19	12	8	6	5	4	4	3	3	3
v46	*Tsuga*	32	17	6	6	4	4	3	3	2	2
v47	*Ulmus*	23	19	8	5	4	4	3	3	3	2

Note: Values represent explained variance expressed in percentages

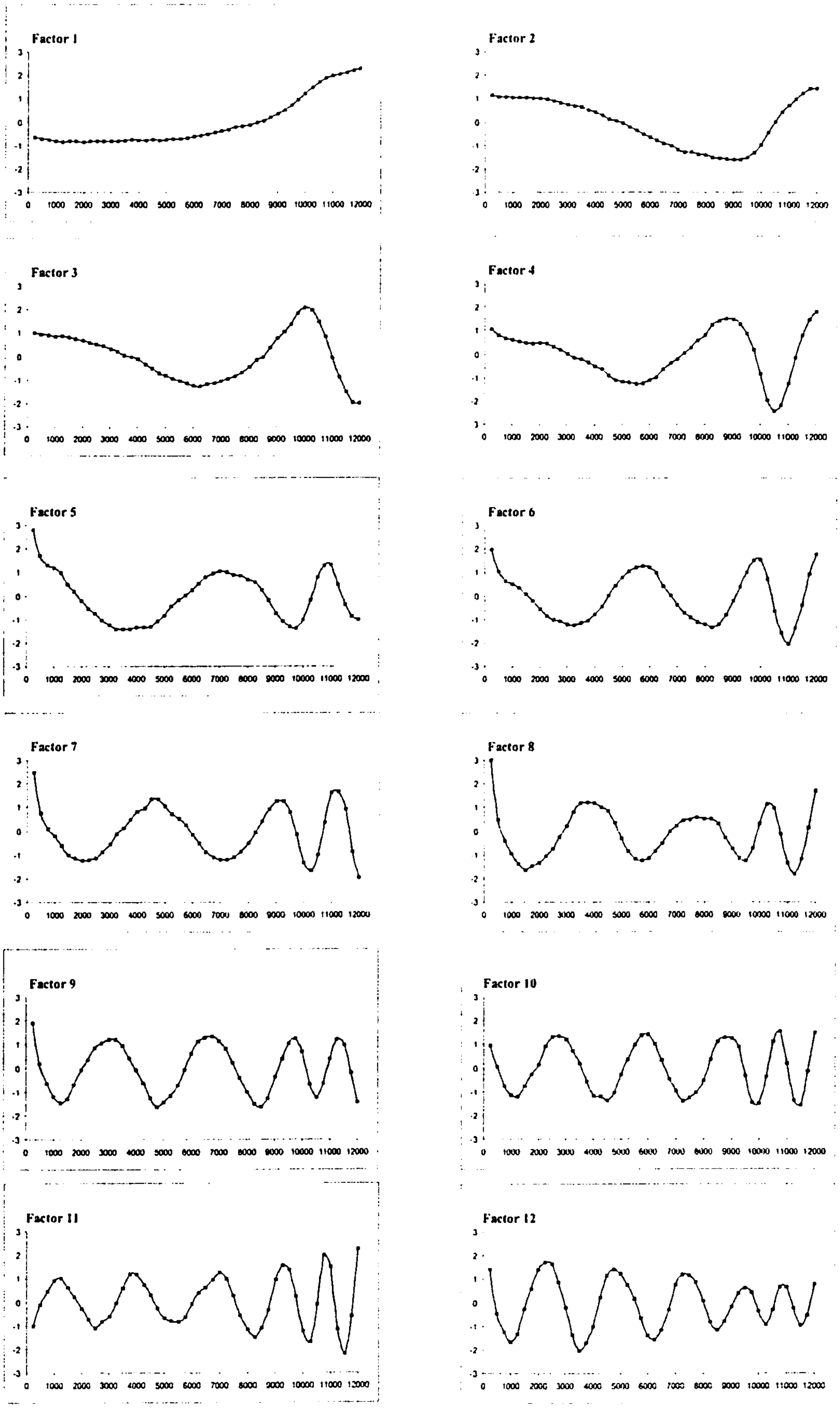


Figure F1.1: PCA on 89 and 47-types pollen sums. X-axis = time (C14); Y-axis= PCA scores.

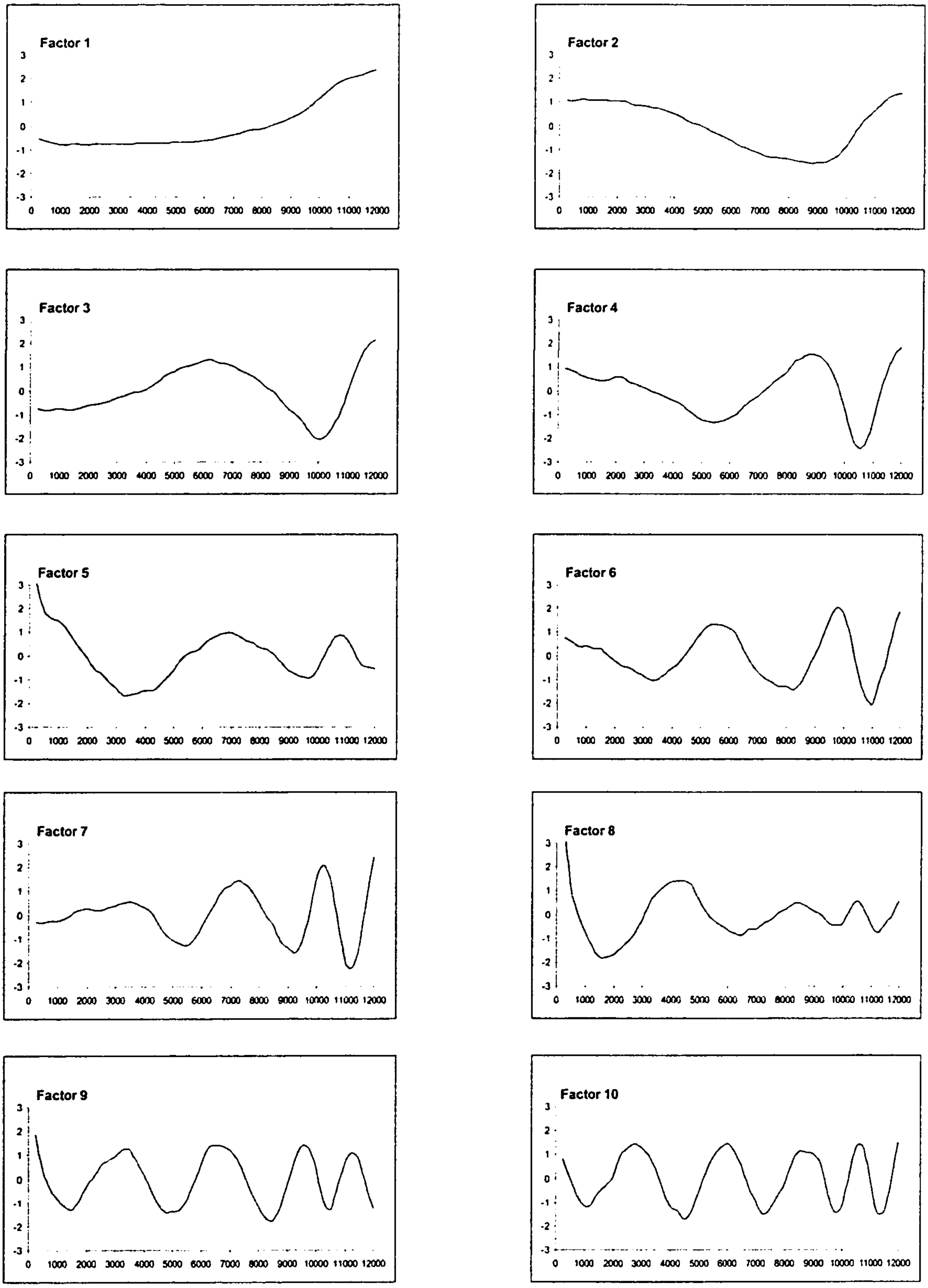


Figure F1.2: PCA on 24-types pollen sum (C in Table F2). X-axis = time (C14); Y-axis= PCA scores.

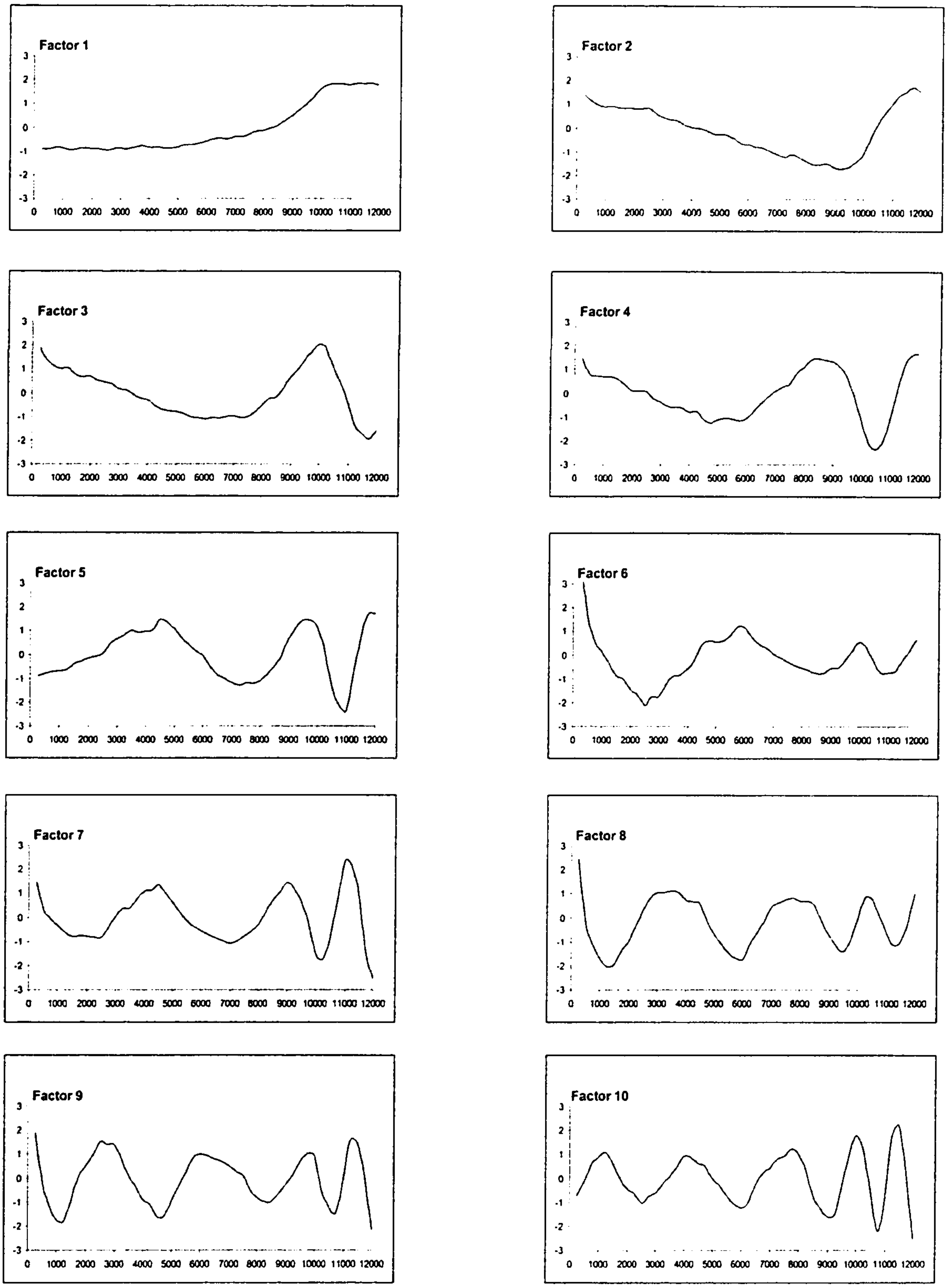


Figure F1.3: PCA on 23-types pollen sum (D in Table F2). X-axis = time (C14); Y-axis= PCA scores.

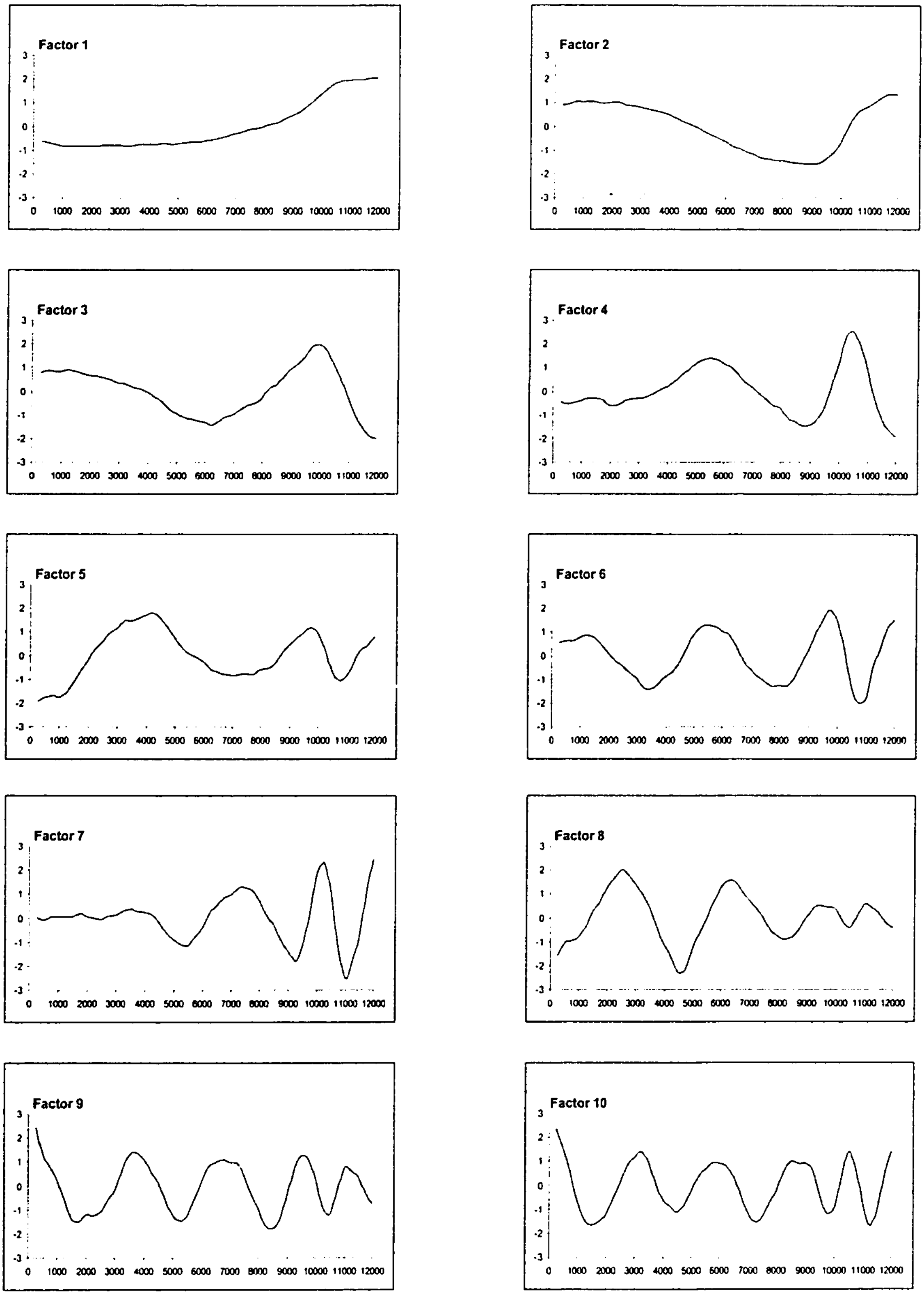


Figure F1.4: PCA on 18-types pollen sum (F in Table F2). X-axis = time (C14); Y-axis= PCA scores.

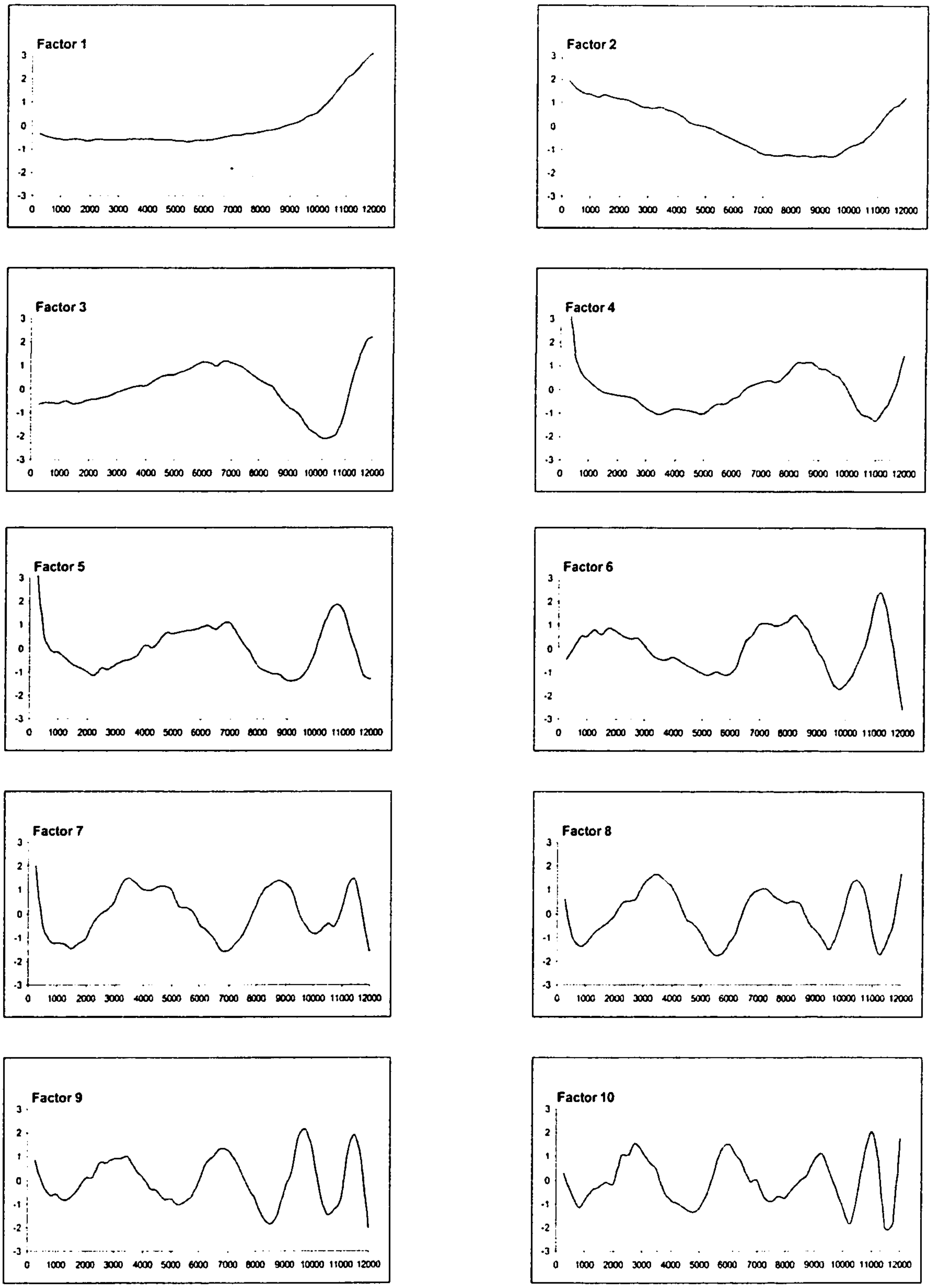


Figure F1.5: PCA on 6-types pollen sum (G in Table F2). X-axis = time (C14); Y-axis= PCA scores.

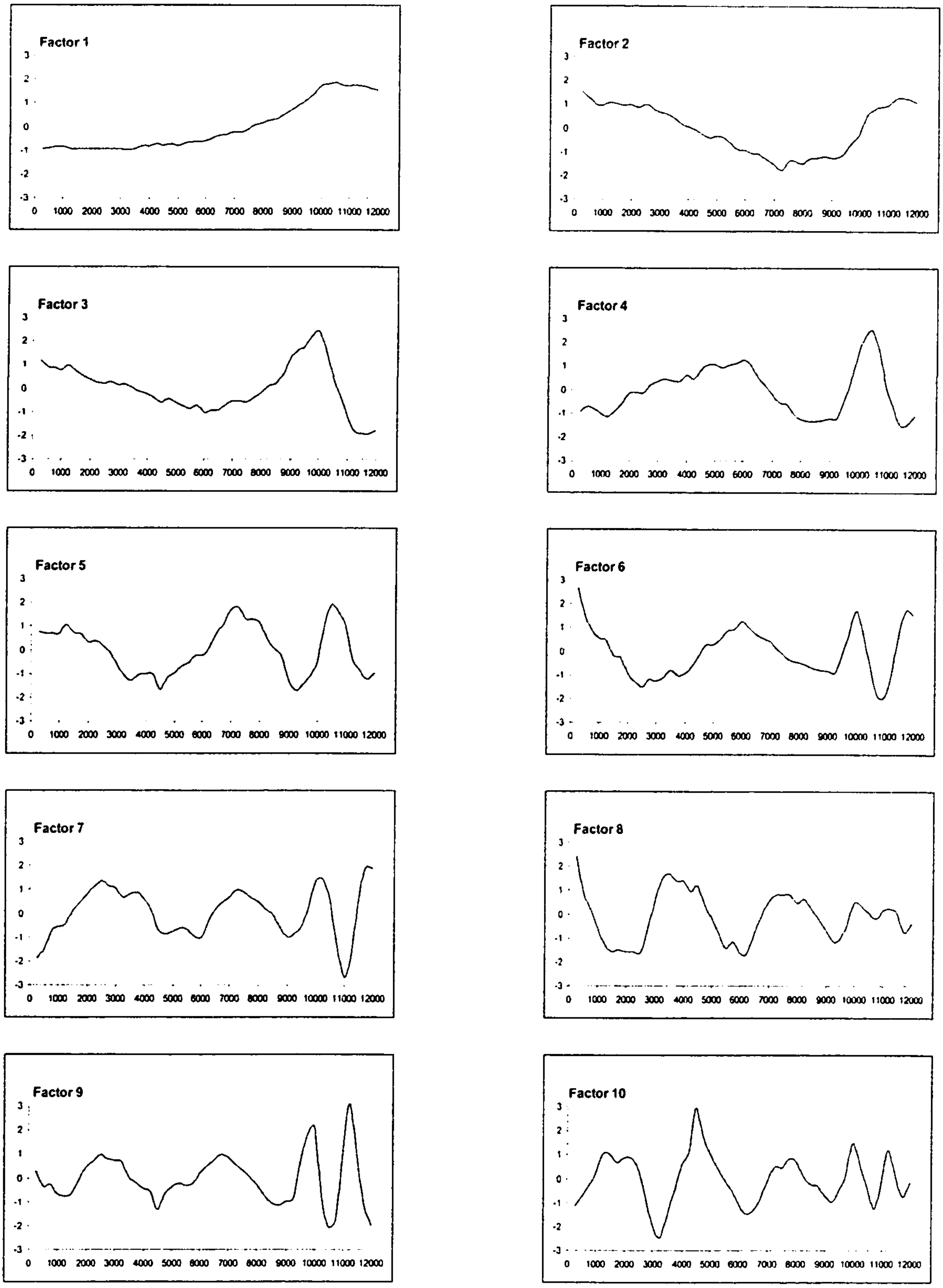


Figure F1.6: PCA on 11-types pollen sum (H in Table F2). X-axis = time (C14); Y-axis= PCA scores.

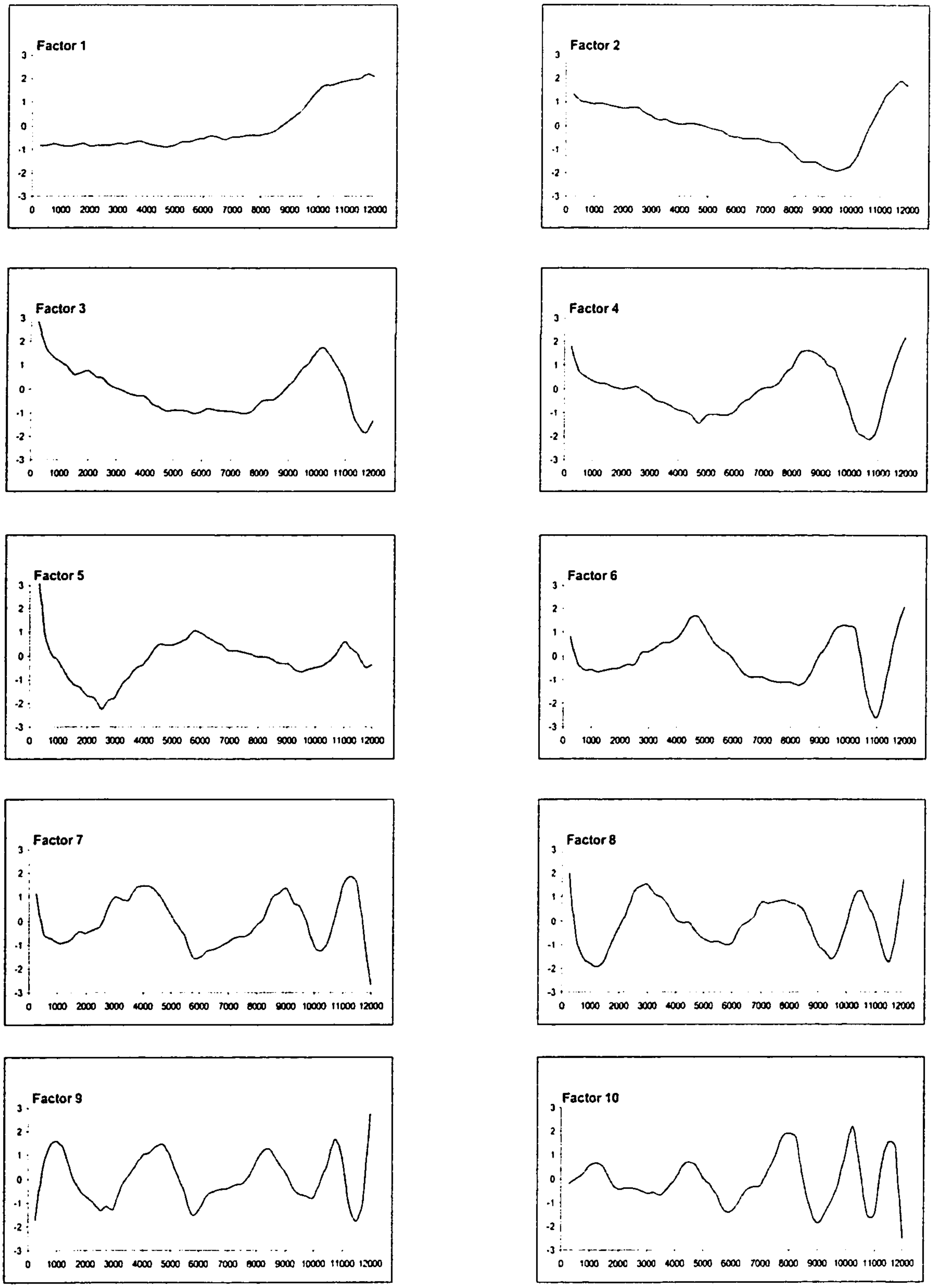


Figure F1.7: PCA on 12-types pollen sum (I in Table F2). X-axis = time (C14); Y-axis= PCA scores.

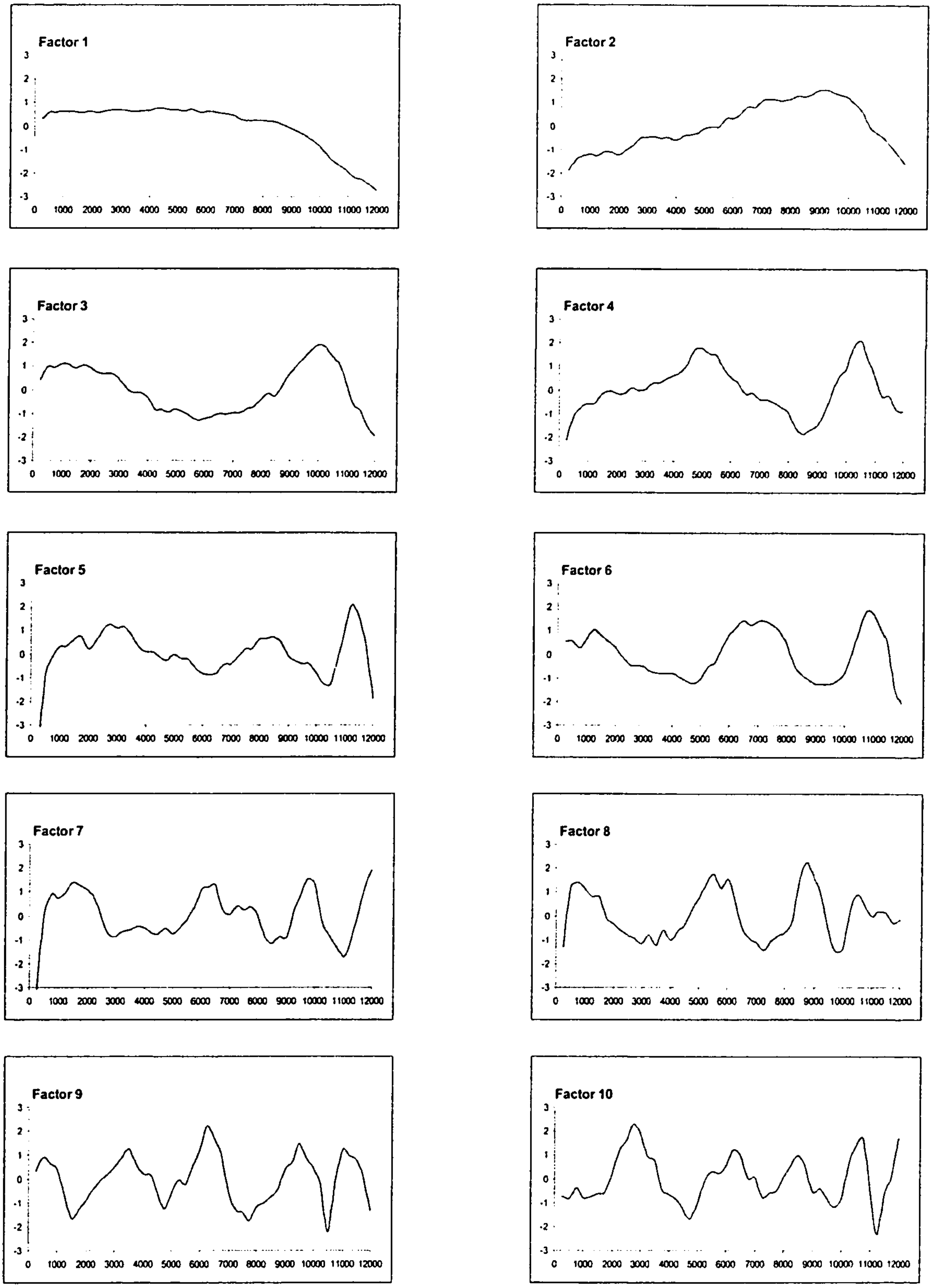


Figure F1.8: PCA on all trees pollen sum through space and time (A in Table F2). X-axis = time (C14); Y-axis= PCA scores.

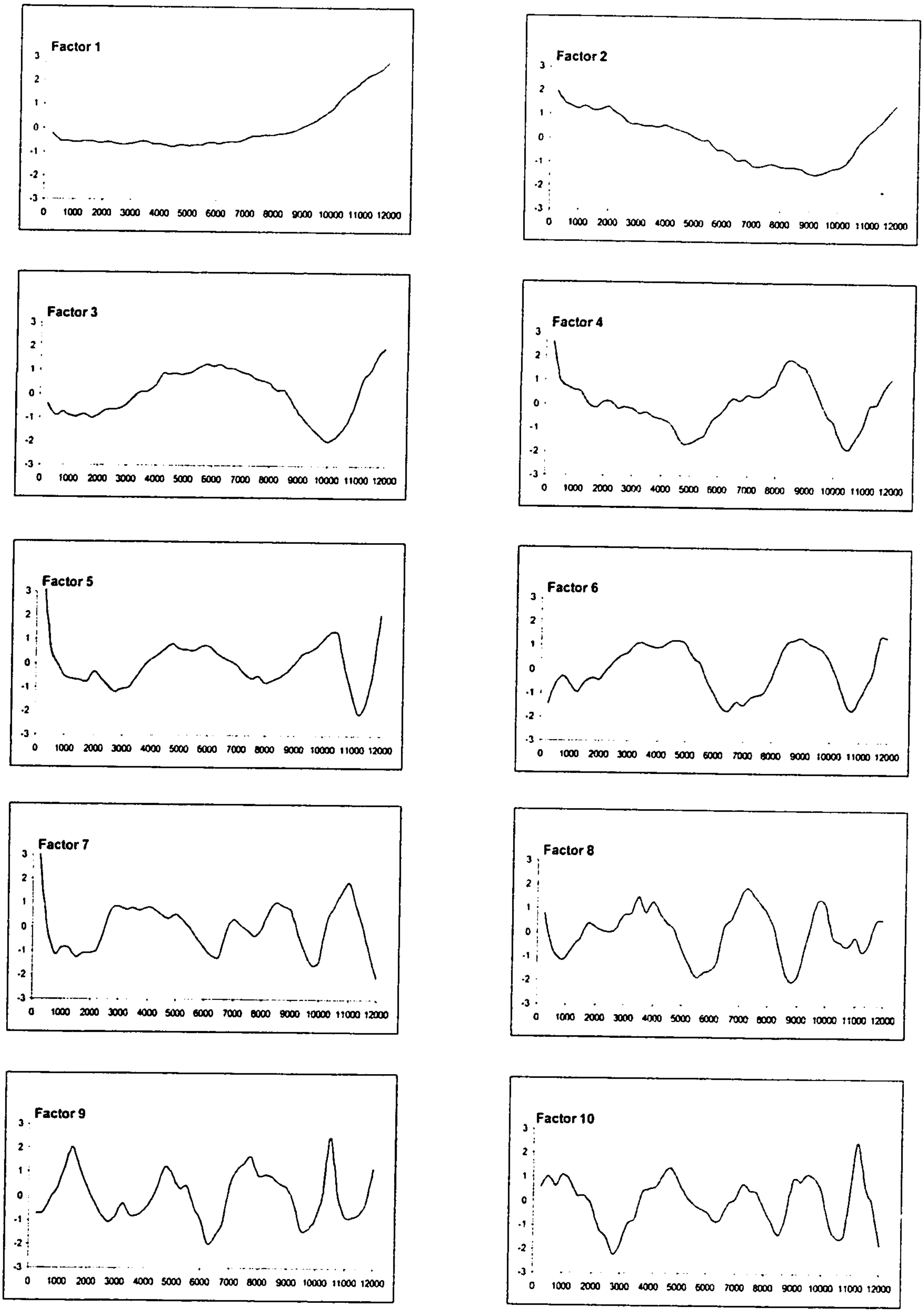


Figure F1.9: PCA on all herbs pollen sum through space and time (B in Table F2). X-axis = time (C14); Y-axis= PCA scores.

