

**No-analog ecological communities since the last glacial maximum
inferred from fossil pollen records for North America**

by

Zihaohan Sang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Forest Biology and Management

Department of Renewable Resources

University of Alberta

© Zihaohan Sang, 2019

Abstract

Given that efforts to reduce greenhouse gas emissions appear to fall short, we need to develop adaptation strategies to mitigate anthropogenic climate change impacts, which may include predicting and managing no-analog ecological communities, for example by selecting suitable species assemblages for ecological restoration projects or in reforestation. To develop analytical tools for the future, this study looks at the past to test if no-analog communities can be detected in the North American fossil pollen record, if they were associated with no-analog climate conditions, and if they can be predicted with species distribution models. The working hypothesis is that no-analog communities that are not associated with no-analog climates, but result from differential migration rates will be the most difficult to predict using species distribution modelling (SDM).

This study contributes a systematic analysis of 51,427 dated pollen community records from 1,384 sites from the Neotoma Paleoecology Database with climate hindcasts for 21000, 16000, 14000, 11000, 6000 years before present from the CCM1 general circulation model. The analysis relies on a multivariate regression tree approach that delineates pollen communities but constrains the clustering of similar pollen records by using climate variables as partitioning criteria. Modern pollen records are also used to train a species distribution model and create hindcasts up the last glacial maximum to test if past no-analog communities can accurately be predicted.

The analysis revealed three pollen communities that were widespread during the last glaciations or during deglaciation but that have no equivalence in today's pollen record. The most distinct no-analog community was a birch parkland ecosystem with

exceptionally high frequencies of *Betula*, *Poaceae*, *Cyperaceae* and forb pollen. This ecosystem also occupied no-analog climate space that was similar to today's arctic tundra with high seasonality and relatively dry conditions, but had a stronger maritime influence leading to higher temperatures during the warmest and wettest quarter of the year.

The second, widespread no-analog community was a previously described spruce woodland with spruce having a dominant frequency, but also included a large diversity of broadleaf species that today do not co-occur in spruce-dominated ecosystems. This community occurred between 14000 and 11000 years before present south of the Great Lakes region. A third potential no-analog community was a montane steppe ecosystem in the Rocky Mountains with lower frequencies of tree pollen and higher frequencies of pollen from grasses and forbs than modern equivalents. This ecosystem was associated with drier conditions than today but overlapped with the current mountain ecosystem both in pollen frequency and climate space.

While species distribution models could hindcast birch presence in novel climate space, they failed to predict the uniquely high frequencies of birch in the observed birch-parkland community. Similarly, the uniquely high frequencies of spruce in spruce woodlands between 14000 to 11000 years before present was not correctly predicted. In both cases, violation of assumptions of species distribution models, in the first case climate equivalence, in the second case lack of consideration of habitat factors other than climate appear to prevent accurate predictions of community composition in the past, and by inference also for the future.

Preface

This thesis is being prepared for submission as a journal article. The contributors are Zihaohan Sang and Andreas Hamann. The study was conceived and designed by myself, with input from AH. I conducted the data collection and initial analysis. I analyzed the data and wrote the first draft of the thesis. AH provided suggestions for additional analytical approaches, contributed to interpreting the results and edited the manuscript.

This thesis was submitted in partial fulfillment of the requirements for a dual-degree for Master of Science, University of Alberta, Edmonton, Alberta, and Master of Forest Sciences, University of Freiburg, Germany.

Acknowledgements

The process developing this thesis has been one of the most challenging yet rewarding processes that I have ever embarked on, and would not have been possible without the guidance of my supervisors Dr. Andreas Hamann. I appreciate the countless hours he spent on sharing me ideas and editing this thesis. He believes in my ability even when this project went wrong and I started to question everything in my life. Without his help and support, I would not able to finish this thesis or the MSc degree.

I am grateful to have Dr. Scott Nielsen and Dr. Rong-cai Yang as my committee members and thank them for accepting the task at (almost) the very last minute to read my thesis and serve on my examining committee.

I would also want to thank my parents who always trying their best to provide me secure financial conditions for my study. I am very lucky to have their support and understanding for these years.

To my lab-mates at the Spatial Information System (SIS) lab, thank you for always being willing to help with any problems, and for the company that made the research environment supportive and encouraging.

Table of Contents

1. Introduction	1
2. Literature review	5
2.1 No-analog paleoecological communities	5
2.2 Paleoclimate reconstructions and no-analog climates.....	7
2.3 Niche theory and model limitations	9
3. Materials and Methods	12
3.1 Pollen data	12
3.2 Climate data.....	13
3.3 Grouping of pollen data into climate periods.....	14
3.4 Constrained clustering of pollen communities.....	15
3.5 Climate space and no analog climate	17
3.6 Species distribution modeling	17
4. Results	18
4.1 Classifying modern pollen data into bioclimatic zone	18
4.2 Appearance of new bioclimatic classes in the past	22
4.3 Changes to climate space since LGM	25
4.4. Can SDMs predict the pollen frequencies of novel communities?	28
5. Discussion	29
5.1 No analog communities and no analog climate	29
5.2 Clustering pollen communities.....	31
5.3 All no analog communities may be difficult to predict.....	33
6. Conclusions	34
7. References	36
8. Appendix	43

List of Tables

- Table 1.** Average pollen percentages for groups shown in Fig. 3, which were determined with a climate-constrained cluster analysis of pollen sites. The averages were calculated across all six periods shown in Fig. 3.24
- Table 2.** Loadings of canonical discriminant analysis of groups from all six periods corresponding to Fig. 5. The first axis (Can1) explains 40.7% variance; the second axis (Can2) explains 28.9%; the third (Can3) and fourth (Can4) explains 16.1% and 7.4%, respectively.27
- Table 3.** Pollen percentages predicted with a species distribution model (Random Forest) for the same groups as shown in Table 1. The model was trained with pollen data from the modern period (0-4 ka), but the averages shown below were calculated across all six periods as in Table 1. Thus, only the no-analog communities (Birch Parkland, Spruce Woodland and Montane Steppe) are completely independent predictions that did not overlap with training data.....28

List of Figures

- Figure 1.** Reconstructed temperature and available pollen sites for six periods from 23,000 years ago to present. Hindcast temperature pattern from CCM1 model for North America are compared with global temperature reconstruction by ice core proxy (Petit *et al.* 1999) between 23,000 years before present (1950) to current. Six reconstructed climate points (black empty circles) from CCM1 model generally fit the moving average curve of the global temperature changes. Red dots represent the sum of available pollen sites for each of the six periods. The number of pollen sites for 200-year intervals are shown in histogram.15
- Figure 2.** (a) Regression tree analysis for modern pollen records, (b) geographic location of samples and clusters from the analysis, and (c) modern ecological zones for comparison (adapted from Oslon *et al.* 2001). Bioclimatic variables used by the regression tree include MAT (mean annual temperature), TWarmQ (mean temperature of the warmest quarter), TColdQ (mean temperature of the coldest quarter), MAP (mean annual precipitation), TWetQ (mean temperature of the wettest quarter), TD (continentality), TS (temperature seasonality), and CMI (climate moisture index). The bars of (a) represent square root transformed taxa percentages for better visualization.20
- Figure 3.** Climate-constrained clusters of pollen sites from a multivariate regression tree analysis for six periods (as shown in Fig.1). The analysis created 22 pollen groups that were summarized into 11 groups in this map for simplicity, but maintaining groupings for no-analog communities (Montane Steppe, Birch Parkland and Spruce woodland).23
- Figure 4.** Changes in frequency of pollen community groups from the last glacial maximum to present, as shown in Fig. 3. Frequency of each pollen group for each period is given as percent of all pollen sites available for each period.24
- Figure 5.** Canonical discriminant analysis of 11 bioclimatic zones with four climate dimensions explaining 93.1% of the total variance in climate variables. The top row is a plot of component 2 over component 1 and the bottom row represents component 4 over component 3 for three selected climate periods. Plots for all six climate periods are provided in Appendix 1.26

1. Introduction

One of the most striking examples of no-analog fossil assemblages comes from the Eocene epoch (hereafter Eocene), where CO₂ concentrations have been substantially higher than today, approaching a maximum of approximately 1,400 ppm during the early Eocene around 50 million years ago and gradually declining to around 600 ppm toward the end of the Eocene 34 million years ago (Anagnostou *et al.* 2016). During this time, fossil finds from the Axel Heiberg Island in Canadian High Arctic (80° N latitude), revealed redwood-dominated, high-biomass forests with fossilized stumps and stems up to 40 m tall and over 1 m diameter at breast height (Williams *et al.* 2003). Eocene temperatures in the high arctic have been estimated to be 30-50 °C warmer than today (Huber & Caballero 2011), resulting in mean annual temperatures of around 15 °C (Greenwood, Basinger & Smith 2010; Eberle & Greenwood 2012), compared to today's mean annual temperatures of about -20 °C at this location (Wang *et al.* 2016). Coastal forest tree species that are common today in the Pacific Northwest therefore found suitable habitat near the arctic circle at this time, but experiencing no-analog climate conditions with extreme seasonality (Eldrett *et al.* 2009; Schubert *et al.* 2012) due to dry, long arctic winters without sunlight for extended periods.

Anthropogenic climate change, even if not of the same magnitude as during the Eocene, may produce similar no-analog communities in the future. CO₂ concentrations toward the end of the century are expected to range between 450 and 1,000 ppm depending on the emission scenarios (Moss *et al.* 2008; Moss *et al.* 2010; Meinshausen *et al.* 2011; van Vuuren *et al.* 2011), i.e. representing only about half the CO₂ concentrations of the Eocene period. Temperatures projection under anthropogenic warming for the Axel Heiberg Island location, range from 10-20 °C warmer than today under those emission scenarios for the end of the century (Taylor,

Stouffer & Meehl 2012; Wang *et al.* 2016), bringing mean annual temperatures to -10 to 0 °C by the end of the century. Under the most pessimistic emission scenario (Representative Concentration Pathway 8.5), Eocene-like climate conditions due to anthropogenic warming are projected for 2150, i.e. just 130 years from now (Burke *et al.* 2018).

Given that efforts to reduce greenhouse gas emissions appear to fall short, we need to develop adaptation strategies to mitigate anthropogenic climate change impacts, which may include predicting and managing no-analog ecological communities, for example by selecting suitable species assemblages for ecological restoration projects or in reforestation (Pedlar *et al.* 2012; Williams & Dumroese 2013). In this study, I will look at the past to test if no-analog communities can be detected in the North American fossil pollen record, and test if those no-analog communities can be predicted with species distribution models, a tool that is widely used to predict potential climate change impacts on species distributions (e.g., Pearson & Dawson 2003; Hijmans & Graham 2006; Thuiller *et al.* 2008), and that may potentially be useful to predict no-analog species assemblages under future scenarios as well.

No-analog fossil assemblages have also been documented under much less pronounced climate change. For example, during the last 4 million years of the Quaternary, CO₂ concentrations were relatively stable and varied between approximately 250 and 450 ppm based on proxy reconstructions, such as isotopes in ocean floor sediments and phytoplankton shell (e.g., Pagani *et al.* 2010; Seki *et al.* 2010; Badger *et al.* 2013). Associated temperature changes ranged from approximately 8 °C colder to around 2 °C warmer in mean annual temperature than today. This has resulted in no-analog ecological communities documented across continents and taxonomic groups, including vascular plants (e.g., Jackson & Williams 2004; Williams & Jackson 2007), mammals (e.g., Graham *et al.* 1996; Graham

2005), beetles (e.g., Coope & Angus 1975), and mollusks (e.g., Roy, Jablonski & Valentine 1995; Preece & Bridgland 2012).

For North America, no-analog communities during the Quaternary have been inferred from macrofossils and fossil pollen records. Overpeck *et al.* (1985; 1992) compared fossil pollen spectra and those derived from modern communities by calculating a dissimilarity metric (squared-chord distance) based on pollen percentages in samples. Several no-analog vegetation types were suggested, especially a large spruce-dominated woodland region with high amount of sedge at the southern of Great Lakes about 12 ka (kilo annum, signifies ‘thousand calendar years before present’). Williams *et al.* (2001; 2004) identified unique vegetation in the eastern North America association centered the Great Lakes, a mixed parkland biome at 14 ka. These communities are characterized by co-occurrences of species with high abundance that are now nearly allopatric, e.g., *Picea* and *Fraxinus*. These vegetation assembles were common during the late-glacial period, but no longer exist today. A no-analog species community was also identified in Beringia based on pollen and macrofossil data (Edwards *et al.* 2005). During the early Holocene (13-10 ka), shrub tundra ecosystems responded to climate warming through a rapid shift from shrub tundra to deciduous forest or woodland, while normally associated conifer species were lacking.

These no-analog communities have been observed during periods of relatively rapid warming and deglaciation, and could simply arise from differential response of species migrating into newly available habitat by migration lag, physiological tolerances, and geographic barriers (e.g., Urban, Tewksbury & Sheldon 2012). Alternatively, these communities could also be the result of no-analog climate conditions at the time. In an analysis for eastern North America, Williams, Shuman and Webb (2001) found that no-analog communities in the Great Lakes forest region

were associated with no-analog climates: the spruce-ash-sedge assemblage occurred under seasonal temperature and moisture regimes that do not exist today. In contrast, Shuman *et al.* (2002) reported no-analog plant associations during 13 to 11 ka to be associated with the rate at which climate change occurred, i.e. vegetation may fail to keep pace with climate (Correa-Metrio *et al.* 2012).

