

**Environmental change and population history  
of North America from the late Pleistocene to the Anthropocene**

Michelle Chaput

Department of Geography, Environment and Geomatics  
University of Ottawa

A thesis submitted in partial fulfillment of the requirements for the  
Doctorate in Philosophy degree in Geography

Examining committee:

James Savelle, McGill University

Elena Ponomarenko, University of Ottawa

Michael Sawada, University of Ottawa

André Viau, University of Ottawa

Supervisor:

Konrad Gajewski, University of Ottawa

© Michelle Chaput, Ottawa, Canada, 2018

## Abstract

The assumption that prehistoric Native American land use practices had little impact on the North American landscape persists in the literature. However, recent research suggests the effects of prehistoric burning, deforestation and agriculture may potentially have been greater than previously considered. To resolve this discrepancy, quantitative estimates of changes in human population size and forest structure and composition over the course of the Holocene are needed. This thesis addresses this need by providing radiocarbon-based paleodemographic reconstructions and pollen-inferred estimates of vegetation change, as well as analyses of associations between the two at both continental and regional scales, from the late Pleistocene to the Anthropocene.

One way to estimate paleodemographic change is to use the number of radiocarbon ( $^{14}\text{C}$ ) dates from a given area to study patterns of human occupation through time. A review of the literature and compilation of existing databases relevant to this method showed there is now sufficient data to study the paleodemographic history of many regions around the world. An analysis of  $^{14}\text{C}$  datasets from North America and Australia compared well with model-based reconstructions of past demographic growth, and provided higher frequency fluctuations in population densities that will be important for future research.

Using a kernel density estimation approach, the first estimates of prehistoric population density for North America were obtained and synthesized into a series of continental-scale maps showing the distribution and frequency of  $^{14}\text{C}$  dates in the Canadian Archaeological Radiocarbon Database (CARD). The maps illustrated the space-time evolution of population and migration patterns, which were corroborated by independent sources of evidence. A methodology based on the statistical evaluation of cross-correlations between population and plant abundance was then developed to analyze the associations between these population estimates and plant communities derived from pollen databases. Periods of high spatial cross-correlation (positive and negative) between population and plant abundance were irregular and did not improve over time, suggesting that ancient human impacts are not discernable at a continental scale, either due to low populations or varying human land use practices. To further examine the relationship between pollen data and human land use at a regional scale, estimates of plant density and landscape openness are needed. The REVEALS (Regional Estimates of VEgetation Abundance from Large Sites) model corrects for the non-linear relationship between pollen production and plant abundance and can therefore be used to map histories of land use and land cover change. The model was applied to pollen records from lake sediments in the deciduous forest of southeastern Quebec. A preliminary analysis comparing these results to population density revealed low population during times of high *Populus* abundance and high population following the appearance of the mixed temperate forest suggesting a discernable human-environment association at regional scales.

Overall, the results of this thesis support the growing body of literature that suggests prehistoric Native Americans impacted their environments and that these impacts can be detected and quantified by integrating archaeological and paleoecological information. However, the timing, location, and intensity of human land use has changed in both space and time, suggesting regional- to local-scale analyses of human-environment interactions are most appropriate for continental North America. The methodology presented here can be used to study additional North American regions for the purpose of developing a continental history of human-environment interaction.

*To Mom, Dad and Mireille,  
for your unconditional love and support,  
and for always making me feel like I could truly accomplish anything I set my mind to,*

*to Josh,  
for always putting a smile on my face,  
and helping me find strength in simply being myself,*

*and to Konrad,  
for being my champion and never losing faith in me,  
and for inspiring me to be a better scientist and a better person*

## Acknowledgements

To my supervisor Konrad Gajewski: I owe it all to you. Without your subtle nudges in the right direction and your unwillingness to let me get away with procrastination, I would have been lost for sure. At times you filled the role of life coach and personal psychologist, and dedicated many hours to discussing the ups and downs of academia, which will forever be remembered. I truly could not have found a better mentor. Merci beaucoup!

I am grateful to my fiancé Josh, my mother Cathy, and my sister Mireille. I love you all unconditionally and will forever be thankful for your encouragement and support. A very special gratitude goes out to my father Guy: you have always been my inspiration, and you always know exactly what to say to lift my heart up. Throughout my studies, my personal mantra has always been *keep calm and call dad*.

Special gratitude goes out to my best friends for making sure I played hooky from time to time, and to my lab mates and fellow graduate students for all the wonderful memories.

I would like to thank Shinya Sugita and Marie-José Gaillard-Lehmdal for your instruction and friendship, Matthew Betts and Andrew Martindale for welcoming me to the CARD team, and Bjoern Kriesche, Volker Schmidt and Rafal Kulik for the opportunity to collaborate.

I would like to mention my committee members Michael Sawada, Elena Ponomarenko, André Viau and James Savelle for all your help, as well as Luisa Veronis and Eric Crighton for giving me opportunities to grow personally and professionally. I thank Chantale Arcand and Nathalie Maras for having all the answers!

Finally, I would like to say a special thank you to Sylvie Thériault for welcoming a shy outsider into the Department with open arms and convincing me I had the skills to complete a degree in geography, and my high school physics teacher Ruth Chmara for recognizing my potential and for showing me women have a place in science.

I thank you all for giving me the chance to learn, I am truly grateful.

## **Financial support**

This work would not have been possible without the financial support of the Ontario Student Assistance Program, the Ontario Graduate Scholarship Program and the University of Ottawa Admission Scholarship Program in the Faculty of Arts. I am especially indebted to my supervisor, Dr. Konrad Gajewski, who generously provided additional funding from a Natural Sciences and Engineering Discovery Grant.

For conference and travel purposes, I received funding from the University of Ottawa Conference Travel Grant Program, the Research and Publications Fund of the Faculty of Arts, the Canadian Association of Palynologists, the Iowa Lakeside Laboratory, the University of Notre Dame in Indiana, the National Centre of Competence in Climate Research, EarthCube, the American Quaternary Association's Denise Gaudreau award and the Polar Continental Shelf Program.

For research purposes, I was awarded the University of Ottawa Geography Graduate Student Scholarship, the Nicole Bégin-Heick Scholarship, and the TD Canada Trust Green Bursary.

The German Academic Exchange Service provided funds to co-authors BK and VS for travel and research purposes abroad.

## Table of contents

Abstract.....	ii
Dedication.....	iii
Acknowledgments.....	iv
Financial support.....	v
Table of contents.....	vi
List of figures.....	x
List of tables.....	xii
1 Introduction.....	1
1.1 History of North America: humans, climate and environment.....	1
1.2 Controversy over prehistoric North American land use impacts.....	5
1.3 Population data .....	8
1.4 Vegetation data .....	10
1.5 Quantifying human-environment interactions.....	12
1.6 The Early Anthropocene Hypothesis.....	13
1.7 Summary of research questions and methods.....	15
1.7.1 Research question #1 .....	16
1.7.2 Research question #2.....	16
1.7.3 Research question #3.....	16
1.7.4 Research question #4.....	17
1.8 Organization of thesis and author contributions.....	17
2 Radiocarbon dates as estimates of ancient human population size.....	21
2.1 Abstract.....	21
2.2 Introduction.....	22
2.3 Archaeological radiocarbon data at a continental scale.....	23
2.3.1 Radiocarbon background and methods .....	23
2.3.2 Interpreting radiocarbon datasets .....	24
2.3.3 Archaeological datasets available online or in publications.....	27

2.4 Regional-scale analyses .....	32
2.5 Data-model comparison.....	32
2.6 Conclusion .....	33
3 Spatio-temporal distribution of Holocene populations in North America.....	41
3.1 Abstract.....	41
3.2 Introduction.....	42
3.3 Methods .....	43
3.4 Distribution of <sup>14</sup> C dates .....	45
3.5 Data verification .....	46
3.6 Results.....	46
3.6.1 Alaska .....	46
3.6.2 Canadian Arctic .....	47
3.6.3 British Columbia and the ice-free corridor.....	47
3.6.4 Eastern United States.....	48
3.6.5 Atlantic Canada .....	48
3.6.6 Central and Western United States.....	49
3.7 Discussion.....	49
4 Estimation of spatio-temporal correlations of prehistoric population and vegetation in North America.....	55
4.1 Abstract.....	55
4.2 Introduction.....	56
4.2.1 Motivation .....	56
4.2.2 Population intensity maps.....	56
4.2.3 Outline .....	57
4.3 The Neotoma Paleoecology Database .....	57
4.3.1 Calibration of radiocarbon ages.....	58
4.3.2 Data selection .....	59
4.4 Temporal interpolation and smoothing of pollen abundances.....	59
4.5 Vegetation intensity maps.....	61

4.5.1 Nonparametric estimation of vegetation intensity maps .....	61
4.5.2 Computation of taxon ranges.....	63
4.5.3 Interpretation of results.....	64
4.6 Correlation analysis of population and vegetation intensities .....	64
4.6.1 Cross-correlations of population and vegetation intensity maps.....	64
4.6.2 Cross-correlations of population and vegetation changes with temporal lag .....	67
4.6.3 Nonparametric confidence bands of cross-correlation functions .....	68
4.7 Interpretation and discussion of correlation results .....	69
4.8 Conclusions.....	71
5 Human-vegetation interactions during the Holocene in North America .....	81
5.1 Abstract.....	81
5.2 Introduction.....	82
5.2.1 North America as a cultural landscape .....	82
5.2.2 North America as a relatively undisturbed landscape .....	83
5.2.3 Resolving the viewpoints of ancient North America .....	84
5.3 Methods .....	86
5.3.1 Data.....	86
5.3.2 Methods .....	86
5.4 Results.....	89
5.4.1 Population.....	89
5.4.2 Boreal and widespread taxa ( <i>Picea</i> and <i>Pinus</i> ) .....	89
5.4.3 Major taxa: <i>Quercus</i> and <i>Castanea</i> .....	90
5.4.4 Major mast trees: <i>Juglans</i> and <i>Carya</i> .....	91
5.4.5 Northern deciduous forest: <i>Fagus</i> and <i>Acer</i> .....	92
5.4.6 Disturbance indicators: <i>Populus</i> and Poaceae .....	93
5.5 Discussion.....	93
5.5.1 Summary of results .....	93
5.5.2 Vegetation history .....	94

5.5.3 Human impacts on tree species over the Holocene at continental scales .....	95
5.5.4 Summary and implications .....	97
6 Relative pollen productivity estimates and changes in Holocene vegetation cover in the deciduous forest of southeastern Quebec, Canada.....	104
6.1 Abstract.....	104
6.2 Introduction.....	105
6.3 Study area .....	107
6.4 Methods .....	108
6.4.1 Pollen data .....	108
6.4.2 Vegetation data.....	109
6.4.3 Extended R-value modelling.....	110
6.4.4 REVEALS modelling.....	111
6.5 Results.....	112
6.5.1 Pollen data .....	112
6.5.2 Vegetation data.....	113
6.5.3 ERV analysis .....	113
6.5.4 REVEALS model .....	114
6.6 Discussion.....	115
6.6.1 Modern vegetation data and interpolation methods .....	116
6.6.2 Pollen productivity and source area .....	116
6.6.3 Vegetation reconstruction.....	117
7 Summary and conclusions .....	133
7.1 Summary.....	133
7.2 Conclusions.....	136
References.....	141
Appendix 1 (Supplementary information for chapter 3).....	170
Appendix 2 (Video for chapter 3).....	196
Appendix 3 (Supplementary information for chapter 5).....	197
Appendix 4 (Videos for chapter 5) .....	208

## List of figures

Figure 2.1	Global distribution of archeological radiocarbon dates.....	35
Figure 2.2	Global frequency distributions of radiocarbon dates.....	36
Figure 2.3	North American frequency distributions of radiocarbon dates .....	37
Figure 2.4	Comparison of CARD and AustArch databases with HYDE3.1 estimates .....	38
Figure 3.1	Locations of sites in the CARD and sampling intensity base map.....	51
Figure 3.2	Radiocarbon Frequency Population Estimates between 0.5 and 13 ka BP .....	52
Figure 3.3	Population estimates at European contact .....	53
Figure 4.1	Population intensity maps for 2.6 – 3.1 ka BP and 9.3 – 9.8 ka BP .....	73
Figure 4.2	Relative pollen abundance for two sites and ten taxa.....	74
Figure 4.3	Analysis of errors when exchanging calibration and interpolation of ages.....	75
Figure 4.4	Vegetation intensity maps for <i>Quercus</i> .....	76
Figure 4.5	Cross-correlations for <i>Quercus</i> and population between 12.7 and 0.75 ka BP .....	77
Figure 4.6	Cross-correlations for 500-year changes in <i>Quercus</i> and population .....	78
Figure 4.7	Examples of confidence bands for cross-correlations .....	79
Figure 4.8	Half of the maximum width of confidence bands for cross-correlations .....	80
Figure 5.1	Neotoma and CARD distribution and mean July temperature anomalies.....	99
Figure 5.2	Overview of methodology related to cross-correlation functions .....	100
Figure 5.3	Diagrams summarizing cross-correlations for ten taxa .....	103
Figure 6.1	Location of observation sites for modern vegetation survey.....	121
Figure 6.2	PCA of modern pollen data from 16 sites .....	122
Figure 6.3	PCA of fossil pollen data from 4 sites.....	123
Figure 6.4	IDW modern vegetation surfaces .....	124
Figure 6.5	PCA of taxon-specific values of mean vegetation cover.....	125
Figure 6.6	Relative Source Area of Pollen .....	126
Figure 6.7	Scatterplots of vegetation and adjusted pollen proportion .....	127
Figure 6.8	Pollen percentages and REVEALS reconstruction .....	128
Figure 6.9	PCA of the pollen data and the REVEALS reconstruction .....	129

Figure 7.1	Study area showing pollen sites and <sup>14</sup> C dates .....	138
Figure 7.2	Pollen percentage diagram, REVEALS diagram and <sup>14</sup> C dates .....	139
Figure 7.3	Pollen diagram from Lake Tortue and <sup>14</sup> C dates .....	140
Figure A1.1	Maps showing effects of choosing different kernel function radii on RFPEs.....	177
Figure A1.2	RFPE maps for the past 13,000 years at 500-year intervals .....	178
Figure A1.3	Comparison of original RFPEs with RFPEs excluding anomalous dates .....	189
Figure A1.4	Comparison of original RFPEs with RFPEs excluding 50% of dates .....	190
Figure A1.5	Histogram of bandwidths based on Scott's rule .....	191
Figure A1.6	Maps comparing nearest neighbour and variable kernel effects .....	192
Figure A1.7	Locations of sites in the CARD described as anomalous .....	193
Figure A1.8	Histograms of dates within 25 km of three sites .....	194
Figure A1.9	Calibration using IntCal04 and IntCal13 .....	195
Figure A3.1	Heat maps summarizing spatial correlations of maps of 10 taxa .....	197

## List of tables

Table 1.1	Summary of North American population estimates .....	8
Table 1.2	Potential start dates for the Anthropocene .....	15
Table 2.1	Summary of global digital radiocarbon databases .....	39
Table 2.2	URLs of online radiocarbon databases .....	40
Table 3.1	Summary of major patterns in population change .....	54
Table 6.1	Lake features for 20 lakes in southeastern Quebec.....	130
Table 6.2	Fall speeds of pollen .....	131
Table 6.3	Relative pollen productivity estimates (PPEs).....	132

# 1 Introduction

## 1.1 History of North America: humans, climate and environment

The North American continent spans latitudes from the Arctic to the subtropics and contains numerous biomes, ecotones and ecosystems. It has a topography ranging from a flat interior to mountainous and icy coasts, a complex network of freshwater lakes and drainage systems, and a climate governed by air masses originating from polar, maritime and tropical latitudes. As a result of this geography, the continent is characterized by a complex environmental history influenced by both natural and anthropogenic forces.

The initial human migration out of eastern Siberia into Beringia and the Americas is estimated to have occurred between 30,000 and 15,000 years before present (BP) via the Bering Land Bridge that previously connected Alaska and Russia. During the Last Glacial Maximum (LGM; ~ 27,000 – 18,000 BP), ocean waters became locked up in vast ice sheets exposing shallow continental shelves forming a 5,000 km passage between the two hemispheres. It is widely accepted that the first inhabitants of the Americas navigated this passage on foot and on rafts along its interior and coastlines (Dixon, 2001; Mandryk et al., 2001; Erlandson et al., 2007), and that their emigration was prompted by a mixture of climate-induced food scarcity and environmental homogenization (Kelly and Todd, 1988), the lure of high coastal productivity (Erlandson et al., 2015), big game hunters tracking prey (Hopkins et al., 1982), a limited sense of geography (Kelly, 1996), conflict over land or food sources (Kelly, 1996), and simple curiosity (Parekh and Childs, 2016). Controversy still exists, however, over how and when this migration proceeded and the rate at which humans were able to occupy and exploit unglaciated lands. Martin (1973) and Surovell et al. (2016) have suggested that the first migrants may have travelled from the southern margins of the North American ice sheets to the southern tip of South America in approximately 1,000 years, but the fossil record of this migration is sparse and new findings may force the reconsideration of this question.

Linguistic, archaeological and DNA analyses suggest humans were present on the Land Bridge approximately 30,000 BP (Nichols, 1990; Campbell, 1997; Hoffecker et al., 2016; Graf and Buvit, 2017). The earliest undisputable evidence of occupation in the continental interior of North America, however, suggests humans arrived no earlier than 15,000 BP (Goebel et al., 2008). Small populations could have survived in Beringia during this 15,000-year time gap, labelled the “Beringian Standstill”, since floral and faunal diversity was high and pockets of forest and tundra persisted throughout the LGM despite harsh climatic conditions (Hoffecker et al., 2014; 2016). However, it is still unclear whether there was a true Beringian Standstill during which Beringians became genetically isolated from Siberians, or if populations were free to travel between Siberia and Beringia and continuously exchange genetic material (Tamm et al., 2007). Further, conflicting genetic evidence suggests there could have either been a single wave of migration (Raghavan et al., 2015; Moreno-Mayar et al., 2018), two waves (Hubbe et al., 2010), or three or more waves (Achili et al., 2013) once Beringians finally began moving into the interior. Regardless of which migration interpretation is correct, the founding population(s) was small (Hey, 2005; Kitchen et al, 2008), and traveled swiftly on foot and on rafts along a Pacific

coastal route (Mandryk et al., 2001; Waguespack et al., 2007; Erlandson et al., 2008; Bodner et al., 2012) and then east towards the remainder of unglaciated North America.

The transition from the late Pleistocene to the early Holocene coincides with the evolution of the Clovis culture beginning around 13,300 BP. Countless Clovis sites have been identified across the continent (O'Brien and Buchanan, 2017). Clovis peoples are known for crafting fluted projectile points with intricate flaking patterns and blade edges used to fashion spears, darts and knives. At this time, air temperatures across much of North America were rapidly increasing and precipitation levels were higher than present (Viau et al., 2006; Ruddiman, 2008), although with considerable variability, such as a short-lived return to glacial conditions during the Younger Dryas (12,900 – 11,700 BP). Beringia, the ice-free corridor between the Laurentide and Cordilleran ice sheets, and much of Atlantic Canada, were dominated by shrub and herb tundra, and taiga, mixed parkland and cool mixed forest covered the majority of western, central and eastern North America, respectively (Williams et al., 2004; Dyke, 2005).

Some sources suggest that this period represents the beginnings of quantifiable (and potentially irreversible) human impacts on North American landscapes and ecosystems (i.e., the Overkill Hypothesis; Martin, 1973; 1984). Clovis peoples were skilled hunters, adept at hunting megafauna such as mastodons, bison, sloths and horses. Overkill, brought about by the efficacy of Clovis spears and perhaps the use of fire to control animal paths, along with the naiveté of the animals to human hunting, led to a rapid decline in the number of megafauna species to the point of eventual extinction (Fowler and Konopik, 2007). This is apparent in studies comparing archaeological and paleoenvironmental records which show strong correlations between initial megafaunal declines across much of the Americas and the first evidence of permanent human occupation (e.g., Surovell et al., 2016). Gradual warming and drying which resulted in novel vegetation communities likely exacerbated the process (Cooper et al., 2015), however the relative contribution of the climate versus direct human impacts to their extinction has yet to be quantified. The resulting loss of an entire trophic level (megaherbivores) may have driven a cascading loss of many megacarnivores, and the structure, functioning and composition of ecosystems dependent on trampling, grazing and regular fertilization changed immensely. In most cases, the loss of megafauna directly preceded intensifications in fire regimes and the rapid replacement of open woodlands with forest, since large populations of megaherbivores tend to suppress above ground biomass (and thus flammable fuels) and maintain open parklands (with few mature trees) due to their browsing habits (Malhi et al., 2016).

During the subsequent Paleoindian period (10,500 – 9,500 BP), humans continued to hunt with modified projectiles that were better suited for smaller prey (Zettler, 2015). Charcoal and radiocarbon analyses suggest they also used fire to encourage the growth of grasses, forbs and shrubs, and to force their prey out of densely forested areas into large open clearings (e.g., Bonnicksen, 2000; Fowler and Konopik, 2007; Anderson et al., 2009; Lightfoot and Cuthrell, 2015). The climate of North America was warming, although the rate of temperature increase was slower than that of previous millennia (Viau et al., 2006). Northeastern North America was still covered by the Laurentide ice sheet, and temperatures in this area remained colder than at

present (Ruddiman, 2008). The biomes continued to shift in response to glacial retreat and stands of mixed forest began forming in the Great Lakes and New England regions. Steppe dominated the central United States and the warm mixed forest was developing in the southeast (Williams et al., 2004).

For many years, Paleoindian fire use has been overlooked in discussions regarding past environmental change in North America. This is largely due to the widely-spread misconception that small and ancient populations lacked the capacity to alter their environments in long-lasting, measurable ways (e.g., see Sauer, 1956; Vale, 2002). However, archaeological and paleobotanical evidence has since demonstrated that prehistoric human activities have caused significant environmental transformations across many parts of the continent (e.g., see Cronon, 1983; Denevan, 1992; Foster, 2000; Denevan, 2011). In eastern North America in particular, numerous case studies suggest pre-agricultural burning allowed disturbance-tolerant, fire-resistant taxa to become dominant at the cost of taxa typical of undisturbed environments or late successional forests (e.g., Abrams, 1992; Munoz and Gajewski, 2010; Munoz et al., 2010). Similar ideas found in other texts also suggest pioneer and disturbance species such as *Pinus*, *Quercus* and *Populus* only covered large areas of North America because Paleoindian fire use created large openings in which these species could grow (e.g., Bonnicksen, 2000). As a result, several authors now argue Paleoindian fire use is a key driver of past changes in vegetation and fire regime, particularly in biotas not prone to burning (e.g., Marlon et al., 2013).

The climate of the Archaic period (8,000 – 2,800 BP) is characterized by both low and high frequency temperature variability. The collapse of the Laurentide ice sheet between 8,500 and 8,200 BP caused a brief cooling of the Northern Hemisphere (Törnqvist and Hijma, 2012), that lasted until 6,000 BP in North America (Viau et al., 2006). A subsequent climatic optimum, dated in North America between 6,000 and 3,200 BP, was accompanied by exponential increases in population density. This population increase continued towards the present despite the onset of cooler and wetter conditions beginning around 3,000 BP (Chaput and Gajewski, 2016; chapter 2). Superimposed on these changes are millennial-scale temperature variations on the order of  $\pm 2^{\circ}\text{C}$  that affected all regions in North America (Viau et al., 2006). During this time, changes occurred in the abundance and location of plant taxa and the location and extent of the major biomes. Following the collapse of the ice sheet, the extent of the tundra and taiga increased to the north, as did the abundance of broad-leaved deciduous species in the southeast (e.g., *Betula*, *Quercus*, *Acer*). Boreal taxa including *Picea* and *Pinus* continued to migrate northward into and across northern Canada in areas previously occupied by ice, and their abundance in the southern portion of their range declined. *Fagus* and *Tsuga* appeared in the New England and Great Lakes regions, and temperate deciduous forests replaced the majority of cool-mixed forests (Williams et al., 2004; Gajewski et al., *subm.*; chapter 5). Species of nut-bearing (e.g., *Castanea*, *Carya*, *Juglans*), mast (e.g., *Quercus*) and fruit-bearing trees (e.g., *Amelanchier* and *Prunus*), and berry-producing shrubs increased in abundance and served as a new source of food for humans (Abrams and Nowacki, 2008).

The Archaic period is also marked by a significant increase in the extent and intensity of human-set fires (Fowler and Konopik, 2007) and a transition from a specialist to a generalist diet (Fagan, 2000). Native Americans began using a more diverse group of foods that included game, fish and a wide variety of plants, seeds, and nuts. They continued to use fire for hunting and clearing purposes, but larger patches of land were now being burned to make room for fields and to encourage the growth of seed, nut and berry-producing species (Fowler and Konopik, 2007). Fluted point technology continued to change, and wooden tools were adopted for digging and grinding (Minnis, 2010). It has been suggested that ongoing population growth coupled with a lack of large mammals capable of being domesticated ultimately led to the shift from a predominantly hunter-gatherer lifestyle to the adoption of agriculture, and this marked the end of the Archaic period (Gupta, 2004).

The employment of agriculture in North America brought about significant changes to the demography and ecology of the entire continent. During the Woodland (2,800 – 1,800 BP) and Mississippian (1,300 – 400 BP) periods, North American cultures in the Midwest and the East focused on building permanent settlements near freshwater sources, and plant cultivation and domestication (Fowler and Konopik, 2007). Slash and burn techniques were used to prepare fields for domesticated species (Ruddiman, 2003), and maize cultivation was established by 2,100 and 2,150 BP in western and eastern North America, respectively (Merrill et al., 2009; Price, 2009). Population sizes continued to increase exponentially due to the new possibilities that arose from staying in one place for extended periods of time (Ruddiman, 2014). Several pollen and lake-level studies from both western and eastern North America indicate that climate became cooler and wetter during the late-Holocene (Shuman et al., 2004; Viau et al., 2006; Shuman et al., 2009a).

When the first European explorers arrived in the late 1400s, at least half of the North American continent was still occupied by hunter-gatherers focused on gathering, fishing and hunting practices rather than full-scale agriculture (possibly due to climatic conditions, topography, cultural preference, or a combination thereof; Lightfoot and Parrish, 2009). Agricultural communities were primarily located in the lower Midwest and the numerous river valleys in the East (Delcourt; 1987; Bellwood, 2005; Fowler and Konopik, 2007; Ruddiman, 2014). The North American population is estimated to have been in the millions at the time of European contact, although numerous estimates have been proposed that vary significantly (see [Table 1.1](#) in section 1.3), and population numbers declined rapidly thereafter. The new colonists brought with them domesticated cattle, horses, turkeys, sheep, goats, pigs and chickens (Bowling, 1942; Speller, 2010). Seeds for fruit trees (e.g., oranges, lemons), vegetables, cereals and sugar cane were also imported along with new weapons, tools and construction materials such as steel and copper. The colonists also brought war, enslavement, and a number of diseases to which indigenous communities had never been exposed. The spread of disease occurred faster than the colonization of the Europeans and over the span of the 16<sup>th</sup> and 17<sup>th</sup> centuries, a possible 50 – 90% of the indigenous population in North America perished (Denevan, 1992).

This demographic collapse coincides with the most dramatic decrease in atmospheric CO<sub>2</sub> and CH<sub>4</sub> in the last 2,000 and 1,000 years, respectively, suggesting a close association between Native American land use and the global carbon cycle during the last several millennia (Stewart, 2002; Ruddiman, 2003; 2014). Ruddiman, and others, suggest that atmospheric greenhouse gas levels decreased in response to the near cessation of human-induced biomass burning and the regeneration of forests following this massive population decline (Ruddiman, 2003; Ferretti et al., 2005; Faust et al., 2006; Abrams and Nowacki, 2008; Nevle et al., 2011; Ruddiman et al., 2011), which may have amplified the global effects of the Little Ice Age (Dull et al., 2010). Further, prior to European contact, it is estimated that Native American fire use could have doubled the number of fires that would have naturally-occurred as a result of lightning strikes (Bonnicksen, 2000). Despite this mounting evidence, many authors still posit that population density would have been too low to have global-scale implications (Ubelaker, 1992; Steckel and Rose, 2005; Liebmann et al., 2016) and that past changes in atmospheric gas levels are not directly associated with Native American land use (Bird et al., 2017).

Although decades old, this debate concerning prehistoric human impacts as a factor in the evolution of ecosystems in North America is ongoing. The following section summarizes the key arguments and points of view that are central to this discussion.

## **1.2 Controversy over prehistoric North American land use impacts**

Views of a “harmonious relationship” between prehistoric North Americans and the environment became popular when Sauer (1956) wrote that, despite obvious land cover changes brought about by activities such as agriculture and deforestation, early North Americans had negligible impacts on the health of ancient ecosystems. This view was supported by population estimates which considered prehistoric North American populations to be quite low (e.g., Mooney, 1928; Denevan, 1992; Ubelaker, 1992). From a spiritual point of view, many authors emphasized the sustainable approach towards flora and fauna and the high regard for the vitality of the Earth exhibited by numerous cultures (e.g., Fertig, 1970; Taylor and Spurlock, 1982). From an ecological perspective, it seemed possible that lower population numbers combined with careful management of the landscape could have meant that Native Americans left very little indication of impact (e.g., Vale, 2002; Munoz et al., 2014), and that climate variability was the primary driver of changes in vegetation structure and composition (e.g., Williams et al., 2004; Bird et al., 2017). These works fostered a view that North America remained “pristine and natural” until post-European times (Denevan, 1992).

In regions where long-term human occupation is well-described, for example in the eastern United States, southern Ontario, and the Great Plains, some authors have concluded that major ecosystem alterations were perhaps only associated with settlements or portions of the landscape that were constantly being used at a high intensity (Vale, 2002). For example, southern Ontario pollen records show significant vegetation changes coincide with an intensification of Native American agriculture, and further changes follows the abandonment of the region (Burden et al., 1986; McAndrews and Boyko-Diakonow, 1989; Clark and Royall, 1995), but these changes

diminish with distance away from the sites (Munoz and Gajewski, 2010). In Tennessee, Chapman et al. (1982) identified changes in pollen and macrofossil assemblages reflecting human selection for specific fruit- and seed-producing taxa, but these changes were restricted to sites within the Little Tennessee River Valley. A review of prehistoric fire-use practices across New England found the effects of human-induced fires to be local, suggesting Holocene forests were predominantly influenced by climate and environmental factors (Russell, 1983). Vale (2002) came to a similar conclusion after studying Native American impacts in the western United States. The overall conclusion from these regional studies using paleoecological records does suggest that human impacts were ecologically important, but it is still unclear if these impacts are measurable at the landscape level. Further, many authors note the difficulty in separating climate-induced vegetation change from human impacts (Gajewski, 1987; 1988; Nordt et al., 1994; Paquette and Gajewski, 2013; Lafontaine-Boyer and Gajewski, 2014).

In contrast, Cronon (1983) argued unequivocally that environmentalists and historians were making a grave mistake by dismissing the magnitude of prehistoric Native American impacts and defended the importance of using paleoecological and historical data to study past environmental change. A number of studies from all across North America have since made similar arguments and combined paleoecological, archaeological and ethnohistoric data to study the implications of land use change (Orme, 1981; Delcourt, 1987; Delcourt and Delcourt, 1987; Whitehead and Sheehan, 1985; Silver, 1990; Doolittle, 1992; Kay, 1994; Krech, 2000; Harkin and Lewis, 2007; Abrams and Nowacki, 2008; Munoz et al., 2010, Courtwright, 2011). Central to these arguments are elevated estimates of population density prior to European settlement (e.g., Denevan, 1992; Warrick, 2008; Peros et al., 2010). This body of research ultimately suggests that the North American landscape was a cultural or humanized landscape millennia before CE 1492 (Zelinsky, 1973; Butzer, 1992; Butzer, 1996; Denevan, 2011).

The studies mentioned above make up only a small subset of the literature in the field of human-environment interactions. However, they illustrate that the debate is ongoing, in part due to lack of data needed to resolve the issue – something that is mentioned by almost all researchers in the field. Political and social interpretations and belief systems also present challenges (Vale, 2002; Harkin and Lewis, 2007), as do disciplinary differences: social scientists and historians tend to argue in favour of a more disturbed landscape, whereas ecologists tend to find evidence for a more natural, less-impacted system.

Two questions that make up an important part of this debate, but that are rarely discussed in the context of these three points of view (i.e., North America as a pristine landscape, North America as a cultural landscape, or some intermediate between these two extremes) and thus deserve some mention here, are (1) how has the rate of quantifiable human-induced environmental change varied over the course of the Holocene, and (2) are all human impacts inherently negative? With respect to the first question, current projections suggest many human environmental impacts will be two to three times stronger within the next 50 years (Tilman et al., 2001), thus it is a plausible assumption that improved adaptation abilities coupled with increasing population numbers would lead to more intensive human land use in prehistoric times

as well. In fact, Boserup (1965) and others (e.g., Ruddiman and Ellis, 2009) have integrated evidence from field studies and modelled estimates and proposed mechanisms to explain land use intensification with increasing population (and a coupled decline in per-capita land use). However, the overall lack of quantifiable estimates of human-induced land cover change (and data on past populations) has hindered attempts to measure variations in rates of change through time. With regards to the second question, opinions differ widely on what constitutes a negative impact, when the first negative human impact occurred, and the potential (or usefulness) of differentiating between early human impacts (which are often assumed to have a neutral effect on the environment) and modern human impacts (which tend to be labelled as entirely negative). Many studies mention the Pleistocene extinctions as the first negative human impacts in the Americas (e.g., Ripple and Van Valkenburgh, 2010), however more recent research initiatives have focused on the possibility of human impacts being both positive and negative regardless of geological time period (e.g., land use strategies which boost production and reproduction within an ecosystem but also simplify food webs and homogenize landscapes; see Western, 2001). More work (and more data) is needed here as well, particularly regarding the consequences of human impact on, for example, energy pathways, nutrient cycles, productivity, albedo and large-scale processes governing biological and biogeochemical cycles (Western, 2001).

Peacock (1998) presented an argument that is still highly relevant to overall debate regarding human impacts in North America:

*“It is no longer acceptable for scientists to assume that prehistoric ecosystems were ‘pristine’, nor is it acceptable to assume that, if humans were present in the past, they necessarily brought about significant environmental alterations. It is not necessary to assume anything; rather, it is essential that existing archaeological and historical data be incorporated into environmental studies, especially those that seek to characterise the long-term development of biophysical landscapes. If archaeological and paleoecological data do not exist for a particular area, those data can be retrieved through integrated research programs designed to explore the nature of human impact over the long-term.”*

At the time of this statement, quantitative estimates of human population density over the Holocene for North America were lacking. However, in recent years, these data have become available (Peros et al., 2010; Gajewski et al., 2011; Martindale et al., 2016; Canadian Archaeological Radiocarbon Database). Although they provide only relative numbers, these data can be used to reconstruct paleodemographic change in North America for the past 13,000 years (Peros et al., 2008), and make maps of regional population density (Chaput et al., 2015; chapter 3). Further, fossil pollen data from sediment cores spanning the continent have recently been compiled in the Neotoma Paleocology Database, which can be used as a proxy for past vegetation abundance (e.g., Williams et al., 2004) enabling the space-time comparison of changes in population and vegetation for the whole continent. The following sections review these data sources and evaluate their role in the field of human-environment research.

### 1.3 Population data

The paleodemographic history of North America has been inferred from ethnographic, archaeological, historical and statistical (model) sources, and estimates of pre-European population size range between 0.5 and 18 million. **Table 1.1** summarizes these estimates:

**Table 1.1** Summary of prehistoric population estimates for North America prior to CE 1491.

<b>Reference (Area north of Mexico only)</b>	<b>Population (000)</b>
Mooney, 1910	1,150
Rivet, 1924	1,148
Sapper, 1924	2,000 – 3,500
Mooney, 1928	1,153
Wilcox, 1931	1,002
Kroeber, 1939	900
Rosenblat, 1945	1,000
Steward, 1945	1,000
Ashburn, 1947	2,000-2,500
Steward, 1949	1,000
Aschmann, 1959	2,240
Driver, 1961	1,000-2,000
Dobyns, 1966	9,800-12,250
Denevan, 1976	4,400
Ubelaker, 1976a	2,171
Thornton and Marsh-Thornton, 1981	7,705
Dobyns and Swagerty, 1983	18,000
Thorton, 1987	7,000
Denevan, 1992	3,790
Ubelaker, 1992	1,900
Sioui, 1999	18,000
Klein Goldewijk et al., 2010	540
Peros et al., 2010	1,200 - 2,600
Ruddiman, 2014	7,000

The discrepancy between estimates has obvious implications concerning the interpretation of the environmental history of the North American continent, including the size of the founding population, population growth rates, the carrying capacity of prehistoric forests, the intensity of human land use and the magnitude of indigenous population loss post-contact. Issues also arise when population estimates are required for global models of environmental change which require population as an input. This includes models which reconstruct past environmental change (e.g., the History Database of the Global Environment model – HYDE; Klein Goldewijk et al., 2011) and models which forecast future scenarios (e.g., the Agriculture and Land Use model – AgLU; Sands and Leimbach, 2003).

Many studies are now focusing on spatio-temporal estimates of environmental change since these provide new insight into population-vegetation dynamics compared to independent analysis of time series or spatial data (Grove, 2011). Recently, spatio-temporal estimates of paleopopulation change have become available (Chaput et al., 2015; chapter 3) which can be used to model paleodemographic growth following human migration into the Americas for the past 13,000 years. These estimates are the first step toward being able to understand continental-scale human impacts on the North American ecosystem during the Holocene as well as demographic growth and migrations in relation to environmental changes.

These estimates are based on radiocarbon ( $^{14}\text{C}$ ) data from the Canadian Archaeological Radiocarbon Database (CARD; Peros et al., 2010; Gajewski et al., 2011; Martindale et al., 2016) and follow the assumption that variations in the number and distribution of dates reflect fluctuations in the relative size and location of past populations. The CARD (now in version 2.1) is an openly-accessible global database which contains over 50,000 uncalibrated  $^{14}\text{C}$  dates for North America (92,000+ dates globally). The CARD was originally created by R. Morlan of the Canadian Museum of History (formerly the Canadian Museum of Civilization) and is now being updated as part of a partnership between the Canadian Museum of History and the Laboratory of Archaeology at the University of British Columbia ([www.canadianarchaeology.ca](http://www.canadianarchaeology.ca)). The vast majority of dates within the CARD are georeferenced and culturally-associated, and many are described in detail (e.g., geographic context, site type, material dated). Dates originate from cultural and paleoenvironmental materials and range from 0 to beyond 40,000 calibrated years BP, although most are from the last 15,000 years.

The use of  $^{14}\text{C}$  data as a proxy for paleodemographic change, first discussed by Wendland in the 1970s and Butzer in the 1980s, and labelled as the “dates-as-data” method by Rick (1987), has proliferated over the last two decades. Numerous studies across the globe have documented the correlation between  $^{14}\text{C}$  data, population fluctuations, and environmental and climate change. For example, Timpson et al. (2014) used observed and simulated summed probability distributions (SPDs) of  $^{14}\text{C}$  dates from twelve regional datasets across Europe. Eight of the twelve datasets indicated population booms immediately following the arrival of agriculture, and all datasets indicated population busts were not climatically-driven. French and Collins (2015) studied  $^{14}\text{C}$  dates and archaeological site counts from southern France and provided evidence that the region was used as a population refugium during the Last Glacial Maximum (LGM). There are further examples from Europe (Gamble et al., 2005; Shennan and Edinborough, 2007), Siberia (Kuzmin and Keates, 2005; Fiedel and Kuzmin, 2007), Ireland (Turney et al., 2006a), northeast Spain (González Sampériz et al., 2009), and Australia (Holdaway and Porch; 1995, Smith and Ross, 2008; Turney and Hobbs, 2006; Williams et al., 2008b, 2010). The increased use of the dates-as-data method, however, has brought to light the limitations and biases inherent in  $^{14}\text{C}$  data, and these are reviewed in detail in chapter 2.

A small number of studies have made use of North American  $^{14}\text{C}$  data from the CARD. In southern Ontario, Munoz and Gajewski (2010) related the density of archaeological sites to nearby pollen records and presented a late-Holocene record of human-induced forest clearance based on the replacement of cool-adapted tree species with herbaceous plants, *Quercus* and

*Pinus*. One of the first regional applications of the CARD showed distinct transitions in forest composition closely related to paleoclimatic fluctuations and prehistoric human history (Fig. 2 in Munoz et al., 2010). The evolution of human population size across North America was further explored by Peros et al. (2010) where the frequency of radiocarbon dates was used as an indicator of population size. In this interpretation, the population level for North America remained low prior to 6,000 BP after which population size increased dramatically. However, when a taphonomic correction was applied, total population was estimated to be much smaller in magnitude and the increase only began after 2,000 BP. Buchanan et al. (2008) used  $^{14}\text{C}$  data to test the “Extraterrestrial Impact Hypothesis” (Firestone et al., 2007) but did not find any correlation between Paleoindian population decline and the suggested timing of the potential impact. They advised caution when referring to a potential late Pleistocene extraterrestrial impact north of the Great Lakes. Steele (2010) used  $^{14}\text{C}$  dates to assess Clovis-age population-front migration speeds across southern North America, and Surovell and Brantingham (2007) and Surovell et al. (2009) used  $^{14}\text{C}$  dates from the United States to develop models for the correction of taphonomic bias. Lastly, North American  $^{14}\text{C}$  data have been studied to understand the association between deglaciation, changes in sea ice extent, variations in sea level, and the paleodemography of Arctic communities (e.g., Dyke, 2004; Dyke and Savelle; 2009; Savelle and Dyke; 2014).

Although  $^{14}\text{C}$  data do not provide absolute population numbers, statistical analyses of large collections of radiocarbon dates can be used as a demographic proxy for prehistoric human population change worldwide (Zahid et al., 2016). This quantitative approach is valuable to the fields of human-environment research and Earth-system modeling, and when combined with environmental reconstructions based on paleoecological data, associations between humans and their environments can be quantified.

#### **1.4 Vegetation data**

The greatest changes to prehistoric North American ecosystems brought about by humans were made with fire (Boyd, 1999; Fowler and Konopik, 2007), and the most robust method of observing these changes is the analysis and interpretation of changes in fossil pollen assemblages (Birks and Birks, 1980; Seppä and Bennett, 2003; Birks and Berglund, 2018). The development and analysis of pollen databases for mapping and paleoclimate reconstructions began in the 1970s with the work of T. Webb III as part of the COHMAP project (Cooperative Holocene Mapping Project; Bernabo and Webb III, 1977; Wright Jr and Bartlein, 1993; COHMAP members, 1988). Webb, along with many students and colleagues including P. Bartlein and I.C. Prentice, studied the relationship between climate variables and the abundance of different taxa by mapping geographical plant abundance in climate space (response surface analysis; Bartlein et al., 1986; Prentice et al., 1991). Their projects provided information about the individualistic behaviour of different tree taxa in time and space in response to Holocene climate changes and served as a starting point for more complex biogeographical studies. Webb made significant contributions to the field of vegetation dynamics, having tested pollen-climate response surfaces

against climate models (Webb et al., 1998), and furthering the understanding of the links between biomes, individual taxa and climate variables.

Similarly, H. and P. Delcourt completed studies integrating paleoecological and paleoethnobotanical evidence of prehistoric human impact on the structure of North American forests (e.g., Delcourt and Delcourt, 1997). Their work revealed patterns in historical agricultural transformation and provided information about the increase in abundance of certain species of trees in specific habitats. They also studied the use of fire by Native Americans in several case studies in the United States (e.g., Delcourt, 1987; Delcourt and Delcourt, 1998) as part of the Quaternary history of the U.S., and species migration patterns and palynological records of areas such as the Mississippi Embayment in the south-central United States (Delcourt et al., 1980).

Although this work provided information about broad scale patterns, it was limited by the use of pollen percentages as evidence of past vegetation. Ultimately, the non-linear relationship between pollen deposition and vegetation abundance, and taxon-specific differences in pollen productivity, limit the ability of percentage data to provide quantitative estimates of past vegetation (Davis, 1963; Sugita, 1994; Sugita et al., 1999). Davis (1963) first addressed this by developing an “R-Value” model as a means of estimating the abundance of vegetation on the landscape quantitatively. Andersen (1970), Prentice and Parsons (1981), and Parsons et al. (1983) identified statistical issues associated with the use of the R-value model and made improvements by correcting for the non-linearity of pollen percentage data, resulting in the Extended R-Value (ERV) model. Building on this, Sugita (2007a; 2007b) developed the Landscape Reconstruction Algorithm (LRA), a two-step framework developed for the quantitative reconstruction of vegetation, which overcomes some of the fundamental problems associated with the R-Value and ERV models. This method uses the ERV model to estimate the abundance of vegetation on the landscape, and accounts for taxon-specific pollen production, pollen dispersal, and pollen source area, ultimately providing a more robust representation of past vegetation cover when compared to pollen percentage studies (Åkesson et al., 2015; Chaput and Gajewski, 2018; chapter 6).

The data required for the analysis of land cover change, and the study of past human-environment relationships in general, comes from paleoenvironmental and paleolimnological studies of fossil pollen originating from lake sediments. In North America, the paleoenvironmental data collected as part of the COHMAP project was eventually incorporated into the North American Pollen Database (Grimm, 2000) and the North American Plant and Macrofossil Dataset (Jackson et al., 1997) which evolved in the comprehensive Neotoma Paleoecology Database ([www.neotomadb.org](http://www.neotomadb.org)). Major outcomes of this collective research include the creation of isopoll maps showing vegetation composition changes through time at different scales (Jackson et al., 2000; Williams et al., 2004), such that predictions about vegetation patterns that would exist in human-free scenarios could be made. At present, the Neotoma Paleoecology Database is the largest paleoecological database available, containing over 1,500 fossil and 2,500 modern pollen records (Whitmore et al., 2005) from thousands of sites spanning North America (Neotoma, 2017), as well as other paleoenvironmental data. The

North American Pollen Database is now a constituent of the Neotoma Database, and new datasets are being uploaded regularly.

## **1.5 Quantifying human-environment interactions: integrating archaeology and paleoecology**

An important unanswered question in the field of human-environment research in North America is the extent of the impact of prehistoric land use activities on the structure, functioning and health of the various ecosystems across all major regions. Corollaries to this question are the relative contribution of Native American land use activities to past global climate change and the role of Native Americans in driving the onset of the early Anthropocene (*sensu* Ruddiman, 2003). One approach towards answering these questions is the integration of archaeological and paleoecological data needed to assess past human-environment relationships.

Local-scale comparisons of archaeological and paleoecological data exist for many North American regions. For example, Lacourse et al. (2007) showed that localized declines in ecologically-important tree stands in the Pacific Northwest were coincident with the arrival of the Haida and the establishment of small communities along the coast. Their analysis highlighted the limited extent of forest clearance relative to other North American regions known to have supported larger populations, but showed forests were disturbed at large distances away from settlements despite populations being small (see Acheson, 1998). Studies of other island populations have reached similar conclusions (e.g., McKechnie, 2007). Boreal North America has recently been described as a cultural landscape, despite previously being viewed as a natural wilderness, where the intensity of human-induced vegetation change is positively correlated with population size and resource demand across the biome (Johnson and Miyanishi, 2012; Danneyrolles et al., 2016). An analysis of pollen and archaeological data from the Cahokia region showed strong correspondence between periods of agricultural expansion (contraction) and increases (decreases) in the abundance of pollen from non-arboreal cultigens characteristic of the Eastern Agricultural Complex (EAC; Munoz et al., 2014). Forest clearance was detected well before the peak in Cahokia civilization.

In southern Ontario, studies of Crawford Lake describe eutrophication and changes in surrounding vegetation coincident with Native American burning and agricultural practices (e.g., Ekdahl et al., 2004). Nearby studies show that pollen records influenced by prehistoric human land use activities can be distinguished from those that were primarily influenced by climate, and that human impacts were most often recorded as a decrease in forest cover and an increase in herbs and other disturbance indicators (Munoz and Gajewski, 2010). The association between forest clearance and human occupation is prevalent in literature from the Great Basin (Kelly, 1997), the Midwest (McLauchlan, 2003; Scharf, 2010) and eastern North America (Delcourt et al., 1986; Abrams and Nowacki, 2015).

Based on these and other case studies, the emerging consensus among scientists working within and between the fields of archaeology and ecology is that an interdisciplinary approach

combining archaeological and paleoecological data important (Peacock, 1998; Kirch, 2005). For example, records of changes in both pollen and  $^{14}\text{C}$  will be necessary in order to resolve the controversy over prehistoric Native American land use impacts and determine whether North America is indeed a cultural landscape (Denevan, 2011). If Native Americans actively (or passively) promoted the growth of mast and fruit trees (e.g., *Carya*, *Juglans*), then the abundance of these trees should increase during periods of high population density across all regions where these taxa are dominant, and this should be reflected in the pollen records from these regions (Dorney, 1981; Wycoff, 1991; MacDougall, 2003; Smith, 2007). Alternatively, if changing climatic conditions periodically led to increases in taxa most commonly used for construction, crafting, medicinal and ceremonial purposes, then this could potentially lead to population spikes and noticeably higher abundances of archaeological remains *in situ*. If extensive burning and forest clearance was practiced in all regions conducive to sustainable agriculture, this should appear as an increase in disturbance taxa such as *Populus* or *Poaceae* or a decrease in late successional taxa such as *Acer* or *Fagus* in regions adjacent to known human occupation.

Related to the need for interdisciplinary approaches is the need for spatiotemporal analyses which combine spatial and temporal methods. Grove (2011) developed a methodology that combined  $^{14}\text{C}$  dates from the Mesolithic (point data) with kernel density estimation techniques and successfully reconstructed changes in human land use patterns across several hundred kilometers in southwestern Europe. Likewise, Trondman et al. (2015) developed a methodology using pollen records from over 600 lakes and bogs and a model of quantitative vegetation reconstruction to map anthropogenic deforestation between 6,000 BP and the present across most of Europe. However, research has yet to be done which combines these two approaches into a single, unified framework. The development of these methods, and the recent availability of North American databases such as the CARD and Neotoma, provides the opportunity to perform this type of research and is an important motivation behind this thesis.

Lastly, there is a need for archaeological and paleoecological datasets which could be used to improve current climate and environmental models of past changes and future scenarios (Pongrantz et al., 2008; Goldewijk et al., 2011; Kaplan et al., 2010). Archaeological data can provide finer-scale information about changes in population density relative to hindcasting methods based on census data (Chaput and Gajewski, 2016; chapter 2), and quantitative reconstructions of past vegetation cover are more informative than the pollen percentage data most often used for modelling purposes (Åkesson et al., 2015; Trondman et al., 2015; Chaput and Gajewski, 2018; chapter 6).

## **1.6 The Early Anthropocene Hypothesis**

Crutzen and Stoermer (2000) proposed the term “Anthropocene” for the current geological epoch during which human activity has dominated global ecosystems and processes (Vitousek et al., 1997; Crutzen, 2002; Sanderson et al., 2002). Although the formal designation of the Anthropocene Epoch has yet to be approved by the International Commission on Stratigraphy, the Anthropocene Working Group has submitted a proposal to the Commission suggesting the Anthropocene officially began in the mid-20<sup>th</sup> century. The majority opinion within the Working

Group is that a plausible global marker would be the stratigraphic layer of radionuclides that is observed in the soils of many regions due to nuclear weapons testing (Zalasiewicz et al., 2017).

An alternative designation of the Anthropocene Epoch follows from the Early Anthropocene Hypothesis. Ruddiman (2003; 2005a) suggested significant anthropogenic environmental change has been occurring since well before the mid-20<sup>th</sup> century. This three-part hypothesis draws attention to the likelihood that indigenous peoples in the Eastern Hemisphere began 1) altering atmospheric CO<sub>2</sub> concentrations approximately 7,000 years ago through early agricultural practices and forest clearing, 2) altering atmospheric CH<sub>4</sub> concentrations approximately 5,000 years ago after the discovery of rice irrigation, and 3) reversing the natural cooling trend that was occurring, ultimately averting the next glaciation due to an increase in unsustainable environmental pressures. Support for this hypothesis comes from Greenland CH<sub>4</sub> and Antarctic CO<sub>2</sub> ice core records which show atmospheric greenhouse gas concentrations increasing since the mid-Holocene, a trend opposite to that of previous interglacials.

Numerous suggestions have been put forward regarding the potential timing of the onset of the global Anthropocene and the most plausible stratigraphic evidence marking the dawning of irreversible, global anthropogenic impacts. **Table 1.2** (modified from Lewis and Maslin, 2015) summarizes some of these suggestions.

Consensus cannot be reached regarding the most reasonable geological time interval for the Anthropocene due to the lack of a globally synchronous signal (“golden spike”) reminiscent of that marking the transition from the Precambrian to the Cambrian Period or the Silurian-Devonian boundary, for example. This has caused some experts to question the overall practicality of officially ending the Holocene in favour of the Anthropocene. Some have proposed that the use of the term Anthropocene should remain informal, and that making a decision now would be premature and imprudent (Ellis, 2017), while others believe a true global impact has yet to come (Wolff, 2014). Others still have focused on the redundancy of having an epoch dedicated solely to humans modifying the global landscape since the Holocene marks a point at which humans occupied all continents and environments (Walker et al., 2012).

Advocates of the proposal made by the Anthropocene Working Group have theorized that global populations would have been much too small to have any traceable effect on the biosphere prior to the 20<sup>th</sup> century, and that the rise in CH<sub>4</sub> in ice core records in the late Holocene can be explained by changes in the Earth’s orbital pattern and insolation changes in the tropics (recreated in climate simulations; Singarayer et al., 2011). Further, there is a lack of variation in Antarctic ice core  $\delta^{13}\text{C}$  records, an unexpected result if vast expanses of forest were indeed cleared and burned and the resulting CO<sub>2</sub> released into atmosphere (since plants preferentially take up  $\delta^{13}\text{C}$  and such activities would likely cause a change in the  $\delta^{12}\text{C}/\delta^{13}\text{C}$  ratio; Elsig et al., 2009).

Conversely, those who support the Early Anthropocene Hypothesis claim the dismissal of the magnitude of prehistoric Native American impacts sets the science back to its pre-twentieth century origins. The cutting and burning of biomass and farming activities are known to release various compounds into the ground and greenhouse gases into the atmosphere (Smith et al.,

2008). It is therefore possible that past forest-clearing and agriculture practiced by Native Americans contributed to changes in global CO<sub>2</sub> and CH<sub>4</sub> budgets, if the populations were large enough and activity levels sufficiently intense. Although the timing and intensity of the global CO<sub>2</sub> and CH<sub>4</sub> increases discussed above may not be directly linked to the demographic history of the North American continent, Ruddiman (2014) advocates that, due to the extensive burning that took place, the “*growing populations of early Americans would have contributed substantially to the observed rise in CO<sub>2</sub> during the last several millennia*” and the rise in CH<sub>4</sub> during the last 5,000 years.

**Table 1.2** Potential start dates and evidence for the Anthropocene Epoch

<b>Event</b>	<b>Date</b>	<b>Geographical extent</b>	<b>Stratigraphic evidence</b>
Late Pleistocene megafauna extinction	50,000 – 10,000 BP	Most continents, global impact	Fossils
Global-scale agriculture	3,000 – 500 BP	Local impacts, observed globally	Fossil pollen, soil composition
New-Old world collision	CE 1492	Europe-Americas, global impact	Decrease in ice core CO <sub>2</sub>
Industrial Revolution	CE 1760	Europe-America, global impact	Charcoal
Nuclear weapons testing	CE 1945	United States, global impact	Radionuclides in tree rings
The Great Acceleration	CE 1950	Local events, global impact	Biodiversity loss, atmospheric CO <sub>2</sub> and CH <sub>4</sub> , manufactured waste, dam construction, population

At present, the Anthropocene remains an informal designation, and it will remain so until the controversy surrounding this proposal is resolved and irrefutable geological evidence is found supporting a true transition between epochs. To this end, quantitative estimates of population and vegetation change are needed across all regions of the world. A summary of how this thesis seeks to contribute to this endeavor, at least for North America, is provided in the following section.

## **1.7 Summary of research questions and methods**

The assumption that prehistoric Native American land use practices had little impact on the North American landscape persists in the literature. However, research has shown that Native Americans did in fact have a considerable impact on lands and ecosystems through their daily activities including burning, deforestation, agriculture, silviculture and irrigation. What remains unknown is the extent of this impact. To resolve this, quantitative estimates of changes in human population size and forest structure and composition in both time and space are needed. This thesis seeks to address this need by providing radiocarbon-based paleodemographic reconstructions and pollen-inferred estimates of vegetation change, as well as analyses of

correlations between the two at both continental- and regional-scales, from the late Pleistocene to the Anthropocene.

*1.7.1 Research Question #1: Are archaeological radiocarbon ( $^{14}\text{C}$ ) dates a robust proxy for paleodemographic change, and is the North American archaeological record comprehensive enough to reconstruct relative population changes from 13,000 years BP to the present?*

Studies of archaeological  $^{14}\text{C}$  dates have shown that they can be used to interpret regional (e.g., Rick, 1987; Gamble et al., 2005; Shennan and Edinborough, 2007; Munoz et al., 2010) and continental (Peros et al., 2010; Williams, 2010) changes in past population density. This method is questioned, however, based on uncertainties and biases associated with archaeological data (Steele, 2010). A literature review was done to assess the state of the art of using  $^{14}\text{C}$  ages of culturally-significant artifacts as a proxy for paleodemographic change (**chapter two**). To test the completeness of the North American archaeological record, time series methods used in Peros et al. (2010) were used with an updated version of the Canadian Archaeological Radiocarbon Database (CARD) that contains data for Canada and the United States (Gajewski et al., 2011). Temporal frequency distributions were used to model paleopopulations, and taphonomic corrections were applied to account for the loss of organic material in the distant past (Surovell et al., 2009). Time series were presented for other areas of the globe and compared to model-based estimates of population change.

*1.7.2 Research Question #2: How has the spatiotemporal distribution of the North American population changed between 13,000 years BP and the present?*

Spatiotemporal modeling methods based on kernel density estimation have previously been applied to archaeological  $^{14}\text{C}$  data in Europe (Collard et al., 2010; Grove, 2011). However, the databases used in these studies were small - a limitation which can potentially amplify the effects of dating errors and spatially and temporally discontinuous datasets. **Chapter three** provides the first spatiotemporal estimates of population density for North America in a series of continental-scale maps which take into account the distribution and frequency of radiocarbon dates at 500-year intervals. The maps also accounted for taphonomic bias (Surovell et al., 2009), edge effects, sampling intensity and  $^{14}\text{C}$  calibration.

*1.7.3 Research question #3: Are prehistoric human impacts on North American vegetation abundance discernable at continental-scales?*

Numerous studies have shown that human impacts on prehistoric vegetation are measurable at local- to regional-scales and discernable in pollen assemblages (e.g., Chapman, 1982; Denevan, 1992; Clark and Royall, 1995). Some studies indicate, however, that this relationship becomes less defined as scale increases (Russell, 1983; Vale, 2002; Munoz and Gajewski, 2014). Continental-scale associations between population and vegetation change have never before been estimated, and this is now possible in North America with the development of the CARD and the Neotoma databases. **Chapter four** introduces a new methodology for cross-correlating archaeological radiocarbon data and pollen-based vegetation data at a continental scale using temporal smoothing techniques and kernel density estimation. This methodology was also used

to test lagged cross-correlations between human population numbers and *Quercus* abundance since human-induced land cover changes may not be immediately recorded in the pollen record.