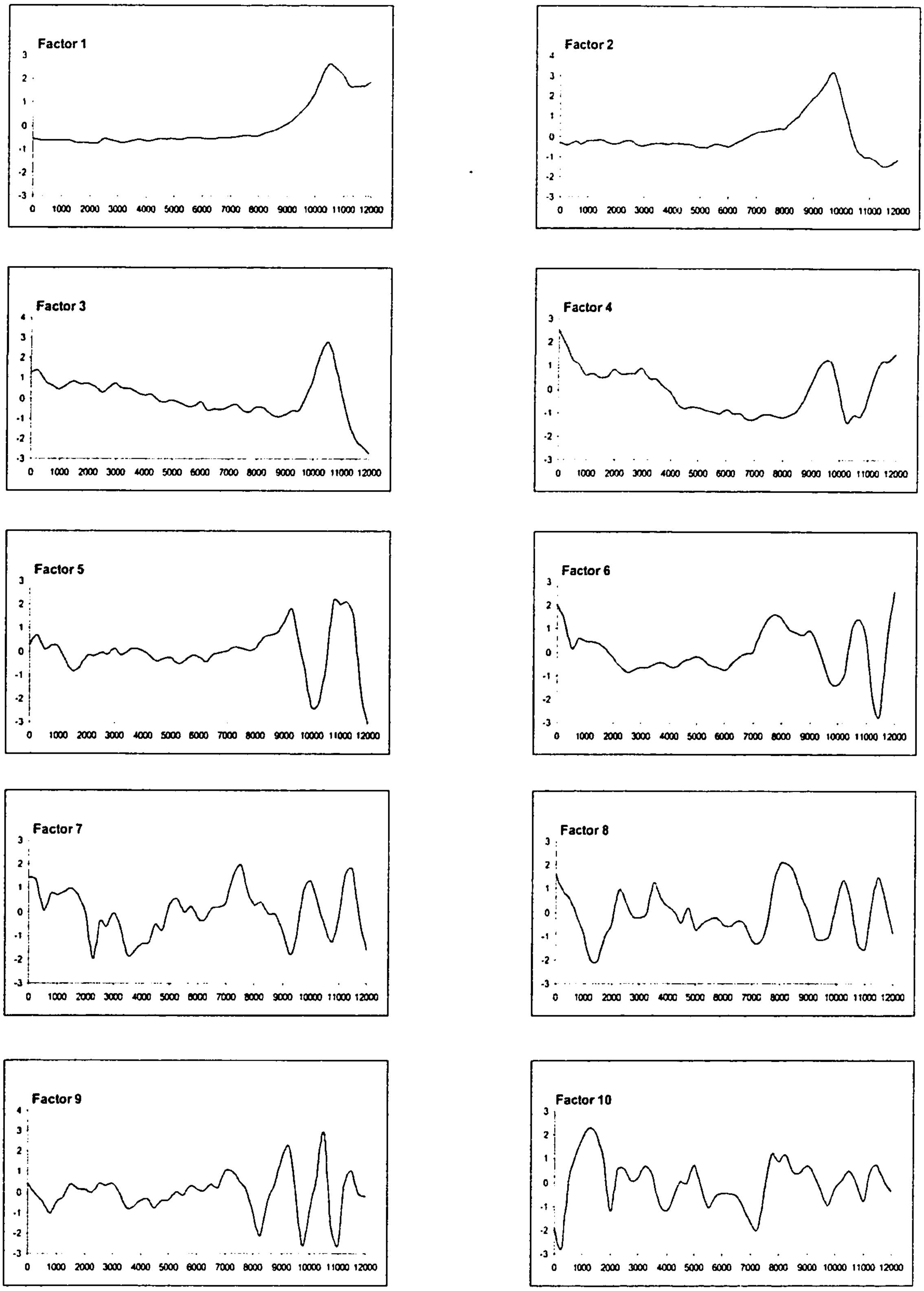


Figure F2.1: PCA on all *Abies* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

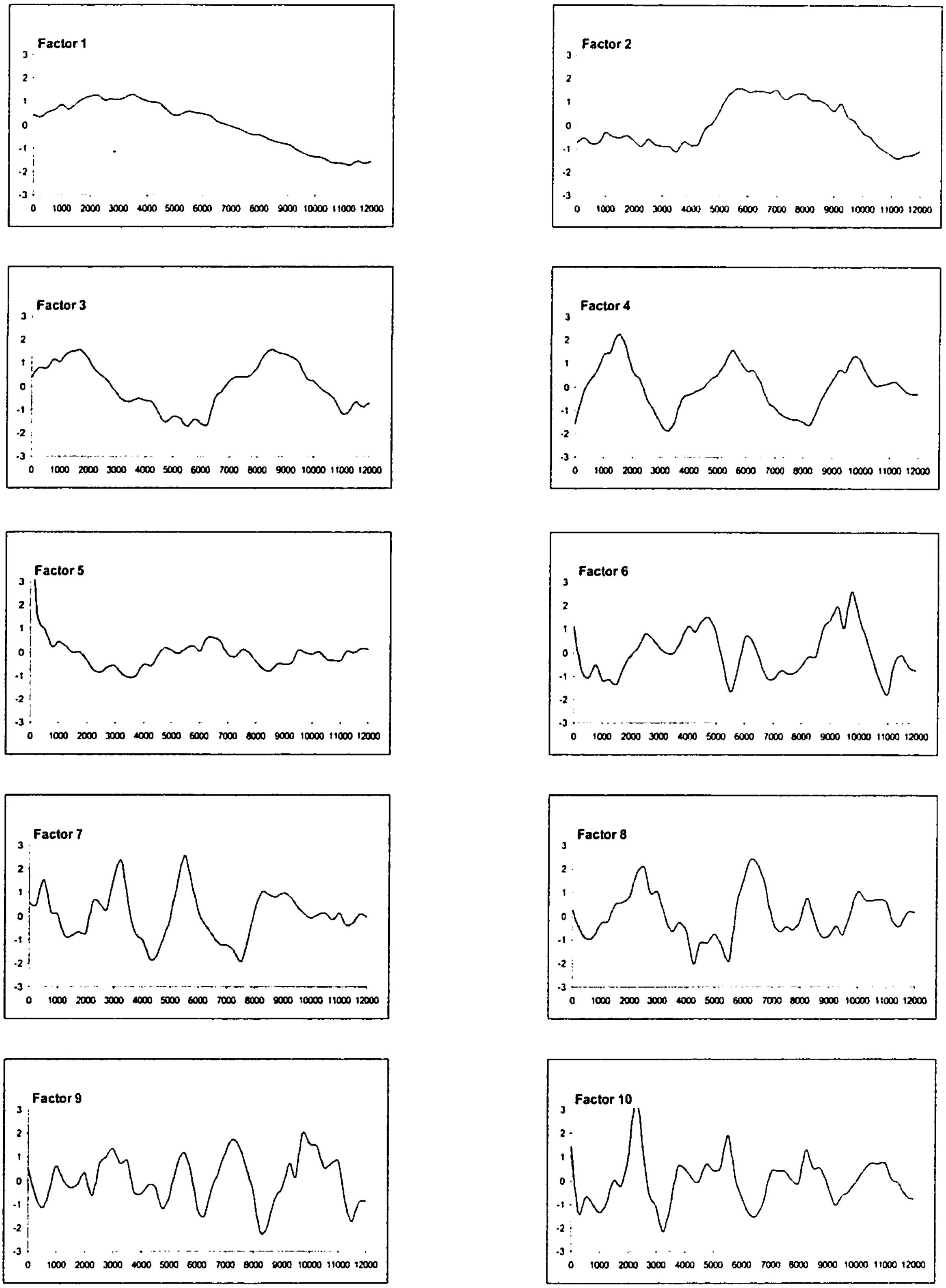


Figure F2.2: PCA on all *Acer* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

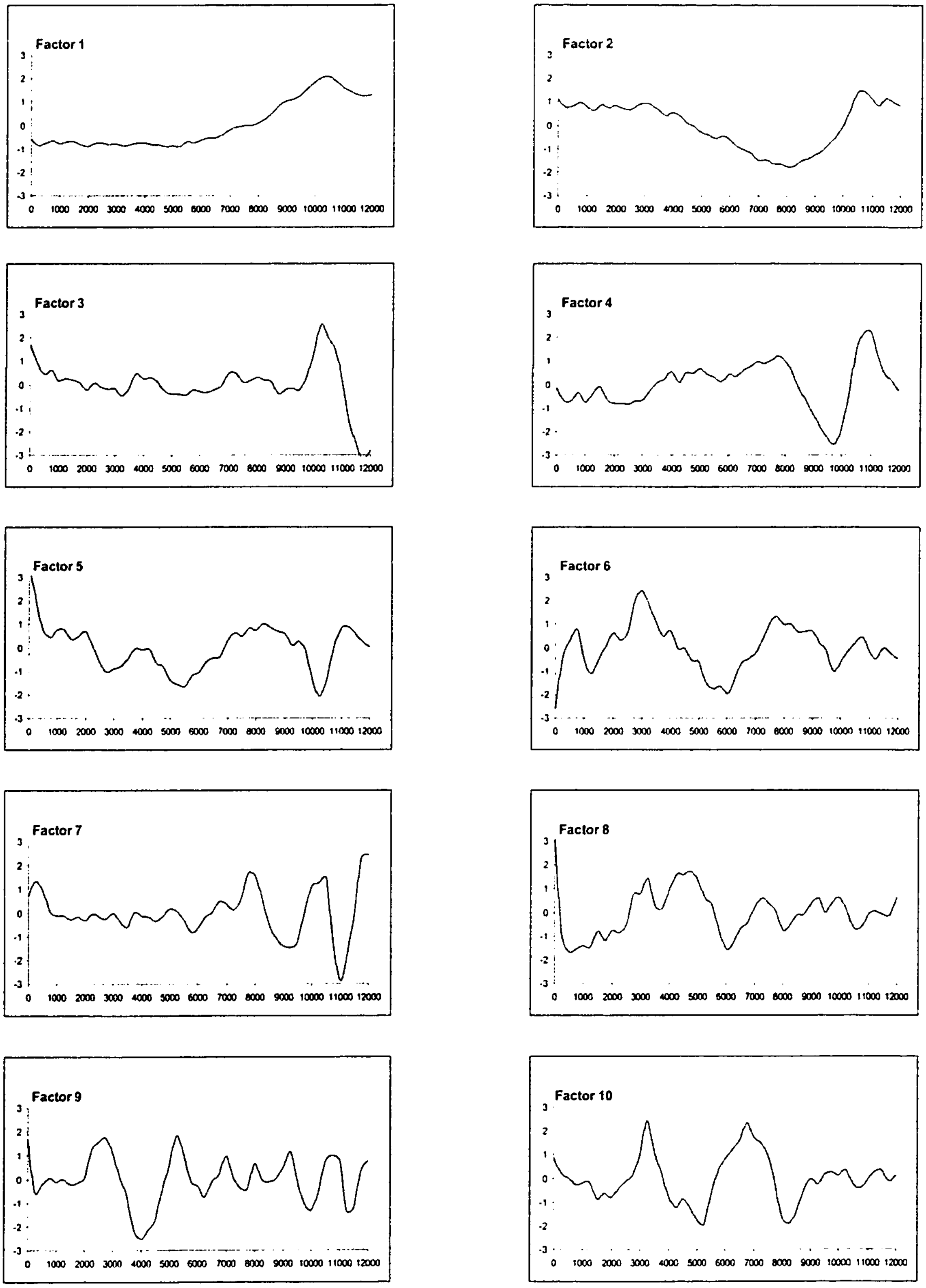


Figure F2.3: PCA on all *Alnus undiff.* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

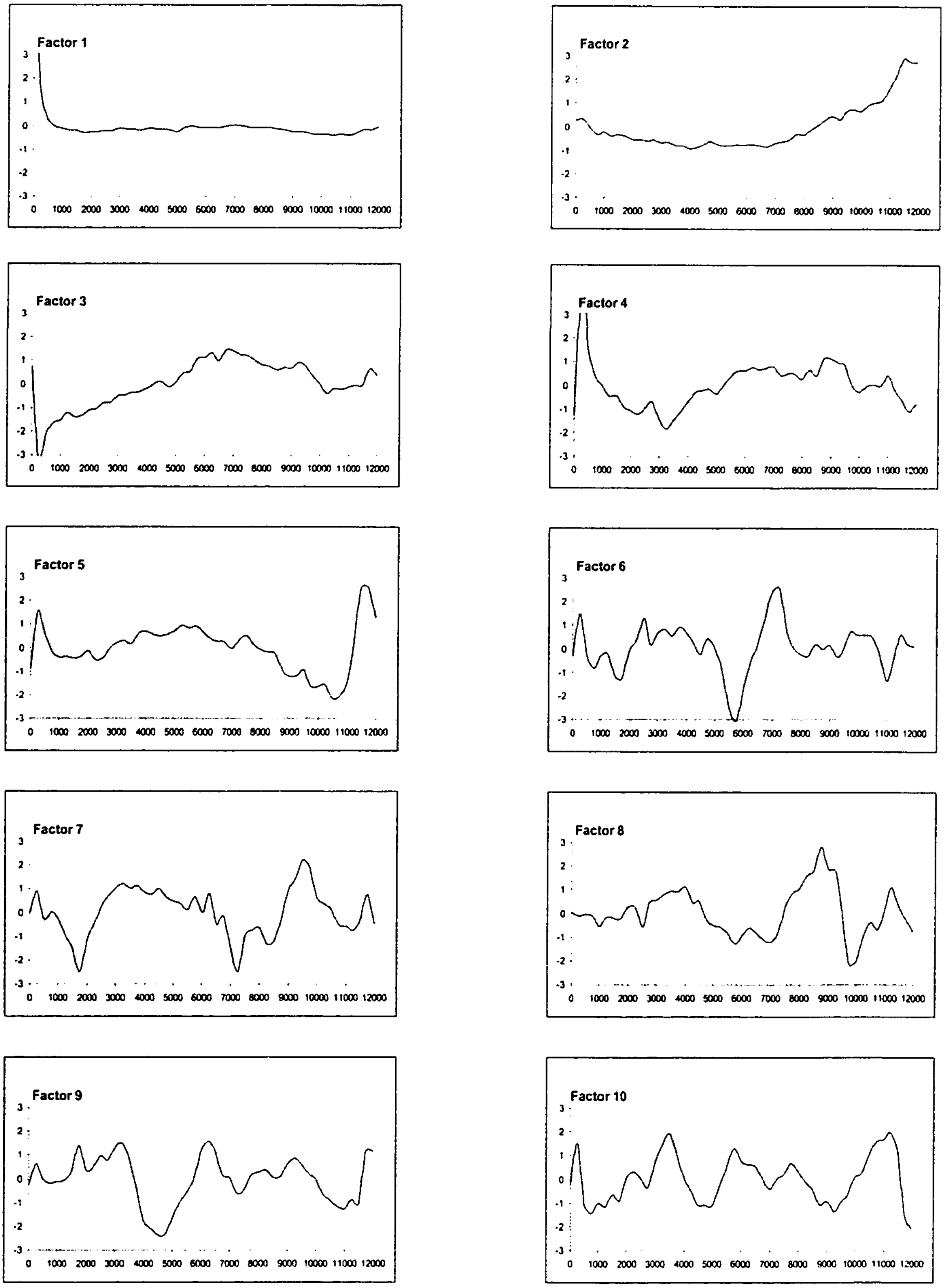


Figure F2.4: PCA on all *Ambrosia* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

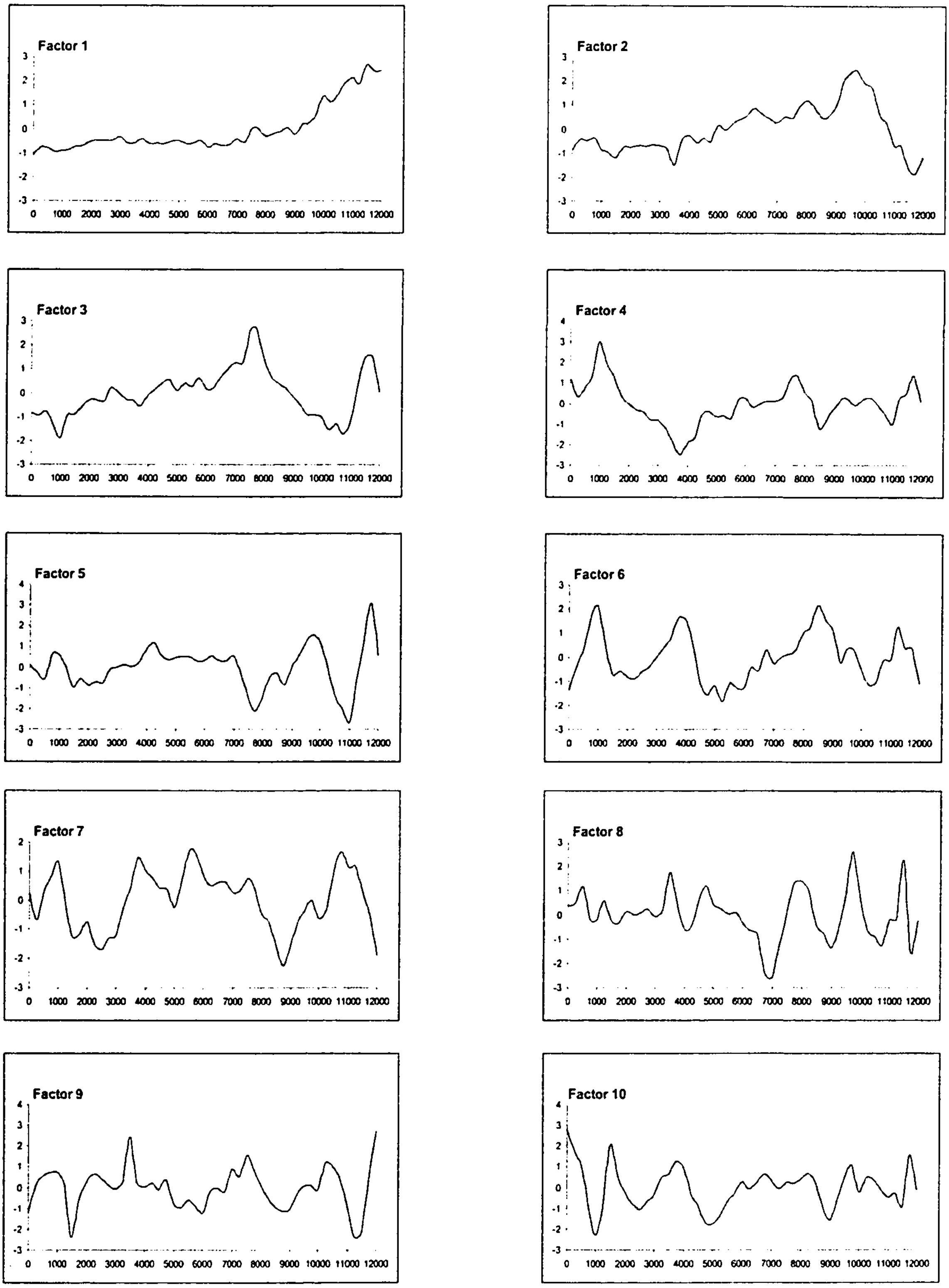


Figure F2.5: PCA on all *Apiaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

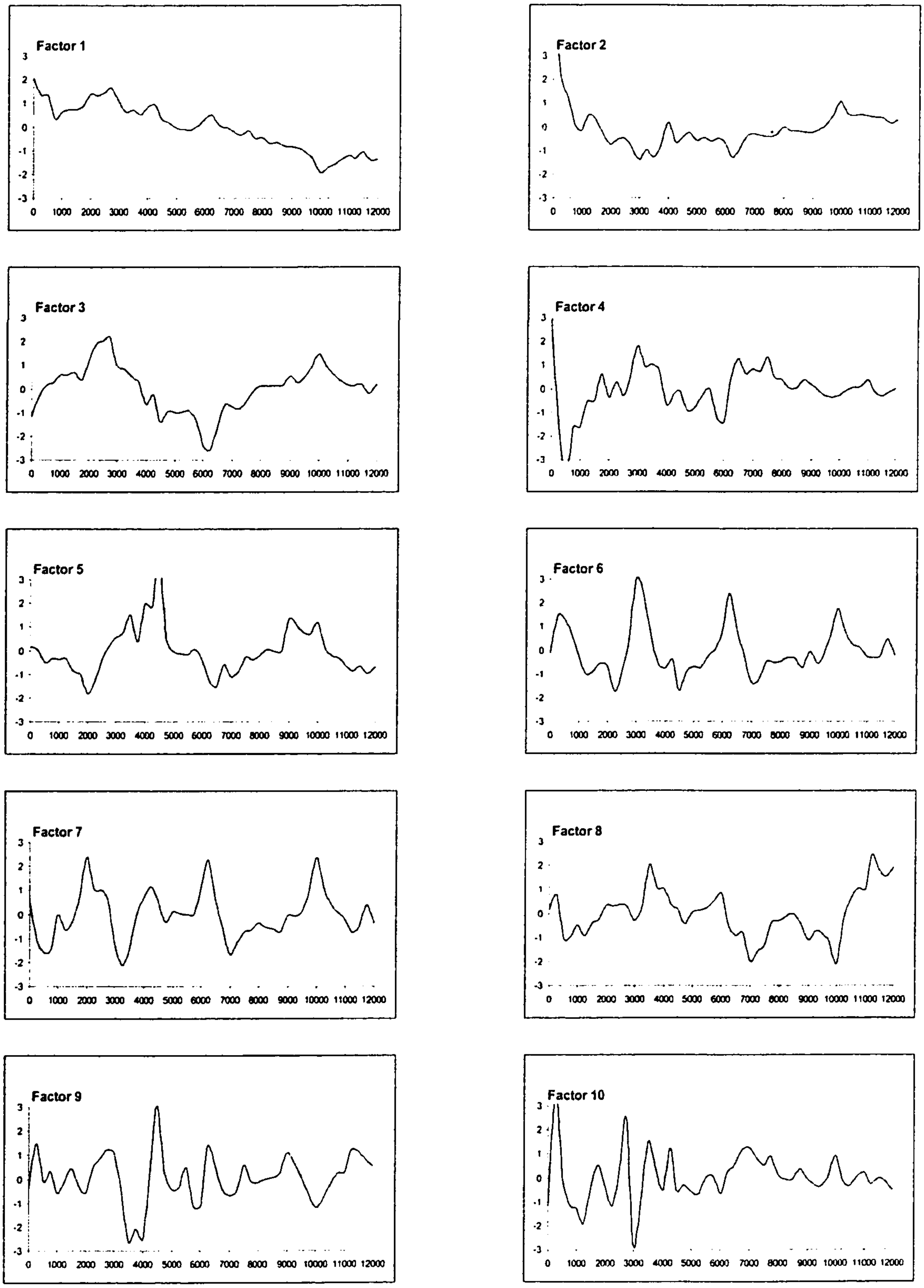


Figure F2.6: PCA on all *Aquifoliaeae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

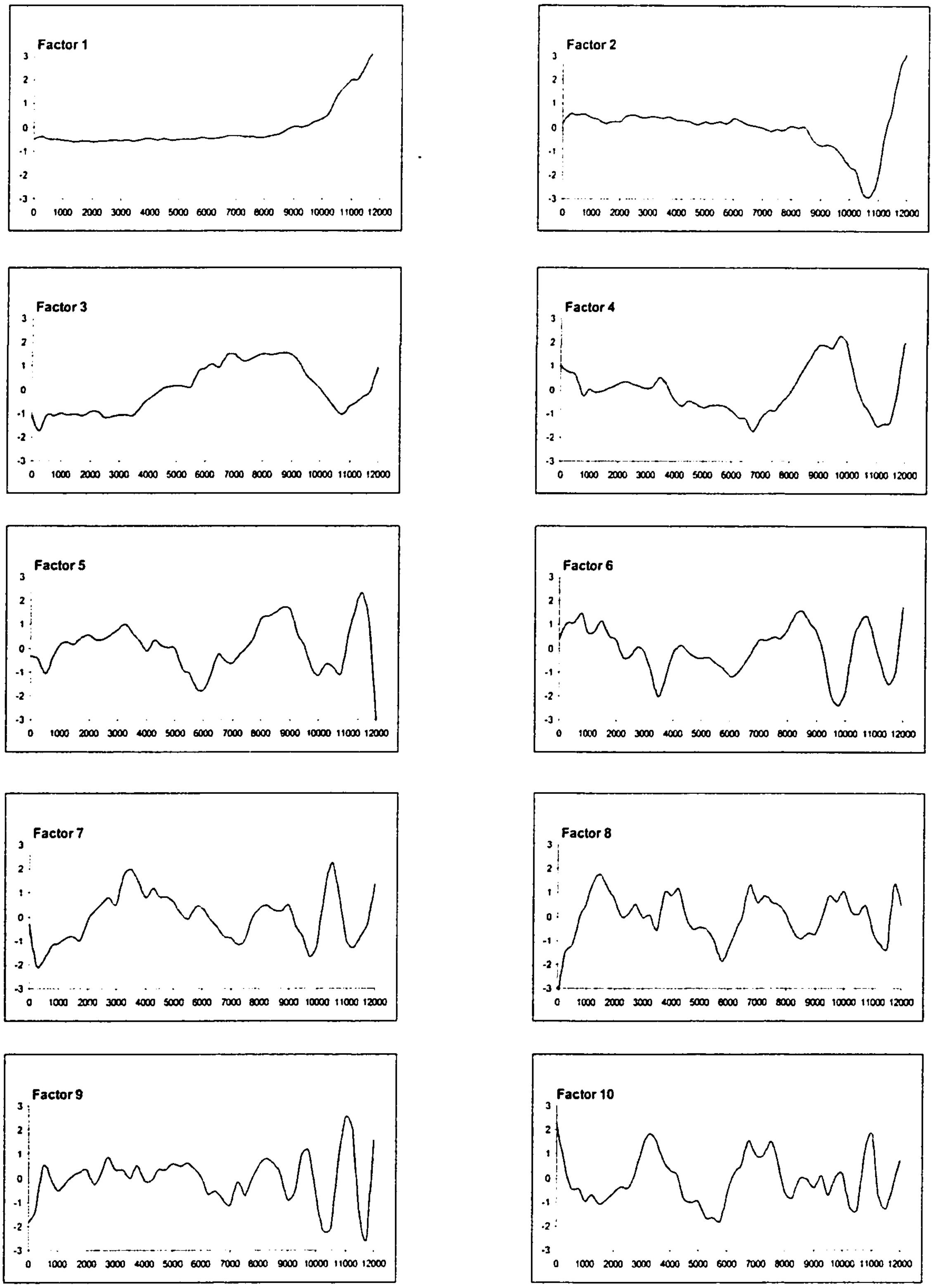


Figure F2.7: PCA on all *Artemisia* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

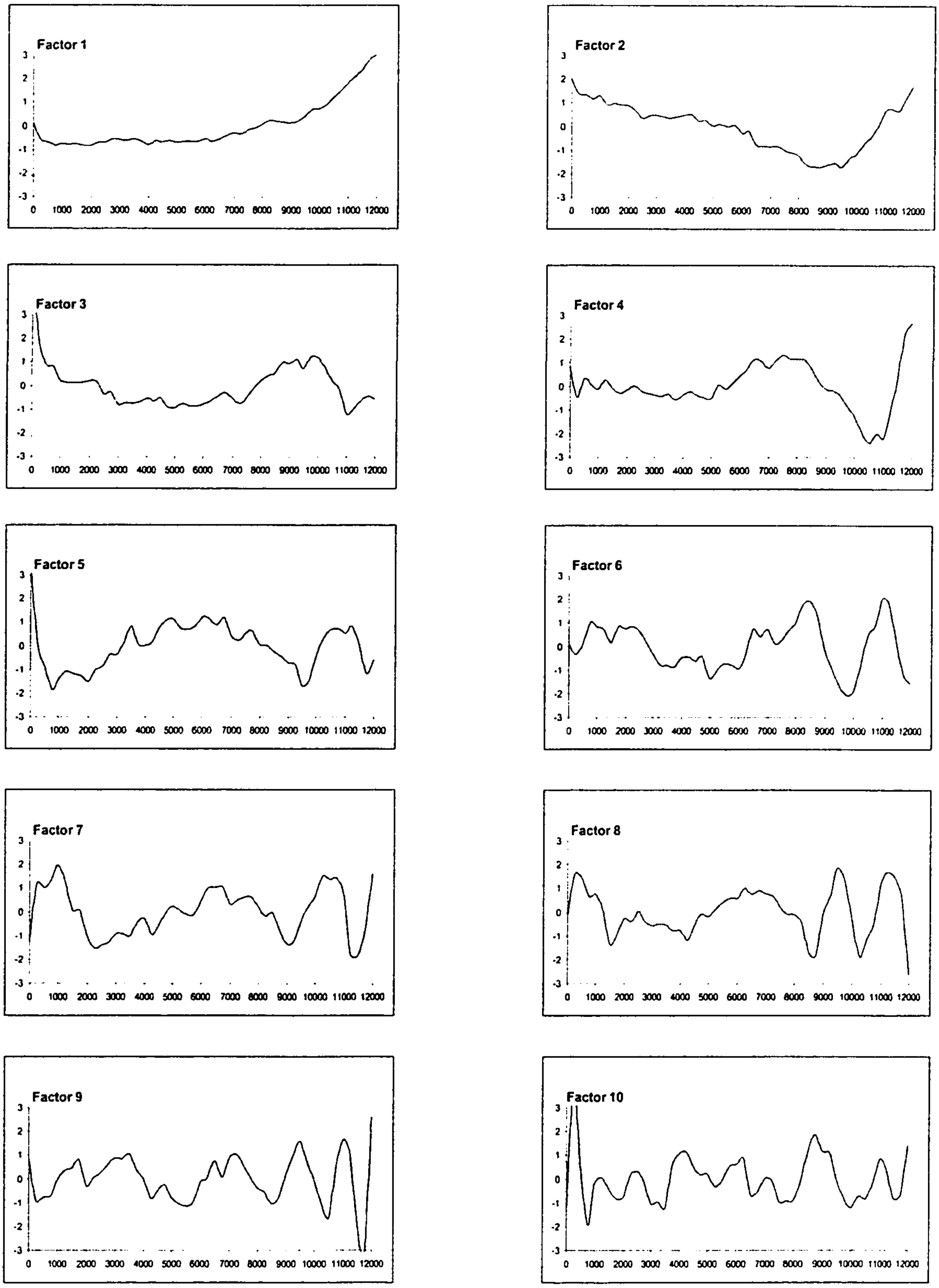


Figure F2.8: PCA on all *Asteraceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

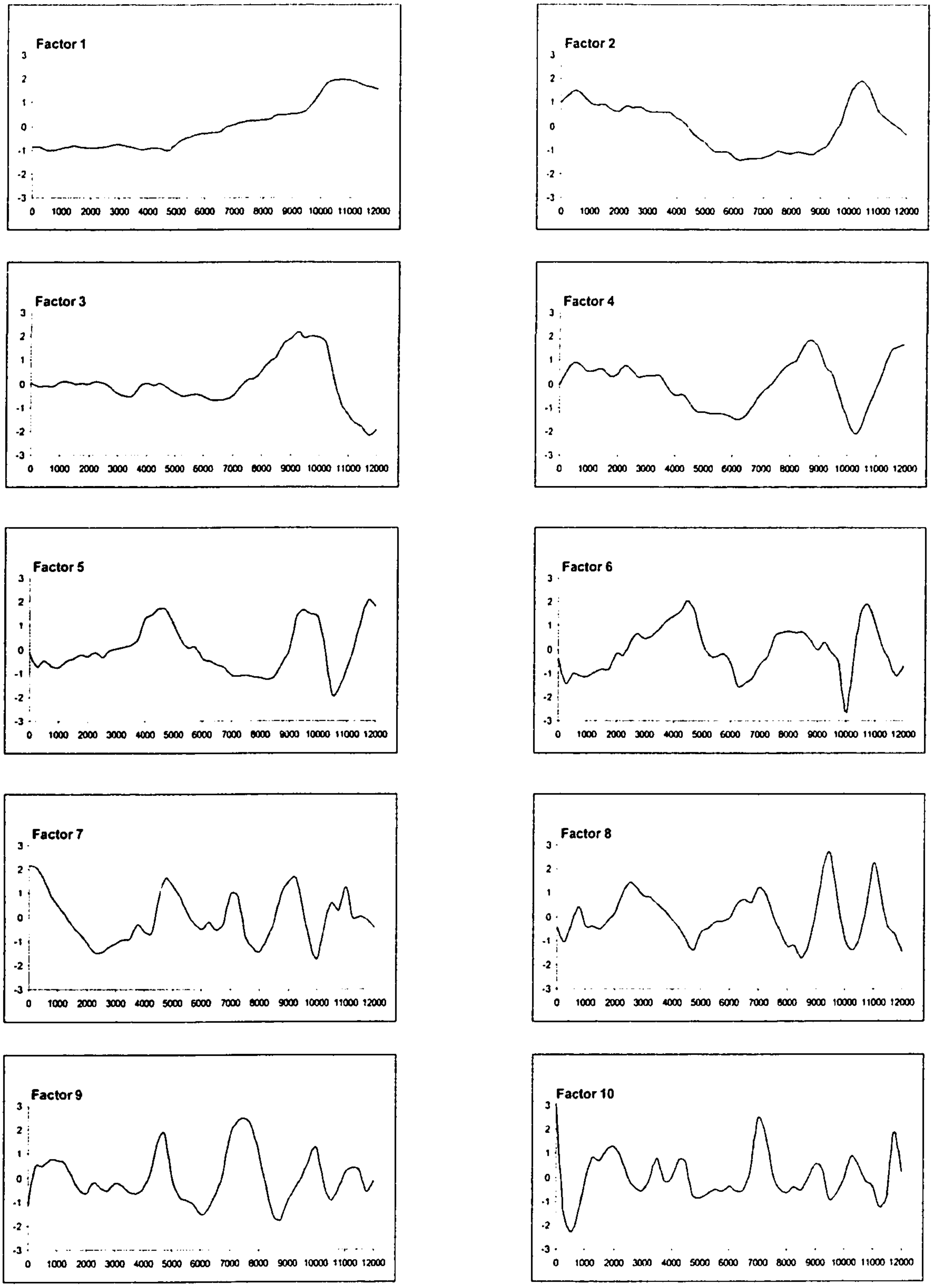


Figure F2.9: PCA on all *Betula* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