Here, I contribute a systematic joint analysis of publicly available fossil pollen data for North America from the Neotoma Paleoecology Database (Williams *et al.* 2018) and climate hind casts for 21, 16, 14, 11, 6 ka from the CCM1 general circulation model (Kutzbach *et al.* 1998). No-analog communities that simply result from different migration rates would be expected during rapid climate warming at the onset of the Holocene, where temperatures increased by 5 °C within 5000 years from 16 to 11 ka (IPCC 2007). Alternatively, no-analog communities might also be associated with no-analog climate conditions either for extended periods during the last glaciation, or when the climate system underwent profound changes driven by the loss of the continental glaciers during deglaciation. Lastly, I test if any no analog communities that emerge from this analysis can be predicted with species distribution models. The expectation is that no-analog communities that result from differential migration rates in response to climate change may be most difficult to predict, since species distribution models do not account for species lagging behind their suitable habitat.

2. Literature review

2.1 *No-analog paleoecological communities*

Past ecological communities with species compositions unlike modern communities are known as no-analog communities. One of the reasons of no-analog communities in the past could be climate conditions that do not exist today. Vegetation associations are not a fixed units that move together in response to climate (Gleason 1926), rather each species responds differently and occupies habitat according to their unique fundamental niche (Hutchinson 1958; Web 1988; Huntley & Webb 1989; Graham & Grimm, 1990; Webb *et al.* 2004; Araujo & Guisan 2006). Individualistic behavior of plant species includes the timing and direction of changes in range and abundance (Williams *et al.* 2004), and therefore plant communities are not permanent under environmental fluctuations (West 1964). Given current limited environmental combinations, the environmental niche of any given species is usually not fully known. Species may expand into unrealized niche portions and unexpected geographical regions (e.g. as introduced species on a different continent) under novel climate combinations (Jackson & Overpeck 2000). Similarly, no-analog paleoecological communities emerged by individualistic species responses to different climate conditions than those observed today.

The first evidence for no-analog communities in North America were described for the eastern United States based on pollen records (Wright Jr, Winter & Patten 1963; Overpeck, Webb & Prentice 1985), and has been summarized by Overpeck, Webb and Webb (1992). Pollen samples for 18 ka suggest boreal forest, mixed forest, and deciduous forest were present in the eastern Appalachian Mountains from south to north. During deglaciation, around 12 ka, a large region south of the Great Lakes was

covered by no-analog spruce-dominant woodland vegetation, characterized by a high percentage of spruce and a variety of eastern hardwood species that today do not co-occur with spruce at high frequencies. At the same time, the pine-dominated forests were restricted to Florida. By 9 ka, prairie ecosystems, previously restricted to the southern US expanded to central North America. Tundra and taiga ecosystems followed the edge of the retreating glaciers. By then, the modern vegetation distribution was generally established.

Also working with Northeastern US forest communities, Williams, Shuman and Webb (2001) added a climatic analysis using paleoclimatic hindcasts of the CCM1 General Circulation Model, and identified that no-analog vegetation associations centered the Great Lakes at 14 ka corresponded to high climate dissimilarity. According to Williams, Shuman and Webb (2001), these communities were characterized by anomalously high abundances of *Betula*, and co-occurrences of taxa now nearly allopatric (e.g., *Picea*, *Fraxinus*), with very low abundances of now common taxa including *Pinus*. Modern associations such as *Fagus–Tsuga* and *Picea–Alnus–Betula* date to the 11 ka, whereas other associations common to the late-glacial period (e.g., *Picea–Cyperaceae–Fraxinus–Ostrya/Carpinus*) no longer exist.

In western North America Whitlock and Bartlein (1997) analyzed vegetation records for a montane site in the western Cascade mountains that dates back 125 ka. They describe transitions between climate conditions that support forests and drier conditions that resulted in steppe ecosystems. During deglaciation, around 11 ka, increased summer temperatures, decreased effective precipitation, caused intensified drought and associated vegetation changes. Novel biomes were also suggested in Beringia based on pollen and macrofossil data (Edwards *et al.* 2005). During the early Holocene (13-10 ka), shrub tundra ecosystems responded to climate warming through a shift from shrub tundra to deciduous forest or woodland. Early-Holocene

vegetation was structurally, and hence functionally, novel compared with current dominant vegetation types.

Looking back much further, using macrofossil rather than pollen records, notable no-analog communities are also documented for the Eocene, 65 million years ago in the High Arctic region, which currently is a polar desert characterized by extremely cold (minimum temperature could reach $-50\text{ }^{\circ}\text{C}$) and dry climate (annual precipitation normally lower than 100 mm, and mostly in form of snow) with up to six-month long, dark expanded winter time. In this region, Williams *et al.* (2003) found redwood-dominated forest during the Eocene based on fossilized treetops and trunks on Axel Heiberg Island. Results suggests a high-productivity forest grew north of the Arctic Circle similar in stature to the modern old-growth forest of the Pacific Northwest of the North America. According to the proxies of macroflora, mean annual precipitation was estimated to be around 1200 mm for the Axel Heiberg Island location (Greenwood, Basinger & Smith 2010) and mean annual temperature for at the Arctic Circle was approximately $13\text{-}15\text{ }^{\circ}\text{C}$ (Tedford & Harington 2003). Fossil from sediments provides evidence that the swamp forests in Canadian Arctic Archipelago not only included abundant conifers but also had pines and broadleaves roughly 40-million years ago (Greenwood & Basinger 1993). Based on the anatomical features of well-preserved fossil, new species of *Larix* (Lepage & Basinger 1991) and *Pinaceae* (LePage 2001) were identified existing during that time.

2.2 Paleoclimate reconstructions and no-analog climates

Past climate conditions can be modeled with General Circulation Models (GCMs) or more recent atmosphere-ocean coupled general circulation models (CGCMs or AOGCMs). While they are primarily meant for future climate projections, the models are also run for the past for the purpose of validation. AOGCMs are mathematical

thermodynamic models of atmosphere and ocean circulation to provide information about the development of global climate when parameters change, such as greenhouse gas concentrations, solar insolation, or in case of paleoclimatic reconstructions continental ice cover. GCMs usually include equations that describe the energy changes that occur when regions of different temperature, pressure, chemical composition, velocities, and their interactions. GCMs simulate atmospheric and ocean processes in three dimensions. The processes are modeled based on three-dimensional grid cells, extending into the atmosphere and ocean at varying resolution, and equations calculate energy and mass exchange among grid cells at hourly time steps, which are later summarized to provide average predictions for climate normal periods (McGuffie & Henderson-Sellers 2005).

Paleoclimatic projections by the NCAR Community Climate Model (CCM1) (Kutzbach *et al.* 1998) for 16 and 14 ka are characterized by highly seasonal temperatures with winter temperatures as much as 10 °C colder than today, but summer temperatures being comparatively warm. The combination of hot summers, cold winters, and low precipitation places the simulated late-glacial climates outside the modern range of climates in North America. The CCM1 simulations indicate that at the last glacial maximum, temperatures for North America were approximately 15 °C colder than today and most regions were also drier (Webb *et al.* 1998). The climate patterns of North America during this time were heavily influenced by the Laurentide ice sheet, which including most of the land area that is now Canada. The ice sheet influenced climate by affecting the albedo, the hydrological cycle, and the atmospheric-ocean circulation (Manabe & Broccoli 1985; Broccoli & Manabe 1987; Clark, Alley & Pollard 1999). The increased albedo of the ice sheets caused very low surface temperatures, which in turn generated a strong anticyclone over the Laurentide ice sheet by cold air sinking over the ice sheet and outward air streams being deflected by the earth's rotation.

The anticyclone brought cold, dense air to the North Atlantic and strong easterly winds along the southern margin of the ice sheet (e.g., Cohmap Members 1988; Bromwich *et al.* 2005). The easterly winds split the westerly jet stream into two distinct branches straddling the Laurentide ice sheet along the northern edge, and the other branch being deflected to the southern part of the North American continent according to Manabe and Broccoli (1985). Along the western coast of the United States and southern Canada, this led to a 30% decrease in precipitation due to the redirection of the mid-latitude jet stream. However, the southern branch of the split jet stream enhanced precipitation in the southern part of the North American continent (e.g., Cohmap Members 1988; Shinn & Barron 1989; Kutzbach *et al.* 1998; Clark, Alley & Pollard 1999).

2.3 Niche theory and model limitations

Species distributions models (SDMs) are used in conservation ecology to infer species vulnerabilities to future climate change. The foundation of SDMs is ecological niche theory that within certain environmental conditions a species can survive and grow. The species niche can be estimated based on the relationship among observed occurrence or abundance records with environmental variables. Subsequently, the modeled bioclimatic envelope is used to predict the potential geographic distribution of species under current and future climate habitat. However, violating the assumptions of SDMs is likely to weaken the accuracy of niche modeling and distribution predictions. One key assumption of niche theory underpinning SDMs is species-environment equilibrium. It is often assumed that organisms are found in the environmental space that best fits their fundamental requirements, and they are in equilibrium with their environment. However, species often incompletely occupy suitable niche space due to disequilibrium (e.g., dispersal rate, migrating lag) or

geographical barriers, and therefore the potential future habitat of these species is incompletely predicted by SDMs.

Although SDMs can quantify the bioclimatic niche of species, observed records of a species only represent its realized niche, which is not necessarily the same as its fundamental niche. Fundamental niche denotes all theoretically environmental conditions suitable for the species to survive, but the realized niche is based on the observed records where a species has been observed to occur. The realized niche includes factors other than climate, such as biotic interactions (e.g., competition), soil conditions, and geographic barriers. These factors also shape the species distribution, and individuals may not be found throughout all climatic suitable space due to these additional factors. Therefore, the realized niche is only a portion of the fundamental niche. The fundamental niche, given current limited environmental combinations, is normally not known and would require experimentation to explore more thoroughly. Thus, SDMs based on empirical relationships among observed species distribution and environment predictors can only modeling its current realized niche.

Species may expand into unrealized niche portion when under novel climate combinations (Jackson and Overpeck, 2000), and SDMs are likely to provide inaccurate results when extrapolating the species distribution under such circumstances. The performance of SDMs on estimating species potential distribution across temporal scale with no modern analog has previously been studied by training SDMs with fossil pollen records to predict modern “no-analog” communities (Veloz *et al.* 2012). Taxa which were abundant in areas with no-analog late glacial climates substantially shifted their realized niches from the late glacial period, such as *Fraxinus*, *Ostrya/Carpinus* and *Ulmus*, SDMs showed low predictive accuracy when projected to modern climates despite demonstrating high predictive accuracy for late glacial pollen distributions. Meanwhile, for other taxa such as *Quercus*, *Picea*, *Pinus*

strobis, the study showed relatively stable realized niches and models for these taxa tended to have higher predictive accuracy when projected to present.

Jackson & Williams (2004) discussed the mechanisms that might limit the usefulness of SDMs due to individualistic responses of plant species, which include migration lags, geographical barriers, soil development and differing distribution during the Last Glacial Maximum (LGM). For example, different species are believed to be in their refugial phases and locations during different parts. Therefore, cold-adapted, high-latitude taxa are more restricted in refugium during interglacials rather than during glacials when temperate, mid-latitude taxa tend to be in refugia (Stewart & Dalen 2008). Therefore, assuming no interactions among taxa with communities during migration, individualistic responses driven by climate-related variables can be predicted for each single taxon and then resembled for community predictions.

One other assumption of SDMs that may be violated is fundamental niche constancy. Genetic adaptation of species could cause a species' fundamental niche to shift, but is rarely documented in response environmental changes. Normally, genetic adaptation occurs over long time periods. Although the climatic niche shift was reported in some species, the rate of niche shift is consistently low. So far, there is no clear evidence of fundamental niche shifts in response to anthropogenic climate change (e.g., Davis & Shaw 2001; Etterson & Shaw 2001), even for species with short generation times and diverse gene pools, such as grasses (Cang, Wilson & Wiens 2016). Thus, fundamental niche constancy is unlikely to significantly compromise SDM projections because evolution occurs over longer time scales. Over longer evolutionary time scales and more pronounced climate shifts, fundamental niche changes have almost certainly taken place, however. For example, forests described during the Eocene near the Arctic Circle must have been genetically adapted to dark arctic winters.

3. Materials and Methods

3.1 Pollen data

I use pollen records for North America since the last glacial maximum from 23,000 calendar years before present (hereafter referred to as 23 ka) to present from the Neotoma Paleocological Database (<http://api.neotomadb.org>), obtained through the *Neotoma* package (Goring *et al.* 2015) for the R programming environment (R Core Team 2018). Because most of the pollen records are not identified to the species level, 25 pollen taxa developed by Gavin *et al.* (2003) were used in this analysis, including *Abies*, *Acer*, *Alnus*, *Betula*, *Carya*, *Corylus*, *Cupressaceae* plus *Taxaceae*, *Cyperaceae*, *Fagus*, *Fraxinus*, *Juglans*, *Larix*, *Ostrya* and *Carpinus*, *Picea*, *Pinus*, *Platanus*, *Poaceae*, *Populus*, Prairie Forbs, *Quercus*, *Salix*, *Tilia*, *Tsuga*, *Ulmus*, and *Liquidambar*. Although fossil pollen records are associated with vegetation composition, taxa with few pollen records may not fully represent their distribution because of the limitation of seed dispersal rate, abundant, etc. Therefore, six taxa (*Corylus*, *Cupressaceae* & *Taxaceae*, *Larix*, *Platanus*, *Salix*, *Populus*) with very low frequency pollen records that do not fit their current distributions mapped by Little and Viereck (1971) were excluded from the analysis.