In **chapter five**, the methodology from **chapter four** was used in an attempt to detect continental-scale associations between population change and nine additional taxa that have large ranges and coincide with human occupation. Taxa with widespread ranges (*Picea*, *Pinus*), characteristic of the eastern deciduous forest (*Quercus*, *Castanea*, *Acer*, *Fagus*), with seed-producing abilities (*Carya*, *Juglans*), and disturbance indicators (*Populus*, *Poaceae*) were included.

#### *1.7.4 Research question #4: Are prehistoric human impacts on North American vegetation abundance discernable at regional-scales?*

The pollen data used in **chapters four** and **five** provide estimates of changes in vegetation composition through time. As such, absolute changes in vegetation cover or land use (e.g., landscape openness) cannot be examined. The REVEALS (Regional Estimates of VEgetation Abundance from Large Sites) model was developed for this purpose and enables the quantification of past changes in plant cover (Sugita, 2007a). This model has been applied extensively in Europe (Hellman et al., 2008; Broström et al., 2008; Soepboer et al., 2010; Trondman et al., 2016; Abraham et al., 2017) and REVEALS reconstructions have been shown to produce more accurate vegetation histories (Åkesson et al., 2015), but it has yet to be tested to the full extent in North America.

This model requires taxon-specific pollen productivity estimates for individual study areas, as well as modern vegetation information. Productivity estimates for North American taxa are currently rare and regional vegetation surveys are difficult to obtain. In **chapter six** a vegetation survey from southeastern Quebec was used to calculate productivity estimates for eight tree taxa, and the REVEALS model was used to estimate regional changes in vegetation cover over the course of the Holocene.

The results of this case study demonstrate the importance of using pollen counts rather than percentages, in association with quantitative modelling methods, for the interpretation of past vegetation change. However, it was not possible to fully answer research question #4 since the study was limited by the availability of pollen samples and consequently the number of taxa analyzed. Ultimately, the REVEALS model improved our understanding of past changes in the relative abundance of different taxa by accounting for pollen over- and under-representation, and produced results that better predicted the evolution of plant communities following the northward retreat of the Laurentide ice sheet.

## **1.8 Organization of the thesis and author contributions**

The thesis is written in article format. **Chapter one** introduces North America as a landscape with a long and complex history of human influence. The ongoing discussion over the nature of prehistoric human-environment interactions is summarized, followed by a review of existing population estimates and vegetation data from the most commonly used sources. The potential

insight that can be gained from integrating paleoenvironmental and archaeological data is reviewed and assessed in the context of the Early Anthropocene Hypothesis.

**Chapter two** is a review of the state-of-the-art practices in using radiocarbon ages of culturally-significant artifacts as a proxy for paleodemographic change. This chapter describes how existing datasets from around the world were accumulated for the first time. It introduces an ongoing initiative to create a single, openly-accessible online database, and highlights the strengths and shortcomings of using radiocarbon dates as a proxy for paleodemographic change. An analysis of comprehensive datasets from North America and Australia highlights the insight gained from this method over alternative population models based on historical estimates and hindcasting methods since high frequency fluctuations in population density are easily discerned.

- **Chapter two** authors: Michelle Chaput and Konrad Gajewski
- The idea for this project came from KG who has previously worked with the CARD. Data collection and preparation, numerical data analysis, GIS analysis, and interpretation of results was completed by MC. MC wrote the paper with contributions from KG.
- This paper has been published in *Anthropocene* (15: 3-12. doi.org/10.1016/j.ancene.2015.10.002).

**Chapter three** provides the first spatiotemporal estimates of population density for North America in a series of continental-scale maps which take into account the distribution and frequency of radiocarbon dates at 500-year intervals. Taphonomic bias and data quality are also addressed. Population density is obtained using kernel density estimation techniques that provide a smoothed, mapped surface. The maps illustrate the evolution of population size, migration patterns and the boom and bust patterns of several population centres. The dates-as-data method does not provide absolute population numbers but does provide relative population densities which are consistent with independent archaeological and ethnohistoric evidence.

- **Chapter three** authors: Michelle Chaput, Bjoern Kriesche, Matthew Betts, Andrew Martindale, Rafal Kulik, Volker Schmidt, Konrad Gajewski
- KG, BK and MC came up with and developed the idea for this research, which is based on the dataset developed by MC for chapter two. BK performed the statistical modeling and MC performed the GIS analysis. MC developed and wrote the paper with contributions from all coauthors. VS, BK and RK verified the statistical results and AM and MB provided archaeological interpretation.
- This paper has been published in the *Proceedings of the National Academy of Sciences USA* (112(39): 12127-12132. doi:10.1073/pnas.1505657112).

The availability of estimates of paleopopulations based on chapter three, coupled with the abundance and availability of digitized fossil pollen records which also span the continent provided the motivation for a continental-scale analysis of the association between ancient populations and vegetation. A methodology was developed (**chapter four**) which centers on the statistical evaluation of cross-correlations between population and plant abundance. Using *Quercus*, an ecologically important taxon in the eastern deciduous forest, as an example,

population changes interpreted from the maps in chapter three are cross-correlated with changes in *Quercus* abundance and range during the last 13,000 years.

- **Chapter four** authors: Bjoern Kriesche, Michelle Chaput, Rafal Kulik, Konrad Gajewski, Volker Schmidt
- MC and KG developed the idea for additional archaeological analyses using the output from chapter three and a statistical methodology developed by BK. Interpretation was based on an exchange of ecological (MC, KG) and statistical (BK, RK, VS) expertise and the manuscript was written by BK, MC and KG.
- This paper is under review with *Geographical Analysis*.

**Chapter five** builds on this methodology by including nine additional taxa from across the continent and analyzing in detail the relation between Holocene populations and plant distribution and abundance. Taxa with widespread ranges (*Picea*, *Pinus*), characteristic of the eastern deciduous forest (*Quercus*, *Castanea*, *Acer*, *Fagus*), with seed-producing abilities (*Carya*, *Juglans*), and disturbance indicators (*Populus*, *Poaceae*) are included. Periods of high spatial cross-correlation (positive and negative) between population and plant abundance are seen, but these associations were irregular and did not improve over time. The results suggest, in most cases, that ancient human impacts are not detectable at a continental scale, either due to low populations or varying human land use practices, implying ancient human-environment relationships were regional rather than continental.

- **Chapter five** authors: Konrad Gajewski, Bjoern Kriesche, Michelle Chaput, Rafal Kulik, Volker Schmidt
- The idea for the paper came from KG, MC and BK to investigate the fundamental questions raised in the introduction of the thesis. BK performed all statistical analyses, with help from VS and RK, MC compiled results and figures, and KG provided ecological expertise. KG wrote the paper, with contributions from MC, BK, VS and RK.
- This paper has been submitted to the *Journal of Archaeological Sciences*.

To test this hypothesis and further examine the relationship between pollen data and human land use at a regional scale, an additional study was done (**chapter six**) using a complementary methodology. The REVEALS (Regional Estimates of VEgetation Abundance from Large Sites) model is used to obtain quantitative estimates of plant cover and provide an indication of the ecological evolution of forests. This model corrects for the non-linear relationship between pollen production and plant abundance and is currently being used internationally by paleoecologists to map global histories of land use and land cover change. In chapter six, REVEALS is applied for the first time in Canada to pollen records from lake sediments in the deciduous forest of southeastern Quebec. The results reflect the tendency for traditional pollen percentage studies to over- or underestimate the presence of certain taxa in assemblages and provide a more accurate representation of the landscape to which population data can be compared. A limitation of this method is the lack of pollen productivity estimates for the different taxa characteristic of this forest, therefore eight major taxa were selected for this study.

In Chapter seven, a preliminary comparison with CARD data shows how these results may improve our understanding of human impacts on the landscape.

- **Chapter six** authors: Michelle Chaput and Konrad Gajewski
- KG met with Shinya Sugita (who developed the REVEALS model) and discussed using the model in North America. MC developed the plan for this paper after attending a workshop hosted by Sugita and meeting with members of the Past Global Changes Landcover6k Working Group at an international conference. Pierre Richard provided the modern vegetation data. MC performed all data preparation, analysis and interpretation. MC wrote the paper with contributions from KG.
- The paper has been published in *Botany* 96(5): 299-317. doi.org/10.1139/cjb-2017-0193.

**Chapter seven** summarizes the research and results that make up the contents of the thesis. Preliminary results are presented from a comparison between the vegetation reconstruction obtained in chapter six and a population model for the region. Overall conclusions and future work are also discussed.

## 2 Radiocarbon Dates as Estimates of Ancient Human Population Size

Michelle Chaput and Konrad Gajewski

### 2.1 Abstract

Archaeological radiocarbon ( $^{14}\text{C}$ ) dates are a fundamental source of information documenting patterns in paleodemographic change from prehistoric to modern times. Several open access databases (Canadian Archaeological Radiocarbon Database, Radiocarbon Palaeolithic Europe Database, CONTEXT, RAdiocarbon Dates ONline, and AustArch) and publications which include lists of dates provide easy access to archaeological  $^{14}\text{C}$  data, presently totalling over 70,000 dates worldwide. Some parts of the world are more extensively sampled than others including North America, Australia and China, whereas in others the databases have not yet been prepared. A comparison of frequency distributions of  $^{14}\text{C}$  dates from North America and Australia to modelled estimates of historical population growth for these continents from the HYDE 3.1 database shows similarities, providing confidence in long-term estimates of population growth using both methods. Our capacity to study global demographic change is currently limited by the spatiotemporal completeness of regional  $^{14}\text{C}$  databases. These results suggest the systematic collection and entry of dates into an openly-accessible, global  $^{14}\text{C}$  database will allow for significant advances to be made in archaeology, anthropology and Quaternary paleogeography.

Keywords: Anthropocene, radiocarbon dates as data, frequency distribution, Holocene, paleodemography, HYDE

## 2.2 Introduction

Estimating ancient human impacts on the landscape or on the global climate (i.e., Early Anthropocene Hypothesis; Ruddiman, 2003) requires information about past population densities, technological development and land use patterns, as well as paleoenvironmental information. Empirical studies associating human and environmental change document interactions in both directions, as human activities modified the environment at many temporal and spatial scales, and past environmental changes affected human population growth and technological adaptations. Model-based studies of the potential effects of Holocene human activity on the environment (Kaplan et al., 2009; 2010; Klein Goldewijk et al., 2010; Pongratz et al., 2008) have provided useful tools for hypothesis-testing and synthesizing available knowledge toward this goal. Scaling human impacts on the environment upward from local studies to quantify global impacts requires appropriate databases and methodologies to analyze them.

In this paper, we are concerned with empirical studies to quantify human impacts on the environment at large spatial and temporal scales. We discuss one key variable needed in this research program: estimating human population sizes over the course of the Holocene at regional and global scales. We approach this problem using data obtained from archaeological studies and accumulated in databases and discuss a methodology to convert these data to quantitative estimates of past populations at regional to continental to global scales.

Estimates of total human population numbers and densities for the Holocene have been made and form part of the History of the Global Environment (HYDE) 3.1 database (Klein Goldewijk et al., 2010; Klein Goldewijk et al., 2011). These are hindcast based on suppositions of historical population numbers from multiple sources (Lahmeyer, 2004; Livi-Bacci, 2007; Maddison, 2001; McEvedy and Jones, 1978) and consequently do not account for short-term population booms or busts nor changes in the spatial distribution of settlements. Although these are the most widely-used estimates available, they need to be verified against independent datasets. Further, in this kind of effort, it can be difficult and, in some cases, not possible to include model parameters which account for demographic drivers such as climate or vegetation change, or factors such as disease, war and famine which can produce high-frequency fluctuations in population curves, but cannot be known except by using some external data input.

In parallel to this model-based approach is an empirical approach which reconstructs past conditions using fossil or ethnographic data, or a combination of both. A large amount of information about human technological adaptations to past environments is provided in archaeological and ethnographic records, and a long history of archaeological investigations and paleoenvironmental studies from around the world has documented Holocene human-environment interactions. Approximate and qualitative estimates of population sizes in the Americas have been attempted through various means including theoretical calculation and ethnographic studies (e.g., Johnson, 2014; Johnson et al., 2015; Newson, 1996), but have arrived at greatly different estimates (Haynes et al., 2007; Waters and Stafford, 2007), and in any event

require verification. Data from networks of archaeological sites enable the mapping and analysis of the magnitude of human impacts on the landscape, and also the potential effects of climate or environmental change on local human activities or population size.

One data source commonly used to estimate past demographic trends are radiocarbon databases obtained from samples collected at archaeological sites (e.g., Peros et al., 2010). Radiocarbon ( $^{14}\text{C}$ ) databases record fluctuations in temporal frequency distributions of  $^{14}\text{C}$  dates, which are considered proportional to human population density, and are therefore indicative of paleodemographic trends (e.g., Anderson et al., 2011; Crombé et al., 2014; Gamble et al., 2005; Rick, 1987; Williams, 2012). These results can then be mapped or their temporal distributions studied in relation to paleoenvironmental data (e.g., pollen records). For example, in North America, summed frequency distributions of  $^{14}\text{C}$  dates have been compared to time series of vegetation and climate in the northeastern United States to postulate potential impacts of environmental change on population size or cultural change (Munoz et al., 2010). Maps of population density through the Holocene are available (Chaput et al., 2015; chapter 3).

If we are to arrive at more definite estimates of ancient population numbers, a necessary input for research on human-environment interactions in the Anthropocene, we must combine the results from current data-based and model-based studies. In this paper, we discuss the methodology of obtaining estimates of the spatiotemporal population distribution of ancient humans using  $^{14}\text{C}$  databases. We assess the current state of regional  $^{14}\text{C}$  datasets which can serve as the base for paleodemographic estimates and evaluate their potential for producing empirical estimates of past population density. Our comparisons of the most complete  $^{14}\text{C}$  databases from North America and Australia with HYDE 3.1 population estimates (Klein Goldewijk et al., 2011) indicate that  $^{14}\text{C}$  data can effectively be used to evaluate model-based estimates of past populations. Thus, they can serve as inputs to studies quantifying human-induced paleoenvironmental change or evaluating the Early Anthropocene Hypothesis.

## **2.3 Archaeological radiocarbon data at a continental scale**

In this section, we present an introduction to the use of  $^{14}\text{C}$  data as a proxy for paleodemographic change. We then discuss briefly the most common issues that arise when analyzing large collections of  $^{14}\text{C}$  data. This is followed by a description of data currently available in openly-accessible format or in some cases from the literature.

### *2.3.1 Radiocarbon background and methods*

The statistical analysis of compilations of  $^{14}\text{C}$  dates has a long history in the fields of Quaternary geology and paleoclimatology (e.g., Dyke, 2005; Gajewski et al., 2006; Wendland and Bryson, 1974). In archaeology, “dates as data” is the common name for the practice of using  $^{14}\text{C}$  dates as a form of data from which spatiotemporal paleodemographic change may be inferred. It was first used in archaeology by Rick (1987) who suggested that an accumulation of culturally-associated

dates in any given area should be related to the degree of human presence on the landscape, if inherent biases and sources of error are accounted for.

Biases and errors fall into two groups. The first group includes errors strictly associated with  $^{14}\text{C}$  dating, including reservoir effects from the oceans and atmosphere (Tull et al., 2013), issues with modern  $^{14}\text{C}$  dating methods (Bowman, 1990), calibration effects (Stuiver et al., 1998), and the usual problems of sample contamination. Although it is necessary to adequately deal with these, many are known and can be corrected, and these corrections are universal to anyone working with Quaternary fossils. These are not discussed further here. The second group is more particular to the dates as data approach, and is more difficult to correct for. It includes biases and errors associated with archaeological studies, such as the taphonomic loss of samples and sample size (Williams, 2012), as well as issues associated with the analysis of databases of these sites (e.g., uneven sampling strategies both within and across sites). Solutions to these issues have been proposed and research is ongoing to better understand and control for these biases (section 2.2).

The association between  $^{14}\text{C}$  dates and human population density has been most-commonly made using temporal frequency distribution plots where numbers of dates are summed and binned into successive time intervals, resulting in a histogram of frequencies of dates through time (e.g., Gajewski et al., 2011; Peros et al., 2010). Since the frequency of dates is interpreted to be proportional to population density, these histograms indicate relative trends in population increase or decrease, but do not provide an absolute numerical estimate unless they can be calibrated, for example, using estimates that overlap reliable census records.

### *2.3.2 Interpreting radiocarbon datasets*

There are a number of considerations when interpreting population signals in  $^{14}\text{C}$  data (e.g., Brown, 2015; Surovell and Brantingham, 2007; Williams, 2012) :

1. **Quality issues:** an initial assessment of data quality and reliability should be done prior to any in-depth analysis. For example, a verification that dates have been logically assigned to the correct cultural phase based on its age range may be necessary (Gajewski et al., 2011). Duplicate laboratory identification numbers and large standard deviations should also be investigated. Flohr et al. (2015) describe a well-documented series of steps that could be followed. More generally, Woods (2015) explains the need for consistent publication practices when reporting  $^{14}\text{C}$  results. Differences in the sampling error of dates obtained using earlier methods (e.g., conventional radiocarbon dating) versus dates obtained using modern methods (e.g., Accelerator Mass Spectrometry) should also be considered. Population estimates are typically interpreted without considering the errors of the calibrated  $^{14}\text{C}$  dates, therefore little information is available on the effects of combining dates that were measured using different methods. In continental-scale studies using large numbers of dates, it is likely that

issues associated with dating methods are less influential, since precision is of greater importance in these cases than accuracy (i.e., if the dataset contains a systematic error, the large sample size ensures that the overall population estimate is still attainable). However, when small datasets are used, accuracy may become more important than precision, and any systematic differences between dating methods may need to be addressed.

2. **Sample size:** as with most point data, a representative sample must first be obtained. It is unclear what constitutes a representative sample of  $^{14}\text{C}$  dates, but in general a larger number of dates is preferable. Williams (2012) suggested that at least 500 dates are needed to confidently discuss dominant population trends. However, Peros et al. (2010) showed that even a small ( $\sim 0.0001\%$ ) sample can be representative of the source population. In another study at a continental scale of North America, Chaput et al. (2015) randomly removed 50% ( $n=16,894$ ) of their data and this did not alter their conclusions related to paleodemography, suggesting sample size requirements may change on a case by case basis. In any event, the statistical field of sampling theory provides a guide to this question (e.g., Cochran, 1963).
3. **Sampling bias:** In some instances, sampling bias issues can be difficult to address. Some areas are studied more, or are more accessible, which would lead to more dates from that region. It is often the case that the sampling strategies used by archaeologists are designed with a particular region, site or culture in mind. This is frequently the result of funding strategies and mandates of broader archaeological programmes. This can lead to an uneven distribution of  $^{14}\text{C}$  dates in space, with sites of greater interest frequently being dated more intensely. One solution to spatial sampling bias is presented in Chaput et al. (2015), but more research is needed on this subject.
4. **Temporal bias:** similarly to spatial sampling bias, there is also temporal bias where a particular time period may be studied in more detail than previous or subsequent periods due to the particular research question. It is also possible that within a given set of  $^{14}\text{C}$  data, several dates may originate from the same “site phase” or context based on occupation events and cultural complexes noticeable in excavation horizons at a single site. Treating these dates as independent samples may create a temporal bias leading to the overrepresentation of a particular time period. Ways of accounting for this bias are discussed at length in Shennan and Edinborough (2007), Shennan et al. (2013) and Peros et al. (2010). For example, using a single pooled mean date per site phase when sample sizes are large enough, or carefully choosing bin sizes to reflect cultural shifts are ways of addressing this issue.

5. **Taphonomic issues:** it is expected that temporal distributions of  $^{14}\text{C}$  dates will show an approximate exponential increase towards the present as populations flourished due to evolving hunting strategies, agriculture, technological innovation and industrialization. However, this expected increase is confounded by the natural taphonomic loss of older materials from the fossil record. Taphonomic corrections have been adapted and applied to data from different regions of the world based on site type and environment (e.g., Chaput et al., 2015; Munoz et al., 2010; Oinonen et al., 2010; Peros et al., 2010; Surovell et al., 2009; Williams, 2012; 2013), but more research on this subject is needed.
6. **Human factors:** different forms of past human activity can result in the deposition of unequal amounts of datable archaeological material largely due to the amount of waste generated and the nature of the material being left behind. For example, changes in cultural (e.g., stone replacing wood in tool manufacturing), settlement (e.g., nomadic movement vs. the construction of permanent structures), alimentation (e.g., hunting and gathering vs. agriculture and pastoralism), depositional (e.g., scattering hunting remains vs. depositing marine shells in mounds and burying other food waste in trenches) and religious (e.g., leaving the deceased in open, unmarked areas vs. using individually marked graves) practices could influence the amount of datable material left at a site. In turn, this could systematically alter the resulting summed probability curve based on these materials. The rate of production of datable material varies temporally and across populations, and the systematic nature of this variation could introduce a bias (e.g., between cultures that construct monoliths vs. those that do not; Torfing, 2015). However, Timpson et al. (2015) detail how to proceed with summed probability distributions despite these issues by, for example, testing explicit hypotheses (i.e., Shennan et al., 2013).
7. **Comparability:** in some instances, and especially in continental-scale analyses, it is assumed that all sites (and thus all dates) are comparable and represent equal population units despite any differences in feature type, site size, culture, or environment. This would imply, for example, that the population represented by a certain number of dates from a site in the High Arctic is equivalent to the population represented by the same number of dates from a site in the Eastern Agricultural Complex. When using comprehensive radiocarbon databases populated by thousands of individual dates for the purpose of comparing, in a general sense, changes in relative population size through time at the continental scale, the inaccuracies associated with this approach are likely masked by the overall population trend. However, finer-scale analyses require that this assumption be dealt with properly. Research which takes into account the importance of distinct archaeological episodes or “periods” as well as the spatial distribution of dates for the study of past populations is currently being done by Palmisano et al. (2018) and Bevan et al. (2017), for example. These studies show

that it is possible to systematically compare dates from different archaeological excavations and reports, however this work must be done on a case-by-case basis, objectively taking into account known details of each population in question.

8. **Classification of dates:** a distinction can be made between dates that are classified as cultural and dates that are classified as paleoenvironmental (Gajewski et al., 2011). Cultural dates refer to those obtained from material directly associated with human activity (e.g., charcoal from a hearth, human bone from a burial feature, charred cooking residue from a pot, etc.). Paleoenvironmental dates are derived from materials not directly associated with human activity (e.g., stratigraphic horizons, paleosols, animal collagen, etc.). When interpreting paleopopulation signals, only cultural dates should be included; paleoenvironmental dates may be used to interpret synchronous environmental transitions such as climate oscillations or deglaciation patterns.
9. **Class of information:** a distinction can be made between data that are classified as archaeological and data that are classified as historical (e.g., Muckelroy, 1976; Rougeulle, 2017). Archaeological data refer to data that have been collected at an excavation site and recorded/described in the site report (e.g., artifacts, features, structures, etc.). Historical data, on the other hand, include documented records of the past or other physical sources of evidence which are generated manually by humans or are manufactured (e.g., censuses, coins, photographs, etc.). Both classes of information are important for reconstructing past populations since archaeological records can be extended into modern times using historical data sources, although the methodology to do so needs to be developed.

### *2.3.3 Archaeological datasets available online or in publications*

To implement the dates as data approach, databases are created containing  $^{14}\text{C}$  dates extracted from an exhaustive sample of archaeological sites obtained from the literature. These georeferenced datasets contain the  $^{14}\text{C}$  dates, along with information needed for their analysis such as cultural association, material dated, and fields pointing to the original reference. These databases may include non- $^{14}\text{C}$  dates, such as those determined using luminescence, oxidizable carbon, uranium-series, electron spin resonance or amino acid racemization methods.

Large datasets of archaeological  $^{14}\text{C}$  dates are currently available for five regions of the world (**Fig. 2.1; Table 2.1**), but spatial coverage in these regions is uneven, and there are areas where data compilation has not begun. In some cases, these datasets are online and can be downloaded; in others, they come from the supplemental information or date lists embedded in original publications. To enable the discussion of these datasets, we converted all conventional (i.e., beta counts) and AMS (accelerator mass spectrometry)  $^{14}\text{C}$  dates to calendar years before present using CALIB v.7.0.4 (Stuiver and Reimer, 1993) and the IntCal13 (Reimer et al., 2013) northern and southern hemisphere, terrestrial and marine calibration curves. The calibrated age assigned

to each  $^{14}\text{C}$  date was the median of the 2 sigma ages. The histograms in [Figure 2.2](#) are based solely on cultural  $^{14}\text{C}$  data; dates from paleoenvironmental materials are excluded.

1. **North America - Canadian Archaeological Radiocarbon Database:** In North America, over 31,600 uncalibrated  $^{14}\text{C}$  dates ([Fig. 2.1](#)) are included in the CARD (Canadian Archaeological Radiocarbon Database v2.0, which includes data from both Canada and the United States; Gajewski et al., 2011). The CARD is a searchable and downloadable database ([Table 2.2](#)) that is constantly being updated and its geographic extent expanded with new submissions from around the world. The vast majority of dates are georeferenced and culturally-associated, and many are described in detail (e.g., geographic context, site type). The development of the database and details about its content is described in Gajewski et al. (2011), and work is in progress to update and further verify the data.

Dates originate from cultural and paleoenvironmental materials and range from 0 to beyond 40,000 cal BP although most are from the last 15 ka ([Fig. 2.2](#)), the period following human colonization. Early to mid-Holocene population numbers are low but increase steadily. There were two periods of marked growth (8 – 2.5 and 2.5 – 0.6 ka) which are coincident with known cultural shifts in Indigenous populations (Peros et al., 2010). There was a decline in the number of dates in the last 500 years, the period after European colonization of the continent. Although it is well known that there was a major population crash due to the spread of European diseases and other factors associated with European colonization, potential sample bias also exists, where archaeologists would not date site phases using radiocarbon if historical data such as artifacts or historical descriptions with known ages are available.

A reconstruction of the relative population size of North America through time has been attempted using the CARD (Peros et al., 2010). Data have been used to investigate the likelihood of the extra-terrestrial impact hypothesis (Buchanan et al., 2008; 2011; Steele, 2010). Studies of the impact of the Younger Dryas climate fluctuation (Miller and Gingerich, 2013) on the population density during this period (Anderson et al., 2011) indicate a dramatic decline in Paleoindian populations as temperatures cooled. The overlap between changes in Paleoindian and megafauna populations in the early Holocene has been explored (Bamforth and Grund, 2012; Boulanger and Lyman, 2014) and a recent spatiotemporal analysis of the CARD which accounted for spatial, temporal and taphonomic bias revealed the dynamics of human dispersal across the continent over the last 13 ka (Chaput et al., 2015). Land use changes associated with the introduction of agriculture had an impact on the forests of southern Ontario in areas where population density was greater (Munoz and Gajewski, 2010).

2. **Australia - AustArch database:** In Australia, 5,522 uncalibrated  $^{14}\text{C}$  and non- $^{14}\text{C}$  dates are included in the AustArch database (Williams et al., 2014), which is available online ([Fig. 2.1](#), [Table 2.2](#)). Dates have been collected from over 1,700 sites from across the continent, and include cultural, geographic, material and site type descriptions. A detailed description of the database is provided in Williams et al. (2014).

The population in Australia shows a gradually rising trend with more abrupt changes at 8 and 5 ka followed by rapid growth until 0.6 ka. These changes have been attributed to human exploitation of land and resources following climatic amelioration in successive parts of the continent (Williams, 2013; 2015c). The decline in the number of dates in the past 300 years is resultant of the decline in Indigenous populations following European colonization in the late 18th century and the spread of European diseases.

The first of three successive editions of the AustArch database was published by Williams et al. (2008a) and contained data from the Australian arid zone. The second (Williams et al., 2010; Williams, 2012) and third (Williams and Smith, 2013) editions added data from the remaining ecological regions in Australia. Several continental studies have been done using this database which provide distributions of relative population size over the last 40 ka (Williams, 2013), locations of potential human refugia before the Pleistocene-Holocene transition (Williams et al., 2013) and compare global climate records to human settlement patterns (Williams et al., 2015a). More recently, human-induced environmental change (Williams et al., 2015b) and hunter-gatherer tendencies have been explored (Williams et al., 2015c).

3. **Europe - The RADON and Radiocarbon Palaeolithic Europe databases:** Two databases of  $^{14}\text{C}$  and non- $^{14}\text{C}$  data are available which cover Europe, and to a lesser extent parts of Siberian Russia and Scandinavia ([Fig. 2.1](#)). The RADON database contains ca. 11,000  $^{14}\text{C}$  dates from the Neolithic and early Bronze Age (Radiocarbon Dates ONLINE; Hinz et al., 2012) and is available on the web ([Table 2.2](#)). The online version allows for calibration (including summed calibrations for multiple dates) and location visualization in Google Earth. These dates have been collected from 80 countries and over 4,000 archaeological sites.

A second database (Radiocarbon Palaeolithic Europe Database v18; Vermeersch, 2015) contains over 18,000 dates from the Paleolithic. This database began as part of an INQUA Congress working group project on European human ecology and cultural adaptation (Vermeersch, 2005) and has since been available online ([Table 2.2](#)) for use and for contribution. It is made up of approximately two thirds  $^{14}\text{C}$  dates and one third non- $^{14}\text{C}$  dates which extend beyond 60 ka. Almost all dates are associated with a cultural phase and come with site, material and methodological

descriptions. Most come from the European Plain, however coastal areas and Siberian Russia are also represented. Successive versions of this database have been used to document and explain European population dynamics during major climate transitions (Bradtmöller et al., 2012; Vermeersch, 2011).

Because the two databases have been developed separately, a combined analysis of the entire time period is not possible at this time (Fig. 2.2). In addition, a temporal sample bias would need to be accounted for in some way (in more recent periods archaeologists would tend not to use  $^{14}\text{C}$  data if a site can be dated using historical means). This limits our ability to interpret the  $^{14}\text{C}$  record from Europe, however broad trends in population prior to the historical period are visible. Population numbers increased rapidly during the Upper Paleolithic from 50 to 38 ka when anatomically modern humans advanced across the region (Fu et al., 2015). Small variations in population growth and decline followed, attributable to changes in both climate and environment (Tallavaara et al., 2015). When the characteristically arid conditions and cold temperatures of the Last Glacial Maximum begin to ameliorate after 21 ka (Strandberg et al., 2011), population numbers increased at which point the lack of overlap between the two data sets from Europe makes further interpretation difficult.

4. **Asia - CONTEXT, Flohr et al. (2015) and Wang et al. (2014) databases:** These data come from three separate datasets from southwest Asia and China. The two databases from southwest Asia, a major area for archaeological research, have been combined in this study (Fig. 2.1). The southwest Asia datasets are the CONTEXT database (Böhner and Schyle, 2002-2006) and data presented in Flohr et al. (2015). CONTEXT contains 2,875  $^{14}\text{C}$  dates from 17 countries and over 400 sites. Flohr et al. (2015) compiled 3,397  $^{14}\text{C}$  dates originally from CONTEXT, but also from PPND (the Platform for Neolithic Radiocarbon Dates; Benz, n.d.) and CalPal (Weninger et al., 2013). CONTEXT is an online database (Table 2.2) and dates extend beyond 40 ka. The data supplied as supplementary information in Flohr et al. (2015) are most numerous in the early Neolithic although they extend beyond 12 ka. Flohr et al. (2015) subjected all dates to a thorough quality assessment and provide information on whether or not a date should be included in time series analyses. There are no doubt numerous additional dates from southwest Asia not yet in these databases.

The third dataset includes 4,656  $^{14}\text{C}$  dates from archaeological sites which are from almost all of China. These dates are available in the supplementary information from Wang et al. (2014). All dates have been laboriously checked for dating and descriptive errors, culturally-associated and georeferenced. The authors describe an exhaustive literature review process which enabled the collection of high-quality data.

Flohr et al. (2015) find no evidence of changes in population size or settlement patterns with respect to the 9.2 and 8.2 ka climate events and suggest rapid adaptation as a possible cause. Results from Wang et al. (2014) show positive correlations between population size and climate amelioration inferred from ice core isotopic records. Li et al. (2015b) expand on this by demonstrating shifts in agricultural practices towards more sustainable crops and efficient technologies were also positively correlated with population growth.

Histograms of data from both parts of Asia are similar in shape (Fig. 2.2). The  $^{14}\text{C}$  records from southwest Asia show maximum values centered on 9 ka whereas the population in China gradually increased after 9 ka following the onset of the Neolithic and Bronze Age periods. All records show a decline in the total number of dates from the mid-Holocene onwards presumably due to the use of historical information to date sites and the early development of writing, although the timing differs in the two areas.

#### 5. **South America:**

We compiled the supplementary information and lists of  $^{14}\text{C}$  and non- $^{14}\text{C}$  dates from several publications to form a preliminary dataset from South America. This database includes over 2,000 cultural dates from studies in Argentina (Barbarena et al., 2015; Bonomo et al., 2015; Gil et al., 2005; Mancini et al., 2013; Prates et al., 2013; Steele and Politis, 2009), Chile (Campbell and Quiroz, 2015; Latorre et al., 2013), Amazonia (Lombardo et al., 2013), Uruguay (López Mazz, 2013), Ecuador (Ogburn, 2012), the Isthmus of Panama (Cooke et al., 2013), Brazil (Bueno et al., 2013) and from greater South America (Barnosky and Lindsey, 2010). This compilation is not exhaustive considering many dates are likely reported in non-accessible journals or archaeological bureau records. In most cases,  $^{14}\text{C}$  data are used to track population expansion following regional environmental change or to infer migration routes following colonization. Many different cultures, site types and biomes have been studied.

South American dates span the last 20 ka; however a lack of dates from the mid-Holocene is apparent (Fig. 2.2). This is likely due in most part to an incomplete database; however several studies suggest regional site abandonment and population decline during this time resulting from climatic deterioration (e.g., Araujo et al., 2005). These studies show that a large-scale analysis of population growth, such as Peros et al. (2010) attempted for North America, should be possible in the near future if the data available in local and regional studies continue to be accumulated.

## 2.4 Regional-scale analyses

Large datasets which are sufficiently representative in both time and space are useful at not only continental-scales, but at regional-scales as well. For example, Williams et al. (2008b) studied paleodemographic variability apparent in the  $^{14}\text{C}$  record in response to ENSO activity at regional-scales across Australia. Studies from South America were discussed above.

Taphonomically-corrected histograms of the dates from the CARD plotted by cultural region (Gajewski et al., 2011; Ubelaker, 1992; 2006) show considerable variability in time (**Fig. 2.3**) and this variability can be related to regional paleoenvironmental information. In a broad sense, these distributions offer insight into the colonization and settlement of these regions. Population gradually increased in the Subarctic, the Plateau and along the northwest coast following colonization. The Great Basin shows periods of growth during the mid- to early Holocene followed by intensive growth in the last millennium. The remaining regions have highly variable  $^{14}\text{C}$  distributions, and century- to millennial-scale relationships can be studied as well. For example, Munoz et al. (2010) used the CARD to relate changes in population and culture to climate variability and vegetation change for the northeastern United States, which has a well-established series of cultural phases. All major cultural transitions could be related to environmental changes, and they were also associated with demographic changes as well. The availability of a relatively complete regional database enables quantitative associations to be made between archaeological and paleoenvironmental changes.

## 2.5 Data-model comparison

Klein Goldewijk et al. (2010) have estimated total population numbers and densities for major regions across the world. These estimates are “hindcast” to 10,000 BC based on growth rates calculated using population estimates from Lahmeyer (2004), Livi-Bacci (2007), Maddison (2001), and McEvedy and Jones (1978) and interpolated where numbers are missing. The resultant database – HYDE 3.1 – has been used as input for global change models (e.g., Hurtt et al., 2011; Kaplan et al., 2010; Strandberg et al., 2014). We compare HYDE population estimates with the taphonomically-corrected (Surovell et al., 2009; Williams et al., 2014)  $^{14}\text{C}$  frequency distributions for North America and Australia (**Fig. 2.4**), as these are the only regions where continental-scale paleopopulation growth curves have been developed (Peros et al., 2010; Williams, 2013). Radiocarbon data are truncated at 15 ka to focus on the time range estimated by Klein Goldewijk et al. (2010).

For North America, the overall trend between the modelled and archaeological data is comparable, although not all periods of population expansion begin at the same time. Klein Goldewijk et al. (2010) show a stable population from the first arrival though to 8 ka, a slow increase until 6 ka, and a rapid increase until 0.5 ka after which population crashes (**Fig. 2.4a**: line). The same tendencies are seen in the CARD data (**Fig. 2.4a**: bars). Overall patterns in the distribution of Australian data are similar to the CARD data. The HYDE estimates show a constant population size from 12 ka until 1.4 ka (**Fig. 2.4b**: line). HYDE correctly estimates a

subsequent step change, but after this HYDE predicts constantly increasing values towards the present when the  $^{14}\text{C}$  data suggest a second step change ca. 1 ka. The AustArch data show relatively low values until 5 ka, an increase between 5 and 2 ka and two step changes between 2 and 1 ka, and a decrease over the past 400 years (Fig. 2.4b: bars).

In general, the broad-scale temporal trends in the  $^{14}\text{C}$  data and modelled population curves are similar, simultaneously validating the paleoreconstruction inputs and highlighting the effectiveness of the dates as data method. However, the  $^{14}\text{C}$  data indicate finer-scale fluctuations which can be the object of further study (cf. Munoz et al., 2010). This divergence is also seen when the two datasets are compared spatially. Maps of population density based on HYDE 3.1 suggest widespread population growth in North America took place exclusively over the last 200 years (Klein Goldewijk et al., 2010). By spatiotemporally mapping  $^{14}\text{C}$  data from the CARD, Chaput et al. (2015) clearly show the unlikeliness of this scenario and demonstrate significant population growth throughout the Holocene and across many different regions. This implies the HYDE approach is reliably estimating broad-scale variations in population, but the  $^{14}\text{C}$  data can provide a more complete representation of population change at varying scales, which could lead to paleodemographic studies of a more detailed nature. This also suggests  $^{14}\text{C}$  databases are useful if not necessary for calibrating future model outputs from, for example, land use and land cover change models.

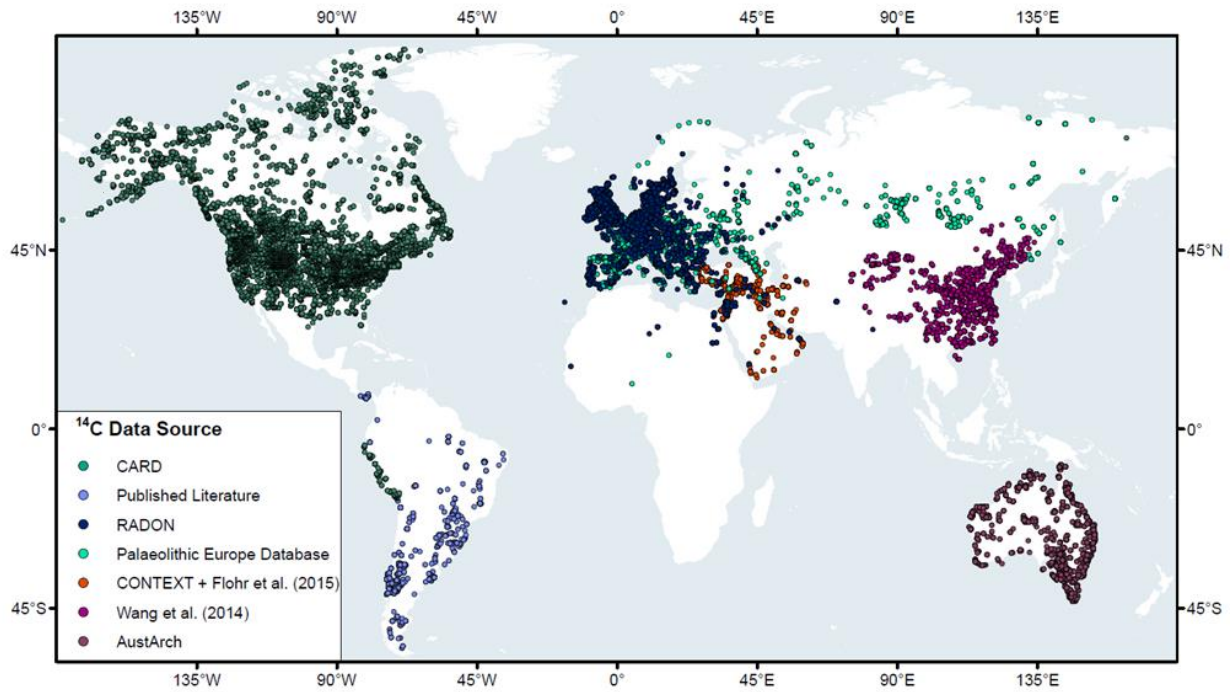
## 2.6 Conclusion

Dates as data is now a well-established method that will continue to be developed as the availability of  $^{14}\text{C}$  databases increases. Scrutiny of the methodology has resulted in improved understanding of the assumptions. Applications in regional and continental studies of summed frequency distributions and mapped point patterns are leading to new insight into human population growth and spatiotemporal distribution. A first comparison between empirical and model-based estimates of ancient populations showed encouraging similarities between both approaches. Our results lend support to both empirical and model-based studies of paleodemographic change and serve as a useful reminder of the assumptions and limits associated with both methods.

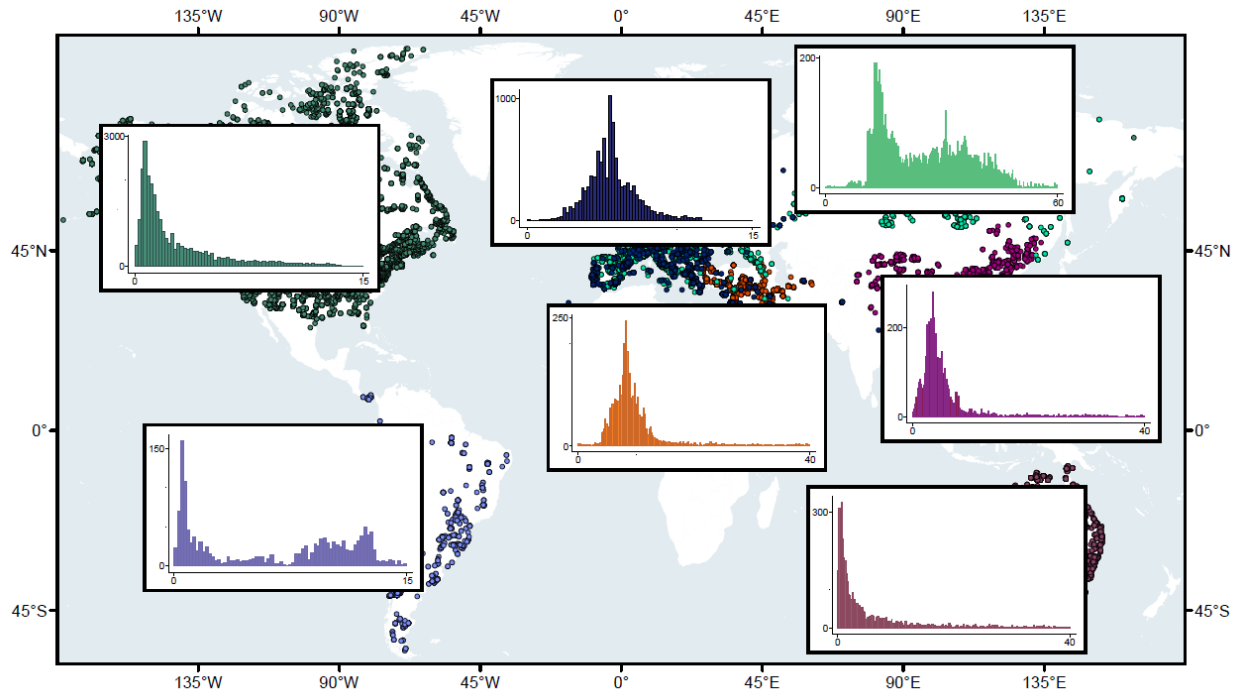
Databases of  $^{14}\text{C}$  data are available for many parts of the world including the Americas, Europe, Australia and parts of Asia. Data from South America, India, Africa, and additional parts of Eurasia are still needed. These data are all available on the web and are openly-accessible for studying paleodemographic change. The individual datasets from which these data originate still require quality assurance verification for issues such as duplicate laboratory sample references or missing and anomalous data. Nonetheless, these databases are an invaluable resource with which we can now estimate population change at varying spatial and temporal scales. The next step will be to use these data to refine estimates of human impacts on the climate and environment in the Anthropocene.

## **Acknowledgements**

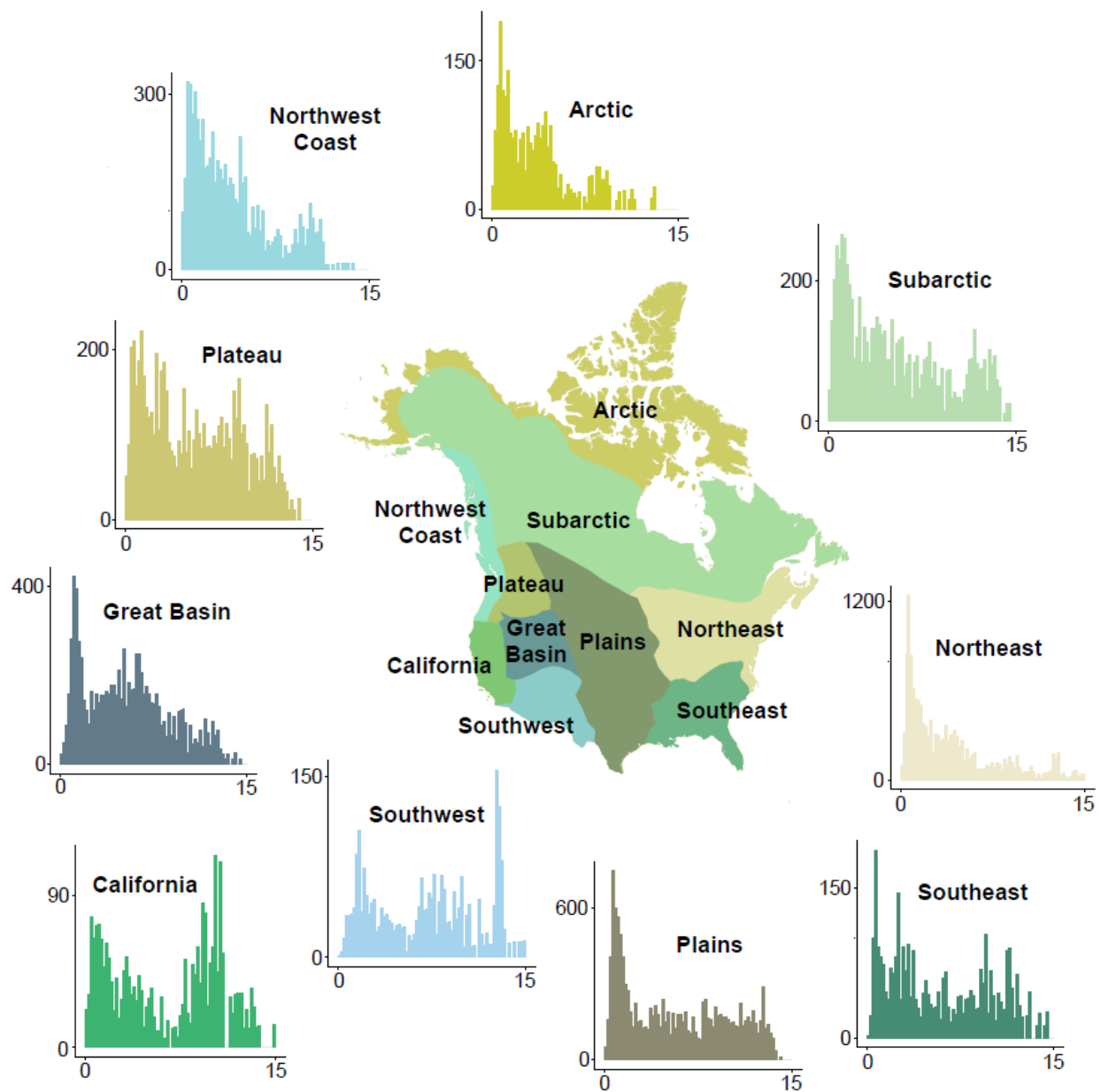
We would like to thank all contributors to the CARD, AustArch, RADON, CONTEXT and the Radiocarbon Palaeolithic Europe Database, as well as individuals who collected and published South American and Asian dates. Assistance from Dr. Andrew Martindale and Dr. Kevan Edinborough with the CARD is also acknowledged. This research was funded by an NSERC Discovery Grant to K.G.



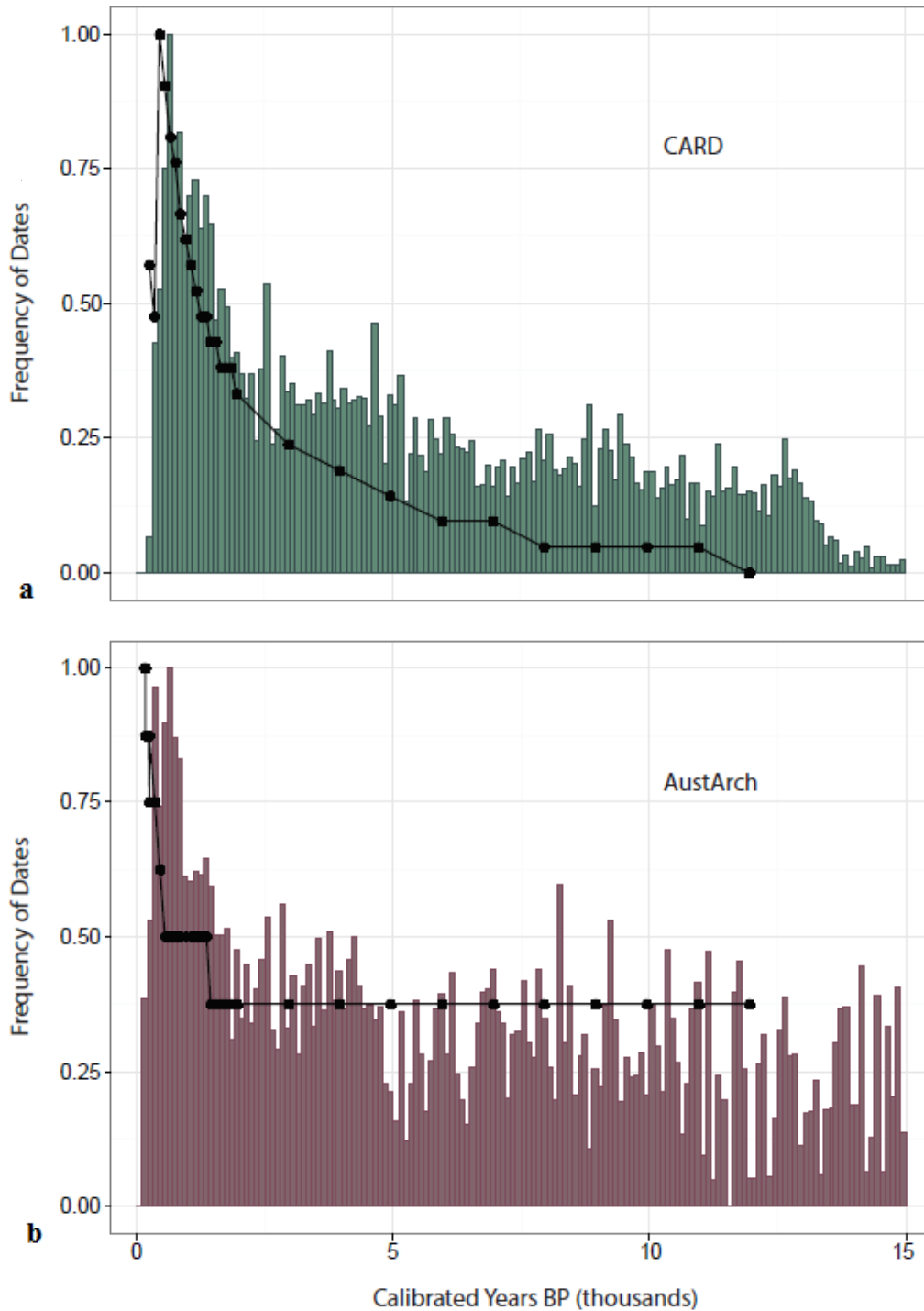
**Figure 2.1** Distribution of archaeological radiocarbon databases publically available on the web or in open literature (**Table 2.1**). Some points are dated using other methods (e.g., luminescence). Locations of dates from Gil (2005), Mancini et al. (2013), Barnosky and Lindsey (2010), Campbell and Quiroz (2015), Latorre et al. (2013) and Lombardo et al. (2013) were georeferenced in ArcGIS from original publication figures.



**Figure 2.2** Frequency distributions of calibrated, cultural  $^{14}\text{C}$  and non- $^{14}\text{C}$  archaeological dates from the databases in **Table 2.1**. Frequencies have not been corrected for taphonomic loss. Individual distributions were truncated at 15, 20, 40 or 60 ka to focus on time periods containing at least 90% of the total number of dates. Bins = 200 years. 0 Cal BP = AD 1950.



**Figure 2.3** Taphonomically-corrected frequency distributions of calibrated, cultural  $^{14}\text{C}$  dates from the CARD, as a function of the cultural regions presented in Ubelaker (1992; 2006) and Gajewski et al. (2011). Bins = 200 years.



**Figure 2.4** Taphonomically-corrected frequency distributions of calibrated, cultural  $^{14}\text{C}$  and non- $^{14}\text{C}$  dates from the CARD (North America) and AustArch (Australia) databases (bars) plotted against HYDE 3.1 (Klein Goldewijk et al., 2010, Klein Goldewijk et al., 2011) population estimates (lines). In both cases, data have been scaled by the respective maximum value of each dataset and truncated at 15 ka. Taphonomic corrections follow Surovell et al. (2009) for CARD and Williams (2012) for AustArch. Bins = 100 years. 0 Cal BP = AD 1950.

**Table 2.1** Summary of databases containing  $^{14}\text{C}$  and non- $^{14}\text{C}$  dates documenting past human presence or cultural activity.

Region	Database	Reference	Total number of dates	Calibrated age range (BP)	Total number of cultural dates in <a href="#">Fig. 2</a>
North America	Canadian Archaeological Radiocarbon Database (CARD)	Gajewski et al., 2011	31,692	Modern -48,656	26,454
South America	Several regional compilations including: Barbarena et al., 2015; Bonomo et al., 2015; Gil et al., 2005; Mancini et al., 2013; Prates et al., 2013; Steele and Politis, 2009; Barnosky and Lindsey, 2010; Campbell and Quiroz, 2015; Latorre et al., 2013; Lombardo et al., 2013; López Mazz, 2013; Ogburn, 2012; Cooke et al., 2013; Bueno et al., 2013		>2000	Modern - 37,542	1,522
Australia	AustArch	Williams et al., 2014	5,522	Modern - 270,000	4,170
Europe	Radiocarbon Palaeolithic Europe Database	Vermeersch, 2015	12,424	Modern - 3,950,000	10,886
Europe	RAdiocarbon Dates ONline (RADON)	Hinz et al., 2012	10,975	Modern -40,866	10,073
SW Asia	CONTEXT	Böhner & Schyle, 2002-2006	2,874	Modern - 49,114	2,760
SW Asia	-	Flohr et al., 2015	3,397	5,723 - 12,865	1,272
China	-	Wang et al., 2014	4,656	Modern – 46,070	4,656

**Table 2.2** URL of online radiocarbon databases.

Database	Website
CARD	<a href="http://www.canadianarchaeology.ca/">http://www.canadianarchaeology.ca/</a>
AustArch	<a href="http://archaeologydataservice.ac.uk/archives/view/austarch_na_2014/">http://archaeologydataservice.ac.uk/archives/view/austarch_na_2014/</a>
RADON	<a href="http://radon.ufg.uni-kiel.de/">http://radon.ufg.uni-kiel.de/</a>
<sup>14</sup> C Palaeolithic Europe	<a href="http://ees.kuleuven.be/geography/projects/14c-palaeolithic/">http://ees.kuleuven.be/geography/projects/14c-palaeolithic/</a>
CONTEXT	<a href="http://context-database.uni-koeln.de/index.php">http://context-database.uni-koeln.de/index.php</a>

### **3 Spatiotemporal Distribution of Holocene Populations in North America**

Michelle Chaput, Bjoern Kriesche, Matthew Betts, Andrew Martindale, Rafal Kulik, Volker Schmidt and Konrad Gajewski

#### **3.1 Abstract**

As the Cordilleran and Laurentide ice sheets retreated, North America was colonized by human populations; however the spatial patterns of subsequent population growth are unclear. Temporal frequency distributions of aggregated radiocarbon ( $^{14}\text{C}$ ) dates are used as a proxy of population size and can be used to track this expansion. The Canadian Archaeological Radiocarbon Database (CARD) contains more than 35,000  $^{14}\text{C}$  dates, and is used in this study to map the spatio-temporal demographic changes of Holocene populations in North America at a continental scale for the past 13,000 years. We use the kernel method which converts the spatial distribution of  $^{14}\text{C}$  dates into estimates of population density at 500-year intervals. The resulting maps reveal temporally distinct, dynamic patterns associated with paleodemographic trends which correspond well to genetic, archaeological and ethnohistoric evidence of human occupation. These results have implications for hypothesizing and testing migration routes into and across North America as well as the relative influence of North American populations on the evolution of the North American ecosystem.

Keywords: Canadian Archaeological Radiocarbon Database, Holocene, North America, Paleoecology, Paleodemography, Kernel Methods, Taphonomy

## 3.2 Introduction

Databases of radiocarbon dates obtained from archaeological sites are used as a source of information about past demographic changes of Holocene populations (Munoz et al., 2010; Peros et al., 2010; Gajewski et al., 2011; Shennan et al., 2013; Boulanger and Lyman, 2014). These are typically analyzed using summed probability distributions for a given region, which are interpreted as time series of site occupancy and inferred population tendency. For example, frequency distributions have been used to estimate continental-scale population growth in North America (North America here means the US and Canada, the extent of our database) over the past 15 ka using the Canadian Archaeological Radiocarbon Database (CARD) (Peros et al., 2010; Gajewski et al., 2011). Regional studies in North America have highlighted the impacts of environmental change on population size using these methods and documented the impact of human activities on the vegetation (Buchanan et al., 2008; Munoz and Gajewski, 2010; Munoz et al., 2010; Steele, 2010; MacDonald et al., 2012; Kelly et al., 2013). Although some studies have attempted to investigate between-region demographic changes for particular time periods (Miller and Gingerich, 2013) or map the point distribution for certain time intervals (Gajewski et al., 2011), the spatio-temporal distribution of demographic growth through the Holocene in North America after the arrival of humans has not been tracked at a continental scale.

The underlying assumption of the “dates as data” approach is that the frequency of  $^{14}\text{C}$  dates is proportional to population density. This relation improves when aggregated data have a sufficient sample size to apply methods to overcome archaeological sampling bias and taphonomic and calibration effects (Rick, 1987). Summed probability distributions have been used to assess paleodemographic trends in the Americas, Europe, Eurasia and Australia (e.g., Gamble et al., 2005; Kuzmin and Keates, 2005; Munoz et al., 2010; Peros et al., 2010; Anderson et al., 2011; Ballenger and Mabry, 2011; Williams, 2012; Miller and Gingerich, 2013; Crombé and Robinson, 2014; Wang et al., 2014; Méndez et al., 2015) and methodological issues have been discussed at length (see Bayliss et al., 2007; Bamforth and Grund, 2012; Williams, 2012) and **Appendix 1** for examples). However, to determine and compare demographic patterns in both space and time, a way is needed to estimate the density of archaeological records at a particular time.