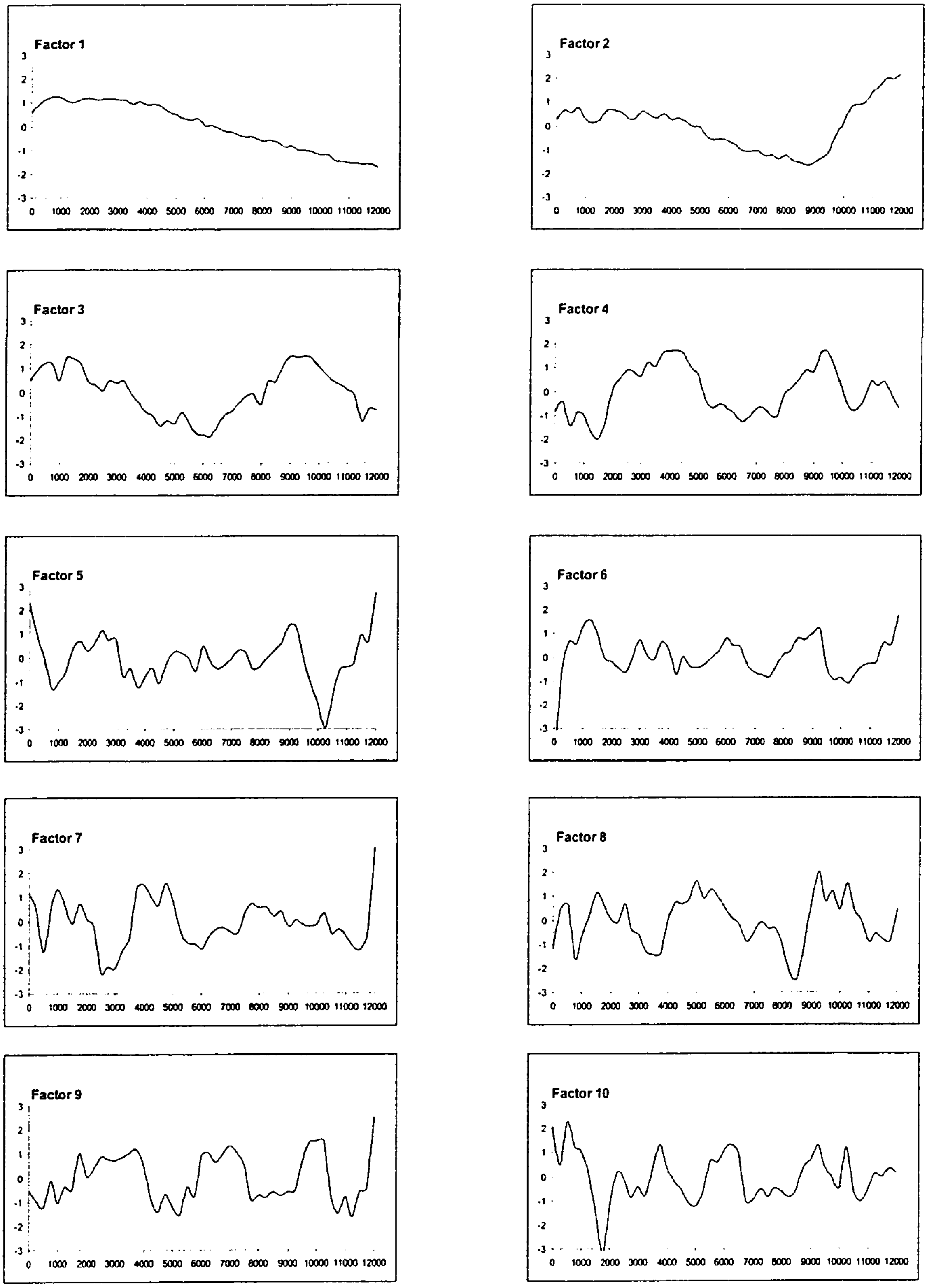


Figure F2.10: PCA on all *Carya* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

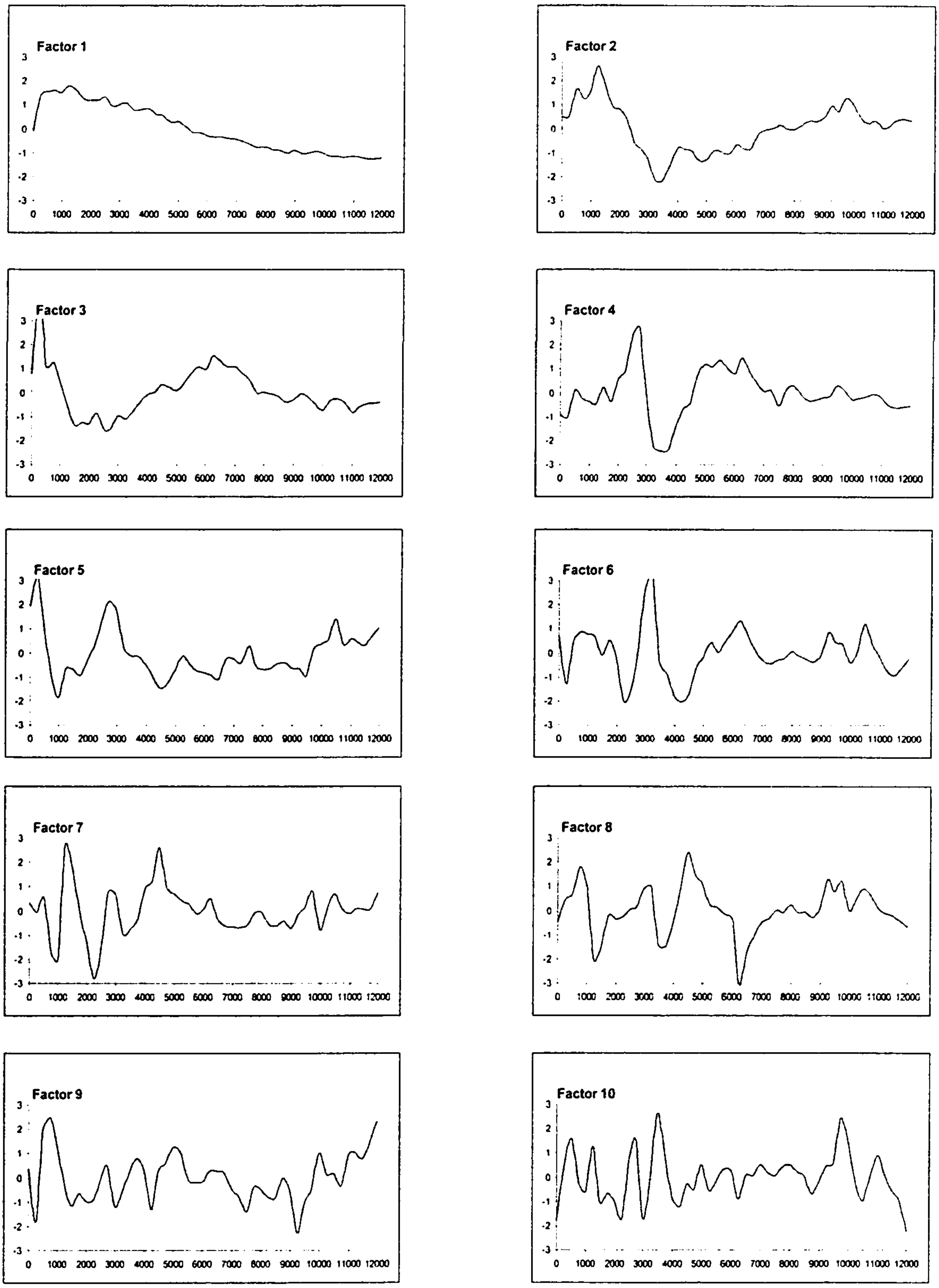


Figure F2.11: PCA on all *Castaneae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

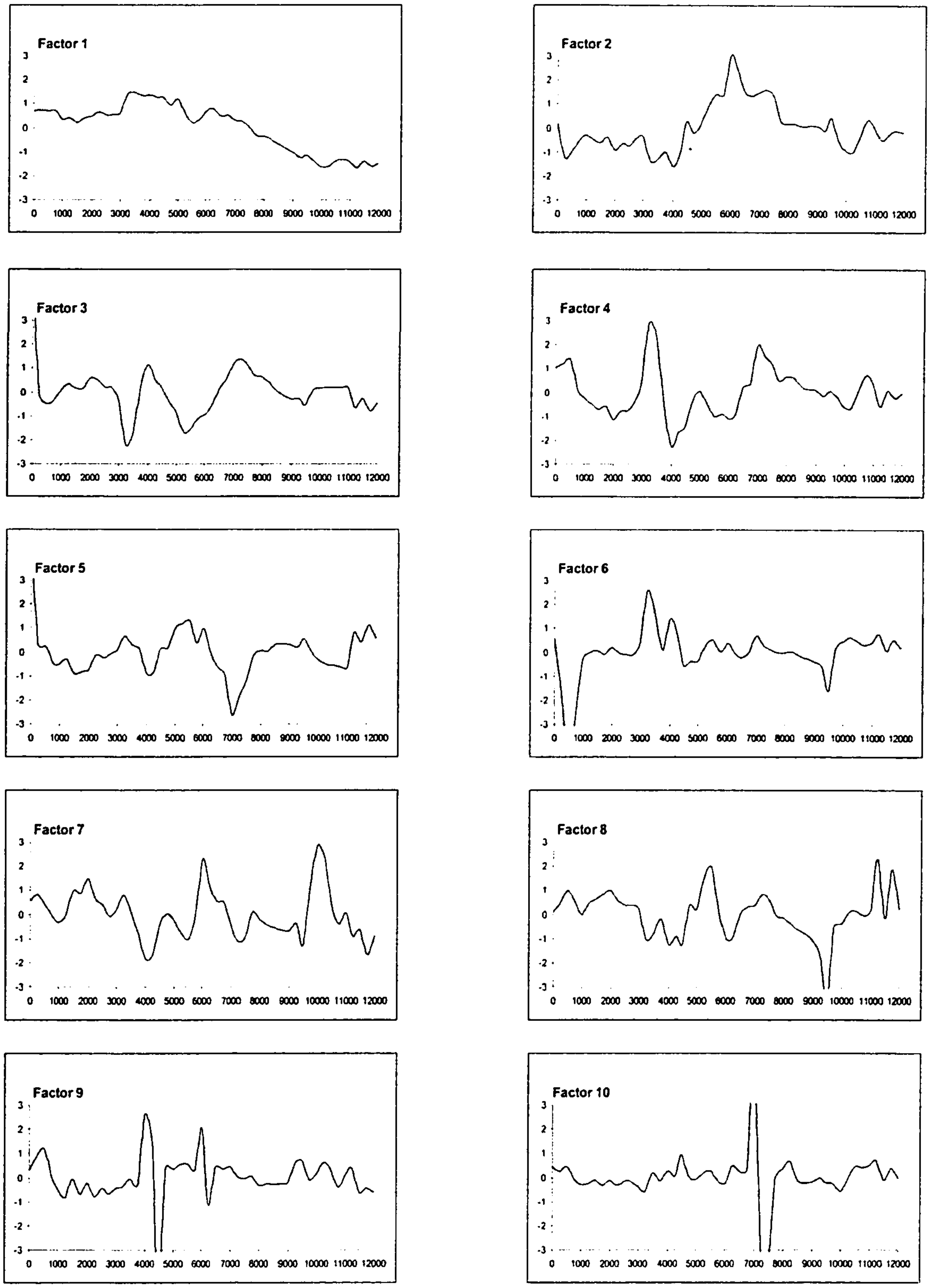


Figure F2.12: PCA on all *Cephalanthus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

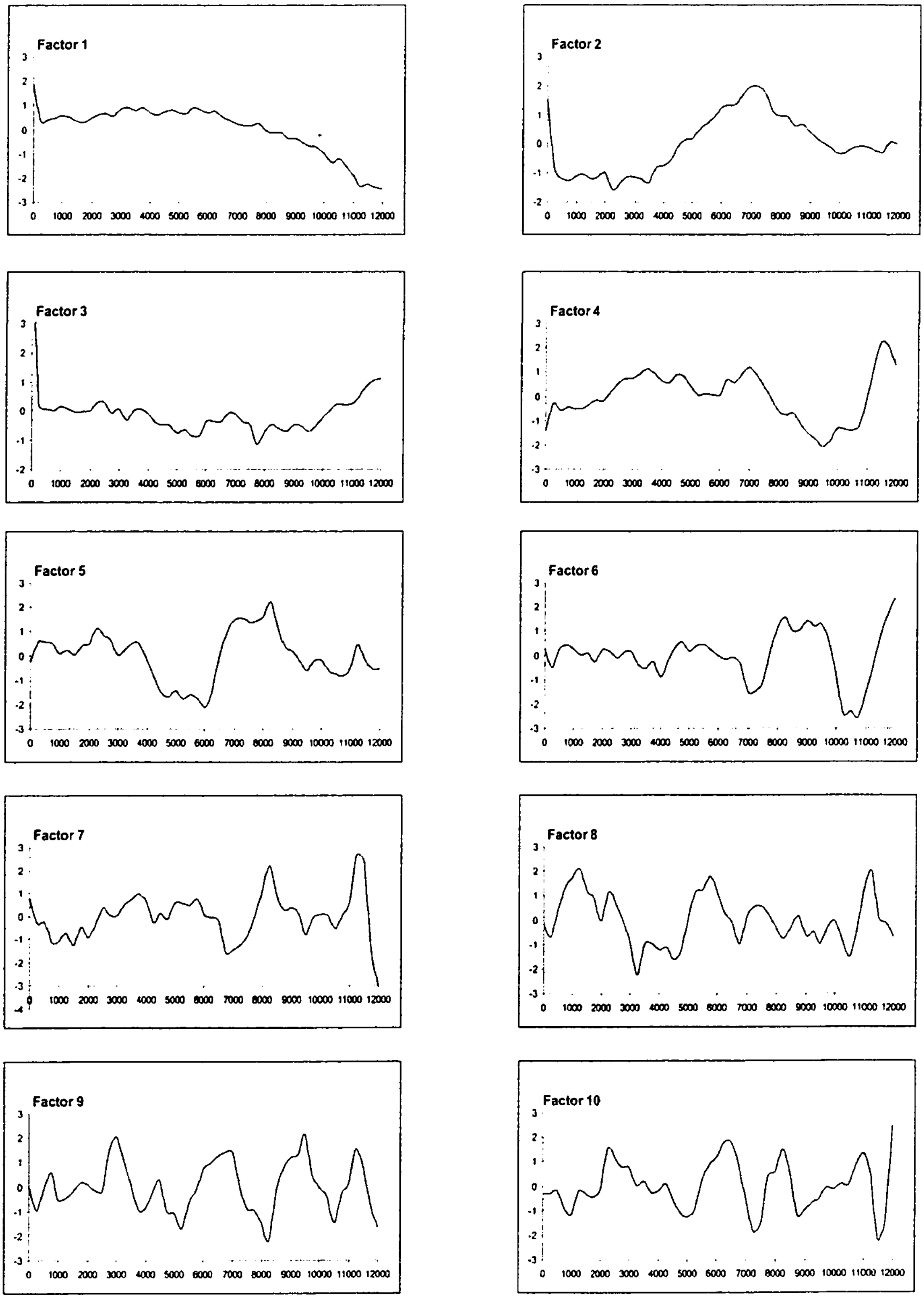


Figure F2.13: PCA on all *Chenopodiaeae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

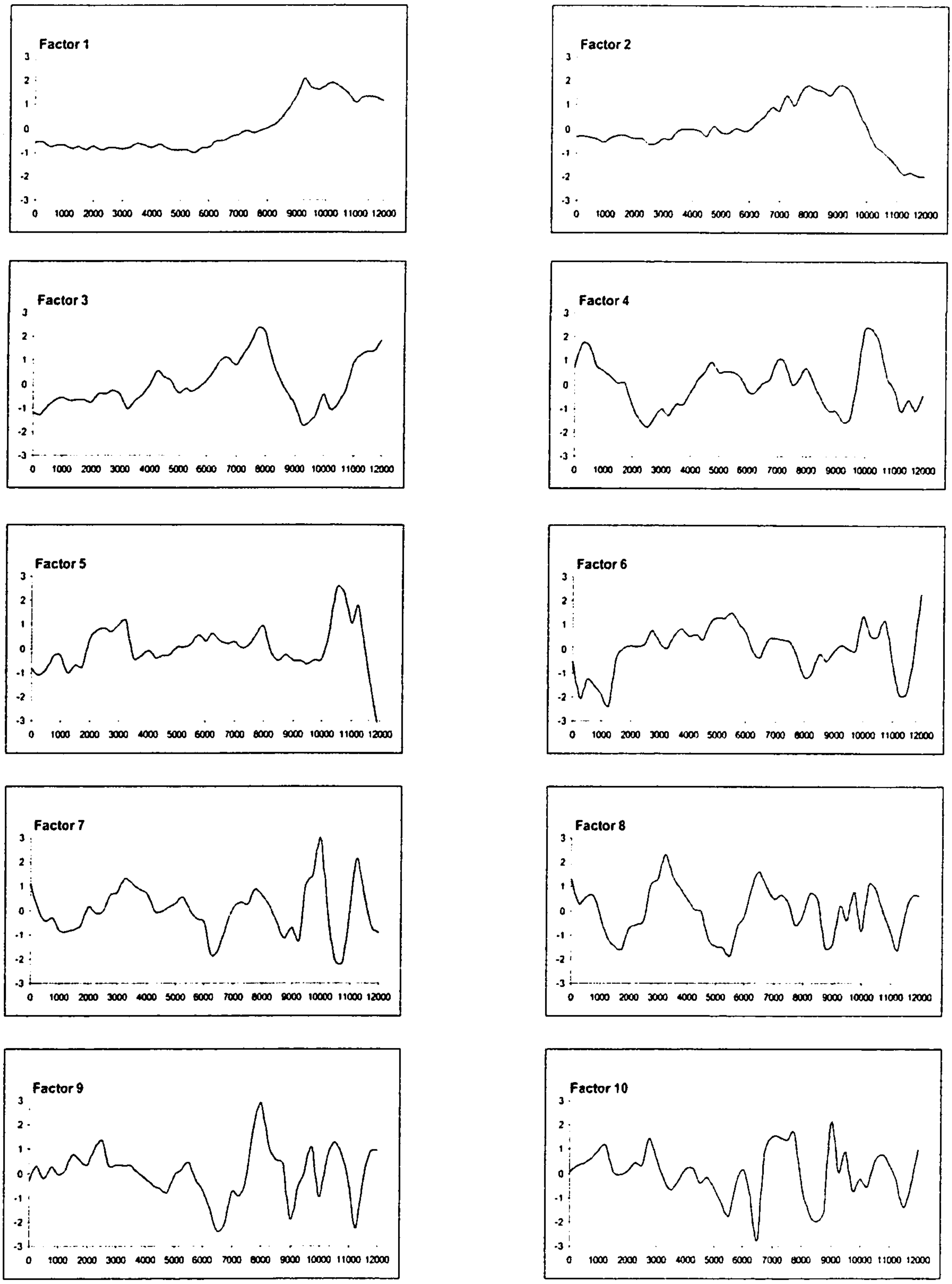


Figure F2.14: PCA on all *Corylus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

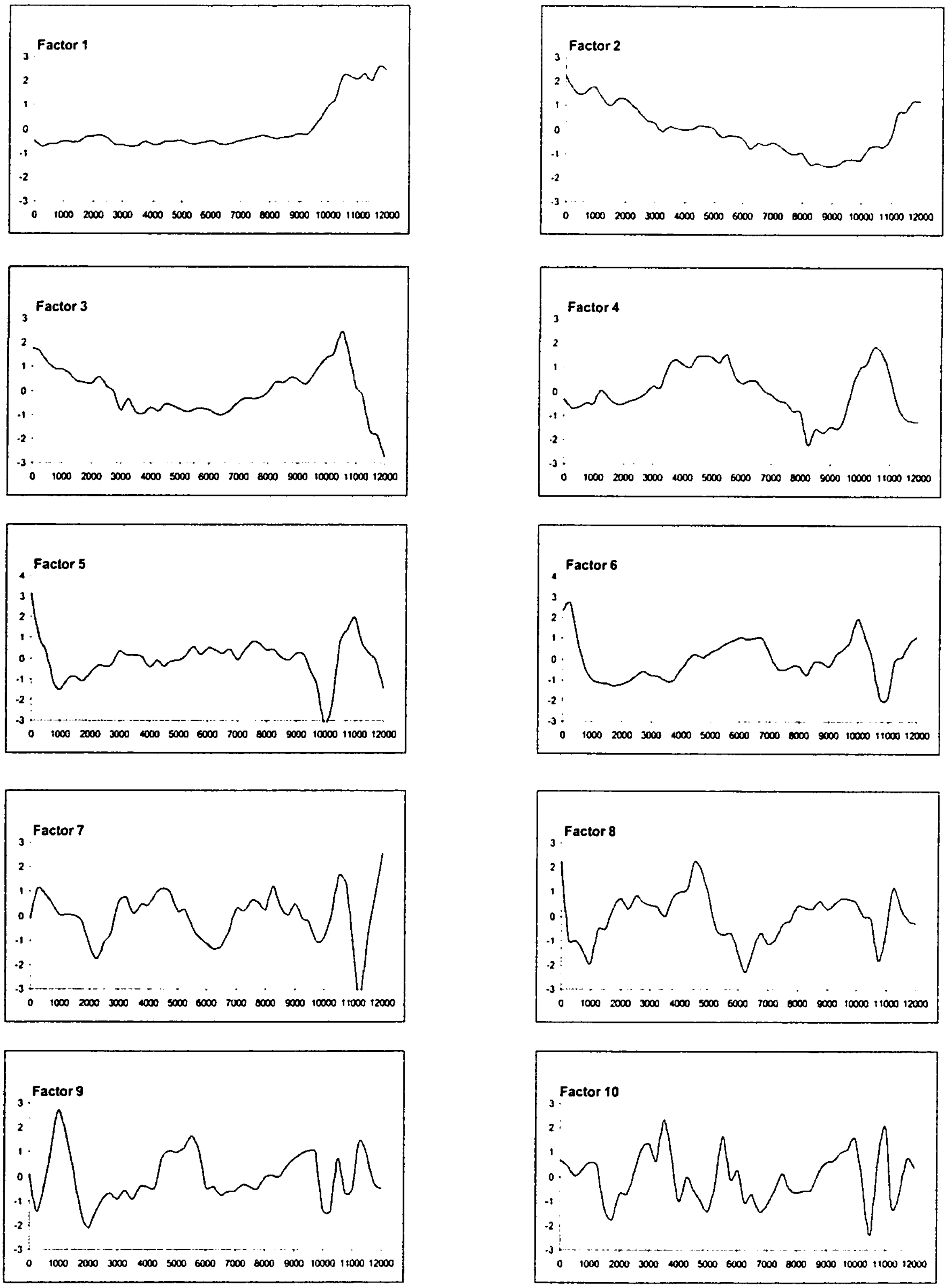


Figure F2.15: PCA on all *Cupressaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

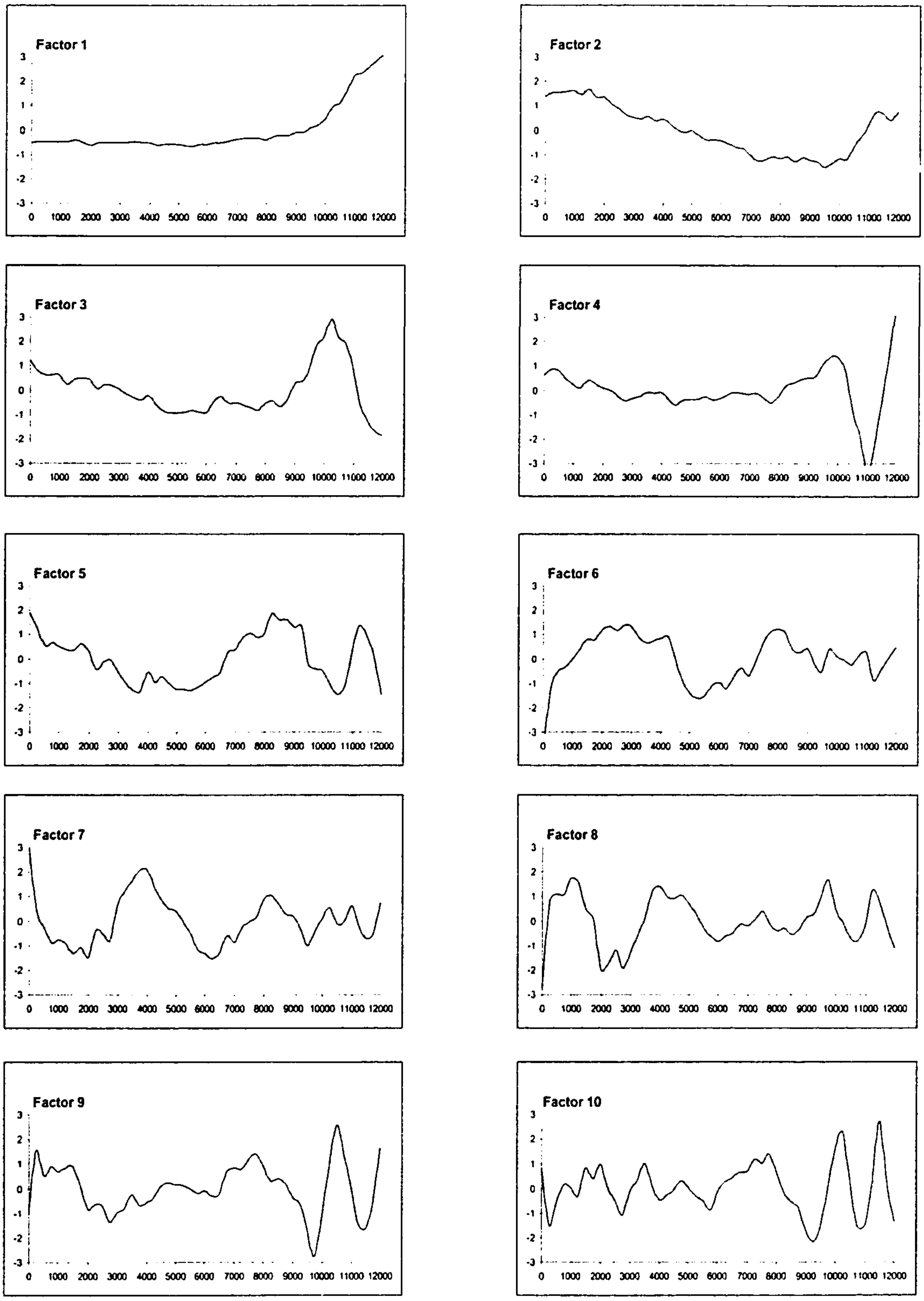


Figure F2.16: PCA on all *Cyperaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

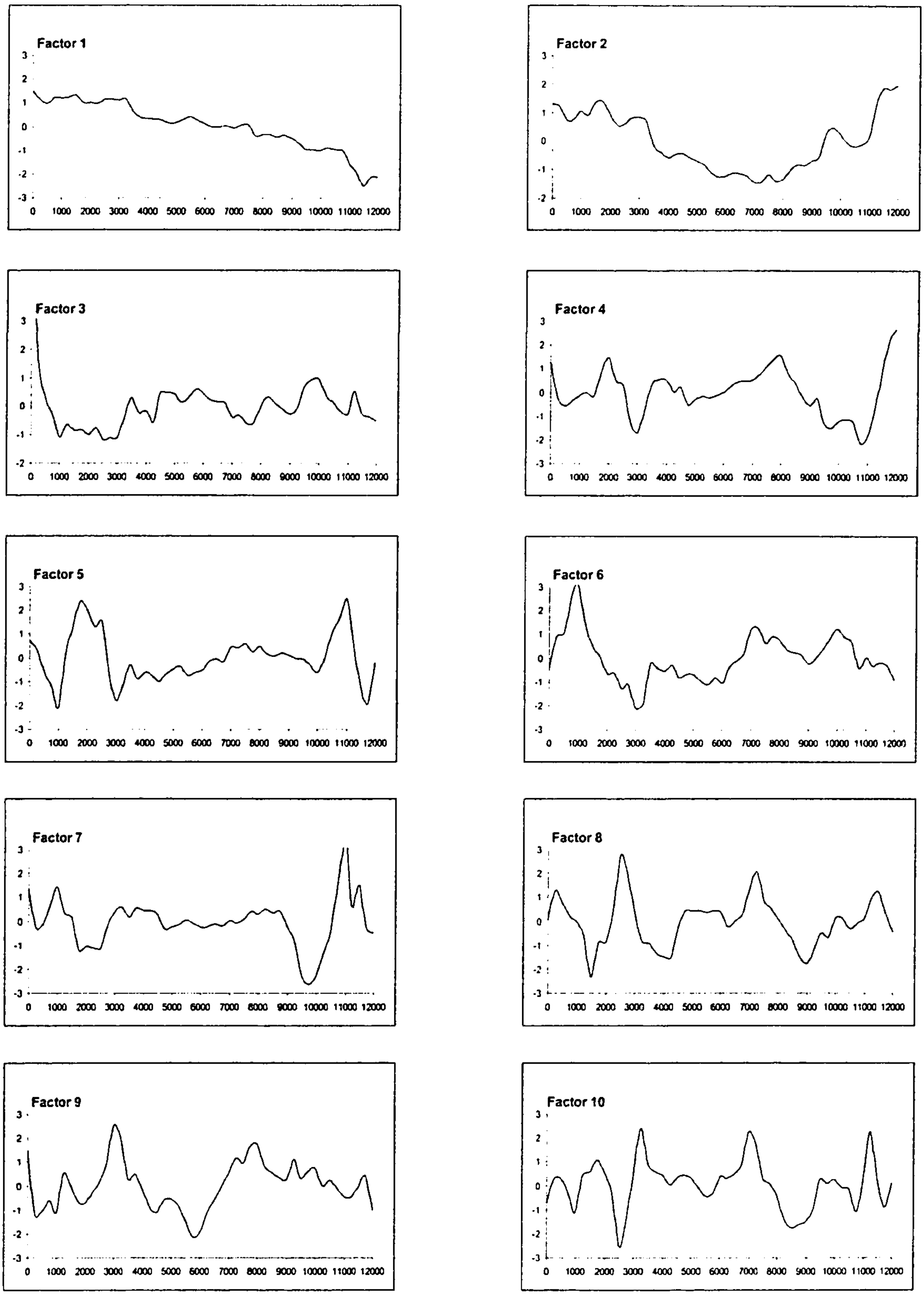


Figure F2.17: PCA on all *Ericaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

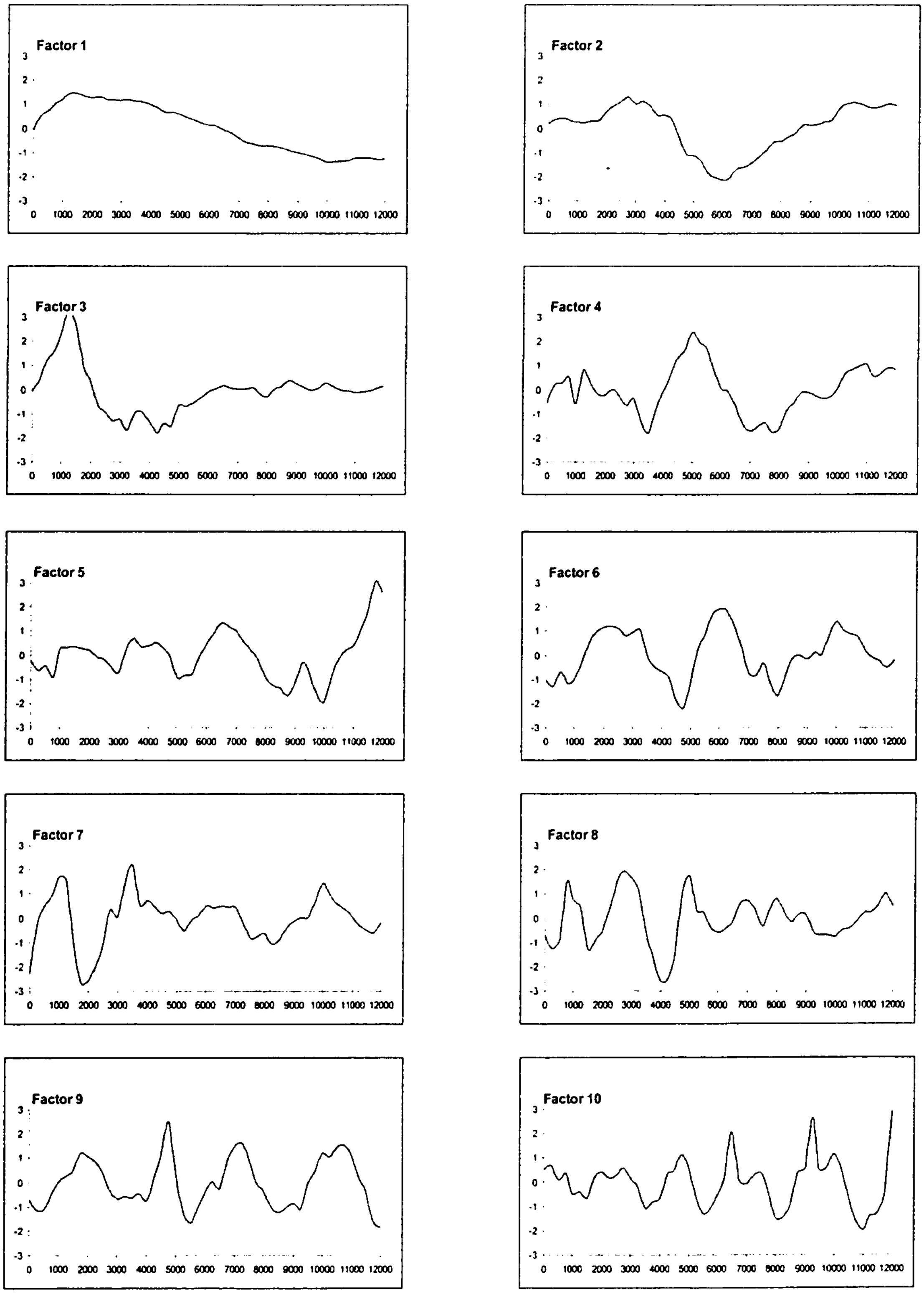


Figure F2.18: PCA on all *Fagus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