In total, the database contained 51,427 dated pollen community records from 1,384 sites after removing records with incomplete geographic coordinates or missing values for dated ages. For samples without calendar years but with radiocarbon dated age, radiocarbon dates were converted to calendar years using the intCal13 calibration curve (Reimer *et al.* 2013) to match the vegetation assembles to the climate given time periods. The numbers of unique sites increases substantially from the last glacial maximum, where few pollen records are available, to present. During the last glacial maximum around only about 100 sites with pollen records were available between 23

ka and 18ka, gradually increasing to about 600 sites at 10 ka and staying relatively constant until the present, with the latest 1000 year interval with comprising 1200 sites (Fig. 1b).

3.2 Climate data

Paleoclimatic data for this study was estimated using the delta method, also referred to as anomaly method, where coarse resolution anomalies for past period are subtracted from a high resolution interpolated climate surface for the present. Past climate reconstructions for the periods 6,000, 11,000, 14,000, 16,000, and 21,000 calendar years before present (defined as 1950) obtained from the general circulation model CCM1 (Community Climate Model v1) (Kutzbach *et al.* 1998). This model was chosen because it provides hindcasts for a number of intermediate climate conditions between the last glacial maximum and the Holocene (Fig 1a). Data from coarse-resolution CCM1 were overlaid on high resolution modern climate data as deviations from the 1961-1990 reference period. The temperature pattern in period 4-0 is general stable, and therefore 1961-1990 climate from ClimateNA v.5.21 (Wang *et al.* 2016) could be used to represent the climate conditions of this period.

Regarding climate variable selection, 19 bioclimatic variables were constructed following the standard equations for worldClim developed by Hijmans *et al.* (2017) with 9 variables, selected to avoid high collinearity: MAT = Annual Mean Temperature (bio1), TS = Temperature Seasonality (standard deviation * 100) (bio4), TD = Temperature Annual Range (bio7), TWetQ = Mean Temperature of Wettest Quarter (bio8), TDryQ = Mean Temperature of Driest Quarter (bio9), TWarmQ = Mean Temperature of Warmest Quarter (bio10), TColdQ = Mean Temperature of Coldest Quarter (bio11), MAP = Annual Precipitation (bio12), PS = Precipitation Seasonality (Coefficient of Variation) (bio15). In addition to these bioclimatic

variables, I also use the growing season precipitation (April to September, GSP), annual heat-moisture index (AHM), and a drought index CMI (Hogg 1997), which was designed to predict the boundary of boreal forests and grassland, with a zero water deficit estimation indicating the transition zone.

3.3 Grouping of pollen data into climate periods

To map the vegetation compositions for each periods, I summarized the pollen data for time periods that correspond to climate reconstructions from the CCM1 for 21, 16, 14, 11, 6, 0 ka. The corresponding intervals for pollen data were chosen at 23-18, 18-15, 15-12, 12-10, 10-4, and 4-0 ka (Fig. 1), so that the climate reconstruction of the GCM represents average conditions for those periods. These intervals were determined using a continuous global reconstruction of climate conditions from ice cores (Petit *et al.* 1999). However, continuous paleoclimate hindcasts are not available, and I consequently have to use one reference climate estimation representing the general climate conditions for each period. Although the absolute values of CCM1 hindcasts for North America indicate climate conditions during the last glacial maximum is lower than the global average (Fig. 1a, cf. right and left scales), the temporal sequence of warming since the last glacial maximum appears to be accurately captured by the general circulation model.

Since pollen cores contain multiple dated records for each of the periods defined above, pollen counts per taxon were summed up across each core and each period, and calculated as taxon percentage for each pollen period and site. This measure reflects the average taxa composition corresponding to each period. Using a percentage of pollen count rather than a presence-absence data as dependent variable for species distribution modeling also makes the analysis more robust against high mobility of pollen exaggerating the taxon's range.

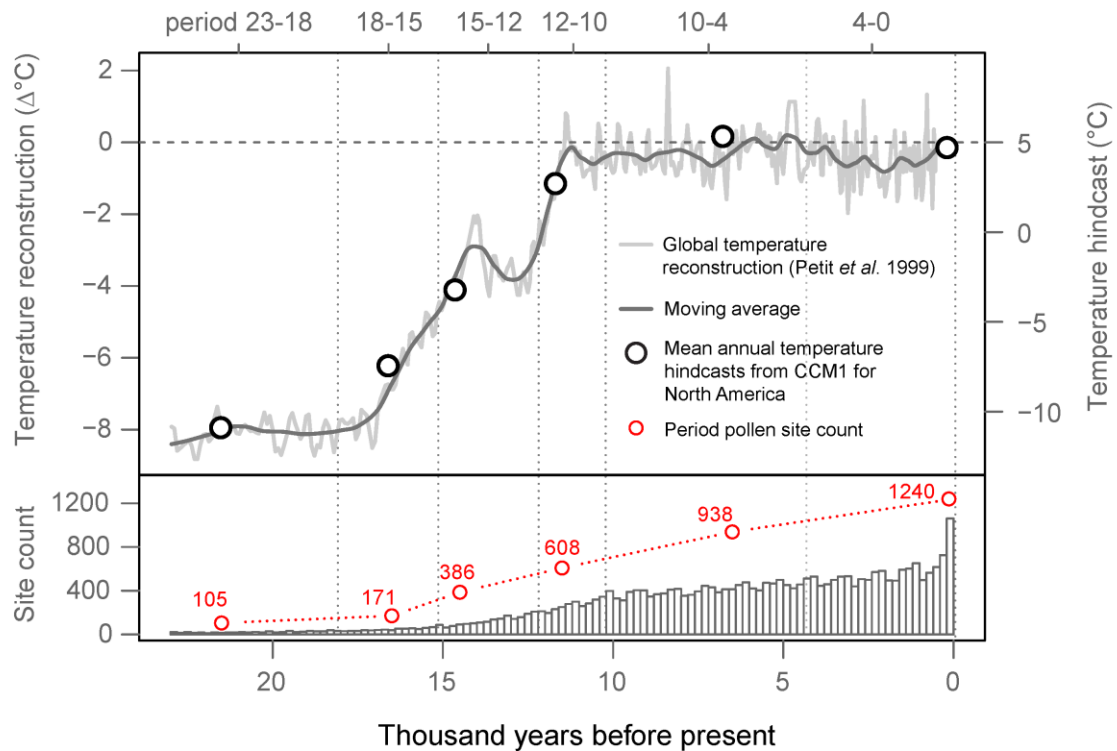


Figure 1. Reconstructed temperature and available pollen sites for six periods from 23,000 years ago to present. Hindcast temperature pattern from CCM1 model for North America are compared with global temperature reconstruction by ice core proxy (Petit *et al.* 1999) between 23,000 years before present (1950) to current. Six reconstructed climate points (black empty circles) from CCM1 model generally fit the moving average curve of the global temperature changes. Red dots represent the sum of available pollen sites for each of the six periods. The number of pollen sites for individual 200-year intervals are shown in histogram.

3.4 Constrained clustering of pollen communities

Since the most pollen records are only available at the level of genera, traditional clustering methods do not yield good measures of similarity due to pollen frequencies alone. For example, a sample from the Pacific Northwest may contain pollen of *Thuja* (red cedar), *Picea* (sitka spruce), *Acer* (bigleaf maple), *Tsuga* (western hemlock) and *Pinus* (shore pine). A sample from the opposite side of the continent may have a

comparable pollen composition with completely different species: *Thuja* (white cedar), *Picea* (red spruce), *Acer* (sugar maple), *Tsuga* (eastern hemlock) and *Pinus* (eastern white pine). To address this problem, I use a constrained clustering technique, multivariate regression trees (MRT) that uses one set of variables (here: climatic characteristics) as predictors for clustering a second set of variables (here: pollen frequencies to infer vegetation communities).

Very similar with regression tree analysis (CART), multivariate regression tree (MRT) is a decision-tree resulting from recursive partitioning of the response data, with indication of the influence of the explanatory variable at each split but for multiple response variables (De'Ath 2002). All possible cut-point of predictors would be tested to minimize the total within-group sum of square deviances (De'Ath 2002). The first split of the tree normally explains the greatest variances of response variables, and the importance of split decrease with smaller branches. Since MRT is suitable for both categorical and continuous data type and no specific assumptions of data distribution, this method is generally robust for explanation and prediction.

Multivariate regression tree analysis was computed by function *mvpert()* of the MVPART package version 1.6-2 (De'Ath 2002) for the R programming environment (R Core Team 2018). In this study, an MRT analysis was run twice using 12 climate variables mentioned above as predictor variables of MRT to form clusters sharing similar vegetation compositions. Firstly, I explore the robustness of the methodological approach by clustering the pollen records with the corresponding climate from 0-4 ka, and compare the results with modern vegetation community distributions to check whether the MRT is capable to make reasonable distinctions among ecological zones or biomes. Subsequently, I use the same method to cluster records from all six paleoclimatic periods to search for novel vegetation communities. It should be noted that the clustering approach does not require novel vegetation

communities to be associated with novel climate conditions. The method simply uses climate variables to delineate clusters, so that the maximum variance in pollen data within clusters is minimized.

3.5 Climate space and no analog climate

To identify if no-analog pollen communities are associated with no-analog climate, I rely on a visual analysis of climate space for 12 bioclimatic variables. Canonical discriminant analysis was used to ordinate pollen samples so that the maximum variance between groups determined by the regression tree analysis is explained. Since the analysis benefits from normally distributed variables, some variables were transformed. The transformations were logarithmic for mean annual precipitation (MAP) and precipitation seasonality (PS), and a square root transformation for growing season precipitation (GSP).

3.6 Species distribution modeling

Because of the restrictions of genus-level pollen data, SDM modeling is also challenging. The climatic envelope of multiple species has to be modeled simultaneously. E.g. the model must be capable of predicting the climate envelope of *Thuja* (red cedar and white cedar), *Picea* (sitka spruce), *Acer* (bigleaf maple), *Tsuga* (western hemlock) and *Pinus* (shore pine, eastern white pine), etc. simultaneously. Random forest is one machine learning method which is widely recognized for its flexibility in species distribution modeling for accommodating interactions of multiple climate variables, and discontinuous niche spaces. By growing a large amount of classification or regression trees from bootstrap subsets of the predictor variables, the model determines the predicted results by summarizing majority vote

over trees (Breiman 2001). This method does not have any assumptions for one particular distribution, which makes it ideal for this study potentially having multimodal data.

For training the RandomForest model implemented with the *randomForest* package (Liaw & Wiener 2002) for the R programming environment, using pollen records from the current period (0-4 ka) and climate data from the 1961-1990 period as training data to build models separately for each taxon. The model was then applied to predict pollen percentages at 0, 6, 11, 14, 16, and 21 ka for the North American continent using gridded climate data obtained with the delta method as described above.

4. Results

4.1 Classifying modern pollen data into bioclimatic zone

A multivariate regression tree analysis of modern pollen samples were split into 12 groups with bioclimatic variables as criteria (Fig. 2a), which explains 53% of the variance in the total dataset consisting of 19 pollen taxa percentages. The clustered groups are generally consistent with the modern ecosystem distributions of North America (*cf.*, Fig. 2b and 2c). Note that closeness of groups in a regression tree analysis does not necessarily imply similarity of communities as in a normal cluster analysis. Rather, the objective of the algorithm is to only minimize variance within clusters, using a second set of variables (here climate) as partitioning criteria. For example, the eastern transition zone between the boreal forest and the eastern mixedwoods is split into two groups in the regression tree: the Eastern Transition (n = 233) and the Forest-boreal Transition (n = 94). The two transition zones therefore end up distant in the regression tree, despite having similar species compositions (Fig. 2a).

Climate variables are often multicollinear, and the regression tree analysis can use more than one variable to separate groups with an identical or very similar grouping. I therefore also report the alternative variables that would lead to similar groupings. The first node (Fig. 2a, #1) of the multiple regression tree roughly separates northern ecosystems along the Canadian-US border, plus Montane Forest ecosystems and Northwest Pacific Coast. This partitioning can be based on mean temperature of warmest quarter (TWarmQ, $R^2=0.17$) or mean annual temperature (MAT, $R^2=0.14$). Groups from warm, southern regions (right branch of the node) tend to have lower conifer but higher hardwood percentages than cold, northern groups (left branch of the node).