We build upon previous work in North America by using an alternative method with data from the CARD, and provide the first estimates of the spatio-temporal distribution of the population of North America during the past 13 ka. The kernel method is used to convert the spatial distribution of  $^{14}\text{C}$  dates into estimates of population density at 500-year intervals with a kernel radius (or bandwidth which we use interchangeably) of 600 km; a bin size and radius which optimize the macro-temporal and -spatial patterning at continental and geological scales (**Appendix 1**). The radius of 600 km was chosen based on the optimal bandwidth (595 km) calculated following Scott’s rule for bandwidth selection (Scott, 1992). Scott’s rule was chosen since this method is optimal for a large  $n$  and does not oversmooth histogram results (Scott, 1979). This bandwidth does not have a simple physical meaning (such as the range of a hunter-gatherer); it is meant to represent where humans could have been (in a probabilistic sense) during

the 500-year interval based on the distribution of dates, since each interval represents many generations of humans. A fixed bandwidth was chosen a) to ensure comparability across maps, b) to allow for the interpretation of results at the scale of the six major North American regions discussed below, and c) because, for the purpose of this study, the radius of human activity should be the same everywhere (ignoring geographical features).

The CARD is unique since it contains 35,905  $^{14}\text{C}$  dates, 33,756 of which have geographical, chronological and descriptive information permitting spatial and temporal analysis. The frequency of dates increases exponentially through time, with a maximum frequency around 0.7 ka cal BP (all future references to dates are in ka cal BP denoted simply as ka). Previous results (Peros et al., 2010), as well as our own cross validation exercises (**Appendix 1**) indicate that this amount and distribution of dates are sufficiently representative to begin studying past populations at a continental scale (Peros et al., 2010).

Based on our method, we hypothesize that the resulting intensity distributions (referred to as Radiocarbon Frequency Population Estimates or RFPEs) capture paleodemographic trends since 1) the RFPEs show clear patterns when methods to reduce taphonomic and sampling bias are applied, and 2) the patterns correlate well with pre-existing archaeological interpretations of human cultural-change across the continent that were derived partly from  $^{14}\text{C}$  dates but also other data sources and inferences, thus representing a partial independent test. The results of this study can be used to model human demographic growth in relation to environmental change and cultural innovation, and to determine human impacts on the landscape or fauna at a continental scale, particularly in relation to the Early Anthropocene Hypothesis (Ruddiman, 2003).

### 3.3 Methods

The Canadian Archaeological Radiocarbon Database (Gajewski et al., 2011) contains 35,905 radiocarbon dates derived from cultural and paleoenvironmental material collected from 9,149 geographically distinct archaeological sites. The selection of dates for this study depended on completeness of the database (locations for 77 dates are missing), location (61 dates are from Russia), dating information (2,071 entries are missing normalized  $^{14}\text{C}$  ages, normalized  $^{14}\text{C}$  errors, or both) and classification (cultural, paleoenvironmental, etc.). Descriptions of the cultural association and type of material dated, and general comments about the entry were also considered, although dates were not eliminated from the study if this information was missing. This resulted in 29,609 cultural dates which we used to create the RFPE maps, and 4087 dates from paleoenvironmental (or unknown) material (**Fig. 3.1a**). The  $^{14}\text{C}$  ages calculated from the CARD (Munoz et al., 2010) have been converted to calendar years (based on CALIB v.5.0.2 and IntCal04) using the median of the 2 sigma ages of the calibration curve (Gajewski et al., 2011). Since the maps are based on intervals of 500 years, any error when using the median as an estimate is much less than the sample interval.

To reduce sampling bias in the CARD, which is a result of inconsistencies in regional- (e.g., Canada vs. United States) and local-scale (e.g., Wyoming) sampling strategies and intensities, a

sampling intensity map (**Fig. 3.1b**) was created using kernel density estimation. Only geographically distinct sites ( $n=7,754$ ) were used to create **Figure 3.1b**; the amount of sampling (intensity) is estimated by considering only the geographic location of each site so that multiple dates at the same site have no influence. All archaeological sites were used in this step regardless of the age of the dated material or its classification (i.e., cultural, paleoenvironmental, unknown). This map is interpreted as the density of sites, in other words, the distribution of sampling sites in space, and not the number of  $^{14}\text{C}$  dates per site. If this remained unaccounted for, regions with higher sampling intensity would always be associated with elevated population estimates.

The underlying assumption of this approach is that a greater number of dates is indicative of a higher population density (Gamble et al., 2004; Gajewski et al., 2011). To produce the RFPE maps, a grid of points every  $0.1^\circ$  between  $23 - 85^\circ\text{N}$  and  $173 - 51^\circ\text{W}$  was created to define the area of interest (North America) in ESRI® ArcMap 10.1. This grid, in combination with ArcGIS© shapefiles of the extent of glacial ice at 1,000 year intervals (Dyke et al., 2002), was used to differentiate between glaciated and inundated areas and land. The period from 0.5 to 13 ka was chosen due to the lack of CARD dates prior to and after this date range. To visualize patterns in older time periods we would need to use a larger smoothing radius, but this tends to over-generalize the resulting patterns (**Fig. A1.1**). The data were examined in 500-year intervals as a compromise to avoid aggregating too many dates and having too few to produce a reliable estimate, with consecutive intervals overlapping every 100 years.

To temporally smooth the data within each interval and account for  $^{14}\text{C}$  dating errors, dates which do not occur directly within the interval are considered if they occur within 400 years prior to or following the interval. The dates within an interval are given a weight ( $w$ ) of 1 whereas the dates within  $\pm 400$  years of an interval are given a weight of  $<1$ . The weighting is achieved by taking a date which occurs outside of an interval and computing the temporal distance to the interval boundary and dividing the result by 400 (i.e., dates which occur at a greater distance from an interval boundary will have a smaller weight). Dates within 400 years prior to or following an interval are considered since they can still have an influence on population intensity (i.e., the estimated density of people at a given time and location is influenced by how many people were there before them) and to account for possible dating errors. When multiple dates exist for a single site within an interval, each date is considered separately.

Kernel density estimation, a method previously applied to archaeological data from Europe (Collard et al., 2010; Grove, 2011), is used to produce the RFPE maps. First, raw population estimates are produced. Generally, these estimates are highly correlated to the sampling intensity since a greater sampling effort produces a larger number of dates which results in higher population estimate (Peros et al., 2010). Therefore, producing RFPE maps based on raw estimates does not truly reflect human activity due to this sampling bias. A more reliable estimate is obtained when the raw population estimates are divided by the sampling intensity (**Figs. 3.2 and A1.2**). In doing so, RFPEs can be interpreted as the relative frequency of dates per site, which should better reflect population intensity than only considering the number of dates.

A similar approach has previously been used to account for a temporal bias (Peros et al., 2010). The values on the maps have been rescaled to one and are plotted in a logarithmic scale, with each colour being approximately three times that of the previous colour. This is done to maintain a consistent scale across all time intervals, and to make temporal changes more apparent, especially during earlier intervals. The resulting maps were corrected using a taphonomic formula that had previously been applied to the CARD data (Surovell et al., 2009; Munoz et al., 2010).

To assess the quality and completeness of the CARD, we ran two additional analyses. In the first, individual dates within the database flagged as “anomalous” (i.e., too young or old) were removed (n=1,419 or 5% of the original data points) and a new series of maps were made. This is representative of a worst-case scenario in which the dates with the highest probability of negatively-influencing the RFPEs are removed thus addressing the potential effect of “bad dates” on the quality of the maps. In the second analysis, a new series of maps is made after 50% of the dates (n=17,953) had been randomly selected and removed, to determine whether or not the amount of data in the CARD is spatially representative. **Figures A1.3** and **A1.4** show a very high degree of similarity between the original RFPE maps and the two new series despite the missing data, putting to rest our concerns regarding the state and usability of the database and its applicability in spatio-temporal analyses (see **Appendix 1** for additional discussion).

### **3.4 Distribution of <sup>14</sup>C dates**

At the continental scale of this study, the distribution of <sup>14</sup>C dates is the result of human activity which is proportional to population density, although data incompleteness, sampling intensity and natural taphonomic processes have an effect as well. There are too few dates in the CARD prior to 13 ka to create reliable population estimates prior to this time using the methods employed in this analysis. Other known issues with the database include a variable temporal and geographic representation, and in particular a lack of data from the southern United States which limits interpretation of results in this area.

As shown in **Figure 3.1a**, some areas are extensively represented (e.g., Wyoming) and others less so (e.g., Texas and Florida). Areas with greater sampling intensity are commonly the result of heightened archaeological interest, mapping objectives and available funding while others relate to issues involved with the construction of the database, including variable effort in data gathering from available literature (Gajewski et al., 2011). This creates a spatial bias which can confound continental-scale maps of population density. To account for this bias and normalize the series of RFPE maps, a spatially-varying sampling intensity was computed using <sup>14</sup>C dates from all sites. Using kernel density estimation, a sampling intensity map was then made which treated multiple dates from the same site in each time interval as one and standardized the influence of sites with higher sampling intensities (**Fig. 3.1b**, details in Methods).

There is a tendency in the CARD for a greater number of more recent archaeological sites since geological deposits tend to disappear over time due to normal taphonomic processes (Gajewski

et al., 2011). We appreciate that taphonomy may vary spatially, but apply a spatially constant taphonomic correction to the maps (Surovell et al., 2009; Peros et al., 2010) since the undertaking of how, where and by how much this should be varied requires more study at regional scales. We propose that areas along some coastlines may require an even greater correction due to the loss of accessible archaeological sites as a consequence of sea level rise, or due to reduced sampling of higher elevation stranded shorelines. Similarly for the Arctic, the kernel function radius of 600 km used to obtain the RFPEs (which is indicative of the spatial extent of the influence of a site on the surrounding intensity) may be too large due to geographic constraints (see “Canadian Arctic” below and [Appendix 1](#)).

### 3.5 Data verification

The only time period where patterns produced using the CARD data may be verified with independent data is just prior to European contact, when ethnohistoric estimates of the spatial distribution of population have been made (Ubelaker, 1992). Although the absolute numbers have been questioned (Thornton, 1987), the relative demographic and spatial patterns should be broadly accurate (Ubelaker, 1992).

The RFPE patterns estimated from 0.4 to 0.6 ka ([Fig. 3.3a](#)) generally correspond to those based on pre-contact estimates of population size ([Fig. 3.3b](#)), with half of the distinct regions in the maps differing by only one level of density. Across a band of northern United States (and British Columbia), where we have high confidence in the data, the estimates from both sources are similar. In the south where the database is known to be incomplete, and in New England, our estimates of population density are much lower. In the north, an area where the ethnohistoric data were known to be incomplete, our estimates differ by only 1 or two levels. In spite of calibration, taphonomic or sampling effects, a broad-scale pattern emerges from the data and a reasonable hypothesis is that it is recording a demographic trend. This is the first verification of the accuracy of the CARD data and suggests that the current number of archaeological records in the database, although not exhaustive, is sufficient to produce a general representation of the demographic history of North America (Peros et al., 2010), if the regional incompleteness of the database is accounted for. Larger data sets and refinements to the method will enhance this kind of modelling, as will updates of the CARD itself.

### 3.6 Results

#### 3.6.1 Alaska

RFPEs fluctuate steadily for the entire study period ([Fig. 3.2](#) and [Fig. A1.2](#)), which is realistic considering Alaska was the location of repeated migrations into the Americas (Reich et al., 2012). Estimates are relatively high in northern and central Alaska in comparison to other parts of North America between 7.5 and 13 ka, with a peak at 11.5 ka. This coincides with an abundance of sites associated with fluted points, characteristic of technologies used by Paleoindian groups after 13.5 ka (Goebel et al., 2013). Starting around 6 ka, a second increase is

centered on the Aleutian Islands and slowly moves east over the next 1.5 ka. This agrees with archaeological evidence suggesting Aleut peoples colonized the central islands ca. 6 ka (West et al., 2007). The stable but relatively low estimates in Alaska just prior to the Aleut occupation may be associated with a well-known “hiatus” in human presence (Rogers et al., 2009). Beginning at 4.5 ka, a 1,000 year-long increase is observed, composed of two centers along the northern and southwestern coasts, coinciding with the suspected migration of early Paleoeskimo from Siberia around 4.5 ka (Gilbert et al., 2008) and the Arctic Small Tool tradition in the eastern Aleutians (Hatfield, 2010). Estimates show a decrease for several hundred years and then begin to increase at the northern coast at 2.5 ka and along the west coast (2 ka), until finally the majority of the Alaskan region appears occupied. The development of Thule culture in Alaska (beginning ca. 1 ka) coincides with the increasing RFPE patterns (Dumond, 2000; Gilbert et al., 2008).

### *3.6.2 Canadian Arctic*

The maps show a rapid increase in RFPEs in coastal northwestern Canada, including the Mackenzie Delta, Banks Island and western Victoria Island around 6.5 ka, and an increase shortly thereafter on Ellesmere Island (**Fig. A1.2**). This predates the Independence I (Bennike and Andreasen, 2008) and early Pre-Dorset (Snow, 2010) occupations by several thousand years and may be due to known dating problems from Arctic sites (Friesen and Arnold, 2008) and temporal smoothing of the maps. However, these dates support recent genetic evidence suggesting there was human activity in the general region around 6,000 years ago (Achilli et al., 2013). By 4 ka, there are signs of occupation across the entire Canadian Arctic, and at 2.5 ka, site density increases to the south. This is due to data from Dorset culture sites which have been found from Victoria Island in the west to Greenland and Newfoundland in the east (Park, 2012). RFPEs on Victoria Island as well as the area surrounding Foxe Basin are likely underestimated, since Pre-Dorset archaeological evidence for the latter suggests a population was centered on this region around 2.5 ka (Friesen, 2007).

An increase in RFPEs along the southernmost Thule migration route from Alaska to western Hudson Bay beginning at 0.75 ka is expected in the context of history of the Thule people (Bennike, 2014). A site in Nunavut at 9.5 ka, in an area which was glaciated at this time, causes an isolated population center in the central Arctic, which persists until 8.1 ka at which point it is joined with the population further west (**Fig. A1.2**) and by which time the ice had retreated. The spatial extent of this population is a product of the kernel radius chosen (600 km); a larger radius would merge this population with its western neighbour. Although these dates appear inconsistent with the ice extent (which is, however, based on radiocarbon data), it is reasonable to hypothesize that humans were found along the edge of the ice at this time.

### *3.6.3 British Columbia and the Ice-free Corridor*

Prior to and during the Last Glacial Maximum (LGM), human migration into the Americas via an interior route was not possible due to the presence of ice sheets. At approximately 13.5 ka an

ice-free corridor opened between the Laurentide and Cordilleran ice sheets (Schurr and Sherry, 2004; Pitblado, 2011). Genetic and cultural evidence suggests a nomadic pre-Clovis population would have been the first to migrate northwards into the corridor from further south around 10.5 ka, following a previous migration south via a Pacific Ocean coastal route (Goebel et al., 2013). In the maps, a high-density population is observed in the ice-free corridor two millennia prior to 10.5 ka (**Fig. A1.2**).

RFPEs remain high until 12 ka when the Cordilleran ice sheet retreated and westward human expansion into this new terrain likely occurred. Around 11.5 ka, RFPEs increased along the entire west coast of British Columbia. This timing is in line with archaeological findings of maritime communities living on the outer coast of Haida Gwaii (Erlandson et al., 2008) and is consistent with the hypothesis that Pleistocene-early Holocene sites in outer coastal regions would today be underwater and thus poorly sampled archaeologically. A mirror isostatic effect, in which glacial loading depresses areas under ice sheets while raising offshore areas beyond the ice margin, strands shorelines along the mainland at much higher elevations, which leads to under-sampling (McLaren et al., 2011). At 7 ka, a second increase at the coast is observed, followed by an increase in the corridor. After 5.5 ka, populations appear to increase in both locations, potentially due to the stabilization of relative sea level on the coast and the development of new trap and tool technologies east of the Rocky Mountains (Maschner, 2012). In the late Holocene, the entire province of British Columbia was densely populated and remained so until 1 ka, reflecting the demographic success of complex hunter-gatherer cultures.

### *3.6.4 Eastern United States*

Beginning around 11 ka, there was an increase in RFPEs in the southeastern United States, which grew larger and expanded northward then westward after 9.5 ka (**Fig. A1.2**). This interpretation is plausible in light of numerous archaeological studies indicating a strong presence of Paleoindians in the southeast prior to 10 ka (Goodyear, 2004). RFPEs decreased east of the Appalachians after 9.5 ka and were centered in Missouri at 8.5 ka, where the presence of Paleoindian populations has been confirmed (Hajic et al., 2007). Estimates fluctuated until 4.9 ka when populations grew to the east and west of the Appalachians as well as in the Middle Atlantic and New England regions. After 3.5 ka, RFPEs greatly increased in the central part of the eastern temperate deciduous forest where studies of numerous sites document intense human occupation in the late Archaic and early Woodland periods (e.g., Dunbar Cave, Tennessee (Simek et al., 2012)). After 2 ka, the highest estimates were in the Ohio-Kentucky region (Cahokia and environs), and this persisted until the time of contact.

### *3.6.5 Atlantic Canada*

Between 12 and 13 ka, the maps show a relatively large Paleoindian population in Atlantic Canada. The Debert site (Nova Scotia) is the only dated site in the CARD at this time, but the dates from this site are consistent with one another and are accepted in archaeological literature (MacDonald, 1968). The RFPEs decreased following this, as temperatures and dryness increased

during the Younger Dryas (Newby et al., 2005). RFPE values fluctuated and began to increase around 10.5 ka, consistent with archaeological evidence from Debert (MacDonald, 1968). At 8.5 ka, RFPEs tripled in Newfoundland and Labrador (**Fig. 3.2**) with the onset of early Maritime Archaic settlement along the coast (5.5 to 8 ka (Bell and Renouf, 2003)). Estimates are relatively low in the Maritime Provinces as many sites were presumably inundated by sea level rise.

There is a slight increase in RFPEs during the Middle Archaic in eastern Quebec and central Labrador, and at 4.5 ka there was a migration into these provinces. A larger population appears to have spread across eastern Hudson Bay and the coastline of Labrador between 3.2 and 4 ka during the late Archaic. The influence of Arctic-adapted Paleoeskimo populations (e.g., Groswater and Dorset sites), the Meadowood complex and late Woodland populations are observed during the last 3 ka, although these values may be overestimated due to a higher number of sites being divided by a low sampling intensity in more recent times. RFPEs decreased in central Quebec but remained high in Newfoundland, Labrador and the area surrounding New Brunswick until 1.5 ka, when estimates increase over Atlantic Canada.

### *3.6.6 Central and Western United States*

The oldest maps contain two centers of relatively high estimates in Arizona and on the Colorado-Kansas border, areas where paleoecological records confirm the presence of Paleoindians (Bahre, 1995). At 12 ka, these centers shifted to Texas, although this area remains anomalous for the entire record due to the relatively large number of dates in the early and mid-Holocene compared to the more recent past (**Fig. A1.2**). By 10 ka, an increase appeared in California and persisted until 7.5 ka, which is likely associated with offshore communities exploiting marine resources (Newsome et al., 2004). Also during this time there was a slight increase in RFPEs in Montana which became joined with larger populations living in the ice-free corridor. The latter half of the Holocene was characterized by an intensification in Idaho, which shifted towards the coast at 4 ka.

## **3.7 Discussion**

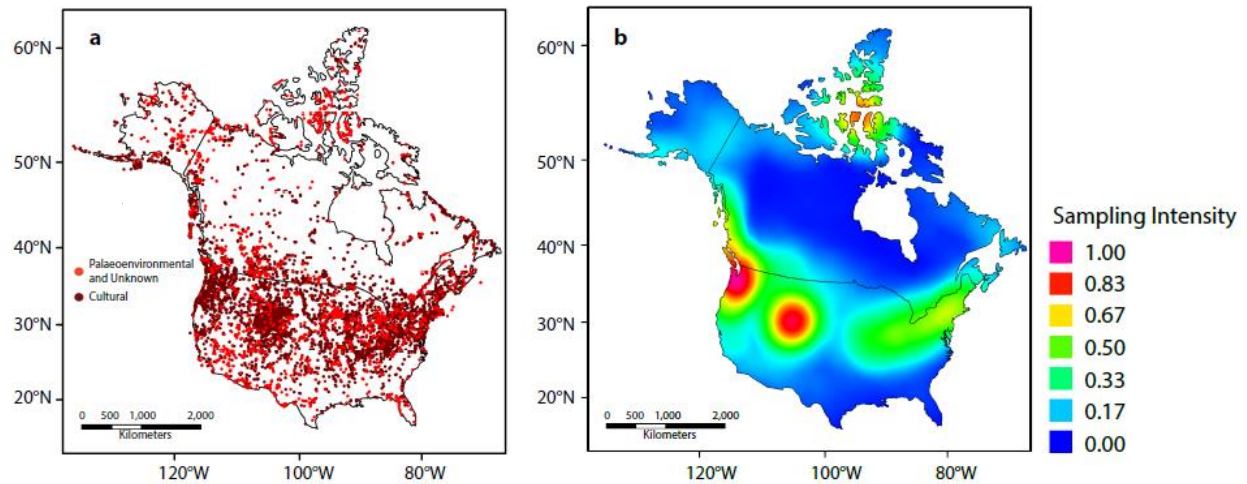
If known issues with the database are accounted for, RFPEs correspond with archaeological, paleoenvironmental and genetic evidence for human migration and dispersal into and across North America (**Table 3.1**). The maps (**Fig. 3.2**, **Fig. A1.2** and **Video A2.1**) show the importance of proximity to the coast for human settlement as well as areas where human activity appears to have been continuous throughout much of the Holocene (northeastern North America, Alaska, western Plains). The maps also show examples where sampling bias (i.e., the quality of the CARD data) affects the RFPEs. For example, the Kivalliq Region (west of Hudson Bay) and the Ungava Region show an overestimation beginning ca. 4 ka. This is connected to selective archaeological work in the region and an extensive research program into the local pre-contact history (Gordon, 1975; Gordon, 1976; Gordon, 1996) combined with comparatively few dates in the CARD prior to this interval.

We have shown that aggregated continental data exhibit macro-demographic patterns comparable to known regional histories, in addition, use of the CARD permits a continental-scale analysis. At 12.5, 9, 4.5, and 1 ka, the overall pattern is close to generally accepted culture-historical reconstructions (Bell and Renouf, 2003; Newsome et al., 2004; Erlandson et al., 2008; Simek et al., 2012; Goebel et al., 2013; Bennike, 2014). There was an overall increase in continental population estimates from 12.5 ka, as expected, followed by a decrease between 0.5 and 1 ka due partly to the taphonomic correction, but also reflecting the decreased use of  $^{14}\text{C}$  dating in the context of historical methods and a lessened archaeological interest in recent periods. The distribution at ca. 0.8-1.2 ka provides a broad example of both current and ancestral population distribution since comparatively little migration has occurred since then, with the exception of the Athapaskan migration into the southwest.

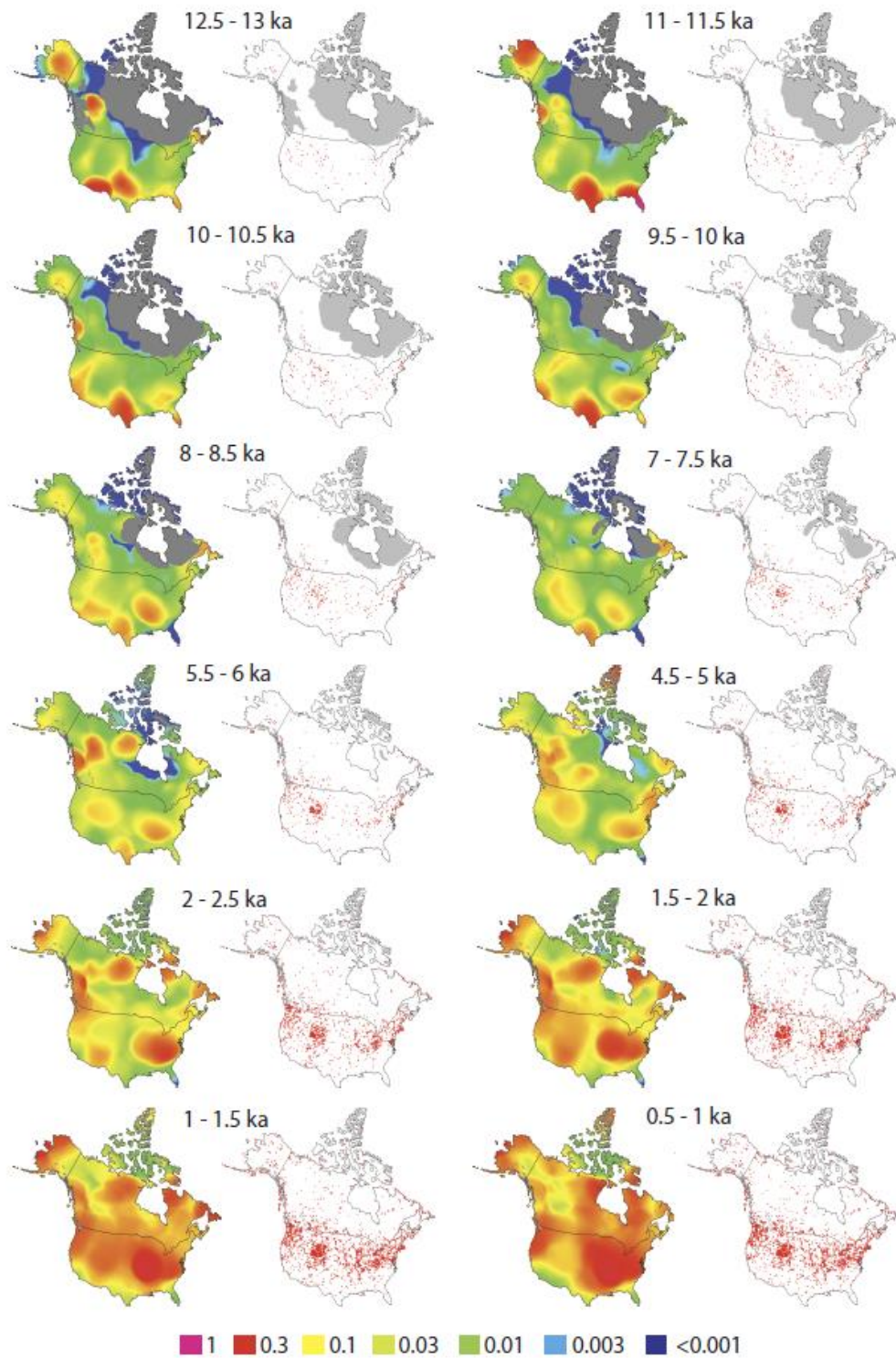
These results illustrate the value in applying advanced statistical methods to aggregate  $^{14}\text{C}$  data from archaeological databases. The accumulation of known  $^{14}\text{C}$  data into facilities such as the CARD should be a priority, as this offers an opportunity to summarize previous research and provide a context for regional and local studies. “Dates as data” (Rick, 1987; Steele, 2010) is a relevant means of investigating and modelling demographic, and thus historical, trends across long periods of time and continental areas. The kernel density method, in combination with statistical approaches to reduce sampling and taphonomic bias, holds potential for accommodating known issues in  $^{14}\text{C}$  datasets that cannot be handled by other statistical analyses developed so far (e.g., averaging the number of observations). These results suggest that the CARD is a highly useful archive of paleodemographic data that can be used to investigate subjects such as migration routes into and across North America as well as a valuable tool for studies linking anthropogenic impacts with post-ice age faunal extinctions, ecosystem decline and changing environmental and climatic conditions.

## **Acknowledgments**

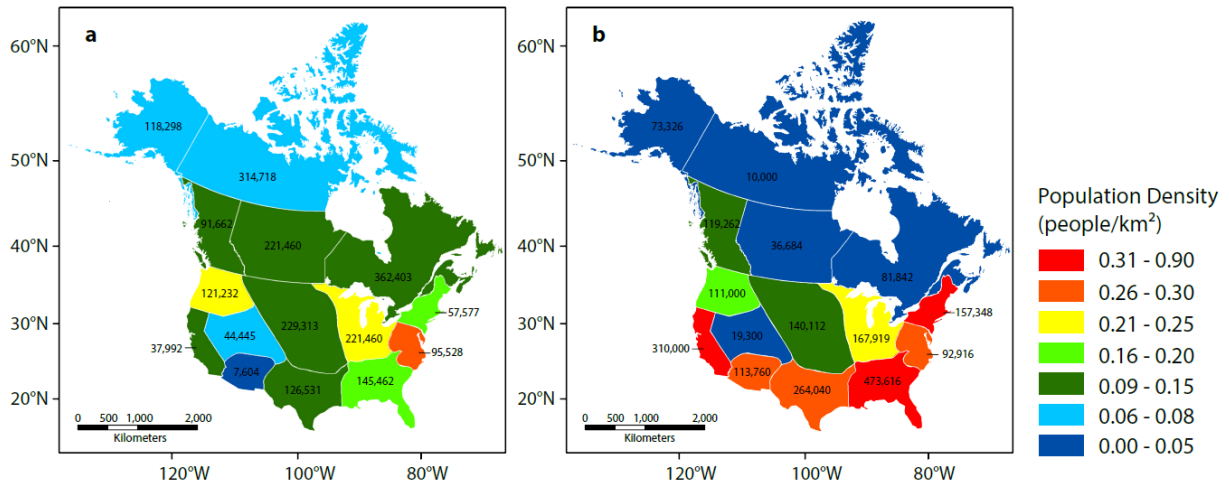
CARD was created by Richard Morlan of the Canadian Museum of Civilization (now the Canadian Museum of History). It was transferred to the University of British Columbia’s Laboratory of Archaeology in 2014. The CARD data are available at <http://www.canadianarchaeology.ca/>. K.G. and R.K. received support from NSERC Discovery Grants. The German Academic Exchange Service (DAAD) provided travel funds for B.K. and V.S. We thank T. Merk for assistance with **Figure 3.3** and T. Brereton for participating in discussions.



**Figure 3.1** Locations of the archaeological sites in the CARD and the sampling intensity base map. Distribution of the CARD sites including dated cultural, paleoenvironmental and unknown material (a). Sampling intensity base map produced using all geographically distinct cultural and paleoenvironmental dates spanning the full age range in the CARD (b), such that multiple dates at the same location are considered as one.



**Figure 3.2** Radiocarbon Frequency Population Estimate (RFPE) maps. A subset of RFPEs and the distribution of the associated CARD dates for selected time intervals between 0.5 and 13 ka. Grey areas show the extent of the ice sheets. The entire series of RFPE maps is shown in **Figure A1.2**.



**Figure 3.3** Comparison between population densities at European contact estimated using the RFPE technique and ethnohistorical data. a, Population density estimates based on the distribution of the CARD sites from 0.4 to 0.6 ka. b, Population density estimates based on work by Ubelaker (1992).

**Table 3.1** Summary of the major patterns seen in the RFPE maps and archaeological history for the different regions mentioned in the text with references.

<b>Region</b>	<b>Time Period (ka cal BP)</b>	<b>RFPE Patterns</b>	<b>Archaeological History</b>	<b>References</b>
Alaska	7.5 - 13	Increase in north, central	Fluted point technology	Goebel et al. 2013
	6	Increase on Aleutian Islands	Aleut colonization	West et al., 2007
	4.5 - 5.5	Increase at northern and southwestern coast	Migration from Siberia, Arctic Small Tool Tradition	Gilbert et al., 2008; Hatfield, 2010
	1	Increase	Thule	Gilbert et al., 2008; Dumond, 2000; Park, 2012
Canadian Arctic	2.5	Increase in southern Arctic	Dorset Culture	
	modern-0.75	Increase along southern Thule migration route	Thule migration beginning ca. 0.8	Bennike, 2014
British Columbia/ Corridor	11.5	Increase along west coast	Maritime community occupation	Erlandson et al., 2008
	5.5	Second increase along west coast and ice-free corridor	Shift in technology	Maschner, 2012
Eastern United States	8.5	Increase in Missouri	Paleoindian occupation	Hajic et al., 2007
	4.9	Increase in mid-Atlantic and New England	Late Archaic and Early Woodland occupation	Simek et al., 2012
	modern-2	Increase in Cahokia	Population rise with maize agriculture	Hart and Lovis, 2013
Atlantic Canada	10.5	General increase	Debert Culture	MacDonald, 1968
	8.5	Large increase in Newfoundland and Labrador	Early Maritime Archaic	Bell and Renouf, 2003
Central/Western United States	7.5	Increase in California	Marine exploitation	Newsome et al., 2004

## **4 Estimation of spatio-temporal correlations of prehistoric population and vegetation in North America**

Bjoern Kriesche, Michelle Chaput, Rafal Kulik, Konrad Gajewski and Volker Schmidt

### **4.1 Abstract**

We discuss a simple methodology to enable a statistical comparison of human population with the vegetation of North America over the past 13,000 years. Nonparametric kernel methods are applied for temporal and spatial smoothing of point data obtained from the Neotoma Paleoecology Database and the Canadian Archaeological Radiocarbon Database, which results in sequences of maps showing the development of population and different plant taxa during the Holocene. The estimation of smooth spatial and spatio-temporal cross-correlation functions is proposed in order to detect relationships between population and vegetation in fixed time intervals. Furthermore, the effects of varying environment on demographic changes as well as potential impacts of populations on plant taxa over time are analyzed. Pointwise confidence bands for cross-correlation functions are computed using nonparametric resampling methods to assess the significance of obtained results. Considering the example of oak, an interpretation of our results for eastern North America shows the value of this methodology.

Keywords: Space-time analysis, Nonparametric estimation, Cross-correlation function

## 4.2 Introduction

### 4.2.1 Motivation

In North America, European colonists had an enormous impact on the landscape since their arrival 400 years ago. For example, they deforested much of the landscape of eastern North America by the early 20<sup>th</sup> century, although much has since regrown (Williams, 1989b). Determining human impacts on the landscape of North America before historical times, however, is more complicated, and is done using (sub)fossil data.

Two opposing viewpoints have been proposed about the nature of the vegetation of North America prior to the arrival of Europeans (e.g., Denevan, 1992; Vale, 2002). A first view is that North America in CE1492 was a “pristine landscape” with low population densities and vegetation largely unaltered by human activities. In this case, the primary factor causing changes in population numbers or in cultural/technological expression would have been climate and environmental changes. However, an alternate view suggests that in large parts of the Americas, First Nations used fire to clear the forest and maintain the prairies, practiced more extensive agriculture, and in general altered the forests and plains through extensive land use.

Although local and regional studies have shown human impacts on the landscape (Delcourt and Delcourt, 2004; Munoz and Gajewski, 2010), it is important to understand this interaction at continental scales (e.g., Jäger and Neuhäusl, 1994). As this has not yet been discussed in the literature, we suggest using simple and intuitive statistical strategies to examine this potential association over the course of the past 13,000 years using centennial-scale time intervals.

### 4.2.2 Population intensity maps

We investigate space-time correlations of population density and vegetation abundances in prehistoric North America based on fossil data. While a methodology for the estimation of vegetation abundance is proposed in section 4.5, we use population estimates that were derived in a previous study (Chaput et al., 2015; chapter 3). The original data come from the Canadian Archaeological Radiocarbon Database (CARD), which is a compilation of radiocarbon measurements that indicate the ages of samples from archaeological sites in North America (e.g., Gajewski et al., 2011). The CARD consists of 35,905 calibrated radiocarbon dates, with 29,609 of them containing suitable geographical, chronological and descriptive information linking them to culturally-distinct human activity. The suggested approach relies on the well-accepted assumption that the frequency of radiocarbon dates is, after accounting for sampling and taphonomic biases, proportional to population activity (the dates as data approach, e.g., Steele, 2010).

At first, a sequence of 121 500-year time intervals was selected, which range from 500-1,000 BP, 600-1,100 BP, 700-1,200 BP up to 12,500-13,000 BP. For each interval, all corresponding radiocarbon dates are selected and spatially smoothed intensities of date counts are computed for

all locations of the North American continent not permanently covered by ice using a nonparametric kernel density estimator with a two-dimensional Epanechnikov kernel and a globally fixed bandwidth of 600 km (the choice of this methodology is justified in Chaput et al. (2015)).

In order to account for spatially inhomogeneous sampling strategies, we also estimate a sampling intensity based on all geographically distinct data locations in CARD (from all intervals) and divide intensities of date counts of each interval by sampling intensities. The resulting population intensities can be understood as (smoothed) numbers of dates per sampling site (falling in the considered time interval) and are interpreted as indicators of population density. In addition to sampling biases, the method also accounts for biases occurring due to taphonomic loss and boundary effects. Furthermore, some temporal smoothing is included in the method.

Two examples of estimated population intensity maps are shown in [Figure 4.1](#). The maps show clear patterns of population activity and we consider that the results capture paleodemographic trends. This is justified by the observation that patterns correlate well with previous archaeological interpretations of population change across the North American continent during the Holocene that were derived partly from radiocarbon dates, but also other data sources and inferences.

#### *4.2.3 Outline*

In section 4.3 we introduce the database used to obtain estimates of vegetation abundance of different taxa over the past 13,000 years, including a discussion of calibrated radiocarbon ages. Section 4.4 describes an intuitive procedure for temporal smoothing and interpolation of pollen percentages to the target ages needed for the correlation analyses. The preparation of spatially smooth vegetation intensity maps is discussed in section 4.5, along with a depiction of examples for one taxon and a comparison to existing results from the literature. In section 4.6 we describe a simple approach to the estimation of spatial cross-correlations of vegetation abundances with population densities and the cross-correlations of changes in vegetation and populations at various temporal lags. This section also introduces a method to compute nonparametric confidence bands of estimated cross-correlations. In section 4.7 we apply the presented methods to a selected taxon and illustrate the insights obtained from this analysis.

### **4.3 The Neotoma Paleoecology Database**

Space-time data of prehistoric pollen abundance used in this study are obtained from the Neotoma Paleoecology Database (Grimm, 2008). Neotoma is a comprehensive compilation of fossil data from the Holocene, Pleistocene, and Pliocene for more than 8,400 sites worldwide.

### 4.3.1 Calibration of radiocarbon ages

For the estimation of vegetation intensities we use pollen data, which are typically acquired as follows. Consider a set of sites (lakes) on the North American continent, at which pollen data are available in Neotoma. Samples are taken along a sediment core for a sequence  $d_0 < \dots < d_n$  of  $n + 1$  depths. Each sample contains a certain number of fossil pollen, which are counted and classified. Next, the age of the sample at each depth needs to be determined. Since radiocarbon dating is rather expensive, it is not used to estimate the ages of all samples. Instead, radiocarbon ages are only determined for a subset  $\{d_{i1}, \dots, d_{ij}\} \subset \{d_0, \dots, d_n\}$  and ages for depths  $d \in \{d_0, \dots, d_n\} \setminus \{d_{i1}, \dots, d_{ij}\}$  are computed using interpolation methods. Ages obtained from the radiocarbon method (or from interpolation) need to be calibrated to be measurable in calendar years BP. As the manual use of standard calibration curves is not suitable for application to such a large number of dates as considered in our study, we suggest converting radiocarbon ages into calibrated ages using a smoothed radiocarbon calibration curve (Grimm, 2008; Fig. 4.3). This simplified calibration is not exact but in Grimm (2008) it is estimated that the probability of the occurring error being less than or equal to 25 years is 0.47 and the probability of the occurring error being less than or equal to 200 years is 0.97. Since this is clearly below the temporal scale of this study, it is extremely unlikely that a significant bias is introduced by using the simplified calibration curve. The result of calibration is a sequence  $a_0, \dots, a_n$  of calibrated radiocarbon ages (the unit being cal years BP; for simplicity we write BP) that correspond to the samples taken at depths  $d_0, \dots, d_n$ .

The correct procedure would be to calibrate radiocarbon ages at depths  $d_{i1}, \dots, d_{ij}$  first, resulting in calibrated ages  $a_{i1}, \dots, a_{ij}$ , and to interpolate ages for depths  $d \in \{d_0, \dots, d_n\} \setminus \{d_{i1}, \dots, d_{ij}\}$  afterwards based on  $a_{i1}, \dots, a_{ij}$ . For the majority of sites, however, ages are interpolated based on uncalibrated ages at  $d_{i1}, \dots, d_{ij}$  before feeding data into Neotoma, and we then calibrate the ages of all  $n + 1$  depths  $d_0, \dots, d_n$  using the smoothed calibration curve.

In general, this exchange of calibration and interpolation leads to an error in calibrated ages. Unfortunately, for those sites at which ages are interpolated before calibration, it does not seem possible to identify, which (uncalibrated) ages were obtained from dating and which from interpolation, making it impossible to eliminate this error. To investigate whether these calibration errors will significantly bias the results of our study, we perform the following comparison. We select 22 independent test samples from the literature, each consisting of a sequence  $d_0, \dots, d_n$  of depths in cm and a sequence  $\bar{a}_0, \dots, \bar{a}_n$  of uncalibrated radiocarbon ages, with  $n$  varying between 4 and 17. For each test sample, we consider the sequence  $\{\delta_1, \dots, \delta_k\} \subset [d_0, d_n]$ , which contains all depths between  $d_0$  and  $d_n$  that are a multiple of 5 cm. We first determine radiocarbon ages for depths  $\delta_1, \dots, \delta_k$  by applying linear interpolation based on  $\bar{a}_0, \dots, \bar{a}_n$  and afterwards calibrate interpolated ages using the smoothed calibration curve.

Next, we first calibrate  $\bar{a}_0, \dots, \bar{a}_n$  and then determine calibrated ages for  $\delta_1, \dots, \delta_k$  by applying linear interpolation. This results in two calibrated radiocarbon ages  $a$  and  $a'$  for each depth  $\delta \in \{\delta_1, \dots, \delta_k\}$ , i.e., in two depth-age curves for each test sample. We find that the vast majority of

absolute errors is smaller than 100 years with absolute errors of more than 300 years occurring extremely rarely (Fig. 4.2). Moreover, the largest errors occur for ages older than 13,000 BP, which are not considered in our analysis. In summary, since the observed differences are small compared to the temporal scale of this study, we can assume that the errors occurring from exchanging calibration and interpolation of radiocarbon ages are negligible in the following.

#### 4.3.2 Data selection

In order to access the Neotoma database for automatic data selection and processing we use the R package *neotoma* (Goring et al., 2015). First, we select all sites from Neotoma which are labeled with the geopolitical ID “Canada” or “United States”. Then, all datasets associated with the sites obtained are loaded and those with the dataset type “pollen” and the collection type “composite” or “core” are selected. Each dataset contains a sequence of  $n + 1$  samples that are taken at depths  $d_0 < \dots < d_n$  and the corresponding ages  $a_0, \dots, a_n$  in cal years BP. The samples contain count data for different taxa. Since we are interested in pollen counts only, we select data with taxon group “vascular plant”, variable element “pollen” or “spore” and variable unit “number of identified specimen (NISP)”.

Count data sometimes cannot be compared directly across sites due to different levels of taxonomic resolution used by the researchers contributing to Neotoma. For example, one analyst might discriminate sub-taxa of *Acer* (maple), such as *Acer rubrum* (red maple) or *Acer saccharinum* (silver maple), while another might simply identify *Acer* to the genus level. To provide comparability, taxa and corresponding pollen data are aggregated using the standardization list suggested in Williams and Shuman (2008).

Finally, the usual practice is to compute relative pollen abundances for all sites, ages and taxa based on pollen counts for all plant taxa, which are much easier to interpret and compare. Altogether, we obtain a set of 1,151 sites, each of which contains a sequence  $a_0, \dots, a_n$  of calibrated radiocarbon ages together with the corresponding relative pollen abundances of 64 taxa.

### 4.4 Temporal interpolation and smoothing of pollen abundances

In order to be able to estimate spatial vegetation intensity maps, relative pollen abundances need to be available at all 1,151 sites with Neotoma data simultaneously for the same years. Since relative pollen abundances of one taxon might vary considerably even during short time periods, some of which are random sampling errors, it is preferable to use smoothing methods over interpolation. In the following, we consider a fixed site, a fixed taxon and the corresponding relative pollen abundances.

With  $(a_0, p_0), \dots, (a_n, p_n)$ , where  $0 \leq a_0 < \dots < a_n$ , we denote the ages of the available samples and the corresponding relative pollen abundances (taking values in  $[0, 1]$ ) for the chosen site and taxon. We interpret  $(a_0, p_0), \dots, (a_n, p_n)$  as (sorted) realizations of some independent and

identically distributed random vectors  $(A_0, P_0), \dots, (A_n, P_n)$  taking values in  $[0, \infty) \times [0, 1]$ . It seems almost impossible to find a parametric representation describing the dependence of random pollen abundances  $P_0, \dots, P_n$  on random ages  $A_0, \dots, A_n$  sufficiently well, which is why the following (nonlinear) relationship is assumed:

$$P_i = p(A_i) + \varepsilon_i \quad \text{for } i = 0, \dots, n, \quad [1]$$

where  $p : [0, \infty) \rightarrow [0, 1]$  with  $p(a) = E(P_0 | A_0 = a)$  for  $a \geq 0$  is the conditional expectation function of  $P_0$  given  $A_0$  and  $\varepsilon_0, \dots, \varepsilon_n$  denote random errors with  $E \varepsilon_0 = \dots = E \varepsilon_n = 0$ . In order to estimate  $p$  by using kernel smoothing, we consider the time intervals  $I_1, \dots, I_n$ , where  $I_i = [a_{i-1}, a_i)$  for  $i = 1, \dots, n-1$  and  $I_n = [a_{n-1}, a_n]$ . We denote by  $h_i = a_i - a_{i-1}$  the length of interval  $I_i$  for  $i = 1, \dots, n$ .

The bandwidth  $h$ , which controls the degree of smoothing in kernel estimators, should be chosen not smaller than the maximum  $\max\{h_1, \dots, h_n\}$ . However, for some sites and taxa, long time periods  $I_i$  without a sample can occur leading to a bandwidth that can cause oversmoothing in other intervals  $I_j, j \neq i$ , eliminating too many details. To avoid such effects, we only consider those intervals  $\{I_{i1}, \dots, I_{ik}\} \subset \{I_1, \dots, I_n\}$  that have a length of not more than 2,000 years and define  $I = I_{i1} \cup \dots \cup I_{ik}$  and  $h = \max\{h_{i1}, \dots, h_{ik}\}$ . For all  $a \in I$  an estimate  $\hat{p}(a)$  of  $p(a)$  can be determined using a (one-dimensional) Nadaraya-Watson estimator with the bandwidth  $h$  (e.g., Wand and Jones, 1995), by

$$\hat{p}(a) = \frac{\sum_{i=0}^n p_i K_1^G \left( \frac{a - a_i}{h} \right)}{\sum_{i=0}^n K_1^G \left( \frac{a - a_i}{h} \right)} \quad \text{for all } a \in I, \quad [2]$$

Where  $K_1^G : \mathbb{R} \rightarrow [0, \infty)$  denotes the one-dimensional Gaussian kernel defined as

$$K_1^G(u) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2}u^2 \right) \quad \text{for all } u \in \mathbb{R}. \quad [3]$$

Motivation for choosing the Gaussian kernel is that it has an unbounded support and is thus expected to provide a smooth estimate even in regions of sparse data. However, for  $a \notin I$ , using the Nadaraya-Watson estimator occasionally results in sudden decreases or increases making it an inappropriate choice. Therefore, we alternatively suggest estimating  $p$  using linear interpolation, which leads to an estimate  $\bar{p}(a)$  of  $p(a)$  for all  $a \in [a_0, a_n]$ . Finally, we set  $\hat{p}(a) = \bar{p}(a)$  for  $a \in [a_0, a_n] \setminus I$ .

The proposed methodology is applied to 10 selected taxa. These include major taxa and those representative or characteristic of the major biomes (*Quercus* (oak): deciduous forest; *Picea* (spruce) and *Pinus* (pine): boreal forest; Poaceae (grass): Prairie), taxa used or potentially

affected by human use (*Carya* (hickory), *Juglans* (walnut, butternut), *Castanea* (chestnut)), disturbance taxa (*Populus* (poplar, aspen)), or representatives of mature forests, i.e., non-disturbance (*Fagus* (beech), *Acer* (maple)). For example, if Native Americans extensively used fire to manage the forests, we would expect a positive association of human population densities and oak tree abundance (Abrams and Nowacki, 2008). If Native Americans managed the forests and planted useful trees, then there should be a spatial correlation of population density and nut trees, as seen on a local level (Wycoff, 1991). Munoz and Gajewski (2010) showed an increase in *Populus* in southern Ontario associated with the introduction of agriculture, and we wanted to see if this occurred at a larger scale. On the other hand, we would also expect no association of spruce with human population densities, as human populations were assumed to be relatively low in boreal forests.

In **Figure 4.3** sample data of relative pollen abundance from *Neotoma* are shown for two sites in North America together with estimates  $\{\hat{p}(a), a \in I\}$  and  $\{\bar{p}(a), a \in [a_0, a_n]\}$  obtained using a Nadaraya-Watson estimator and linear interpolation, respectively. We find that choosing the smoothing parameter  $h$  as explained above leads to estimates that capture the temporal development very precisely (even for abrupt increases or decreases, see, e.g., the cyan curve in **Fig. 4.3**, top), whereas noise is also eliminated quite well, see, e.g., the yellow and the blue curve in **Fig. 4.3** (top). In **Figure 4.3** (bottom) we have that  $I \neq [a_0, a_n]$  since for ages between 3,500 BP and 6,500 BP interpolation is preferred over smoothing due to sparse data.

## 4.5 Vegetation intensity maps

### 4.5.1 Nonparametric estimation of vegetation intensity maps

We now address the question how spatially smooth vegetation maps can be determined. We consider the sequence of years 750 BP, 850 BP, . . . , 12,750 BP, which corresponds to the midpoints of the 500-year intervals considered for the nonparametric estimation of population intensity maps sketched in section 4.2.2. The estimation procedure is applied to each year and taxon separately. Let  $y$  be the considered year, i.e.,  $y \in \{750, 850, \dots, 12,750\}$ , and let  $W$  denote the set of all locations on the North American continent that are not covered by ice in year  $y$ .

At first, we need to identify at which sites pollen abundances for the year  $y$  are available. In order to do that we consider the 1,151 sites with *Neotoma* data and check for each site individually whether  $\mathbf{y} \in [a_0, a_n]$ , with  $a_0$  and  $a_n$  being the site-specific minimal and maximal available calibrated radiocarbon ages introduced in section 4.3.1. This results in a sequence  $(s_1, \pi_1), \dots, (s_m, \pi_m)$ , with  $m \leq 1,151$ , where  $s_1, \dots, s_m$  denote the sites in North America with available pollen abundances  $\pi_1, \dots, \pi_m$  in year  $y$  that were estimated using the kernel estimator in equation 2 or linear interpolation (for time intervals with sparse data). Typically,  $m$  is higher in more recent years since for older time periods there is a larger number of sites with  $a_n < y$ .

As relative pollen abundances may vary considerably among closely located sites (due to noise)

and as we are instead interested in large scale patterns of spatial vegetation intensities, it seems more suitable to apply nonparametric spatial smoothing methods than interpolation methods such as kriging (e.g., Diggle and Ribeiro Jr., 2007; Cressie and Wikle, 2011). Furthermore, since vegetation intensity maps are compared to population intensity maps in section 4.6, we aim to provide as many similarities in the corresponding estimation procedures as possible.

We suppose that  $(s_1, \pi_1), \dots, (s_m, \pi_m)$  can be interpreted as realizations of some independent and identically distributed random vectors  $(S_1, \Pi_1), \dots, (S_m, \Pi_m)$  with values in  $W \times [0, 1]$ . Again, it seems impossible to find a parametric representation that models the relationship between sites  $S_1, \dots, S_m$  and relative pollen abundances  $\Pi_1, \dots, \Pi_m$  sufficiently well, which is why we suppose that

$$\Pi_i = \pi(S_i) + \varepsilon'_i \quad \text{for } i = 1, \dots, m, \quad [4]$$

where  $\pi : W \rightarrow [0, 1]$  with  $\pi(s) = E(\Pi_1 | S_1 = s)$  for  $s \in W$  denotes the conditional expectation function of  $\Pi_1$  given  $S_1$  and  $\varepsilon'_1, \dots, \varepsilon'_m$  are random errors with  $E \varepsilon'_1 = \dots = E \varepsilon'_m = 0$ .

An extensive literature has demonstrated that pollen can be used as a quantitative index of past plant abundance (Birks and Birks, 1980), which is why the field  $\{\pi(s), s \in W\}$  is considered as a map of expected intensities of the considered taxon in the following. However, as pollen percentages may be over- or underrepresented in comparison with the abundance of the plant on the landscape, we are not reconstructing vegetation, but rather mapping the change in time and space.

Pollen representation should not affect our analysis as we are relating the spatial pattern of the pollen, and its change over time, to the spatial pattern of population abundance later on. We suggest to estimate  $\{\pi(s), s \in W\}$  based on  $(s_1, \pi_1), \dots, (s_m, \pi_m)$  using a two-dimensional Nadaraya-Watson estimator (e.g., Härdle et al., 2004). Accordingly, an estimate  $\hat{\pi}(s)$  of  $\pi(s)$  is determined by

$$\hat{\pi}(s) = \frac{\sum_{i=1}^m \pi_i K_2^E(s, s_i, h)}{\sum_{i=1}^m K_2^E(s, s_i, h)} \quad \text{for all } s \in W, \quad [5]$$

where  $K_2^E : W \times W \times (0, \infty) \rightarrow [0, \infty)$  with

$$K_2^E(s_1, s_2, h) = \left(1 - \frac{d(s_1, s_2)^2}{h^2}\right) \mathbf{1}_{[0, h)}(d(s_1, s_2)) \quad \text{for all } s_1, s_2 \in W \quad [6]$$

denotes a scaled two-dimensional Epanechnikov kernel,  $h > 0$  is the bandwidth controlling the degree of smoothing and  $d(s_1, s_2)$  denotes the geographic distance of two locations  $s_1, s_2 \in W$  (Diggle and Ribeiro Jr., 2007, Sec. 2.7).

We set  $h = 600$  km in equation 5 to ensure comparability to the population intensity maps described in section 4.2.2, which is also the reason for choosing the Epanechnikov kernel over a kernel with unbounded support (such as the Gaussian kernel). However, in contrast to the estimation of population intensity maps, we do not need to account for sampling biases and errors occurring due to taphonomic loss. On the one hand, the denominator in equation 5 prevents the estimator from being influenced by inhomogeneous sampling strategies. On the other hand, relative pollen abundances do not contain taphonomic biases as pollen does not degrade over the time-scale of our study. Computing  $\{\hat{\pi}(s), s \in W\}$  for all  $y \in \{750, \dots, 12,750\}$  results in a sequence of estimated vegetation intensity maps.

#### 4.5.2 Computation of taxon ranges

Most taxa analyzed in this paper have regions of typical occurrence (e.g., the eastern and southern U.S. for *Quercus*). Only in those regions do the estimated vegetation intensities have significantly positive values, whereas in the remaining areas intensities are zero or very close to zero (due to few pollen grains being transported by wind or caused by data and measurement errors). When comparing population and vegetation intensity maps in section 4.6, it seems reasonable to only take into account the taxon range  $\xi \subset W$ . For example, if correlations between population activity and the intensity of *Quercus* are analyzed, only regions in the eastern and southern U.S. should be considered. In the western U.S. and in Canada (i.e., outside the range of *Quercus*), estimated vegetation intensities will be very close to zero although population activity is found in various regions, which will dilute correlation results.

For that purpose, we suggest an approach to determine (temporally varying) estimates of the taxon range  $\xi$ . Let  $\{\hat{\pi}(s), s \in W\}$  be the estimated vegetation intensity map of a selected taxon for a given year  $y$ . Then, we suggest computing an estimate  $\hat{\xi}$  of the taxon range  $\xi \subset W$  by

$$\hat{\xi} = \{s \in W : \hat{\pi}(s) \geq u \cdot \max\{\hat{\pi}(s), s \in W\}\}, \quad [7]$$

with a suitable threshold  $u \in (0, 1)$ , i.e., the estimated range  $\hat{\xi}$  consists of those locations on the North American continent, at which the local vegetation intensity is at least  $u$  times the global maximum of the vegetation intensity map.

In order to determine an optimal choice for the threshold  $u$ , the taxon range is estimated for the most current year ( $y = 750$  BP) for thresholds 0.05, 0.1, 0.15, 0.2, 0.25 and 0.3, and a visual comparison to the taxon's modern range is made (Thompson et al., 1999). We observe that for all 10 taxa considered in section 4.4 the threshold  $u = 0.2$  provides the best overall match. Since the estimate  $\hat{\xi}$  as defined in equation 7 always depends on the maximum estimated vegetation intensity of the current year (and thus changes over time), we suppose that the suggested approach is also able to capture the typical taxon ranges for older years, which is confirmed by the examples shown in section 4.5.3.

### 4.5.3 Interpretation of results

As an example illustrating the results of the methodology discussed here, we use the history of *Quercus* (oak) over the past 13,000 years. Oak is the characteristic species of the eastern deciduous forest and is a general indicator of the extent of this ecosystem. It is a major source of food for both humans and game animals (McShea and Healy, 2002). Prior to 12,000 BP, high values of *Quercus* were restricted to Florida and the southern Gulf States. The range expanded and moved north and westward after 10,750 BP, see [Figure 4.4](#) (top). At a smaller scale, the maps capture the lower values of *Quercus* in the upper slopes of the Appalachian Mountains, and *Quercus* became more abundant especially west of this mountain chain. During this time, *Quercus* remained abundant across all southern states.

Beginning around 9,000 BP, *Quercus* decreased in the south, and remained at higher values in a band across the mid-Atlantic States, see [Figure 4.4](#) (center). During the period between 8,000-7,000 BP, maximum oak abundance remained in the region. Note that many sites have high values of oak pollen to the north of the region of maximum abundance, but other sites in proximity have lower values, so the map shows a decrease in regional abundance in this region. This pattern remained for the next few thousand years, followed by a slight decrease in abundance in the past 1,500 years, see [Figure 4.4](#) (bottom).

Maps depicting the abundance of the pollen of tree taxa through time for eastern North America have been produced several times since the 1970s, as the database increased in size and methods to produce the maps have evolved, see the references in Williams et al. (2004). The latest version (Williams et al., 2004) used the North American Pollen Database with 759 sites, which has been incorporated and expanded in Neotoma. They made maps every 1,000 to 2,000 years using tri-cubic distance weighting to average pollen data from a  $300 \text{ km} \times 300 \text{ km} \times 500 \text{ m}$  (vertical) window to a 50 km grid. There is a very close visual correspondence of our maps to those of Williams et al. (2004); all of the features discussed above are seen in both sets of maps suggesting that the methodology proposed in the present paper is able to provide reliable indicators of vegetation abundance.

A more sophisticated approach to the inference of local vegetation intensities for tree taxa is discussed in Paciorek and McLachlan (2009), where Bayesian models and methods have been also used to estimate uncertainties in obtained results. However, we do not consider this to be necessary in our study as estimated vegetation intensity maps correspond particularly well with those of the existing literature, see above. Furthermore, the proposed methodology is closely related to that we used when estimating population intensity maps.