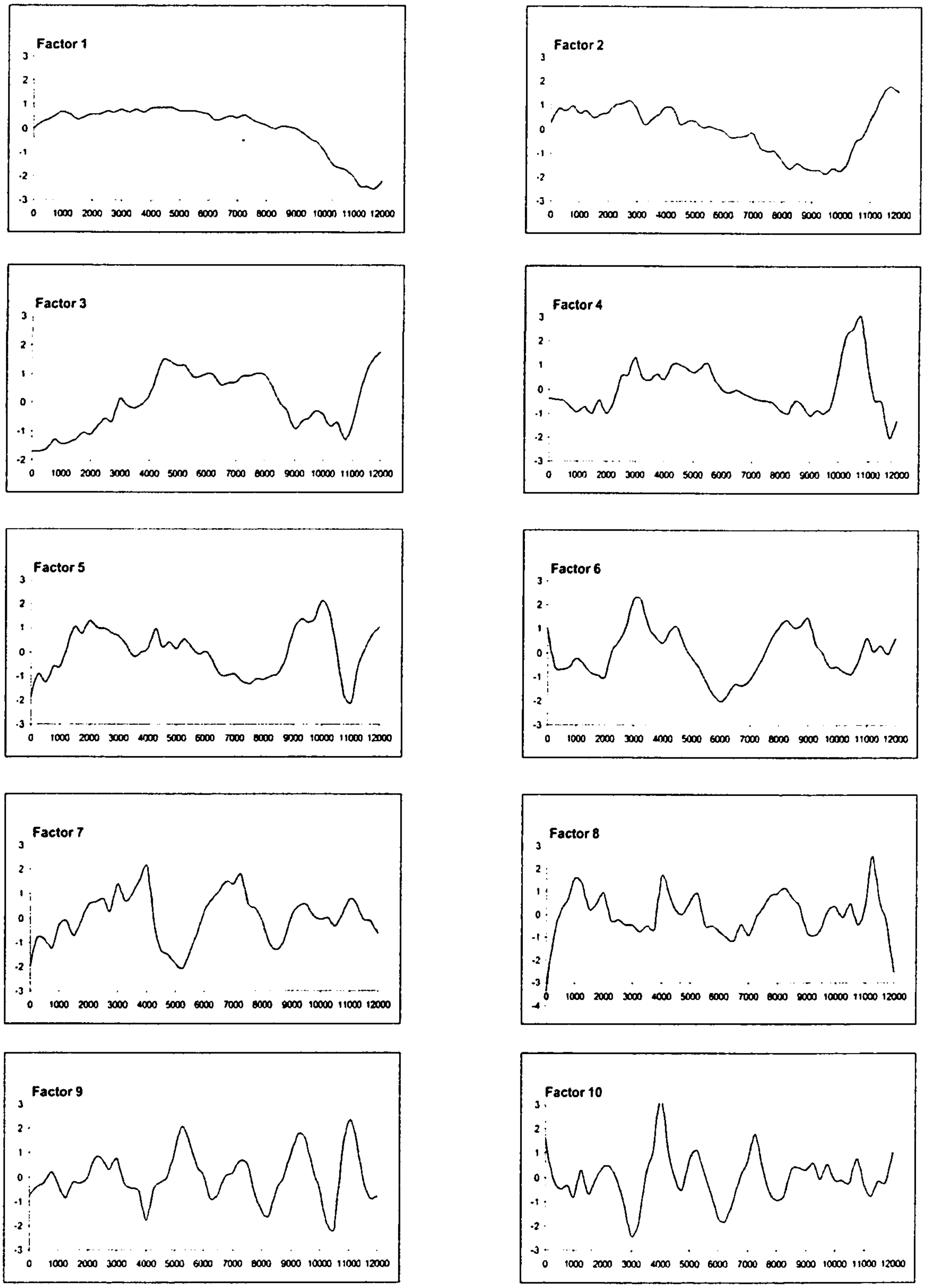


Figure F2.19: PCA on all *Fraxinus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

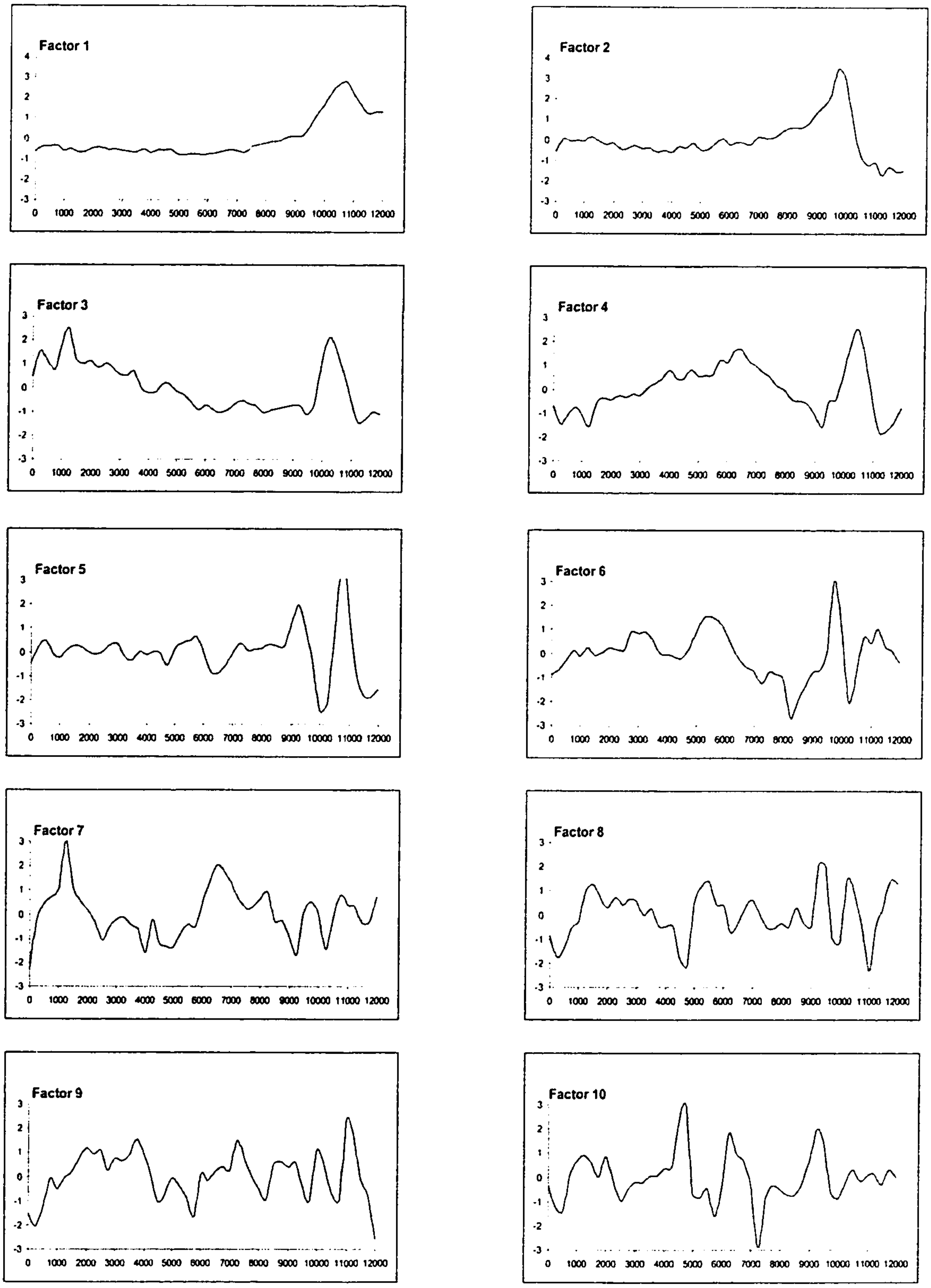


Figure F2.20: PCA on all *Larix/Pseudotsuga* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

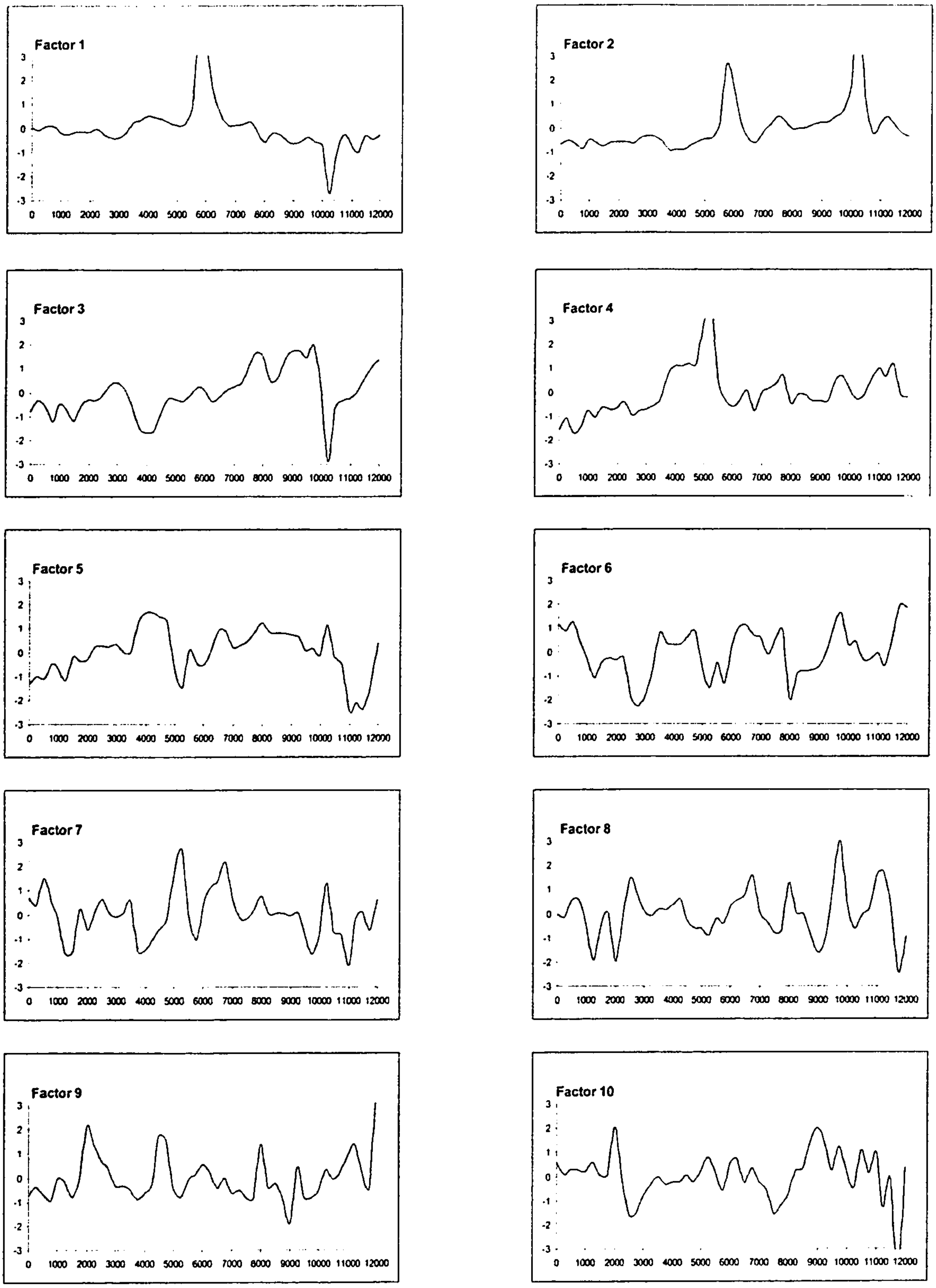


Figure F2.21: PCA on all *Liliaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

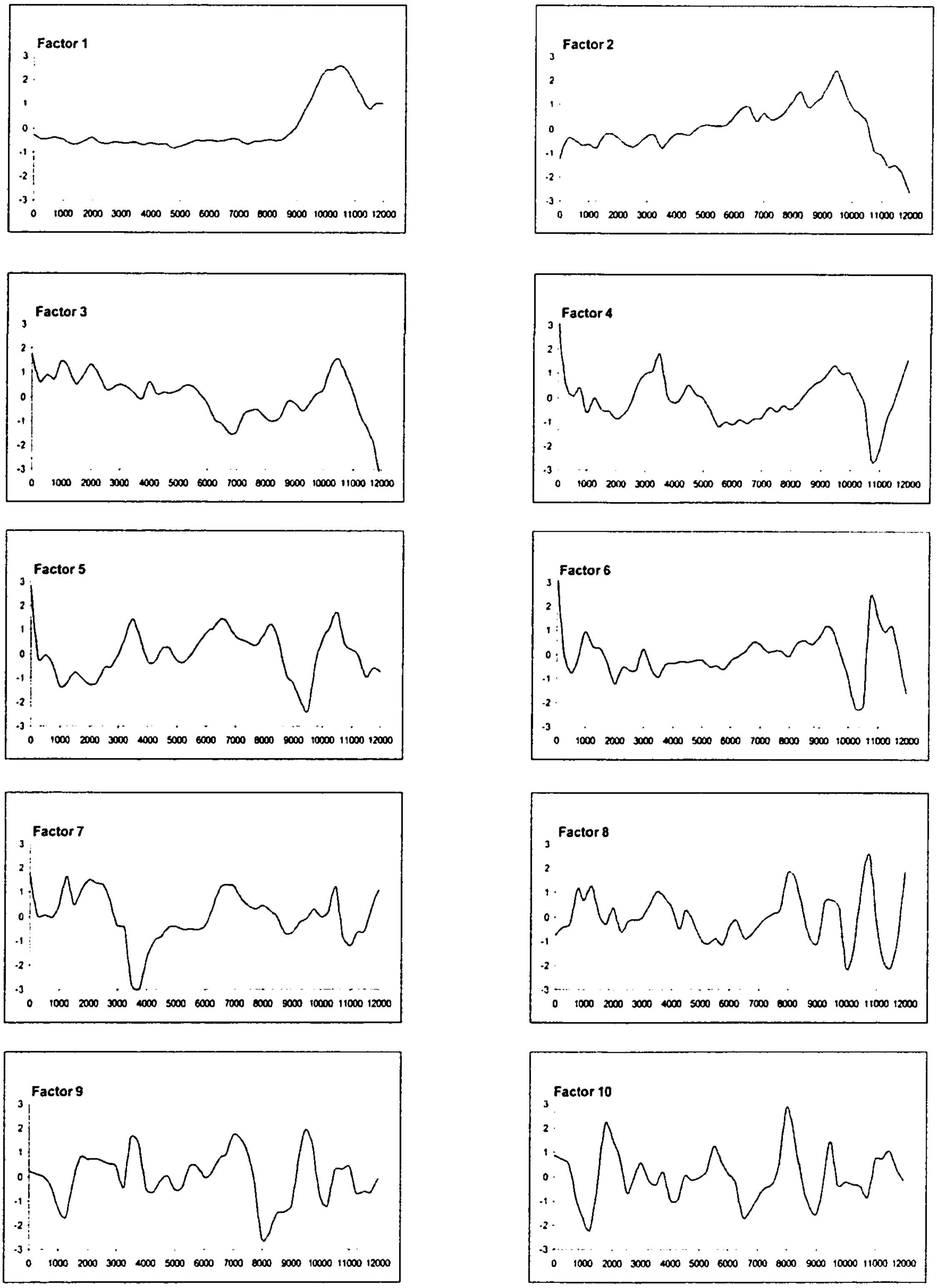


Figure F2.22: PCA on all *Lycopodium* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

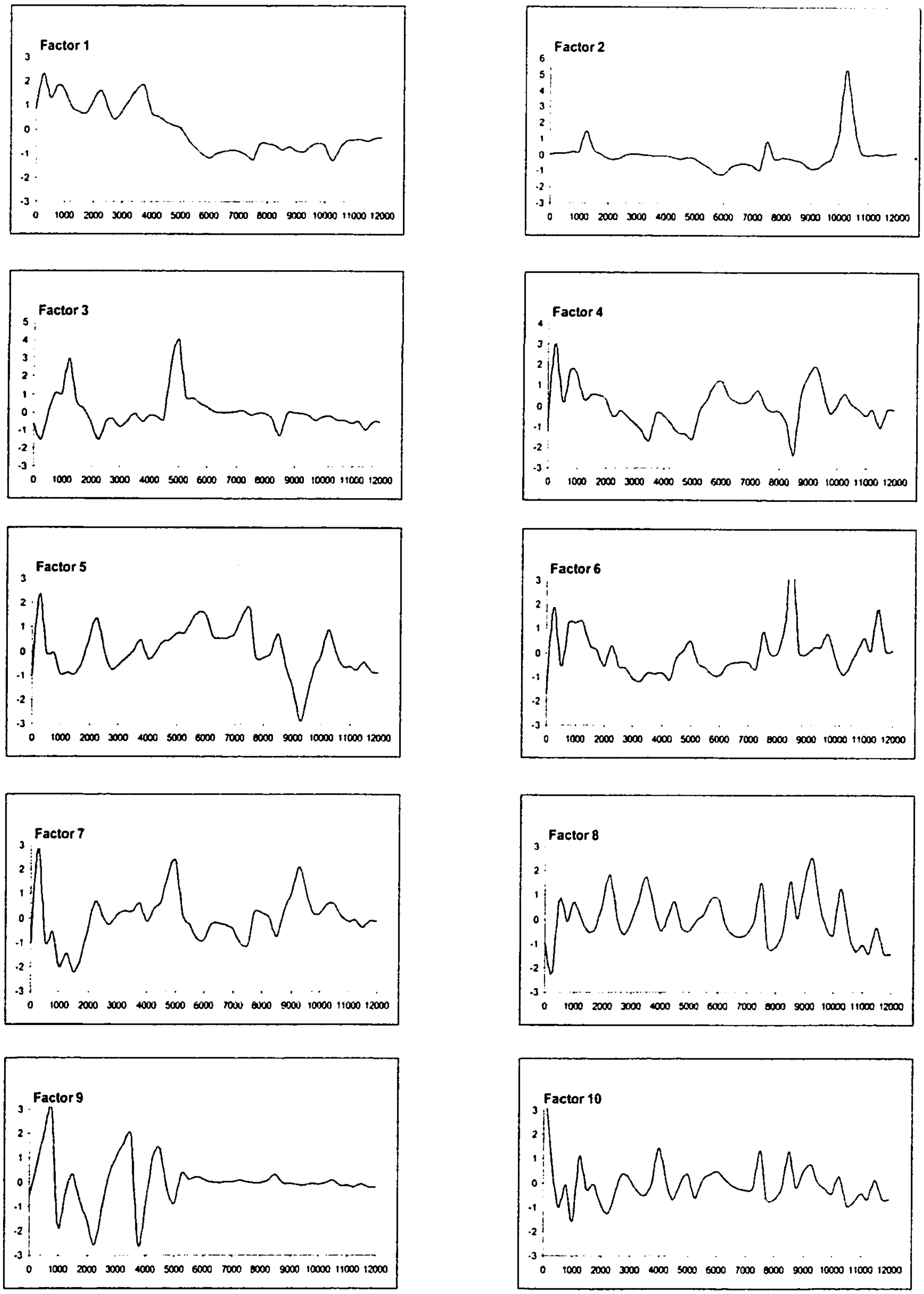


Figure F2.23: PCA on all *Magnoliaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

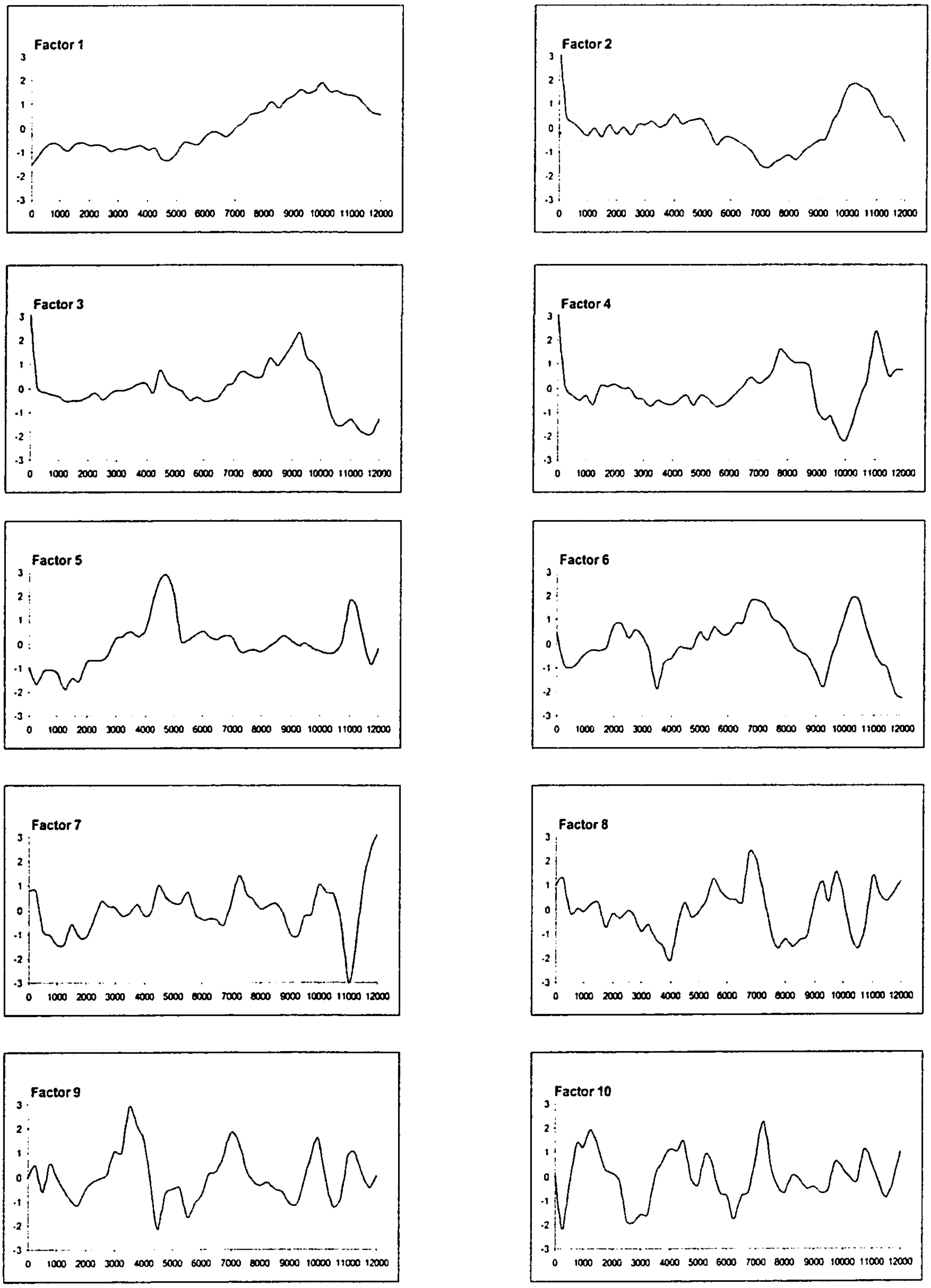


Figure F2.24: PCA on all *Myrica* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

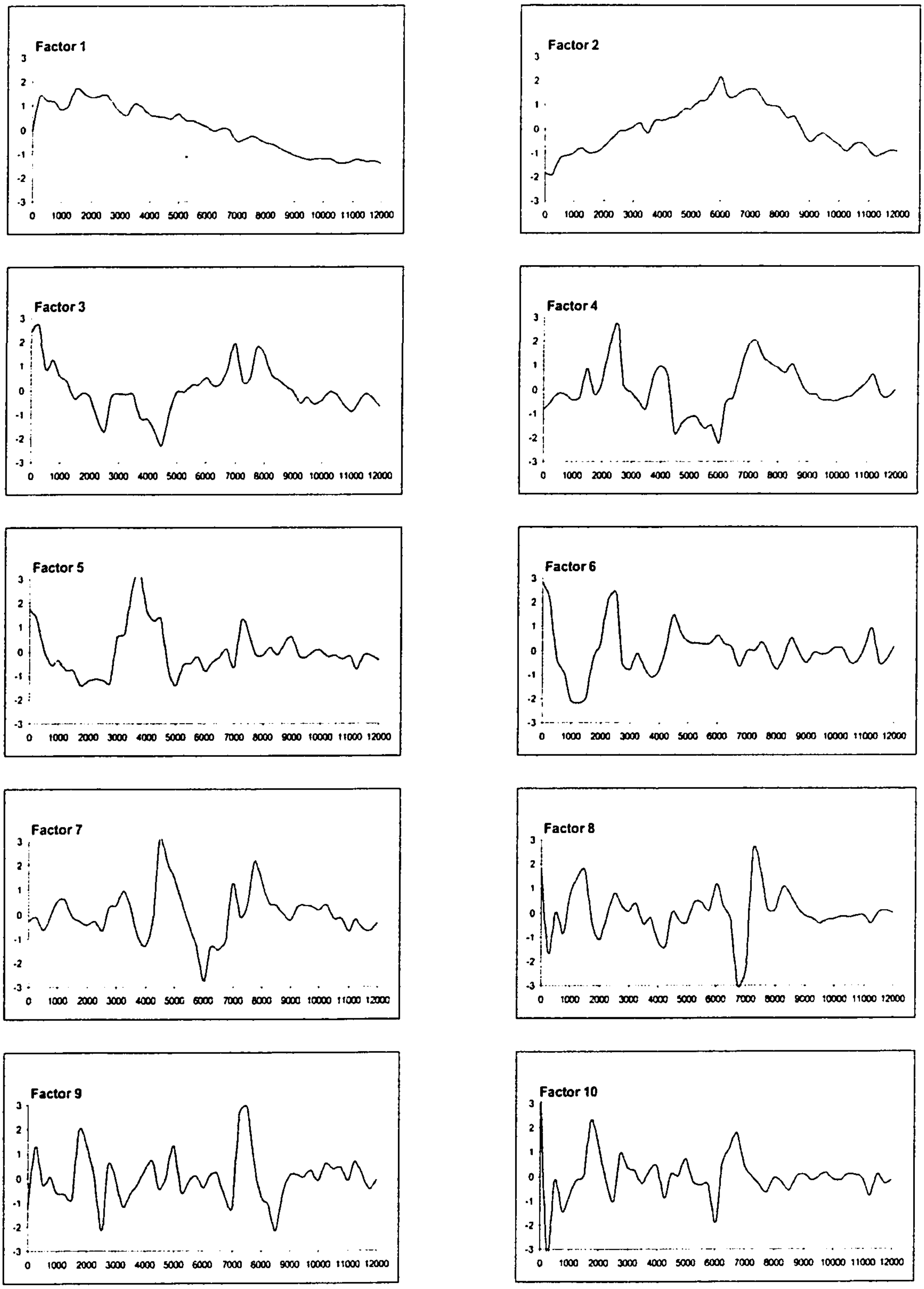


Figure F2.25: PCA on all *Nyssa* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

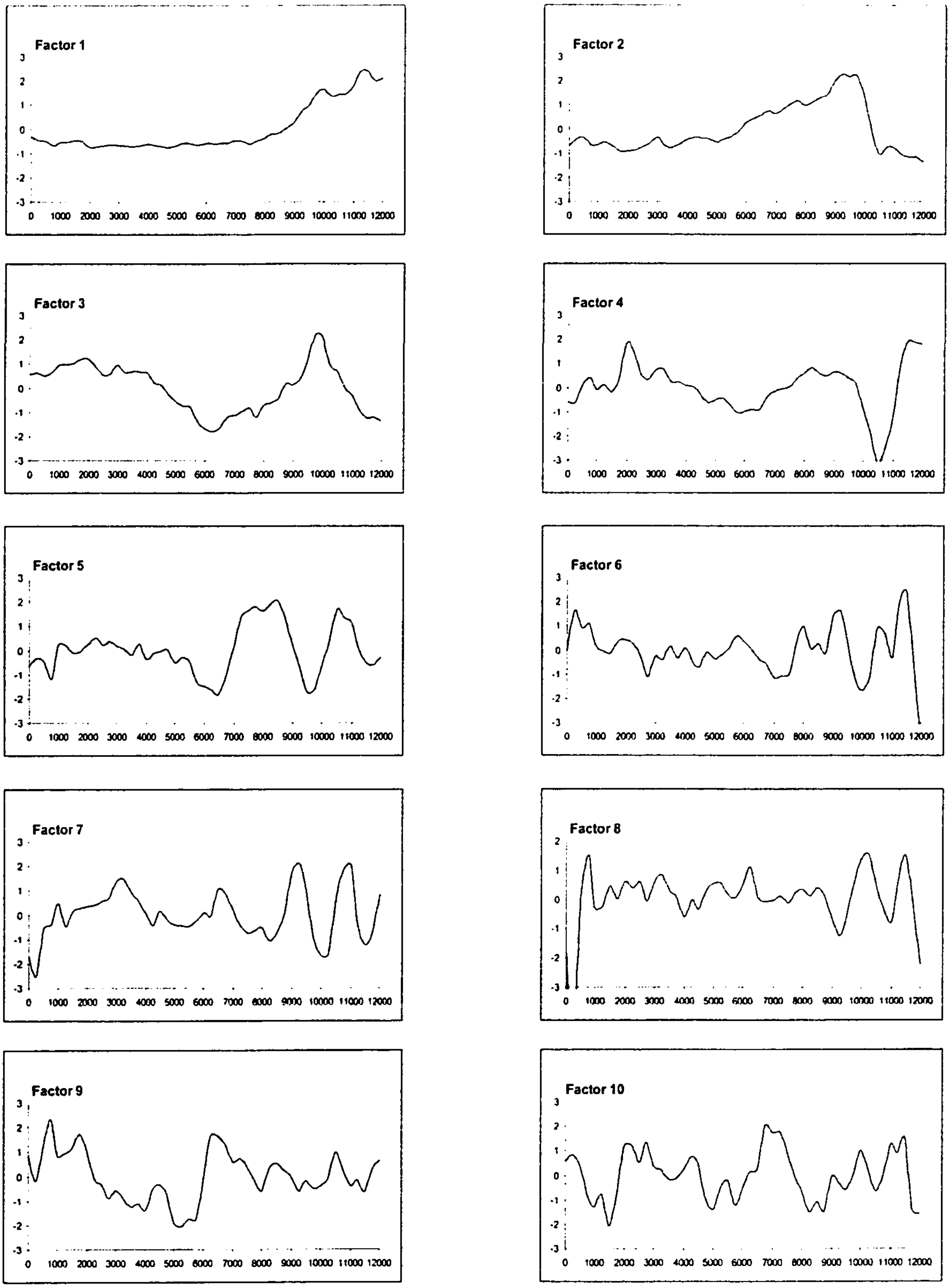


Figure F2.26: PCA on all *Ostrya/Carpinus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

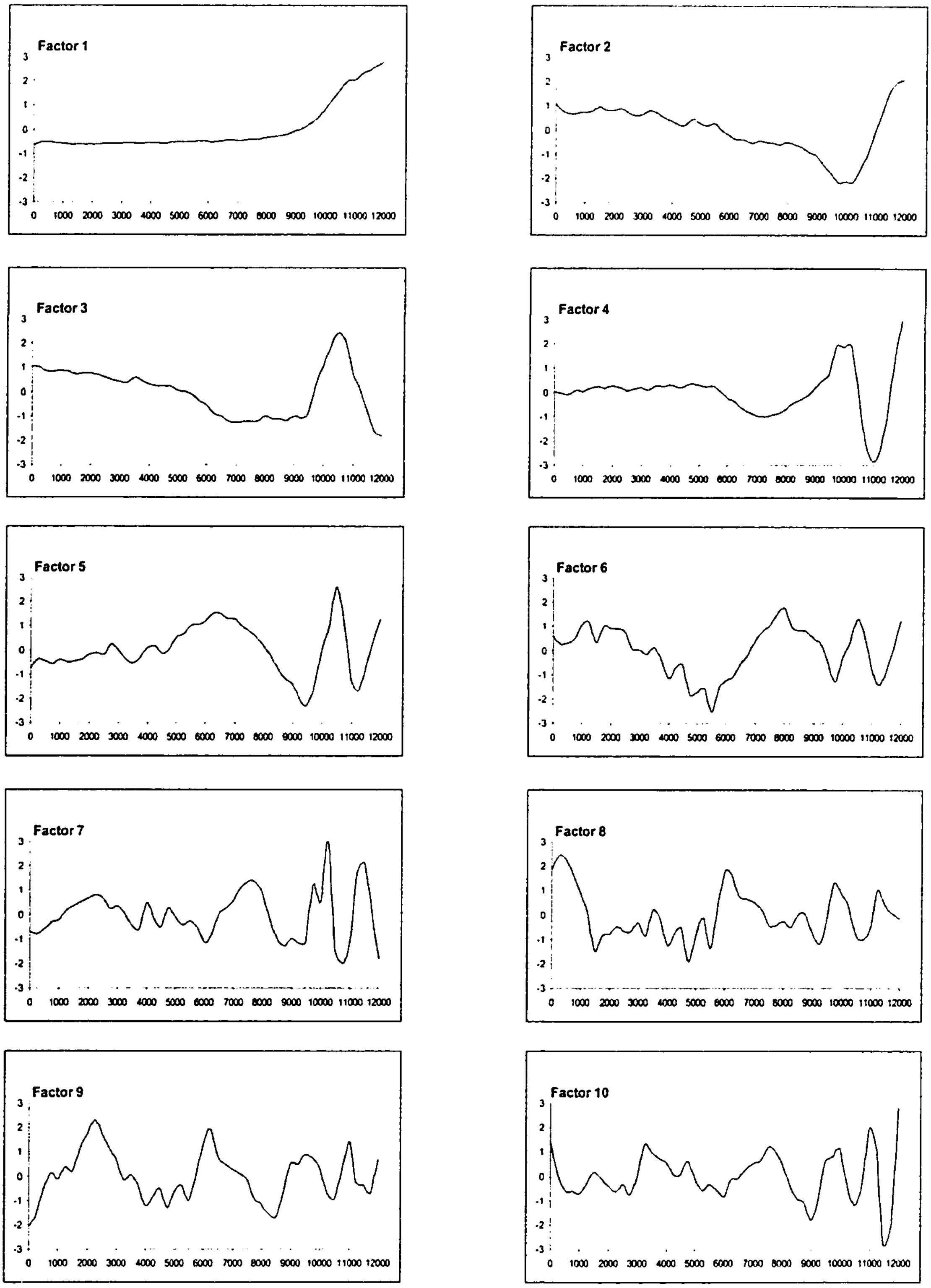


Figure F2.27: PCA on all *picea undiff.* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