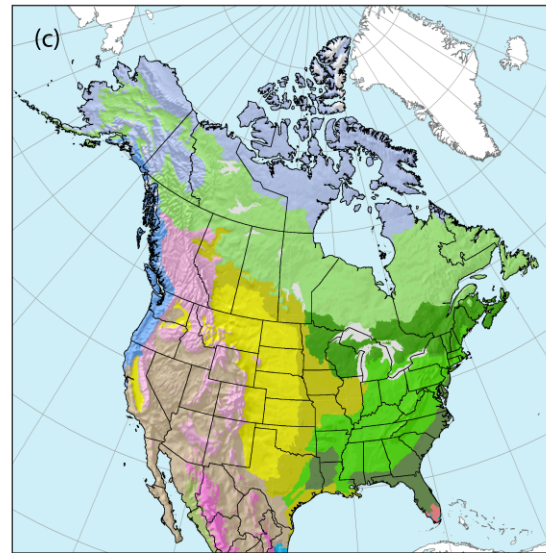
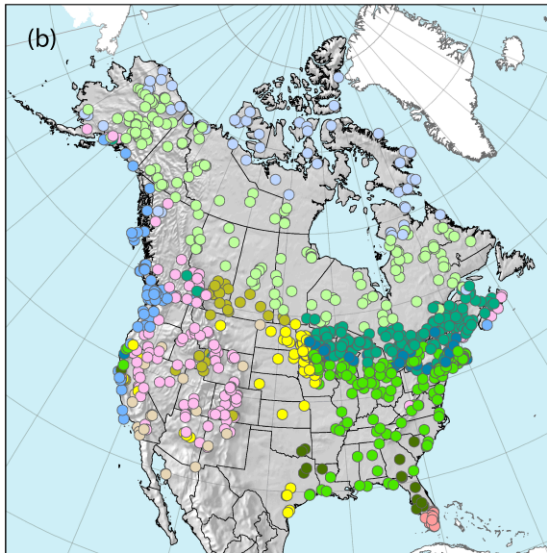
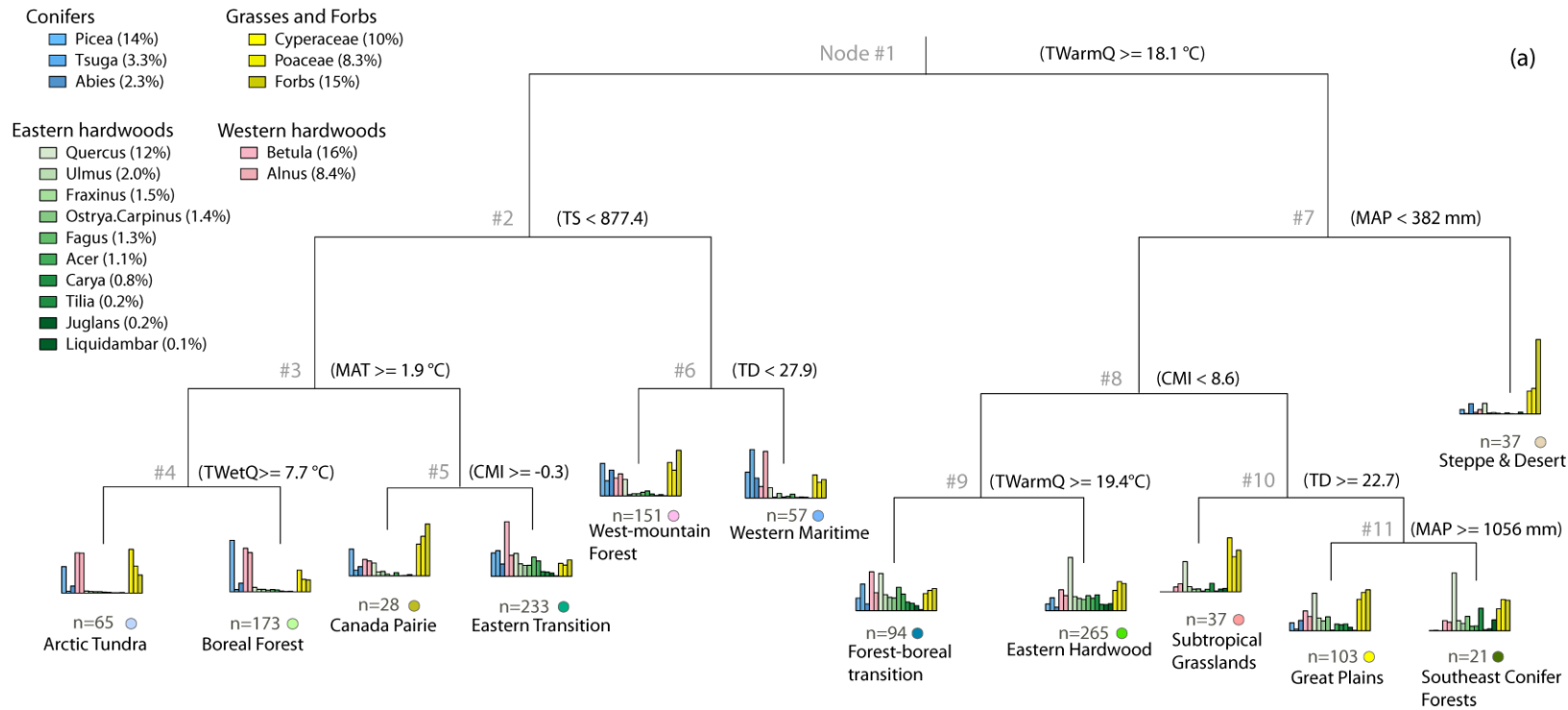


Figure 2. (a) Regression tree analysis for modern pollen records, (b) geographic location of samples and clusters from the analysis, and (c) modern ecological zones for comparison (adapted from Oslon *et al.* 2001). Bioclimatic variables used by the regression tree include MAT (mean annual temperature), TWarmQ (mean temperature of the warmest quarter), TColdQ (mean temperature of the coldest quarter), MAP (mean annual precipitation), TWetQ (mean temperature of the wettest quarter), TD (continentality), TS (temperature seasonality), and CMI (climate moisture index). The bars of (a) represent square root transformed taxa percentages for better visualization.

Temperature-related variables explain the main divisions of northern or high-elevation six groups (Fig. 2a, Node #2). Sites from western North America, including Montane Forests (n = 151) and Pacific Maritime (n = 57), are separated from the northern sites by low temperature seasonality (TS, $R^2=0.14$) or mean temperature of the driest quarter (TDryQ, $R^2=0.12$). These two western groups show similar pollen compositions with median, even percentages of conifer, western hardwood and grasses, but the Maritime sites have visibly more *Alnus* and *Tsuga*, and less grasses and forbs. Node #3 separates two high-latitude groups from mid-latitude ones based on differences of mean annual temperature (MAT, $R^2=0.21$) or mean temperature of the coldest quarter (TColdQ, $R^2=0.20$). Arctic Tundra and Boreal Forest are abundant in both *Betula* and *Alnus*, and the Boreal Forest group is distinguished by a high spruce pollen percentage. Drought indices (CMI or AHM) account for a small amount of vegetation variation in node #5 (2.4%) to separate Eastern Transition from Canada Prairie by more broadleaf species.

For the southern groups on the right side of the regression tree, precipitation and moisture generally account for the subsequent partitioning. Low mean annual precipitation (MAP, $R^2=0.22$) or annual heat-moisture index (AHM, $R^2=0.21$) distinguishes 37 sites as steppe and desert in node #7. Node #8 is based on drought indices CMI, ($R^2=0.13$) or AHM ($R^2=0.13$) and separates wetter eastern hardwood ecosystems from drier grasslands. In the group with the warmest climate conditions, subtropical grasslands, i.e. the Everglades of southern Florida, are separated by low continentality in Node #10 (with alternative variables being TD, MAT, and TColdQ) and is characterized by high pollen percentage of oaks and grasses. In Node #11, lower mean annual precipitation weakly separates the Great Plains from the Southeast Conifer Forests.

4.2 Appearance of new bioclimatic classes in the past

In an expanded analysis, I now apply the same procedure for pollen data with their corresponding paleoclimatic characterization for six paleoclimatic periods from the last glacial maximum to modern conditions. In general, the regression tree analysis will first delineate bioclimatic zones that represent many samples, and thereby explain a large portion of the overall variance in pollen records. To discover groupings that are represented by fewer pollen records in the past (Fig. 1), I allow more groups than in the above tree for modern pollen data. I have chosen 22 groups, beyond which I did not find any additional interpretable communities. However, for concise reporting I subsequently summarize some of the 22 groups that neither showed novel communities or novel climate conditions. To give an example from the regression tree in Fig. 2, the two boreal-broadleaf transition forests could be summarized for simplicity of analysis and interpretation. In the following, I focus on three no-analog communities and the remaining groups summarized into eight broader zones for reference.

The three no-analog communities can be described as Birch Parkland (Figure 3, orange circles), Spruce Woodland (Fig. 3, turquoise circles), and Montane Steppe (Fig. 3, purple circles). The Birch Parkland appears to be the most striking no-analog community, distinguished by exceptionally high pollen percentages of *Birch* and also differs from current Arctic Tundra and Boreal Forest ecosystems by low pollen percentages of *Alnus* from Boreal forest and Arctic Tundra (Table 1). Secondly, the analysis revealed the emergence of unusual combinations of spruce and hardwood species, previously described by Overpeck, Webb and Webb (1992). This ecosystem is dominated by spruce with frequencies comparable to modern boreal forests, but eastern broadleaf taxa also co-exist at higher percentages. The bioclimatic zone was widespread during 11 ka to 14 ka but has no current analog (Fig. 4). Lastly, a

marginally differentiated mountain ecosystem was delineated that may or may not be labeled as no-analog. The community may be corresponding to a previously described Montane Steppe pollen group (Whitlock & Bartlein 1997), with more forbs and slightly lower pollen percentages for tree species such as *Tsuga*, *Alnus* and other broadleaf taxa than the current Montane Forest ecosystem. The Montane Steppe bioclimatic zone is prevalent at 11 and 14 ka (Fig. 4).

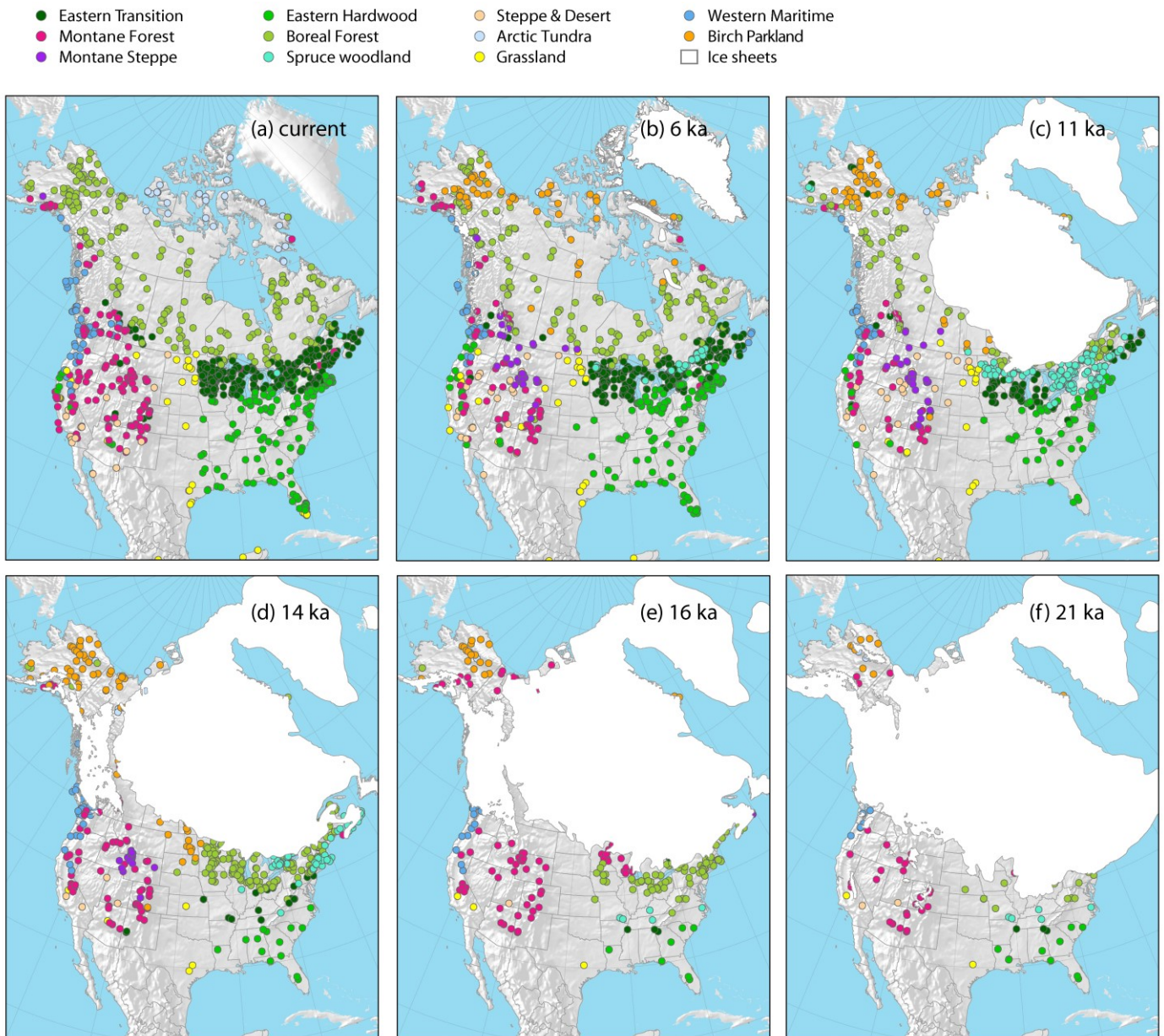


Figure 3. Climate-constrained clusters of pollen sites from a multivariate regression tree analysis for six periods (as shown in Fig. 1). The analysis created 22 pollen groups that were summarized into 11 groups in this map for simplicity, but maintaining groupings for no-analog communities (Montane Steppe, Birch Parkland and Spruce woodland).

Table 1. Average pollen percentages for groups shown in Fig. 3, which were determined with a climate-constrained cluster analysis of pollen sites. The averages were calculated across all six periods shown in Fig. 3.