## 4.6 Correlation analysis of population and vegetation intensities

### 4.6.1 Cross-correlations of population and vegetation intensity maps

We aim to investigate whether significant correlations between vegetation and population intensities can be found. In the following, we again fix a taxon and a year  $y \in \{750, 850, \dots, 12,750\}$  and consider the corresponding map  $\{\hat{\pi}(s), s \in \mathcal{X}\}$  of estimated vegetation intensities

restricted to the estimated taxon range  $\hat{\xi}$ . Furthermore, let  $\{\hat{\lambda}(s), s \in \hat{\xi}\}$  denote the corresponding estimated population intensity map for the 500-year time period  $[y - 250, y + 250]$ , which is obtained using a similar nonparametric smoothing method (Chaput et al., 2015; chapter 3). In statistical estimation theory, it is common to model estimators as random elements, which is why  $\hat{\xi}$  is considered to be a realization of some random closed subset  $\Xi$  of  $W$  (Chiu et al., 2013). Furthermore, the estimates  $\{\hat{\pi}(s), s \in \hat{\xi}$  and  $\{\hat{\lambda}(s), s \in \hat{\xi}\}$  can be interpreted as realizations of two random fields  $\Pi = \{\Pi(s), s \in W\}$  and  $\Lambda = \{\Lambda(s), s \in W\}$  restricted to  $\Xi$ .

To infer probabilistic properties we assume that  $\Pi$  and  $\Lambda$  are jointly second-order stationary and isotropic (Wackernagel, 2003). This means that  $E \Pi(s) = \mu_{\Pi}$  and  $E \Lambda(s) = \mu_{\Lambda}$  for all  $s \in W$ , and that for any pair of locations  $s_1, s_2 \in W$  the covariances  $\text{cov}(\Pi(s_1), \Pi(s_2))$ ,  $\text{cov}(\Lambda(s_1), \Lambda(s_2))$  and  $\text{cov}(\Pi(s_1), \Lambda(s_2))$  only depend on the geographic distance  $d(s_1, s_2)$  between  $s_1$  and  $s_2$ . Clearly such assumptions are hardly realistic on a continental scale. However, when estimating cross-correlation functions we only consider the restriction of the random fields to the estimated taxon range, which covers a small, geographically homogeneous region for most taxa (in particular for *Quercus*).

For the analysis of relationships between vegetation and population intensities we consider cross-covariance and cross-correlation functions. Let  $r_{max} > 0$  denote the maximum distance of any two locations in  $W$ , i.e.,  $r_{max} = \max\{d(s_1, s_2), s_1, s_2 \in W\}$ . As  $\Pi$  and  $\Lambda$  are assumed to be jointly second-order stationary and isotropic, the cross-covariance function  $C_{\Pi\Lambda} : [0, r_{max}] \rightarrow \mathbb{R}$  of  $\Pi$  and  $\Lambda$  is defined by  $C_{\Pi\Lambda}(r) = \text{cov}(\Pi(s_1), \Lambda(s_2))$  for any  $s_1, s_2 \in W$  with  $r = d(s_1, s_2)$  (Genton and Kleiber, 2015). Since cross-covariance functions are difficult to interpret, they are typically normalized to obtain cross-correlation functions. Using that the variances  $\sigma_{\Pi}^2 = \text{var} \Pi(s)$  and  $\sigma_{\Lambda}^2 = \text{var} \Lambda(s)$  do not depend on  $s \in W$  and by assuming that  $\sigma_{\Pi}^2, \sigma_{\Lambda}^2 > 0$ , the cross-correlation function  $\rho_{\Pi\Lambda} : [0, r_{max}] \rightarrow [-1, 1]$  of  $\Pi$  and  $\Lambda$  can be defined by

$$\rho_{\Pi\Lambda}(r) = \frac{C_{\Pi\Lambda}(r)}{\sqrt{\sigma_{\Pi}^2 \sigma_{\Lambda}^2}} \quad \text{for all } r \in [0, r_{max}]. \quad [8]$$

In order to give an estimator of the cross-correlation function  $\rho_{\Pi\Lambda}$ , we first need to choose a finite sequence  $t_1, \dots, t_k \in \hat{\xi}$  of sample points, at which the values  $\hat{\pi}(t_1), \dots, \hat{\pi}(t_k)$  and  $\hat{\lambda}(t_1), \dots, \hat{\lambda}(t_k)$  are determined to be used for estimation. For that purpose, we suggest to generate a realization of a homogeneous Poisson point process in  $W$  with intensity  $\alpha = 0.0003$  (e.g., Chiu et al., 2013; Diggle, 2014), and use those points of the process as sample points  $t_1, \dots, t_k$  that fall into the considered estimated taxon range  $\hat{\xi}$ . The intensity  $\alpha = 0.0003$  seems suitable as, on average, we get a number of sample points (3,876 in  $W$  in our example) that is high enough to ensure a reliable estimation but still allows computations to be done in a reasonable time. A homogenous Poisson process is chosen over an inhomogeneous process since the latter would disproportionately focus on regions with high values and strong correlations and ignore regions with low values; our intention here is to study both.

The most intuitive approach is to first estimate the cross-covariance function  $C_{\Pi\Lambda}$  using the method of moments (e.g., Genton and Kleiber, 2015). However, this method usually provides estimates for certain distance classes only leading to discontinuous and unstable estimated cross-covariance functions. Also the fitting of parametric models, as advised in most geostatistical applications to obtain smooth and continuous estimates (Montero et al., 2015), is not suitable here since it seems impossible to find a model that can be fitted adequately for all taxa and time periods. A more appropriate alternative is to consider a nonparametric kernel approach for the estimation of cross-covariance functions (similar to the Nadaraya-Watson estimator used in sections 4.4 and 4.5.1). This kind of estimator is frequently used in point process statistics for the estimation of mark correlation functions (Illian et al., 2008) and has been successfully applied and interpreted in a geographical context (e.g., Shimatani, 2002; Ledo et al., 2011). Let  $r_{est} > 0$  be chosen in such a way that for each  $r \leq r_{est}$  the number of pairs of sampling points  $t_i, t_j \in \{t_1, \dots, t_k\}$  in  $\mathcal{E}$  with approximate distance  $r \approx d(t_i, t_j)$  is reasonably large. Then, an estimate  $\hat{C}_{\Pi\Lambda}(r)$  of  $C_{\Pi\Lambda}(r)$  based on  $\hat{\pi}(t_1), \dots, \hat{\pi}(t_k)$  and  $\hat{\lambda}(t_1), \dots, \hat{\lambda}(t_k)$  can be computed by

$$\hat{C}_{\Pi\Lambda}(r) = \frac{\sum_{i=1}^k \sum_{j=1}^k (\hat{\pi}(t_i) - \hat{\mu}_{\Pi})(\hat{\lambda}(t_j) - \hat{\mu}_{\Lambda}) K_1^E\left(\frac{r-d(t_i, t_j)}{h}\right)}{\sum_{i=1}^k \sum_{j=1}^k K_1^E\left(\frac{r-d(t_i, t_j)}{h}\right)} \quad \text{for all } r \in [0, r_{est}], \quad [9]$$

Where  $\hat{\mu}_{\Pi} = \frac{1}{k} \sum_{i=1}^k \hat{\pi}(t_i)$  and  $\hat{\mu}_{\Lambda} = \frac{1}{k} \sum_{i=1}^k \hat{\lambda}(t_i)$  are standard estimates of  $\mu_{\Pi}$  and  $\mu_{\Lambda}$  and  $K_1^E : \mathbb{R} \rightarrow [0, \infty)$  denotes the one-dimensional Epanechnikov kernel with bandwidth  $h > 0$ , which is defined as

$$K_1^E(u) = \frac{3}{4}(1 - u^2)\mathbf{1}_{(-1,1)}(u) \quad \text{for all } u \in \mathbb{R}. \quad [10]$$

Comparisons of estimates based on different bandwidths have shown that  $h = 20$  is a good choice to obtain smooth functions without eliminating important details. Using different types of kernel functions has a negligible effect on obtained results. Finally, a plug-in estimate  $\hat{\rho}_{\Pi\Lambda}(r)$  of the cross-correlation function  $\rho_{\Pi\Lambda}(r)$  is given by

$$\hat{\rho}_{\Pi\Lambda}(r) = \frac{\hat{C}_{\Pi\Lambda}(r)}{\sqrt{\hat{\sigma}_{\Pi}^2 \hat{\sigma}_{\Lambda}^2}} \quad \text{for all } r \in [0, r_{est}], \quad [11]$$

Where  $\hat{\sigma}_{\Pi}^2$  and  $\hat{\sigma}_{\Lambda}^2$  are estimates of the variances  $\sigma_{\Pi}^2$  and  $\sigma_{\Lambda}^2$ . In order to obtain stable estimated cross-correlation functions it is recommended not to consider the standard moment estimates here. Instead, by using that  $\sigma_{\Pi}^2 = \text{cov}(\Pi(s), \Pi(s))$  for all  $s \in W$ , an estimate  $\hat{\sigma}_{\Pi}^2$  can be computed according to equation 9 for  $r = 0$  with  $\hat{\pi}(t_1), \dots, \hat{\pi}(t_k)$  and  $\hat{\mu}_{\Pi}$  instead of  $\hat{\lambda}(t_1), \dots, \hat{\lambda}(t_k)$  and  $\hat{\mu}_{\Lambda}$ .

The estimate  $\hat{\sigma}^2_{\Lambda}$  is determined analogously. The value  $\hat{\rho}_{\Pi\Lambda}(r)$  for any  $r \in [0, r_{est}]$  can be interpreted as the estimated correlation of the vegetation intensity and the population intensity at two arbitrary locations in the estimated taxon range  $\hat{\xi}$  separated by distance  $r$ .

#### 4.6.2 Cross-correlations of population and vegetation changes with temporal lag

An even more interesting question that arises when analyzing correlations of vegetation and population is how certain taxa responded to changes in population and vice versa. For that purpose, we suggest to estimate and interpret cross-correlations of changes in vegetation and population intensity maps.

We fix a taxon and a year  $y \in \{1,000, 1,100, \dots, 12,500\}$ . Let  $\hat{\pi}^{(y)}(s)$  denote the estimated change in vegetation intensity at location  $s \in W$  between the years  $y + 250$  and  $y - 250$  (in years BP), which is computed based on the estimated vegetation intensities for the years  $y + 250$  and  $y - 250$ . Analogously, let  $\hat{\lambda}^{(y)}(s)$  be the estimated change of population intensity at  $s$  in the same period (or, to be more precise, between the 500-year intervals  $[y, y + 500]$  and  $[y - 500, y]$ , whose midpoints correspond to  $y + 250$  and  $y - 250$ ).

We only consider locations  $s$  that fall into the intersection  $\hat{\xi}^{(y)}$  of the estimated taxon ranges corresponding to the years  $y + 250$  and  $y - 250$  in the following. By  $\bar{t}_1, \dots, \bar{t}_k$  we denote those sample points (obtained from a realization of a Poisson point process, see section 4.6.1) that fall into  $\hat{\xi}^{(y)}$ . We again suppose  $\hat{\xi}^{(y)}$  to be a realization of some random closed subset  $\Xi^{(y)}$  of  $W$ . Furthermore, we interpret the fields  $\{\hat{\pi}^{(y)}(s), s \in \hat{\xi}^{(y)}\}$  and  $\{\hat{\lambda}^{(y)}(s), s \in \hat{\xi}^{(y)}\}$  as realizations of some random fields  $\Pi^{(y)} = \{\Pi^{(y)}(s), s \in W\}$  and  $\Lambda^{(y)} = \{\Lambda^{(y)}(s), s \in W\}$  restricted to  $\Xi^{(y)}$  with  $\Pi^{(y)}(s), \Lambda^{(y)}(s)$  taking values in  $\mathbb{R}$  for  $s \in W$ .

It is conceivable that a change in vegetation intensity between the years  $y + 250$  and  $y - 250$  does not affect the change in population intensity in the same period but in a period a few hundred years later (or that the vegetation change is influenced by a population change some hundred years earlier). Therefore, we also analyze the cross-correlation function of  $\Pi^{(y)}$  and  $\Lambda^{(y+\delta)}$  for a temporal lag  $\delta \in \{-1,000, -900, \dots, 900, 1,000\}$ . At first, we again assume that the random fields  $\Pi^{(y)}$  and  $\Lambda^{(y+\delta)}$  are jointly second-order stationary and isotropic. Accordingly, the cross-covariance function  $C_{\Pi^{(y)}\Lambda^{(y+\delta)}} : [0, r_{max}] \rightarrow \mathbb{R}$  of  $\Pi^{(y)}$  and  $\Lambda^{(y+\delta)}$  is defined by  $C_{\Pi^{(y)}\Lambda^{(y+\delta)}}(r) = \text{cov}(\Pi^{(y)}(s_1), \Lambda^{(y+\delta)}(s_2))$  for  $s_1, s_2 \in W$  with  $r = d(s_1, s_2)$ . The cross-correlation function  $\rho_{\Pi^{(y)}\Lambda^{(y+\delta)}} : [0, r_{max}] \rightarrow [-1, 1]$  is defined as

$$\rho_{\Pi^{(y)}\Lambda^{(y+\delta)}}(r) = \frac{C_{\Pi^{(y)}\Lambda^{(y+\delta)}}(r)}{\sqrt{\sigma_{\Pi^{(y)}}^2 \sigma_{\Lambda^{(y+\delta)}}^2}} \quad \text{for all } r \in [0, r_{max}], \quad [12]$$

where  $\sigma_{\Pi^{(y)}}^2 = \text{var } \Pi^{(y)}(s)$  and  $\sigma_{\Lambda^{(y+\delta)}}^2 = \text{var } \Lambda^{(y+\delta)}(s)$  for any  $s \in W$ . For the estimation of  $\rho_{\Pi^{(y)}\Lambda^{(y+\delta)}}$  consider a distance  $r^{(y)}_{est} > 0$  such that for each  $r \leq r^{(y)}_{est}$  the number of pairs of sampling points  $\bar{t}_i, \bar{t}_j \in \{\bar{t}_1, \dots, \bar{t}_k\}$  in  $\hat{\xi}^{(y)}$  with approximate distance  $r \approx d(\bar{t}_i, \bar{t}_j)$  is large enough. For the computation of estimates  $\hat{\rho}_{\Pi^{(y)}\Lambda^{(y+\delta)}}(r)$  of the cross-correlations  $\rho_{\Pi^{(y)}\Lambda^{(y+\delta)}}(r)$  for all  $r \in [0, r^{(y)}_{est}]$

we suggest to use the estimators proposed in section 4.6.1 based on  $\hat{\pi}^{(y)}(\bar{t}_1), \dots, \hat{\pi}^{(y)}(\bar{t}_k)$  and  $\hat{\lambda}^{(y+\delta)}(\bar{t}_1), \dots, \hat{\lambda}^{(y+\delta)}(\bar{t}_k)$  instead of  $\hat{\pi}(t_1), \dots, \hat{\pi}(t_k)$  and  $\hat{\lambda}(t_1), \dots, \hat{\lambda}(t_k)$ .

### 4.6.3 Nonparametric confidence bands of cross-correlation functions

In order to determine which of the values of estimated cross-correlation functions should be considered as significantly different from zero, we suggest constructing pointwise confidence bands using non-parametric resampling methods.

We consider two approaches: a subsampling method and a bootstrap method (e.g., Chernick and LaBudde, 2011). We describe how the methods can be applied to estimated cross-correlations of vegetation and population intensity maps obtained in section 4.6.1. An application to cross-correlations of vegetation and population changes, see section 4.6.2, works analogously. Let  $\{\hat{p}_{\Pi\Lambda}(r), r \in [0, r_{est}]\}$  be the estimated cross-correlation function of some fixed year and taxon (with estimated range  $\bar{\xi}$ ) obtained according to equation 11 based on estimated vegetation intensities  $\hat{\pi}(t_1), \dots, \hat{\pi}(t_k)$  and estimated population intensities  $\hat{\lambda}(t_1), \dots, \hat{\lambda}(t_k)$  at locations  $t_1, \dots, t_k \in \bar{\xi}$ .

In the subsampling approach with subsampling proportion  $\beta \in (0, 1)$  a random Bernoulli experiment with success probability  $\beta$  is performed for each sample point  $t \in \{t_1, \dots, t_k\}$  (independently for different sample points) to decide whether the data pair  $(\hat{\pi}(t), \hat{\lambda}(t))$  is retained or dropped. Based on all retained data pairs another estimate  $\{\hat{p}^{(1)}_{\Pi\Lambda}(r), r \in [0, r_{est}]\}$  of  $\{\rho_{\Pi\Lambda}(r), r \in [0, r_{est}]\}$  is computed.

When applying the bootstrap approach,  $\{\hat{p}^{(1)}_{\Pi\Lambda}(r), r \in [0, r_{est}]\}$  is determined based on  $\hat{\pi}(t'_1), \dots, \hat{\pi}(t'_k)$  and  $\hat{\lambda}(t'_1), \dots, \hat{\lambda}(t'_k)$ , where the  $k$  sample points  $t'_1, \dots, t'_k$  are drawn randomly with replacement from the set of original sample points  $\{t_1, \dots, t_k\}$ .

The difference between both approaches can be interpreted as follows. In the subsampling approach, the cross-correlation function is re-estimated based on a reduced set of sampling points, i.e., we are interested in how the estimates change when it is assumed that some of the data are overrepresented or erroneous and are thus dropped. In the bootstrap approach, we do not only drop some of the sample points but also provide the estimated intensities at some other sample points with a higher weight to consider the case that those intensities might be underrepresented in the data. In both approaches the spatial correlation structure of the data is not violated since estimated vegetation and population intensities are still associated with the same geographical locations as before. The resampling procedure is repeated 5,000 times, resulting in a sample  $\hat{p}^{(1)}_{\Pi\Lambda}(r), \dots, \hat{p}^{(5,000)}_{\Pi\Lambda}(r)$  of estimates for each  $r \in [0, r_{est}]$ . Then, based on this sample a confidence interval  $[\theta^{(\gamma)}(r), \bar{\theta}^{(\gamma)}(r)]$  of level  $\gamma \in (0, 1)$  can be computed for  $\hat{p}_{\Pi\Lambda}(r)$ , where  $\theta^{(\gamma)}(r)$  denotes the empirical  $(1 - \gamma)/2$  quantile and  $\bar{\theta}^{(\gamma)}(r)$  the empirical  $1 - (1 - \gamma)/2$  quantile of the sample  $\hat{p}^{(1)}_{\Pi\Lambda}(r), \dots, \hat{p}^{(5,000)}_{\Pi\Lambda}(r)$ . The confidence interval  $[\theta^{(\gamma)}(r), \bar{\theta}^{(\gamma)}(r)]$  is constructed in such a way that it contains  $\gamma \cdot 100\%$  of the estimates  $\hat{p}^{(1)}_{\Pi\Lambda}(r), \dots, \hat{p}^{(5,000)}_{\Pi\Lambda}(r)$ .

Finally, the functions  $\{\theta^{(\gamma)}(r), r \in [0, r_{est}]\}$  and  $\{\bar{\theta}^{(\gamma)}(r), r \in [0, r_{est}]\}$  describe the (lower and upper) boundaries of pointwise confidence bands of level  $\gamma$  for the estimated cross-correlation

function  $\{\hat{p}_{\Pi\Lambda}(r), r \in [0, r_{est}]\}$ . We consider an estimated cross-correlation  $\hat{p}_{\Pi\Lambda}(r)$  to be significantly different from zero, if  $0 \neq [\theta^{(\gamma)}(r), \bar{\theta}^{(\gamma)}(r)]$ , where usually  $\gamma = 0.95$  or  $\gamma = 0.99$  is chosen.

#### 4.7 Interpretation and discussion of correlation results

In this section we focus on the history of *Quercus* as a (more detailed) interpretation for all taxa considered in section 4.4 would go beyond the scope of this paper.

**Figure 4.5** shows estimated cross-correlation functions  $\hat{p}_{\Pi\Lambda}$  for all years  $y \in \{750, 850, \dots, 12,750\}$ , where warm colors indicate positive and cold colors negative cross-correlations. Prior to 10,550 BP both *Quercus* and population intensities were very high in Florida and decreased to the north and westward, leading to high cross-correlations. During the period between 10,550 BP and 6,850 BP *Quercus* was moving northwards, while the population showed a complex pattern of changes. For example, relatively low values of *Quercus* in Texas while population intensities were high and decreasing values of population in Florida while *Quercus* remained stable contributed to low cross-correlations. This is a time period of quite some variability in the climate (Viau et al., 2006), although incompleteness of the population database cannot be ruled out.

Between 6,850 BP and 3,650 BP, the late Archaic cultural period, the cross-correlations were generally high. The distribution of *Quercus* did not change much during this time, and population gradually increased in the central portion of the continent, the region of maximum *Quercus* intensities. Fluctuations in the population caused the cross-correlations to increase and decrease during this time. In the past 3,650 years, the Woodland Period, cross-correlations were uniformly high. Populations decreased in New England and greatly increased, within the central portion of the range, which continued to be the area of maximum *Quercus* abundances. Potential issues with incompleteness of the CARD database, or low sample density in the pollen database toward the west, may have contributed to values of the cross-correlations in the earlier part of the record. The spatial association of the two variables in the late Holocene suggests optimal conditions for human population growth during this time.

As explained in Section 4.6.2 we also estimate cross-correlation functions of changes in vegetation and population intensity during 500-year intervals. Additionally, a temporal lag is introduced, which implies that changes in *Quercus* in a certain interval are also compared to population changes in an interval starting and ending  $\delta$  years earlier/later with  $\delta \in \{-1,000, -900, \dots, 900, 1,000\}$ .

In order to allow for a thorough analysis of relationships between vegetation and population changes it is desirable to summarize cross-correlations for all temporal lags in one figure. For all  $y \in \{1000, \dots, 12,500\}$  and  $\delta \in \{-1,000, -900, \dots, 900, 1,000\}$  we compute the mean of the estimated cross-correlations  $\{\hat{p}_{\Pi}^{(y)}_{\Lambda}^{(y+\delta)}(r), r = 30, 35, 40, \dots, 200 \text{ km}\}$  (to reflect local associations only) and summarize computed means in **Figure 4.6**. Mean cross-correlations on the bold diagonal line correspond to the temporal lag  $\delta = 0$ , values above the diagonal line correspond to a negative temporal lag and values below the line to a positive lag.

Several periods of positive cross-correlations between changes in intensities of *Quercus* and population are seen, notably between 11,000 and 10,000 BP, around 6,000-5,000 BP, and in the past 2,000 years. Negative cross-correlations are found, e.g., between 7,000 and 6,000 BP.

The largest positive cross-correlations are found for changes in *Quercus* in the period between 11,000 and 9,500 BP and changes in population in 11,300-10,000 BP (for all positive and small negative temporal lags). This suggests that *Quercus* and population changed similarly over a period of more than 1,000 years. During this time, both were decreasing in Florida, while increasing to the north (especially *Quercus*). This was a time of rapid warming in North America (Viau et al., 2006). During this time, people of the Paleoindian culture relied on a diverse set of resources including hunting different animals and gathering a variety of food (Fagan, 2000), and the increased diversity of the forest would have provided a suitable habitat to enable the increase of the population. However, the sample size used for estimating population intensities, especially to the south in Florida is small, so more definitive interpretation awaits further development of the database.

A period with low to medium negative associations between 8,000-6,500 was a time when both of the variables showed a complex series of changes. During this time, major changes in *Quercus* occurred in the west, due to a climate becoming moister, whereas changes in population occurred in various places, including to the east. In addition, a decrease in *Quercus* abundance in Florida and the southeast while population was increasing in the region contributed to the negative correlation. Further work is needed to ensure this is a real signal and not a function of lack of data; both datasets have low site densities in this region. The major changes in population at this time therefore are most likely not driven by climate changes which would have been directly affecting *Quercus*. An increase in sedentism, and use of a diverse resource base by Middle Archaic peoples (Fagan, 2000) would have enabled adaptation to numerous environments, and contributed to a lower association with *Quercus* abundance. Delcourt and Delcourt (2004) have speculated that people may have thinned tree populations to encourage acorn yields, at least in some regions, and this would also contribute to a negative correlation.

In the recent past, agriculture and more sedentary lifestyles became increasingly established, with large cities in some areas and increasing impact on the environment (e.g., Fagan, 2000; Delcourt and Delcourt, 2004; Munoz and Gajewski, 2010). The population greatly increased across most of the area, especially in the past 2,000 years and there were small increases in oak abundance, especially to the west. Agriculture and *Quercus* abundance would both be favored in areas with optimal environmental conditions, which could explain positive cross-correlations between change in *Quercus* in 3,000-2,500 BP and population in 2,500-1,500 BP as well as between change in *Quercus* in 1,500-1,000 BP and population in 2,500-1,500 BP.

To conclude the interpretation and discussion we present some example results of confidence bands for cross-correlation functions as proposed in section 4.6.3, where we focus on cross-correlations of changes in *Quercus* and population. Due to their time-consuming computation, confidence bands were only provided every 1,000 years for temporal lag  $\delta = 0$  using the bootstrap method and the subsampling method with success probabilities  $\beta_1 = 0.25$ ,  $\beta_2 = 0.5$  and

$\beta_3 = 0.75$ . In **Figure 4.7** confidence bands based on the subsampling method with  $\beta_2 = 0.5$  for one cross-correlation function are shown.

Cross-correlations for distances up to 550 km are significantly different from zero at significance level 0.99. Using the subsampling method with  $\beta_1 = 0.25$  or  $\beta_3 = 0.75$  results in wider or narrower bands, respectively. The bootstrap approach leads to bands that are almost identical to subsampling with  $\beta_2 = 0.5$ . In order to give a more general recommendation on how large (absolute) cross-correlations of changes in *Quercus* and population should be in order to be considered as significantly different from zero, we compute the maximum width of the confidence band (the difference of the upper and the lower bound of level 0.95) for all distances between 30 and 200 km (estimates for smaller distances tend to be occasionally unstable) for a sequence of estimated cross-correlation functions and divide them by 2 (**Fig. 4.8**). For example, the maximum of the green line indicates that when relying on the subsampling method with  $\beta_2 = 0.5$  (which is a reasonably conservative choice) all cross-correlations greater than 0.35 (or smaller than -0.35) can be assumed to be significant. Note that this approach assumes the confidence bands to be symmetric, which is the case (at least approximately) for the vast majority of estimated cross-correlation functions.

#### 4.8 Conclusions

The aim of this study was to find out if there was a relation between human population intensities and plant abundance intensities over the past 13,000 years in North America, using estimates of human population density from archaeological databases and plant abundance data from a pollen database. The methodology discussed in this paper is based on creating a comparable series of maps for the two datasets using simple nonparametric kernel approaches, and analyzing the spatio-temporal cross-correlations between them. Pointwise confidence bands of cross-correlation functions are proposed to investigate the significance of the obtained results.

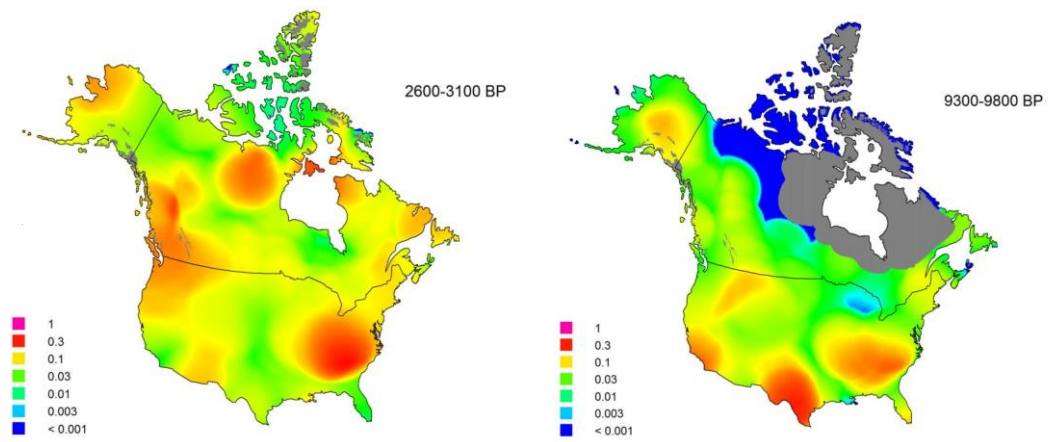
Using an example of the relation between population densities in eastern North America and oak pollen abundance (an index of tree abundance through time), we could identify times in the past when there were both positive and negative associations between changes in the intensities of these two features. Inspection of the maps indicates times and locations where individual points seem to have undue influence on the correlations, especially in areas with few sites in the database. Future work refining the method could be to determine site density needed to more reliably estimate the intensities, especially to enable identification of overly influential or outlier points. Local analyses with smaller smoothing bandwidths may be effective here. In a companion study (Gajewski et al., 2017) the other taxa for which results have been prepared are analyzed in more detail. Since the community ecology of the tree species is well studied, and relative human use of the various taxa is also understood, consistencies between the taxa should help to identify the interpretable signal in these results. Finally these results will be interpreted in the context of more detailed regional studies to ensure that database limitations are not driving the correlations.

The relation between human population growth and ecosystem changes is complex, and causation can go in both directions. In addition, the human use of resources or human influence

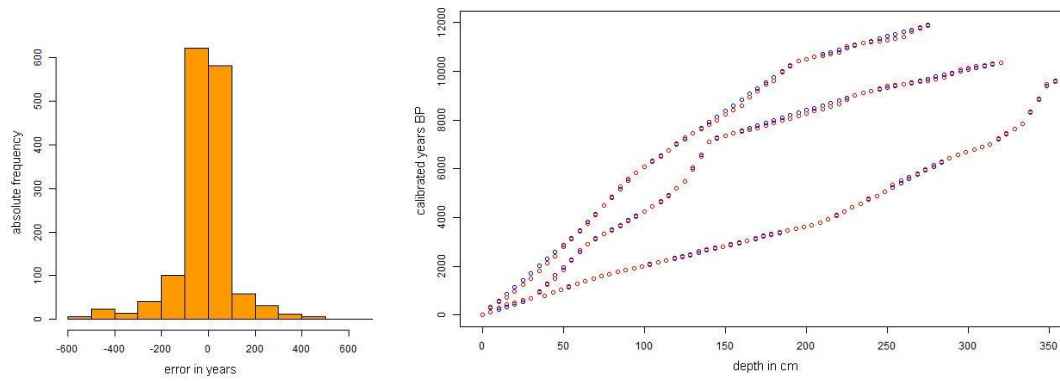
on the environment has changed through time as cultures evolve. This paper provides a method to quantify the relation between population and vegetation and show how they have changed through time and space. The approach is the first attempt to quantify this relation on continental scales and in that way contributes to our understanding of human-environment interactions in North America. Furthermore, the methodology has wide applications in geographic and environmental sciences, where large databases consisting of repeated measurements at a spatial system of sites are being accumulated.

### **Acknowledgments**

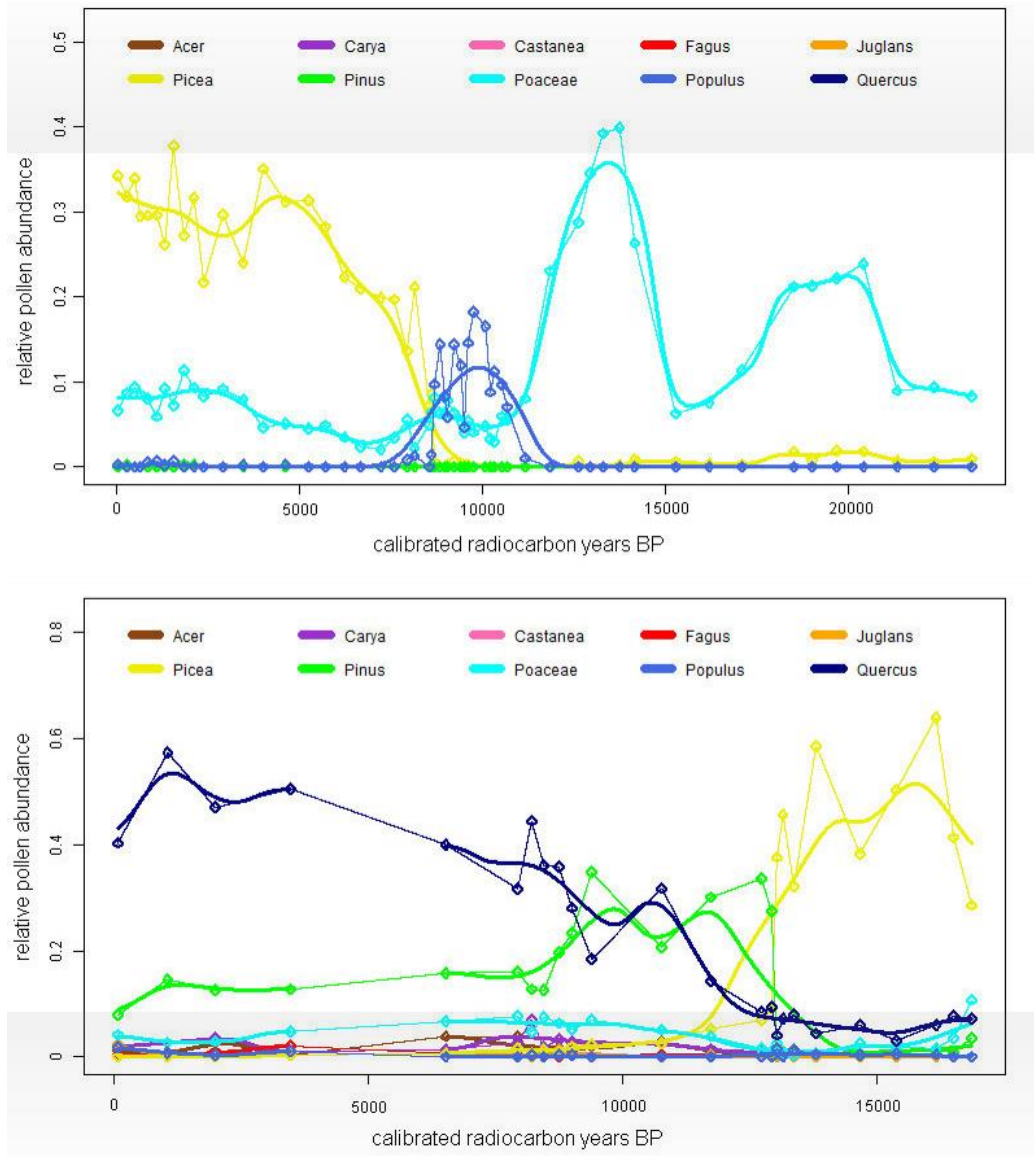
CARD was created by Richard Morlan of the Canadian Museum of Civilization (now the Canadian Museum of History). It was transferred to the University of British Columbia's Laboratory of Archaeology in 2014. The CARD data are available at <http://www.canadianarchaeology.ca/>. The Neotoma data are available at [www.neotomadb.org](http://www.neotomadb.org). K.G. and R.K. received support from NSERC Discovery Grants. The German Academic Exchange Service (DAAD) provided travel funds for B.K. and V.S.



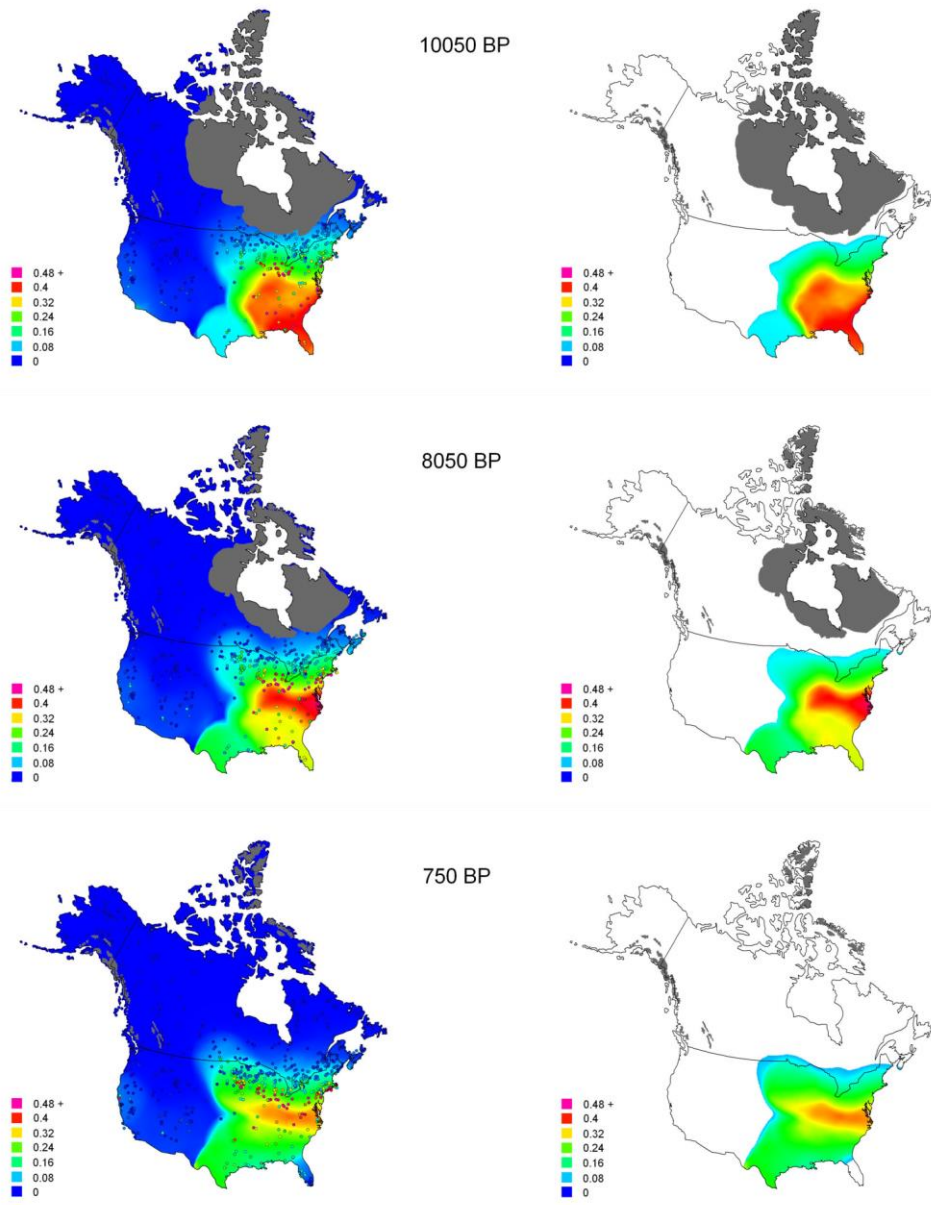
**Figure 4.1** Population intensity maps estimated from the CARD for two time intervals. Intensities are colored according to a logarithmic scale in the interval  $[0, 1]$ . Gray colors indicate areas permanently covered by ice.



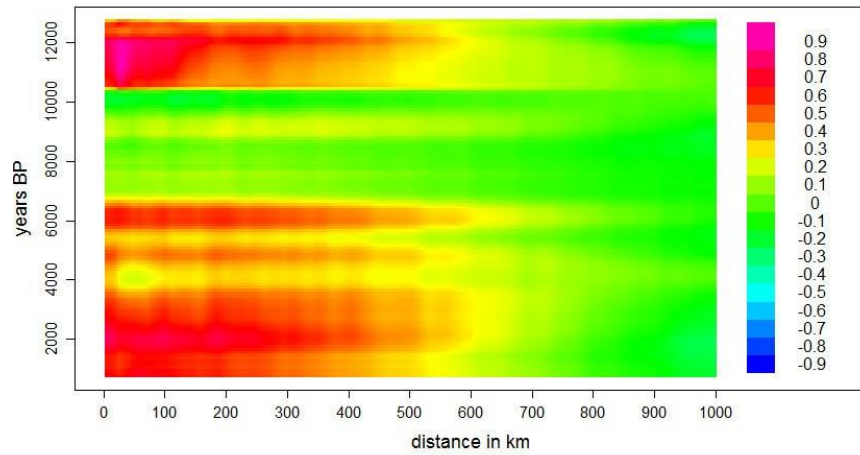
**Figure 4.2** Analysis of errors occurring when exchanging calibration and interpolation of radiocarbon ages. Left: histogram of errors for all chosen depths in all test samples. Right: depth-age-curves for three test samples showing the differences between correct procedure (blue) and exchanging calibration and interpolation (red).



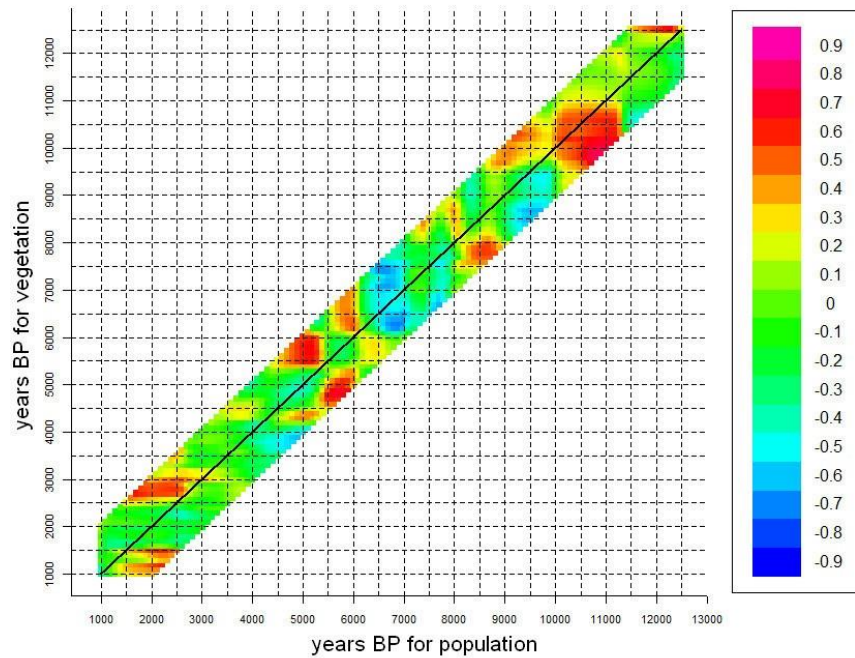
**Figure 4.3** Relative pollen abundances for two different sites in North America and 10 selected taxa: data from Neotoma (points), estimates obtained using linear interpolation (thin lines) and estimates obtained using a Nadaraya-Watson estimator (bold lines).



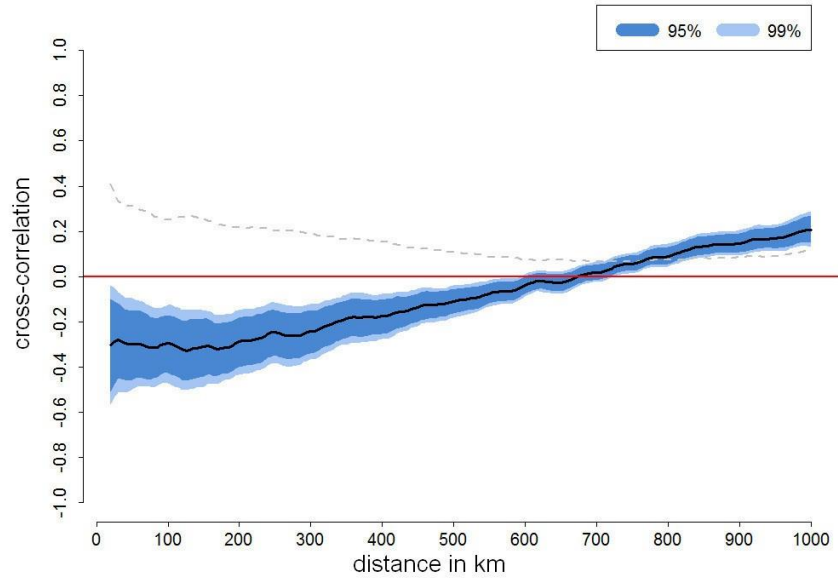
**Figure 4.4** Estimated vegetation intensity maps showing the relative abundance of *Quercus* for three example years. The left-hand side illustrates intensities within the entire observation window  $W$  together with temporally smoothed abundances of *Quercus* at all sites from Neotoma with available data in the considered year. On the right only vegetation intensities within the taxon range estimated with threshold 0.2 are shown. Gray colors indicate areas permanently covered by ice.



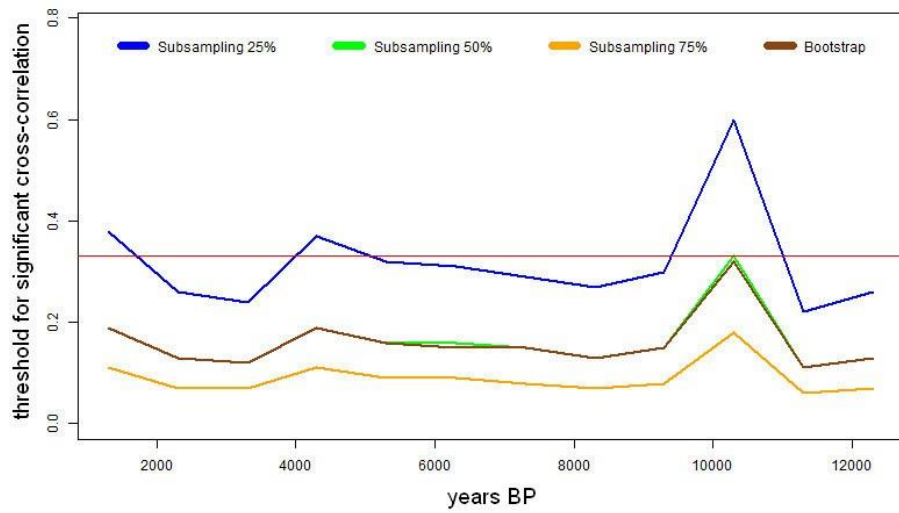
**Figure 4.5** Estimated cross-correlation functions for vegetation intensity maps of *Quercus* and population intensity maps for time intervals between 12750 BP and 750 BP.



**Figure 4.6** Mean cross-correlations for 500-year changes in vegetation intensity maps of *Quercus* and population intensity maps. Means are computed based on cross-correlations for distances between 30 and 200 km. Values on the abscissa correspond to the midpoints of the 500-year intervals of change in population intensity and values on the ordinate correspond to the midpoints of the 500-year intervals of change in vegetation intensity.



**Figure 4.7** Estimated cross-correlation function of changes in *Quercus* and population between 9550 and 9050 BP together with pointwise confidence bands of levels 0.95 and 0.99 computed based on the subsampling method with success probability  $\beta = 0.5$ . The gray dashed line indicates the width of the confidence band (i.e., the difference of the upper and the lower bound) of level 0.95.



**Figure 4.8** Half of the maximum width of confidence bands for cross-correlation functions of changes in *Quercus* and population. Bands are computed using four different approaches and maxima are determined among all widths for differences between 30 and 200 km.

## **5 Human-vegetation interactions during the Holocene in North America**

Konrad Gajewski, Bjoern Kriesche, Michelle Chaput, Rafal Kulik and Volker Schmidt

### **5.1 Abstract**

Between the initial colonization of North America and the European settlement period, Native American land use practices shaped North American landscapes and ecosystems in a significant manner. A critical question is the quantification of the extent of Native American impacts on the land, and how these shaped the distributions of flora and fauna. The present study addresses this question by estimating the spatial correlation between continental-scale records of fossil pollen and archaeological radiocarbon data, and provides a detailed analysis of the spatiotemporal relationship between paleo-populations and ten important North American pollen taxa. Maps of relative Native American population density, based on the Canadian Archaeological Radiocarbon Database, are compared to maps of plant abundance, based on pollen records from the Neotoma Paleoecology Database, using nonparametric kernel estimators and cross-correlation techniques. Periods of high spatial cross-correlation (either positive or negative) between population and plant abundance were identified, but these associations were intermittent and did not increase towards the present. In many cases, high values of population corresponded with high values of a particular taxon in one region, but simultaneously corresponded with low values in other regions, lessening the overall correlation between the two fields. This analysis suggests that human impacts were not significant enough to be identified at a continental scale, either due to low populations or land use, implying significant impacts of ancient human activities on the vegetation were regional rather than continental.

Keywords: Anthropocene, forest history, pollen, archaeology, spatio-temporal cross-correlation, kernel smoother, Canadian Archaeological Radiocarbon Database, Neotoma Paleoecology Database

## 5.2 Introduction

Human activity has had a significant impact on the environment, especially the vegetation. European colonists deforested much of North America by the early 20th century, although it has since regrown in many areas (e.g., Cronon, 1983; Williams, 1989a; Whitney, 1994; Dyer 2006), and the impact of this is easily discerned in the fossil pollen record (e.g., McAndrews, 1988). Broad-scale impacts of human activity of Native Americans on the vegetation are less easily determined, and in the past it was generally assumed that the human footprint in North America prior to European colonization was insignificant. This view was based on the assumption that the total population in the Americas was relatively small and the impacts on the environment were minimal, or at best widely scattered (Denevan, 1992). Changes in human population numbers or in cultural and technological expression would have been primarily driven by climate and environmental changes in this case (e.g., Munoz et al., 2010; Foster, 2012).

There is now an extensive literature on human-environment interactions before European settlement in North America, approached through environmental historians and geographers, archaeologists, and ecologists (e.g., Cronon, 1983; Silver, 1990; Doolittle, 2000; Krech, 2000; Kay and Simmons, 2002; Delcourt and Delcourt, 2004; Harkin and Lewis, 2007). This literature documents extensive human impact on the environment, and instead considers human activity to have been a major determinant of the vegetation distribution prior to European settlement. A critical issue is quantification of the extent of the Native American impact on the landscape, through deliberate or incidental activities such as use of fire, agriculture and silviculture (Denevan, 1992; Vale, 2002; Tulowiecki and Larsen, 2015). Extreme interpretations range from the view that the majority of North America was extensively altered, versus local impacts in a relatively sparsely populated continent. Although few people today accept the idea of a “primeval wilderness” or “endless expanse of forest” (e.g., Braun, 1950) prior to European arrival, the impact of Native American activities as well as the nature of the human-environment relationship is under discussion (Denevan, 1992; Krech, 2000; Kay and Simmons, 2002; Vale, 2002; Harkin and Lewis, 2007).

### *5.2.1 North America as a cultural landscape*

At one extreme is a viewpoint that the population in 1492CE was very high and that the technology was advanced enough such that human impact was extensive across North America (Denevan, 1992) making the continent a “cultural landscape”. Recent work on the population history of North America suggests greater numbers than previously thought, from 0.5-18 million persons, although these numbers are controversial and questioned (Denevan, 1992; Warrick, 2008; Peros et al., 2010).

The data for these conclusions comes from several sources. Archaeological work has suggested large population concentrations in some areas and potentially large impacts on the land in others (Doolittle, 2000). Environmental historians have noted reports by the first European settlers of open forests, of burning by Native Americans to encourage game and make the forests easier to travel through, and active or abandoned agricultural fields (e.g., Cronon, 1983; Silver, 1990; Krech, 2000; Kay and Simmons, 2002; McShea and Healy, 2002; Harkin and Lewis, 2007).

These reports have been extrapolated to suggest that much of the landscape was open parkland. Although these studies provide a snapshot of only one point in time, they are plausible and extensive. However, as Russell (1983) noted, colonial-era descriptions of open lands may be biased, as they were describing their own settlement areas, which were frequently built near Native villages. Potential biases of the settlers (e.g., encouraging immigration) make interpretation difficult.

Anthropologists interviewed Native Americans about their lifestyles and practices, and these studies indicate local adaptations to the vegetation, including burning of the prairies or forests (e.g., Stewart, 2002; Courtwright, 2011). On the other hand, these also are a snapshot from memory, and in most cases, recollections of practices long abandoned.

Land survey records provide a snapshot of the pre-European vegetation and these have been analysed to determine Native American impacts at the time of contact. For example, Dorney (1981) and Foster et al. (2004) compared vegetation composition near and away from Native American villages; although finding some differences in tree composition near settlements, they concluded there was little impact on the forest, although fires may have been more frequent. A summary of several studies of land survey records by Abrams and Nowacki (2008) showed associations of higher abundance of mast trees (oak, hickory, chestnut) with Native settlement sites. Note, however, that if some estimates of the spread of European diseases are accepted, many lands would have been abandoned and in various stages of succession by the time of these surveys. The major settlement of the United States west of the Appalachians occurred during the transition from the Little Ice Age to the modern climate era, but the relative effect of climate and land-use has not been sufficiently studied for this time period.

In their summary, Abrams and Nowacki (2008) specifically attempted to determine the extent of Native American impact on the forests and grasslands of eastern North America and concluded “we believe that the vast majority of vegetation in the eastern USA was managed directly or indirectly by Native Americans, especially through the use of fire” (pg. 1134). Similar conclusions have been discussed by many others (e.g., Pyne, 1982; Cronan, 1983; Silver, 1990; Krech, 2000; Kay and Simmons, 2002; Stewart, 2002; Harkin and Lewis, 2007; Courtwright, 2011), and this viewpoint has been increasingly accepted (Krech, 2000; Kay and Simmons, 2002; Harkin and Lewis, 2007).

### *5.2.2 North America as a relatively undisturbed landscape*

At the other extreme, lower population numbers, perhaps combined with careful management of the landscape, could have meant that Native Americans left a small footprint on the land or in any event, a landscape with only regional or lesser indication of impact (e.g., Vale, 2002; Munoz et al., 2014). This view seems to be held by paleoecologists, whose studies emphasize the importance of climate variability on the vegetation composition (e.g., Williams et al., 2004).

An intermediate interpretation is that extensive alteration of the landscape may be only associated with villages or portions of the landscape that were used extensively. A series of studies of the vegetation of the mountainous western North America concluded that human impact, including the setting of fires, was more local than regional (Vale, 2002), as did a study of

the eastern deciduous forest (Munoz et al., 2014). In Iroquoian regions, upland areas were more greatly affected (Tulowiecki and Larsen, 2015), whereas Delcourt & Delcourt (2004) found greater impact in bottomlands of the Little Tennessee River Valley. The intensity of impact was dependent on distance from settlements in a complex way and would depend on population density and landscape characteristics (Vale, 2002; Munoz et al., 2014; Tulowiecki et al., 2015).

Paleoenvironmental studies at local or regional scales have shown strong associations of tree species composition with human activity in time and space, although the intensity of activity has been hard to estimate (e.g., Burden et al., 1986; Delcourt et al., 1986; McAndrews and Boyko-Diakonow, 1989; Clark and Royall, 1995; Fuller et al., 1998; Parshall and Foster, 2002; McLauchlan, 2003; Delcourt and Delcourt, 2004; Munoz and Gajewski, 2010). Several regions have been the subject of more intensive study using paleoenvironmental methods. In southern Ontario, pollen records showed changes in the forest coincident with the introduction of Native American agriculture and again following abandonment of the region (Burden et al., 1986; McAndrews and Boyko-Diakonow, 1989; Clark and Royall, 1995), and this seems to scale to the regional level (Munoz and Gajewski, 2010). In New England, studies of the past millennium (e.g., Fuller et al., 1998; Parshall and Foster, 2002) have generally concluded that Native American impacts are local, but difficult to separate from climate impacts. Faison et al. (2006) did not find evidence of human-induced openness in the forests of southern New England until European settlement. Delcourt and Delcourt (2004) presented paleoecological and archaeobotanical evidence of long-term impacts of human populations on the forest composition and landscape-level structure in Tennessee and surrounding areas; the impacts greatly increased after 3200 BP. Vale (2002) summarized the evidence for Native American impact on the environment in western United States mountain ranges, and the various authors in this volume generally concluded that the impacts, including that of fire, were local and not particularly extensive.

The overall conclusion from these regional studies using paleoecological records does suggest that human impact was strong in areas, but it is unclear if it extended over the entire landscape. However, most of these studies are restricted to the past ~1000 years, and interpretation is confounded by climate changes that have occurred during this time, notably the Medieval Warm Period and Little Ice Age (Gajewski, 1987, 1988; Paquette and Gajewski, 2013; Lafontaine-Boyer and Gajewski, 2014), which have impacted the forests and whose timing is frequently (and coincidentally?) similar to changes in human impacts. Many of these studies have low temporal resolution, frequently don't study herbaceous or shrub pollen, and in any event show relative composition and not plant density.

### *5.2.3 Resolving the viewpoints of ancient North America*

The discussion has gone back and forth without resolution, in part due to lack of data needed to resolve the issue. Political and social interpretations and belief systems also confuse the issue (Vale, 2002; Harkin and Lewis, 2007), including disciplinary differences: social scientists and historians tend to accept a more impacted landscape, whereas ecologists tend to view a less-altered landscape. Lack of resolution has many sources, including: (a) rehashing the same limited datasets and difficulty in knowing which historical reports are reliable, as well as extrapolating to

a regional scale from local reports, (b) studies based on ethnographic evidence report events that had occurred long ago, (c) estimates of Native American population at contact are in dispute, as is the change in population between the 1500s and French-English settlement, leading to reports and land surveys documenting land may have been already altered, and (d) coincidence in time of potential climate and human-caused changes.

One component that has been lacking in the study of human-environmental relationships, and this has been noted as far back as by Day (1953), is consistently-derived and quantitative estimates of human population density over the Holocene for North America. However, in recent years, these data have become available (Canadian Archaeological Radiocarbon Database or CARD; Peros et al., 2010; Gajewski et al., 2011; Martindale et al., 2016; [Fig. 5.1b](#)). Although it provides only relative numbers, it has been used, for example, in a reconstruction of the time evolution of the population of North America for the past 15000 years (Peros et al., 2008) which is associated with climate variations ([Fig. 5.1c](#)). These paleodemographic estimates have been mapped, depicting the time-space evolution of relative population density (Chaput et al., 2015; chapter 3). Fossil pollen data from sediment cores spanning the continent ([Fig. 5.1a](#)) are available in the Neotoma Paleoecology Database, and a large literature has demonstrated that these are a proxy for past populations of trees (e.g., Williams et al., 2004). The availability of databases of past vegetation (pollen records) and past human population numbers in time and space provides an opportunity to study this question at a continental scale.

In this paper we approach the question of large-scale human-vegetation interactions by comparing estimates of Native American populations over the past 13000 years (Chaput et al., 2015) to data on the distribution of trees during the same time period. If it is true that North America is a cultural landscape, then this should be seen in the fossil record. Here we ask two questions. First, did Native Americans and their activities affect the vegetation in a quantitative, measurable way at regional to continental scales? For example, if Native Americans planted or encouraged the growth of particular tree species, such as *Carya* or *Juglans* (Wycoff, 1991; Abrams and Nowacki 2008), then the abundance of these trees should increase after a population increase. Alternatively, did changes in the vegetation affect human activity, including population growth? If tree populations that supported wildlife and produced consumables for Native Americans, such as *Quercus* or *Castanea* (e.g., Wang et al., 2013) became more abundant, this would mean a more reliable food source, encouraging larger populations. If there was extensive burning or clearing through agriculture, this should appear as an increase in disturbance taxa such as *Populus* or *Poaceae*, or decrease in late successional taxa such as *Acer* or *Fagus*, as seen on a regional scale in Ontario in the late Holocene (Munoz and Gajewski, 2010).

Kriesche et al. (subm.) described a methodology developed for this purpose, and briefly discussed the example of *Quercus*; here we investigate several other key taxa in greater detail. Although the migration, distribution and abundance of the major tree taxa of eastern and northern North America has been studied on numerous occasions (e.g., Davis, 1976; 1981; Delcourt and Delcourt, 1987; Jacobsen et al., 1987; Webb, 1988; Webb et al., 1993; Williams et al., 2004), we update this body of work using an expanded pollen database, and include taxa not previously studied, especially in the more recent of these references. The present study is the first

to investigate the pollen data in association with estimates of human paleo-populations over the course of the Holocene.