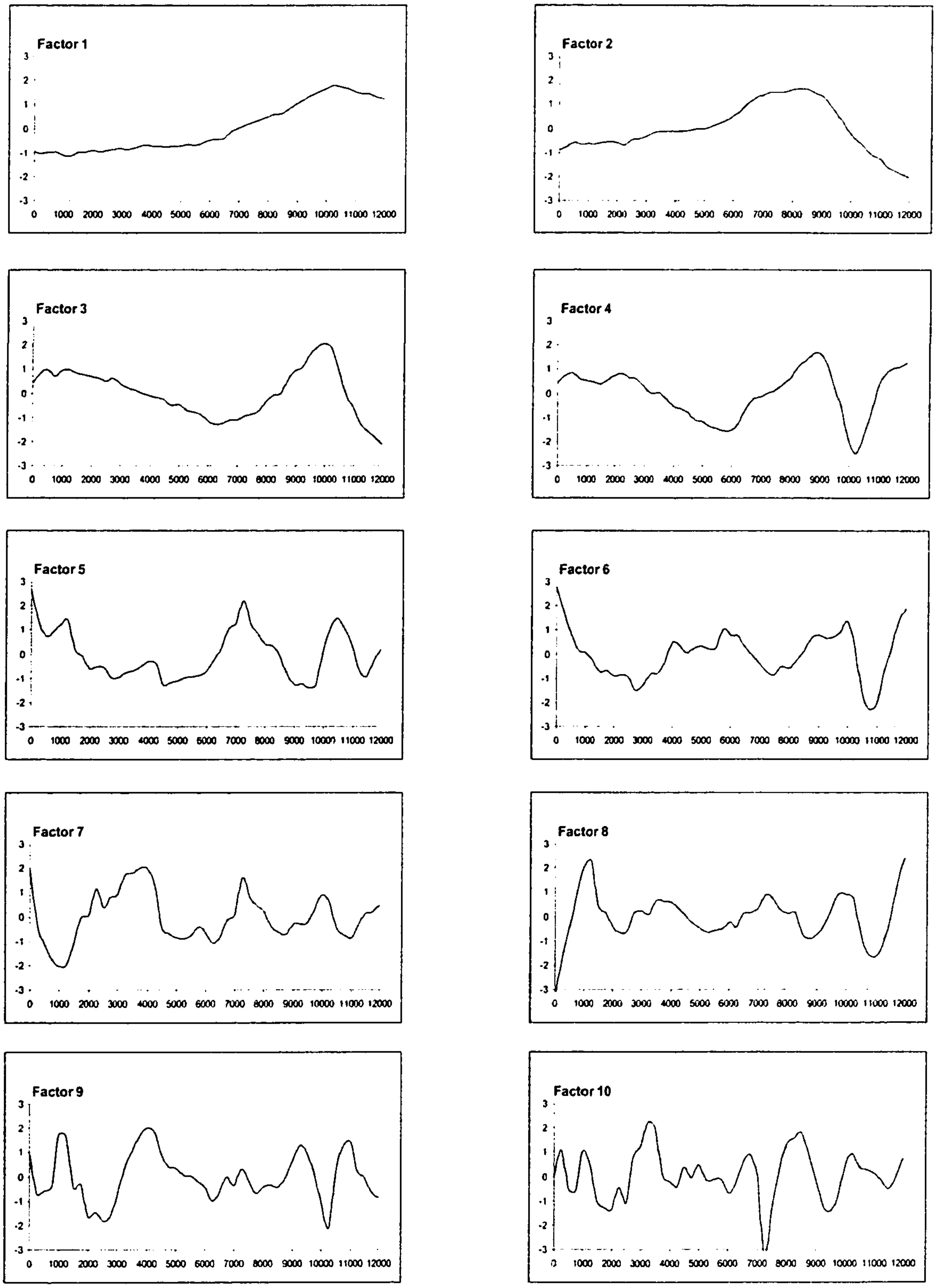


Figure F2.28: PCA on all *Pinus undiff.*. taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

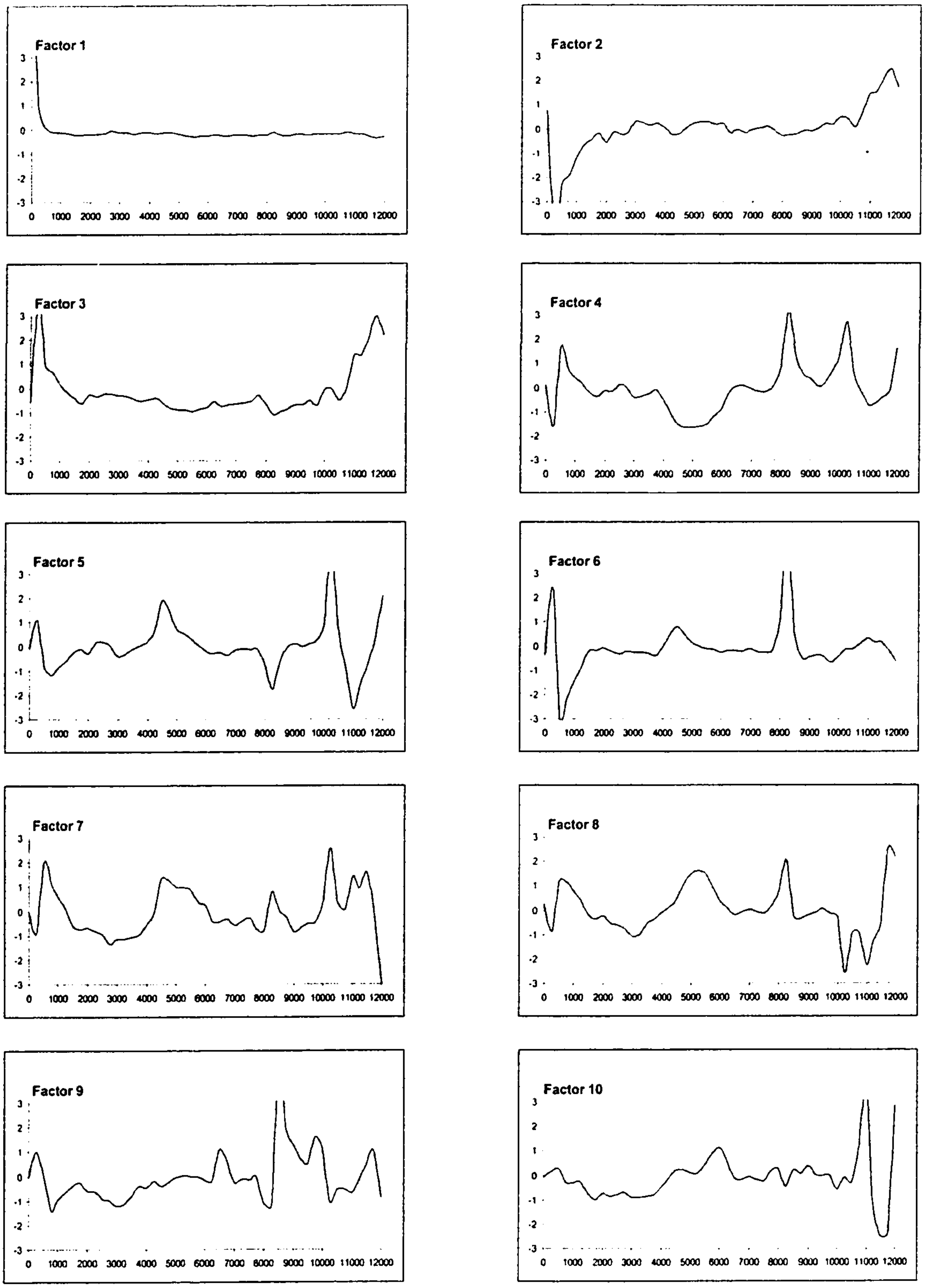


Figure F2.29: PCA on all *Plantaginaceae.* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

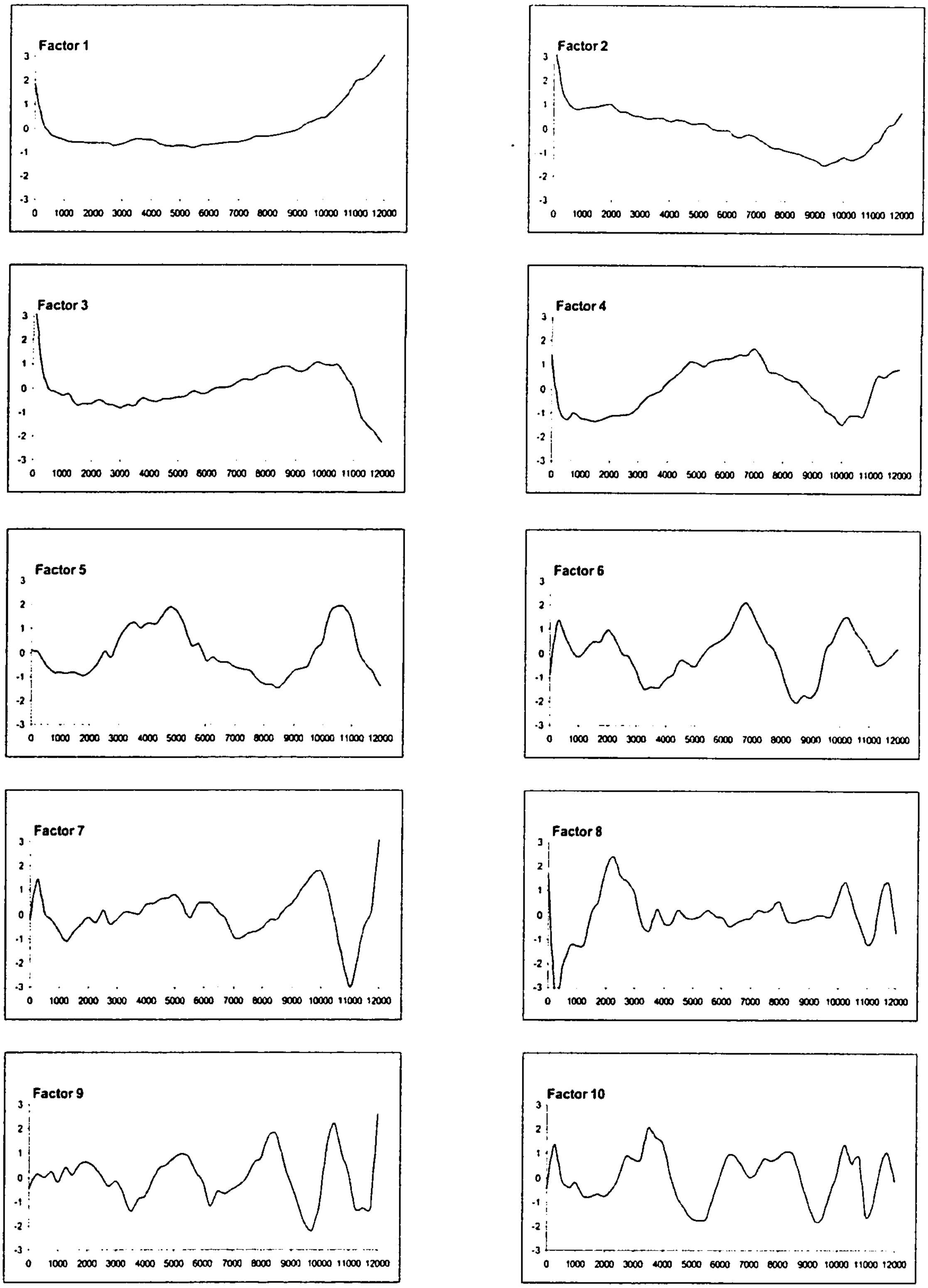


Figure F2.30: PCA on all *Poaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

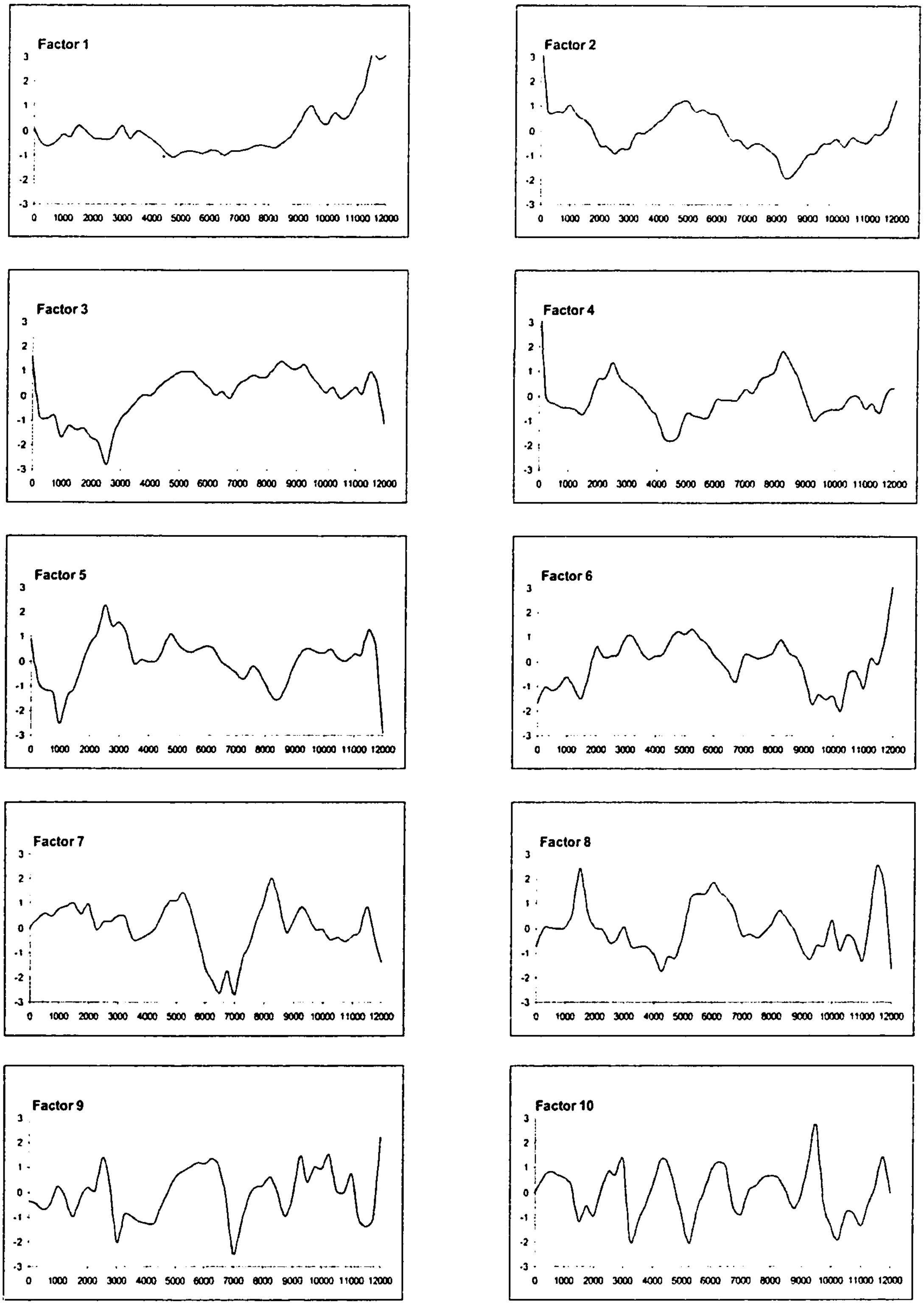


Figure F2.31: PCA on all *Polygonum* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

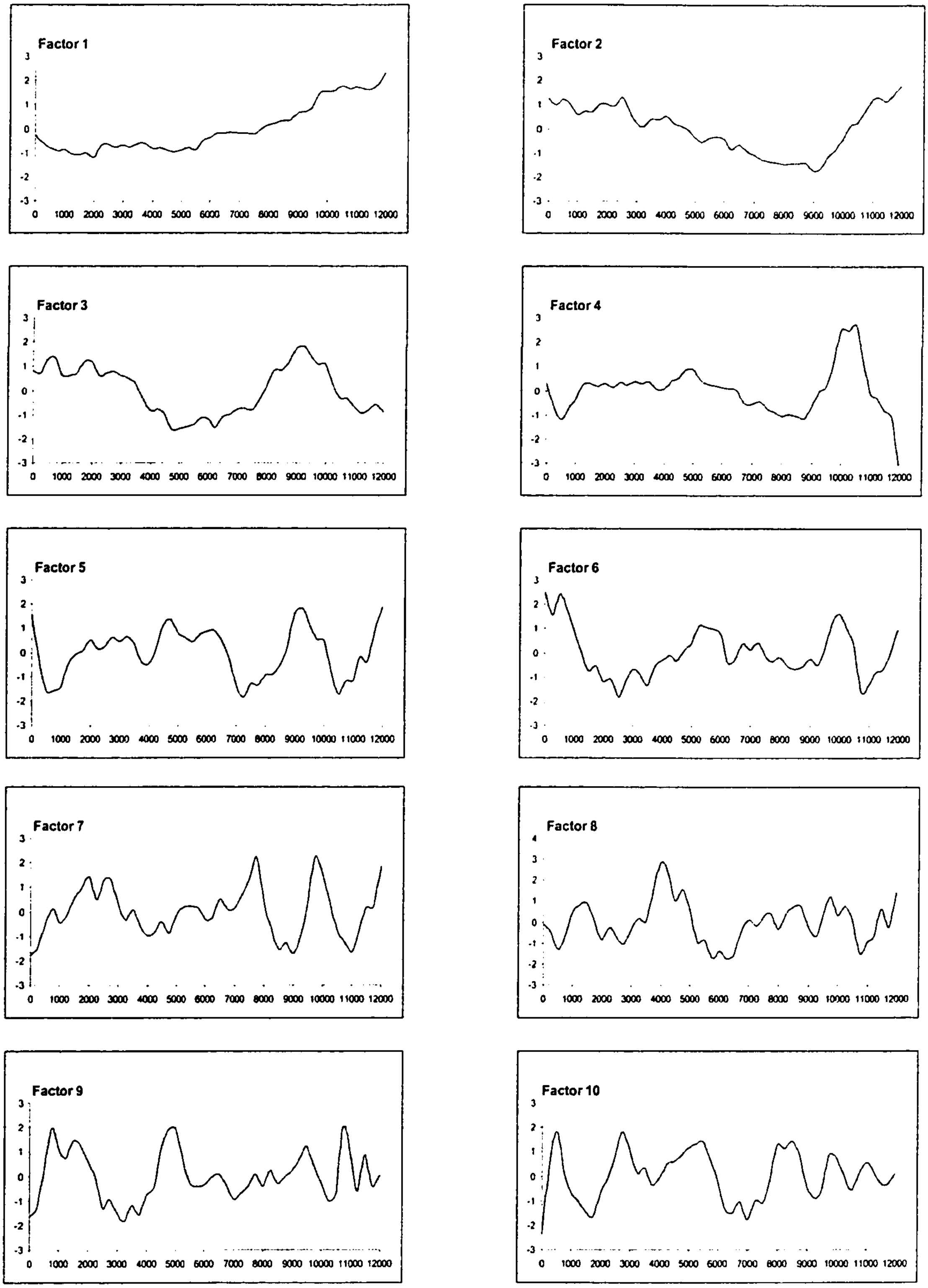


Figure F2.32: PCA on all *Polypodiaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

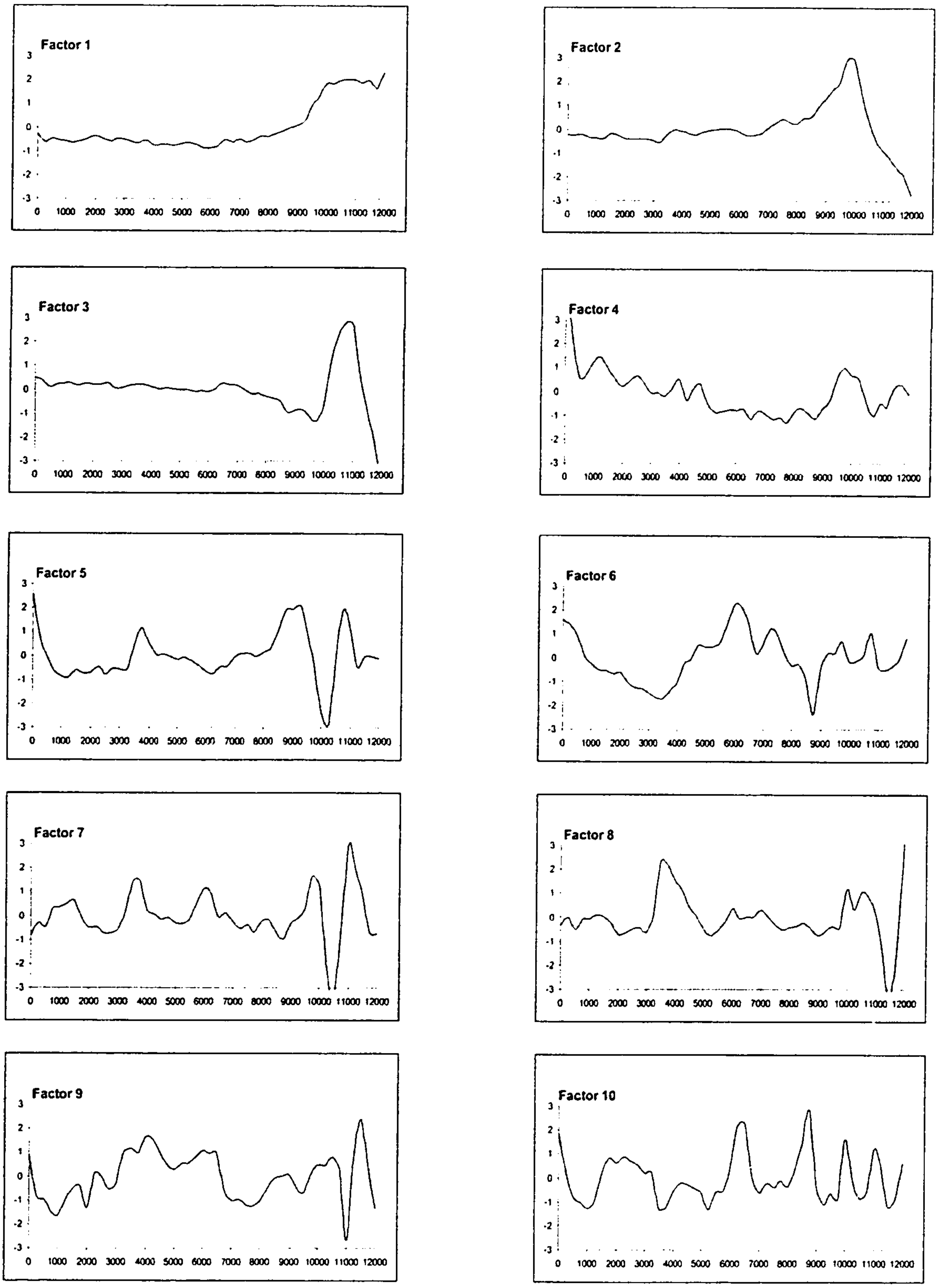


Figure F2.33: PCA on all *Populus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

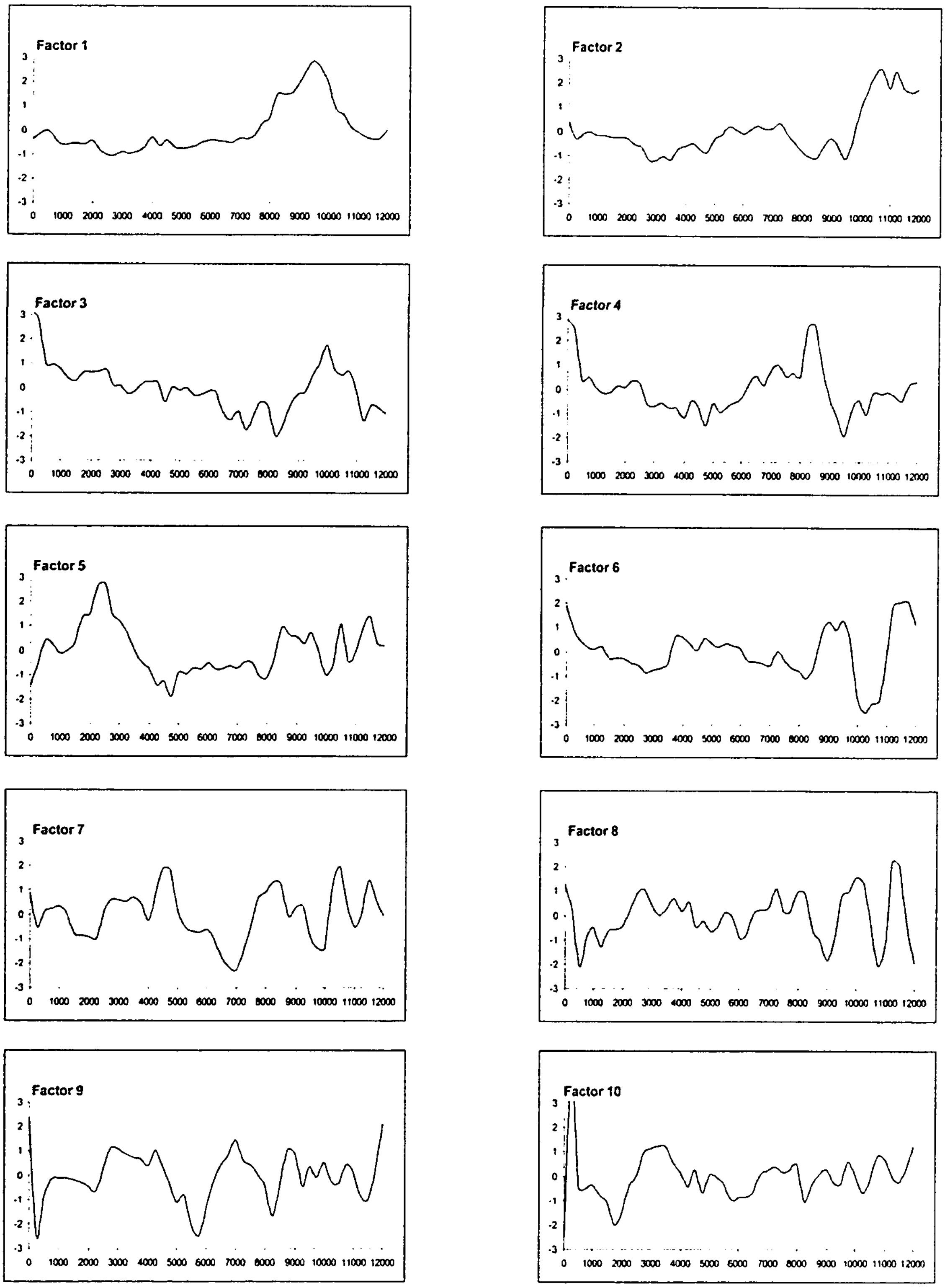


Figure F2.34: PCA on all *Pteridium* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

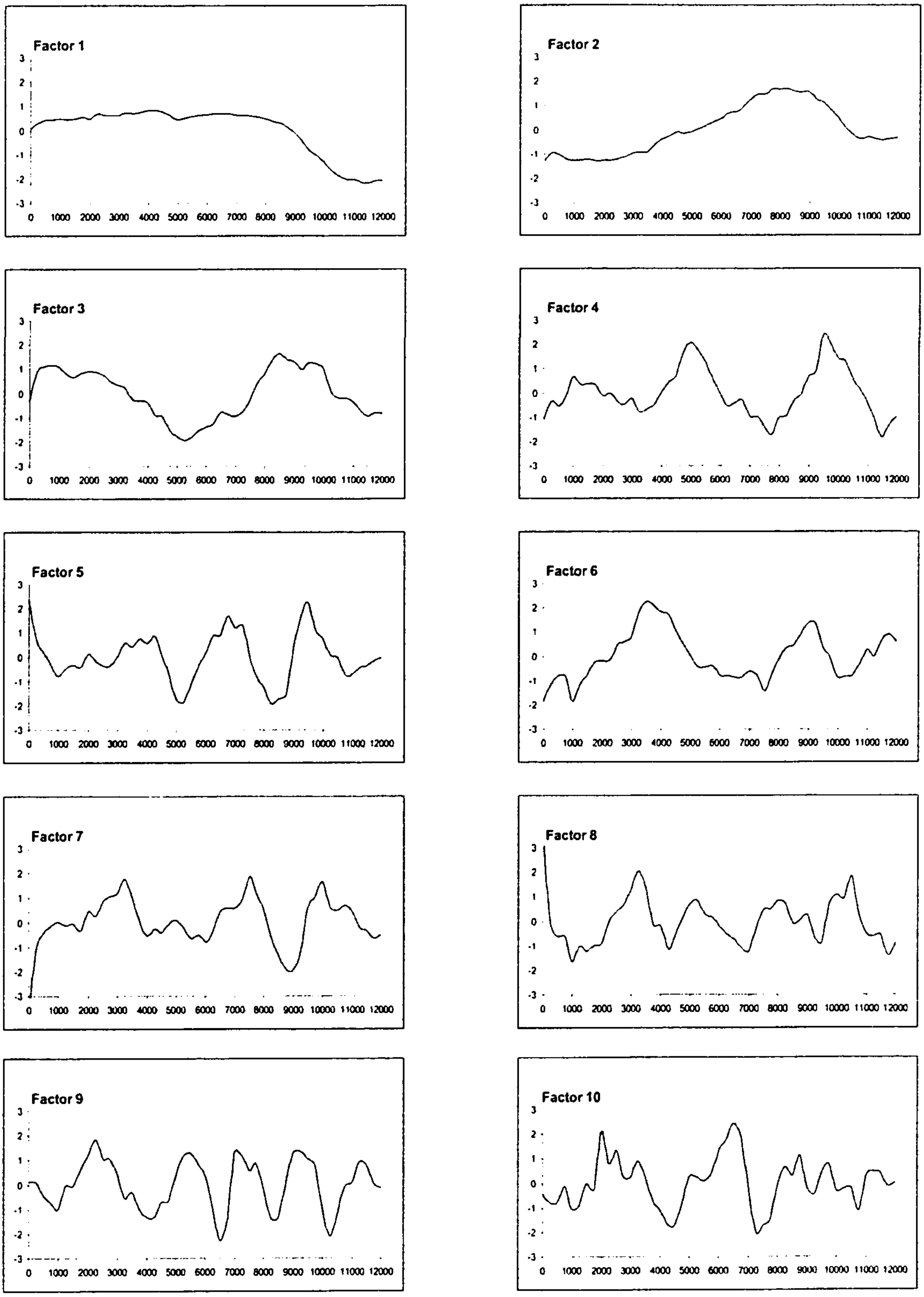


Figure F2.35: PCA on all *Quercus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

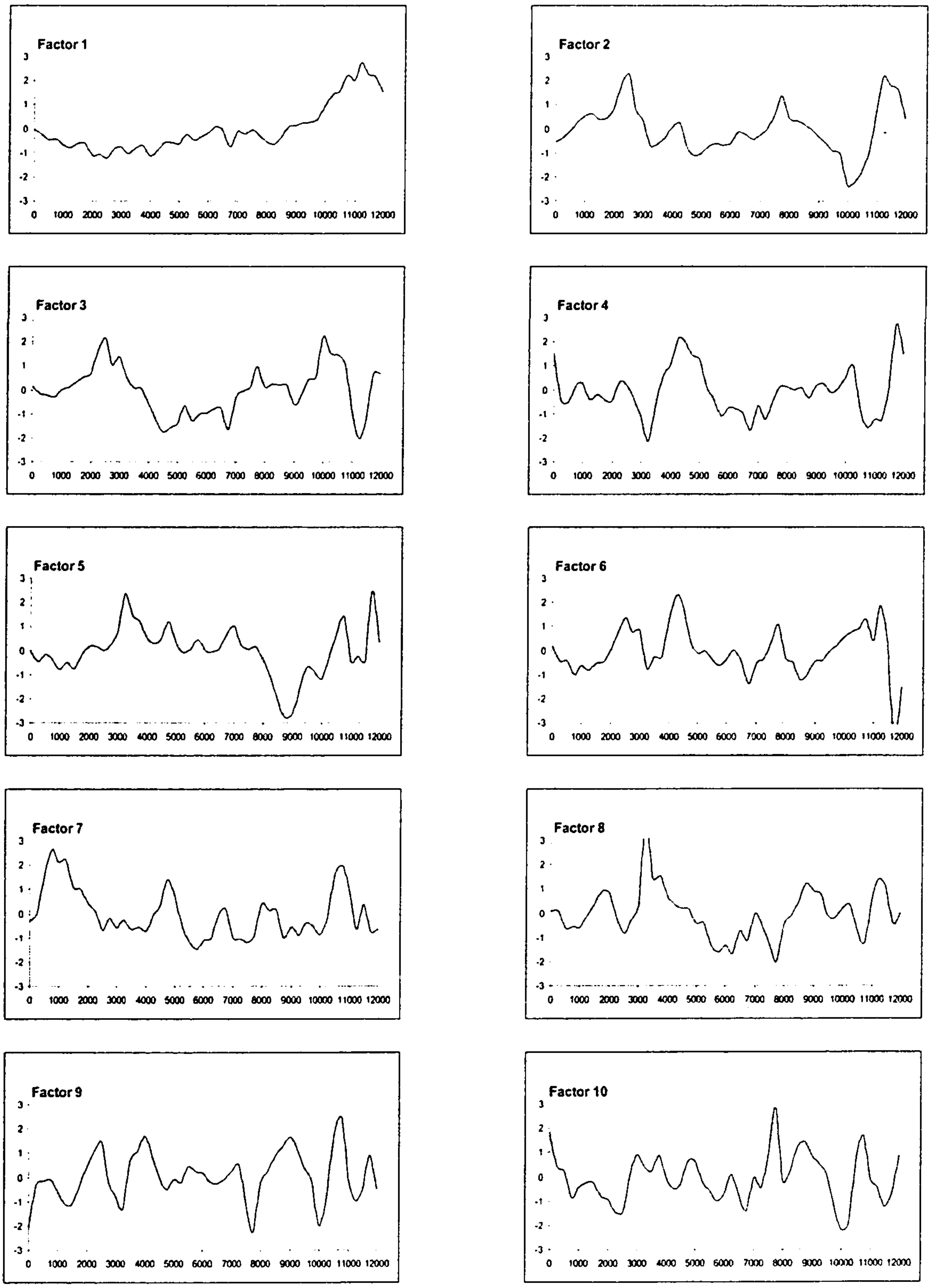


Figure F2.36: PCA on all *Ranunculaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