Ecogroup	Arctic Tundra	Birch Parkland	Boreal Forest	Western Maritime	Eastern Hardwood	Eastern Transition	Spruce Woodland	Grassland	Montane Forest	Montane Steppe	Steppe & Desert
<i>Picea</i> (spruce)	7.90	9.46	32.28	11.12	2.61	6.16	25.65	3.17	13.43	14.43	1.09
<i>Tsuga</i> (hemlock)	0.01	0.00	0.38	23.20	2.70	6.07	3.28	0.15	2.13	0.14	0.02
<i>Abies</i> (fir)	0.03	0.03	1.42	10.10	1.48	1.46	2.36	1.39	4.96	4.77	1.94
<i>Betula</i> (birch)	16.06	40.94	22.76	1.91	4.51	23.05	19.07	1.36	5.65	4.14	0.38
<i>Alnus</i> (alder)	12.46	6.55	12.93	37.41	4.59	4.36	4.24	7.60	8.78	4.08	0.26
<i>Quercus</i> (oak)	0.02	0.27	3.37	1.08	39.73	17.72	12.97	11.22	2.46	0.66	2.74
<i>Ulmus</i> (elm)	0.03	0.24	1.20	0.02	2.46	4.58	3.83	1.43	0.12	0.05	0.15
<i>Fraxinus</i> (ash)	0.01	0.09	1.65	0.14	1.80	2.48	2.69	1.20	0.44	0.01	0.08
<i>Ostrya, Carpinus</i> (hornbeam)	0.04	0.04	1.09	0.01	1.84	2.96	2.81	0.27	0.10	0.03	0.42
<i>Fagus</i> (beech)	0.00	0.00	0.25	0.01	2.36	3.62	0.75	0.01	0.13	0.00	0.00
<i>Acer</i> (maple)	0.00	0.02	0.30	0.16	1.47	2.84	1.44	0.27	0.27	0.11	0.06
<i>Carya</i> (Hickory)	0.00	0.01	0.17	0.00	4.14	0.77	0.43	0.56	0.02	0.01	0.01
<i>Tilia</i> (linden)	0.00	0.00	0.04	0.01	0.21	0.64	0.23	0.03	0.00	0.00	0.00
<i>Juglans</i> (walnut)	0.01	0.05	0.07	0.04	0.43	0.33	0.27	0.12	0.03	0.00	0.03
<i>Liquidambar</i> (sweetgum)	0.00	0.00	0.00	0.00	0.77	0.01	0.01	0.03	0.00	0.00	0.00
<i>Cyperaceae</i> (sedge)	35.32	22.02	11.06	5.42	8.25	5.95	10.49	18.32	11.88	12.21	12.65
<i>Poaceae</i> (grasses)	17.21	11.76	4.88	3.54	10.25	7.51	3.97	23.50	10.90	11.84	11.18
Prairie, Forbs	10.90	8.50	6.17	5.83	10.39	9.50	5.50	29.40	38.69	47.51	68.99

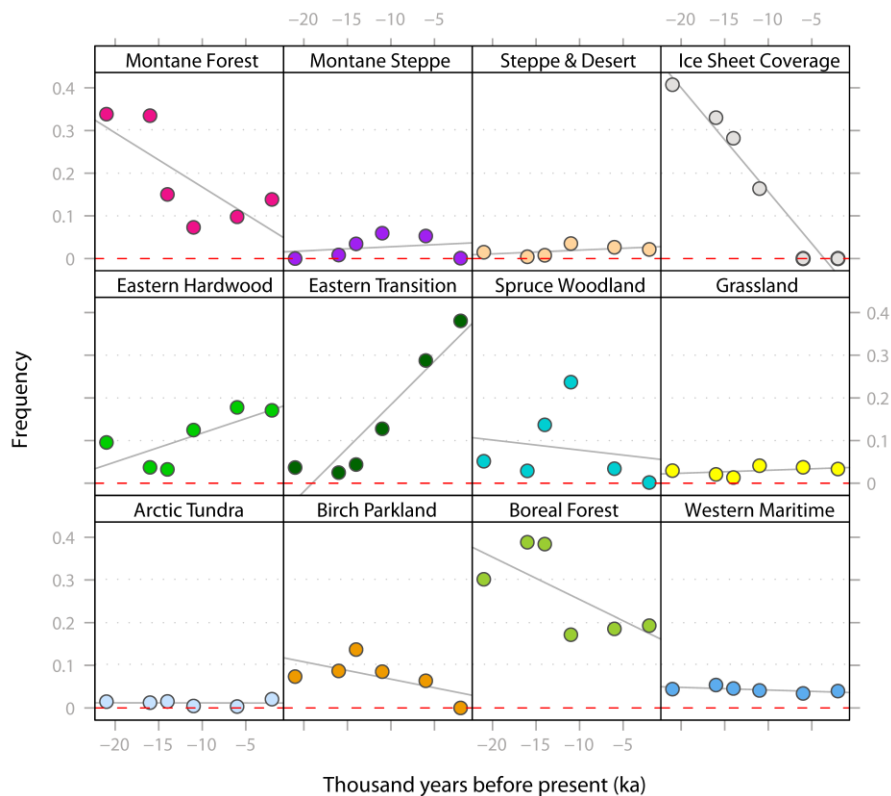


Figure 4. Changes in frequency of pollen community groups from the last glacial maximum to present, as shown in Fig. 3. Frequency of each pollen group for each period is given as percent of all pollen sites available for each period.

4.3 Changes to climate space since LGM

Pollen communities are ordinated in four climate dimensions that maximize the variance explained among bioclimatic zones, and explain 93.2% of the total variance in climate variables (Fig. 5). High values for Axis 1 (Table 2 and Fig. 5a right side) are associated with high temperature seasonality (TD, TS) and dryness (MAP, CMI pointing in the opposite direction) accounting for 40.7% of the total variance. Axis 2 represents temperature variables accounting for 28.9% of the total variance, with warm ecosystems located toward the bottom half of Fig. 5c. Axis 3 and 4 explain 16.1% and 7.4% of total variances, respectively. Axis 3 represents precipitation seasonality, with low growing season precipitation causing high values in precipitation seasonality (Fig. 5c and Table 2). Axis 4 describes cool overall climate conditions associated with favorable growing period, i.e. during the warmest and wettest quarter of the year (Fig. 5c and Table 2).

Two of the no-analog pollen communities appear to correspond to no-analog climate conditions. The no-analog Birch Parkland pollen community (orange circles) appears to occupy a similar climate space to the current Arctic Tundra (light blue circles) for Axes 1 and 2 representing cold, seasonal, dry conditions (cf., Fig. 5a and b). Note that in these two graphs the birch parkland does not overlap with maritime ecosystems at all (Fig. 5a and b, dark blue circles). The situation changes when looking at Axes 3 and 4. Here, the birch parkland overlaps with maritime conditions (Fig. 5d and e, dark blue circles) having relatively warm conditions during the warmest and wettest quarter of the year and low precipitation seasonality. The uniqueness of this birch parkland climate combination may allow for *Betula* to dominate the landscape when it would be arctic tundra under drier and colder growing season conditions.

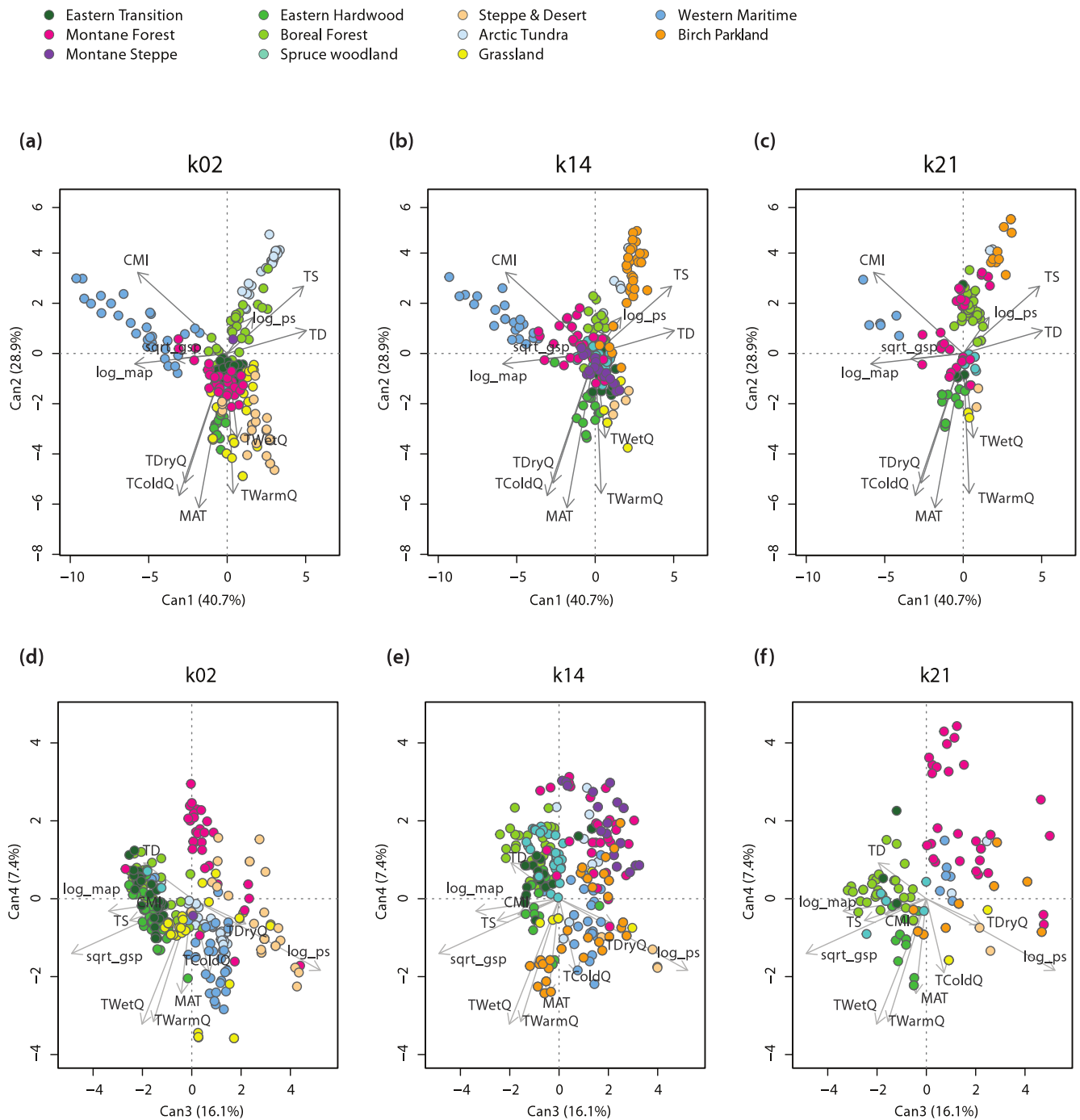


Figure 5. Canonical discriminant analysis of 11 bioclimatic zones with four climate dimensions explaining 93.1% of the total variance in climate variables. The top row is a plot of component 2 over component 1 and the bottom row represents component 4 over component 3 for three selected climate periods. Plots for all six climate periods are provided in Appendix 1.

Table 2. Loadings of canonical discriminant analysis of groups from all six periods corresponding to Fig. 5. The first axis (Can1) explains 40.7% variance; the second axis (Can2) explains 28.9%; the third (Can3) and fourth (Can4) explains 16.1% and 7.4%, respectively.

Climate variable	Can1	Can2	Can3	Can4
Mean Annual Temperature (MAT)	-0.26	-0.88	-0.06	-0.35
Mean Temperature of Warmest Quarter (TWarmQ)	0.05	-0.80	-0.22	-0.45
Mean Temperature of Wettest Quarter (TWetQ)	0.09	-0.48	-0.29	-0.46
Mean Temperature of Coldest Quarter (TColdQ)	-0.44	-0.81	0.10	-0.27
Mean Temperature of Driest Quarter (TDryQ)	-0.38	-0.74	0.31	-0.09
Temperature Annual Range (TD)	0.72	0.13	-0.28	0.13
Temperature Seasonality (TS)	0.69	0.38	-0.36	-0.08
Climate Moisture Index (CMI)	-0.82	0.46	-0.24	-0.06
Mean Annual Precipitation (log_MAP)	-0.84	-0.06	-0.48	-0.05
Growing Season Precipitation (sqrt_GSP)	-0.48	-0.03	-0.70	-0.20
Precipitation Seasonality (log_PS)	0.24	0.21	0.74	-0.26

The no-analog Montane Steppe pollen community is generally warmer (Axis 2) and drier during the growing season (Axis 1) than current Montane Forest ecosystems (cf. Fig. 5a and b). It does, however, overlap the climatic space with the current Montane Forest group, and is therefore not clearly distinguished from modern montane ecosystems. The third no-analog community, spruce woodland, appears less distinguished in our plots. It showed overlap of climatic space with Eastern Transition ecosystem but a slightly higher precipitation seasonality (Fig. 5b). This bioclimatic zone appears in the center of the ordination for all axes and for all periods from the last glacial maximum to present. Plots of all paleoclimatic periods are provided on Appendix 1.

4.4. Can SDMs predict the pollen frequencies of novel communities?

Lastly, I used a species distribution model approach to hindcast pollen frequencies of individual taxa, based on modern samples, to test if past no-analog communities can be predicted by species distribution models. While species distribution models could hind-cast birch presence in novel climate space, they failed to predict the uniquely high frequencies of birch in the observed birch parkland community. Similarly the uniquely high frequencies of spruce in the 14 to 11 ka spruce woodlands was not correctly predicted. Generally, the uncertainty in projections for the past exceeded the degree to which no-analog communities were distinguished from current ecosystems.

Table 3. Pollen percentages predicted with a species distribution model (Random Forest) for the same groups as shown in Table 1. The model was trained with pollen data from the modern period (0-4 ka), but the averages shown below were calculated across all six periods as in Table 1. Thus, only the no-analog communities (Birch Parkland, Spruce Woodland and Montane Steppe) are completely independent predictions that did not overlap with training data.