## 5.3 Methods

### 5.3.1 Data

For estimates of past vegetation, we used pollen data from the Neotoma Paleoecology Database (Fig. 5.1a; [www.neotomadb.org](http://www.neotomadb.org)); all North American (north of Mexico) data were extracted for this study. We used estimates of human population in space and time (Chaput et al., 2015; chapter 3) based on the CARD database (Fig. 5.1b; Gajewski et al., 2011; [www.canadianarchaeology.ca](http://www.canadianarchaeology.ca)). All radiocarbon and pollen data were calibrated (Chaput et al., 2015; Kriesche et al., *subm.*), and ages discussed in ka (1000 years before present). Details of data extraction and preparation of the maps and figures are explained in Chaput et al. (2015; population) and Kriesche et al. (*subm.*; pollen).

### 5.3.2 Methods

To illustrate the series of steps involved in the methodology (Fig. 5.2), we include an example that shows how population is compared to one tree taxon (*Quercus*) for one time period (0.75ka). In total, figures were produced for 121 time periods and 10 taxa: boreal and widespread taxa (*Picea*, *Pinus*), major genera of the eastern deciduous forests (*Quercus*, *Castanea*), important food sources (*Carya*, *Juglans*), important trees of closed forests of the eastern deciduous forests (*Acer*, *Fagus*), and disturbance indicators (*Populus*, Poaceae). The results were obtained as follows:

- a) Chaput et al. (2015) produced a series of maps showing the spatial distribution of human population of North America through time. Each map was based on data from a 500-year window (Fig. 5.2a), and subsequent maps were produced at overlapping intervals of 100 years: 0.5-1.0 ka, 0.6-1.1 ka, ..., 12.5-13.0 ka. Smooth population maps were prepared from the point clouds (radiocarbon dates from archaeological sites; a subset of Figure 5.1b using data from the appropriate interval) using a kernel density estimator, which appropriately accounts for biases from inhomogeneous sampling strategies, taphonomic loss, and boundary effects. The present study is based on the period 0.75 ka - 12.75 ka, where the central age of the 500-yr period is used as the age index. This resulted in a total of 121 maps.
- b) A series of comparable maps showing the pollen percentage distributions was prepared for each of the 10 taxa. First, pollen percentages were temporally smoothed to the target ages for each site individually. Then, an estimator from kernel regression (Kriesche et al., *subm.*) was used to determine the spatial distribution of pollen percentages at the central age of each 500-year time period (Fig. 5.2b), resulting in a comparable set of 121 maps per taxon. To enable comparability, the same configurations of the estimators that were used in step (a) were used for the population maps, except that the population was illustrated on a logarithmic scale and the pollen percentages on a linear scale.

- c) Since it would be illogical to compare population and pollen values from outside the taxon range (where pollen percentages are practically zero), only the values of population (**Fig. 5.2c, left map**) within the range of the taxon (**Fig. 5.2c, right map**) were studied. The range of a plant taxon at a given year was determined as the region where estimated pollen percentages have a value of at least 0.2 times the maximum of pollen percentages in that time-slice (Kriesche et al., *subm.*). The threshold of 0.2 was determined by visual comparison of the taxon ranges for most recent maps (computed with several different thresholds) with the modern range of the taxon. In the remainder of the study, only estimates of vegetation abundance and population within this range were used for correlation and cross-correlation analysis. Animations were prepared showing pollen abundances through time with the data (**Videos A4.1 – A4.10**), for both the entire continent and restricted to the taxon range. Animations were also produced comparing estimated pollen percentages for each taxon with population estimates within the taxon range (**Videos A4.11 – A4.20**). These estimated values were used in the following for correlation analyses.
- d) Spatial correlation functions of both population and pollen percentage maps (separately) were estimated for each time period; an example is shown of *Quercus* and population for the time period centred on 0.75 ka (**Figs. 5.2d, 5.2e**). Note that positive correlations are caused in part by the smoothing parameter of the kernel methods, but the strength of the functions also reflect the degree of spatial homogeneity in the underlying archaeological and paleoecological data.
- e) Step (d) produced 121 graphs for each taxon and for population. To facilitate interpretation and discussion of these results, all of the correlation functions for each taxon within the range and for population were summarized in one graph. These “heat maps” (**Figs. 5.2f, 5.2g**) show the spatial autocorrelation of the taxon or population through time.
- f) Spatial cross-correlation functions were estimated that describe statistical relationships between estimates of population and pollen abundance of each taxon (both being restricted to the estimated taxon range). There are 121 graphs for each taxon-population pair (**Fig. 5.2h** is an example for 0.75 ka), and the series of each taxon were summarized in one plot (**Fig. 5.2i**) which consists of 121 cross-correlation functions, which identify times of strong spatial correlation between population and taxa.
- g) To determine if changes in population (**Fig. 5.2j**) and taxon abundance (**Fig. 5.2m**) were correlated, difference maps were produced subtracting maps with a temporal distance of 500 years to provide 116 maps of 500-year changes in population and taxon abundance (in 100-year steps). Animations summarizing and comparing the changes in pollen percentages and population were prepared (**Videos A4.21 – A4.30**).

These animations compare population and pollen changes that correspond to the same 500-year change interval and also to intervals that are shifted by the temporal lags +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years since it seems plausible that vegetation does not respond to population change immediately but some hundred years later or vice versa. A positive lag means that vegetation change corresponds to a more recent 500-year change interval than change in population and a negative lag means the opposite.

- h) Similarly to the actual maps, spatial correlation functions were estimated for the difference maps of vegetation abundance (**Fig. 5.2k**) and population (**Fig. 5.2n**) for each 500-year change interval, and the series of correlation functions were summarized as heat maps (**Figs. 5.2l, 5.2o**). Cross-correlations between the difference maps for vegetation abundance and population were also prepared for lags of -1000, -900, ..., -100, 0, 100, ..., 900, 1000 years; selected examples are shown (**Figs. 2p** (lag 0), **5.2q** (+100), **5.2r** (-100), **5.2v** (+300), **5.2w** (+500), **5.2x** (+1000), **5.2ab** (-300), **5.2ac** (-500), **5.2ad** (-1000)) and the comparable summaries for all time periods are shown as heat maps below the individual example cross-correlation graphs. Inspection of these figures shows that in general we obtained similar cross-correlation results for similar lags, which arise in part from the temporal smoothing parameters, but also due to the association between the fields. For more distinct lags, however, larger differences in cross-correlation functions are observed.
- i) Finally, the mean cross-correlations between differences in taxon abundance and population at small distances (between 30 and 200 km) are summarized in a matrix plot. These are the most expressive results summarizing all estimated statistical cross-correlations between changes in the considered taxon and population (for all temporal lags) in one graph to enable a meaningful interpretation (**Fig. 5.3**).
- j) **Appendix 3** presents a taxon-by-taxon summary of the major results. A one-page set of heat maps is provided for each of the ten taxa which includes a summary of the cross-correlations between difference maps (a), correlation functions of the maps of taxon abundance (b), cross-correlations of maps of taxon abundance and population surfaces (c), correlation functions of differences in taxon abundance and cross-correlation of these maps with population (d) at zero lag (e), lag of +100 years (f), lag of -100 years (g), lag of +300 years (h), lag of -300 years (i), lag of +500 years (j), lag of -500 years (k), lag of +1000 years (l), lag of -1000 years (m).
- k) Nonparametric resampling methods were used to assess the significance of obtained correlation results (Kriesche et al., *subm.*). For each taxon a minimal correlation coefficient has been determined such that all estimated (cross-) correlations with higher absolute values are considered to indicate a real signal. These minimal values are ca. 0.15 (*Picea*, *Pinus*), 0.2 (*Acer*, *Carya*, *Juglans*), 0.25 (*Fagus*), 0.3 (*Castanea*, *Poaceae*, *Quercus*) and 0.4 (*Populus*).

## 5.4 Results

### 5.4.1 Population

The population maps are discussed in detail in Chaput et al. (2015). Prior to 10 ka, there were population maxima in the southern United States and northwest of North America. Maxima remained in the south in the early Holocene, and population tended to increase in the west over the course of the Holocene. In the eastern deciduous forest region, there was an early maximum of population density in the Gulf States and Texas; it then moved northward and was west of Appalachians by 8 ka. After 6 ka the population maximum moved eastward, and by 4 ka, there was a band of high population density along the Atlantic seaboard north of Georgia through to the Maritimes. In the past 3 ka, population decreased in New England, and a large maximum in population was centred over Tennessee and Kentucky. Population increased across all of North America in the past 2 ka (Chaput et al., 2015; chapter 3).

### 5.4.2 Boreal and widespread taxa: *Picea* and *Pinus*

The range and abundance maps depicting the distribution of *Picea* through time (**Videos A4.1 – A4.10**) closely resemble those presented in Williams et al. (2004). In the early Holocene, *Picea* was found just south of the Laurentide ice sheet, with maximum values in the Midwest and a high spatial autocorrelation (**Fig. 5.3, Appendix 3**). The range moved northward and westward across a broad front, with maximum values in the Northwest Territories by 10 ka, resulting in high values of the spatial autocorrelation of the difference maps (**Videos A4.21 – A4.30**). During this time, large increases in *Picea* occurred in the northwestern part of range, but there were both positive and negative changes in population depending on the region, reducing the cross-correlation of the difference maps of *Picea* and population (**Fig. 5.3, Appendix 3**). An abrupt decrease in the correlation function at ~7 ka occurred when values were decreasing in Colorado and there were 2 maxima in western Canada, in part caused by a lack of sites between them. Zones of increasing or decreasing values were sometimes in phase or not with population, so the cross-correlation of *Picea* with population was low. The range of *Picea* reached a minimum ~5 ka with the southern limit located further northward than before or after, and low values in Colorado, leading to a decrease in the spatial autocorrelation function and in the difference function. Through the late Holocene there were some small changes in abundance and limit, for example a southward expansion of the southern limit in the past 2 ka.

There was no clear association with population or changes in population. During the period between ~5.5-3.5 ka both *Picea* and population increased in the east, leading to a brief period of high positive cross-correlations. In the past 1 ka, changes in *Picea* were relatively small, whereas population was increasing across the entire range; the cross-correlation of the change maps remained non-significant as a consequence.

Again, our migration maps for *Pinus* (**Videos A4.11 – A4.20**) look similar to those of Williams et al. (2004). Our oldest maps show high values in Florida and in a band south of the ice sheet along the US-Canadian border from Great Lakes eastward. Between 7.5-6.5 ka, *Pinus* decreased in the east and increased in the Prairie Provinces, leading to a larger spatial autocorrelation of the difference maps. The southern pines increased in abundance by 5 ka, and *Pinus* had reached

Yukon and the Northwest Territories in the west. Throughout the Holocene, there were high values of *Pinus* in the western mountains. The spatial scale was large prior to 10 ka, as *Pinus* pollen was broadly distributed, but maxima were not that large (**Appendix 3**). From 10-9 ka, the scale decreased, as there were several centres in northern Ontario and in the mountains with very high values, thereby decreasing the scale although the overall range did not change substantially. In the past 3 ka, the range of *Pinus* pollen increased, especially in the central part of the continent. There are no clear correlations with population change maps (**Fig. 5.3, Appendix 3**).

#### 5.4.3 Major taxa: *Quercus* and *Castanea*

*Quercus* was discussed in Kriesche et al. (subm.) as an example of application of the methodology. At 12 ka there was abundant *Quercus* in the Gulf States, with a maximum in Florida. After 11 ka, the range moved northward both west and east of Appalachians, and by the mid-Holocene it was found across the present-day range, with a maximum from Illinois to Virginia. There was a slight decrease in abundance in the late Holocene, and as a result, the spatial autocorrelation was largest between 10-5 ka. Although there was some association between *Quercus* and population (**Fig. 5.3**), it is probably because *Quercus* is representative of the deciduous forest rather than through any kind of causal relation between *Quercus* and population.

The migration front of *Castanea* was mapped by Davis (1976, 1981), but values of the abundance across the range have not been subsequently studied. In our maps, the range of *Castanea* was mapped in Texas and Louisiana in the earliest part of the record (**Videos A4.1 – A4.10**). Through much of the early and mid-Holocene, a site with positive values of *Castanea* in an area of low site density causes the range to be mapped in Texas; this is probably a different species from *Castanea dentata* that is the primary subject of this study. By 10 ka, there were also high values centred over Virginia, and this eastern part of the range increased in size at 8.8 ka. *Castanea* then decreased in abundance and the range extended northward after 6.8 k, with maximum abundance along the Appalachians. The range, as depicted on these maps, reached New England by 5 ka. Our maps show a faster migration northward than depicted by Davis (1976, 1981; see also Foster et al., 2002), but in general concord with regional interpretations of its history (e.g., Paillet, 2002).

High values of the spatial autocorrelation and cross-correlation of population and *Castanea* prior to ~5 ka (**Appendix 3**) are difficult to interpret as the pollen record is dominated by sites in Texas, resulting in a discontinuous range (**Videos A4.11 – A4.20**). The cross-correlation between *Castanea* and population abundances is relatively high and positive in the late Holocene (**Appendix 3**), as both variables have high values in the central portion of the range. In the past 2 ka, the population maximum values are found in the western portion of the range, reducing the cross-correlation somewhat. *Castanea* is one taxon where the positive and negative lag correlations differed (**Fig. 5.3**), but this is mostly in the older, less reliable section of the graph. Between 5 ka and 4 ka, *Castanea* decreased or did not change much as it moved northward, whereas population was increasing (**Videos A4.21 – A4.30**). Between 3.5 and 2.5 ka, both pollen and population increased, especially in the southern part of the range. During the past 2 ka, *Castanea* changed little, except in the northern part of the range (Paillet, 2002) where it

increased, whereas population was increasing greatly across the range, so there was little cross-correlation.

#### 5.4.4 Major mast trees: *Juglans* and *Carya*

Prior to 12 ka, maximum values of *Juglans* were found from Texas through Arkansas, with scattered values of over 1% across the eastern United States and Canada and in the southwestern United States (**Videos A4.1 – A4.10**). This broadly concurs with results of Delcourt and Delcourt (1987). Shortly after 12 ka, the area mapped as high values was located in a broad area from Louisiana to the Great lakes, west of the Appalachians, and the spatial autocorrelation increased (**Appendix 3**). This rapid change in the range suggests *Juglans* was growing, although rare, across a relatively broad area, and increased in abundance in response to climate warming (Viau et al., 2006; Fig. 1c), as human population density was low (Peros et al., 2008; Chaput et al., 2015). By 9-10 ka, *Juglans* decreased in the south, with maximum values centred over the region of Indiana, Illinois, Missouri and Kentucky and by 7 ka, there were a number of sites with percentages approaching 2%. The range was at a minimum at this time (**Videos A4.1 – A4.10**), as seen in the spatial autocorrelation (**Appendix 3**). The maximum remained in this area until the present.

In the early Holocene, the cross-correlation of *Juglans* with population was relatively small, since the maximum of the population relative density was in the south, whereas maximum *Juglans* was in the central part of the eastern United States (**Fig. 5.3**). Similarly, the difference maps showed a complex sequence of changes that were rarely aligned between populations and *Juglans* percentages. Through most of the mid- to late Holocene, the range of *Juglans* changed little. The population maxima were to the south and east of that of *Juglans*. Since the trees seemed to increase first, this does not suggest that Native populations caused an increase in abundance in this species, although the association may suggest that it was exploited. In the late Holocene, population increased across the entire range of *Juglans*, and maximum values were overlapping with the pollen maximum, displaced a bit to the west. High values of population associated with high and low values of *Juglans* contributed to the lower spatial autocorrelation. This preceded the major increase in agriculture in the region.

Prior to 12 ka, *Carya* pollen were found in the southeast, with a slight maximum over Arkansas and the lower Mississippi valley (**Videos A4.1 – A4.10**). The range gradually extended northward, especially west of Appalachians, and extending to the Great Lakes by 9 ka. From 7-4 ka it increased in abundance in the Southwest, with not much change in range, and through much of the Holocene, it was found across a broad band from Texas to Pennsylvania. Some of these tendencies could be seen in Delcourt and Delcourt (1987), Davis (1976, 1981) or Williams et al. (2004).

As *Carya* moved northward and was found at moderate values across a large range, the maximum in population was also broadly coincident, leading to high cross-correlations between 10-7.5 ka (**Appendix 3**). In the mid-Holocene, high values of population sometimes corresponded to high values of *Carya*, but in other regions to low values, so the overall spatial correlation was around zero. The spatial cross-correlations of the difference maps can be divided

into three periods, the early Holocene, the mid-Holocene between 8-4 ka and the late Holocene (**Appendix 3**), with few consistent associations in the early or late Holocene. An alternation of positive and negative cross-correlations in the mid-Holocene is driven in part by very large increases in *Carya* from 7-5 ka in Texas, accompanied with increases or decreases in population, and this may tend to overly affect the correlations. *Carya* decreased in abundance in the southern part of the range at times when population did as well, leading to some periods of high cross-correlation (**Fig. 5.3**). Between 4.5-3.5 ka, *Carya* increased in the northern central part of the range, and large population increases in the same area contribute to the band of high correlation in **Figure 5.3**. In the past 2 ka, *Carya* changed less in the northern part of the range while at times decreasing in Texas; during this time population was greatly increasing in the central portion of the range, leading to overall low cross-correlations.

#### 5.4.5 Northern deciduous forest: *Fagus* and *Acer*

*Acer* and *Fagus* were found in the northern part of the deciduous forest in the mid- to late Holocene. In both cases, pollen are apparently rarely produced in the southern part of the range, or are depressed by high percentages of other taxa, and the range is underestimated in the south (**Videos A4.1 – A4.10**; Williams et al., 2004).

As found earlier (Davis, 1981a; Delcourt and Delcourt, 1987; Williams et al., 2004), both *Acer* and *Fagus* were found in the lower Mississippi Valley in the late glacial and early Holocene. At 13 ka, the range of *Acer* was extensive, with maximum values in south-central United States (**Videos A4.1 – A4.10**). *Acer* moved north quickly, with significant values even up to the retreating ice sheet. The spatial scale of *Acer* decreased after 6 ka (**Appendix 3**) and by the mid-Holocene, maximum values were centred on Ontario; this continued until the present, with a slight reduction in last 2 ka. The stability of the range in the mid- to late Holocene, and general increase in abundance led to relatively high cross-correlations with population at certain times, but human population maxima also shifted, leading to alternating periods of positive and negative correlation. For example, in the mid-Holocene, there was a minimum in population in the area where *Acer* was most abundant, leading to weak cross-correlations (**Appendix 3**). Generally the maximum in population was found to the south of the region of most abundant *Acer* (**Fig. 5.3, Appendix 3**). The period of large positive cross-correlations in the difference maps from 8-9 ka (*Acer*) and 7-8 ka (population) occurred when *Acer* was increasing greatly in Ontario, and population was quite constant across all of range (**Videos A4.21 – A4.30**). In the past 2 ka, *Acer* increased in the southwest portion of its range, when population was also increasing, leading to positive correlation. Given the stability of a large maximum in *Acer* centred over southern Ontario through much of the Holocene in both the absolute and difference maps, the lack of a comparable feature in population (in either a negative or positive sense) suggests little relation of these two variables.

In the late glacial, *Fagus* was restricted to south of Virginia, leading to a small spatial scale (**Appendix 3**). The range expanded quickly northward, and by 10 ka there were high values in western Pennsylvania; this remained until the present. In the past 8.5 ka, large values of *Fagus* pollen were restricted to the north of Virginia, with most abundant values in Pennsylvania and New York. In the oldest part of the record, when *Fagus* was restricted to the south, there was

little association of *Fagus* and population intensity, and this continued in the period between 11-8.5 ka as it migrated rapidly northward. In the period around 8 ka, *Fagus* difference maps were positive, centred on the range for long time periods, whereas population increased or decreased in various portions of the range. Large positive values of cross-correlation (**Appendix 3**) occurred with large increases in *Fagus* and small but smooth values of population. The large positive association between 4.5-6 ka (*Fagus*) and 4-5 ka (population) occurred when both were increasing in the northeastern portion of the range. After 4 ka, *Fagus* increased in Ontario and Quebec, at times when population was decreasing and at other times when increasing. Increases in *Fagus* and population in the past 3 ka in the western part of the range lead to positive cross-correlations.

#### 5.4.6 Disturbance indicators: *Populus* and *Poaceae*

In the early Holocene, *Populus* was found in Beringia and in a band just south of the ice sheet across the continent leading to a high spatial scale (**Appendix 3**); this band followed the retreating ice. The range of significant values of *Populus* decreased rapidly between 10-8 ka, seen in the maps as well as in the spatial scale (**Videos A4.1 – A4.10, Appendix 3**) and almost no pollen were found in sediment subsequently (Peros et al., 2008). Throughout the Holocene, there was an area with relatively high values in the Southwest. Cross-correlations of *Populus* and population were high in the early Holocene, as both were abundant in Beringia and Alberta (**Videos A4.11 – A4.20**). After 8 ka, when the range of *Populus* was very small and broken into disjunct regions, the autocorrelation functions decreased substantially, as did the cross-correlation with population. In the mid-Holocene, during periods of strong cross-correlation of the difference maps (**Fig. 5.3; Appendix 3**), *Populus* was restricted to a small area of the Southwest.

Before 9 k, *Poaceae* pollen was found above the threshold in several areas from the Arctic to Florida. Over the course of the Holocene, it increased in abundance in the present-day prairie region, as this biome developed after 8 ka. After ~10 ka, the spatial scale of *Poaceae* greatly increased, as the range consolidated into a large band in the central US and Canada. High values of *Poaceae* and population, especially in the southern plains, and low values elsewhere led to some high spatial cross-correlations. The cross-correlation decreased after 5 ka, as maximum *Poaceae* values were found in the central portion of the range, whereas the maximum population was around the edges of the range. Similarly, large values of the difference maps were centred in the Plains, whereas population changes frequently occurred along the edge of the range, leading to little significant cross-correlation.

## 5.5 Discussion

### 5.5.1 Summary of results

Although periods of high spatial cross-correlation between population and plant abundance were identified, these are not consistent, nor do they concur with expected interpretation. We should expect increasing cross-correlations in the recent past, with increasing populations (Peros et al., 2008) and expansion of agriculture, but in fact, this was not observed. If extensive human-caused fires really did push the forests of eastern North America toward more *Quercus* and disturbance

taxa across a large area (below), we would not expect high positive correlations with *Acer* and *Fagus*. The increase of *Populus*, as agriculture developed in Ontario (Munoz and Gajewski, 2010) does not seem to scale upwards to the continental scale. Given the importance of mast species as a food source, we would have expected a higher correlation of *Juglans* and *Carya* with population than was seen.

Several possibilities can be proposed as explanation: (a) there was little or no relation; human impact was not great enough to have an influence at this scale, either due to low populations or the kinds of activities. Any large impact of human activity on vegetation was regional rather than continental; (b) the impact of Native American activity on the vegetation was ubiquitous, but not associated with changes in the population; (c) we did not see the associations since our method is too insensitive, or results are too much determined by aspects of the data such as a low signal to noise ratio;

- a) our conclusion is that there is presently little indication of continental-scale impacts of human activities on the vegetation of North America over the course of the Holocene. This does not mean there was no impact of Native Americans, but that it was local and regional. Vale (2002) provided a classification of land use by Native Americans useful for this discussion, and more work at regional scales, as discussed in the Introduction, could document this interaction.
- b) although this is a possibility, it cannot be investigated at present. Ongoing work attempting to estimate vegetation density, and new data on charcoal and pollen from more sites could help resolve this question.
- c) we identified some taxa, especially those that are underrepresented in the pollen rain, where problems with estimating the spatial fields for part of the time period may make interpretation more difficult. However, for many taxa the fields are reasonable, coherent with a large body of previous work, described above, and associations should be observable. Inadequacies with the databases, especially low site density or incomplete archaeological data in some regions may be reducing the sensitivity of the study. Future work at regional scales, or using different averaging periods could determine the robustness of these results.

### 5.5.2 Vegetation history

The methodology provides reasonable estimates of the tree ranges, which compare well with modern distribution tree maps and with previous studies of the postglacial migration of tree taxa. A series of studies have mapped the migration of major tree taxa, including Davis (1976, 1981), who mapped the migration front of several taxa from the eastern deciduous forest, and Bernabo and Webb (1977), Delcourt and Delcourt (1987; based on 162 pollen sites), Jacobsen et al. (1987; 211 sites at 6 ka), Webb (1988), Webb et al. (1993; 328 sites in eastern North America) and Williams et al. (2004; 759 sites in eastern and boreal North America) who presented contour maps or grids of pollen percentages using various methods. In this paper we used an updated database (Neotoma Paleocology Database) to (a) extend these maps across all of North

America, although many taxa are regionally restricted, (b) present some taxa not studied earlier, and (c) produce maps using a new method, where the abundances are estimated using a kernel density method. Overall, the maps we presented resembled previous work, although they provided more detail due to an increased number of sites, suggesting our methodology is robust. This new analysis and methodology enabled further study of the Holocene vegetation history including (a) an analysis of the spatial autocorrelation of the mapped patterns, (b) mapping of the temporal changes of the taxa using difference maps, and (c) analysis of the spatial cross-correlation of taxon abundance and population. Studying the spatial scale through the autocorrelation function and the difference maps provides useful summaries of dynamics in space and time of the populations. Our spatial analysis shows that not only does each taxon have a different migration history (e.g., Davis, 1976; Webb, 1988; Williams et al., 2004) but the range extension and times of rapid change differ as well, in response to the particular ecological conditions that affect it.

In some cases, a few sites do seem to have an overly large influence on the results. An example is *Castanea* in the older part of the record. *Castanea* pollen in sediment cores is relatively under-represented compared to the abundance of trees on the landscape (Foster et al., 2002), so estimating its range and abundance has greater error due to sampling problems. This over-influence occurs in spite of the smoothing used in the extraction of the data from the raw counts in the database. On the one hand, sites could be selected, and assumed outliers dropped. However, based on the literature, there is little justification for this; for example, the values of *Castanea* or *Carya* in the region of Texas are not unusual (Bryant and Holloway, 1985). The large influence on the spatial fields is rather a consequence of the low spatial resolution of sites in south-central United States. A priority for future work would be to increase the density of pollen records from less densely-sampled regions.

### *5.5.3 Human impacts on tree species over the Holocene at continental scales*

Overall, there was little relation between changes in *Picea* or *Pinus* and changes in population. This is not surprising, as even in the recent past, the human population density of boreal regions was low and agriculture was not possible. The wide range of these taxa, extending over many different cultures and climates can lead to regional positive and negative correlations between population and the taxon cancelling and leaving little correlation.

We could also see little relation between the disturbance indicators and population. Although local and regional paleoecological studies have shown the impact of agriculture on long term vegetation dynamics (Delcourt, 1987; McAndrews and Boyko-Diakonow, 1989; Clark and Royall, 1995; Delcourt and Delcourt, 2004; Munoz and Gajewski, 2010), these impacts apparently do not scale to the entire eastern forest region. For example, Munoz and Gajewski (2010) found increases in *Populus* in pollen diagrams with indications of agriculture, but not at sites with no indications of agriculture in the area surrounding the lake, and suggested a regional change in the forests toward taxa characteristic of disturbance were associated with Native American agriculture since 1200CE in southern Ontario. We could not identify this signal at the larger scale of this study. However, *Populus* decreased from a widespread distribution in the early Holocene, to almost no pollen being recorded in lake sediments after 8 ka (Peros et al.,

2008). It seems that the values of *Populus* pollen are too small today to analyze at a continental scale, and little relation with population was seen in the early Holocene.

In the forest, we would expect an increase in grasses and other herbaceous plants after the introduction of agriculture, or as a consequence of extensive burning. Agriculture was prevalent in two regions of North America: the southwest, where irrigation canals and other structures are seen (Doolittle, 1992; Krech, 2000), and in the east, where traces of agriculture are abundant (Doolittle, 1992; Doolittle, 2000). In the northeast, the major impact would have been after maize became widely used in the past 1000 years, although the origins of agriculture go much farther back (Smith, 2006). The continual clearing of land and abandoning and moving to new locations could have had an enormous impact on the vegetation, especially if fire was used to aid in clearing the forests, although Doolittle (1992) has questioned some of the assumptions about this aspect. However, we could not identify such an impact.

It is widely claimed that the prairie or especially the prairie-forest border was maintained by fires (e.g., Pyne, 1982; Abrams, 1992; Pyne, 2007; Courtwright, 2011) and that these were augmented, if not almost entirely caused, by humans. Evidence for this, in addition to the reports of the first settlers about the extensive burning, was that the forest tended to grow back when fire was suppressed. However, the actual data used for this conclusion is not clear. For example the often cited Gleason (1913) is based on argument rather than data, and the settlement of the prairies occurred during the transition from the Little Ice Age to the modern period; this climate change may have been the cause of dis-equilibrium of climate and prairie in the ecotone. Again, we could not identify a clear association of population with the extent of the prairie, and more work is needed to understand the human role in prairie fires.

In the eastern deciduous forest region, we have the most data, an extensive historical literature and several regions of extensive paleoecological study (see above). The importance of human-caused fires has been emphasized by many authors and whether or not the pre-European forests were continuous or open parklands is not resolved (Pyne, 1982; Cronan, 1983; Silver, 1990; Krech, 2000; Kay and Simmons, 2002; Stewart, 2002; Harkin and Lewis, 2007; Abrams and Nowacki, 2008; Courtwright, 2011). *Quercus* would be favoured by fires caused by human activity (Abrams, 1992; McShea and Healy, 2002; Black et al., 2006; Tulowiecki and Larsen, 2015), however, we did not see an association. Due to the ability to sprout, *Castanea* increased after European settlement (Russell, 2006) and there is speculation that it may have been favoured by fires associated with Native American populations in New England (Foster et al., 2002), but we see little evidence in our results. There are many reports of Native ground fires and this is supposed to have affected the large scale nature of the forests (e.g., Silver, 1990; Stewart, 2002). Paleoecological studies find little evidence for fires (e.g., Clark and Royall, 1996; Lafontaine-Boyer and Gajewski, 2014), although there are few available records, especially to the south. We might expect *Acer* and *Fagus* to decrease if growth of mast trees were being encouraged by Native American activities such as fire or planting, as these are late successional canopy trees and can reproduce in the shade. Positive correlations of these two taxa with population would be expected if populations were having little impact on the forests. The overall positive association of *Fagus* and *Acer* abundances with population may suggest that any human disturbance was

insufficient for wide-scale conversion of the forest into more early successional states, as shown, for example in southern Ontario (Munoz and Gajewski, 2010). Either the disturbance was localized on parts of the landscape or insufficient to affect the entire forest at the scale of this study. An alternative interpretation is that factors that favoured the maintenance of these taxa also maintained populations of Native Americans.

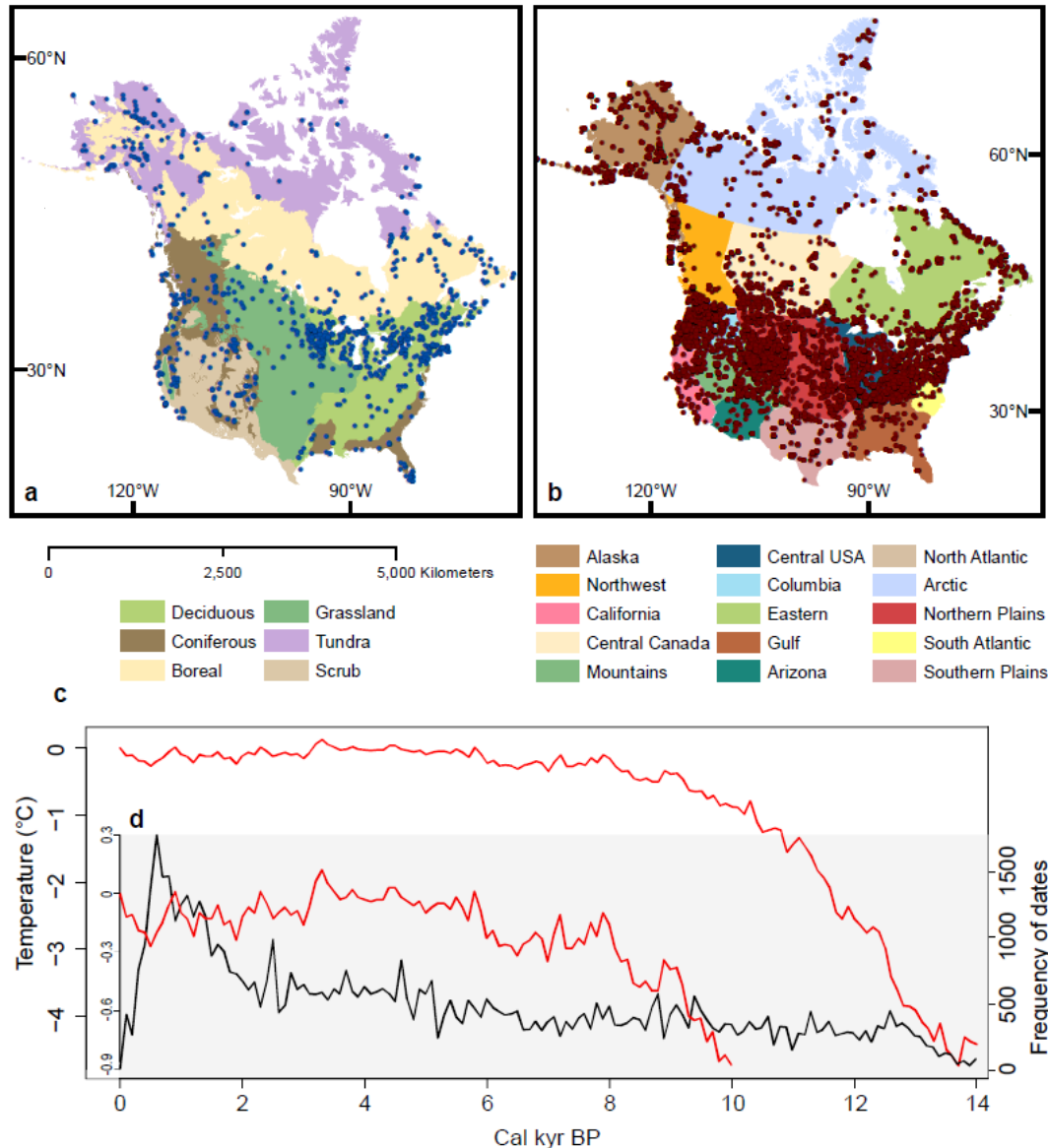
An extensive literature has noted the importance of acorns and nuts for human consumption as well as for animal populations they would have hunted (e.g., Delcourt and Delcourt, 1979; Silver, 1990; Wycoff, 1991; Scarry, 2003). There is a possibility of Native American dispersal of nut trees either deliberately or at least more rapidly than would have occurred naturally (Dorney, 1981; Wycoff, 1991; MacDougall, 2003; Smith, 2007). There are also reports of large populations of mast-producing trees in association with villages, due to either intentional or accidental growth of trees planted or discarded (Silver, 1990; Black et al., 2006). Today, *Carya* is not regenerating across large areas in eastern North America (Cowden et al., 2013). *Carya* would increase in abundance within the range not only at times of appropriate temperature and moisture, but also with sufficient canopy disturbance, and we may expect an increase when population increased. However, increased drought or other factors could also increase canopy disturbance. The lack of a clear association with population (**Fig 5.3**) may indicate a lack of such disturbances. The clearest signal of a potential association between these two taxa and population is found in the northeast in the middle Holocene, and a regional study could help determine its significance.

#### *5.5.4 Summary and Implications*

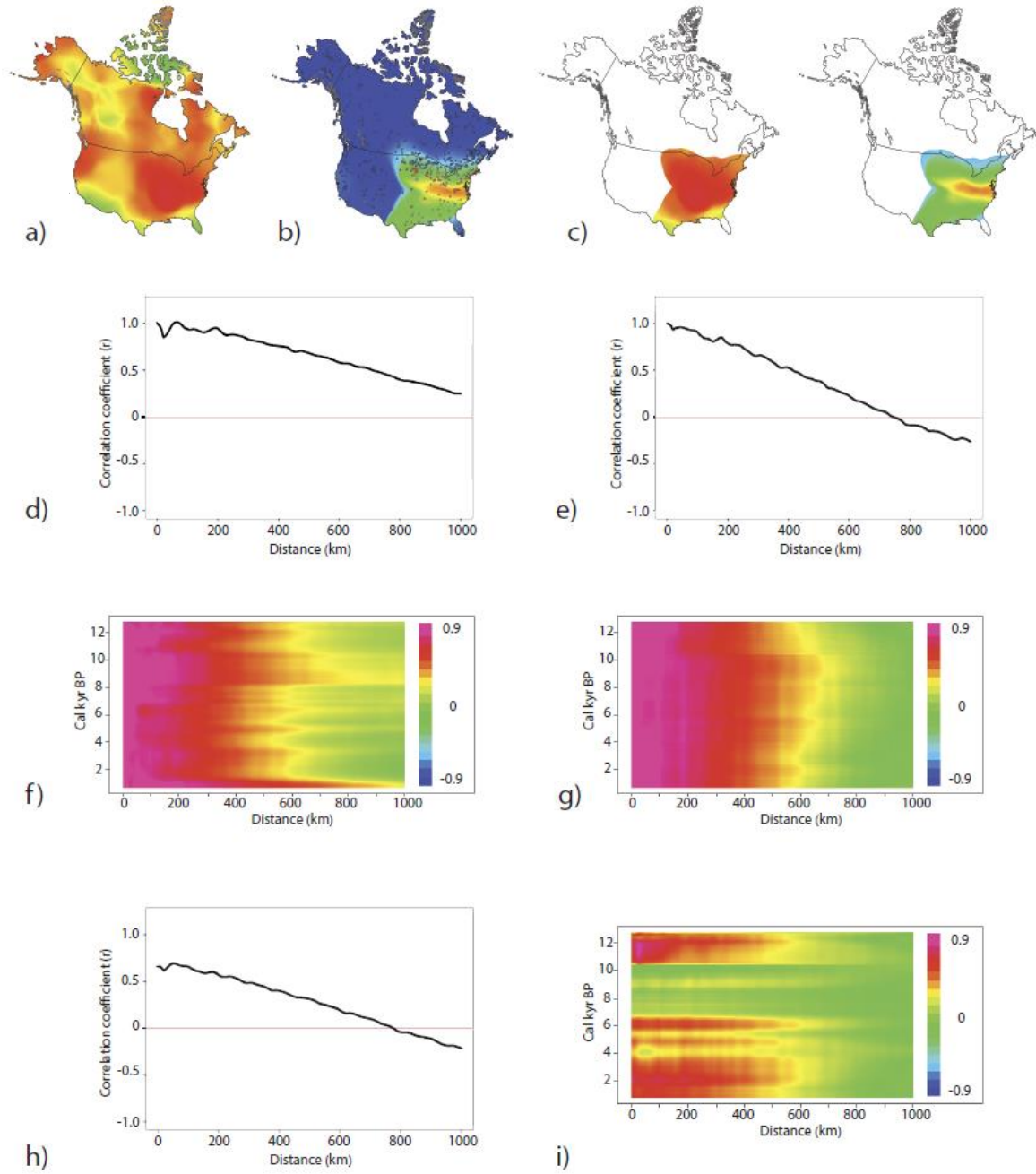
Ruddiman (2014) has suggested that human impacts on the global climate (carbon cycle) began to be measurable thousands of years ago, a concept which has been called the Early Anthropocene Hypothesis. Much of this impact was due to the introduction of rice agriculture in Asia, and extensive deforestation and other activities in Eurasia. The question we asked here was: did impacts of Native North Americans contribute to this early Anthropocene? For example, it is speculated that reforestation of tropical Americas following the massive depopulation due to diseases introduced by Europeans may have contributed to the Little Ice Age cooling through drawdown of carbon dioxide concentrations (Dull et al., 2010; Ruddiman, 2014, his Fig 20-7). Even if the impacts were not sufficient to affect global carbon concentrations, large human populations could have significant impacts on North American vegetation through influences on, for example, ecological succession, forest openness, forest composition and tree ranges, or the location of ecotones such as the prairie-forest boundary. The results of our study do not suggest a significant impact; if they had occurred, we could not identify these at a continental scale. Regional analyses have indicated some impacts (e.g., Delcourt and Delcourt, 2004; Munoz et al., 2010), but these do not appear to scale up to the continental scale. Further regional studies may illustrate the complex interactions of humans and their environment on long timescales, and better identify the spatial extent of impacts. The databases are continually being improved, and future work could attempt this analysis on a more complete dataset.

## **Acknowledgments**

Pollen data are available from the Neotoma Paleoecology Database ([www.neotomadb.org](http://www.neotomadb.org)), and the paleo-population data are available from the Canadian Archaeological Radiocarbon Database ([www.canadianarchaeology.ca](http://www.canadianarchaeology.ca)). We acknowledge those who contributed data to both databases, as well as database coordinators, especially Eric Grimm, Simon Goring and Andrew Martindale. This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Deutscher Akademischer Austauschdienst (DAAD).



**Figure 5.1** a) Map of the major biomes of North America (Olson et al., 2001) with locations of pollen records extracted from the Neotoma Paleocology Database ([www.neotomadb.org](http://www.neotomadb.org)). Each point is one multivariate time series of pollen percentages (vegetation) at that site. Note that the record may vary in length. b) Map of the major cultural regions of North America (Mooney, 1928) with points from the CARD database (Gajewski et al., 2011; Martindale et al., 2016; [www.canadianarchaeology.ca](http://www.canadianarchaeology.ca)). Each point is one radiocarbon date from an archaeological site. Note there may be many points stacked at any location. c) Mean July temperature anomaly of North America derived from the pollen records in [Figure 5.1a](#) (Viau et al., 2006). This curve shows the complete record for the past 14 ka. d) The same temperature data as in [Figure 5.1c](#) but truncated at 10 ka and expanded to better show the variability in the Holocene. This is plotted against taphonomically-corrected human population (black curve) of North America for the past 14 ka, estimated using the data from [Figure 5.1b](#); see Peros et al., 2010, for details.



**Figure 5.2** Overview of methodology used in this study; see Methods section for details.

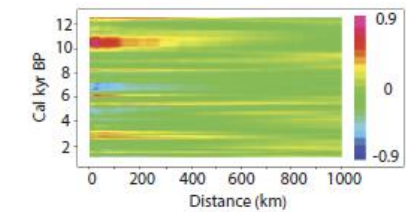
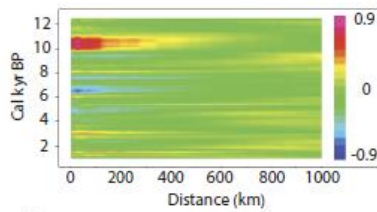
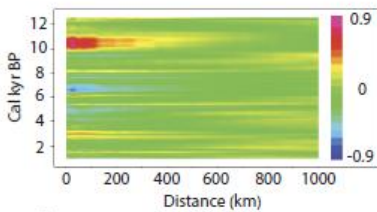
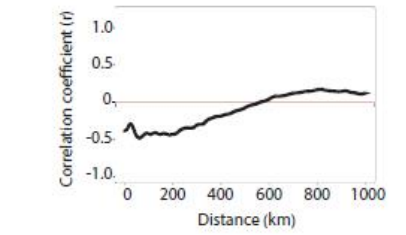
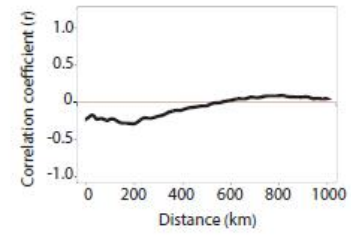
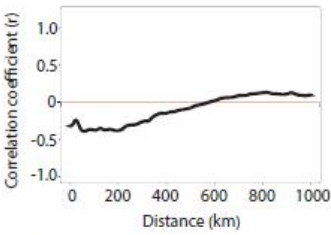
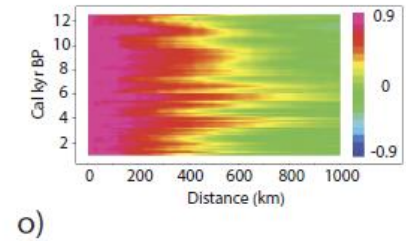
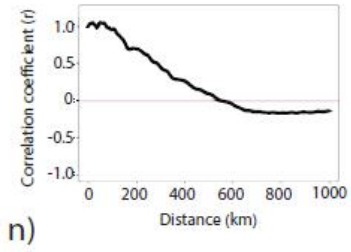
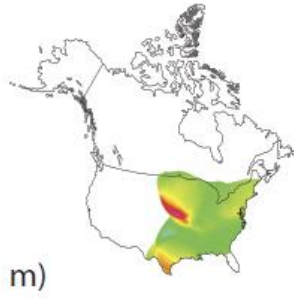
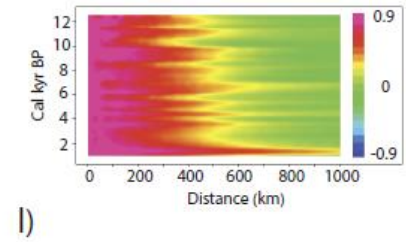
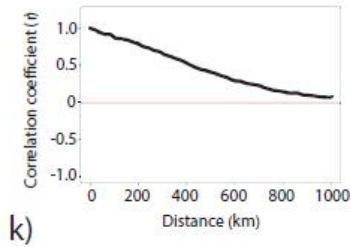
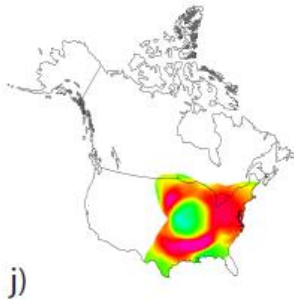


Figure 5.2 Continued.

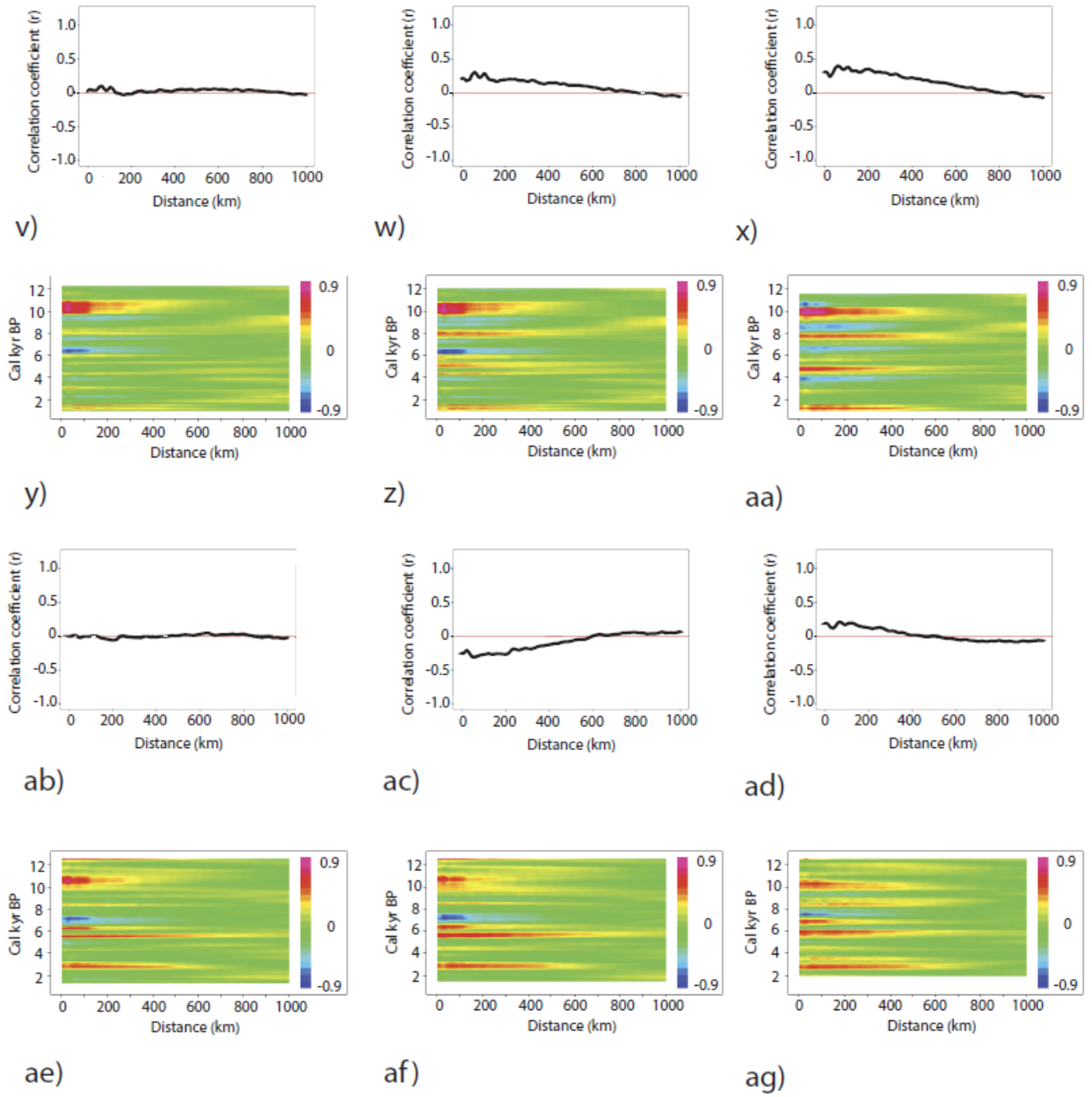
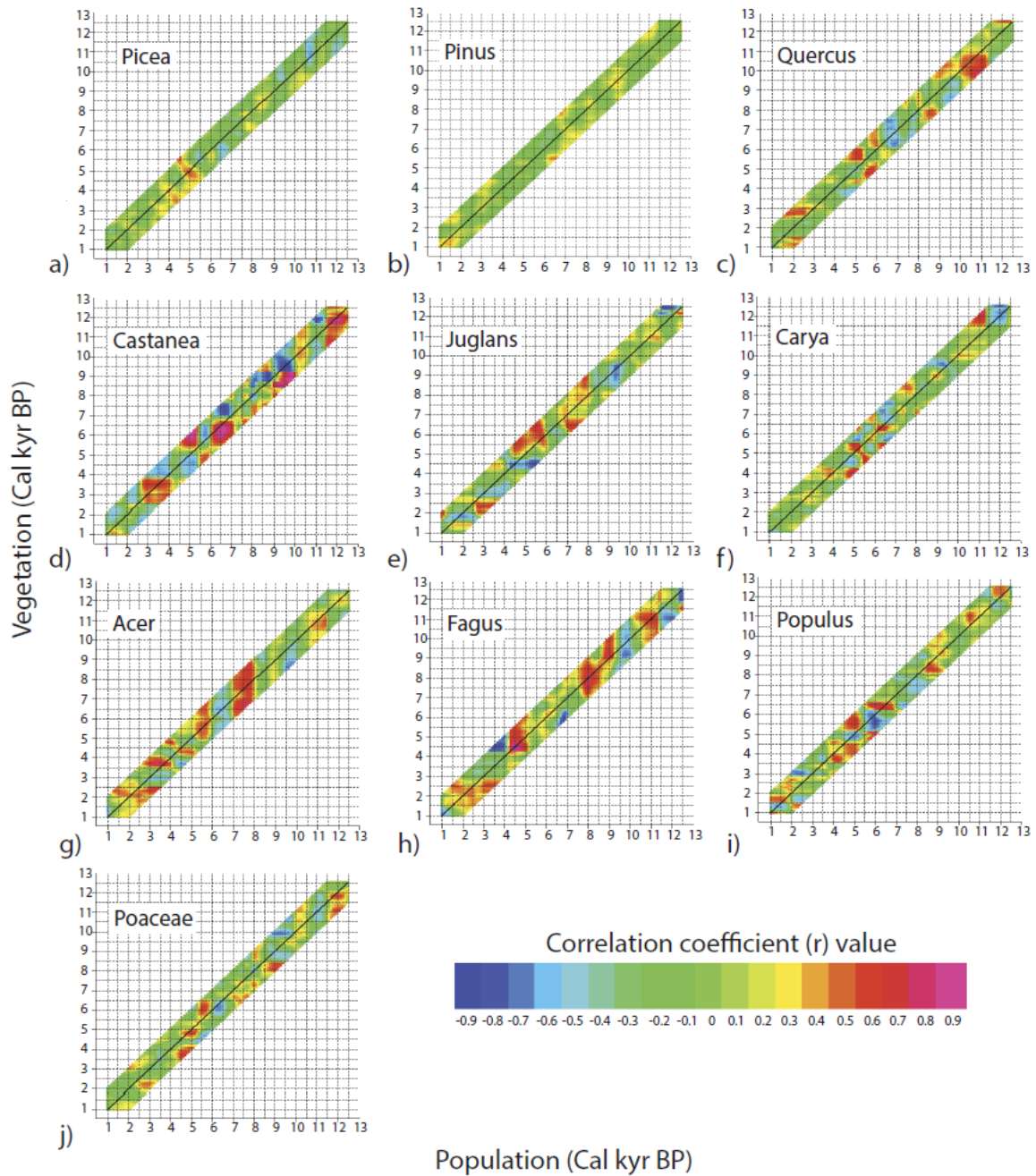


Figure 5.2 Continued.



**Figure 5.3** Diagrams summarizing the cross-correlations between differences maps of taxon abundance and population. For each of the spatial cross-correlation functions (121 for each taxon), the average correlation for distances between 30 and 200 km was extracted, for zero lag, +/-100 years, +/-200 years, ..., +/-1000 years, and these plotted on the one diagram. A positive lag means that vegetation change corresponds to a more recent 500-year change interval than change in population and a negative lag means the opposite.

## 6 Relative pollen productivity estimates and changes in Holocene vegetation cover in the deciduous forest of southeastern Quebec, Canada

Michelle Chaput and Konrad Gajewski

### 6.1 Abstract

The Regional Estimates of VEgetation Abundance from Large Sites (REVEALS) model was used to quantify Holocene changes in vegetation cover in the deciduous forest of southeastern Quebec, Canada. The Extended R-Value (ERV) model was used to obtain relative Pollen Productivity Estimates (PPEs) for eight tree taxa and to determine the Relevant Source Area of Pollen (RSAP) for lakes in this ecosystem. Modern vegetation was estimated using pollen data from 16 small (<0.5 km<sup>2</sup>) lakes and a species-level vegetation survey of southern Quebec. The RSAP was estimated to be within 1600 m of the lakes. *Tsuga*, *Fagus* and *Quercus* were the most productive taxa, and *Populus* and *Acer* were the lowest. Reconstructed changes in absolute vegetation cover show a high abundance of *Picea* followed by *Populus* in the early Holocene. The reconstructed values for *Populus* suggest that it was widely distributed across the landscape. *Abies* and *Acer* were dominant on the landscape during the late to mid-Holocene and an increase in *Picea* during the Neoglacial is more significant than in percentage diagrams. The REVEALS results provide estimates of land-cover change that are more realistic and informative than the use of pollen percentages alone.

Keywords: Relative pollen productivity estimates (PPEs), REVEALS model, pollen-vegetation relationship, Holocene, Quebec, North America

## 6.2 Introduction

Traditional paleoecological studies depend on pollen percentage diagrams, pollen concentration or pollen accumulation rates, which provide insight into the vegetation history of a region. However, the non-linear relationship between pollen deposition, vegetation abundance and taxon-specific differences in pollen productivity hinders the use of these data to obtain quantitative estimates of past vegetation (Sugita, 1994; Sugita et al., 1999).

The Landscape Reconstruction Algorithm (LRA), a two-step framework developed for the quantitative reconstruction of vegetation, overcomes some of these fundamental problems (Sugita, 2007a, 2007b). The first step of the LRA involves reconstructing regional vegetation by transforming pollen percentage data from large ( $\geq 0.5 \text{ km}^2$ ) lakes into vegetation cover. This step uses the Regional Estimates of VEgetation Abundance from Large Sites (REVEALS) model (Sugita, 2007a), which is the model used in this study. The second step uses the Local Vegetation Estimates (LOVE) model to quantify the vegetation within a few square kilometers of small lakes in the same region (Sugita, 2007b). The LOVE model has been tested elsewhere (e.g., Fyfe et al., 2013; Cui et al., 2014; Hjelle et al., 2015) but is not used here.

The LRA uses an Extended R-Value (ERV) model (Andersen, 1970; Prentice, 1985; Prentice et al., 1987; Prentice, 1988), which stems from the R-Value model first proposed by Davis (1963) for estimating the abundance of vegetation on the landscape. The ERV is used to model the pollen-vegetation relationship by correcting for the non-linearity of pollen percentage data. The LRA also accounts for taxon-specific pollen production, pollen dispersal, and pollen source area, ultimately providing a more robust representation of past vegetation cover when compared to pollen percentage studies (Åkesson et al., 2015). Although widely used in Europe (e.g., Broström et al., 2008; Hellman et al., 2008; Mazier et al., 2012; Baker et al., 2016) and elsewhere (Wang et al., 2011), this framework has rarely been tested in North America.

In this paper, we apply the first step of the LRA using pollen data from lakes in the temperate deciduous forest of southeastern Quebec. There have been numerous studies of the postglacial vegetation history of southern Quebec, resulting in an extensive set of pollen records and descriptions of regional vegetation change throughout the Holocene (e.g., Anderson, 1985; Richard, 1995). A lack of information related to the pollen productivity rates of various North American tree, herb and grass species has, however, impeded progress toward the quantification of past vegetation cover. We address this by analyzing the modern pollen/modern vegetation relationship and modelling the association between fossil pollen counts from small lakes and changes in regional tree abundance using REVEALS. This model was originally developed to reconstruct regional (i.e.,  $106 \text{ km}^2$ ) vegetation based on pollen records from lake sediment cores from either one or many large ( $\geq 0.5 \text{ km}^2$ ) lakes within an area of homogenous vegetation but has since been shown to work equally well with multiple small sites (Sugita et al., 2010; Fyfe et al., 2013; Trondman et al., 2015).

In addition to pollen data, REVEALS requires specification of pollen site radii, pollen fall speeds, wind speed, and relative pollen productivity estimates (PPEs). The pollen productivity of individual taxa can vary based on regional climate, topography, soil properties, species characteristics, growing season and environmental history (Broström et al., 2008; Li et al., 2015a; Baker et al., 2016); thus, PPEs need to be obtained on a regional basis if they are to produce accurate reconstructions. Standard sets of PPEs are now available in the literature for parts of North America, Europe and Asia (e.g., Sugita et al., 1999; Broström et al., 2008; Mazier et al., 2012). For regions where PPEs are not yet available, they can be calculated using modern pollen records and modern vegetation data. Modern pollen records come from pollen rain that collects in lake sediments, or in moss, litter or surface soils. Modern vegetation data are obtained from forest inventories, land surveys, satellite images or extrapolations based on stratified random samples (e.g., 2009., 2008; Baker et al., 2016).

Several studies incorporating PPEs and the REVEALS model have been done in Europe (e.g., Hellman et al., 2008; Broström et al., 2008; Soepboer et al., 2010; Trondman et al., 2016; Abraham et al., 2017), Asia (e.g., Wang et al., 2011; Xu et al., 2014; Niemeyer et al., 2015; Xu et al., 2016), and the United States (Calcote, 1995; Sugita, 2010; Commerford et al., 2013) to quantify changes in forest openness and taxon-specific cover through time. The output from these studies is being used to inform climate modelling projects (Pirzamanbein et al., 2014; Gaillard et al., 2015; Trondman et al., 2015) and land cover and land use change models (Gaillard et al., 2008, 2010; Klein Goldewijk et al., 2011).