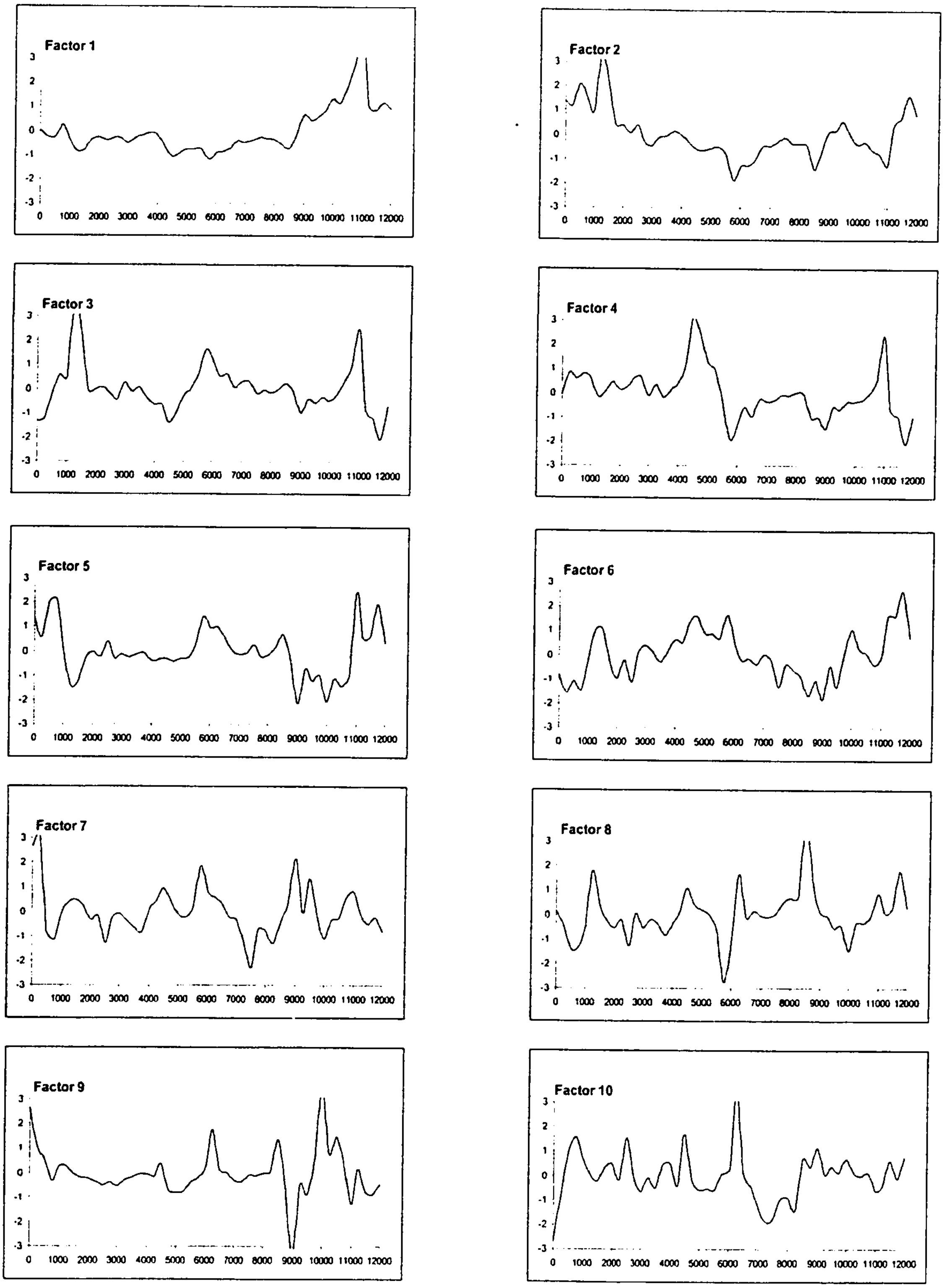


Figure F2.37: PCA on all *Rubiaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

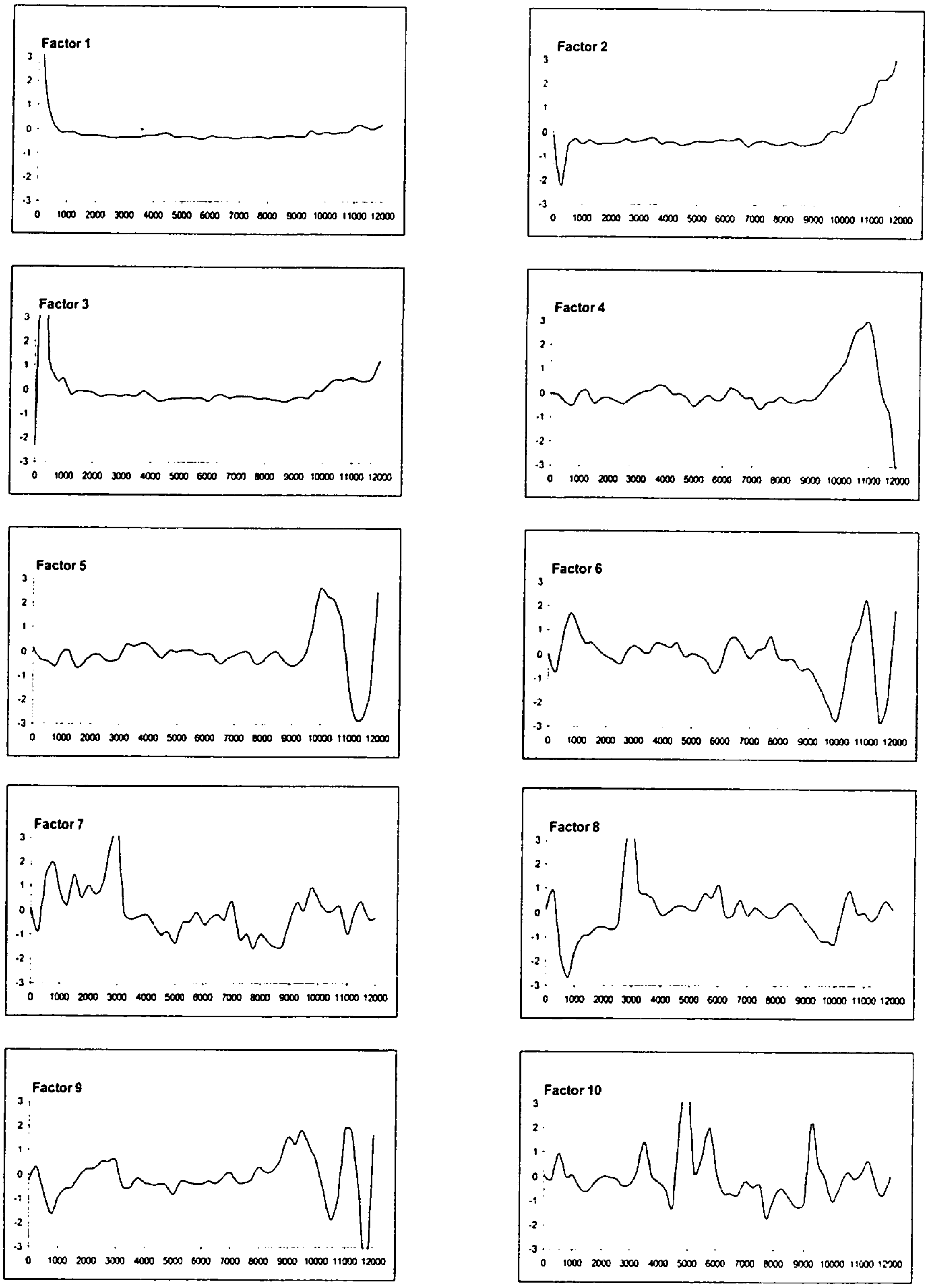


Figure F2.38: PCA on all *Rumex* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

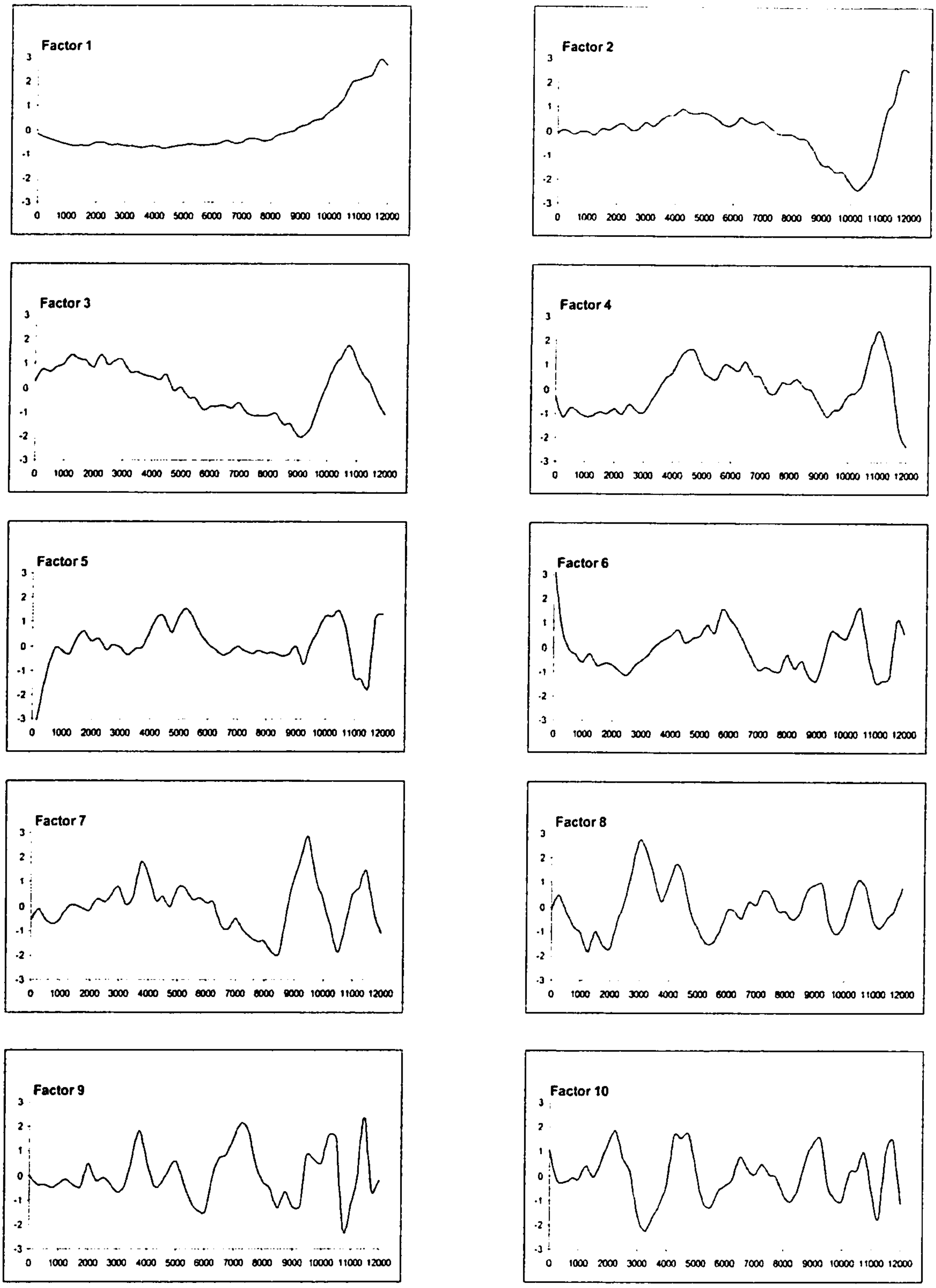


Figure F2.39: PCA on all *Salix* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

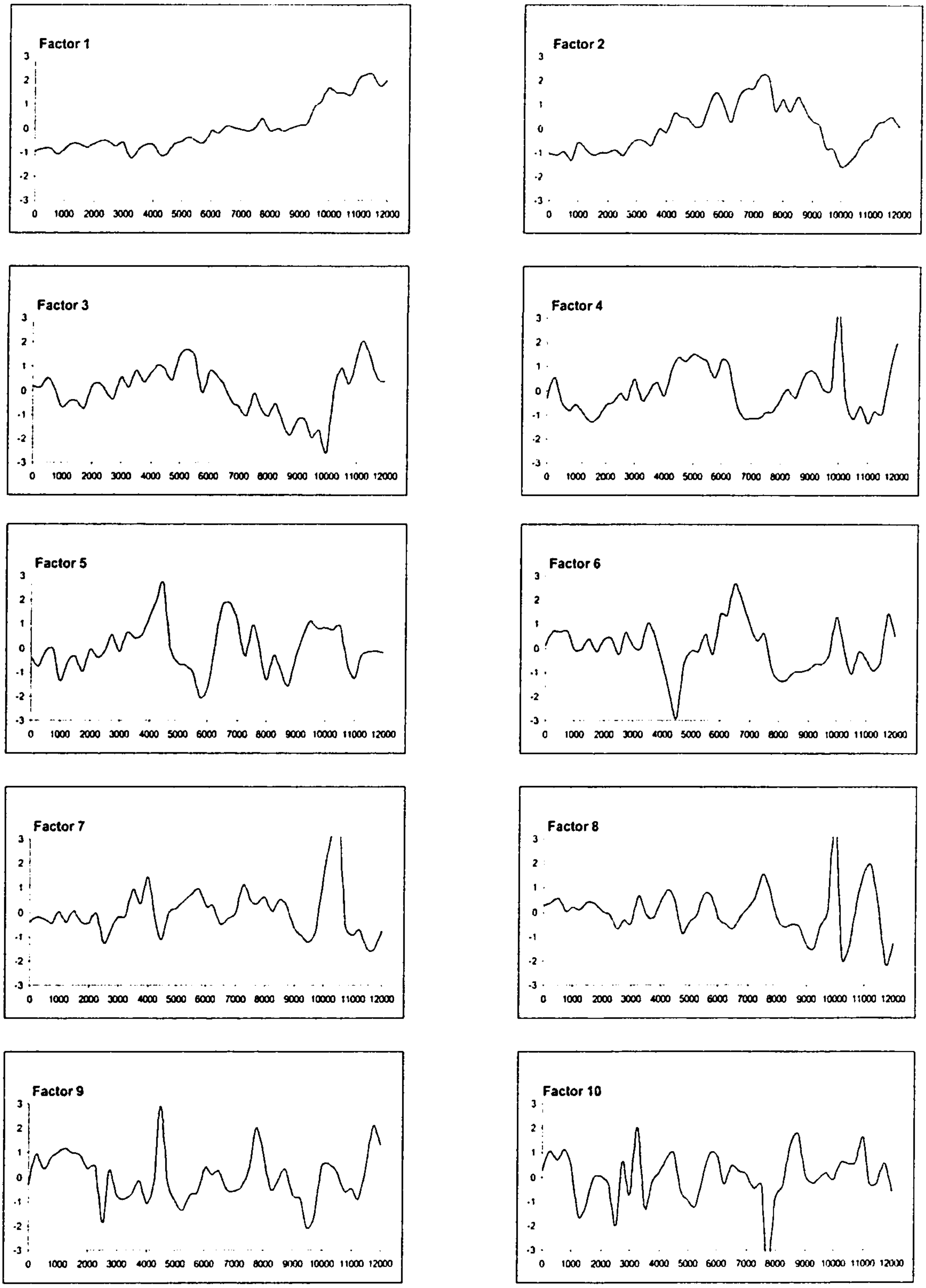


Figure F2.40: PCA on all *Sarcobatus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

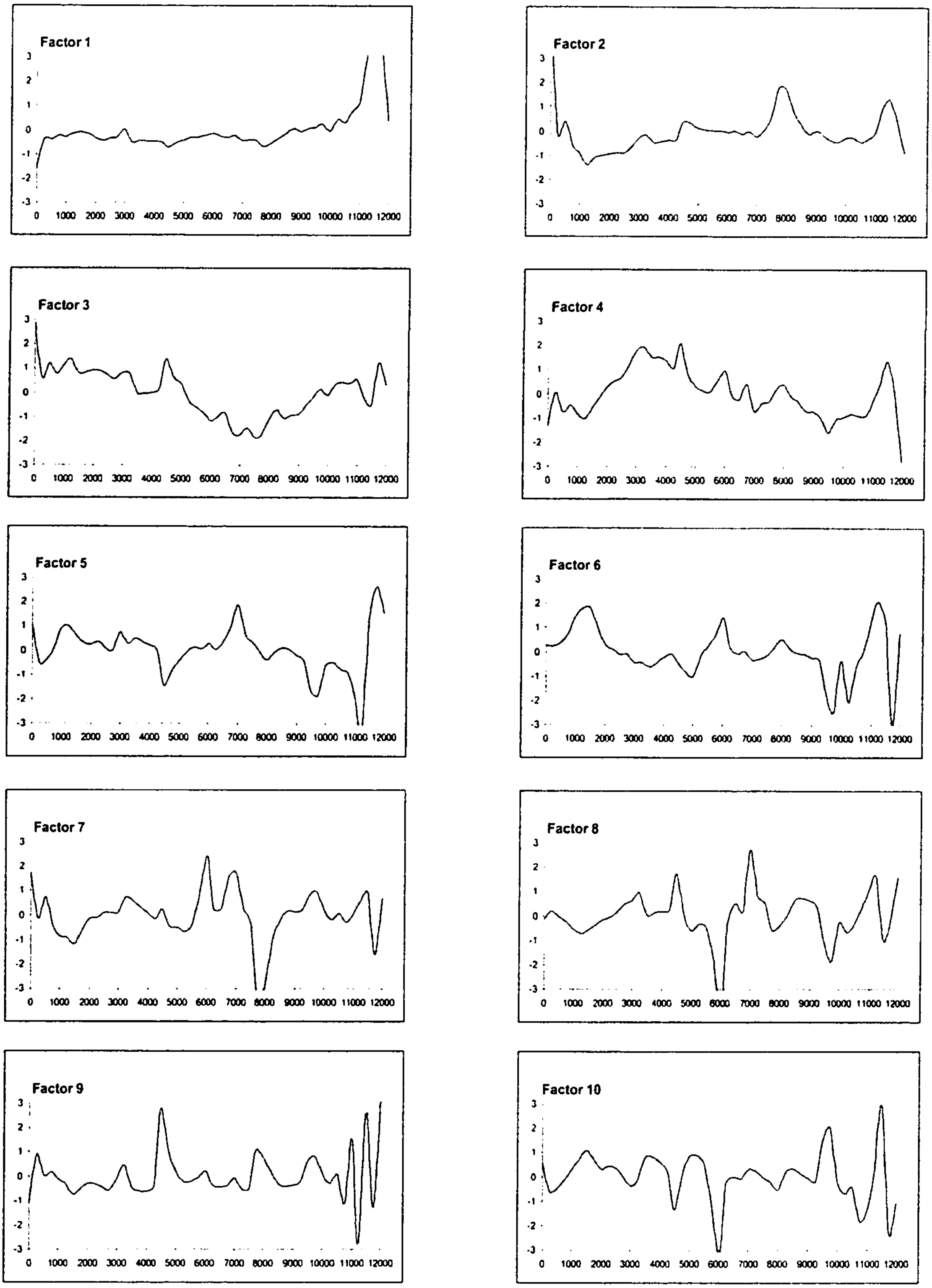


Figure F2.41: PCA on all *Saxifragaceae* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

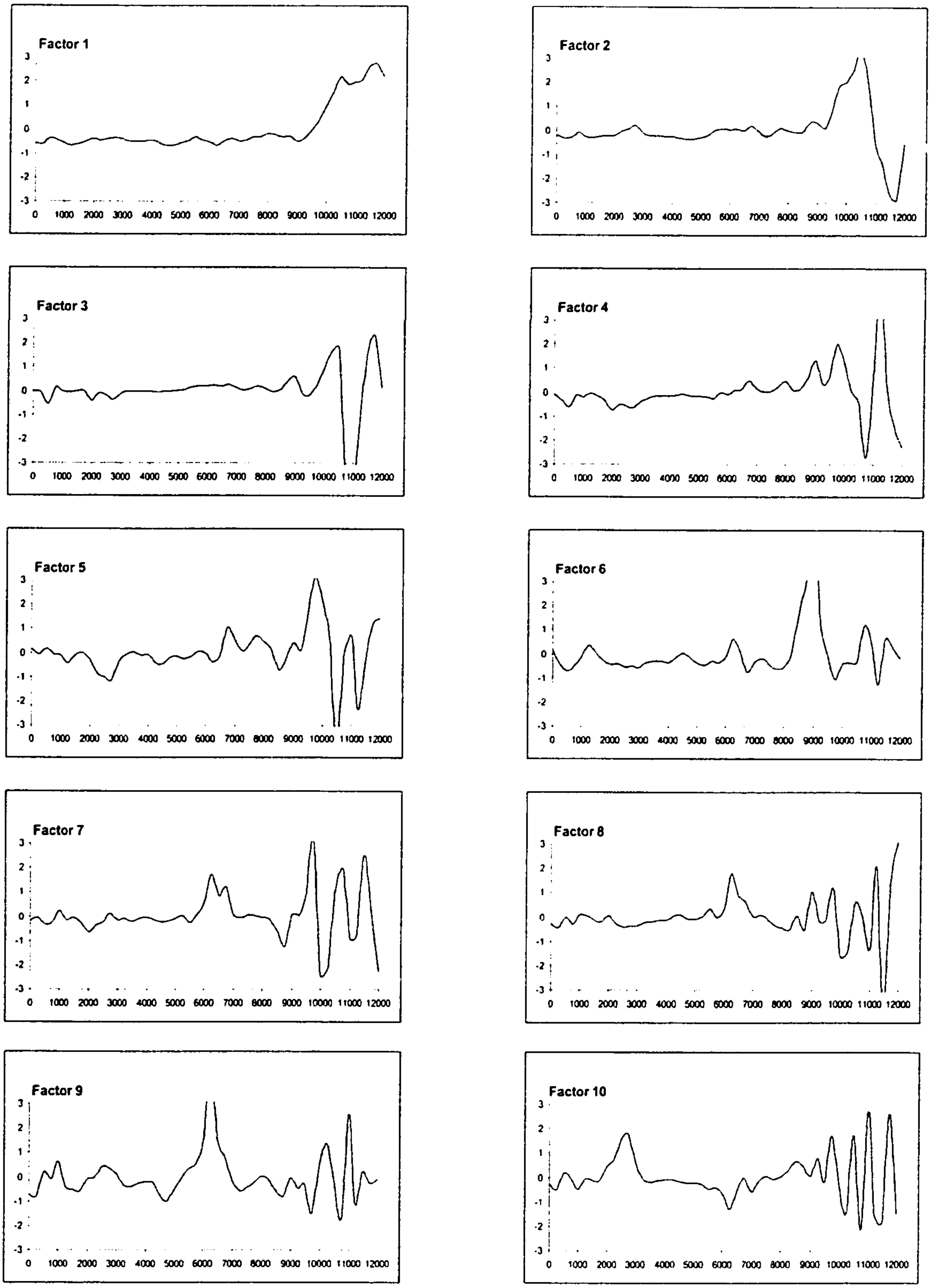


Figure F2.42: PCA on all *Shepherdia* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

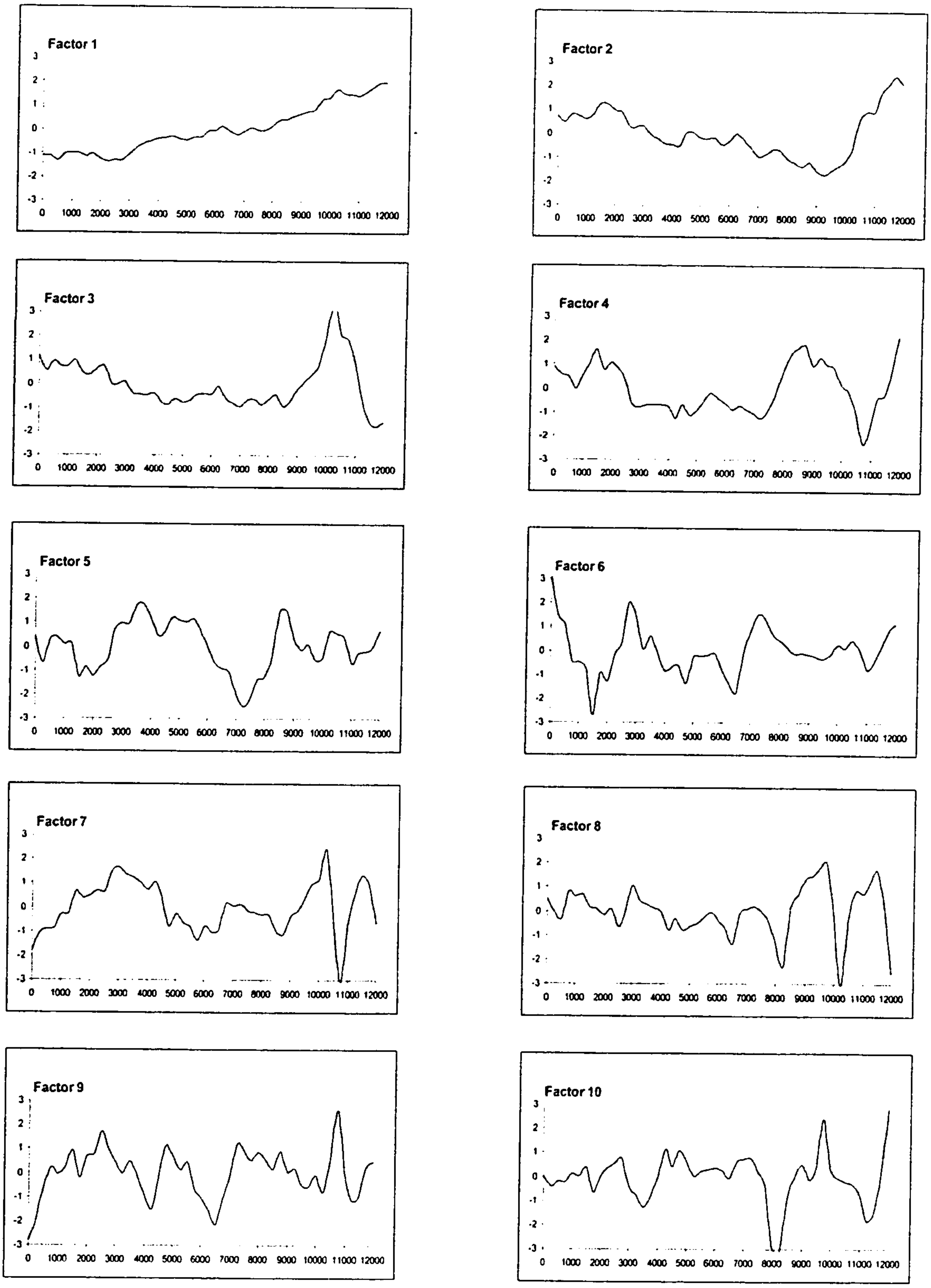


Figure F2.43: PCA on all *Sphagnum* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

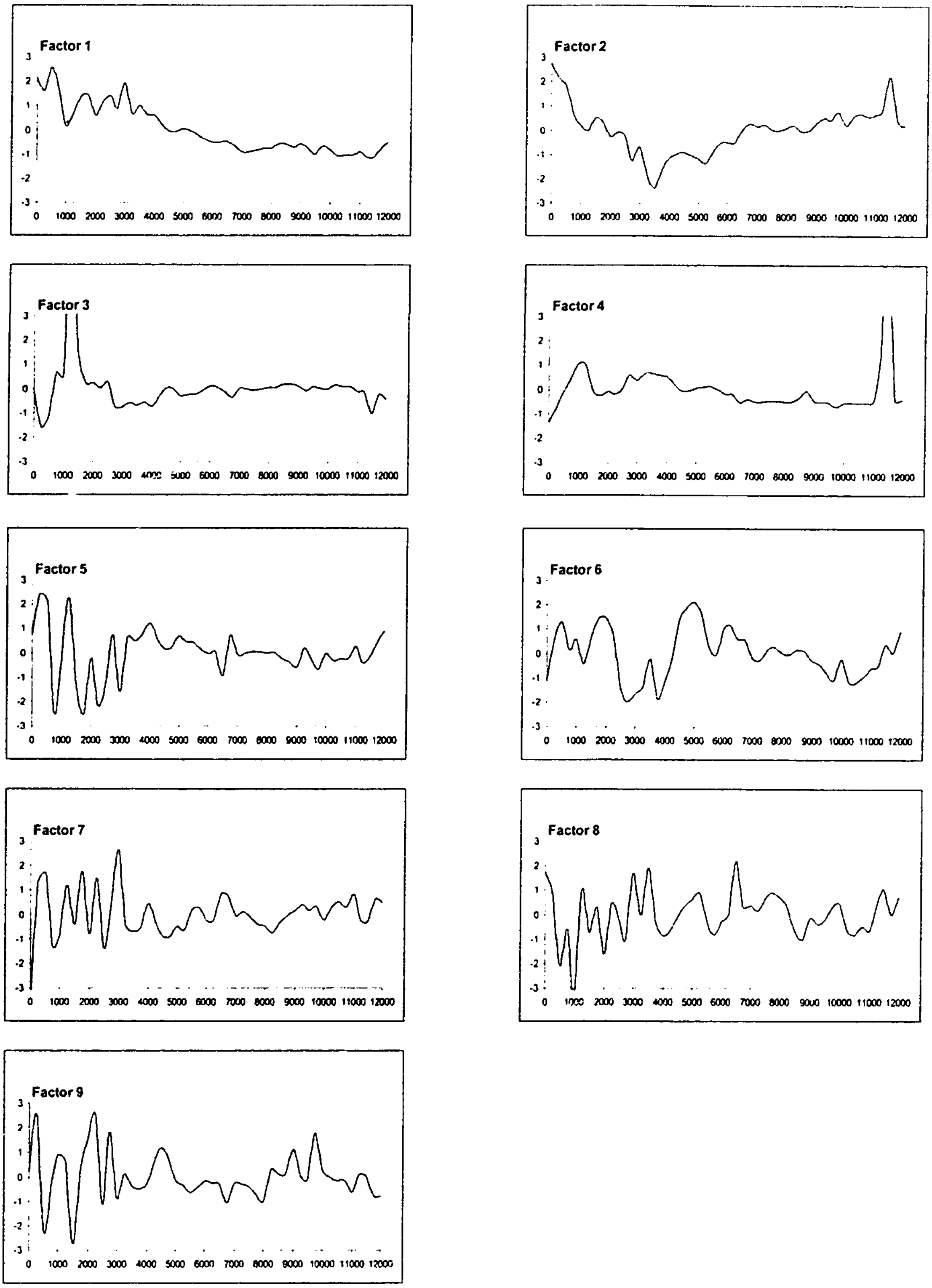


Figure F2.44: PCA on all *Taxodium* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

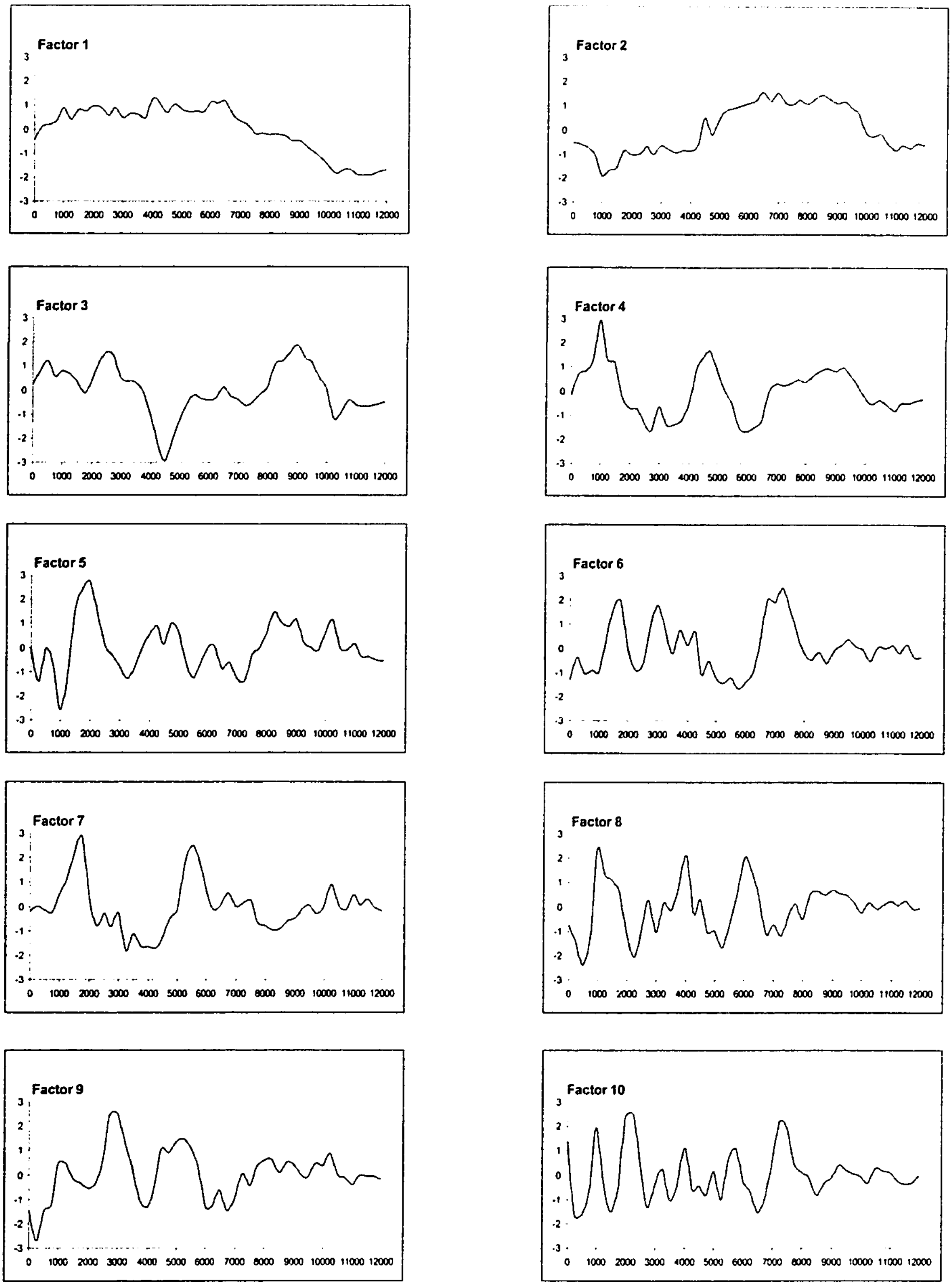


Figure F2.45: PCA on all *Tilia* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

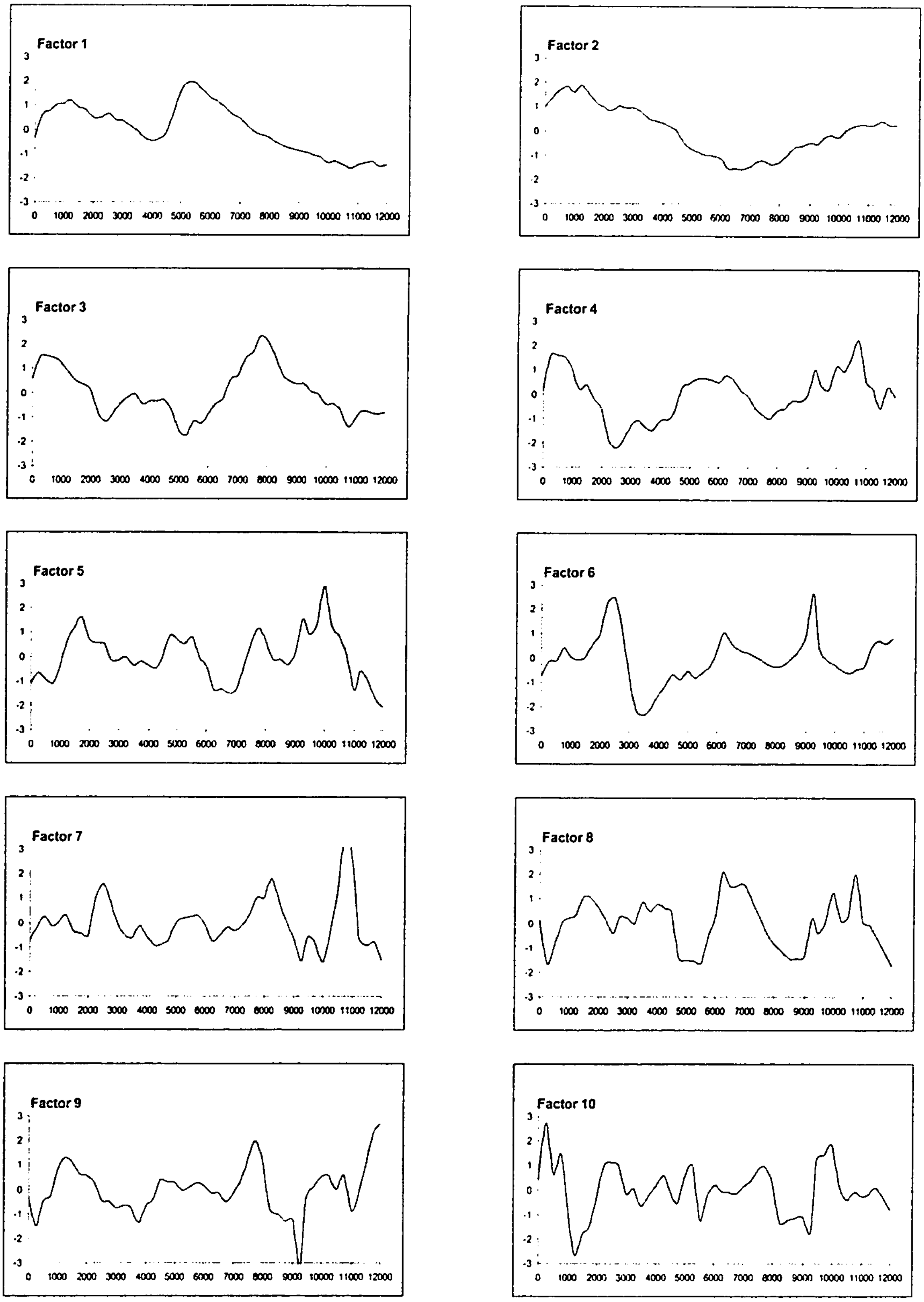


Figure F2.46: PCA on all *Tsuga* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