Ecogroup Taxon (%)	Arctic Tundra	Birch Parkland	Boreal Forest	Western Maritime	Eastern Hardwood	Eastern Transition	Spruce Woodland	Grassland	Montane Forest	Montane Steppe	Steppe & Desert
<i>Picea</i> (spruce)	10.41	18.99	30.80	12.45	1.20	5.47	14.29	3.41	14.29	17.45	2.45
<i>Tsuga</i> (hemlock)	0.11	0.32	0.56	33.84	3.68	6.28	3.41	0.45	3.07	1.62	0.58
<i>Abies</i> (fir)	0.30	0.55	1.78	8.38	0.71	1.21	2.03	0.96	4.23	2.98	1.64
<i>Betula</i> (birch)	16.11	22.67	26.93	4.96	6.67	23.27	23.22	3.69	13.43	20.54	3.20
<i>Alnus</i> (alder)	20.91	24.84	21.81	20.40	4.57	5.26	9.35	4.82	17.55	14.53	3.39
<i>Quercus</i> (oak)	0.13	0.83	0.90	1.72	37.31	17.98	11.76	14.13	2.22	1.83	4.85
<i>Ulmus</i> (elm)	0.07	0.17	0.30	0.07	2.19	2.85	2.09	1.18	0.26	0.61	0.45
<i>Fraxinus</i> (ash)	0.03	0.11	0.27	0.22	1.88	1.74	1.23	1.09	0.21	0.29	0.33
<i>Ostrya</i> . <i>Carpinus</i> (hornbeam)	0.06	0.12	0.19	0.08	1.17	2.17	1.23	0.48	0.15	0.25	0.18
<i>Fagus</i> (beech)	0.02	0.06	0.46	0.30	3.19	4.32	2.26	0.27	0.67	0.61	0.17
<i>Acer</i> (maple)	0.04	0.08	0.27	0.32	1.75	2.52	1.61	0.44	0.35	0.47	0.19
<i>Carya</i> (Hickory)	0.01	0.03	0.04	0.10	4.18	0.81	0.79	1.13	0.06	0.05	0.41
<i>Tilia</i> (linden)	0.01	0.02	0.02	0.01	0.28	0.57	0.35	0.13	0.03	0.06	0.07
<i>Juglans</i> (walnut)	0.01	0.01	0.02	0.02	0.54	0.28	0.20	0.12	0.03	0.03	0.06
<i>Liquidambar</i> (sweetgum)	0.00	0.00	0.00	0.01	0.65	0.04	0.04	0.08	0.01	0.00	0.06
<i>Cyperaceae</i> (sedge)	33.32	19.60	9.18	10.53	7.98	6.35	8.05	18.52	18.29	12.74	16.77
<i>Poaceae</i> (grasses)	14.08	7.25	3.34	3.22	11.28	9.54	9.51	21.70	8.85	8.10	13.61
Prairie.Forbs	4.37	4.36	3.11	3.36	10.76	9.36	8.59	27.40	16.32	17.83	51.59

5. Discussion

5.1 No analog communities and no analog climate

Three no-analog pollen communities which were widespread during the last glaciations or during deglaciation were identified (Fig. 3). The most distinct no-analog community was a birch parkland ecosystem with exceptionally high frequencies of *Betula*, *Poaceae*, *Cyperaceae* and forb pollen. The second, widespread no modern analog community was a previously described spruce woodland with spruce having a role as dominant as in today's boreal forest, but a large variety of broadleaf species also present at lower frequencies. This community generally matches the no-analog community described by Overpeck *et al.* (1992). Unlike detail-rich, local scale study by Williams *et al.* (2001), this study is based on general pollen percentages grouped by macro-climatic conditions, and could not provide the same detail as Williams *et al.* (2001), with the communities they describe being likely also influenced by local site factors. The third no-analog community, Montane Steppe, which the analysis produced was not very distinctive in either composition or climate conditions. It was included here because the group was quite large and widespread, and it is consistent with previous research by Whitlock and Bartlein (1997) who analyzed vegetation records for a montane site in the western Cascade mountains that dates back 125 ka. They describe repeated transitions between climate conditions that support forests versus drier conditions that resulted in steppe ecosystems. During deglaciation, around 11 ka, high summer temperatures, decreased effective precipitation, caused intensified drought and associated vegetation changes. In our analysis, pollen records have a higher proportion of forbs rather than trees, which seem to be a plausible ecological response to warmer and drier climate conditions during the growing season for a montane ecosystem that was more widespread during deglaciation than today.

One plausible explanation for the emergence of no-analog communities is that they are the result of climatic conditions that do not exist today. This appears to be the case for the birch parkland bioclimatic zone, which currently has no equivalent climatic space (Fig. 5). Williams, Shuman and Webb (2001) also found that the timing and spatial distribution of vegetation and climate dissimilarity indicates the appearance of no-analog vegetation communities is due to now extinct environmental conditions, rather than different plant taxa showing different migration lag during rapid climate change. Although other factors may also cause the emergence of no-analog communities, climate might be the primary factor of vegetation transitions (Whitlock & Bartlein 1997).

Our analysis, however, also provides counter examples. The other two no-analog vegetation communities show no distinct associations with no-analog climate and largely overlap with the climatic niches of modern bioclimatic zones (Fig. 5). The spruce woodland community centering on the Great Lakes region during deglaciation (Fig. 3) might be driven by environmental factors other than climate. Newly exposed soils following glacier retreat might have unusual characteristics and soil microbial communities (e.g., Nemergut *et al.* 2007; Schmidt *et al.* 2008). Vegetation community compositions evolve from young deglaciated areas until soil conditions mature and stabilize (e.g., Finegan 1984; Chapin *et al.* 1994).

Lastly, the individual responses of species to climate change could explain no-analog compositions. Due to differences of climate tolerances, dispersal ability, and migration rates, transitional no-analog vegetation communities may emerge (e.g., Urban, Tewksbury & Sheldon 2012). This may not be the most plausible explanation for the spruce woodland no-analog community which was characterized by high frequencies of spruce, while also including hardwoods associated with warmer climate conditions.

Spruce dispersal rates are not particularly high, but the species is capable to grow under relatively poor soil and nutrient conditions that were likely present following deglaciation.

5.2 Clustering pollen communities

Analysis of pollen records to infer ecological communities either from current or from fossil records poses some challenges. Species differ in the amount of pollen that they produce, some species distribute their pollen far beyond their range and therefore do not represent the local vegetation community, pollen of some species is perishable and not well preserved in the fossil record, and pollen identification is often only possible at the genus or family level.

This means that generally pollen communities should only serve as an indicator for actual communities that correspond to the pollen groups (i.e. the frequencies of species on Fig. 2a and b do not truly reflect modern species compositions shown in and 2c). Nevertheless, the principles of pollen analysis assumes that there is constant and interpretable relationship between the pollen deposited in the sediment and the local vegetation (Ritchie 1995). Similarly, it has to be assumed that no-analog pollen assemblages are originated from no-analog vegetation communities.

Pollen identification has limited taxonomic resolution, and many ecologically important taxa or species have indistinguishable pollen grains. The low-resolution fossil information limits inferences on the type of corresponding plant communities, and may also lead to incorrect paleoclimate reconstructions based on pollen data (e.g., Birks & Birks 2000). Further, the pollen-vegetation relationship varies among taxa and sites due to landscape characters and dispersal capabilities (e.g., Fagerlind 1952; Goring *et al.* 2013; Dawson *et al.* 2016). For example, Schofield *et al.* (2007)

reported *Betula*, *Salix* and *Poaceae* are over-represented in pollen assemblages whereas *Juniperus* which is poorly represented.

In order for a quantitative analysis of pollen communities to work, I had to make a number of subjective choices in this analysis. First, I excluded six taxa with low frequency pollen that did not conform well to their current ranges (*Corylus*, *Cupressaceae* & *Taxaceae*, *Larix*, *Platanus*, *Salix*, *Populus*). Second, I also excluded pollen records of the genus *Pinus*, which had the most abundant and widespread pollen records, but occurred far beyond its current distribution, e.g. reaching high percentages in the Canada prairies where no pine species exist. It has previously been noted that species with high production and wide dispersal of pollen represent a complicating factor for paleocommunity reconstruction (Delcourt, Delcourt & Webb 1984; Faegri, Kaland & Krzywinski 1989).

As an alternative to excluding the species, I also tried transformations that diminish the influence of common species in community ecology analysis, i.e. a Hellinger transformation (square root of taxa percentage per site), standardization (normalize individual species frequencies – i.e. columns of a standard data table), and a Wisconsin transformation (normalize columns then rows). None of these methods yielded satisfactory results for clustering modern pollen communities with *Pinus* included. The transformations also did not improve bioclimatic zones when compared to ecological delineations after *Pinus* was removed. I therefore used the untransformed percentages of taxa pollen per sites with no pine records. Nevertheless, I used a Hellinger transformation for visualization of taxa composition for groups, but not for analysis (Fig. 2a).

5.3 All no analog communities may be difficult to predict

In this analysis, a species distribution modeling failed to predict the unique composition of the no-analog bioclimatic zones. This is likely to be caused by violating species distribution model assumptions. For the birch parkland, there is no modern training data to make predictions, and the model makes predictions by extrapolating beyond currently known climate space. In the case of the spruce woodlands, the key model deficiency may be the lack of including habitat parameters other than climate. If young soils after glacial retreat are a major factor in determining community composition, then model projections cannot be expected to be accurate. For the Montane Steppe bioclimatic zone, the model predicts very similar pollen frequencies as for the modern montane ecosystems. The model under-predicts forbs and over-predicts *Alnus* and *Betula* pollen percentages for both ecosystems, indicating that this bioclimatic zone is generally not well modeled. It appears to infer high pollen percentages of *Alnus* and *Betula* from other cold ecosystems such as boreal forests and arctic tundra.

Another key assumption underpinning species distribution models that could have compromised the accuracy of pollen percentage predictions for no-analog communities is that species distributions and environmental conditions must be in equilibrium (e.g., Pearson & Dawson 2003; Araujo & Guisan 2006; Thuiller *et al.* 2008). It is often assumed that organisms are found in the environmental space that best fits their fundamental requirements, and they are in equilibrium with their environment. Hence, species incompletely occupy suitable niche space due to disequilibrium (e.g., dispersal rate, migrating lag) cannot be captured by SDMs.

However, in this study I did not find an example that could illustrate this effect. At the time-scales used in this analysis, climate summaries and pollen summaries spanning

several thousand years for each paleoclimatic period, migration lags may not play a dominant role. The birch parkland could be such an example, because birch species are pioneer species with large seed production and far ranging seed dispersal. It could have colonized the available habitat before other, slower migrating species. However, the counter argument would be that the birch parkland also existed during the stable last glacial maximum climate conditions. It also exists today on other continents, and this analysis also showed that it is associated with unique climate conditions.

6. Conclusions

In summary, this thesis contributes a broad analysis of North American fossil pollen data and I discovered one no-analog community that was clearly linked to no-analog climate conditions: a birch parkland that was characterized by cold maritime climate conditions that does not exist today. A second widespread no-analog community that was previously described for the great lakes region, a spruce woodland, was not associated with no-analog climate conditions and may have been driven by young soils following deglaciation. In both cases, violation of assumptions of SDMs, in the first case climate equivalence, in the second case lack of consideration of habitat factors other than climate appear to prevent accurate predictions of community composition in the past, and by inference also for the future.

In general, novel climate conditions appear to be the main driver of the emergence of novel ecological communities, and predictions that rely on extrapolations for the future appear challenging with species distribution models that model the realized niche based on climate conditions only. Furthermore, future climate change is expected to be orders of magnitude faster than historical climate change since the last glacial maximum. This means that drivers of no-analog communities caused by the

disequilibrium would likely become more important in the future with differential migration rates being the primary driver of ecosystem composition in newly available habitat. Predictions of no-analog communities would therefore require models that incorporate non-climatic factors, species' environmental tolerances, as well as mechanistic modelling components to take migration rates into consideration. Additionally, human land use, landscape fragmentation, biological invasions, increases in atmospheric CO₂, may interact with novel climate conditions to yield yet more ecological uncertainty, which may become too complex to predict.

7. References

- Anagnostou, E., John, E.H., Edgar, K.M., Foster, G.L., Ridgwell, A., Inglis, G.N., Pancost, R.D., Lunt, D.J. & Pearson, P.N. (2016) Changing atmospheric CO₂ concentration was the primary driver of early Cenozoic climate. *Nature*, **533**, 380.
- Araujo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677-1688.
- Badger, M.P., Schmidt, D.N., Mackensen, A. & Pancost, R.D. (2013) High-resolution alkenone palaeobarometry indicates relatively stable p CO₂ during the Pliocene (3.3–2.8 Ma). *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **371**, 20130094.
- Birks, H.H. & Birks, H.J.B. (2000) Future uses of pollen analysis must include plant macrofossils. *Journal of Biogeography*, **27**, 31-35.
- Breiman, L. (2001) Random forests. *Machine learning*, **45**, 5-32.
- Broccoli, A.J. & Manabe, S. (1987) The influence of continental ice, atmospheric CO₂, and land albedo on the climate of the last glacial maximum. *Climate Dynamics*, **1**, 87-99.
- Bromwich, D.H., Toracinta, E.R., Oglesby, R.J., Fastook, J.L. & Hughes, T.J. (2005) LGM summer climate on the southern margin of the Laurentide Ice Sheet: wet or dry? *Journal of Climate*, **18**, 3317-3338.
- Burke, K.D., Williams, J.W., Chandler, M.A., Haywood, A.M., Lunt, D.J. & Otto-Bliesner, B.L. (2018) Pliocene and Eocene provide best analogs for near-future climates. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, 13288-13293.
- Cang, F.A., Wilson, A.A. & Wiens, J.J. (2016) Climate change is projected to outpace rates of niche change in grasses. *Biology Letters*, **12**.
- Chapin, F.S., Walker, L.R., Fastie, C.L. & Sharman, L.C. (1994) Mechanisms of Primary Succession Following Deglaciation at Glacier Bay, Alaska. *Ecological Monographs*, **64**, 149-175.
- Clark, P.U., Alley, R.B. & Pollard, D. (1999) Northern Hemisphere ice-sheet influences on global climate change. *Science*, **286**, 1104-1111.
- Cohmap Members (1988) Climatic changes of the last 18,000 years: observations and model simulations. *Science*, 1043-1052.
- Coope, G. & Angus, R. (1975) An ecological study of a temperate interlude in the middle of the last glaciation, based on fossil Coleoptera from Isleworth, Middlesex. *The Journal of Animal Ecology*, 365-391.