In North America, PPEs have been calculated for a small number of trees and herbs in the tall-grass prairie of Kansas (Commerford et al., 2013) and the hardwood forests of Minnesota (Calcote, 1995) using fossil pollen data from forest hollows and small ponds, and modern vegetation data from forest surveys and vegetation maps. The complete LRA has been tested using pollen data from Michigan and Wisconsin (Sugita et al., 2010). However, more studies are needed due to the vastness and topographic complexity of the continent and the diversity of biomes, ecological communities and plant species that exist there.

The primary objectives of this study are to (1) obtain PPEs for 8 major taxa in the eastern deciduous forest using pollen data from small lakes, (2) test the use of spatial interpolation methods and large-scale vegetation survey data as inputs to compute PPEs, and (3) present quantitative estimates of past vegetation cover based on REVEALS in relation to pollen percentage diagrams from the region. This study is the first attempt to calculate PPEs and quantify regional vegetation composition using REVEALS in the temperate deciduous forest of eastern Canada. The vegetation history of this area is well-known due to numerous pollen studies (e.g., Davis and Webb III, 1975; Richard, 1978; Webb III et al., 1978; Savoie and Richard, 1979; Anderson, 1985; Richard, 1995; Lavoie and Filion, 2001; Paquette and Gajewski, 2013; Lafontaine-Boyer and Gajewski, 2014) but changes in vegetation cover have yet to be quantified.

We calculate PPEs for *Abies*, *Acer*, *Fagus*, *Picea*, *Pinus*, *Quercus* and *Tsuga* relative to *Betula* using modern pollen data from the uppermost cm of lake sediments from 16 small lakes in

southeastern Quebec, and use fossil pollen records from four small lakes with the REVEALS model to study changes in vegetation throughout the Holocene. An innovation in this study is the use of an extensive spatial dataset of absolute, species-specific canopy cover for our modern vegetation estimates (Saucier et al., 1998; MRNFQ, 2006). We interpolate values from the vegetation survey in a GIS to quantify the vegetation in the vicinity of the lakes for which modern pollen data were available.

### 6.3 Study area

The study area covers  $5 \times 10^4$  km<sup>2</sup> of the temperate deciduous forest within the region of the St. Lawrence Lowlands in southeastern Quebec, Canada (Fig. 6.1). It is composed of temperate deciduous taxa and is located just south of the boreal-temperate forest ecotone of Quebec. Reconstructions of the Laurentide ice sheet deglaciation show that recession began in the south around 13,000 years BP, and the entire area was ice-free by 7,000 years BP (Dyke, 2004), following the collapse of the Hudson Bay dome of the ice sheet around 8,000 years BP (Shuman et al., 2002). Continental freshwater runoff from melting ice and a temporary embayment of the Atlantic Ocean led to the formation of the brackish Champlain Sea which flooded the St. Lawrence and Champlain Valleys to the south of the study region between 13,000 and 9,000 BP (Parent and Occhietti, 1999; Cronin et al., 2008). The underlying geology of the area is Precambrian igneous and metamorphic rock characteristic of the Grenville geological province (Rivers et al., 1989). The southern portion of the region covers low-lying flatlands with moraines and small hills while the northern portion abuts the Laurentian mountain chain and is higher in elevation by several hundred meters (Fig. 6.1c). The area has a humid continental climate, with large seasonal temperature differences. Summers are warm and humid (mean July temperature of 19.3 °C; precipitation 121.4 mm) and winters are cold and relatively dry (mean January temperature of -12.8 °C; precipitation 86.6 mm).

The distribution of trees and plant associations in the study area has changed significantly since the late Pleistocene. During the early stages of the post-glacial period, the southern portion of the study area was dominated by herbaceous and shrub tundra (Muller and Richard, 2001) while the northern portion remained covered by the Laurentide ice sheet. Boreal taxa including *Picea*, *Pinus*, and *Abies* quickly migrated into the area during the warming trend of the early Holocene (Williams et al., 2004). At that time, hardwood species including *Acer*, *Betula*, and *Quercus* were also present but were less abundant than the boreal taxa (Muller and Richard, 2001). A distinct increase in *Populus* during the early Holocene is also documented (Richard and Poulin, 1976; Mott, 1978; Peros et al., 2008). Following the collapse of the Hudson Bay dome of the ice sheet, the modern boreal forest became established across the entire area (Huntley and Webb III, 1988; Shuman et al., 2002), although this formation was not stationary and the ranges of individual taxa varied significantly through time. A decrease in *Picea* beginning ca. 9,500 BP and a subsequent increase in the abundance of hardwood species including *Acer*, *Quercus* and *Betula* marked the formation of the cool mixed forest (Huntley and Webb III, 1988; Williams et al., 2004), after which the *Tsuga-Acer-Betula* forest appeared and persisted until the present day (Ritchie, 1987).

Land survey studies indicate that at the time of European settlement, *Abies*, *Betula*, *Juniperus*, *Picea*, and *Pinus* were growing in the area (Boucher et al., 2009). *Acer* and *Quercus* trees occurred in low numbers and in unevenly distributed patches. Logging began following French settlement in the early 1600s (Lessard, 2014), and greatly accelerated in the late 19<sup>th</sup> century to industrial scale. The forest permitted extensive timber trade and market activities due to its proximity to the St. Lawrence River system which expedited the movement of the trees. *Picea* and *Pinus* were most often selected for trade (Dupuis et al., 2011) because of the quality of the wood and growth patterns of the trees. Today, the forest is semi-open with dense patches of mixed broadleaf (*Quercus*, *Acer*, *Betula*) and coniferous (*Pinus*, *Abies*, *Picea*) taxa; the entire area is dotted with small ponds, bogs and extensive peatlands.

## 6.4 Methods

The sections below outline the steps taken to obtain a reconstruction of Holocene vegetation change in the study area, for seven tree taxa relative to *Betula*. The choice of pollen sites, and a summary of the modern and fossil pollen data, and the modern vegetation data, are described in sections 3.1 and 3.2. Relative PPEs were calculated using the ERV method and modern pollen data from 16 small lakes (section 3.3). These were then used as input, together with fossil pollen records from four small lakes in the area, in a single REVEALS model run (section 3.4).

### 6.4.1 Pollen data

Modern pollen data were obtained from the North American Modern Pollen Database (Whitmore et al., 2005; [www.lpc.uottawa.ca/data/modern/index.html](http://www.lpc.uottawa.ca/data/modern/index.html)) and used to calculate relative Pollen Productivity Estimates (PPEs). Fossil data from the Neotoma Paleoecology Database ([www.neotomadb.org](http://www.neotomadb.org)) were used to reconstruct past vegetation using the REVEALS model. Sites from these databases were chosen if they (1) were within the provincial boundary of Quebec, (2) were within the boundary of the temperate continental forest as defined by the North American Forest Commission (NAFC, 2011), (3) were lakes, (4) received pollen primarily by wind, i.e. without large river inputs, and (5) were within a 200 km distance from the shore of the St. Lawrence River and Estuary (to maintain forest homogeneity). This resulted in a study area of less than 10<sup>6</sup> km<sup>2</sup> in size, which is considered the maximum spatial extent for application of the REVEALS model (Abraham et al., 2014; Trondman et al., 2016).

The selection process resulted in 16 lake sites with modern pollen records, and four lakes with fossil pollen records spanning the Holocene (locations in [Fig. 6.1c](#)). Most lakes were relatively isolated, accessible only by a single road with limited modern human influence; however some were located in close proximity to small towns or were used for year-round recreational activities. Lake radii varied between 17 and 214 m ([Table 6.1](#)), thus all sites are considered small according to the categorization outlined in Sugita (2007a).

The taxa chosen were *Abies*, *Acer*, *Betula*, *Fagus*, *Picea*, *Populus*, *Quercus*, and *Tsuga*. These taxa are major components of the deciduous forest of southern Quebec and are found throughout

the study area. Together with *Pinus* (which is not included here), they consistently made up a significant portion of the total pollen sum across all sites included in this study. Although the study region is presently a temperate area, conditions were significantly different in the past, thus *Picea* was chosen since it is a key taxon in early and late Holocene forests and exhibited changes in abundance throughout the Holocene, including an increase during the Neoglacial. *Populus* was included since previous studies have shown a large increase in abundance during the early Holocene. *Fagus*, *Quercus* and *Tsuga* were studied since they are key indicators of the development of the modern deciduous forest. Lastly, *Abies*, *Acer* and *Betula* were chosen due to their north-south gradients in regional abundance. *Pinus* is arguably a key taxon with similar importance but it was left out since it lacked a spatial gradient.

To evaluate site-to-site differences and provide a summary of the data, a Principal Components Analysis (PCA) of the 16 modern pollen records was done. This helped assess the expected correlation between sites and see if there were any extreme values (outlying pollen records) that would need to be removed from the analysis. A second PCA was performed using data from the fossil pollen assemblages from the four lakes. Pollen percentages were calculated on the sum of the eight taxa and analyzed using the statistical software R (version 3.3.2; R Core Team, 2016) and the devtools, ggplot2, ggbiplot and rgl packages (Wickham, 2009; Vu, 2009; Adler et al., 2016; Hadley and Chang, 2016).

#### 6.4.2 Vegetation data

A comprehensive inventory of the forests of the province of Quebec covering more than 700,000 km<sup>2</sup> was used for vegetation data (Saucier et al., 1998; MRNFQ, 2006). In southern Quebec (Fig. 6.1c), 11,111 observation points of 400 m<sup>2</sup> were surveyed to obtain species-specific canopy-layer densities (Government of Quebec, 1994). For each observation point, the density of tree canopy cover of each species was visually estimated; the total density of cover may exceed 400 m<sup>2</sup> due to multiple vegetation layers. Estimations of density cover for 6 vegetation layers based on the height of each tree trunk (moss and lichen, herbs, < 1 m, 1 – 4 m, 4 – 7 m, and > 7 m) were made in 8 classes: 0%, < 1%, <5%, 6 – 25%, 26 – 40%, 41 – 60%, 61 – 80% and 81 – 100%.

For the purposes of this study, density cover was rounded to 0%, 1%, 5%, 10%, 30%, 50%, 70%, or 90% and only data from the uppermost vegetation layer were used. Most trees at this height would be mature and actively producing pollen and pollen from the upper canopy is most easily transported to lakes. Percentages of cover were then converted to square meters since absolute values are preferred when applying REVEALS. In this study, the REVEALS model was run by grouping together species belonging to the same genus and summing their respective densities for each observation point.

The vegetation data were used to create single taxon interpolated raster layers in ArcMap v.10.4. Interpolations were done using Inverse Distance Weighting (IDW) methods in the Geostatistical Analyst extension with a maximum of three nearest neighbours to ensure the output surfaces were not over-smoothed. To test the effects of (1) smoothing by different interpolation methods

and (2) the number of neighbouring data points, sensitivity analyses were done that included different combinations of interpolation methods and numbers of neighbours. Although the methods did not have much of an effect on the output, the number of neighbours did, since it directly influenced the smoothness of the surface. The final IDW parameters that were used were (1) a power of one, (2) a variable radius, and (3) a maximum number of three nearest neighbours. With IDW, the inverse of the distance is raised to a power, thus a low power creates an interpolated surface that is smoother. A variable radius ensures the chosen number of nearest neighbours is always used. Three nearest neighbours were used to improve processing speed and reduce the influence of distant points with poor spatial correlation on interpolated values. Mapped surfaces of interpolated plant canopy cover were made for all eight taxa.

At each of the 16 sites with modern pollen records ([Fig. 6.1c](#)), consecutive rings 200 m in width were drawn beginning at 100 m (the average radius of a lake) out to 3100 m using the Buffer tool in ArcMap. Two-hundred-meter-wide rings were chosen since the forest contains large open areas, the observation points from the vegetation survey were on average eight km apart, and smaller widths caused abnormal likelihood function scores. The mean abundance of each taxon in each consecutive 200-meter-wide ring between 100 and 3,100 m around each lake site was extracted using the Zonal Statistics tool and the Python plugin in QGIS v2.14.3 and input into the ERV analysis. These values were then compared locally and across all 16 lakes in the study area using a PCA.

#### 6.4.3 Extended R-Value modelling

A revised version (Sugita, 1994) of the Extended R-Value (ERV) model (Parsons and Prentice, 1981; Prentice and Parsons, 1983) was used to estimate relative Pollen Productivity Estimates (PPEs) for individual taxa as well as the Relevant Source Area of Pollen (RSAP *sensu* Sugita, 1994) using a maximum likelihood method. Three ERV submodels exist which differ in the way they treat the constant background pollen component. With ERV submodel 1, the background component is the taxon-specific proportion of pollen loading coming from beyond the RSAP in relation to the total pollen loading on the lake surface (Parsons and Prentice, 1981). In ERV submodel 2, the background component is the ratio of species-specific pollen loading coming from beyond the RSAP to total plant abundance for the whole assemblage within the RSAP (Prentice and Parsons, 1983). ERV submodel 3 assumes constant background pollen loading from beyond the RSAP (Sugita, 1994). ERV submodels 1 and 2 are most appropriate when vegetation data are expressed in percentages; ERV submodel 3 can be used when vegetation data are absolute values.

To run an ERV submodel, several parameters must be defined. Our PPEs were calculated in version 1.3.0 of the ERV model program (Sugita, 1994) using:

- a lake radius of 100 m for all sites
- previously published taxon-specific fall speeds of pollen ([Table 6.2](#))
- an average wind speed of 3 m/s

- *Betula* as a reference taxon
- an inverse distance weighting function ( $1/d$ ) for approximating pollen dispersal

In this study, the Sugita (1993) pollen deposition model for lakes and ponds was used, and a  $1/d$  distance weighting function was chosen since this is the simplest way to weight vegetation data and it is often found to be the most appropriate choice. Further, a function of  $1/d$  assumes above-canopy air flow is the only method of transport which is likely in this case. Because *Betula* displayed the largest gradient in both the pollen and vegetation data across the study region, this taxon was chosen as the reference taxon for ERV analysis (i.e., PPE=1).

It is possible to compare the goodness of fit between the observed data and the modelled data produced by the three different submodels since the ERV program provides likelihood function scores (Sugita, 1994). The submodel with the best fit will have the lowest likelihood function score. The RSAP is most often estimated using a moving-window regression approach, the distance at which the likelihood function curve reaches an asymptote, or the distance at which the difference between likelihood values becomes 0.1 or less over 50 m (Commerford et al., 2013). In this study, the size of the consecutive 200 m rings within which distance-weighted plant abundance was measured made it difficult to use such methods. Thus, the RSAP was estimated based on the nature of the likelihood function scores *and* the relationship between PPEs and distance from the lakes. The distance at which (1) an asymptote appears in the likelihood function scores, and (2) the PPEs calculated for each 200 m increment are most stable, was chosen as the maximum distance for the RSAP. We calculated PPEs three times, using the three different submodels, and the resulting likelihood function scores were compared.

#### 6.4.4 REVEALS modelling

The REVEALS model was developed to quantitatively reconstruct vegetation cover at a regional scale using fossil pollen counts. Detailed descriptions of the model as well as its assumptions and limitations are found in Sugita (2007a), Hellman et al. (2008) and Soepboer et al. (2010).

The REVEALS model requires pollen counts to be summed into broad time windows that are consistent across sites. For this purpose, age-depth chronologies for sites Geai, Manitou, Sam and Tortue (Savoie and Richard, 1979; Gauthier, 1981; Neotoma, 2017; Richard, unpublished) were established based on the IntCal13 calibration curve and the Bacon age-modelling program (Reimer et al., 2013; Blaauw and Christen, 2011). To ensure each time window contained pollen counts from as many of the four sites as possible, pollen counts were aggregated into 300-year bins spanning the last 13,000 years. The REVEALS 4.2.2 program (Sugita, 2007a) was then run on each of the four sites using:

- site-specific lake radii ([Table 6.1](#))
- previously published taxon-specific fall speeds of pollen ([Table 6.2](#))
- estimates of PPEs obtained from ERV analysis
- the taxon-specific variance and covariance in PPEs obtained from ERV analysis
- an average wind speed of 3 m/s, and

- a maximum extent of 50 km away from lake edges for incoming pollen (Mazier et al., 2012)

This resulted in an estimate of vegetation cover per taxon, per lake, for each time interval of 300 years. These were then averaged to obtain an estimate of vegetation cover per taxon for each time interval that was representative of the region as a whole. The average of the fossil pollen data from lakes Geai, Manitou, Sam and Tortue and the average REVEALS reconstruction were analyzed stratigraphically and with a PCA.

## 6.5 Results

### 6.5.1 Pollen data

Modern pollen assemblages from 16 small lakes in the deciduous forest of southeastern Quebec (Fig. 6.1) are primarily composed of *Betula*, *Picea*, *Tsuga* and *Fagus*, which together make up over 80% of the pollen sum of the eight taxa across all sites. *Quercus* and *Abies* were also important components of the pollen assemblages, but to a lesser extent. Three *Betula* species (*B. alleghaniensis*, *B. papyrifera* and *B. populifolia*) and three *Quercus* species (*Q. alba*, *Q. macrocarpa* and *Q. rubra*) occur in the study area although some of these are less abundant in the north. Three species of *Acer* (*A. spicatum*, *A. saccharum*, *A. rubrum*) occur across the area and an additional three (*A. pensylvanicum*, *A. saccharinum* and *A. nigrum*) are only found in the south. Three species of *Populus* (*P. balsamifera*, *P. grandidentata* and *P. tremuloides*) and one species of *Abies* (*A. balsamea*), *Tsuga* (*T. canadensis*) and *Fagus* (*F. grandifolia*) occur across the entire range. *Picea glauca* and *Picea mariana* co-exist across the entire range, and *Picea rubens* is found to the east along the St. Lawrence River and Estuary.

A principal components analysis (PCA) of the modern pollen revealed no outliers in the pollen assemblages from the 16 sites (Fig. 6.2), and a north-south gradient is apparent. The loadings for *Betula*, *Abies* and *Picea* on the left side of the graph align with the lakes from the central and northern part of the study region.

The four lakes sites used with the REVEALS model were chosen for their central location within the study area (Fig. 6.1c), as well as the length, quality and similarity in taxonomic composition. To assess similarity, the fossil records from each lake were individually analyzed using a PCA (Fig. 6.3a-d). Older samples align with *Populus*, *Picea* and *Quercus*, and younger samples show increases in the abundance of deciduous taxa following the development of the temperate-mixed forest. The PCA loadings (Fig. 6.3a-d) show only slight differences in the associations between the taxa presumably caused by topographic effects on pollen dispersal and geographic location. However, time series of the PCA scores (Fig. 6.3e) show that the pollen diagrams of these lakes are largely similar suggesting the use of these lakes with REVEALS is appropriate.

### 6.5.2 Vegetation data

The north-south gradient visible in the modern pollen data from the 16 lakes (Fig. 6.2) is also seen in the maps of interpolated modern vegetation cover (Fig. 6.4). For example, *Fagus*, *Quercus* and *Tsuga* are more abundant in the southern portion of the study area, and *Picea* and *Abies* are characteristic of the north. Values of *Abies* are relatively low, perhaps in part since we only used the upper canopy values from the vegetation survey. *Acer* and *Fagus* tend to grow in the St. Lawrence Valley, although the range of *Acer* is much wider. *Populus* and *Betula* are abundant across the entire area, although *Populus* stands occur in patches.

The broad-scale gradient in the vegetation is retained when the vegetation data (mean abundance in consecutive 200 m wide rings within the area 100 to 3100 m away from the lake) is extracted from the area around each lake (Fig. 6.5). There is some variation in forest composition within the area directly surrounding each lake, suggesting that the interpolation method is not over-smoothing; in addition, site-to-site differences in vegetation composition are quite prominent. Thus, the maps (Fig. 6.4) and the ordination (Fig. 6.5) capture both the large-scale biogeographic trends, as well as regional and sub-regional patterns in the data.

### 6.5.3 ERV analysis

The ERV model was run three times using the different submodels, to test which description of the background pollen component was most suitable for calculating relative Pollen Productivity Estimates (PPEs) and the Relevant Source Area of Pollen (RSAP) in this study region. The ERV submodels use a distance-weighting function (in this case  $1/d$ ) which causes the maximum likelihood curve to decrease and approach an asymptote as the distance from the sites increases. The RSAP is thus the area between the lake and the distance at which the asymptote is reached (Commerford et al., 2013).

Submodel 1 did not produce a very smooth likelihood curve (Fig. 6.6), and the likelihood values were the highest among the three submodels, so it was not used in further analysis. It is possible that the site-to-site variation in background pollen was too great for use with this submodel (Commerford et al., 2013). The results from submodels 2 and 3 were more plausible, and each produced a likelihood curve from which the RSAP could be estimated (Fig. 6.6). Because the likelihood curve from submodel 2 had the lowest values it was chosen as the best fit for these data. Ideally, the likelihood values should stabilize to an asymptote, even with submodel 2 which was the most appropriate for our data, but Figure 6.6 shows that this is not the case. This is most likely due to the size of the 200 m rings used to calculate distance-weighted plant abundance. In this case, there are too few data points to use the moving-window regression method to calculate the RSAP (e.g., Sugita, 2007b) thus the RSAP was identified based on (1) the slope of the likelihood curve (Fig. 6.6), and (2) the nature of the PPE values calculated for each 200 m increment away from the lakes, which were most stable at 1600 m. This, combined with previous results from similar ecosystems that showed RSAP values ranging between 300 and 1700 m,

justified our choice for a 1600 m RSAP. We suggest the relationship between pollen counts and vegetation cover does not improve beyond a distance of 1600 m in this ecosystem.

**Figure 6.7** shows the relationship between adjusted pollen and vegetation data based on ERV analysis for the eight taxa. The abundance of *Betula* varied the greatest across the study area based on both the pollen and the vegetation data. *Quercus* and *Acer* also had large variations in abundance between sites in both the pollen and vegetation data, respectively. *Fagus* and *Tsuga* show almost no variation in vegetation abundance, but the proportion of pollen grains varied substantially across sites. This is possibly a product of the reproduction strategies of *Fagus* and *Tsuga*.

*Fagus* occasionally reproduces vegetatively but can also produce large amounts of pollen under ideal growing conditions. This could lead to different pollen production across geographically-separated but homogenous vegetation zones. In this study, the relationship between adjusted *Fagus* pollen and vegetation could vary spatially depending on the local reproduction strategy of the trees, resulting in the curve shown in **Figure 6.7**. *Tsuga* tends to grow in moist areas along lake edges and in valleys and tends to produce very high amounts of pollen. *Tsuga* trees may have infrequently been recorded in the vegetation survey because of this, leading to low variation in the vegetation data – but high pollen counts make it an important component in pollen assemblages (**Fig. 6.7**). The abundance of the remaining taxa, inferred from both the pollen and the vegetation data, varied moderately in comparison.

#### 6.5.4 REVEALS model

The REVEALS estimates of vegetation cover, which were calculated using the PPEs from ERV submodel 2 (**Table 6.3**), were compared to the original pollen percentage data from lakes Geai, Manitou, Sam and Tortue. Because we are interested in the regional vegetation history of the area, the results from the four lakes were averaged and plotted as a stratigraphic diagram (**Fig. 6.8**) then evaluated using PCA (**Fig. 6.9**). The calculated mean in **Figure 6.8b** represents taxon-specific regional vegetation cover estimated by REVEALS every 300 years (pollen counts from all four lakes were used when possible; exceptions occurred when pollen counts were only available for 2-3 lakes in a given time interval). REVEALS is generally run with a time window of 100 – 500 years (e.g., Fyfe et al., 2013; Trondman et al., 2015; Mehl and Hjelle, 2016), placing our choice of a 300-year time interval within the typical range.

Although the pollen data and the REVEALS reconstruction provide broadly similar records of regional vegetation change, the REVEALS results show some important differences that enable a modification of the interpretation of the original pollen diagram. The average pollen diagram computed from the four sites shows an early period during which there is an abundance of *Picea* (12,000 – 11,000 cal BP), which is followed by peaks in *Populus*, *Quercus*, and *Tsuga* around 10,500, 8,000 and 7,500 cal BP, respectively (**Fig. 6.8a**). The PCA reflects this, with *Populus* and *Picea* positively associated with the first component and *Acer*, *Fagus* and *Tsuga* negatively associated. *Abies*, *Betula* and *Quercus* are positively associated with the second component (**Fig.**

**6.9a**). Scores of the older samples tend to be negative on component 1 and younger samples are positive (**Fig. 6.9a**).

These trends are characteristic of the region, and are well documented (e.g., Mott, 1978; Williams et al., 2004; Munoz et al., 2010; Peros et al., 2010). The peak in *Populus* during the early Holocene is significantly higher in the REVEALS estimates; it remained in low abundance for the remainder of the Holocene (**Fig. 6.8b**). The percentage of *Betula* remained high throughout the Holocene, a trend which has also been previously identified (Gauthier, 1981; Labelle and Richard, 1981; Comtois, 1982).

The REVEALS reconstruction suggests *Acer* and *Abies* have been important components of the canopy since 9,000 and 10,000 cal BP, respectively, although the abundance of *Abies* is more variable (**Fig. 6.8b**). Both are underrepresented in the pollen data (**Fig. 6.8a**). *Betula* was very abundant, indicated by the REVEALS estimates and the percentage data, but the former shows a gradual increase to towards maximum abundance between 10,000 and 5,000 cal BP and a continuous decrease thereafter (**Fig. 6.8b**). *Betula* is more closely associated with the deciduous taxa of the late Holocene than the mid-Holocene rise in *Abies* abundance (**Fig. 6.9b**). This is quite different from the percentage data, which imply *Betula* abundance did not change much in the past 9,000 years (**Fig. 6.8a**). *Fagus*, *Quercus* and *Tsuga* are overrepresented in the pollen data relative to the vegetation on the landscape, but the overall trends remain the same. The REVEALS results show *Quercus* was more important later than estimated by the pollen record (**Fig. 6.9b**). The trends in *Picea* remain the same as in the pollen data, however the REVEALS estimates show it was more prominent on the landscape, and its increase in abundance during the Neoglacial is more significant than in the percentage diagram (**Fig. 6.8b**).

## **6.6 Discussion**

In this study, we compare pollen percentage data with REVEALS model-based reconstructions of vegetation cover in the southeastern deciduous forest of Quebec for the last 13,000 years. The REVEALS model has been used in many studies of various European forests over the last two decades and this work has led to the compilation of standard sets of pollen productivity estimates (PPEs) for these regions (e.g., Broström et al., 2008; Mazier et al., 2012). The present study is the first attempt to use the REVEALS model in a North American deciduous forest, and the first time the REVEALS model has been applied in Canada. We obtained PPEs for 7 tree taxa relative to *Betula* based on pollen data from 16 small lakes in the vicinity of the St. Lawrence River using spatial interpolation methods and large-scale vegetation survey data, and compared quantitative estimates of past vegetation cover to pollen percentage diagrams from the region.

Our focus was on eight tree taxa characteristic of the deciduous forest, a limitation imposed by the number of surface pollen samples available in regional databases, which meant a component of the landscape (i.e., herbs, grasses, other trees) was left out. Nevertheless, these eight taxa are major contributors to the regional pollen profile and we reconstruct a significant component of the pollen rain. For most taxa, the relation between plant abundance and pollen percentages is as expected based on previous qualitative research, with a few exceptions that we describe below,

following our discussion of using vegetation survey data and spatial interpolation methods to estimate PPEs.

### 6.6.1 Modern vegetation data and interpolation methods

A detailed estimate of the relative abundance of individual taxa in the region surrounding the study lakes is required when using ERV submodels to obtain PPEs. Previous studies have often used “walking-in-circles” field survey methods (e.g., Broström et al., 2004), stratified random sampling and quadrat methods (e.g., Bunting et al., 2005; Abraham and Kozáková, 2012), aerial photographs and satellite imagery (e.g., Broström et al., 2008), or digitized vegetation maps compiled previously (e.g., Hjelle et al., 2015). In contrast, we used a vegetation survey that provided species-level observations of absolute tree cover across an area greater than 50,000 km<sup>2</sup>. This presented a rare opportunity to use a systematically-collected, large-scale vegetation dataset.

The nature of the vegetation survey allowed us to efficiently quantify the vegetation in the vicinity of the pollen sites using spatial interpolation. This provided a more detailed analysis than could have been obtained by sampling within a select few quadrats and extrapolating to the region surrounding the lakes. The distance between observation points and the interpolation method meant that the consecutive rings drawn between 100 and 3100 m away from the lakes needed to be large (i.e., 200 m wide). This provided suitable resolution for the application of the REVEALS model but it was not possible to use the moving-window regression method for the RSAP calculation. Instead, the RSAP (1600 m) was identified based on the nature of the log likelihood function (Fig. 6.6) and the slope of the PPE values calculated for each 200 m increment away from the lakes plotted against distance. The large width of the rings will down-weight the vegetation closer to the lakes (since the vegetation data is smoothed across a large area), and up-weight the vegetation farther away from the lakes.

### 6.6.2 Pollen Productivity and Source Area

The ERV model calculates taxon-specific pollen productivity values by incorporating pollen counts, pollen fall speed, lake size, distance from the edge of the lake, and wind speed. The REVEALS model then produces adjusted plant cover values based on taxon-specific pollen productivities. Thus, even if the pollen sites are small, and even if the landscape is heterogeneous, site-to-site differences in background pollen values should be minimal, and pollen grains coming from beyond the RSAP should have little impact on pollen assemblages. In other words, the correlation between plant abundance and pollen loading on a lake surface does not improve beyond the distance of the RSAP. Previous studies in environmental settings similar to the forests of southern Quebec have shown that RSAP values ranging between 300 and 1700 m are characteristic of lakes approximately 200 m in diameter (Sugita, 1994; Sugita et al., 1999; Nielsen and Sugita, 2005). The estimated RSAP of 1600 m for the temperate deciduous forest of southern Quebec is consistent with this range and suggests the vegetation surveys interpolated from sites 1.5-2 km away from the lake should be suitable for studies of this kind.

The PPEs presented here (Table 6.3) are calculated relative to *Betula* (PPE = 1) using ERV submodels 2 and 3. Broström et al. (2008) and Calcote (1995) describe *Betula* as a relatively

high producer of pollen that readily disperse. It is common for Poaceae to be chosen as a reference taxon (Broström et al., 2008) since it is abundant and widespread. However, our study focused on the dominant tree taxa of the deciduous forest, and *Betula* had the largest range of values in the pollen percentages and vegetation data, thus it was chosen as the reference taxon.

In our study, *Abies*, *Fagus*, and *Quercus* all had higher PPEs than *Betula*. This is in line with previous studies where these taxa were most often identified as intermediate pollen producers, depending on the regional setting (Sugita et al., 1999; Broström et al., 2008; Mazier et al., 2012). The least productive taxa were *Acer*, *Picea* and *Populus*. Low productivity of *Acer* and *Populus* was identified in previous studies (Calcote, 1995; Peros et al., 2008; Commerford et al., 2013). *Picea* tends to be highly productive in mountain regions (Broström et al., 2008) and was not very abundant within our study area (Fig. 6.4), thus a low PPE for the deciduous forest was expected. Our highest PPEs were associated with *Tsuga* since it is largely overrepresented in the pollen data but has a lower than average presence in the canopy. The standard error was low for *Tsuga* suggesting the PPEs are likely accurate and not outliers (Commerford et al., 2013).

These PPEs are the first obtained for the temperate deciduous forest of eastern Canada and are crucial to enable interpretation of past vegetation from pollen records. Future studies incorporating additional modern pollen samples will allow for additional taxa to be studied, and a more complete vegetation reconstruction involving trees, herbs and grasses.

### 6.6.3 Vegetation reconstruction

The history of the vegetation of southeastern Quebec has been extensively studied using pollen percentage data (e.g., Davis and Webb III, 1975; Richard, 1978; Webb III et al., 1978; Savoie and Richard, 1979; Lavoie and Filion, 2001; Paquette and Gajewski, 2013; Lafontaine-Boyer and Gajewski, 2014). However, new insight comes from the REVEALS estimates since they provide a means of interpreting changes in plant cover and forest composition which is not confounded by pollen over- and underrepresentation and taxon-specific productivity patterns (Fig. 6.8).

When compared to the pollen percentage data, the REVEALS reconstruction shows that prior to 11,500 cal BP, the landscape was dominated by *Picea* growing in cooler temperatures (Fig. 6.8). The reconstructed values for *Picea* reach a maximum of >70% whereas the maximum indicated by the pollen data is closer to 50%. Richard (1978; 1985) and Labelle and Richard (1981) have noted that southeastern Quebec was predominantly open spruce forest at this time, and this is consistent with the REVEALS estimates. Future work involving shrubs and herbaceous taxa, would be needed to confirm this interpretation. A decline in *Picea* coincided with the Younger Dryas cold period at which point the entire range of *Picea* began shifting northwards and to the west (Williams et al., 2004).

At this time, *Populus*, a pioneer species of disturbed sites, increased in abundance (Fig. 6.8b). *Populus* today has a range across North America and tends to be underrepresented in pollen spectra since its grains are fragile and degrade quickly in lake sediments (Richard, 1977b; Hadden, 1978; Campbell, 1999). The REVEALS results presented here suggest that *Populus* made up more than half of the canopy for a period of almost 2,000 years (Fig. 6.8b). Previous

studies of the vegetation history of Quebec have interpreted the *Populus* zone as extensive stands of aspen on the landscape, where it may have been the dominant tree on the landscape during a period of rapid environmental change (e.g., Richard and Poulin, 1976; Richard, 1977a, 1977b; Mott, 1978; Richard, 1978). Delcourt and Delcourt's (1987) summary of Holocene tree dynamics in and around the St. Lawrence River and Estuary suggested *Populus* comprised 60 – 80% of the trees at this time. Likewise, Peros et al. (2010) showed that *Populus* was abundant in sites ranging across northern North America prior to and after the initial onset of the Younger Dryas, and again during the Younger Dryas-early Holocene transition. They proposed that prior to 9,000 BP, *Populus* abundance was determined in part by climate, but also that it rapidly colonized land after disturbance or immediately after deglaciation. Further, increased biomass burning and fire activity in North America at the end of the Younger Dryas suggests conditions were ideal for the establishment and growth of drought-resistant *Populus* stands (Marlon et al., 2009; Peros et al., 2010). In the late Holocene, the decline in *Populus* was due to competition with other taxa. The *Populus* zone shown in the REVEALS reconstruction therefore is in line with previous work, which suggests extensive populations of aspen across the landscape during this time period. The difference in *Populus* values between the REVEALS reconstruction and the pollen percentages seen here are plausible, especially given that *Populus* seeds are widely dispersed, and that it rapidly colonizes after disturbances (Gonzalez et al. 2010).

When comparing the maximum *Populus* values in the pollen percentage data and the REVEALS reconstruction, differences in *Populus* abundance on the order of 40% can be seen (Fig. 6.8). Previous studies have also identified such differences in pollen percentage data and REVEALS reconstructions. For example, Åkesson et al. (2015) studied over 22 key taxa in a southeastern Swedish landscape and observed differences in landscape openness of 21 – 45% depending on the time period. Wang et al. (2011) looked at four key taxa from the Tibetan Plateau (*Artemisia*, *Chenopodiaceae*, *Poaceae*, *Cyperaceae*) and identified changes of up to 40% between pollen and vegetation reconstructions. Therefore, although the differences between the pollen percentages and the REVEALS reconstruction are large, they are within the limits seen in other regions and for other taxa, and the ecological interpretation is plausible.

Beginning around 9,000 cal BP, warming temperatures led to the formation of the eastern temperate deciduous forest (Williams et al., 2004). The REVEALS reconstruction shows a mixed forest with a canopy largely composed of *Abies*, *Acer* and *Betula*, but also containing other hardwood and coniferous taxa. This association persisted for the remainder of the Holocene. *Abies* produces large pollen grains that fall rapidly during dispersal (Godwin, 1984) and *Acer* trees are pollinated by both wind and insects. These characteristics mean these taxa are typically underrepresented in pollen assemblages; the REVEALS results suggest that *Abies* and *Acer* were major components of the canopy (Fig. 6.8b). Mott (1977) previously noted low *Acer* and *Abies* counts in the region but suggests these taxa were consistently important on the landscape since the mid-Holocene. Conversely, *Fagus* and *Quercus* were not as abundant on the landscape, as was shown in the pollen percentage data.

Around 8,200 years BP, there was a cold event characterized by a brief global reduction in temperature and atmospheric CO<sub>2</sub> (Thomas et al., 2007) and this was followed by a 2000-year period of cool temperatures in North America (Viau et al., 2008). *Tsuga* increased in abundance at this time, due to increased moisture availability and decreased fire activity, a trend which is

also seen in the northeastern United States (Munoz et al., 2010). The “*Tsuga* decline” beginning around 5,500 cal BP (Williams et al., 2004) has been attributed to insect activity (Davis, 1981b) or changes in climate (Shuman et al., 2004), and is clearly seen in the vegetation reconstruction (Fig. 6.8b). *Tsuga* again began to increase around 3,000 cal BP, due to a rise in regional moisture availability (Davis et al., 1992). The overrepresentation of *Tsuga* grains is well known (Bradshaw and Webb III, 1985), and the REVEALS reconstruction suggests that this tree was never abundant in the forests of the region.

The REVEALS reconstruction shows maximum *Betula* values between 7,000 and 4,500 cal BP, and this interpretation differed from the pollen percentage record. The increasing *Betula* values prior to 5,500 cal BP and the decreasing values thereafter changes our interpretation of this taxon in relation to prior pollen studies from the region (Gauthier, 1981; Labelle and Richard, 1981; Comtois, 1982) which showed a relatively stable but abundant presence of *Betula* throughout most of the Holocene (Fig. 6.8a). Comtois (1982) and Whitehead (1979) hypothesized this was partially due to the overrepresentation of highly productive *Betula* trees relative to other moderately productive taxa, but that site-specific changes in pollen influx could also be a factor.

In the past 2,000 years, an increase in *Picea* is indicative of Neoglacial cooling, but this increase and the overall abundance of *Picea* becomes more prominent when considering the REVEALS reconstruction. The increase in *Picea* has been observed in pollen records across a large part of eastern North America in the area of the boreal-temperate forest ecotone and is likely reflecting increases in both *Picea mariana* (black spruce) and *Picea glauca* (white spruce) in response to Neoglacial cooling (Gajewski, 1987; Gajewski et al., 1987; Richard, 1994; Paquette and Gajewski, 2013; Lafontaine-Boyer and Gajewski, 2014). Alternatively, increases of black spruce in bogs may have contributed to the regional pollen rain (Richard et al., 1992). The large-scale increase in *Picea* across much of the deciduous forest and northwards does suggest a regional increase in spruce in response to the climate variations of the Neoglacial.

The REVEALS model not only improved our understanding of past changes in the relative abundance of different taxa by accounting for pollen over- and under-representation (Fig. 6.8), but also produced results that better predicted the evolution of plant communities following the northward retreat of the Laurentide ice sheet (Fig. 6.9). For example, the pollen data suggest there was an early Holocene association between *Picea*, *Populus* and *Quercus*, with *Quercus* becoming an important part of the canopy shortly after the increase in *Populus* (Fig. 6.9a). *Quercus* also increased in southern Maine during the Little Ice Age, which was not expected (Gajewski, 1987), as the present range of *Quercus* is generally to the south of *Picea*.

The REVEALS reconstruction makes a distinction between the early dominance of *Picea* and *Populus* and a later, more gradual increase in *Quercus* (Fig. 6.9b). The model is likely accounting for the long-distance dispersal of wind-dispersed pollen grains from abundant *Quercus* stands growing in the central and northern United States. Similarly, the pollen data suggest a mid-Holocene association between *Betula* and *Abies* (Fig. 6.9a), but the REVEALS reconstruction shows *Betula* becoming a major component of the canopy at a later time in a predominately deciduous landscape (Fig. 6.9b). It is likely that the latter is capturing the strong association between yellow birch (*Betula alleghaniensis*) and eastern hemlock (*Tsuga*

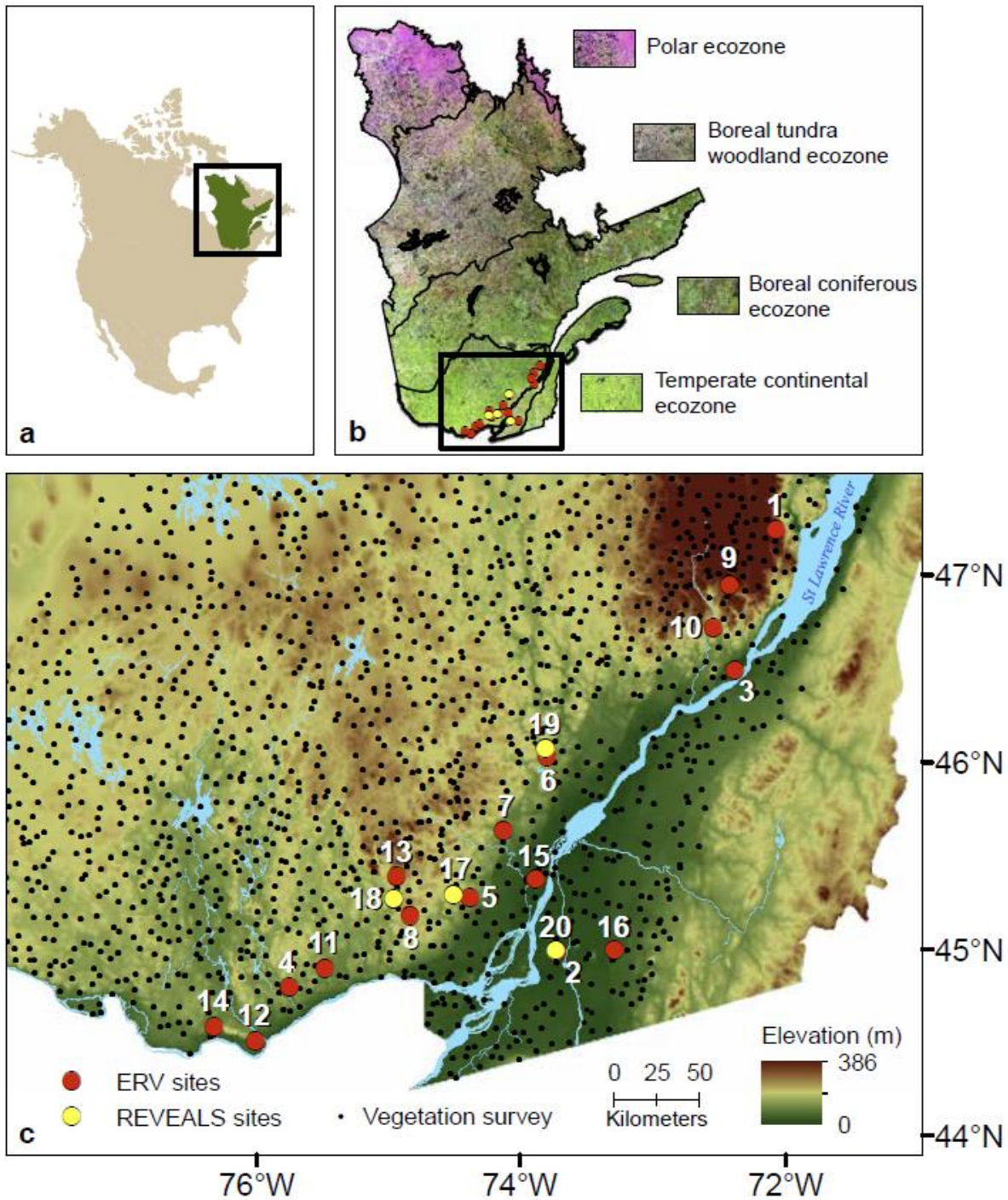
*canadensis*), both of which have similar ecological requirements for seed germination (Burns and Honkala, 1990a, 1990b).

This study is the first quantitative analysis of changes in Holocene vegetation cover in the temperate deciduous forest of eastern Canada based on the Regional Estimates of VEgetation Abundance from Large Sites (REVEALS) model (Sugita, 2007a). The REVEALS-based reconstruction of land cover change accounts for taxon-specific differences in pollen productivity, dispersal, and deposition, and identifies the over- and underrepresentation of certain pollen grains, producing results that are more representative of past forest dynamics than traditional pollen percentage diagrams. Access to an extensive, systematic vegetation survey of the entire southern region of Quebec presented a rare opportunity to use absolute values of present-day, species-specific canopy cover to calculate relative Pollen Productivity Estimates (PPEs). The PPEs ranged from 0.29 (*Acer*) to 17.14 (*Tsuga*) and the Relevant Source Area of Pollen (RSAP) for small lakes in this semi-open landscape was estimated to be 1600 m based on pollen counts from 16 small lakes that averaged 200 m in diameter.

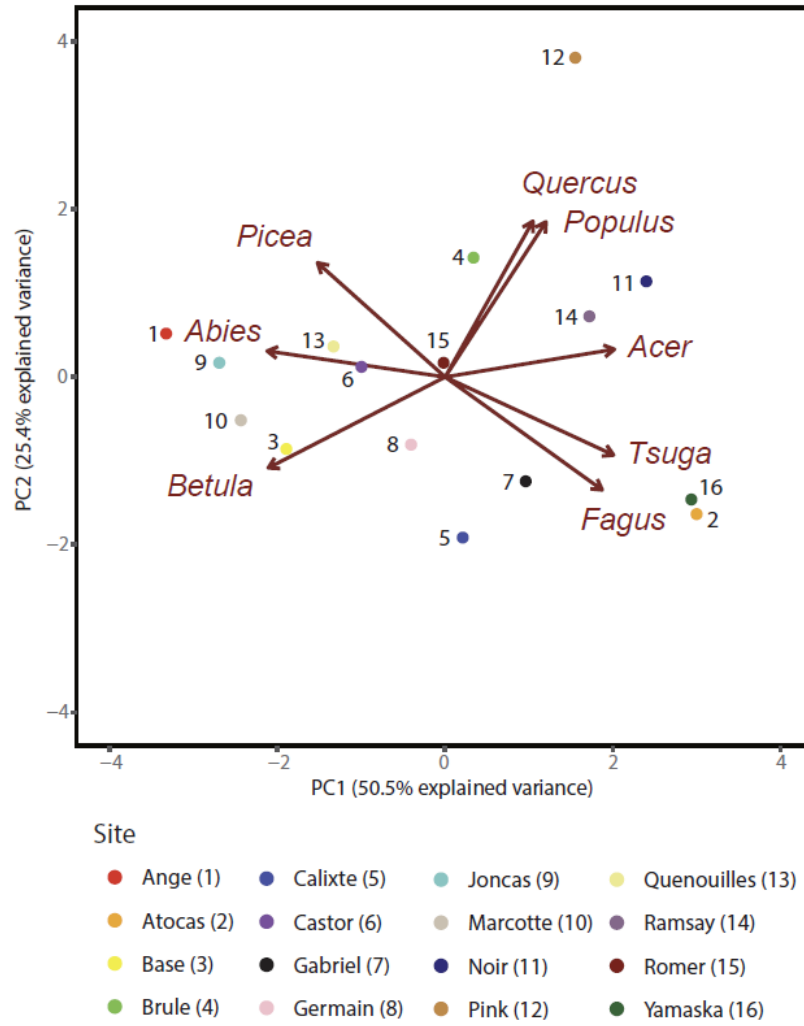
Future research testing alternative interpolation methods and adjusting the associated parameters would be constructive, especially in different environmental settings across the area of Quebec that is covered by the vegetation survey. Using a landscape ecology approach to estimate vegetation abundance including slope and orientation will allow for the inclusion of environmental effects in quantitative studies and better estimates of the tree abundance in the area surrounding the lakes. In addition, future work can use additional layers in the canopy, as the vegetation survey includes data from several canopy layers, thereby enabling a more precise estimate of canopy openness during the various time periods of the Holocene.

## **Acknowledgements**

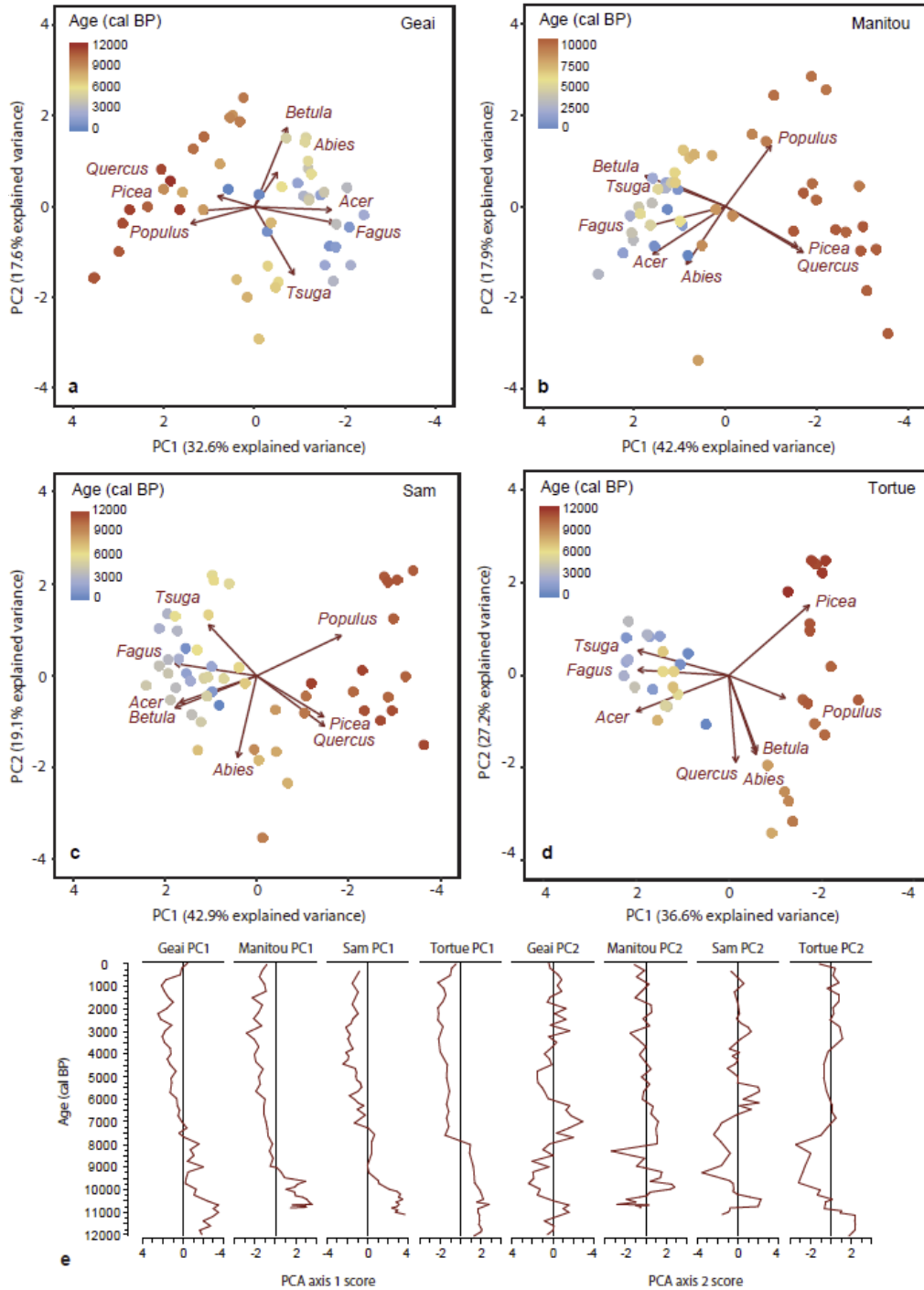
The authors wish to thank the developers of the North American Modern Pollen Database and the Neotoma Paleoecology Database, as well as all those who have contributed published and unpublished data, especially Pierre Richard. We thank the members of the Past Global Changes LandCover6k Working Group for their helpful discussions. We also thank Martin Blaauw for assistance with the Bacon program and Marie-José Gaillard-Lemdhahl for assistance with REVEALS. This research is funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) held by K.G. M.C. received funding from the Ontario Graduate Scholarship program and the Canadian Association of Palynologists.



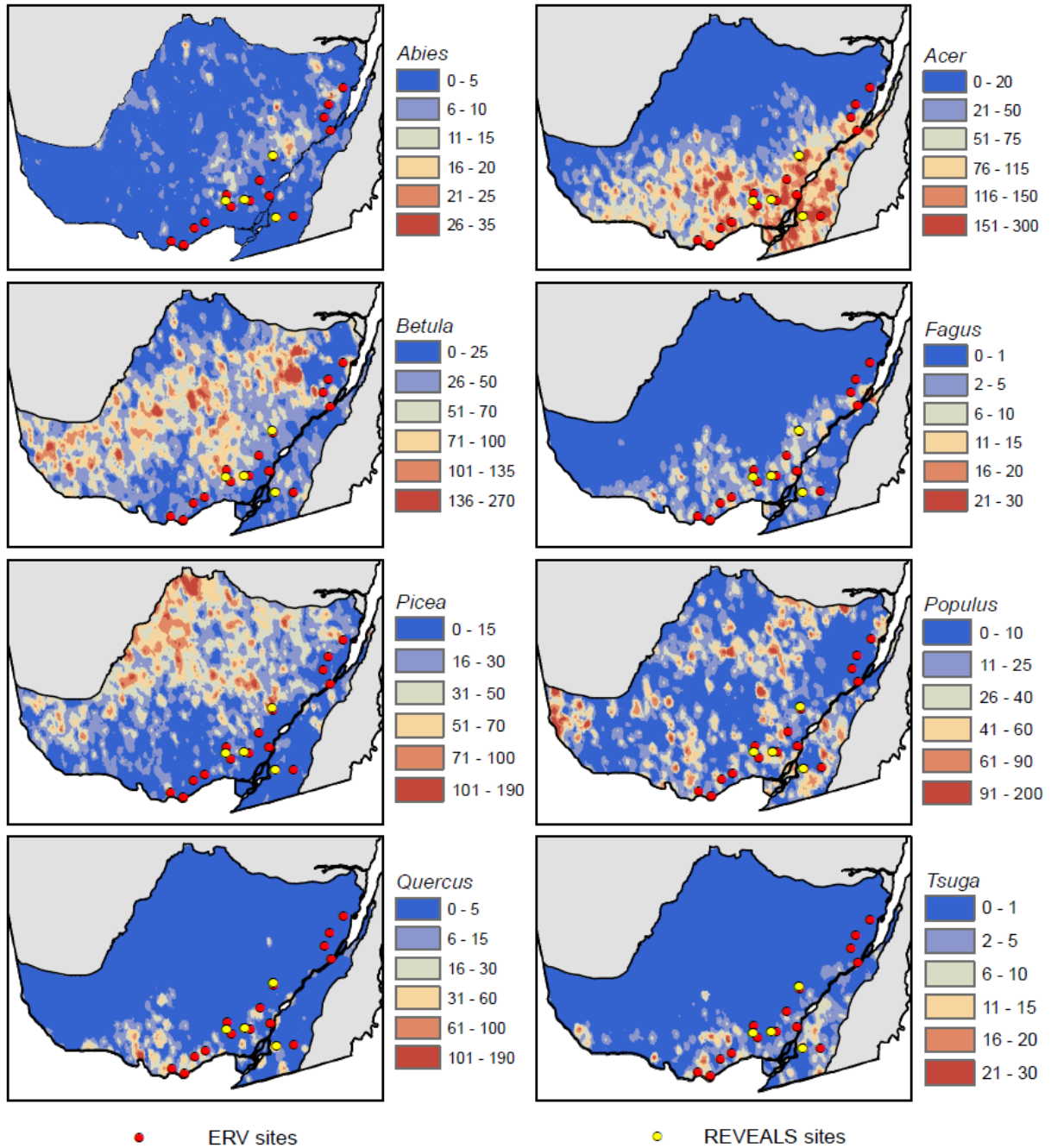
**Figure 6.1** The location of the province of Quebec within North America (a) and the variation in forest types within Quebec (b). The locations of observations sites from the modern vegetation survey (c, black dots; Government of Quebec, 1994; Saucier et al., 1998; MRNFQ, 2006) span the extent of the deciduous forest. The locations of the lake sites used for ERV analysis (c, red dots) and the lake sites used for REVEALS (c, yellow dots) are located within 200 km of the St. Lawrence Seaway



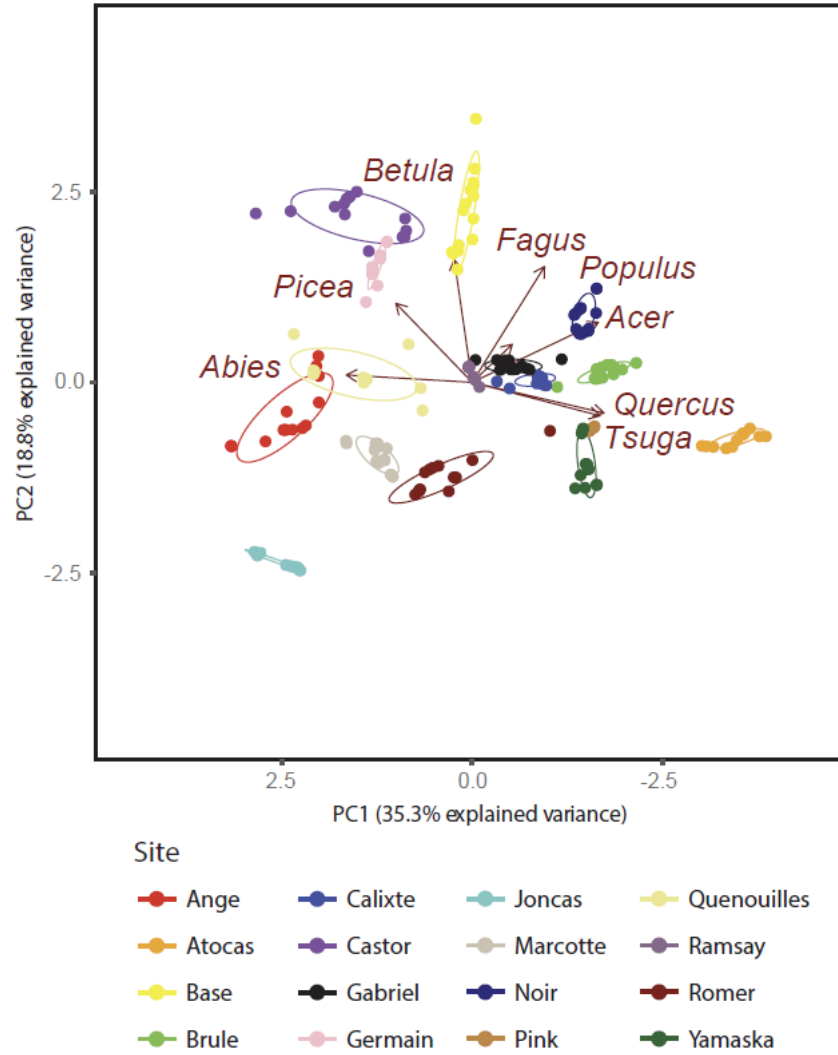
**Figure 6.2** Principal Component Analysis (PCA) of modern pollen data from 16 small lakes in southeastern Quebec used for ERV analysis.