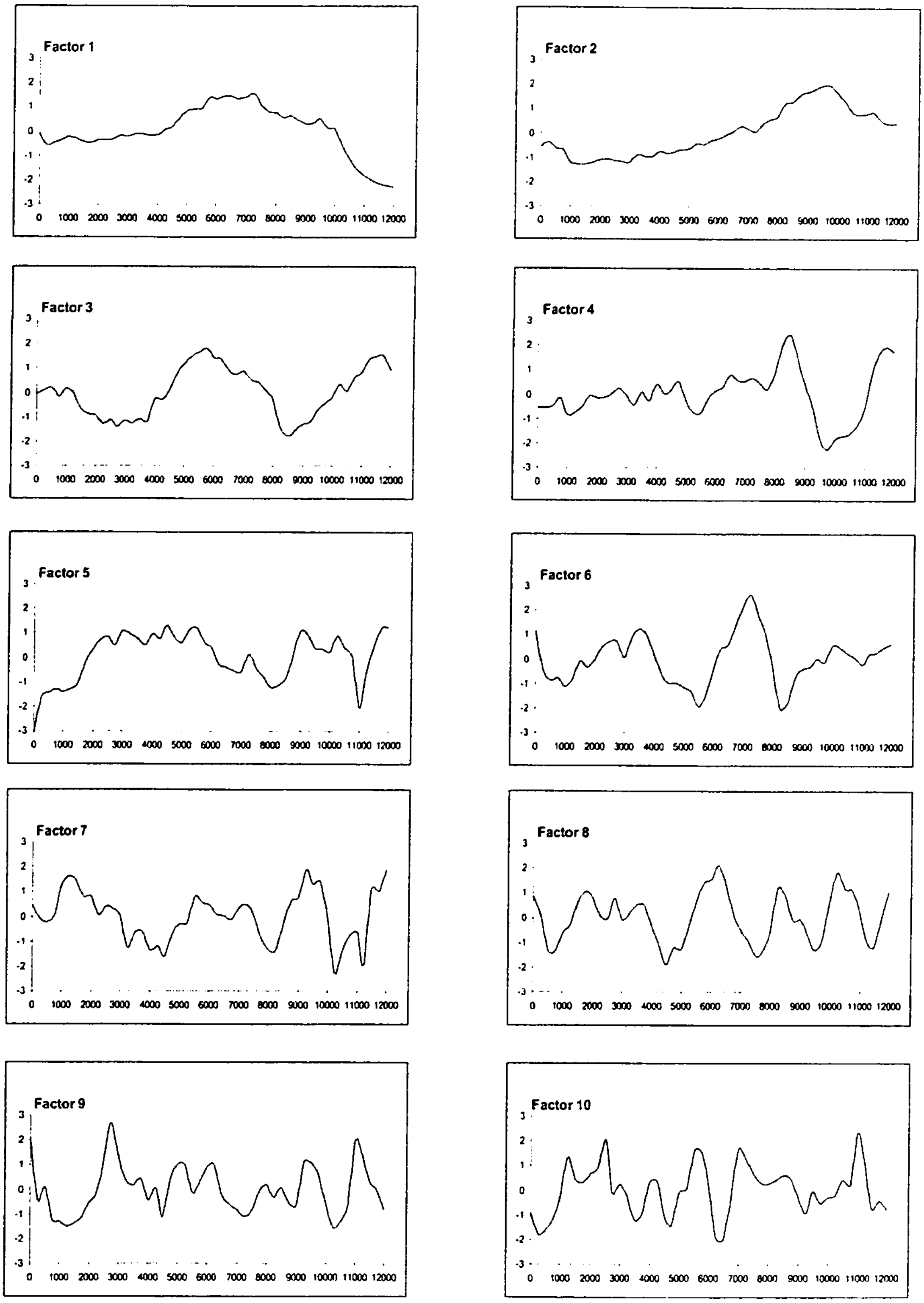


Figure F2.47: PCA on all *Ulmus* taxon in space and time. X-axis = time (C14); Y-axis= PCA scores.

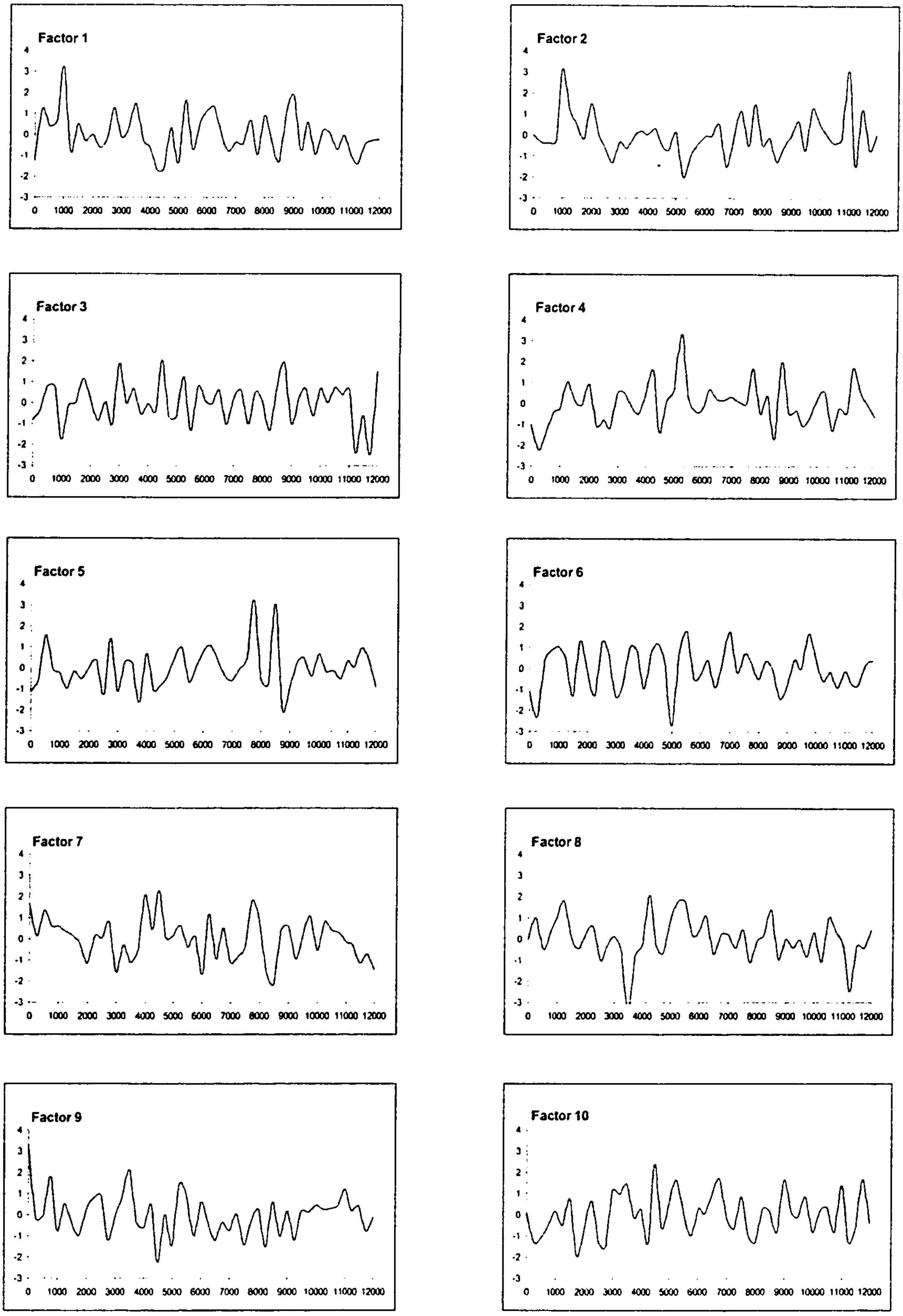


Figure F3: PCA on a random dataset (gamma distribution). X-axis = time (C14); Y-axis= PCA scores.

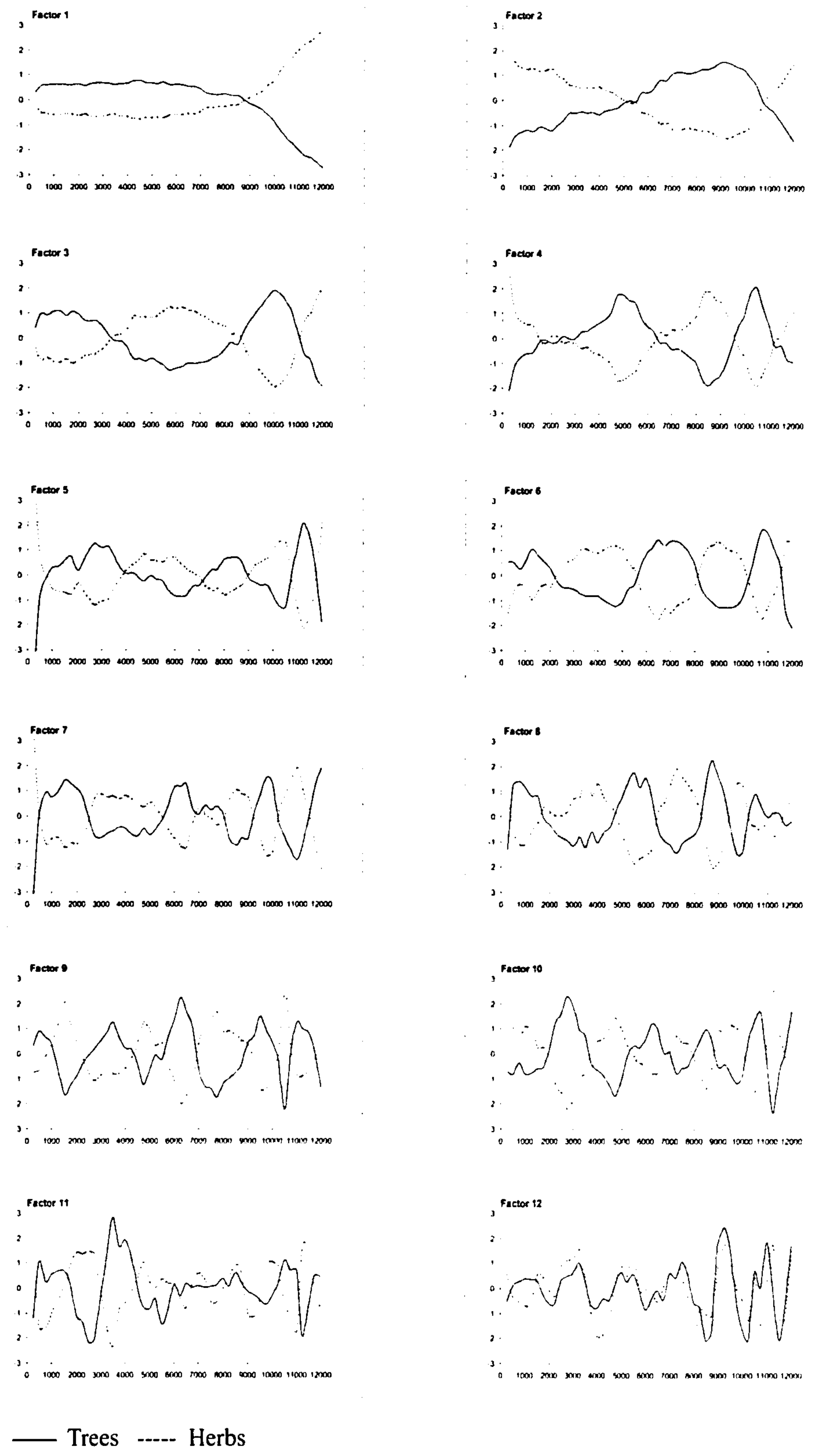


Figure F4: Trees and herbs inverse relationship for first 12 factors using subgroup A and B in Table F2. X-axis=time; y-axis= factor scores. Notice the mirror image until factor 10. The relationship begins to break down in factor 11 and more evident in factor 12.

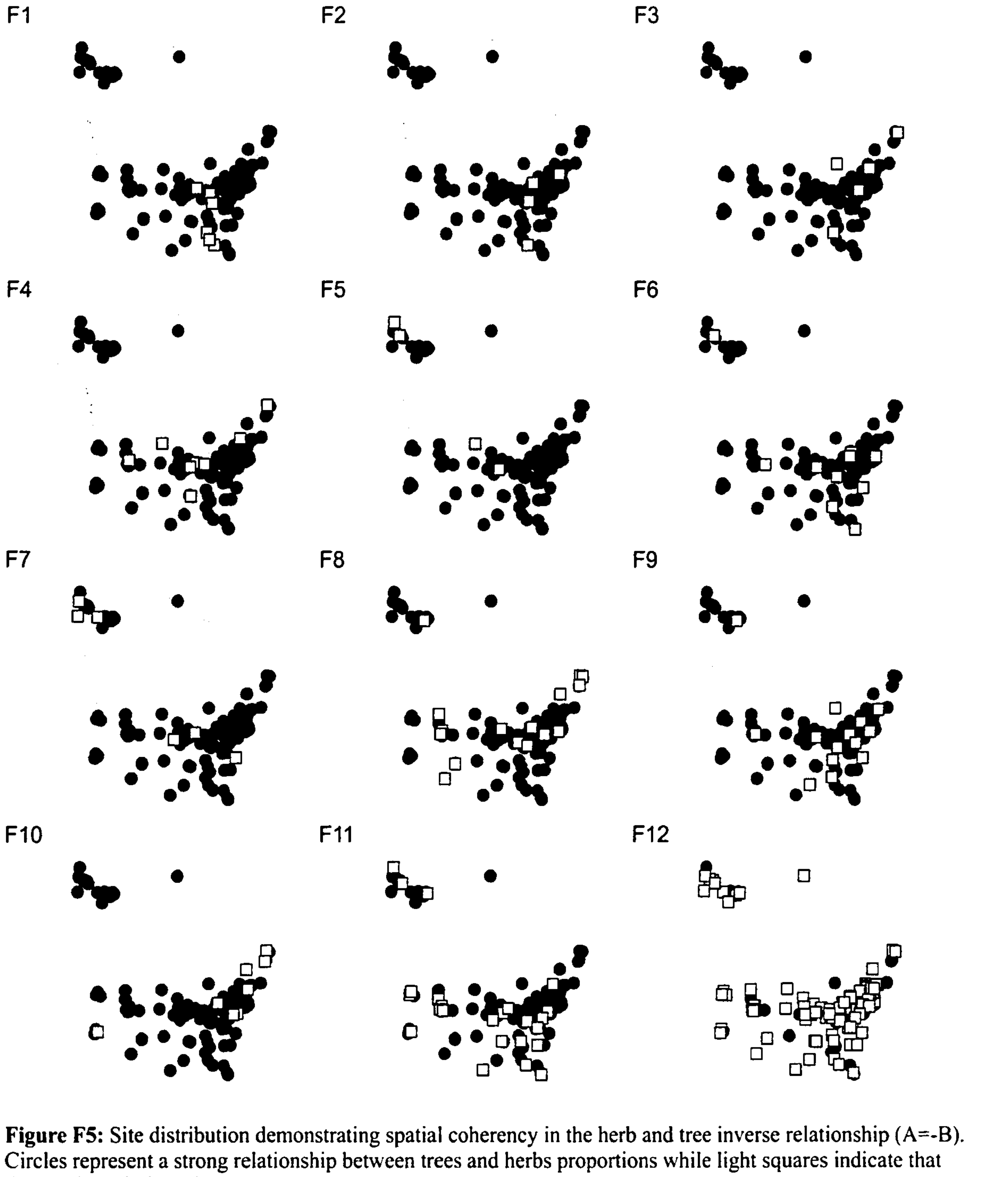


Figure F5: Site distribution demonstrating spatial coherency in the herb and tree inverse relationship (A=-B). Circles represent a strong relationship between trees and herbs proportions while light squares indicate that the relationship is no longer true.

APPENDIX G – "R" SOFTWARE PROGRAMS FOR MAT

Interpolate.fossil.f

/* This R function interpolates the fossil cores to target dates using a linear interpolation (e.g. at 100 years interval).
Developped by M.C. Sawada, 2001; applied by A.E. Viau */

```
function(fossilDataFrame, colFirstTaxa = 10, colLastTaxa = 98, colCurrDates = 8, colEnum = 1, colSiteNum = 2,
colSiteName = 3, colLong = 4, colLat = 5, target.dates = targets)
{
  holdnames <- as.vector(unlist(names.f(fossilDataFrame)))  #
  #get enumbers and the date column plus the columns with the proportions/percentages of each taxon.
  datainterp1 <- as.matrix(fossilDataFrame[, c(colEnum, colCurrDates, colFirstTaxa:colLastTaxa)]) #
  #make no dimnames to speed up calculations
  dimnames(datainterp1) <- NULL  #
  #get list of enumbers for each site
  enumbers <- as.numeric(unique(datainterp1[, 1]))  #
  newcolend <- dim(datainterp1)[2]
  outtable <- vector("numeric")  #
  tstart <- proc.time()[1]
  #ttime <- proc.time()[1] - currtime
  #      currtime <- proc.time()[1]

  for(i in 1:length(enumbers)) {
    datainterp <- datainterp1[which(datainterp1[, 1] == enumbers[i]), ]  #
  #  write(c("prick", "dick"), "c:\\asdf.txt", ncolumns = 2, append = T)
    if(length(dim(datainterp)) == 2) {
      interpolatedPrcts <- data.frame(apply(datainterp[, 3:newcolend], 2, approx.multi.f, datainterp[, 2],
target.dates)) #
      inter <- cbind(ENUM = rep(fossilDataFrame[which(fossilDataFrame[, colEnum] == enumbers[i])[1],
colEnum], length(target.dates)), SITENUM = rep(
        fossilDataFrame[which(fossilDataFrame[, colEnum] == enumbers[i])[1], colSiteNum],
length(target.dates)), NAME = rep(fossilDataFrame[which(
        fossilDataFrame[, colEnum] == enumbers[i])[1], colSiteName], length(target.dates)), LONGITUDE =
rep(fossilDataFrame[which(fossilDataFrame[, colEnum
        ] == enumbers[i])[1], colLong], length(target.dates)), LATITUDE =
rep(fossilDataFrame[which(fossilDataFrame[, colEnum] == enumbers[i])[1], colLat],
        length(target.dates)), AGECALBP = target.dates, interpolatedPrcts)
      #write.table(inter, "c:\\xnt.txt", sep = ",")
      outtable <- rbind(outtable, inter[which(is.na(inter[, 8]) == F), ])  #
    }
    write(proc.time()[1] - tstart, "c:\\timeinterp1.txt", append = T)  #
    tstart <- proc.time()[1]
  }
  write(paste("final", proc.time()[1] - tstart), "c:\\timeinterp1.txt", append = T)  #
  outtable
}
```

Calc.dist.2.f

/* This R function calculates the squared chord distances from the modern pollen set for each time slice in the fossil interpolated dataset. The output yields the modern sites which represent the best modern analog for a given fossil spectra at a given time.

Developped by M.C. Sawada, 2001 and applied by A.E. Viau. */

```
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))
{
  currtime <- proc.time()[1]
  nColsMatrix <- length(inFossil[, 1])
  LocMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
  PosMinRow <- matrix(NA, nrow = TopNAnalogs, ncol = nColsMatrix)
  bestAnalog <- 0
  colNames <- vector("character")
  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])) {
    ttime <- proc.time()[1] - currtime
    currtime <- proc.time()[1]
    write(ttime, "c:\\time150.txt", ncolumns = 1, append = T)
    data2 <- data3[i, ]
    sqdistVec <- (data2^0.5 - data1^0.5)^2
    #x <- apply(sqdistVec,2,sum)
    x = colSums(sqdistVec)
    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]
    LocMinRow[, i] <- x
    PosMinRow[, i] <- topn
  }
  colNames <- sitenames
  dimnames(LocMinRow) <- list(NULL, colNames)
  dimnames(PosMinRow) <- list(NULL, colNames)
  write(c("Final Time", as.character(proc.time()[1] - currtime)), "c:\\time.txt", ncolumns = 2, append = T)
  list(x = inFossil[, LongFossil], y = inFossil[, LatFossil], sqdist = LocMinRow, position = PosMinRow)
}
```

NAPD Fossil Recon: Data Description – March 8th 2002

MatreconMar62002.csv Contains

1 E.	31 Celtis	61 Magnoliaceae	91 Selaginella	121 MARCH	151 LONG1
2 Site.	32 Cephalanthus	62 Malvaceae	92 Shepherdia	122 APRIL	152 LAT1
3 Descriptor	33 Cercocarpus	63 Myrica	93 Sphagnum	123 MAY	153 LONG2
4 LonDD	34 Chenopodiaceae.Amaranthaceae	64 Nyssa	94 Taxodium	124 JUNE	154 LAT2
5 LatDD	35 Chrysolepis	65 Onagraceae	95 Thalictrum	125 JULY	155 LONG3
6 Elevation	36 Clethra	66 Ostrya.Carpinus	96 Tilia	126 AUGUST	156 LAT3
7 Sample.	37 Corylus	67 Oxyria	97 Tsuga	127 SEPTEMBER	157 LONG4
8 calMean	38 Cupressaceae	68 Papaveraceae	98 Ulmus	128 OCTOBER	158 LAT4
9 c14AgeBP	39 Cyperaceae	69 Picea.undiff	99 JanT	129 NOVEMBER	159 LONG5
10 Abies	40 Dryas	70 Picea.glauca	100 JulT	130 DECEMBER	160 LAT5
11 Acer	41 Ephedra	71 Picea.mariana	101 AnnPrecip	131 sq1st	161 LONG6
12 Alnus.undiff	42 Ericaceae.undif	72 Pinus.undif	102 GDD0	132 sq2nd	162 LAT6
13 Alnus.crispa	43 Ericaceae.Chamadaphne.Ledum	73 Pinus.strobus	103 GDD5	133 sq3rd	163 LONG7
14 Alnus.rugosa	44 Ericaceae.Empetrum	74 Plantaginaceae	104 Tcld	134 sq4th	164 LAT7
15 Ambrosia	45 Ericaceae.Vaccinium	75 Platanus	105 Twm	135 sq5th	165 LONG8
16 Amorpha	46 Eriogonum	76 Poaceae	106 MIPT	136 sq6th	166 LAT8
17 Apiaceae	47 Euphorbiaceae	77 Polygonum	107 PJANUARY	137 sq7th	167 LONG9
18 Aquifoliaceae	48 Fabaceae	78 Polypodiaceae	108 PFEBRUARY	138 sq8th	168 LAT9
19 Arceuthobium	49 Fagus	79 Populus	109 PMARCH	139 sq9th	169 LONG10
20 Asteraceae	50 Fraxinus	80 Potentilla	110 PAPRIL	140 sq10th	170 LAT10
21 Artemisia	51 Juglans	81 Prosopis	111 PMAY	141 Pos1st	
22 Betula.undif	52 Larix.Pseudotsuga	82 Pteridium	112 PJUNE	142 Pos2nd	
23 Brassicaceae	53 Larrea	83 Quercus	113 PJULY	143 Pos3rd	
24 Cactaceae	54 Liliaceae	84 Ranunculaceae	114 PAUGUST	144 Pos4th	
25 Campanulaceae	55 Liquidambar	85 Rhamnaceae	115 PSEPTEMBER	145 Pos5th	
26 Caprifoliaceae	56 Lycopodium.undiff	86 Rubiaceae	116 POCTOBER	146 Pos6th	
27 Carya	57 Lycopodium.annotinum	87 Rumex	117 PNOVEMBER	147 Pos7th	
28 Caryophyllaceae	58 Lycopodium.clavatum	88 Salix	118 PDECEMBER	148 Pos8th	
29 Castanea	59 Lycopodium.complanatum	89 Sarcobatus	119 JANUARY	149 Pos9th	
30 Ceanothus	60 Lycopodium.selago	90 Saxifragaceae	120 FEBRUARY	150 Pos10th	

FIELD DESCRIPTIONS

Column Range	Description
1-9	From Fossil Database
10-98	Proportions of taxa for the best modern analog for each level.
99-106	Bartlein's Climate Recon for each level
107-118	Leeman's & Cramer Monthly Precipitation
119-130	Leeman's & Cramer Monthly Temperature
131-140	Squared Chord Distances for 1st Analog (col 131) to 10th Best Modern Analog (Col 140)
141-150	These fields each contain the row positions within the file ModernMarch82002Percent.csv of the 1st (col 141) to 10th (col 150) best modern analogs. Within the file ModernMarch82002Percent.csv there is a [KEY] field that contains the numbers 1-4590. These are used to locate a particular modern sample. If you want to reconstruct the 2nd best modern analog then do a one-to-one join from the modern field called [KEY] to the field [Pos2nd] in the fossil reconstruction table MatreconMar62002.csv.

MODERNMARCH82002PERCENT.CSV CONTAINS

Pos Name	Pos Name	Pos Name	Pos Name	Pos Name	Pos Name
1 KEY	26 Carya	51 Larix.Pseudotsuga	76 Polygonum	101 GDD0	126 SEPTEMBER
2 ID	27 Caryophyllaceae	52 Larrea	77 Polypodiaceae	102 GDD5	127 OCTOBER
3 ENUM	28 Castanea	53 Liliaceae	78 Populus	103 Tcld	128 NOVEMBER
4 SITENUM	29 Ceanothus	54 Liquidambar	79 Potentilla	104 Twm	129 DECEMBER
5 LONG	30 Celtis	55 Lycopodium.undiff	80 Prosopis	105 MIPT	
6 LAT	31 Cephalanthus	56 Lycopodium.annotinum	81 Pteridium	106 PJANUARY	
7 ELEV	32 Cercocarpus	57 Lycopodium.clavatum	82 Quercus	107 PFEBRUARY	
8 GTOPELV	33 Chenopodiaceae.Am.	58 Lycopodium.complanatum	83 Ranunculaceae	108 PMARCH	
9 Abies	34 Chrysolepis	59 Lycopodium.selago	84 Rhamnaceae	109 PAPRIL	
10 Acer	35 Clethra	60 Magnoliaceae	85 Rubiaceae	110 PMAY	
11 Alnus.undiff	36 Corylus	61 Malvaceae	86 Rumex	111 PJUNE	
12 Alnus.crispa	37 Cupressaceae	62 Myrica	87 Salix	112 PJULY	
13 Alnus.rugosa	38 Cyperaceae	63 Nyssa	88 Sarcobatus	113 PAUGUST	
14 Ambrosia	39 Dryas	64 Onagraceae	89 Saxifragaceae	114 PSEPTEMBER	
15 Amorpha	40 Ephedra	65 Ostrya.Carpinus	90 Selaginella	115 POCTOBER	
16 Apiaceae	41 Ericaceae.undif	66 Oxyria	91 Shepherdia	116 PNOVEMBER	
17 Aquifoliaceae	42 Ericaceae.Cha.ledum	67 Papaveraceae	92 Sphagnum	117 PDECEMBER	
18 Arceuthobium	43 Ericaceae.Empetrum	68 Picea.undiff	93 Taxodium	118 JANUARY	
19 Asteraceae	44 Ericaceae.Vaccinium	69 Picea.glauca	94 Thalictrum	119 FEBRUARY	
20 Artemisia	45 Eriogonum	70 Picea.mariana	95 Tilia	120 MARCH	
21 Betula.undif	46 Euphorbiaceae	71 Pinus.undif	96 Tsuga	121 APRIL	
22 Brassicaceae	47 Fabaceae	72 Pinus.strobus	97 Ulmus	122 MAY	
23 Cactaceae	48 Fagus	73 Plantaginaceae	98 JanT	123 JUNE	
24 Campanulace	49 Fraxinus	74 Platanus	99 JulT	124 JULY	
25 Caprifoliacea	50 Juglans	75 Poaceae	100 AnnPrecip	125 AUGUST	

FIELD DESCRIPTIONS

Field	Description
KEY	Primary Key (Relates to the file ModernMarch82002Percent.csv and col 141 to 150)
ID	Primary Key
ENUM	GPD enumber for GPD sites, -9999 for WW2000 and KG2000 sites
SITENUM	Original Site number from GPD (!= Enumber), Key Field from WW2000 and KG2000
SITENAME	Site names from original datasets. Site name from WW2000 is [Name1]+[Name2] fields
LONG	Longitude in full precision from original databases
LAT	*Laititude in full precisions from orignal databases*
ELEV	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	Elevation extracted from GTOPO30 for all sites (see end of paper for details on extraction).
9-97	Proportions of taxa for the best modern analog for each level.
98-105	Bartlein's Climate Recon for each level
106-117	Leeman's & Cramer Monthly Precipitation
118-129	Leeman's & Cramer Monthly Temperature

PROCEDURE FOR RUNNING MAT FOR NAPD IN R GUI SOFTWARE

Restoring the DataSet

In R:

load("c:\\andre.txt")

ls() #gives list of all objects in the session.

Data:

fossil.feb2002.percent #_dget("c:\\fossi.txt")
modern.feb2002.percent #_dget("c:\\mod.txt")

Functions:

Interpolation

To interpolate fossil pollen proportions to target dates use the function <interpolate.fossil.f> with these arguments:

> **interpolate.fossil.f**(fossilDataFrame, colFirstTaxa = 10, colLastTaxa = 98, colCurrDates = 8, colEnum = 1, colSiteNum = 2, colSiteName = 3, colLong = 4, colLat = 5, target.dates = targets)

targets: e.g., a vector of dates at which to interpolate fossil pollen proportions.

```
targets<-c(2350, 2450, 8200)
```

will provide fossil pollen proportions as spectra linearly interpolated to these target dates for all sites in the fossilDataFrame.

Example:

```
fossil.may2002.interp250=interpolate.fossil.f(fossil.feb2002.percent)
```

Fossil Reconstructions

To get a reconstruction, the first step is to run the function <calc.dist.2.f> using raw-modern file (inModern) and either a raw fossil file or an interpolated to target date fossil file (inFossil). The function has these arguments:

```
calc.dist.2.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))
```

Example:

```
recon.may250_calc.dist.2.f(fossil.may2002.interp250,modern.feb2002.percent,4,5,8,96,4,5,7,95,1
0, sitenames = as.vector(unlist(fossil.may2002.interp250[, 3])))
```

Session

```
names.f(fossil.feb2002.percent) #get names of columns in dataframes
names.f(modern.feb2002.percent)
args(interpolate.fossil.f) #gives a list of arguments for functions
```

Making Target Dates

```
> new.targets_seq(0,2000,250)
> new.targets
[1]  0 250 500 750 1000 1250 1500 1750 2000
```

Doing Interpolation

```
fossilInterp<-
interpolate.fossil.f(fossil.feb2002.percent[1:100,],10,98,8,1,2,3,4,5,seq(0,2000,250))
```

Subset example: fossil.feb2002.percent[1:100,c(1,5,9)] #Subsets Rows 1:100 and columns 1,5 and 9.

Write Ouput to CSV file

write.table(fossilInterp,"c:\\andretest.txt",sep=",",quote=F,row.names=F)

Reconstruction

args(calc.dist.2.f)
recon.andre.test_*calc.dist.2.f*(fossilInterp,modern.feb2002.percent,4,5,8,96,4,5,7,95,*10*, sitenames = as.vector(unlist(fossilInterp [, 3]))) #green 10 means 10 best analogs returned

names.f(recon.andre.test)

```
    [,1] [,2]
[1,] "1" "x"      # The x coordinate of the fossil level
[2,] "2" "y"      # Same for y
[3,] "3"  "sqdist"  # Each row contains the squared chord distance for the nth best modern analog for all sites.
[4,] "4"  "position" # Each row contains the position of the nth best modern analog for all sites.
```

Creating Recon.

1. Goto modern file < modern.feb2002.percent> and determine what variables you want to extract for each fossil level – get their column numbers .g., LC July Temperature is column 123
2. Add columns of reconstructed variables to the fossil data by using the best analog's position in the modern file as determined in the reconstruction$position row e.g., if the reconstruction is called recon.andre.test then the best modern analog would be in recon.andre.test$position[1,] the first row of the "position" component of the reconstruction.

Example, appending July Temperature Reconstructions for 1 and 2nd best modern analogs to the fossilInterp table and saving this in the variable <july.recon.andre.test>.

july.recon.andre.test<-
cbind(fossilInterp[,1:6],recJul=modern.feb2002.percent[recon.andre.test$position[1,],123],recJul2nd = modern.feb2002.percent[recon.andre.test$position[2,],123])

Output to table:

write.table(july.recon.andre.test,"c:\\andrJulRec.txt",sep=",",quote=F,row.names=F)

APPENDIX H – SUPPLEMENTAL INFORMATION ON CLIMATE VARIABILITY

INTRODUCTION

There is still debate on the causes and nature of the recent century-long temperature increase (Jones *et al.*, 1999; Mann *et al.*, 1999; IPCC, 2001). This dramatic recent temperature change has been related to human-induced CO_2 concentrations (e.g. IPCC, 2001). With the increasing concentrations of greenhouse gases (GHGs), such as CO_2, CH_4, NO_x, CFCs and others modifying the atmosphere's composition, we still do not know the full implications this will have on our future climate. If a climate change is to occur due to a human-induced "enhanced greenhouse effect", we still cannot answer how or what time or what spatial scales these changes will operate in. Simply put, our present models and knowledge of the climate system do not allow confident forecasting on how fast and extensive the climate will change outside the boundaries of natural climate variability.