- Correa-Metrio, A., Bush, M.B., Cabrera, K.R., Sully, S., Brenner, M., Hodell, D.A., Escobar, J. & Guilderson, T. (2012) Rapid climate change and no-analog vegetation in lowland Central America during the last 86,000 years. *Quaternary Science Reviews*, **38**, 63-75.
- Davis, M.B. & Shaw, R.G. (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673-679.
- Dawson, A., Paciorek, C.J., McLachlan, J.S., Goring, S., Williams, J.W. & Jackson, S.T. (2016) Quantifying pollen-vegetation relationships to reconstruct ancient forests using 19th-century forest composition and pollen data. *Quaternary Science Reviews*, **137**, 156-175.
- De'Ath, G. (2002) Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology*, **83**, 1105-1117.
- Delcourt, P.A., Delcourt, H.R. & Webb, T. (1984) *Atlas of mapped distributions of dominance and modern pollen percentages for important tree taxa of eastern North America*. American Association of Stratigraphic Palynologists Foundation.
- Eberle, J.J. & Greenwood, D.R. (2012) Life at the top of the greenhouse Eocene world-A review of the Eocene flora and vertebrate fauna from Canada's High Arctic. *Geological Society of America Bulletin*, **124**, 3-23.
- Edwards, M.E., Brubaker, L.B., Lozhkin, A.V. & Anderson, P.M. (2005) Structurally novel biomes: A response to past warming in Beringia. *Ecology*, **86**, 1696-1703.
- Eldrett, J.S., Greenwood, D.R., Harding, I.C. & Huber, M. (2009) Increased seasonality through the Eocene to Oligocene transition in northern high latitudes. *Nature*, **459**, 969-U991.
- Etterson, J.R. & Shaw, R.G. (2001) Constraint to adaptive evolution in response to global warming. *Science*, **294**, 151-154.
- Fægri, K., Kaland, P.E. & Krzywinski, K. (1989) *Textbook of pollen analysis*. John Wiley & Sons Ltd.
- Fagerlind, F. (1952) *The real significance of pollen diagrams*.
- Finegan, B. (1984) Forest succession. *Nature*, **312**, 109.
- Gavin, D.G., Oswald, W.W., Wahl, E.R. & Williams, J.W. (2003) A statistical approach to evaluating distance metrics and analog assignments for pollen records. *Quaternary Research*, **60**, 356-367.
- Gleason, H.A. (1926) The Individualistic Concept of the Plant Association. *Bulletin of the Torrey Botanical Club*, **53**, 7-26.

- Goring, S., Dawson, A., Simpson, G., Ram, K., Graham, R., Grimm, E. & Williams, J. (2015) Neotoma: A programmatic interface to the Neotoma Paleocological Database. *Open Quaternary*, **1**.
- Goring, S., Lacourse, T., Pellatt, M.G. & Mathewes, R.W. (2013) Pollen assemblage richness does not reflect regional plant species richness: a cautionary tale. *Journal of Ecology*, **101**, 1137-1145.
- Graham, R.W. (2005) Quaternary mammal communities: relevance of the individualistic response and non-analogue faunas. *The Paleontological Society Papers*, **11**, 141-158.
- Graham, R.W., Lundelius, E.L., Graham, M.A., Schroeder, E.K., Toomey, R.S., Anderson, E., Barnosky, A.D., Burns, J.A., Churcher, C.S., Grayson, D.K., Guthrie, R.D., Harington, C.R., Jefferson, G.T., Martin, L.D., McDonald, H.G., Morlan, R.E., Semken, H.A., Webb, S.D., Werdelin, L. & Wilson, M.C. (1996) Spatial response of mammals to late quaternary environmental fluctuations. *Science*, **272**, 1601-1606.
- Greenwood, D.R. & Basinger, J.F. (1993) Stratigraphy and Floristics of Eocene Swamp Forests from Axel-Heiberg Island, Canadian Arctic Archipelago. *Canadian Journal of Earth Sciences*, **30**, 1914-1923.
- Greenwood, D.R., Basinger, J.F. & Smith, R.Y. (2010) How wet was the Arctic Eocene rain forest? Estimates of precipitation from Paleogene Arctic macrofloras. *Geology*, **38**, 15-18.
- Hijmans, R.J. & Graham, C.H. (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272-2281.
- Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2017) dismo: Species distribution modeling.
- Hogg, E.H. (1997) Temporal scaling of moisture and the forest-grassland boundary in western Canada. *Agricultural and Forest Meteorology*, **84**, 115-122.
- Huber, M. & Caballero, R. (2011) The early Eocene equable climate problem revisited. *Climate of the Past*, **7**, 603-633.
- IPCC, C.C. (2007) The physical science basis. *Contribution of working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, **996**.
- Jackson, S.T. & Overpeck, J.T. (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194-220.
- Jackson, S.T. & Williams, J.W. (2004) Modern analogs in Quaternary paleoecology: Here today, gone yesterday, gone tomorrow? *Annual Review of Earth and Planetary Sciences*, **32**, 495-537.

- Kutzbach, J., Gallimore, R., Harrison, S., Behling, P., Selin, R. & Laarif, F. (1998) Climate and biome simulations for the past 21,000 years. *Quaternary Science Reviews*, **17**, 473-506.
- LePage, B.A. (2001) New species of *Picea* A. Dietrich (Pinaceae) from the middle Eocene of Axel Heiberg Island, Arctic Canada. *Botanical Journal of the Linnean Society*, **135**, 137-167.
- Lepage, B.A. & Basinger, J.F. (1991) A New Species of *Larix* (Pinaceae) from the Early Tertiary of Axel-Heiberg Island, Arctic Canada. *Review of Palaeobotany and Palynology*, **70**, 89-111.
- Liaw, A. & Wiener, M. (2002) Classification and regression by randomForest. *R news*, **2**, 18-22.
- Little, E.L. & Viereck, L.A. (1971) *Atlas of United States trees*. US Dept. of Agriculture, Forest Service.
- Manabe, S. & Broccoli, A. (1985) The influence of continental ice sheets on the climate of an ice age. *Journal of Geophysical Research: Atmospheres*, **90**, 2167-2190.
- McGuffie, K. & Henderson-Sellers, A. (2005) *A climate modelling primer*. John Wiley & Sons.
- Meinshausen, M., Smith, S.J., Calvin, K., Daniel, J.S., Kainuma, M.L.T., Lamarque, J.F., Matsumoto, K., Montzka, S.A., Raper, S.C.B., Riahi, K., Thomson, A., Velders, G.J.M. & van Vuuren, D.P.P. (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213-241.
- Moss, R., Babiker, W., Brinkman, S., Calvo, E., Carter, T., Edmonds, J., Elgizouli, I., Emori, S., Erda, L. & Hibbard, K. (2008) Towards new scenarios for the analysis of emissions: Climate change, impacts and response strategies. Intergovernmental Panel on Climate Change Secretariat (IPCC).
- Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, J.F.B., Nakicenovic, N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson, A.M., Weyant, J.P. & Wilbanks, T.J. (2010) The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747-756.
- Nemergut, D.R., Anderson, S.P., Cleveland, C.C., Martin, A.P., Miller, A.E., Seimon, A. & Schmidt, S.K. (2007) Microbial community succession in an unvegetated, recently deglaciated soil. *Microbial Ecology*, **53**, 110-122.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C. & Loucks, C.J., (2001) Terrestrial Ecoregions of the World: A New Map of Life

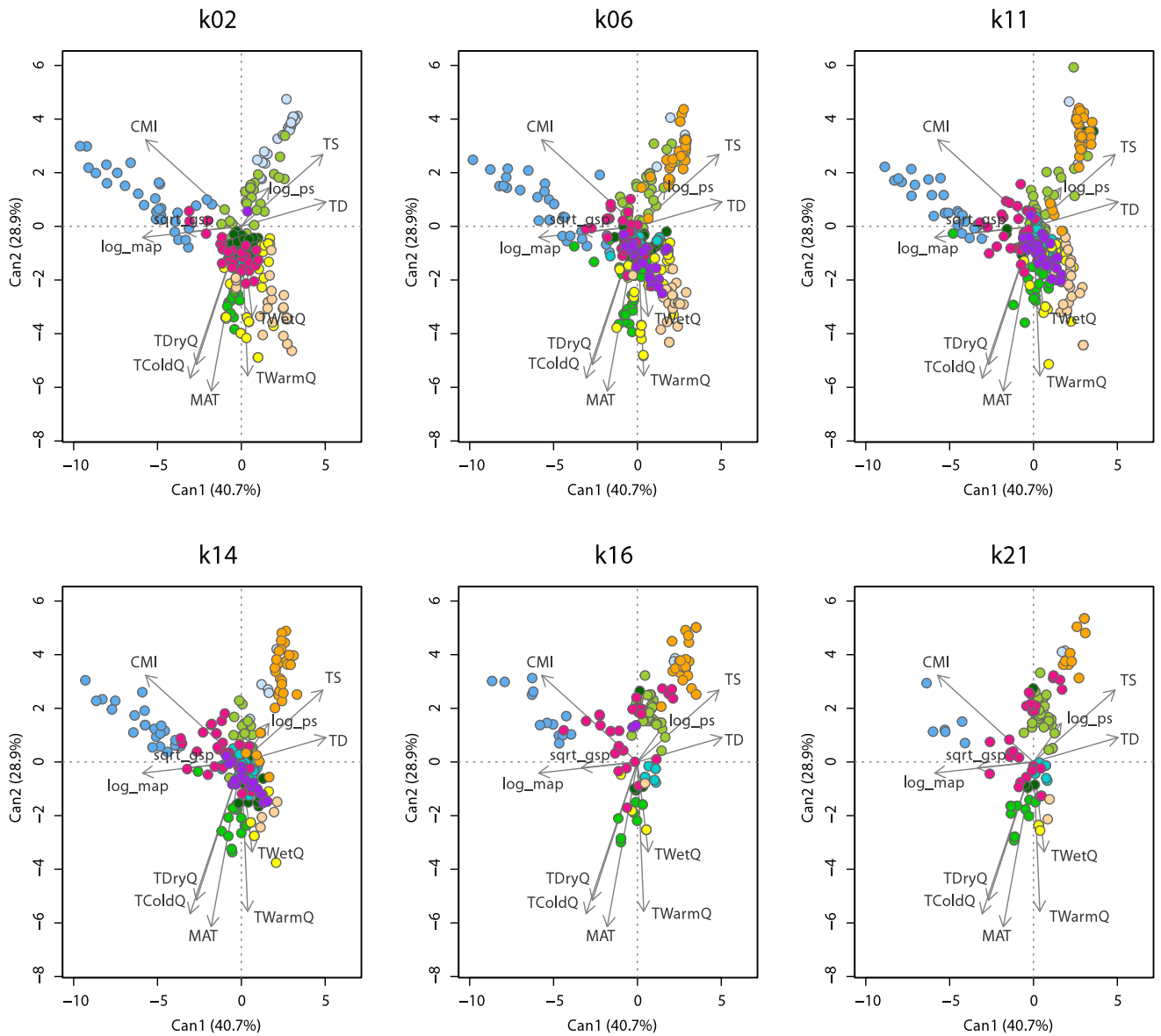
- on EarthA new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, **51**, 933-938.
- Overpeck, J.T., Webb, R.S. & Webb, T. (1992) Mapping Eastern North-American Vegetation Change of the Past 18 Ka - No-Analogs and the Future. *Geology*, **20**, 1071-1074.
- Overpeck, J.T., Webb, T. & Prentice, I.C. (1985) Quantitative Interpretation of Fossil Pollen Spectra - Dissimilarity Coefficients and the Method of Modern Analogs. *Quaternary Research*, **23**, 87-108.
- Pagani, M., Liu, Z.H., LaRiviere, J. & Ravelo, A.C. (2010) High Earth-system climate sensitivity determined from Pliocene carbon dioxide concentrations. *Nature Geoscience*, **3**, 27-30.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361-371.
- Pedlar, J.H., McKenney, D.W., Aubin, I., Beardmore, T., Beaulieu, J., Iverson, L., O'Neill, G.A., Winder, R.S. & Ste-Marie, C. (2012) Placing Forestry in the Assisted Migration Debate. *Bioscience*, **62**, 835-842.
- Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V.Y., Lorius, C., Pepin, L., Ritz, C., Saltzman, E. & Stievenard, M. (1999) Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature*, **399**, 429-436.
- Preece, R. & Bridgland, D.R. (2012) *Late Quaternary Environmental Change in North-west Europe: Excavations at Holywell Coombe, South-east England: Excavations at Holywell Coombe, South-east England*. Springer Science & Business Media.
- R Core Team (2018) R Foundation for Statistical Computing; Vienna, Austria: 2014. *R: A language and environment for statistical computing*, 2013.
- Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Ramsey, C.B., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Hafliðason, H., Hajdas, I., Hatte, C., Heaton, T.J., Hoffmann, D.L., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., Manning, S.W., Niu, M., Reimer, R.W., Richards, D.A., Scott, E.M., Southon, J.R., Staff, R.A., Turney, C.S.M. & van der Plicht, J. (2013) Intcal13 and Marine13 Radiocarbon Age Calibration Curves 0-50,000 Years Cal Bp. *Radiocarbon*, **55**, 1869-1887.
- Roy, K., Jablonski, D. & Valentine, J.W. (1995) Thermally Anomalous Assemblages Revisited - Patterns in the Extraprovincial Latitudinal Range Shifts of Pleistocene Marine Mollusks. *Geology*, **23**, 1071-1074.