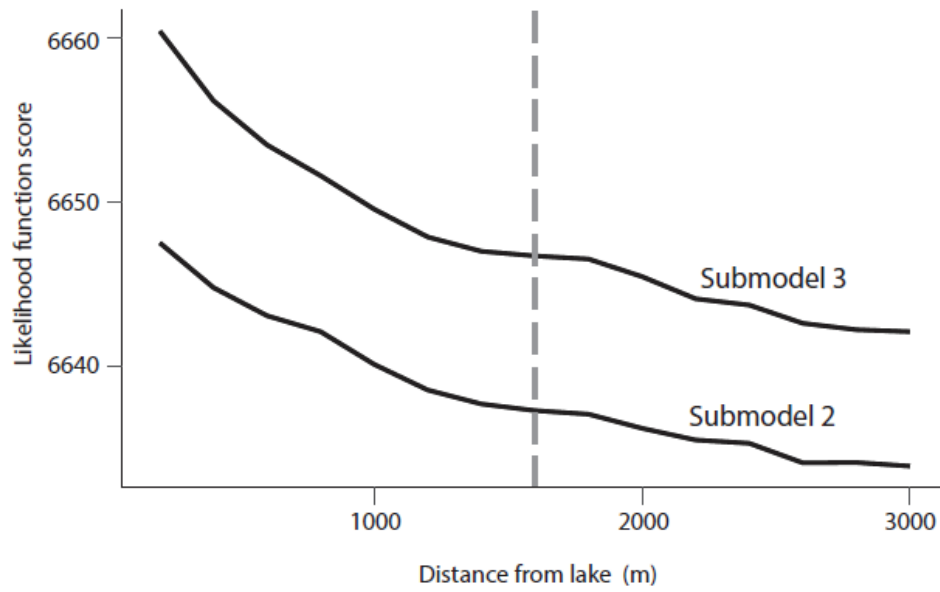
**Figure 6.3** Principal Component Analysis (PCA) of fossil pollen data from lakes Geai (a), Manitou (b), Sam (c) and Tortue (d) in southeastern Quebec used with the REVEALS model. Ages are in calibrated radiocarbon years before present. The scores (red lines) from the first and second principal components are also shown (e). Note the absolute values of the first and second principal components from Geai Lake are reversed to facilitate comparison. The zero line is also shown (thin black line).



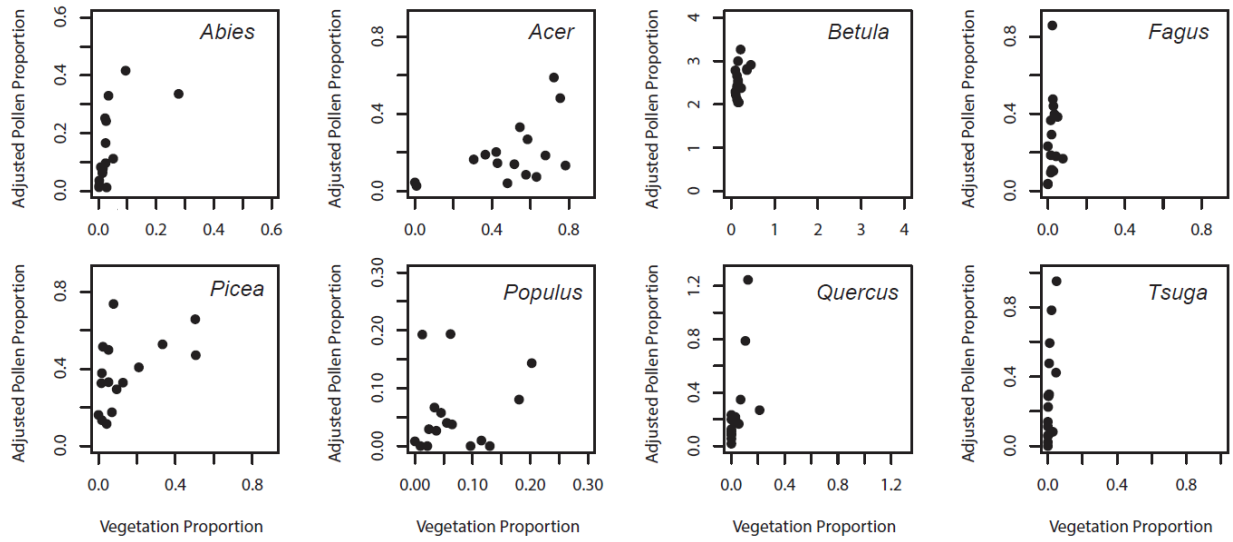
**Figure 6.4** Modern vegetation surfaces developed through Inverse Distance Weighting (IDW) interpolation of absolute values of taxon-specific canopy cover. Parameters included a power of one, a variable radius and a maximum number of three nearest neighbours. Blue indicates lower values of vegetation cover, red indicates higher vegetation cover. Values are in square meters.



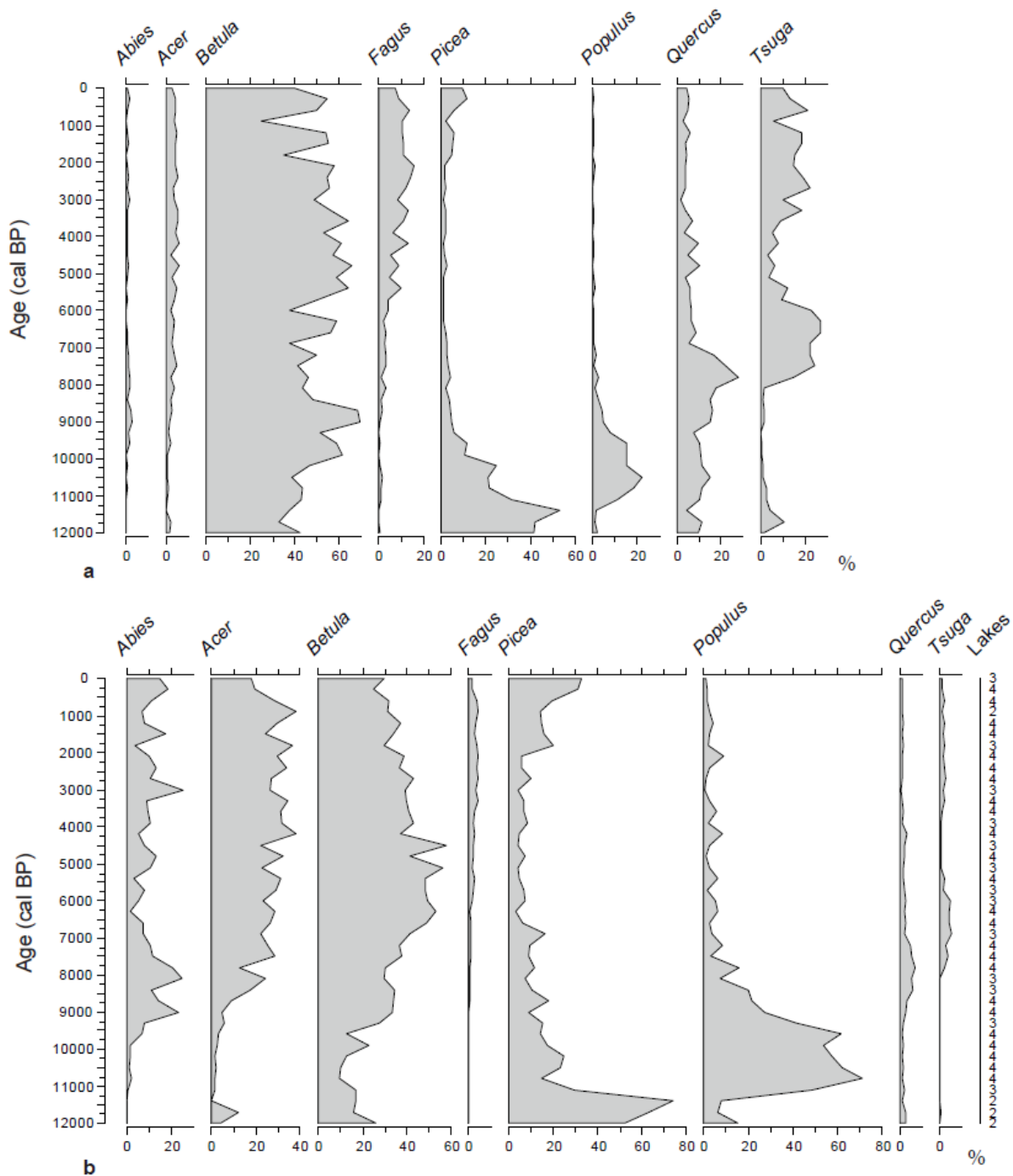
**Figure 6.5** Principal Component Analysis (PCA) of taxon-specific values of mean vegetation cover in consecutive 200 m wide rings within 100 and 3100 m away from the 16 lake sites used in ERV analysis. The 95% confidence limit is indicated by the ellipses.



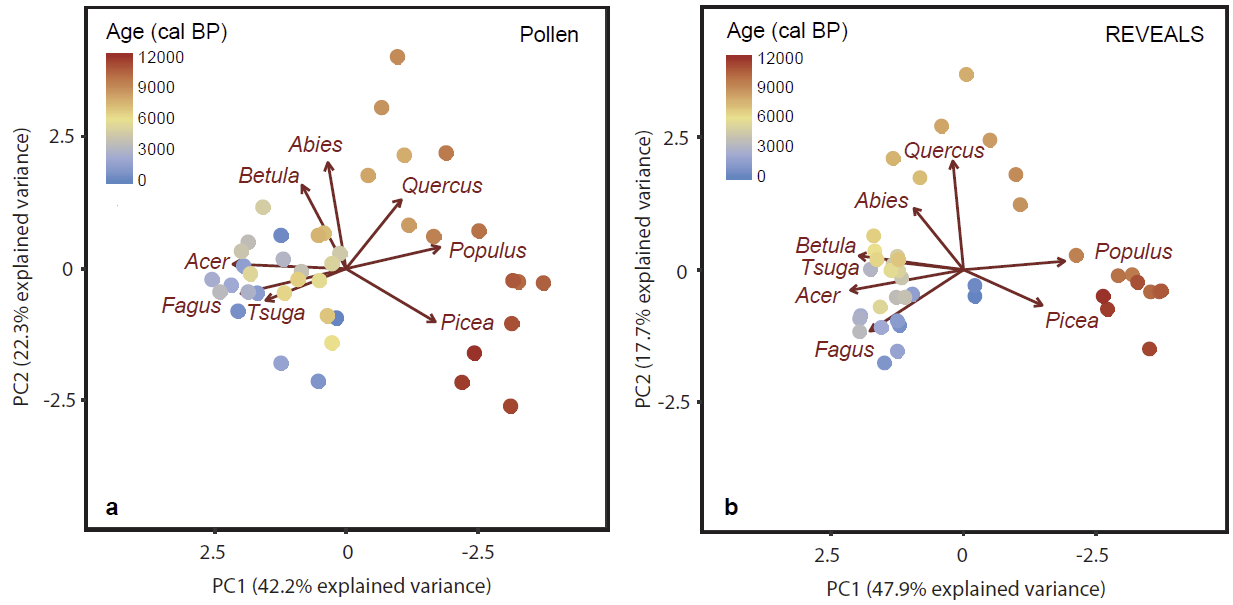
**Figure 6.6** Likelihood function scores for ERV submodels 2 and 3 based on a  $1/d$  distance weighting function. The RSAP is shown at 1600 m by the dashed line.



**Figure 6.7** Scatterplots of vegetation and adjusted pollen proportion at the RSAP (1600 m) for the eight taxa in the analysis based on ERV submodel 2.



**Figure 6.8** A pollen percentage diagram representative of the study region was created using the mean of the original pollen percentages from lakes Geai, Manitou, Sam and Tortue (a). The REVEALS estimates (b) were obtained by running the model using the original fossil pollen counts from lakes Geai, Manitou, Sam and Tortue, and calculating a mean value representing taxon-specific regional vegetation cover every 300 years. Mean vegetation cover is expressed as a percentage specific to the spatial scale of the study area (i.e.,  $6 \times 10^4 \text{ km}^2$ ) and is based on the sum of the cover of the eight taxa studied.



**Figure 6.9** Principal Component Analyses (PCAs) of the original pollen data shown in Figure 6.8a (a) and the REVEALS reconstructions shown in Figure 6.8b (b). Ages are in calibrated radiocarbon years before present.

**Table 6.1** Lake features for 20 small lakes in southeastern Quebec. Sites in bold are used for REVEALS analysis only.

Site Number	Site Name	Latitude (°N)	Longitude (°W)	Radius (m)	Altitude (m)	Number of <sup>14</sup> C measurements	Original References
1	Ange	47.4780	70.6834	134	640	-	Labelle & Richard, 1981
2	Atocas	45.5439	73.3121	44	120	-	Gauthier, 1981
3	Base	46.7908	71.3322	166	11	-	Richard, unpublished; Neotoma, 2017
4	Brule	45.7211	75.4450	190	254	-	Lafontaine-Boyer & Gajewski, 2014
5	Calixte	45.9582	73.8690	62	265	-	Richard, unpublished; Neotoma, 2017
6	Castor	46.6141	72.9812	124	206	-	Richard, unpublished; Neotoma, 2017
7	Gabriel	46.2770	73.4761	142	256	-	Richard, 1977a
8	Germain	45.9450	74.3704	91	469	-	Savoie & Richard, 1979
9	Joncas	47.2575	71.1759	136	786	-	Richard, 1977a
10	Marcotte	47.0547	71.3976	162	498	-	Labelle & Richard, 1981
11	Noir	45.7756	75.1351	214	177	-	Paquette & Gajewski, 2013
12	Pink	45.4682	75.8077	187	164	-	Mott, 1978
13	Quenouilles	46.1740	74.3967	69	406	-	Savoie & Richard, 1979
14	Ramsay	45.5979	76.0996	184	204	-	Mott, 1978
15	Romer	45.9637	73.3265	145	18	-	Comtois, 1982
16	Yamaska	45.4721	72.8659	17	336	-	Richard, unpublished; Neotoma, 2017
<b>17</b>	<b>Geai</b>	45.9962	73.9932	59	367	1	Richard, unpublished; Neotoma, 2017
<b>18</b>	<b>Manitou</b>	46.0601	74.4665	82	448	3	Savoie & Richard, 1979
<b>19</b>	<b>Sam</b>	46.6538	72.9774	89	220	6	Richard, unpublished; Neotoma, 2017
<b>20</b>	<b>Tortue</b>	45.5511	73.3152	80	151	4	Gauthier, 1981

**Table 6.2** Fall speeds of pollen and references for the eight tree genera included in this study.

Taxon	Fall Speed (m/s)	Reference
<i>Abies</i>	0.120	Mazier et al. 2012
<i>Acer</i>	0.056	Mazier et al. 2012
<i>Betula</i>	0.024	Mazier et al. 2012
<i>Fagus</i>	0.057	Mazier et al. 2012
<i>Picea</i>	0.056	Mazier et al. 2012
<i>Populus</i>	0.027	Commerford et al. 2013
<i>Quercus</i>	0.035	Mazier et al. 2012
<i>Tsuga</i>	0.056	Gaillard et al. 2008

**Table 6.3** Relative pollen productivity estimates (PPEs) and standard errors of PPEs for the eight taxa included in this study, calculated at the distance of the RSAP (1600 m) using an average wind speed of 3m/s, the 1/d pollen dispersal model, and a mean lake radius of 100 m.

Taxon	PPE		Standard Error	
	Submodel 2	Submodel 3	Submodel 2	Submodel 3
<i>Abies</i>	1.47	2.53	0.38	3.00
<i>Acer</i>	0.29	0.18	0.03	0.04
<i>Betula</i>	1.00	1.00	0.00	0.00
<i>Fagus</i>	7.19	4.22	1.52	1.84
<i>Picea</i>	0.60	0.46	0.16	0.14
<i>Populus</i>	0.11	0.09	0.15	0.08
<i>Quercus</i>	3.28	1.72	0.35	0.44
<i>Tsuga</i>	17.14	10.94	1.64	2.95

## 7 Summary and Conclusions

### 7.1 Summary

There is controversy about the extent to which prehistoric Native American land use practices impacted North American ecosystems, and consequently, the regional and potentially global climate (chapter 1). To reconcile the various viewpoints, and to begin to quantify the impacts of prehistoric populations on North American ecosystems throughout the Holocene, two major sources of information are necessary: reliable estimates of human population in space and time, and detailed records of vegetation change at various spatial scales. The objectives of this thesis were to synthesize quantitative paleodemographic and paleoenvironmental records for the past 13,000 years in North America and use these records to estimate continental- and regional-scale associations between paleopopulations and past ecosystems. The radiocarbon ( $^{14}\text{C}$ ) data were obtained from the Canadian Archaeological Radiocarbon Database (CARD), which contains more than 30,000 North American dates, most of which are associated with a culture and archaeological context. A second continental-scale database, the Neotoma Paleoecology Database, contains fossil pollen data from lake sediment cores which can be used as an indicator of vegetation change through the Holocene and late glacial. The availability of these two databases provided an opportunity to test the association between changes in population density and forest composition at a continental scale for the first time in North America.

In chapter two, the use of archaeological  $^{14}\text{C}$  dates derived from cultural materials as a proxy for paleopopulation density was assessed and tested against population estimates obtained from the HYDE3.1 population model (Klein Goldewijk et al., 2011). The  $^{14}\text{C}$  data generated a robust continental-scale record of paleodemographic change comparable to the modeled estimates; however, the  $^{14}\text{C}$  data indicated finer-scale temporal fluctuations and provided more detail when mapped at regular intervals than the modeled estimates. This provided the motivation for chapter three, in which the  $^{14}\text{C}$  data were used to obtain the first spatio-temporal estimates of North American population change at 500-year intervals from 13,000 BP to the present. The resulting maps clearly showed the rapid colonization of multiple regions spanning the continent following initial migration, and patterns in migration and settlement were analogous to those interpreted from ethnohistoric, genetic and archaeological records.

Chapters four and five describe and apply a methodology that was developed to evaluate cross-correlations between  $^{14}\text{C}$  data and pollen data. As expected, there was little relation between changes in boreal taxa (e.g., *Picea* or *Pinus*) and changes in population, and some suggestion of positive associations of the abundance of species from the deciduous forest (e.g., *Fagus* and *Acer*) with population. We saw intermittent relations between other taxa (e.g., *Quercus*, *Castanea*, *Carya* and *Juglans*) and population, and little relation between disturbance indicators (e.g., *Populus* and *Poaceae*) and population, which conflicted with the body of literature that emphasizes the role of human land use in influencing ecological succession, forest openness, forest composition and tree ranges.

These results implied known associations between specific taxa and population may not all scale up to the continental scale, and that this methodology could offer substantial insight if tested at a regional scale where both population and paleoenvironmental data are sufficient. However, well-known issues with using pollen percentages to deduce the abundance of plant species in the past, such as some species contributing proportionately more pollen than others, could not be addressed with this methodology. This provided the motivation for chapter six, in which pollen counts were used to estimate past vegetation abundance at a regional scale using the Landscape Reconstruction Algorithm (LRA). This method permitted the reconstruction of vegetation abundance at a regional scale while correcting for the non-linear relationship between pollen production and tree abundance. The resulting reconstruction included eight tree taxa and enabled a modification of the interpretation of the original pollen diagrams from the eastern deciduous forest of North America. Although this study did show the potential of this method to aid in the interpretation of pollen records, it was limited by the number of surface pollen samples available in regional databases, which meant a component of the landscape (i.e., herbs, grasses, and other trees) was left out. To improve upon this work, additional pollen records (from this region and eventually others) and additional taxa are needed. Once these data are obtained, comprehensive vegetation reconstructions can be compared to population histories to characterize past human-vegetation associations for different ecosystems at regional scales, and preliminary results indicate that this would be a promising research avenue.

In a first step towards continuing this research program, the pollen percentage diagram and vegetation reconstruction presented in chapter 6 were compared to regional population density curves based on  $^{14}\text{C}$  data. This was done within a 300 and 150 km radius of Lake Tortue, which is centrally located within the study area for chapter 6 (Figs. 7.1 and 7.2). When *Populus* was abundant in the early Holocene (i.e., 11,500 – 9,400 cal BP), no human presence is recorded in the archaeological data (Fig. 7.2). The reconstruction (Fig. 7.2, bottom) suggests *Populus* was more dominant on the landscape than previous interpretations assumed (Fig. 7.2, top), which could explain the hiatus in human presence at this time. *Populus* seeds are easily dispersed via wind and water, and germinate rapidly following disturbances (Gonzalez et al., 2010). An abundance of *Populus* is indicative of an environment in which the landscape was changing rapidly in response to climatic changes and the effects of deglaciation (Peros et al., 2010). Thus, extensive stands of drought-resistant aspen likely thrived while climate, fire and hydrological regimes were shifting (Richard, 1977a, 1977b, 1978; Peros et al., 2010), while other taxa were less able to respond to such changes. This type of environment would not be conducive to human activities and populations may have retreated south until conditions became more stable. Once *Populus* began to decrease in abundance, the vegetation became more diverse. the environment began to stabilize, and people would have been able to make use of the area once again, although at this time a nomadic lifestyle would have meant populations passed through the area instead of settling permanently.

Second, the reconstruction suggests *Betula* began to decrease in abundance after approximately 6,300 cal BP (Fig. 7.2, bottom), whereas the pollen percentages show no change in *Betula* for most of the mid- to late Holocene (Fig. 7.2, top). The decrease in *Betula* however, coincides with an increase in population density, which can be explained considering the environmental

context at that time. By 6,000 cal BP, the biomes were more-or-less in their modern positions since continental deglaciation was nearly complete (Dyke, 2005). White birch (*Betula papyrifera*), a disturbance taxon, declined significantly after 6,000 cal BP, causing the overall decline in *Betula*, although yellow birch (*Betula alleghaniensis*) persisted in smaller populations. The establishment of a closed temperate deciduous forest in southern Quebec, and environment characterized by warm summers, long growing seasons, mild winters and short-lived snow cover, were optimal living conditions for hunter-gatherers. Access to terrestrial and aquatic food sources, resources for shelter and hunting, and eventually agriculture and sedentism, led to a continued increase in population density in this region (Fig. 7.2).

Lastly, *Fagus grandifolia* increased in abundance after 6,300 cal BP, during the time that *Betula papyrifera* decreased. This change in *Fagus*, however, is only observed in the reconstruction and is greatly masked in the percentage diagram (Fig. 7.2, bottom). This is also indicative of the establishment of the temperate deciduous forest, since *Fagus* is a key component of the modern closed forest, and its increase coincides with an increase in population. Previous research indicates that beech nuts were an important component of hunter-gatherer diets (Bonnicksen, 2000) and it is possible that humans were selecting for its growth later in the record (Delcourt and Delcourt, 2004). More generally, *Fagus* is a component of the closed forest, along with taxa such as *Betula* (yellow birch) and *Acer* (maple), and this suggests relatively warm, moist, and stable conditions with few natural fires, and these conditions are capable of maintaining significant wildlife and other resources. Climatic and environmental conditions promoted the growth of both the forest and the local human population.

Numerous studies have described clear associations between human land use practices and changes in vegetation composition at regional scales. Munoz and Gajewski (2010) identified a period of increased weed, grass and *Populus* growth following prehistoric human land clearance in southern Ontario. Munoz and Gajewski also described how pollen sites located within areas of high archaeological site density were most influenced by human activity. Abrams and Nowacki (2008) studied numerous land survey records and noted associations between high mast tree abundance and proximity to human settlement sites. Several studies of Crawford Lake (ON) indicate an increase in Poaceae, *Portulaca* and maize pollen during Iroquoian occupation, followed by an increase in *Ambrosia*, *Plantago* and *Rumex* after European land clearance (e.g., McAndrews, 1988). Although the preliminary results presented above lack the detail presented in the studies by Munoz et al. (2010), Abrams and Nowacki (2008) and McAndrews (1988), they are in line with the overall conclusion that regional associations between human land use and vegetation are discernable at this scale. Further, the reconstruction indicates that the use of quantitative methods for vegetation reconstruction has the potential to uncover finer details regarding vegetation histories in comparison with pollen percentage diagrams. Although based on a subset of taxa, the reconstruction presented here (Fig. 7.2) allows us to fine-tune previous interpretations of Holocene vegetation change in southern Quebec. The pollen diagram from Lake Tortue shows how future research should focus on incorporating additional trees and non-arboreal species since they are important contributors to the overall pollen signal (Fig. 7.3).

In their study of human influence on Holocene vegetation in southern Ontario, Munoz and Gajewski (2010) emphasized that landscapes that were managed by prehistoric populations could be identified by an increase in *Populus* following a disturbance. The methodology discussed here can be applied to southern Ontario to better understand the results of this study, as the high *Populus* values seen in the reconstruction (Fig. 7.2) suggest such increases could have been even larger than previously thought. Therefore, the next steps towards a more profound understanding of the vegetation history of eastern North America will focus on the collection of additional pollen data and the calculation of pollen productivity estimates (PPEs) for all plant taxa found in this ecosystem, such that quantitative vegetation reconstructions may be done and statistically compared to  $^{14}\text{C}$  data. An alternative step would be to consider dates within a smaller radius of a single pollen site with known Holocene human occupation (e.g., Fig. 7.1, zone 3). The continued use of the databases and approaches discussed in this thesis will enable the development of similar estimates for other regions in North America and a better understanding of human-environment interactions over the course of the Holocene.

## 7.2 Conclusions

Understanding land use and land cover change in response to early human pressures is now identified as a top research priority and is essential for Earth system modellers, institutions focused on sustainability, and decision-makers (Moran & Lopez, 2016). The discussion continues concerning the extent of Native American impacts on the North American landscape (Delcourt and Delcourt, 2004), and if Native American land use could have contributed to changes in the global climate and carbon cycle (Ruddiman, 2014). Although many experts advocate an early Anthropocene, and believe humans began changing the North American landscape through early hunting and burning practices, some do not and favour the view that humans only began influencing ecosystems and climate in North America in the post-Industrial Era (Zalasiewicz et al., 2017).

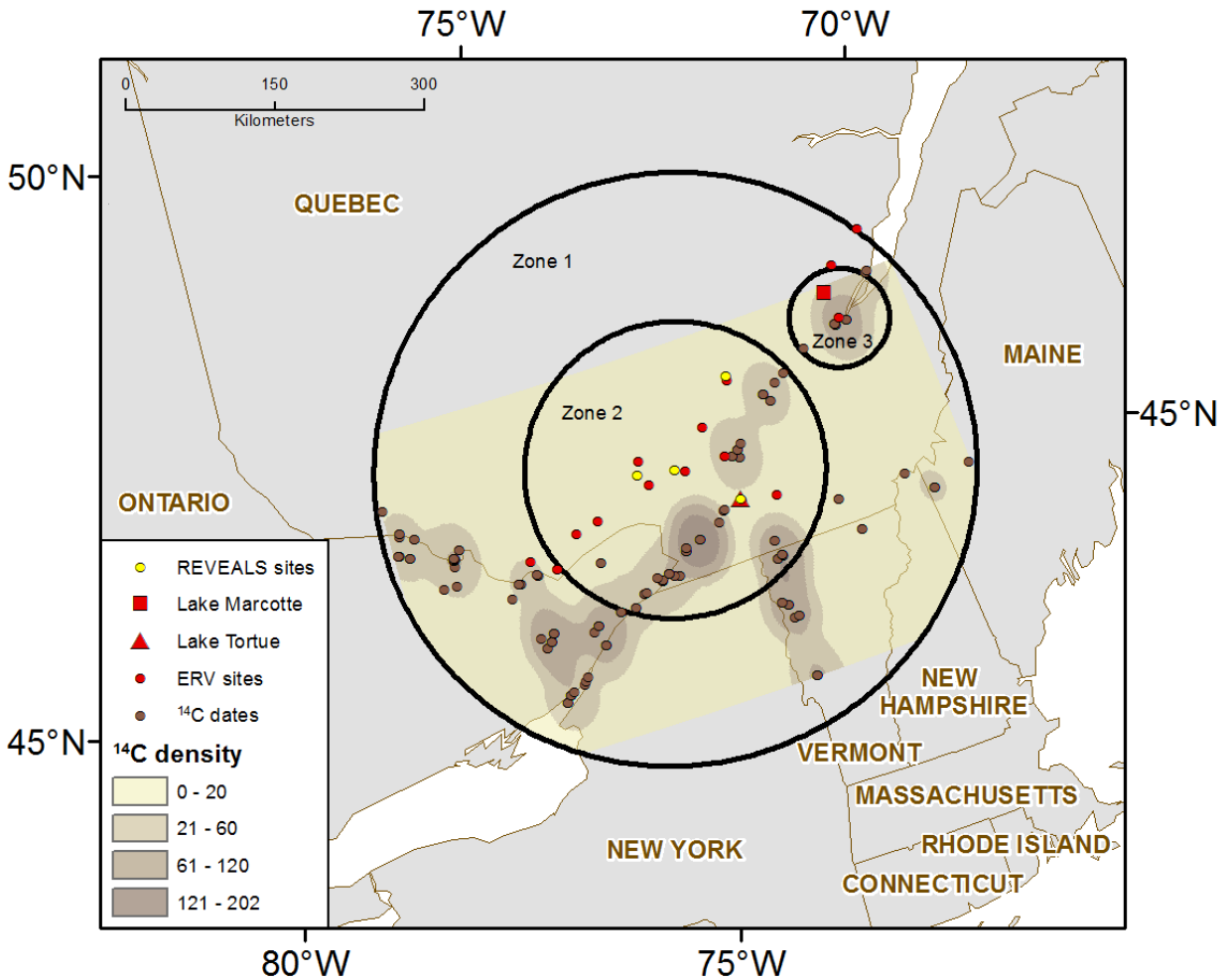
This thesis has demonstrated that, due to the successful compilation and digitization of global archaeological and paleoenvironmental information over the past few decades, North American radiocarbon and fossil pollen databases are now sufficiently complete to begin the process of quantifying the impacts of early humans on their environments. The Neotoma Paleoecology Database and the Canadian Archaeological Radiocarbon Database contain pollen and  $^{14}\text{C}$  records that can be used in the reconstruction of past vegetation abundance and population density for North America at the present time, and globally in the future. Based on the results of this thesis, we now have a continuous record of relative changes in human population density from the late Pleistocene to the present that can be compared to the vegetation history of North America.

As discussed in chapters 4 and 5 and section 7.1, comparisons of population and vegetation data using kernel estimation density approaches are highly promising, since they can allow us to arrive at reconstructed land use and land cover distributions through time in relation to demographic trends. Further, these approaches allow for the correction of issues such as the inhomogeneous spatial distribution of sites, taphonomic bias in archaeological records, and

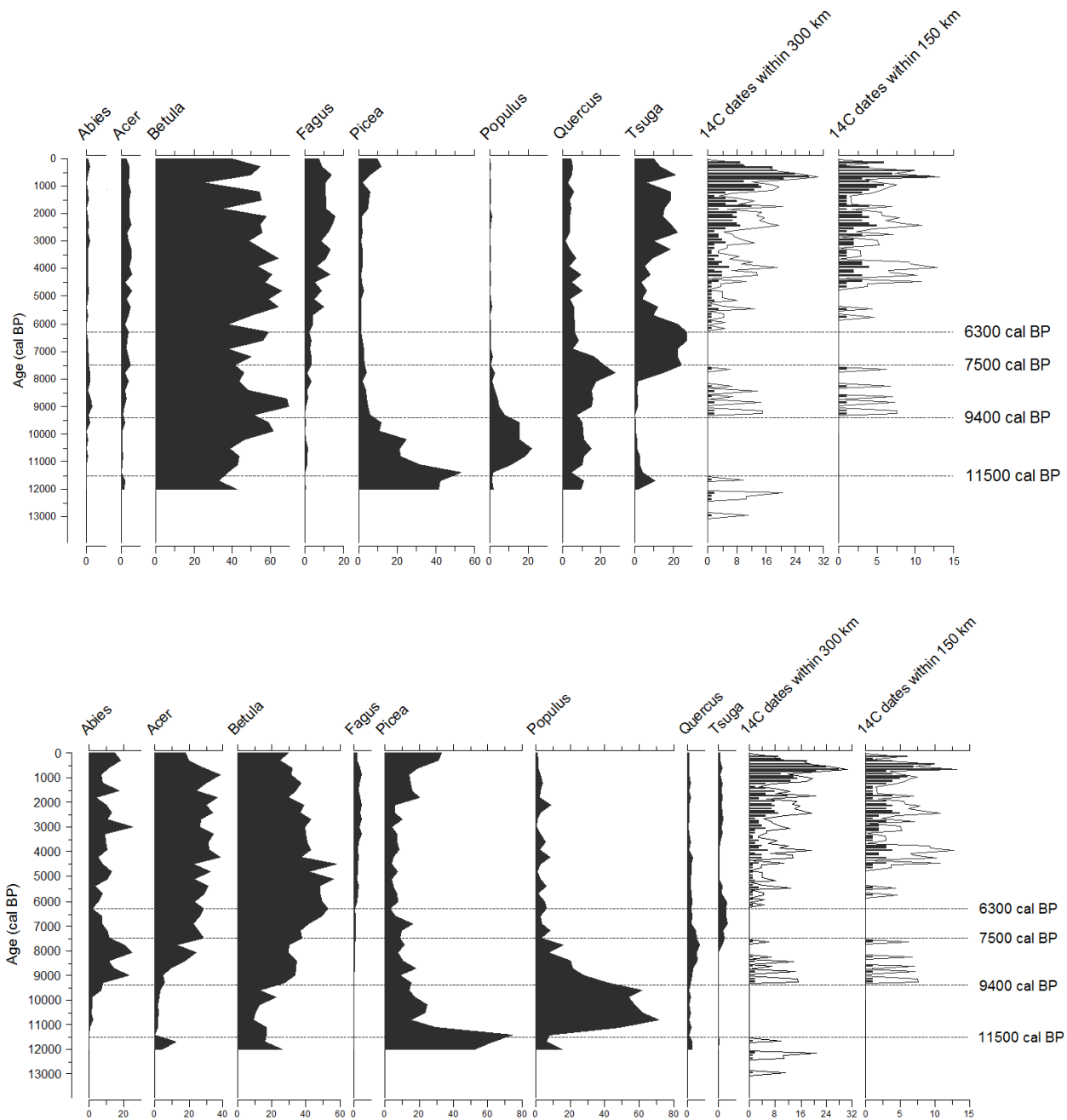
dating inconsistencies. It is possible that prehistoric Native American impacts were not sufficient to affect global concentrations of greenhouse gases, fueling the argument that relatively small populations have minor ecological implications, however small populations could conceivably have significant impacts on vegetation by changing, for example, natural succession patterns, forest openness, forest composition, and species ranges. The results from chapters 4 and 5 suggest a measurable, continental-scale impact is unlikely, but that significant human impacts could have occurred at regional scales, and that regional associations between people and certain economically-important taxa are discernable beyond the areas of intense human activity. Further regional studies are now needed to test this hypothesis and better identify the spatial extent of Native American impacts. The recent development and validation of quantitative models capable of providing new information on the spatial distribution of past vegetation will aid in this process. These models will allow for current vegetation histories from all major North American ecosystems to be revisited and improved upon. The case study from southeastern Quebec clearly showed additional information can be gained from reconstructing absolute vegetation cover in lieu of abundance and that this approach can deal with uncertainties associated with percentage data.

Although large radiocarbon datasets are commonly used as a paleodemographic proxy, these results are continually being refined. Future work needs to systematically address issues associated with the dates as data method, such as taphonomic issues, human factors, and comparability and classification issues (see chapter two). This thesis presents the first attempt at using the Canadian Archeological Radiocarbon Database in a continental-scale spatiotemporal analysis however further research is needed to test the effectiveness and feasibility of the methods used at alternate scales using smaller datasets.

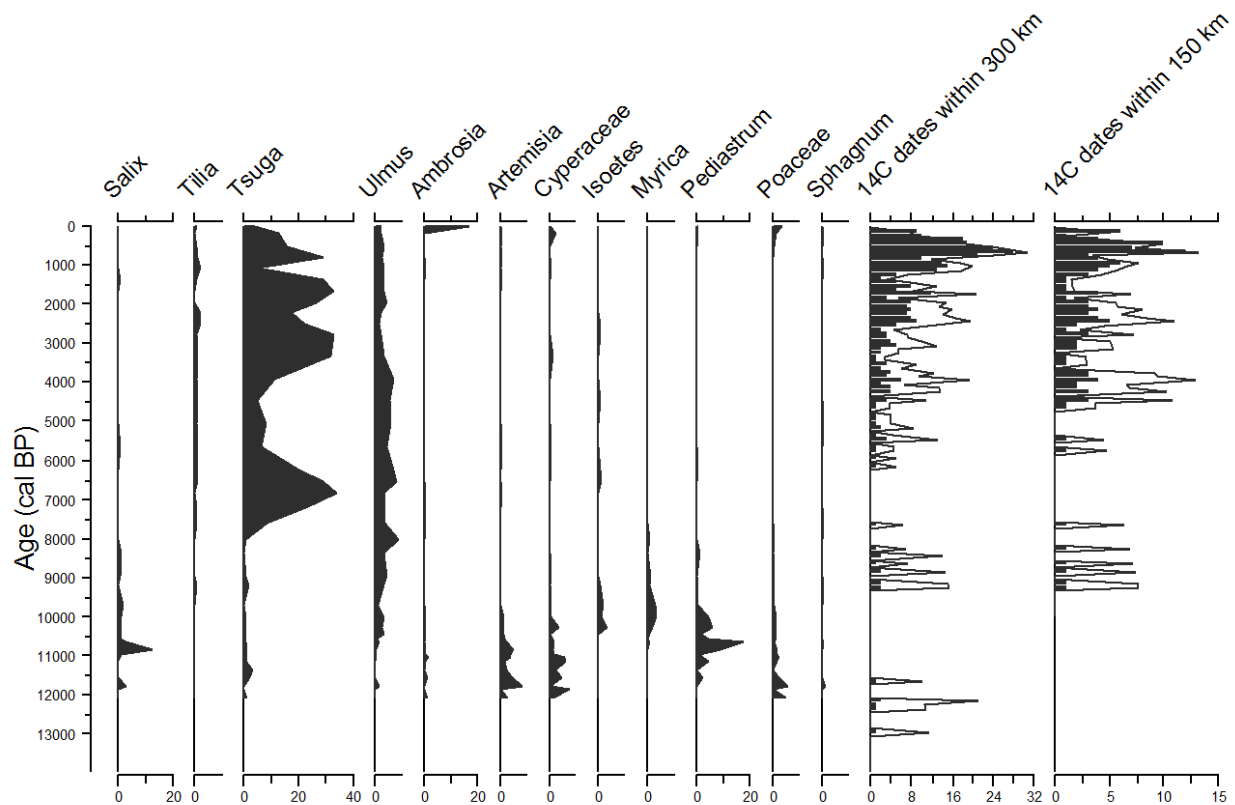
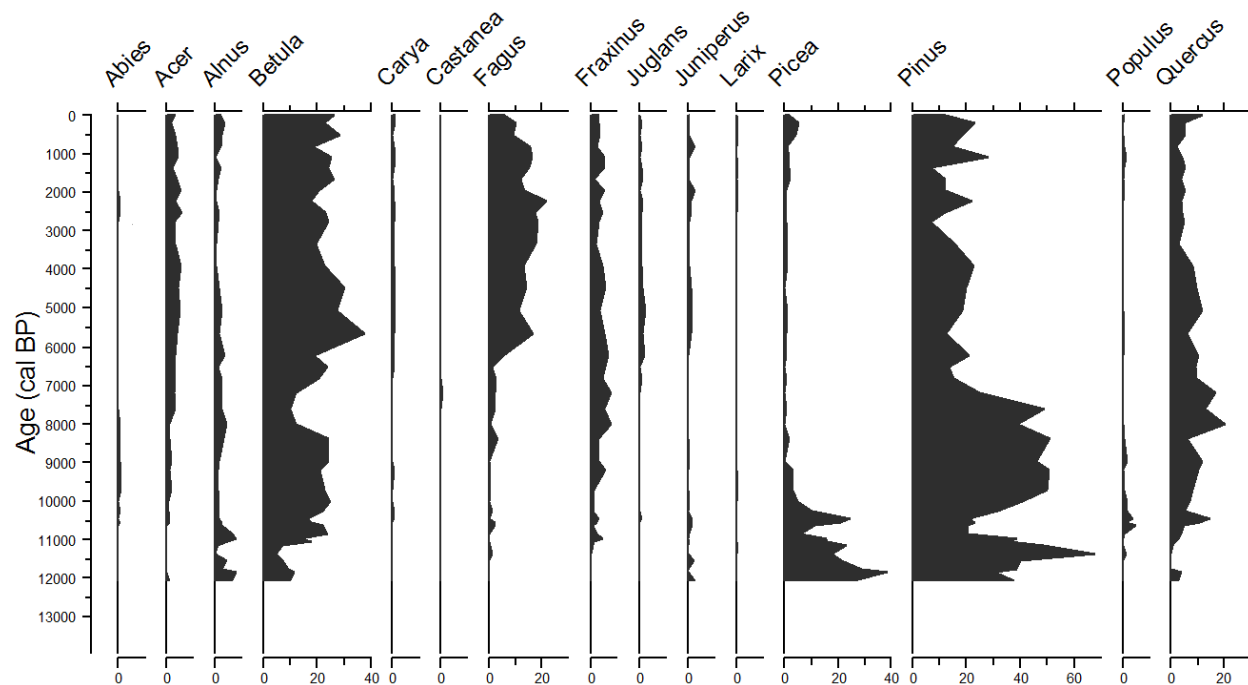
Despite these limitations, the results of thesis support the growing body of literature that suggests prehistoric Native Americans impacted their environments and that these impacts can be detected and quantified by integrating archaeological and paleoecological information. However, the timing, location, and intensity of human land use has changed in both space and time, suggesting regional- to local-scale analyses of human-environment interactions are most appropriate for continental North America. This thesis has demonstrated that we now have the data and methods necessary to estimate Holocene land use and land cover change in North America. Once a sufficient number of regional analyses are complete, the results can be synthesized into a continental history. It is imperative that we continue to integrate archaeological and paleoenvironmental data in the investigation of cultural and environmental change in all regions of the world.



**Figure 7.1** The study area (chapter six) showing sites where fossil (yellow circles) and modern (red circles, red square, red triangle) pollen assemblages were used, as well as the distribution of  $^{14}\text{C}$  dates available in the CARD (brown circles) superimposed on a smoothed density surface of these dates. Zone 1 indicated by the largest black circle represents a 300 km radius around the center of the study region encompassing Lake Tortue (red triangle). Zone 2 represents a 150 km radius around the center of the study region, and zone 3 represents a 50 km radius that encompasses Lake Marcotte (red square).



**Figure 7.2** The pollen percentage diagram from chapter 6 created using the mean of the original pollen percentages from lakes Geai, Manitou, Sam and Tortue (top), and the REVEALS estimates (bottom) from chapter 6, and calculating a mean value representing taxon-specific regional vegetation cover every 300 years. Mean vegetation cover is expressed as a percentage specific to the spatial scale of the study area (i.e.,  $6 \times 10^4 \text{ km}^2$ ) and is based on the sum of the cover of the eight taxa studied. The columns on the right show the raw histograms of  $^{14}\text{C}$  dates (black bars) and the taphonomically-corrected paleopopulation curves (lines) based on data from within a 300 and 150 km radius around the center of the study region.



**Figure 7.3** Pollen diagram from Lake Tortue. The columns on the bottom right show the histograms of  $^{14}\text{C}$  dates (black bars) and the taphonomically-corrected paleopopulation curves (black lines) based on data from within a 300 and 150 km radius around the center of the study region.

## References

- Abraham, V., and Kozáková, R. 2012. Relative pollen productivity estimates in the modern agricultural landscape of Central Bohemia (Czech Republic). *Review of Palaeobotany and Palynology* 179: 1-12.
- Abraham, V., Ouskova, V., and Kunes, P. 2014. Present-Day Vegetation Helps Quantifying Past Land Cover in Selected Regions of the Czech Republic. *PLoS ONE* 9(6): e100117.
- Abraham, V., Novák, J., Houfková, P., Petr, L., and Dudová, L. 2017. A Landscape Reconstruction Algorithm and pedoanthracological data reveal Late Holocene woodland history in the lowlands of the NE Czech Republic. *Review of Palaeobotany and Palynology* 244: 54-64.
- Abrams, M.D. 1992. Fire and the development of oak forests. *BioScience* 42: 346-353.
- Abrams, M.D., and Nowacki, G.J. 2008. Native americans as active and passive promoters of mast and fruit trees in the eastern USA. *Holocene* 18: 1123–1137.
- Acheson, S.R., 1998. In the wake of the ya'áats'xaatgáay ['Iron People']: A study of changing settlement strategies among the Kung'it Haida. *British Archaeological Reports, International Series, No. 711, Oxford.*
- Achilli, A., Perego, U.A., Lancioni, H., Olivieri, A., Gandini, F., Hooshar Kashani, B., Battaglia, V., Grugni, V., Angerhofer, N., Rogers, M.P., Herrera, R.J., Woodward, S.R., Labuda, D., Smith, D.G., Cybulski, J.S., Semino, O., Malhi, R.S., and Torroni, A. 2013. Reconciling migration models to the Americas with variation of North American native mitogenomes. *Proceedings of the National Academy of Sciences USA* 110(35): 14308-14313.
- Adler, D., Murdoch, D., Nenadic, O., Urbanek, S., Chen, M., Gebhardt, A., Bolker, B., Csardi, G., Strzelecki, A., Senger, A., Eddelbuettel, D., and the R Core Team. 2016. Rgl: sD visualization using OpenGL. R package version 0.96.0. Available at <https://CRAN.R-project.org/package=rgl>.
- Åkesson, C., Nielsen, A.-B., Broström, A., Persson, T., Gaillard, M.-J., and Berglund, B.E. 2015. From landscape description to quantification: A new generation of reconstructions provides new perspectives on Holocene regional landscapes of SE Sweden. *Holocene* 25(1): 178-193.
- Andersen, S.T. 1970. The relative pollen productivity and pollen representation of North European trees, and correction factors for tree pollen spectra. *Danmarks Geologiske Undersøgelse. II. RÆKKE. No. 96, 99.*
- Anderson, T. 1985. Late-Quaternary pollen records from eastern Ontario, Quebec and Atlantic Canada. Pp 281-326 in: V. Bryant Jr and R. Holloway (Eds). *Pollen records of Late-Quaternary North American sediments. American Association of Stratigraphic Palynologists. Dallas, Texas.*
- Anderson, N.J., Bugmann, H., Dearing, J.A., and Gaillard, M.-J. 2007. Linking palaeoenvironmental data and models to understand the past and to predict the future. *Trends in Ecology and Evolution* 21: 696-704.
- Anderson, R., Starratt, S., Brunner Jass, R., and Pinter, N. 2009. Fire and vegetation history on Santa Rosa Island, Channel Islands, and long-term environmental change in southern California. *Journal of Quaternary Science* 25(5): 782-797.

- Anderson, D.G., Goodyear, A.C., Kennett, J., and West, A. 2011. Multiple lines of evidence for possible human population decline/settlement reorganization during the Early Younger Dryas. *Quaternary International* 242: 570-583.
- Araujo, A.G.M., Neves, W., Piló, L.B., and Atui, J.P.V. 2005. Holocene dryness and human occupation in Brazil during the “Archaic Gap”. *Quaternary Research* 64: 298-307.
- Aschmann, H. 1959. The central desert of Baja California: Demography and Ecology. *Ibero-Americana* 42.
- Ashburn, P.M. 1949. *The ranks of death: a medical history of the conquest of America*. Coward-McCann, New York.
- Bahre, C.J. 1995. Human impacts on the grasslands of southeastern Arizona. In M.P. McLaran and Van Devender, T.R. (Eds). *The Desert Grassland*. University of Arizona Press, United States of America.
- Baker, A.G., Zimny, M., Keczyński, A., Bhagwat, S.H., Willis, K.J., and Latałowa, M. 2016. Pollen productivity estimates from old-growth forest strongly differ from those obtained in cultural landscapes: Evidence from the Białowieża National Park, Poland. *Holocene* 26(1): 80-92.
- Ballenger, J.A.M., and Mabry, J.B. 2011. Temporal frequency distributions of alluvium in the American Southwest: taphonomic, paleohydraulic, and demographic implications. *Journal of Archaeological Science* 38(6): 1314-1325.
- Bamforth, D.B., and Grund, B. 2012. Radiocarbon calibration curves, summed probability distributions, and early Paleoindian population trends in North America. *Journal of Archaeological Science* 39: 1768-1774.
- Barbarena, R., Prates, L., and De Porrás, M.E. 2015. The human occupation of northwestern Patagonia (Argentina): paleoecological and chronological trends. *Quaternary International* 356: 111-126.
- Barnosky, A.D., and Lindsey, E.L. 2010. Timing of Quaternary megafaunal extinction in South America in relation to human arrival and climate change. *Quaternary International* 217: 10-29.
- Bartlein, P.J., Prentice, I.C., and Webb III, T. 1986. Climatic response surfaces from pollen data for some eastern North American taxa. *Journal of Biogeography* 13:35–57.
- Bayliss A., Ramsey, C.B., van der Plicht, J., and Whittle, A. 2007. Bradshaw and Bayes: Towards a timetable for the Neolithic. *Cambridge Archaeological Journal* 17(1): 1-28.
- Bell, T., and Renouf, M.A.P. 2003. Prehistoric cultures, reconstructed coasts: Maritime Archaic Indian site distribution in Newfoundland. *World Archaeology* 35(3): 350-370
- Bellwood, P. 2005. *First Farmers: The Origins of Agricultural Societies*. Blackwell, Oxford.
- Bennike, O., and Andreasen, C. 2008. Radiocarbon dating of musk-ox (*Ovibos moschatus*) remains from northeast Greenland. *Polar Rec* 41(4): 305-310
- Bennike, O. 2014. Radiocarbon dating of musk-ox (*Ovibos moschatus*) bones from the Thule region, northwest Greenland. *The Polar Record* 50(2): 113-118
- Benz, M., n.d. PPND – the Platform for Neolithic Radiocarbon Dates. Ex Oriente, available at [http://www.exorient.org/associated\\_projects/ppnd.php](http://www.exorient.org/associated_projects/ppnd.php).
- Bernabo, J.C., and Webb III, T. 1977. Changing patterns in the Holocene pollen record from northeastern North America: a mapped summary. *Quaternary Research* 8:64-96.
- Bevan, A., Crema, E., Li, X., and Palmisano, A. 2013 Intensities, interactions and uncertainties: some new approaches to archaeological distributions. In: Bevan, A. and Lake, M. (eds.),

- Computational Approaches to Archaeological Space, 27–52. Walnut Creek: Left Coast Press.
- Bird, B.W., Wilson, J.J., Gilhooly III, W.P., Steinman, B.A., and Stamps, L. 2017. Midcontinental Native American population dynamics and late Holocene hydroclimate extremes. *Scientific Reports* 7: 41628.
- Birks, H.J.B., and Birks, H. 1980. *Quaternary Palaeoecology*. Caldwell: Blackburn Press.
- Birks, H.J.B., and Berglund, B.E. 2018. One hundred years of Quaternary pollen analysis 1916–2016. *Vegetation History and Archaeobotany* 27(2): 271–309.
- Black, B.A., Ruffner, C., and Abrams, M. 2006. Native American influences on the forest composition of the Allegheny Plateau, northwest Pennsylvania. *Canadian Journal of Forestry Research* 36:1266–1275.
- Blaauw, M., and Christen, J.A. 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Analysis* 6: 457–474.
- Bodner, M., Perego, U.A., Huber, G., Fendt, L., Rock, A.W., Zimmermann, B., Olivieri, A., Gomez-Carballa, A., Lancioni, H., Angerhofer, N., Bobillo, M.C., Corach, D., Woodward, S.R., Salas, A., Achilli, A., Torroni, A., Bandelt, H.-J., and Parson, W. 2012. Rapid coastal spread of First Americans: novel insights from South America's Southern Cone mitochondrial genomes. *Genome Research* 22: 811–820.
- Böhner, A., and Schyle, D., 2002–2006. Radiocarbon CONTEXT database 2002–2006. Available at <http://context-database.uni-koeln.de/index.php>.
- Bonnicksen, T.M. 2000. *America's ancient forests: From the Ice Age to the Age of Discovery*. New York: John Wiley and Sons.
- Bonomo, M., Angrizani, R.C., Apolinaire, E., and Noelli, F.S., 2015. A model for the Guaraní in the La Plata Basin and littoral zone of southern Brazil. *Quaternary International* 356: 54–73.
- Boserup, E. 1965. *The conditions of agricultural growth*, Aldine. NY.
- Boucher, Y., Arseneault, D., and Sirois, L. 2009. Logging history (1820 – 2000) of a heavily exploited southern boreal forest landscape: Insights from sunken logs and forestry maps. *Forest Ecology and Management* 258: 1359–1368.
- Boulanger, M.T., and Lyman, R.L., 2014. Northeastern North American Pleistocene megafauna chronologically overlapped minimally with Paleoindians. *Quaternary Science Reviews* 85: 35–46.
- Bowling, G.A. 1942. The introduction of cattle into Colonial North America. *Journal of Dairy Science* 25(2): 129–154.
- Bowman, S. 1990. *Radiocarbon Dating*. University of California Press, Berkeley and Los Angeles.
- Boyd, R. 1999. *Indians, Fire and the land in the Pacific Northwest*. Oregon State University Press, Corvallis, OR, USA.
- Bradshaw, R.H.W., and Webb III, T. 1985. Relationships between contemporary pollen and vegetation data from Wisconsin and Michigan, USA. *Ecology* 66(3): 721–737.
- Bradtmöller, M., Pastoors, A., Weninger, B., and Weninger, G.-C. 2012. The repeated replacement model – rapid climate change and population dynamics in Late Pleistocene Europe. *Quaternary International* 247: 38–49.
- Braun, E.L. 1950. *Deciduous forests of Eastern North America*. The Blakiston Company, Blakiston, Philadelphia, USA.
- Bronk Ramsey, C. 2008. Deposition models for chronological records. *Quaternary Science*

- Reviews* 27(1-2): 42–60.
- Bronk Ramsey, C., Scott, E.M., and van der Plicht, J. 2013. Calibration for archaeological and environmental terrestrial samples in the time range 26-50 ka cal BP. *Radiocarbon* 55(4): 2021-2027.
- Broström, A., Sugita, S., and Gaillard, M.-J. 2004. Pollen productivity estimates for the reconstruction of past vegetation cover in the cultural landscape of southern Sweden. *Holocene* 14(3): 368-381.
- Broström, A., Sugita, S., Gaillard, M.-J., and Pilesjö, P. 2005. Estimating spatial scale of pollen dispersal in the cultural landscape of southern Sweden. *Holocene* 15: 252–262.
- Broström, A., Nielsen, A.B., Gaillard, M.-J., Hjelle, K., Mazier, F., Binney, H., Bunting, J., Fyfe, R., Meltsov, V., Poska, A., Räsänen, S., Soepboer, W., von Stedingk, H., Suutari, H., and Sugita, S. 2008. Pollen productivity estimates of Key European plant taxa for quantitative reconstruction of past vegetation: a review. *Vegetation History and Archaeobotany* 17: 461-478.
- Brown, W.A., 2015. Through a filter, darkly: population size estimation, systematic error, and random error in radiocarbon-supported demographic temporal frequency analysis. *Journal of Archaeological Science* 53: 133-147.
- Bryant, V.M., Jr., and Holloway, R.G. 1985. A Late-Quaternary paleoenvironmental record of Texas: an overview of the pollen evidence. Pp 39-70 in: V.M. Bryant, Jr. and R.G. Holloway, (Eds). *Pollen records of Late Quaternary North American sediments*. American Association of Stratigraphic Palynologists Foundation, Dallas, Texas, USA.
- Buchanan, B., Collard, M., and Edinborough, K. 2008. Paleoindian demography and the extraterrestrial impact hypothesis. *Proceedings of the National Academy of Sciences USA* 105: 11651-11654.
- Buchanan, B., Hamilton, M., Edinborough, K., O'Brien, M.J., and Collard, M. 2011. A comment on Steele's (2010) "radiocarbon dates as data: quantitative strategies for estimating colonization front speeds and event densities". *Journal of Archaeological Science* 38: 2116-2122.
- Bueno, L., Schmidt Dias, A., and Steele, J. 2013. The Late Pleistocene/Early Holocene archaeological record in Brazil: a geo-referenced database. *Quaternary International* 301: 74-93.
- Bunting, M.J., Armitage, R., Binney, H.A., and Waller, M. 2005. Estimates of 'relative pollen productivity' and 'relevant source area of pollen' for major tree taxa in two Norfolk (UK) woodlands. *Holocene* 15(3): 459-465.
- Burden, E.T., McAndrews, J.H., and Norris, G. 1986. Palynology of Indian and European forest clearance and farming in lake sediment cores from Awenda Provincial Park, Ontario. *Canadian Journal of Earth Sciences* 23:43-54.
- Burns, R.M., and Honkala, B.H. 1990a. *Silvics of North America: Volume 1. Conifers*. United States Department of Agriculture (USDA), Forest Service, Agriculture Handbook 654.
- Burns, R.M., and Honkala, B.H. 1990b. *Silvics of North America: Volume 2. Hardwoods*. United States Department of Agriculture (USDA), Forest Service, Agriculture Handbook 654.
- Butzer, K.W. 1992. The Americas before and after 1492: an introduction to current geographical research. *Annals of the Association of American Geographers* 82: 345-368.
- Butzer, K. W. 1996. Ecology in the long view: settlement histories, agrosystemic strategies, and ecological performance. *Journal of Field Archaeology* 23: 141-150.

- Calcote, R. 1995. Pollen source area and pollen productivity: evidence from forest hollows. *Journal of Ecology* 83(4): 591-602.
- Campbell, L. 1997. American Indian languages: the historical linguistics of Native America. New York: Oxford University Press.
- Campbell, I.D. 1999. Quaternary pollen taphonomy: examples of differential redeposition and differential preservation. *Palaeogeography, Palaeoclimatology, Palaeoecology* 149: 246-256.
- Campbell, R., and Quiroz, D. 2015. Chronological database for southern Chile (35°30'-42°S), ~33000 BP to present: human implications and archaeological biases. *Quaternary International* 356: 39-53.
- Chapman, J., Delcourt, P.A., Cridlebaugh, P.A., Shea, A.B., and Delcourt, H.R. 1982. Man-land interaction: 10,000 years of American Indian impact on native ecosystems in the lower Little Tennessee River valley, eastern Tennessee. *Southeastern Archaeology* 1(2): 115-121.
- Chaput, M.A., Kriesche, B., Betts, M., Martindale, A., Kulik, R., Schmidt, V., and Gajewski, K. 2015. Spatiotemporal Distribution of Holocene Populations in North America. *Proceedings of the National Academy of Sciences USA* 112: 12127-12132.
- Chernick, M.R., and LaBudde, R.A. 2011. *An Introduction to Bootstrap Methods with Applications to R*. J. Wiley & Sons, Hoboken.
- Chiu, S.N., Stoyan, D., Kendall, W.S., and Mecke, J. 2013. *Stochastic Geometry and its Applications* (3rd ed.). J. Wiley & Sons, Chichester.
- Clark, J.S., and Royall, P.D. 1995. Transformation of a northern hardwood forest by aboriginal (Iroquois) fire: charcoal evidence from Crawford Lake, Ontario, Canada. *Holocene* 5:1-9.
- Clark, J.S., and Royall, P.D. 1996. Local and regional sediment charcoal evidence for fire regimes in presettlement north-eastern North America. *Journal of Ecology* 84:365-382.
- Cochran, W.G. 1963. *Sampling Techniques*, 2nd Ed. J. Wiley and Sons, New York.
- COHMAP members, 1988. Climatic changes of the last 18,000 years: observations and model simulations. *Science* 241: 1043-1052.
- Collard, M., Edinborough, K., Shennan, S., and Thomas, M.G. 2010. Radiocarbon evidence indicates that migrants introduced farming to Britain. *Journal of Archaeological Science* 37: 866-870.
- Commerford, J.L., McLauchlan, K.K., and Sugita, S. 2013. Calibrating vegetation cover and grassland pollen assemblages in the Flint Hills of Kansas, USA. *American Journal of Plant Science* 4: 1-10.
- Comtois, P.G. 1982. Histoire holocène du climat et de la végétation à Lanoraie (Quebec). *Canadian Journal of Earth Science* 19:1938-1952.
- Cooke, R., Ranere, A., Pearson, G., Dickau, R. 2013. Radiocarbon chronology of early human settlement on the Isthmus of Panama (13,000-7000 BP) with comments on cultural affinities, environments, subsistence, and technological change. *Quaternary International* 301: 3-22.
- Cooper, A., Turney, C., Hughen, K.A., Brook, B.W., McDonald, H.G., and Bradshaw, C.J.A. 2015. Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover. *Science* 349(6248): 602-606.
- Courtwright, J. 2011. *Prairie fire: A Great Plains history*. University of Kansas Press, Lawrence, Kansas, USA.