An increase in the concentration of GHGs results in more of the outgoing energy from the Earth's surface to be absorbed by the atmosphere. One of the consequences of an enhanced greenhouse effect is higher average temperature due to the warming of the lower atmosphere (IPCC, 1995, 2001). However, a basic fundamental problem concerns the way GHGs affect the flow of energy through the climate system.

The climate has changed in the past on different time and space scales and it will continue to change in the future. Some large-scale climate changes operate over timescales of thousands of years and therefore are hardly perceptible during a human lifetime. An alternative model is that climate change occurs rather suddenly in a step-like function rather than gradually (Wendland and Bryson, 1974; Taylor *et al.*, 1993). Some large-scale climate changes seem to have occurred synchronously in different regions of the world on decadal to millennial timescales. These climate "switches" from one mode to another may have affected various regions and their associated cultures. It is this natural variability in the climate that poses major difficulties in understanding the impact of recent anthropogenic forcing on the climate. In any event, understanding the climate system is fundamental to policy makers and scientific advancement whether an anthropogenic or natural forcing will bring the climate system to "switch" to its next mode.

NATURE OF THE PROBLEM

In order to understand the mechanisms and physical processes responsible for the climate, it is necessary to describe the system as well as its key components and behaviour. The climate can be considered as the "averaged weather" over a period of time. Broadly speaking, the geographical distribution of climates, based on temperature and precipitation, show wärm and moist climates in the tropics, warm and very dry in the subtropics, temperate and moist climate in mid-latitudes and cold and dry climates in polar and sub-polar regions. However, this description does not imply that on local to regional scales, these climates prevail as rigorously as described above. Other key elements, such as land-sea contrast, topography, nature of the surface, vegetation and snow and ice have to be considered. The climate, as a thermo-hydrodynamic system, is complex due to the non-linear interactions between its subsystems. The subsystems include the atmosphere, hydrosphere, including the oceans and freshwaters, cryosphere, the underlying lithosphere, and biosphere, including both marine and terrestrial biosystems. Although these subsystems are very different, they are all linked together by fluxes of mass, energy, and momentum (Saltzman, 1985).

A system in which the transfers of matter and energy are allowed between the system and exterior is called an "open" system (Peixoto and Oort, 1992). A cascading system is a chain of open subsystems, which are dynamically linked by a cascade of matter or energy so that the output of mass or energy from one subsystem becomes the input for the next subsystem (Peixoto and Oort, 1992). The input can also be partially stored in one of the subsystems controlling the amounts of mass or energy available as output for the next subsystem. However, this modulation may also come about from external forcings.

These form an internal system embedded in an external one consisting of the remaining subsystems. If we assume that the time scale of the behaviour of the internal system is much smaller than the external, then the external system can be considered in a steady state (Peixoto and Oort, 1992). The external system may then influence the behaviour of the internal leading to a forced adjustment of its properties. This is often referred to as an external forcing. For example, an initial perturbation within one subsystem may eventually fade away or amplify. If it amplifies, the perturbation will result in a new and completely different state of the internal system. However, if an almost periodic external forcing causes the perturbation, it may lead to an oscillatory response of the internal system (Peixoto and Oort, 1992).

The climate system may undergo spontaneous reorganisation into quasi-stable states on various time scales (Rind, 1999). These variations may be caused by external forcings, or they may be independent and due to internal feedbacks and instabilities. Termed "free variations", these are associated with nonlinearities inside the system, which may create the conditions for the growth or damping of small instabilities (Peixoto and Oort, 1992). Negative feedbacks will have the influence of stabilizing or damping the effect while positive feedbacks can amplify an initial perturbation independently of any external forcings.

The climate system is characterized by subsystems that are dynamically linked to form a complex system modulated by both external and internal forcings. These subsystems deserve a brief description. The atmosphere is a gaseous mixture that envelops almost uniformly the surface of the Earth. The atmosphere can be divided into several layers, which differ in composition, temperature and stability. However, 99% of the mass of the atmosphere lies below approximately 30 km which includes the troposphere and mid-stratosphere. The concentration of nitrogen, oxygen and other inert gases are remarkably uniform up to the mesopause. There are also variable components, such as water vapour, carbon dioxide, methane and other trace gases found predominantly in the troposphere. The composition of the atmosphere is further complicated by the presence of clouds (liquid water and ice crystals), dust, aerosols and volcanic ashes. The concentrations of these aerosols also vary in time and space. The atmosphere's response time to perturbations is the shortest of any of the other climate subsystems. That is, the time it takes for the system to re-equilibrate itself to a new state or previous state after a perturbation is less than the other climate system components. The response time of the atmosphere is on the order of days to weeks and is due to its relatively large compressibility and low specific heat and density. The atmosphere is set in motion primarily by differential heating by the Sun. This is a complex process because the motions of the atmosphere are influenced by factors such as the rotation of the earth, thermodynamics and topographical conditions. However, the atmosphere is organized on a global scale.

The hydrosphere includes all water in the liquid phase dispersed on the earth. It includes the oceans, seas, lakes, rivers and ground water sources. Two-thirds of the surface of the Earth is covered by water so the oceans absorb most of the incoming solar radiation. Because of their sheer mass and energy storage capability, the oceans act as a regulator of near surface temperatures. The circulation of the oceans is more sluggish than in the atmosphere, and the response or relaxation time for the ocean varies vertically. In the upper mixed layer (100 m) it is from week to months with seasonal variations in the thermocline and centuries or millennia in the deep ocean (thermohaline overturning due to density variations associated with changes in temperature and salinity). The basic role of oceans in the climate system is the redistribution of heat from the low latitude regions toward colder mid-latitude and polar regions via surface ocean currents associated with wind patterns. The atmosphere and oceans are strongly coupled. Air-sea interactions occur on many time and space scales through the exchange of matter, energy and momentum. The primary exchange of matter between the oceans and the atmosphere is in the form of water vapour through evaporation processes. Once in the atmosphere it is temporarily stored in the form of latent heat, which supplies part of the energy for the hydrological cycle leading to condensation, precipitation, and runoff, which are redistributed to the oceans via runoff.

Water in the solid phase constitutes the cryosphere. Continental snow and ice (ice sheets and alpine glaciers), permafrost and sea ice represent the cryosphere. In addition, it represents the largest reservoir of fresh water on the earth. Its importance to the climate system results mainly from its high reflectivity or albedo for solar radiation and its low thermal conductivity. The cryosphere acts as an insulator for the underlying land and water surface, preventing it from losing heat to the atmosphere and thereby having a cooling effect on regional and local atmospheric temperatures. In addition to seasonal variations of the cryosphere, large changes may occur over much longer time scales. For example, ice sheets do not vary rapidly but play a major role on climatic change over tens of thousands of years, such as glacial-interglacial cycles. Because of their size, ice sheets develop their own internal dynamics. Occasionally, they may collapse, resulting in massive iceberg surges. Alpine glaciers move slowly down valleys under the influence of gravity and may reach a critical point where they melt and disappear over timescales of centuries, giving rise to fresh meltwater pulses, a potential important factor affecting coastal ocean salinity conditions.

The lithosphere includes the continents and the ocean floors. The continents interact with the atmosphere by mechanically influencing air motions while the ocean floor partly determines deep ocean circulation. The lithosphere has the longest response time of all components of the climatic system. There are strong interactions between the lithosphere and the atmosphere through the transfer of mass and sensible heat. The transfer of mass occurs mainly in the form of water vapour, rain and snow and to a lesser extent in the form of dust and aerosols, such as from volcano eruptions. The soil moisture of the active layer in the continental lithosphere has an influence on the local energy balance affecting the rate of evaporation, the surface albedo, and the thermal conductivity of the soil. However, the lithosphere can be regarded as a semi-permanent feature of the climatic system because of its markedly slow response time.

Finally, the biosphere comprises the vegetation, terrestrial fauna and ocean flora and fauna. Vegetation alters the underlying surface roughness, albedo, evaporation, runoff and field capacity of the soil. In addition, the biosphere influences the carbon dioxide balance in the atmosphere and oceans through photosynthesis and respiration. Overall, the biosphere is sensitive to climate changes and paleoclimate reconstructions have been successfully accomplished using terrestrial and marine fossils. The response time of the biosphere varies greatly, from decades to millennia. The biosphere also includes humans who influence the climate system through agriculture, land use and pollution.

The estimated time and space scales of the climate system vary widely from one subsystem to another. The atmospheric boundary layer (interface between the surface and atmosphere) varies from minutes to hours; the free atmosphere from weeks to possibly months; the upper mixed layer of the ocean from weeks to years; the deep ocean waters from decades to millennia; sea ice from weeks to decades; inland waters and vegetation from months to centuries; glaciers on the order of centuries; ice sheets on the order of millennia and beyond; and tectonic phenomena on the order of tens of millions of years.

Due to the complexity of the internal climatic system, it is convenient to consider a hierarchy of internal systems. Simply put, for certain time scales, one or more subsystems must be included in the internal system while the others are considered external forcings modulating the behaviour of the internal system. Thus, the whole climate system must be regarded as continuously evolving parts of a system with some leading and others lagging in time.

The dynamics of the climate system as a whole results from two main external forcings that govern its global behaviour. These are solar radiation and gravity, although solar radiation is the primary external forcing since it provides almost all the energy that drives the climate system. However, even if a ordered external forcing is present, the internal system (ie. climate) is always subject to chaotic or stochastic fluctuations in time and space because of the nonlinearities embedded within the system as discussed above. It is possible to consider a large ensemble of climate states corresponding to the same external forcing. A climate change will then be defined as a departure of a particular climate state for a given time interval from a group of equivalent states (Peixoto and Oort, 1992).

To pursue this idea, let us start with a given initial state: The solutions of the equations that govern the dynamics of a nonlinear system result, for an infinite interval of time, in a set of infinitely long-term statistics (Peixoto and Oort, 1992). If climate is considered from a mathematical viewpoint, it is possible that a particular set of initial conditions may not give rise to a unique climatic state. Two or more distinct set of statistics may result from a single set of initial conditions and this is termed an *intransitive* system (Lorenz, 1976). If, on the other hand, there is only one unique climatic state corresponding to different sets of boundary conditions, that system is said to be *transitive* (Peixoto and Oort, 1992). If there are different sets of statistics that a transitive system may assume in its evolution from different initial boundary conditions through a long, but finite, period of time, the system is called *almost intransitive* (Lorenz, 1976; Peixoto and Oort, 1992). This suggests that climate may be essentially stable until some external factor causes a perturbation in the system. This perturbation may only be a short-term phenomenon, after which the system recovers. However, the resulting climate state may not be exactly the same, even though the boundary conditions are essentially identical to those before the perturbation. This may be due to internal feedbacks or instabilities involving the different components of the climate system. For example, the glacial and interglacial periods may be manifestations of an almost intransitive system (Peixoto and Oort, 1992).

In conclusion, climate variability results from complex interactions of forced and free variations because the climate system is a dissipative, highly non-linear system with many sources of instabilities. This makes it difficult to identify and to obtain a straightforward interpretation of cause and effect.

POTENTIAL MECHANISMS OF CLIMATE CHANGE

Understanding the causes of climate change depends on the scale of the climate variation. The so-called Milankovitch variations (changes in the seasonal and latitudinal distribution of radiation at the top of the atmosphere) can explain reasonably well climate changes over tens of thousands of years, such as glacial-interglacial oscillations. However, these orbital variations cannot explain sudden and abrupt climates change over decadal to millennial scales. It appears likely that the causes of climate change over decade to millennia scales may be linked to internal mechanisms that act as amplifiers of small initial change in solar output variability into much larger shifts in temperature and aridity (Bond *et al.*, 2001; Shindell *et al.*, 1999; Lean *et al.*, 1992; Haigh; 1996; Crowley, 2000; Rind and Overpeck, 1993).

THE OCEAN THERMOHALINE CIRCULATION (THC).

The apparent coherence between records from both hemispheres and from different environments requires a mechanism that could cause changes in the global circulation at time-scales of century to millennial-scales. The most accepted mechanism came from changes in the global ocean "conveyer belt" of the North Atlantic thermohaline circulation and North Atlantic Deep Water (NADW) formation (Broecker *et al.*, 1989, 1995, 1997).

The global ocean "conveyer belt" involves the transport of warm surface waters to northern latitudes and the return of deep ocean waters southward. The properties of the thermohaline circulation are controlled by the properties of seawater, which is a function of both its temperature and salinity (Broecker *et al.*,, 1989). The global conveyer belt involves an interaction between both surface currents and deep-water formation. Warm surface waters from the equatorial Pacific travel through the Indian ocean and around South Africa into the Southern Atlantic ocean and across to the Gulf of Mexico during which time the salinity increases because of evaporative losses (Broecker, 1997; Lehman, 1993). Eventually the waters flow into the North Atlantic via the Gulf Stream and North Atlantic Drift. Upon reaching these regions, the surface water cools, and in combination with high salinities becomes dense enough to sink into the deep ocean (NADW formation) (Street-Perrott and Perrott, 1990; Lehman, 1993). The "pull" exerted by this deep-water formation is thought to help maintain the strength of the warm Gulf Stream, ensuring a current of warm tropical water into the North Atlantic. Interactions with the atmosphere send mild air masses across the European continent (Rahmstorf *et al.*, 1996). The NADW then flows back to the margins of Antarctica where it is distributed to the Pacific and Indian oceans, upwells and returns via surface ocean currents.

Shutting off the NADW could induce a colder climate in the Northern Hemisphere as well as upset the global conveyer belt. If the sinking process in the North Atlantic were to diminish or cease altogether, the weakening of the Gulf Stream would mean colder winters in Europe (Broecker, 1995). However, the presence of the Gulf Stream does not produce markedly warmer summers in Europe. In fact it tends to produce milder summers, so shutting off the Gulf Stream does not in itself explain why summers also became colder during sudden climate shifts. We must look at other interactions of the climate systems under the diminished or cessation of NADW formation.

One possibility is that sea ice forms more readily in colder winter waters due to the shutoff of the Gulf Stream (Lohmann and Gerdes, 1998). A lid of sea ice over the North Atlantic would prevail for a longer part of the year and reflect incoming solar radiation, leading to cooler summers over much of Europe, as well as to colder winters (Overpeck *et al.*, 1997). With cooler summers, the snow would last longer into the spring, further cooling the climate by reflecting back solar radiation. This would result in a colder and drier European climate. The air that reached Europe would carry less moisture, since it would come from a cold sea-ice surface rather than from warm Gulf Stream waters. These colder climates could be sustained for longer periods of time because colder summers would diminish the melting process and allow snow to build up, thus further cooling the climate.

However, this poses the question of how NADW can be shut down. The trigger for a sudden shutdown of NADW formation must be found in a decrease in density of surface waters in the areas of sinking in the North Atlantic Ocean, to critical values that would not allow the sinking. Such a decrease in density should result from changes in salinity (addition of freshwater from continental runoff, precipitation and meltwater) and/or increased temperatures (Rahmstorf *et al.*, 1996). Ocean and climate modelling studies suggest that small increases in freshwater fluxes to the North Atlantic could cause NADW formation in the North Atlantic to diminish or cease (Manabe and Stouffer, 1995, 1997; Broecker, 1998; Rahmstorf, 1994; Weaver, 1994; Jones, 1991; Sakai and Peltier, 1996).

During glacial periods, meltwater fluxes could have triggered a shutoff or decrease in NADW formation leading to abrupt and cold events. Among the possible source of meltwater flux into the North Atlantic are Heinrich events (massive surge of icebergs into the North Atlantic) that could have been sufficient to cause a cessation of NADW formation. The massive input of freshwater input would dilute the dense and relatively saline Gulf Stream and form a temporary "lid" that stops the sinking. This allows more extensive sea-ice to develop across the North Atlantic, thus preventing the ocean "conveyor belt" from starting up again. However, other more subdued events of meltwater fluxes may have only diminished NADW formation and that these reductions in NADW could also provide a mechanism for Dansgaard-Oeschger events (Sakai and Peltier, 1996). This scenario could also explain how some of these events have a global extent (cessation of NADW), whereas others have only regional impact (diminished NADW).

Other possible sources of freshwater fluxes into the North Atlantic include the sudden evacuation of very large ice-dammed lakes and diversion of meltwater from the North American Laurentide ice sheet through the Gulf of St-Lawrence, as seems to have occurred during the Younger Dryas cold reversal (Berger and Jansen, 1995).

The sudden shutoff could presumably occur in the opposite direction and re-establish NADW formation in the North Atlantic. For example, if warmer conditions caused sea ice to melt back, the NADW could just as rapidly start up again. It has been suggested that following an initial cooling caused by a freshwater flux to the North Atlantic; the evaporation in the tropical Atlantic could result in an increase in the salinity of Atlantic ocean surface water. Unable to sink into the North Atlantic because of the sea ice lid the salinity would increase to a point where sinking began to occur anyway at the edges of the lid of sea ice resulting in a more southern circulation (Broecker, 1995). This would result in a renewed northward flux of warm water and air, thus leading to warmer climates in the mid-latitudes and reestablishment of the conveyer belt.

As for interhemispheric synchrony, a possible scenario is presented. Antarctica receives most of its heat from Gulf Stream water that sinks and flows southward down the western side of the deep Atlantic basin and then partially resurfaces near the margin of Antarctica. This water is much warmer than the continent and helps melt back some of the sea ice that forms around Antarctica. If the NADW were switched off, that would lead to a greater sea ice extent around Antarctica, leading to more loss of energy through a greater surface albedo, further cooling the region. However, a problem arises here because of the time it take for NADW to reach Antarctica, so it could only produce an effect a few centuries later. Other climate feedback mechanisms are required to produce closely synchronized changes.

Rapid climate transitions during interglacial periods however show no evidence of increased supply of melting icebergs (McManus *et al.*, 1994). However, millennial-scale climate changes may be linked to the North Atlantic's thermohaline circulation. One or more of the following factors may likely be the cause: (1) enhanced precipitation over the North Atlantic due to the greater amount of solar energy reaching the northern latitudes; (2) increased continental runoff and sea-ice retreat; and (3) other feedback mechanisms as will be discussed below. This dissertation has provided evidence that variations in solar output can cause large-scale ocean-atmospheric reorganization. However, we are still unable to isolate the processes acting as amplifiers, although NADW formation may be a key process (Bond *et al.*, 2001; Broecker, 1997)

Atmospheric composition

The atmosphere's chemical composition is an important determinant of the troposphere's thermal budget. When atmospheric composition is altered, changes occur in the thermal budget. However, feedback processes within the climate system complicate the analysis of the extent of climate change that follows a change in atmospheric composition. Here we look at the effects on climate of the main GHGs such as CO_2, CH_4, H_2O and atmospheric aerosols.

Water vapour is the most important greenhouse gas in terms of concentration and variability. Therefore, water vapour can play a major role as an amplifier or stabilizing feedback mechanism within the climate system. In addition, it plays a major role in the global hydrological cycle. It has been established that as the mean global temperature rises, evapotranspiration over land and the oceans increases, especially in equatorial regions. This results in a net increase of water vapour flux into the atmosphere where it is distributed to higher latitudes through the general circulation. In turn, a higher concentration of water vapour in the atmosphere can amplify or moderate an initial warming by cloud feedback processes. Broecker (1997) suggested that water vapour might act as an amplifier of rapid climate changes. For example, water vapour could have amplified an initial THC perturbation in the North Atlantic basin to be a globally synchronous event with very little lags (Broecker, 1997).

Another greenhouse gas that has been intensely studied is carbon dioxide (CO_2). Analysis of bubbles in ice cores shows that at the peak of glacial phases, the atmospheric CO_2 was about 30% lower than during interglacials (Jouzel *et al.*, 1996). These changes in CO_2 concentrations are thought to result from marine and terrestrial biotic activity. However, the marine and terrestrial ecosystems have opposite effects on atmospheric carbon concentrations. In cold phases, terrestrial ecosystems tend to decrease activity and release carbon to the atmosphere acting as a source while an increase in biotic activity and ocean carbon storage act as a sink. The lower CO_2 concentrations resulting from greater ocean carbon storage would further cool the atmosphere in glacial periods, and allow more snow and ice to accumulate on land. Relatively rapid changes in climate, occurring over a few thousand years, could have resulted from changes in atmospheric CO_2 concentrations (Broecker, 1997).

The problem with invoking CO_2 as a cause in rapid climate changes is that their concentrations seem to have varied too slowly (Jouzel *et al.*, 1996) However, an increase in CO_2 could be very significant, as persistent increases in global mean temperatures may ultimately influence large-scale processes in the coupled ocean-atmosphere system (Manabe and Stouffer, 1993). For example, the thermohaline circulation may have undergone abrupt changes in response to rising temperatures due to increased CO_2 levels, leading to ice-sheet and sea-ice melt, resulting in fresh meltwater fluxes into the North Atlantic. This would be a prerequisite in cessation of NADW (Broecker, 1997; Sakai and Peltier, 1996). The actual importance of CO_2 in terms of the climate system is still under debate because of the complex feedbacks processes involved.

Another less important greenhouse gas is methane (CH_4), whose behaviour closely parallels that of atmospheric CO_2. Methane is primarily produced in swamps and bogs where anaerobic conditions exist. Other sources of methane include animal activity, methane calthrates and permafrost. Methane levels seem to have increased rapidly in concentration in association with climate change, reaching Holocene levels in around 150 years or less during the global warming at the end of the Younger Dryas, around 11,500 years ago (Taylor *et al.*, 1997). However, such a sudden rise in atmospheric methane was probably not important in affecting climate because of its overall low concentration in the atmosphere. It has been suggested that another mechanism, involving massive amounts of methane released from the ocean floors or permafrost melt, in conjunction with an increase in CO_2, could have resulted in rapid warming phases. However, because of the relatively short residence time of methane (around 10 years) in the atmosphere and resolution of the trapped air bubbles in ice cores (10^3 yr), it is difficult to infer such methane pulses. Furthermore, Taylor *et al.* (1997) showed that sudden bursts of methane were not present in their analysis of ice cores from areas where the ice-sheet built up particularly rapidly giving a detailed time resolution. What is important here is the close parallel connection between CO_2 and CH_4 behaviours during cold and warm phases.

Aerosols include dust, sea salt and other particulates present in the atmosphere. Their presence in the atmosphere results from natural, and more recently, anthropogenic sources. For instance, they are blown off the surface of deserts or dry regions or injected into the atmosphere from volcano eruptions. Aerosols affect climate in two ways, directly through scattering and absorbing of solar radiation and indirectly through modification of cloud properties (i.e. some aerosols act as condensation nuclei for cloud droplets). Ice cores from Greenland (Mayewski *et al.*, 1997; Taylor *et al.*, 1997) and Antarctica (Jouzel *et al.*, 1996) show greater concentrations in dust particles during colder phases. This suggests that there was more dust in the atmosphere during colder phases than during warmer phases. Under drier conditions, the area of desert increases, as does the mean wind speed, thereby sending more dust into the atmosphere where it may amplify the cooling by blocking sunlight. The inverse has a significant climate effect as well, where less dust could serve to warm the atmosphere and these conditions would then further decrease the amount of dust blown to the atmosphere.

SURFACE ALBEDO

Another feedback mechanism within the climate system is the surface albedo; the ability of a surface to reflect the Sun's radiation. Presently, about a third of the heat received from the Sun is reflected back to space, and changes in this proportion have the potential to strongly influence global climates. An enhanced greenhouse effect will tend to warm the surface, and because water vapour is an important greenhouse gas, it will amplify the warming and thus evaporation will increase. However, increases in the number of clouds resulting from higher evaporation rates may act either to amplify the warming through the greenhouse effect or reduce it by an increase in planetary albedo. The effect that predominates depends on the height and type of clouds and varies greatly geographically and seasonally.

Ice and snow albedo feedbacks may lead to amplification or stabilization of temperature changes in high latitudes. For example, sea ice, which has a high albedo, will have a major effect in cooling regional and global climates. Sea ice is very sensitive to temperature changes and a small warming may lead to a decrease in sea ice extent. Increased open water, which has a low albedo, will absorb more heat. The increased warming and evaporation will lead to the formation of low clouds, offsetting the change in surface albedo. This is also true for snow albedo processes, where snow will tend to cool local to regional climates and darker land surfaces will have an effect of amplifying a warming.

Changes in vegetation also affect surface albedo, evapotranspiration and roughness. Slow changes over centuries to millennia in surface albedo can thus be an important amplifier or trigger for a major change in global temperature. These slow changes may be caused by variations in the earth's orbit affecting summer incoming radiation in high northern latitudes, gradual changes in atmospheric composition, changes in vegetation cover, or a combination of all the above.

VARIATIONS IN SOLAR RADIATION INPUT

A major background factor causing climate changes on tens of thousands of years seem to be the "Milankovitch orbital rhythms" in seasonal and latitudinal radiation distribution (Imbrie *et al.*, 1993). Although the insolation values vary gradually over thousands of years, the climate response may be abrupt if these cause a critical threshold to be surpassed, which causes other internal climate mechanisms to amplify.

There are three Milankovitch cycles. The "eccentricity" which is the variation in Earth's orbit around the Sun (from nearly circular to elliptical) occurs at periodicities of about 100 ka. Obliquity is the variation in the degree of tilt of earth's axis, which occurs at about a 42 ka cycle. The last cycle is "precession of equinoxes", which is the timing of the seasons relative to the earth's elliptical orbit (nearer and furthest from the Sun) and this cycle has periodicities of 19 and 23 ka. The latter two cycles alter the relative amount of solar radiation reaching the Northern and Southern Hemispheres during summer and winter. When summer insolation in the Northern Hemisphere is strong, rapid global transition from glacial to interglacial conditions tend to occur (Wright *et al.*, 1993). However, it is generally accepted that other internal climate feedback mechanisms directly or indirectly linked to the seasonal insolation changes must be involved in bringing a transition from a glacial phase into interglacial conditions.

The same analogy to the Milankovitch cycles can be used for decadal to millennial-scale climate variability where the Sun's radiative output would vary on different frequencies, therefore acting as a weak external forcing superimposed on Milankovitch orbital forcings. The sun-climate relationship has long been debated because the magnitude of change of solar output is weak. Model simulations suggest that less that 0.1 % in total irradiance change could presumably affect stratospheric ozone processes which in turn would, through downward-propagation, influence tropospheric circulation and thus surface climates (Haigh, 1996, Shindell *et al.*, 1999). A weak solar output decrease on the order of 0.24% on multi-decadal scales has been shown to affect surface climates using sun-like stars as proxies (Lean *et al*, 1992). It is now recognised that solar output variability is much greater in the UV spectrum than the visible (i.e. SOHO experiment on solar wind) (IPCC, 2001). These small variations would in turn affect stratospheric ozone production, therefore warming or cooling the stratosphere, which would influence upper-tropospheric wave patterns and amplify, through internal feedback mechanisms, the impact on surface climates.

In summary, the terrestrial climate has varied significantly and continuously on time scales ranging from years to the age of the Earth. The variability of climate results primarily from forced variations, which are the response of the climate system to external forcings and of free variations due to internal instabilities and feedbacks between the various components of the climate system. The external forcings of climate variations include variability in solar irradiance and orbital parameters (eccentricity, axial precession and obliquity of the ecliptic). The climate response to external and internal perturbations is associated with many positive and negative feedbacks mechanisms such as ocean dynamics, atmospheric composition, surface albedo and other strong interactions between the climate system components. These processes can lead to instabilities or oscillations of the system that can operate independently or introduce strong modifications to external forcings (Peixoto and Oort, 1992), and these cause abrupt changes in the climate on many timescales.

REFERENCES

Berger, W.H. and Jansen, E. 1995. Younger Dryas episode: ice collapse and super-fjord heat pump. In Troelstra, S.R., van Hinte, J.E. and Ganssen, G.M., editors, The Younger Dryas. Koninklijke Nederlandse Akademie van Wetenschappen, Afdeling Natuurkunde, Eerste Reeks 44: 61-105.

Bond, G., Kromer, B., Beer, J., Muscheler, R., Evans, N.E., Showers, W., Hoffman, S., Lotti-Bond, R., Hajdas, I. and Bonani, G. 2001. Persistent solar influence on North Atlantic climate during the Holocene: Science 294: 2130-2136.

Broecker, W.S. 1995. The glacial world according to Wally. New York: Eldigio Press.

Broecker, W.S. 1997. Thermohaline circulation, the Achilles heel of our climate system: will manmade CO2 uset the current balance? Science 278: 1582-88.

Broecker, W.S. 1998. Paleocean circulation during the last deglaciation: a bipolar seesaw? Paleoceanography 13: 119-21.

Broecker, W.S., and Denton, G.H. 1989. The role of ocean-atmosphere reorganizations in glacial cycles. Geochimica et Cosmochimica Acta 53: 2465-2501.

Crowley, 2000. Causes of climate change over the past 1,000 years, Science 289: 270-277.

Haigh, J.D. 1996. The impact of solar variability on climate. Science, 272: 981-984.

Imbrie, J., Berger, A., Boyle, E.A., Clemens, S.C., Duffy, A., Howard, W.A., Kukla, G., Kutzbach, J., Martinson, D.G., McIntyre, A., Mix, A.C., Molfino, B., Morley, J.J., Peterson, L.C., Pisias, N.G., Prell, W.G., Raymo, M.E., Shackleton, N.J. and Toggweiler, J.R. 1993. On the structure and origin of major glaciation cycles. 2. The 100 000 year cycle. Paleoceanography 8: 699-735.

IPCC Climate Change: 1995. The Science of Climate Change. Houghton, J.T, Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A., and Maskell, K. (Eds.). Cambridge University Press. 572 pp.

IPCC Climate Change: 2001. The Scientific Base. Houghton, J.T, Ding, Y.,Griggs, D.J., Noguer, M., van Der Linden, P.J., Dai, X., Maskell, K., and Johnson, C.A. (Eds.). Cambridge University Press. 572 pp.

Jones, G.A, 1991. A stop-start ocean conveyor. Nature 349, 364-365.

Jones, P.D., New, M., Parker, D.E., Martin, S. and Rigor, I.G. 1999. Surface air temperature and its changes over the past 150 years. Reviews of Geophysics 37: 173-199.

Jouzel, J., Waelbroeck, C., Malaize, B., Bender, M., Petit, J.R., Stievenard, M., Barkov, N.I., Barnola, J.M., King, T., Kotlakov, V.M., Lipenkov, V., Lorius, C., Raynaud, D., Ritz, C. and Sowers, T. 1996. Climatic interpretation of the recently extended Vostok ice records. Climate Dynamics 12: 513-21.

Lean, J., Skumanich, A., and White, O. 1992. Estimating the sun's radiative output during the Maunder minimum. Geophysical Research Letters 19(15): 1591-1594.

Lehman, S. 1993. Ice sheets, wayward winds and sea change. Nature 365: 108-110.

Lohmann, G. and Gerdes, R. 1998. Sea ice effects in the sensitivity of the thermohaline circulation. American Meteorological Society 11: 2789-2803.

Lorenz, E. N. 1976. Nondeterminstic theories of climatic change. Quaternary Research 6: 495-506.

Manabe, S. And Stouffer, R.J. 1993. Century-scale effects of increased atmospheric CO2 on the ocean-atmosphere system. Nature 364: 215-217.

Manabe, S., and Stouffer, R.J. 1997. Coupled ocean-atmosphere model response to freshwater input: Comparison to Younger Dryas event. Paleoceanography 12: 321-336.

Manabe, S. And Stouffer, R.J. 1995. Simulation of abrupt climate change induced by freshwater input to the North Atlantic. Nature 378: 165-167.

Mann, M.E., Bradley, R.S. and Hughes, M.K. 1999. Northern Hemisphere temperatures duing the past millennium: Inferences, uncertainties, and limitations. Geophysical Research Letters 26: 759-762.

Mayewski, P.A., Meeker, L.D., Twickler, M.S., Whitlow, S., Yang, Q., Lyons, W.B. and Prentice, M. 1997. Major features and forcing of high-latitude Northern Hemisphere atmospheric circulation using a 110 000 year-long glaciochemical series. Journal of Geophysical Research 102: 26 345-65.

McManus, J.F., Bond, G.C., Broecker, W.S., Johnsen, S., Labeyrie, L. and Higgins, S. 1994. High resolution climate records from the North Atlantic during the last interglacial. Nature 371: 326-29.

Overpeck, J., Hughen, K., Hardy, D., Bradley, R., Case, R., Douglas, M., Finney, B., Gajewski, K., Jacoby, G., Jennings, A., Lamoureux, S., Lasca, A., MacDonald, G., Moore J., Retelle, M., Smith, S., Wolfe, A. and Zielinski, G. 1997. Arctic environmental change of the last four centuries. Science 278: 1251-1256.

Peixoto, J. P., and Oort, A.H. 1992. Physics of Climate. American Institute of Physics. New York 520 pp.

Rahmstorf, S. 1994. Rapid climate transitions in a coupled ocean-atmosphere model. Nature 372: 82-85.

Rahmstorf, S., Marotzke, J. and Willebrand, J. 1996. Stability of the thermohaline circulation. In Kraus, W., editor, The warm watersphere of the North Atlantic Ocean, Berlin: Gebruder Borntrager: 129-57.

Rind, D. and Overpeck, J. 1993. Hypothesized causes of decade-to century-scale climate variability: climate model results. Quaternary Science Reviews 12: 357-74.

Rind, D., 1999. Complexity and Climate. Science 284: 105-107.

Sakai, K. And Peltier, W.R. 1996. Dansgaard-Oeschger ocsillations in a coupled atmosphere-ocean climate model. Americam Meteorological Society 10: 949-970.

Saltzman, B. 1985. Paleoclimate modelling. In: Palroclimate analysis and modelling. A.D. Hetch (Ed.) John Wiley and Sons, New York. 445 pp.

Shindell, D., D. Rind, N. Balachandran, J. Lean, and P. Lonergan, 1999. Solar cycle variability, ozone, and climate. Science 284: 305-308.

Street-Perrott, F.A., Perrott R.A 1990. Abrupt climate fluctuations in the tropics: the influence of Atlantic ocean circulation. Nature 343: 607-612.

Taylor, K.C., Alley, R.B., Doyle, G.A., Grootes, P.M., Mayewski, P.A., Lamorey, G.W., White, J.W.C. and Barlow, L.K. 1993. The "flickering switch" of late Pleistocene climate change. Nature 361: 432-36.

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: 825-27.

Weaver, A. J., Hughes, T.M.C. 1994. Rapid interglacial fluctuations driven by North Atlantic ocean circulation. Nature 367: 447-450.

Wendland, W.M., and Bryson, R.A., 1974. Dating climatic episodes of the Holocene: Quaternary Research 4: 9-24.

Wright, H.E. Jr., Kutzbach, J.E., Webb, T.III., Ruddiman, W.F., Street-Perrott, F.A., Bartlein, P.J. 1993. Global climates since the last glacial maximum.. The University of Minnesota Press. 569 pp.