- Schmidt, S.K., Reed, S.C., Nemergut, D.R., Grandy, A.S., Cleveland, C.C., Weintraub, M.N., Hill, A.W., Costello, E.K., Meyer, A.F., Neff, J.C. & Martin, A.M. (2008) The earliest stages of ecosystem succession in high-elevation (5000 metres above sea level), recently deglaciated soils. *Proceedings of the Royal Society B-Biological Sciences*, **275**, 2793-2802.
- Schubert, B.A., Jahren, A.H., Eberle, J.J., Sternberg, L.S.L. & Eberth, D.A. (2012) A summertime rainy season in the Arctic forests of the Eocene. *Geology*, **40**, 523-526.
- Seki, O., Foster, G.L., Schmidt, D.N., Mackensen, A., Kawamura, K. & Pancost, R.D. (2010) Alkenone and boron-based Pliocene pCO₂ records. *Earth and Planetary Science Letters*, **292**, 201-211.
- Shinn, R.A. & Barron, E.J. (1989) Climate sensitivity to continental ice sheet size and configuration. *Journal of Climate*, **2**, 1517-1537.
- Shuman, B., Thompson, W., Bartlein, P. & Williams, J.W. (2002) The anatomy of a climatic oscillation: vegetation change in eastern North America during the Younger Dryas chronozone. *Quaternary Science Reviews*, **21**, 1777-1791.
- Stewart, J.R. & Dalen, L. (2008) Is the glacial refugium concept relevant for northern species? A comment on Pruett and Winker 2005. *Climatic Change*, **86**, 19-22.
- Taylor, K.E., Stouffer, R.J. & Meehl, G.A. (2012) An Overview of Cmpip5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93**, 485-498.
- Tedford, R.H. & Harington, C.R. (2003) An Arctic mammal fauna from the Early Pliocene of North America. *Nature*, **425**, 388-390.
- Thuiller, W., Albert, C., Araujo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgely, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E. (2008) Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, **9**, 137-152.
- Urban, M.C., Tewksbury, J.J. & Sheldon, K.S. (2012) On a collision course: competition and dispersal differences create no-analogue communities and cause extinctions during climate change. *Proceedings of the Royal Society B-Biological Sciences*, **279**, 2072-2080.
- van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J. & Rose, S.K. (2011) The representative concentration pathways: an overview. *Climatic Change*, **109**, 5-31.
- Veloz, S.D., Williams, J.W., Blois, J.L., He, F., Otto-Bliesner, B. & Liu, Z.Y. (2012) No-analog climates and shifting realized niches during the late quaternary:

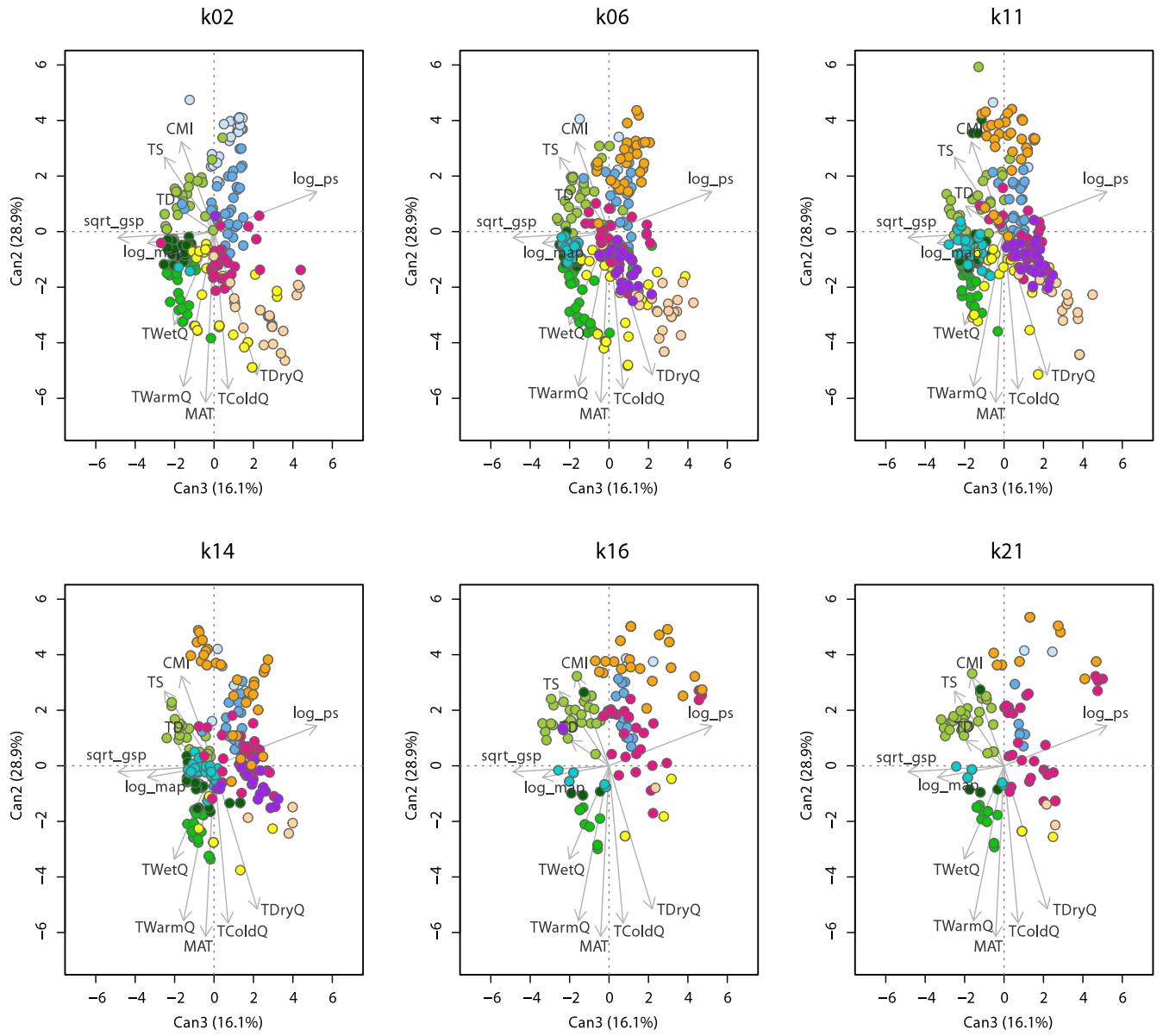
- implications for 21st-century predictions by species distribution models. *Global Change Biology*, **18**, 1698-1713.
- Wang, T.L., Hamann, A., Spittlehouse, D. & Carroll, C. (2016) Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *Plos One*, **11**.
- Whitlock, C. & Bartlein, P.J. (1997) Vegetation and climate change in northwest America during the past 125 kyr. *Nature*, **388**, 57-61.
- Williams, C.J., Johnson, A.H., LePage, B.A., Vann, D.R. & Sweda, T. (2003) Reconstruction of Tertiary Metasequoia forests. II. Structure, biomass, and productivity of Eocene floodplain forests in the Canadian Arctic. *Paleobiology*, **29**, 271-292.
- Williams, J.W., Grimm, E.C., Blois, J.L., Charles, D.F., Davis, E.B., Goring, S.J., Graham, R.W., Smith, A.J., Anderson, M., Arroyo-Cabrales, J., Ashworth, A.C., Betancourt, J.L., Bills, B.W., Booth, R.K., Buckland, P.I., Curry, B.B., Giesecke, T., Jackson, S.T., Latorre, C., Nichols, J., Purdum, T., Roth, R.E., Stryker, M. & Takahara, H. (2018) The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource. *Quaternary Research*, **89**, 156-177.
- Williams, J.W. & Jackson, S.T. (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, **5**, 475-482.
- Williams, J.W., Shuman, B.N. & Webb, T. (2001) Dissimilarity analyses of late-Quaternary vegetation and climate in eastern North America. *Ecology*, **82**, 3346-3362.
- Williams, J.W., Shuman, B.N., Webb, T., Bartlein, P.J. & Leduc, P.L. (2004) Late-quaternary vegetation dynamics in north america: Scaling from taxa to biomes. *Ecological Monographs*, **74**, 309-334.
- Williams, M.I. & Dumroese, R.K. (2013) Preparing for Climate Change: Forestry and Assisted Migration. *Journal of Forestry*, **111**, 287-297.
- Wright Jr, H.E., Winter, T.C. & Patten, H.L. (1963) Two pollen diagrams from southeastern Minnesota: problems in the regional late-glacial and postglacial vegetational history. *Geological Society of America Bulletin*, **74**, 1371-1396.

8. Appendix

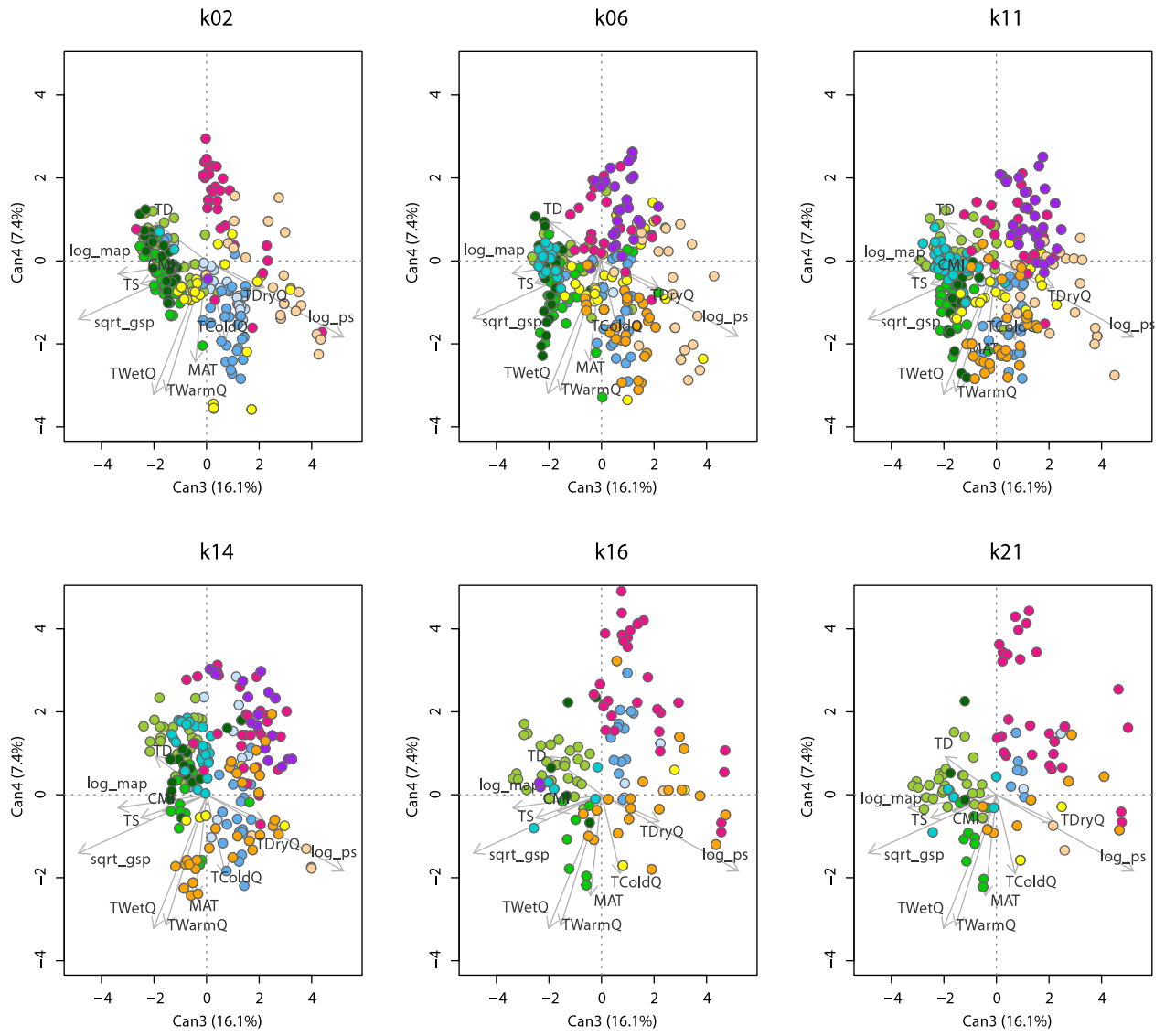
- Eastern Transition
- Eastern Hardwood
- Steppe & Desert
- Western Maritime
- Montane Forest
- Boreal Forest
- Arctic Tundra
- Birch Parkland
- Montane Steppe
- Spruce woodland
- Grassland



Appendix 1. Canonical discriminant analysis of 11 bioclimatic zones with four climate dimensions explaining 93.1% of the total variance in climate variables. Component 2 over component 1 for all six periods.



Appendix 1 continued. Component 2 over component 3 for all six periods.



Appendix 1 continued. Component 4 over component 3 for all six periods.