- Cowden, M., Hart, J., and Buchanan, M. 2014. Canopy accession strategies and climate responses for three *Carya* species common in the Eastern Deciduous Forest. *Trees* 29:223-235.
- Cressie, N.A.C., and Wikle, C.K. 2011. *Statistics for Spatio-Temporal Data*. J. Wiley & Sons, Hoboken.
- Crombé, P., and Robinson, E. 2014. <sup>14</sup>C dates as demographic proxies in Neolithisation models of northwestern Europe: a critical assessment using Belgium and northeast France as a case-study. *Journal of Archaeological Science* 52: 558-566.
- Cronin, T.M., Manley, P.L., Brachfeld, S., Manley, T.O., Willard, D.A., Guilbault, J.-P., Rayburn, J.A., Thunell, R., and Berke, M. 2008. Impacts of post-glacial lake drainage events and revised chronology of the Champlain Sea episode 13–9 ka. *Palaeogeography, Palaeoclimatology, Palaeoecology* 262: 46-60.
- Cronon, W. 1983. *Changes in the Land: Indians, colonists and the ecology of New England*. Hill & Wang, New York, USA.
- Crutzen, P.J. 2002. Geology of mankind. *Nature* 415(6867): 23.
- Crutzen, P.J., and Stoermer, E.F. 2000. The Anthropocene. *Global Change Newsletter*, 41: 17-18. Available at [www.igbp.net/news/opinion/opinion/haveweenteredtheanthropocene](http://www.igbp.net/news/opinion/opinion/haveweenteredtheanthropocene).
- Cui, Q.-Y., Gaillard, M.-J., Lemdahl, G., Stenberg, L., Sugita, S., and Zernova, G. 2014. Historical land-use and landscape change in southern Sweden and implications for present and future biodiversity. *Ecology and Evolution* 4: 3555-3570.
- Danneyrolles, V., Arseneault, D., and Bergeron, Y. 2016. Pre-industrial landscape composition patterns and post-industrial changes at the temperate–boreal forest interface in western Quebec, Canada. *Journal of Vegetation Science* 27: 470-481.
- Davis, M.B. 1963. On the theory of pollen analysis. *American Journal of Science* 261: 897–912.
- Davis, R.B., and Webb III, T. 1975. The contemporary distribution of pollen in eastern North America: a comparison with vegetation. *Quaternary Research* 5: 395-434.
- Davis, M.B. 1976. Pleistocene biogeography of temperate deciduous forests. *Geoscience and Man* 13:13-26.
- Davis, M.B. 1981a. Quaternary history and the stability of forest communities. Pages 132-153 in D.C. West, H.H. Shugart, and D.B. Botkin, editors. *Forest Succession*, Springer Verlag, New York, USA.
- Davis, M.B. 1981b. Outbreaks of forest pathogens in Quaternary History. Pp 216-227 in: D. Bharadwaj, D. Vishnu-Mittre and H. Maheshwari (Eds). *Proceedings of the Fourth International Palynological Conference, Volume 3*. Birbal Sahni Institute of Paleobotany, Lucknow.
- Davis, M.B., Sugita, S., Calcote, R.R., Frelich, L.E. 1992. Effects of invasion by *Tsuga canadensis* on a North American forest ecosystem. Pp 34-44 in: A.J. Teller, P. Mathy and J.N.R. Jeffers (Eds). *Responses of forest ecosystems to environmental changes: 1<sup>st</sup> European Symposium on Terrestrial Ecosystems: Forests and Woodland*. Springer.
- Day, G. 1953. The Indian as an ecological factor in the northeastern forest. *Ecology* 34:329-346.
- Delcourt, H. 1987. The impact of prehistoric agriculture and land occupation on natural vegetation. *Trends in Ecology and Evolution* 2:39-44.
- Delcourt, P.A., and Delcourt, H.R. 1979. Late Pleistocene and Holocene distributional history of the deciduous forest in the Southeastern United States. *Veröffentlichungen des Geobotanischen Institutes der ETH, Stiftung Rübel, Zürich*.

- Delcourt, P.A., Delcourt, H.R., Brister, R.C., and Lackey, L.E. 1980. Quaternary vegetation history of the Mississippi Embayment. *Quaternary Research* 13(1): 111-132.
- Delcourt, P.A., and Delcourt, H.R. 1987. Long-term forest dynamics of the temperate zone. Springer Verlag, New York, USA.
- Delcourt, P.A., Delcourt, H.R., Criddlebaugh, P., and Chapman, J. 1986. Holocene ethnobotanical and paleoecological record of human impact on vegetation in the Little Tennessee River Valley, Tennessee. *Quaternary Research* 25: 330-349.
- Delcourt, P.A., and Delcourt, H.R. 1998. Paleoecological insights on conservation of biodiversity: a focus on species, ecosystems and landscapes. *Ecological Applications* 8(4): 921-934.
- Delcourt, H.R., and Delcourt, P.A. 1997. Pre-columbian Native American use of fire on southern Appalachian landscapes. *Conservation Biology* 11: 1010-1014.
- Delcourt, P.A., and Delcourt, H.A. 2004. *Prehistoric Native Americans and Ecological Change*. Cambridge University Press, Cambridge.
- Denevan, W.M. 1976. Introduction. Pp 1912 in: W.M. Denevan (Ed). *The Native Population of the Americas in 1492*. University of Wisconsin Press, Madison.
- Denevan, W. 1992. The pristine myth: the landscape of the Americas in 1492. *Annals of the Association of American Geographers* 82: 369-385.
- Denevan, W. 2011. The "Pristine Myth" revisited. *The Geographical Review* 101(4): 576-591.
- Diggle, P.J. and Ribeiro, Jr., P.J. 2007. *Model-Based Geostatistics*. Springer, New York.
- Diggle, P.J. 2014. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns* (3rd ed.). Boca Raton: Chapman & Hall/CRC.
- Dixon, E.J. 2001. Human colonization of the Americas: timing, technology and process. *Quaternary Science Reviews* 20: 277-299.
- Dobyns, H.F. 1966. Estimating aboriginal American population: an appraisal of techniques with a new hemispheric estimate. *Current Anthropology* 7: 395-416.
- Dobyns, H.F., and Swagerty, W.R. 1983. *Their Number Become Thinned: Native American Population Dynamics in Eastern North America*. Knoxville, TN: University of Tennessee. Page 42.
- Doolittle, W.E. 1992. Agriculture in North America on the eve of contact: a reassessment. *Annals of the American Association of Geographers* 82: 386-401.
- Doolittle, W.E. 2000. *Cultivated landscapes of Native North America*. Oxford University Press, Oxford, England.
- Dorney, J.R. 1981. The impact of Native Americans on presettlement vegetation in southeastern Wisconsin. *Wisconsin Academy of Sciences, Arts and Letters* 69:26-36.
- Driver, H.E. 1961. *Indians of North America*. University of Chicago Press, Chicago.
- Dull, R.A., Nevle, R.J., Woods, W.I., Bird, D.K., Avnery, S., and Devenan, W.M. 2010. The Columbian Encounter and the Little Ice Age: Abrupt land use change, fire, and greenhouse forcing. *Annals of the American Association of Geographers* 100: 755-771.
- Dumond, D.E. 2000. The Norton Tradition. *Arctic Anthropology* 37(2): 1-22.
- Dupuis, S., Arseneault, D., and Sirois, L. 2011. Change from pre-settlement to present-day forest composition reconstructed from early land survey records in eastern Quebec, Canada. *Journal of Vegetation Science* 22: 564-575.
- Dyer, J.M. 2006. Revisiting the Deciduous Forests of Eastern North America. *Bioscience* 56: 341-352.

- Dyke, A.S., Andrews, J.T., Clark, P.U., England, J.H., Miller, G.H., Shaw, J., and Veillette, J.J. 2002. The Laurentide and Innuitian ice sheets during the Last Glacial Maximum. *Quaternary Science Reviews* 21(1-3): 9-31
- Dyke, A.S. 2004. An outline of North American Deglaciation with emphasis on central and northern Canada. Pp 373-424 in: J. Ehlers and P. L. Gibbard (Eds). *Quaternary Glaciations- Extent and Chronology*, Part II, Elsevier.
- Dyke, A.S. 2005. Late Quaternary vegetation history of northern North America based on pollen, macrofossil and faunal remains. *Géographie Physique et Quaternaire* 59: 211-262.
- Dyke, A.S., and Savelle, J.M. 2009. Paleoeskimo demography and sea-level history, Kent Peninsula and King William Island, Central Northwest Passage, Arctic Canada. *Arctic* 62(4): 371-392.
- Ekdahl, E.J., Teranes, J.L., Guilderson, T.P., Turton, C.L., McAndrews, J.H., Wittkop, C.A., and Stoermer, E.F. 2004. Prehistorical record of cultural eutrophication from Crawford Lake, Canada. *Geology* 32(9): 745-748.
- Ellis, E.C. 2017. Physical geography in the Anthropocene. *Progress in Physical Geography: Earth and Environment* 41(5): 525-532.
- Elsig, J., Schmitt, J., Leuenberger, D., Schneider, R., Eyer, M., Leuenberger, M., Joos, F., Fischer, H., and Stocker, T.F. 2009. Stable isotope constraints on Holocene carbon cycle changes from an Antarctic ice core. *Nature* 461: 507-510.
- Erlandson, J.M., Graham, M.H., Bourque, B.J., Corbett, D., Estes, J.A., and Steneck, R. 2007. The Kelp Highway Hypothesis: marine ecology, the Coastal Migration Theory, and the peopling of the Americas. *Journal of Island and Coastal Archaeology* 2: 161-174.
- Erlandson, J.M., Moss, M.L., and Des Lauriers, M. 2008. Life on the edge: early maritime cultures of the Pacific Coast of North America. *Quaternary Science Reviews* 27(23-24): 2232-2245.
- Erlandson, J.M., Braje, T.J., Gill, K.M., and Graham, M.H. 2015. Ecology of the Kelp Highway: did marine resources facilitate human dispersal from northeast Asia to the Americas? *Journal of Island and Coastal Archaeology* 10: 392-411.
- Fagan, B.M. 2000. *Ancient North America: the archaeology of a continent*. London & New York: Thames & Hudson.
- Faison, E.K., Foster, D.R., Oswald, W.W., Hansen, B.C.S., and Doughty, E. 2006. Early Holocene openlands in southern New England. *Ecology* 84:2537-3547.
- Faust, F.X., Gnecco, C., and Mannstein, H. 2006. Evidence for the postconquest demographic collapse of the Americas in historical CO<sub>2</sub> levels. *Earth Interactions* 10: 1-14.
- Ferretti, D.F., Miller, J.B., White, J.W.C., Ehteridge, D.M., Lasseym K.R., Lowe, D.C., MacFarling Meure, C.M., Dreier, M.F., Trudinger, C.M., van Ommen, T.D., and Langenfelds, R.L. 2005. Unexpected changes to the global methane budget over the past 2000 years. *Science* 309: 1714-1717.
- Fertig, F. 1970. Child of nature, the American Indian as an ecologist. *Sierra Club Bulletin* 55: 4-7.
- Fiedel, S.J., and Kuzmin, Y.V. 2007. Radiocarbon date frequency as an index of intensity of Paleolithic occupation of Siberia: Did humans react predictably to climate oscillations? *Radiocarbon* 49(2):741-56.
- Firestone, R.B., West, A., Kennett, J.P., Becker, L., Bunch, T.E., Revay, Z.S., Schultz, P.H., Belgya, T., Kennett, D.J., Erlandson, J.M., Dickenson, O.J., Goodyear, A.C., Harris, R.S., Howard, G.A., Kloosterman, J.B., Lechler, P., Mayewski, P.A., Montgomery, J., Poreda,

- R., Darrah, T., Que Hee, S.S., Smith, A.R., Stich, A., Topping, W., Wittke, J.H., and Wolbach, W.S. 2007. Evidence for an extraterrestrial impact 12,900 years ago that contributed to the megafaunal extinctions and the Younger Dryas cooling. *Proceedings of the National Academy of Science USA* 104(41): 16016-16021.
- Flohr, P., Fleitmann, D., Matthews, R., Matthews, W., and Black, S., 2015. Evidence of resilience to past climate change in southwest Asia: early farming communities and the 9.2 and 8.2 events. *Quaternary Science Reviews* 136: 23-39.
- Foster, D.R. 2000. From bobolinks to bears: interjecting geographical history into ecological studies, environmental interpretation, and conservation planning. *Journal of Biogeography* 27(1): 27-30.
- Foster, D., Clyden, S., Orwig, D., Hall, B., and Barry, S. 2002. Oak, chestnut and fire: climatic and cultural controls of long-term forest dynamics in New England, USA. *Journal of Biogeography* 29: 1359-1379.
- Foster III, H.T., Black, B., and Abrams, M. 2004. A witness tree analysis of the effects of Native American Indians on the pre-European settlement forests in east-central Alabama. *Human Ecology* 32: 27-47.
- Foster, W.C. 2012. *Climate and Culture Change in North America AD900-1600*. University of Texas Press, Austin, Texas, USA.
- Fowler, C., and Konopik, E. 2007. The history of fire in the southern United States. *Human Ecology Review* 14(2): 165-176.
- Fredh, D., Mazier, F., Brag e, P., Lager as, P., Rundgren, M., Hammarlund, D., and Brostr om, A. 2017. The effect of local land-use changes on floristic diversity during the past 1000 years in southern Sweden. *Holocene* 27(5): 694-711.
- French, J.C., and Collins, C. 2015. Upper Palaeolithic population histories of Southwestern France: a comparison of the demographic signatures of 14C date distributions and archaeological site counts. *Journal of Archaeological Science* 55: 122-134
- Friesen, T.M. 2007. Hearth rows, hierarchies and Arctic hunter-gatherers: the construction of equality in the Late Dorset period. *World Archaeology* 39(2): 194-214.
- Friesen, T.M., and Arnold, C.D. 2008. The timing of the Thule migration: new dates from the western Canadian Arctic. *American Antiquity* 73(3): 527-538.
- Fu, Q., Hajdinjak, M., Moldovan, O.T., Constantin, S., Mallick, S., Skoglund, P., Patterson, N., Rohland, N., Lazaridis, I., Nickel, B., Viola, B., Pr ufer, K., Meyer, M., Kelso, J., Reich, D., and P aabo, S. 2015. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* 524: 216-219.
- Fuller, J.L., Foster, D.R., McLachlan, J., and Drake, N. 1998. Impact of human activity on regional forest composition and dynamics in central New England. *Ecosystems* 1:76-95.
- Fyfe, R.M., Twiddle, C., Sugita, S., Gaillard, M.-J., Barratt, P., Caseldine, C.J., Dodson, J., Edwards, K.J., Farrell, M., Froyd, C., Grant, M.J., Huckerby, E., Innes, J.B., Shawn, H., and Waller, M. 2013. The Holocene vegetation cover of Britain and Ireland: overcoming problems of scale and discerning patterns of openness. *Quaternary Science Reviews* 73: 132-148.
- Gaillard, M.-J., Sugita, S., Bunting, M.J., Middleton, R., Brostr om, A., Caseldine, C., Giesecke, T., Hellman, S.E.V., Hicks, S., Hjelle, K., Langdon, C., Nielsen, A.-B., Poska, A., von Stedingk, H., Veski, S., and POLLANDCAL members. 2008. The use of modelling and simulation approach in reconstructing past landscapes from fossil pollen data: a review

- and results from the POLLANDCAL network. *Vegetation History and Archaeobotany* 14: 419-443.
- Gaillard, M.-J., Sugita, S., Mazier, F., Trondman, A.-K., Brostrom, A., Hickler, T., Kaplan, J.O., Kjellstrom, E., Kokfelt, U., Kunes, P., et al. 2010. Holocene land-cover reconstructions for studies on land cover-climate feedbacks. *Climate of the Past* 6: 483-499.
- Gaillard, M.-J. and LandCover6k Steering group members. 2015. LandCover6k: Global past anthropogenic land cover change and its role in past climate. *PAGES MAGAZINE* 23(1): 38-39.
- Gajewski, K. 1987. Climatic impacts on the vegetation of eastern North America for the past 2000 years. *Vegetatio* 68:179-190.
- Gajewski, K., Swain, A.M., and Peterson, G.M. 1987. Late Holocene pollen stratigraphy in four northeastern United States lakes. *Géographie physique et quaternaire* 41: 377-386.
- Gajewski, K. 1988. Late-Holocene climate changes in eastern North America estimated from pollen data. *Quaternary Research* 29:255-262.
- Gajewski, K., Viau, A.E., Sawada, M., Atkinson, D.E., and Fines, P., 2006. Synchronicity in climate and vegetation transitions between Europe and North America during the Holocene. *Climatic Change* 78: 341-361.
- Gajewski, K., Munoz, S., Peros, M., Viau, A., Morlan, R., and Betts, M. 2011. The Canadian Archaeological Radiocarbon Database (CARD): Archaeological <sup>14</sup>C dates in North America and their paleoenvironmental context. *Radiocarbon* 53: 371-394.
- Gajewski, K., Kriesche, B., Chaput, M.A., Kulik, R., and Schmidt, V. *In preparation*. Human-vegetation interactions during the Holocene in North America.
- Gamble, C., Davies, W., Pettitt, P., and Richards, M. 2004. Climate change and evolving human diversity in Europe during the last glacial. *Philosophical Transactions of the Royal Society B* 359: 243-254.
- Gamble, C., Davies, W., Pettitt, P., Hazelwood, L., and Richards, M. 2005. The archaeological and genetic foundations of the European population during the Late Glacial: Implications for 'agricultural thinking'. *Cambridge Archaeological Journal* 15: 193-223.
- Gauthier, R. 1981. Histoire de la colonisation végétale postglaciaire des Montérégiennes : deux sites du mont Saint-Bruno. Thesis. Université de Montréal, Montréal, Quebec, Canada.
- Genton, M.G., and Kleiber, W. 2015. Cross-covariance functions for multivariate geostatistics. *Statistical Science* 30(2): 147-163.
- Gil, A., Zárata, M., and Neme, G. 2005. Mid-Holocene paleoenvironments and the archaeological record of southern Mendoza, Argentina. *Quaternary International* 132: 81-94.
- Gilbert, M.T.P., Kivisild, T., Gronnow, B., Andersen, P.K., Metspalu, E., Reidla, M., Tamm, E., Axelsson, E., Gotherstrom, A., Campos, P.F., Rasmussen, M., Metspalu, M., Higham, T.F.G., Schwenninger, J.-L., Nathan, R., De Hoog, C.-J., Koch, A., Moller, L.N., Andreasen, C., Meldgaard, M., Villems, R., Bendixen, C., and Willerslev, E. 2008. Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. *Science* 320(5884): 1787-1789.
- Gleason, H.A. 1913. The relation of forest distribution and prairie fires in the Middle West. *Torreyia* 13: 173-181.
- Godwin, H. 1984. History of the British Flora. Cambridge University Press.
- Goebel, T., Waters, M.R., and O'Rourke, D.H. 2008. The late Pleistocene dispersal of modern humans in the Americas. *Science* 319: 1497-1502.

- Goebel, T., Smith, H.L., DiPietro, L., Waters, M.R., Hockett, B., Graf, K.E., Gal, R., Slobodin, S.B., Speakman, R.J., Driese, S.G., and Rhode, D. 2013. Serpentine Hot Springs, Alaska: results of excavations and implications for the age and significance of northern fluted points. *Journal of Archaeological Science* 40(12): 4222-4233.
- González-Sampérez, P., Utrilla, P., Mazo, C., Valero-Garcés, B., Sopena, M.C., Morellón, M., Sebastián, M., Moreno, A., and Martínez-Bea, M. 2009. Patterns of human occupation during the early Holocene in the Central Ebro Basin (NE Spain) in response to the 8.2 ka climatic event. *Quaternary Research* 71(2): 121-132.
- Goodyear, A.C. 2004. Evidence for Pre-Clovis sites in the eastern United States. Pp 103-112 in: *Paleoamerican Origins: Beyond Clovis*. R. Bonnicksen, B.T. Lepper, D. Stanford, M.R. Waters (Eds). Center for the Study of First Americans, Texas A&M University, College Station, United States of America.
- Gordon, B.C. 1975. *Of Men & Herds in Barrenland Prehistory* (National Museum of Man Mercury Series, Archaeological Survey of Canada Paper No. 28. Ottawa: Canadian Museum of Civilization.
- Gordon, B.C. 1976. *Migod – 8000 Years of Barrenland prehistory*. National Museum of Man Mercury Series, Archaeological Survey of Canada Paper No.56. Ottawa: Canadian Museum of Civilization.
- Gordon, B.C. 1996. *People of Sunlight; People of Starlight: Barrenlands Archaeology in the Northwest Territories of Canada*. National Museum of Man Mercury Series, Archaeological Survey of Canada Paper No.154. Ottawa: Canadian Museum of Civilisation.
- Goring, S., Dawson, A., Simpson, G.L., Ram, K., Graham, R.W., Grimm, E.C., and Williams, J.W. 2015. Neotoma: a programmatic interface to the Neotoma paleoecological database. *Open Quaternary* 1(2): 1–17.
- Graf, K.E., and Buvit, I. 2017. Human Dispersal from Siberia to Beringia: Assessing a Beringian Standstill in Light of the Archaeological Evidence. *Current Anthropology* 58(17): S583-S603.
- Grimm, E.C. 2000. Global Pollen Database. IGBP PAGES/World Data Centre for Paleoclimatology, NOAA/NGDC Paleoclimatology Program, Boulder, CO, USA <http://www.ngdc.noaa.gov/paleo/pollen.html>.
- Grimm, E.C. 2008. Neotoma - an ecosystem database for the Pliocene, Pleistocene and Holocene. *Illinois State Museum Scientific Papers E Series* volume 1.
- Government of Quebec. 1994. *Le Point d'Observation Écologique: Normes Techniques*. Ministry of Natural Resources. 116 p.
- Grove, M. 2011. A spatio-temporal kernel method for mapping changes in prehistoric land-use patterns. *Archaeometry* 53(5): 1012-1030.
- Gupta, A.K. 2004. Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Current Science* 87(1): 54-59.
- Hadden, K. 1978. Investigation of the registration of poplar pollen in lake sediments. Thesis, University of Toronto, Toronto, Ontario, Canada.
- Hajic, E.R., Mandel, R.D., Ray, J.H., Lopinot, N.H. 2007. Geoarchaeology of stratified Paleoindian deposits at the Big Eddy Site, Southwest Missouri, USA. *Geoarchaeology* 22(8): 891-934
- Hall, P., Fisher, N.I., and Hoffmann, B. 1994. On the nonparametric estimation of covariance functions. *Annals of Statistics* 22: 2115–2134.

- Härdle, W., Müller, M., Sperlich, S., and Werwatz, A. 2004. *Nonparametric and Semiparametric Models*. Berlin: Springer.
- Harkin, M.E., and Lewis, D.R. 2007. *Native Americans and the Environment: perspectives on the Ecological Indian*. University of Nebraska Press, Lincoln, Nebraska, USA.
- Hart, J.P., and Lovis, W.A. 2013. Reevaluating what we know about the histories of maize in northeastern North America: a review of current evidence. *Journal of Archaeological Research* 21(2): 175-216
- Haslett, J. and Parnell, A. 2008. A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society Series C* 57(4): 399–418.
- Hatfield, V.L. 2010. Material culture across the Aleutian Archipelago. *Human Biology* 82(5-6): 525-556.
- Haynes, G., Anderson, D.G., Ferring, C.R., Fiedel, S.J., Grayson, D.K., Haynes Jr., C.V., Holliday, V.T., Huckell, B.B., Kornfeld, M., Meltzer, D.J., Morrow, J., Surovell, T., Waguespack, N.M., Wigand, P., and Tohe, R.M., 2007. Comment on “Redefining the age of Clovis; implications for the peopling of the Americas”. *Science* 317: 320b.
- Hellerstein, J.M. 2008. Quantitative data cleaning for large databases. White paper, United Nations Economic Commission for Europe.
- Hellman, S., Gaillard, M.-J., Broström, A., and Sugita, S. 2008. The REVEALS model, a new tool to estimate past regional plant abundance from pollen data in large lakes: Validation in southern Sweden. *Journal of Quaternary Science* 23(1): 21-42.
- Hey, J. 2005. On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biology* 3(6): e193.
- Hinz, M., Furrholt, M., Müller, J., Raetzl-Fabian, D., Rinne, C., Sjögren, K.-G., and Wotzka, H.-P., 2012. RADON – Radiocarbon dates online 2012. Central European database of <sup>14</sup>C dates for the Neolithic and Early Bronze Age. Available at: <http://radon.ufg.uni-kiel.de/>.
- Hjelle, K.L., Mehl, I.K., Sugita, S., and Andersen, G.L. 2015. From pollen percentage to vegetation cover: evaluation of the Landscape Reconstruction Algorithm in western Norway. *Journal of Quaternary Science* 30(4): 312-324.
- Hoffecker, J.F., Elias, S.A., and O’Rourke, D.H. 2014. Out of Beringia? *Science* 343:979–980.
- Hoffecker, J.F., Elias, S.A., O’Rourke, D.H., Scott, G.R., and Bigelow, N.H. 2016. Beringia and the Global Dispersal of Modern Humans. *Evolutionary Anthropology* 25: 64-78.
- Holdaway, S., and Porch, N. 1995. Cyclical patterns in the Pleistocene human occupation of Southwest Tasmania. *Archaeology in Oceania* 30(2): 74-82.
- Hopkins, D.M., Matthews, J.V., Schweger, S.E., and Young, S.B. 1982. *Paleoecology of Beringia*. New York: Academic Press.
- Hubbe, M., Neves, W.A., and Harvati, K. 2010. Testing Evolutionary and Dispersion Scenarios for the Settlement of the New World. *PLoS ONE* 5(6): e11105.
- Huntley, B. and T. Webb III (eds.). 1988. *Vegetation History*. Vol. 7 in *Handbook of Vegetation Science*. Kluwer Academic Publ., Dordrecht, The Netherlands.
- Hurttt, G.C., Chini, L.P., Frolking, S., Betts, R.A., Feddema, J., Fischer, G., Fisk, J.P., Hibbard, K., Houghton, R.A., Janetos, A., Jones, C.D., Kinermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., Van Vuuren, D.P., and Wang, Y.P., 2011. Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change* 109: 117-161.

- Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. 2008. *Statistical Analysis and Modelling of Spatial Point Patterns*. Chichester: J. Wiley & Sons.
- Jackson, S.T., Overpeck, J.T., Webb, III, T., Keattch, S.E., and Anderson, K.H. 1997. Mapped plant-macrofossil and pollen records of late Quaternary vegetation change in eastern North America. *Quaternary Science Reviews* 16(1): 1-70.
- Jackson, S.T., Webb, R.S., Anderson, K.H., Overpeck, J.T., Webb, III, T., Williams, J.W., and Hansen, B.C.S. 2000. Vegetation and environment in Eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews* 19(5): 489-508.
- Jacobson, G.L. Jr., Webb III, T., and Grimm, E.C. 1987. Patterns and rates of vegetation change during the deglaciation of eastern North America. Pages 277-288 in W.F. Ruddiman, and H.E Wright, Jr., editors. *North America and adjacent oceans during the last deglaciation*. Geological Society of America, Geology of North America, Vol. K.3, Boulder, Colorado, USA.
- Jäger, K.-D., and Neuhäusl, R. 1994. Interactions between natural environment and Neolithic man in Central Europe - an investigation based on comparative studies on vegetation and settlement with special emphasis on the view of natural science. In B. Frenzel (Ed.), *Evaluation of Land Surfaces Cleared from Forests in the Roman Iron Age and the Time of Migrating Germanic Tribes Based on Regional Pollen Diagrams*, pp. 75–81. Stuttgart: Gustav Fischer Verlag.
- Johnson, E.A., and Miyanishi, K. 2012. The boreal forest as a cultural landscape. *Annals of the New York Academy of Sciences* 1249: 151-165.
- Johnson, A.L. 2014. Exploring adaptive variation among Hunter-gatherers with Binford's Frames of Reference. *Journal of Archaeological Research* 22: 1-42.
- Johnson, A., Gil, A., Neme, G., and Freeman, J. 2015. Hierarchical method using ethnographic data sets to guide archaeological research: Testing models of plant intensification and maize use in Central Western Argentina. *Journal of Anthropological Archaeology* 38: 52-58.
- Kaplan, J.O., Krumhardt, K.M., and Zimmermann, N. 2009. The prehistoric and preindustrial deforestation of Europe. *Quaternary Science Reviews* 28: 3016-3034.
- Kaplan, J.O., Krumhardt, K.M., Ellis, E.C., Ruddiman, W.F., Lemmen, C., and Goldewijk, K.K. 2010. Holocene carbon emissions as a result of anthropogenic land cover change. *Holocene* 21: 775-791.
- Kay, C.E. 1994. Aboriginal overkill: the role of Native Americans in structuring western ecosystems. *Human Nature* 5: 359-398.
- Kay, C.E., and Simmons, R.T. 2002. *Wilderness and Political Ecology*. University of Utah Press, Salt Lake City, Utah, USA
- Kelly, R.L., and Todd, L.W. 1988. Coming into the country: early Paleoindian hunting and mobility. *American Antiquity* 53(2): 231-244.
- Kelly, R.L. 1996. Ethnographic analogy and migration to the western hemisphere. In: Akazawa, T., Szathmary, E.J.E. (Eds.), *Prehistoric Mongoloid Dispersals*. Oxford University Press, Oxford, pp. 228-240.
- Kelly, R.L. 1997. Late Holocene Great Basin Prehistory. *Journal of World Prehistory* 11(1): 1-49.
- Kelly, R.L., Surovell, T.A., Shuman, B.N., Smith, G.M. 2013. A continuous climatic impact on Holocene human population in the Rocky Mountains. *Proceedings of the National Academy of Sciences USA* 110(2): 443-447.

- Kirch, P.V. 2005. Archaeology and global change: The Holocene record. *Annual Review of Environmental Resources* 30: 409–440.
- Kitchen, A., Miyamoto, M.M., and Mulligan, C.J. 2008. A three-stage colonization model for the peopling of the Americas. *PLoS ONE* 3(2): e1596.
- Klein Goldewijk, K., Beusen, A., and Janssen, P. 2010. Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. *Holocene* 20: 565-573.
- Klein Goldewijk, K., Beusen, A., Van Drecht, G., and De Vos, M. 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecological Biogeography* 20: 73-86.
- Krech, S. 2000. *The Ecological Indian: myth and history*. W.W. Norton and Company, New York, New York, USA.
- Kriesche, B., Chaput, M.A., Kulik, R., Gajewski, K., and Schmidt, V. *Submitted*. Estimation of spatio-temporal correlations of prehistoric population and vegetation in North America. *Geographical Analysis*.
- Kroeber, A. L. 1939 Cultural and natural areas of native North America. University of California Publications in American Archaeology and Ethnology 38. Berkeley, p. 134.
- Kuzmin, Y.V., and Keates, S.G. 2005. Dates are not just data: Paleolithic settlement patterns in Siberia derived from radiocarbon records. *American Antiquity* 70(4): 773-789.
- Labelle, C., and Richard, P.J.H. 1981. Végétation tardiglaciaire et postglaciaire au sud-est du Parc des Laurentides, Quebec. *Géographie physique et Quaternaire* 35:345-359.
- Lacourse, T., Mathewes, R.W., and Hebda, R.J. 2007. Paleoecological analyses of lake sediments reveal prehistoric human impact on forests at Anthony Island UNESCO World Heritage Site, Queen Charlotte Islands (Haida Gwaii), Canada. *Quaternary Research* 68: 177-183.
- Lafleur, B., Paré, D., Munson, A.D., and Bergeron, Y. 2010. Response of northeastern North American forests to climate change: Will soil conditions constrain tree species migration? *Environmental Reviews* 18: 279-289.
- Lafontaine-Boyer, K., and Gajewski, K. 2014. Vegetation dynamics in relation to late-Holocene climate variability and disturbance, Outaouais, Quebec, Canada. *Holocene* 24: 1515-1526.
- Lahmeyer, J. 2004. Populstat database. Growth of the population per country in a historical perspective, including their administrative divisions and principal towns. Available at: <http://www.populstat.info/>.
- Latorre, C., Santoro, C.M., Ugalde, P.C., Gayo, E.M., Osorio, D., Salas-Egaña, C., De Pol-Holz, R., Joly, D., and Rech, J.A. 2013. Late Pleistocene human occupation of the hyperarid core in the Atacama Desert, northern Chile. *Quaternary Science Reviews* 77: 19-30.
- Lavoie, M., and Filion, L. 2001. Holocene Vegetation Dynamics of Anticosti Island, Quebec, and Consequences of Remoteness on Ecological Succession. *Quaternary Research* 56: 112-127.
- Ledo, A., Condés, S., and Montes, F. 2011. Intertype mark correlation function: a new tool for the analysis of species interactions. *Ecological Modelling* 222(3): 580–587.
- Lessard, G. 2014. Quebec, ville de bois. *The Forestry Chronicle* 90(3): 330-333.
- Lesure, R. 2008. The Neolithic demographic transition in Mesoamerica? Larger implications of the strategy of relative chronology. *The Neolithic demographic transition and its consequences* ed Bocquet-Appel JP (Springer, New York), pp. 107-138.
- Lewis, S.L., and Maslin, M.A. 2015. Defining the Anthropocene. *Nature* 519: 171-180.

- Li, Y., Nielsen, A.B., Zhao, X., Shan, L., Wang, S., Wu, J., and Zhou, L. 2015a. Pollen production estimates (PPEs) and fall speeds for major tree taxa and relevant source areas of pollen (RSAP) in Changbai Mountain, northeastern China. *Review of Palaeobotany and Palynology* 216: 92-100.
- Li, H., An, C., Fan, W., Dong, W., Zhao, Y., and Wang, H. 2015b. Population history and its relationship with climate change on the Chinese Loess Plateau during the past 10,000 years. *Holocene* 25, 1144-1152.
- Liebmann, M.J., Farella, J., Roos, C.I., Stack, A., Martini, S., and Swetnam, T.W. 2016. Native American depopulation, reforestation, and fire regimes in the Southwest United States, 1492–1900 CE. *Proceedings of the National Academy of Sciences USA* 113(6): E696-E704.
- Lightfoot, K., and Parrish, O. 2009. *California Indians and their Environment: An introduction*. California Natural History Guides. University of California Press.
- Lightfoot, K., and Cuthrell, R. 2015. Anthropogenic burning and the Anthropocene in late-Holocene California. *The Holocene* 25(10): 1581–1587.
- Livi-Bacci, M. 2007. A concise history of world population. Fourth edition. Oxford: Blackwell Publishing.
- Lombardo, U., Szabo, K., Capriles, J.M., May, J.-H., Amelung, W., Hutterer, R., Lehndorff, E., Plotzki, A., and Veit, H., 2013. Early and middle Holocene hunter-gatherer occupations in Western Amazonia: the hidden shell middens. *PLoS ONE* 8: e72746.
- López Mazz, J.M. 2013. Early human occupation of Uruguay: radiocarbon database and archaeological implications. *Quaternary International* 301: 94-103.
- MacDonald, G.F. 1968. Debert: A Palaeo-Indian site in central Nova Scotia. *Anthropology Papers No. 16*. (Ottawa: National Museums of Canada. 207 p.).
- MacDonald, G.M., Beilman, D.W., Kuzmin, Y.V., Orlova, L.A., Kremenetski, K.V., Shapiro, B., Wayne, R.K., and Van Valkenburgh, B. 2012. Pattern of extinction of the woolly mammoth in Beringia. *Nature Communications* 3: 893.
- MacDougall, A. 2003. Did Native Americans influence the northern migration of plants during the Holocene? *Journal of Biogeography* 30: 633-647.
- Maddison, A. 2001. *The world economy: a millennial perspective*. Paris: OECD.
- Malhi, Y., Doughty, C.E., Galetti, M., Smith, F.A., Svenning, J.-C., and Terborgh, J.W. 2016. Megafauna and ecosystem function from the Pleistocene to the Anthropocene. *Proceedings of the National Academy of Sciences USA* 113(4): 838-846.
- Mancini, M.V., Franco, N.V., and Brook, G.A. 2013. Palaeoenvironment and early human occupation of the southernmost South America (South Patagonia, Argentina). *Quaternary International* 299: 13-22.
- Mandryk, C.A.S., Josenhans, H., Fedje, D.W., and Mathewes, R.W. 2001. Late Quaternary paleoenvironments of Northwestern North America: implications for inland versus coastal migration routes. *Quaternary Science Reviews* 20: 301-314.
- Martin, P. 1973. The discovery of America. *Science* 179 (4077): 969-974.
- Martin, P. 1984. Prehistoric Overkill: The Global Model. In *Quaternary Extinctions: A Prehistoric Revolution*, edited by P.S. Martin and R.G. Klein. 354-403. University of Arizona Press.
- Martindale, A., Morlan, R., Betts, M., Blake, M., Gajewski, K., Chaput, M., Mason, A., and Vermeersch, P. 2016. Canadian Archaeological Radiocarbon Database (CARD 2.1). Available at [www.canadianarchaeology.ca](http://www.canadianarchaeology.ca).

- Marlon, J.R., Bartlein, P.J., Daniau, A.-L., Harrison, S.P., Maezumi, S.Y., Power, M.J., Tinner, W., and Vanni re, B. 2013. Global biomass burning: a synthesis and review of Holocene paleofire records and their controls. *Quaternary Science Reviews* 65: 5-25.
- Maschner, H.D.G. 2012. Archaeology of the northwest coast. *North American Archaeology* eds Pauketat TR, Loren DD (Blackwell Publishing, Oxford), pp. 160-172.
- Mazier, F., Gaillard, M.-J., Kuneš, P., Sugita, S., Trondman, A.-K., and Brostr m, A. 2012. Testing the effect of site selection and parameter setting on REVEALS-model estimates of plant abundance using the Czech Quaternary Palynological Database. *Review of Palaeobotany and Palynology* 187: 38-49.
- McAndrews, J.H. 1988. Human disturbance of North American forests and grasslands: the fossil pollen record. Pages 673-697 in B. Huntley, and T. Webb III, editors. *Vegetation History* volume of Handbook of Vegetation Sciences series, Kluwer, Utrecht, Netherlands.
- McAndrews, J.H., and Boyko-Diakonow, M. 1989. Pollen analysis of varved sediment at Crawford Lake, Ontario: evidence of Indian and European farming. Pages 528-530 in R.J. Fulton, editor. *Quaternary Geology of Canada and Greenland*, Geological Survey of Canada, Geology of Canada no 1.
- McEvedy, C., and Jones, R. 1978. *World atlas of population history*. Hammondsworth: Penguin Books Ltd.
- McKechnie, I. 2007. Investigating the complexities of sustainable fishing at a prehistoric village on western Vancouver Island, British Columbia, Canada. *Journal for Nature Conservation* 15(3): 208-222.
- McLaren, D., Martindale, A., Fedje, D., and Mackie, Q. 2011. Relict shorelines and shell middens of the Dundas Archipelago. *Canadian Journal of Archaeology* 35(1): 86-116.
- McLauchlan, K. 2003. Plant cultivation and forest clearance by prehistoric North Americans: pollen evidence from Fort Ancient, Ohio, USA. *Holocene* 13: 557-566.
- McShea, W., and Healy, W. 2002. *Oak Forest Ecosystems*. Baltimore: John Hopkins University Press.
- M endez, C., Gil, A., Neme, G., Nuevo Delaunay, A., Cortegoso, V., Huidobro, C., Duran, V., and Maldonado, A. 2015. Mid Holocene radiocarbon ages in the Subtropical Andes (~29 -35 S), climatic change and implications for human space organization. *Quaternary International* 356: 15-26.
- Merrill, W.M., Hard, R.J., Mabry, J.B., Fritz, G.J., Adams, K.R., Roney, J.R., and MacWilliams, A.C. 2009. The diffusion of maize to the southwestern United States and its impact. *Proceedings of the National Academy of Sciences USA* 106(50): 21019-21026.
- Miller, D.S., Gingerich, J.A.M. 2013. Regional variation in the terminal Pleistocene and early Holocene radiocarbon record of eastern North America. *Quaternary Research* 79: 175-188.
- Minnis, P.E. 2010. *People and plants in ancient Eastern North America*. University of Arizona Press.
- Montero, J.-M., Fern ndez-Avil s, G., and Mateu, J. 2015. *Spatial and Spatio-Temporal Geostatistical Modeling and Kriging*. Chichester: J. Wiley & Sons.
- Mooney, J. 1910 Population. In: *Handbook of American Indians North of Mexico*, F. W. Hodge, ed. B. A. E. Bull., 30: 286-287, Part 2. Washington.
- Mooney, J. 1928. The aboriginal population of America north of Mexico. *Smithsonian Miscellaneous Collections* 80 ed Swanton JR, pp. 1-40.

- Moran, E.F., and Lopez, M.C. 2016. Future directions in human-environment research. *Environmental Research* 144: 1-7.
- Moreno-Mayar, J.V., Potter, B.A., Vinner, L., Steinrücken, M., Rasmussen, S., Terhorst, J., Kamm, J.A., Albrechtsen, A., Malaspinas, A.-S., Sikora, M., Reuther, J.D., Irish, J.D., Malhi, R.S., Orlando, L., Song, Y.S., Nielsen, R., Meltzer, D.J., and Willerslev, E. 2018. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* 553: 203-207.
- Mott, R.J. 1978. Populus in late-Pleistocene pollen spectra. *Canadian Journal of Botany* 56: 1021-1031.
- MRNFQ. 2006. Normes d'inventaire Forestier: placettes-échantillons permanentes, 2006 edn. Direction des Inventaires Forestiers, Ministère des Ressources Naturelles et de la Faune, Quebec.
- Muckelroy, K. 1976. The Integration of Historical and Archaeological Data concerning an Historic Wreck Site: The 'Kennemerland'. *World Archaeology* 7(3) : 280-290.
- Muller, S.D., and Richard, J.H. 2001. Post-glacial vegetation migration in conterminous Montréal Lowlands, southern Quebec. *Journal of Biogeography* 28: 1169-1193.
- Munoz, S.E., and Gajewski, K. 2010. Distinguishing prehistoric human influence on late-Holocene forests in southern Ontario, Canada. *Holocene* 20: 967-981.
- Munoz, S.E., Gajewski, K., and Peros, M.C. 2010. Synchronous environmental and cultural change in the prehistory of northeastern United States. *Proceedings of the National Academy of Sciences USA* 107: 22008-22013.
- Munoz, S.E., Mladenoff, D., Schroeder, S., and Williams, J. 2014. Defining the spatial patterns of historical land use associated with indigenous societies of eastern North America. *Journal of Biogeography* 41: 2195-2210.
- Neotoma. 2017. Paleocology Database. Available at [www.neotomadb.org](http://www.neotomadb.org).
- Nevle, R.J., Bird, D.K., Ruddiman, W.F., and Dull, R.A. 2011. Neotropical human-landscape interactions, fire, and atmospheric CO<sub>2</sub> during European conquest. *Holocene* 21(5): 853-864.
- Newby, P., Bradley, J., Spiess, A., Shuman, B., and Leduc, P. 2005. A Paleoindian response to Younger Dryas climate change. *Quaternary Science Reviews* 24(1-2): 141-154.
- Newson, L.A. 1996. The population of the Amazon basin in 1492: a view from the Ecuadorian headwaters. *Transactions of the Institute of British Geographers* 21: 5-26.
- Newsome, S.D., Phillips, D.L., Culleton, B.J., Guilderson, T.P., and Koch, P.L. 2004. Dietary reconstruction of an early to middle Holocene human population from the central California coast: insights from advanced stable isotope mixing models. *Journal of Archaeological Science* 31(8): 1101-1115.
- Nichols, J. 1990. Linguistic diversity and the first settlement of the New World. *Language* 66: 475-521.
- Nielsen, A.B., and Sugita, S. 2005. Estimating relevant source area of pollen for small Danish lakes around AD 1800. *Holocene* 15: 1006-1020.
- Niemeyer, B., Klemm, J., Pestryakova, L.A., and Herzschuh, U. 2015. Relative pollen productivity estimates for common taxa of the northern Siberian Arctic. *Review of Palaeobotany and Palynology* 221: 71-82.
- Nordt, L., Boutton, C.T.W., Hallmark, C.T., and Waters, M.R. 1994. Late Quaternary vegetation and climate changes in central Texas based on the isotopic composition of organic carbon. *Quaternary Research* 41: 109-120.

- North American Forest Commission (NAFC). 2011. *Forests of North America*. Vector digital data. Commission for Environmental Cooperation: Montreal, Quebec, Canada. Available at <http://www.cec.org/tools-and-resources/map-files/north-american-forests-2011>.
- O'Brien, M.J., and Buchanan, B. 2017. Cultural learning and the Clovis colonization of North America. *Evolutionary Anthropology* 26(6): 270-284.
- Ogburn, D.E. 2012. Reconceiving the chronology of Inca imperial expansion. *Radiocarbon* 54: 219-237.
- Oinonen, M., Pesonen, P., and Tallavaara, M. 2010. Archaeological radiocarbon dates for studying the population history in eastern Fennoscandia. *Radiocarbon* 52: 393-407
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Jura, Y., Lamoureux, J.F., Wettengel, W.W., Hedao, P., and Kassem, K.R. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience* 51: 933-938.
- Orme, B. 1981. *Anthropology for Archaeologists*. London: Duckworth.
- Paciorek, C.J., and McLachlan, J.S. 2009. Mapping ancient forests: Bayesian inference for spatio-temporal trends in forest composition using the fossil pollen proxy record. *Journal of the American Statistical Association* 104(486): 608-622.
- Paillet, F.L. 2002. Chestnut history and ecology of a transformed species. *Journal of Biogeography* 29: 1517-1530.
- Palmisano, A., Bevan, A., and Shennan, S. 2018. Regional Demographic Trends and Settlement Patterns in Central Italy: Archaeological Sites and Radiocarbon Dates. *Journal of Open Archaeology Data* 6(1), p.2.
- Paquette, N., and Gajewski, K. 2013. Climatic change causes abrupt changes in forest composition, inferred from a high-resolution pollen record, Southwestern Quebec, Canada. *Quaternary Science Reviews* 75: 169-180.
- Parekh, R., and Childs, E.W. 2016. *Stigma and Prejudice: Touchstones in Understanding Diversity in Healthcare*. Current Clinical Psychiatry. Springer International Publishing: Humana Press.
- Parent, M., and Occhietti, S. 1999. Late Wisconsin deglaciation and Champlain Sea invasion in the St. Lawrence Valley, Quebec. *Géographie Physique et Quaternaire* 53: 117-135.
- Park, R.W. 2012. Adapting to a frozen coastal environment. Chapter 10 in: T. Pauketat (Ed). *North American Archaeology*. Oxford University Press, New York.
- Parshall, T., and Foster, D.R. 2002. Fire on the New England landscape: regional and temporal variation, cultural and environmental controls. *Journal of Biogeography* 29: 1305-1317
- Parsons, R.W., and Prentice, C.I. 1981. Statistical approaches to R-values and the pollen-vegetation relationship. *Reviews in Palaeobotany and Palynology* 32: 127-152.
- Pauketat, T.R., and Loren, D.D. *North American Archaeology*. Blackwell Publishing, Oxford. Pp. 113-123.
- Peacock, E. 1998. Historical and applied perspectives on prehistoric land use in eastern North America. *Environment and History* 4(1): 1-29.
- Peros, M., Gajewski, K., and Viau, A. 2008. Continental-scale tree population response to rapid climate change, competition and disturbance. *Global Ecology and Biogeography* 17: 658-669.
- Peros, M.C., Munoz, S.E., Gajewski, K., and Viau, A.E. 2010. Prehistoric demography of North America inferred from radiocarbon data. *Journal of Archaeological Science* 37: 656-664.

- Pirzamanbein, B., Lindström, J., Poska, A., Sugita, S., Trondman, A.-K., Fyfe, R., Mazier, F., Nielsen, A.B., Kaplan, J.O., Bjune, A.E., Birks, H.J.B., Giesecke, T., Kangur, M., Latalowa, M., Marquer, L., Smith, B., Gaillard, M.-J. 2014. Creating spatially continuous maps of past land cover from point estimates: a new statistical approach applied to pollen data. *Ecological Complexity* 20:127–141.
- Pitblado, B.L. 2011. A tale of two migrations: reconciling recent biological and archaeological evidence for the Pleistocene peopling of the Americas. *Journal of Archaeological Research* 19(4): 327-375.
- Pongratz, J., Reick, C., Raddatz, T., and Claussen, M. 2008. A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles* 22: GB3018.
- Prates, L., Politis, G., and Steele, J. 2013. Radiocarbon chronology of the early human occupation of Argentina. *Quaternary International* 301: 104-122.
- Prentice, I.C., and Parsons, R.W. 1983. Maximum likelihood linear calibration of pollen spectra in terms of forest composition. *Biometrics* 39: 1051-1057.
- Prentice, I.C. 1985. Pollen representation, source area, and basin size: toward a unified theory of pollen analysis. *Quaternary Research* 23: 76–86.
- Prentice, I.C., Berglund, B.E. and Olsson, T. 1987. Quantitative forest-composition sensing characteristics of pollen samples from Swedish lakes. *Boreas* 16: 43-54.
- Prentice, I.C. 1988. Records of vegetation in time and space: the principles of pollen analysis. In Huntley, B. and Webb, T., III, editors, *Vegetation history*. Kluwer Academic Publishers, 1742.
- Prentice, I.C., Bartlein, P.J., Webb III, T. 1991. Vegetation and climate change in eastern North America since the Last Glacial Maximum. *Ecology* 72(6): 2038-2056.
- Price, T.D. 2009. Ancient farming in eastern North America. *Proceedings of the National Academy of Sciences USA* 106(16): 6427-6428.
- Pyne, S.J. 1982. *Fire in America*. Princeton University Press, Princeton, New Jersey, USA.
- Pyne, S.J. 2007. *Awful Splendor: a Fire History of Canada*. University of British Columbia Press, British Columbia, Vancouver, Canada.
- R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienne, Austria. Available at [www.R-project.org](http://www.R-project.org).
- Raghavan, M., Steinrücken, M., Harris, K., Schiffels, S., Rasmussen, S., DeGiorgio, M., Albrechtsen, A., Valdiosera, C., Ávila-Arcos, M.C., Malaspinas, A.-S., Eriksson, A., Moltke, I., Metspalu, M., Homburger, J.R., Wall, J., Cornejo, O.E., Moreno-Mayar, V., Korneliussen, T.S., Pierre, T., Rasmussen, M., Campos, P.F., de Barros Damgaard, P., Allentoft, M.E., Lindo, J., Metspalu, E., Rodríguez-Varela, R., Mansilla, J., Henriksen, C., Seguin-Orlando, A., Malmström, H., Stafford Jr., T., Shringarpure, S.S., Moreno-Estrada, A., Karmin, M., Tambets, K., Bergström, A., Xue, Y., Warmuth, V., Friend, A.D., Singarayer, J., Valdes, P., Balloux, F., Lebreiro, I., Vera, J.L., Rangel-Villalobos, H., Pettener, D., Luiselli, D., Davis, L.G., Heyer, E., Zollikofer, C.P.E., Ponce de León, M.S., Smith, C.I., Grimes, V., Pike, K.-A., Deal, M., Fuller, B.T., Arriaza, B., Standen, V., Luz, M.F., Ricaut, F., Guidon, N., Osipova, L., Voevoda, M.I., Posukh, O.L., Balanovsky, O., Lavryashina, M., Bogunov, Y., Khusnutdinova, E., Gubina, M., Balanovska, E., Fedorova, S., Litvinov, S., Malyarchuk, B., Derenko, M., Mosher, M.J., Archer, D., Cybulski, J., Petzelt, B., Mitchell, J., Worl, R., Norman, P.J., Parham, P., Kemp, B.M., Kivisild, T., Tyler-Smith, C., Sandhu, M.S., Crawford, M., Villems, R., Smith, D.G.,

- Waters, M.R., Goebel, T., Johnson, J.R., Malhi, R.S., Jakobsson, M., Meltzer D.J., Manica, A., Durbin, R., Bustamante, C.D., Song, Y.S., Nielsen, R., and Willerslev, E. 2015. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* 349(6250): aab3884.
- Rayburn, J.A., Franzi, D.A., Knuepfer, P.L.K. 2007. Evidence from the Lake Champlain Valley for a later onset of the Champlain Sea and implications for late glacial meltwater routing to the North Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology* 246: 62-74.
- Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M.V., Rojas, W., Duque, C., Mesa, N., García, L.F., Triana, O., Blair, S., Maestre, A., Dib, J.C., Bravi, C.M., Bailliet, G., Corach, D., Hünemeier, T., Bortolini, M.C., Salzano, F.M., Petzl-Erler, M.L., Acuña-Alonzo, V., Aguilar-Salinas, C., Canizales-Quinteros, S., Tusié-Luna, T., Riba, L., Rodríguez-Cruz, M., Lopez-Alarcón, M., Coral-Vazquez, R., Canto-Cetina, T., Silva-Zolezzi, I., Fernandez-Lopez, J.-C., Contreras, A.V., Jimenez-Sanchez, G., Gómez-Vázquez, M.J., Molina, J., Carracedo, A., Salas, A., Gallo, C., Poletti, G., Witonsky, D.B., Alkorta-Aranburu, G., Sukernik, R.I., Osipova, L., Fedorova, S.A., Vasquez, R., Villena, M., Moreau, C., Barrantes, R., Pauls, D., Excoffier, L., Bedoya, G., Rothhammer, F., Dugoujon, J.-M., Larrouy, G., Klitz, W., Labuda, D., Kidd, J., Kidd, K., Di Rienzo, A., Freimer, N.B., Price, A.L., and Ruiz-Linares, A. 2012. Reconstructing Native American population history. *Nature* 488: 370-375.
- Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Hafliðason, H., Hajdas, I., Hatté, 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., and Van Der Plicht, J., 2013. IntCal13 and Marine 13 radiocarbon age calibration curves 0-50,000 years cal BP. *Radiocarbon* 55: 1869-1887.
- Richard, P., and Poulin, P. 1976. Un diagramme pollinique au Mont des Éboulements, région de Charlevoix, Quebec. *Canadian Journal of Earth Sciences* 13 : 145-156.
- Richard, P. 1977a. Histoire post-wisconsinienne de la végétation du Quebec méridional par l'analyse pollinique. Service de la recherche, Direction générale des forêts, Ministère des Terres et Forêts du Quebec. Publications et rapports divers. Tome 1, xxiv + 312 p.; tome 2, 142 p.
- Richard, P. 1977b. Végétation tardiglaciaire au Québec méridional et implications paléoclimatiques. *Géographie physique et Quaternaire* 31 : 161-176.
- Richard, P. 1978. Histoire tardiglaciaire et postglaciaire de la végétation au mont Shefford, Quebec. *Géographie physique et Quaternaire* 32 : 81-93.
- Richard, P. 1985. Couvert végétal et paléoenvironnements du Quebec entre 12000 et 8000 ans BP. *Recherches amérindiennes au Quebec* 15 : 39-56.
- Richard, P., Larouche, A.C., and Lortie, G. 1992. Paléophytogéographie et paléoclimats postglaciaires dans l'ouest du Bas-Saint-Laurent, Quebec. *Géographie physique et Quaternaire* 46: 151-172.
- Richard, P. 1994. Postglacial palaeophytogeography of the eastern St. Lawrence River Watershed and the climatic signal of the pollen record. *Palaeogeography, Palaeoclimatology, Palaeoecology* 109 : 137-161.
- Richard, P. 1995. Le couvert végétal du Quebec-Labrador il y a 6000 ans BP : essai. *Géographie physique et Quaternaire* 49 : 117-140.

- Rick, J.T. 1987. Dates as data: an examination of the Peruvian Preceramic radiocarbon record. *American Antiquity* 52: 55-73.
- Ripple, W.J., and Van Valkenburgh, B. 2010. Linking top-down forces to the Pleistocene megafaunal extinctions. *Bioscience* 60 : 516-526.
- Ritchie, J. C. 1987. Postglacial Vegetation of Canada. Cambridge University Press, Cambridge
- Rivers, T., Martignole, J., Gower, C.F., and Davidson, A. 1989. New tectonic divisions of the Grenville Province, southeast Canadian Shield. *Tectonics* 8(1): 63-84.
- Rivet, P. 1924 Langues Amkricaines. In: Les Langues du Monde. A. Meillet and Marcel Cohen, eds. Collection Linguistique. Paris: La Sociktk de Linguistique de Pans, pp. 597-712.
- Rogers, J.S., Yarborough, M.R., and Pendleton, C.L. 2009. An Anangula period core-and-blade site on Amaknak Island, eastern Aleutians. *Alaska Journal of Anthropology* 7(1): 153-165.
- Rosenblat, A. 1945 La poblacion indigena de America desde 1492 hasta la actualidad. Buenos Aires: Institucion Cultural EspaAola.
- Rougeulle, A. Medieval Qalhāt, historical vs archaeological data. *Arabian Humanities* 9. URL : <http://journals.openedition.org/cy/3442> ; DOI : 10.4000/cy.3442.
- Ruddiman, W.F. 2003. The anthropogenic greenhouse era began thousands of years ago. *Climatic Change* 61: 261-293.
- Ruddiman, W.F. 2005. How did humans first alter global climate? *Scientific American* 292(3): 46-53.
- Ruddiman, W.F. 2008. Earth's climate: past and future. Second Edition. New York: W.H. Freeman.
- Ruddiman, W., and Ellis, E.C. 2009. Effect of per-capita land use changes on Holocene forest clearance and CO2 emissions. *Quaternary Science Reviews* 28: 3011-3015.
- Ruddiman, W.F., Kutzbach, J.E., and Vavrus, S.J. 2011. Can natural or anthropogenic explanations of late-Holocene CO2 and CH4 increases be falsified? *Holocene* 21: 865–879.
- Ruddiman, W.F. 2014. Earth Transformed. W.H. Freeman and Company, New York, New York, USA.
- Russell, E.W.B. 1983. Indian-set fires in the forests of northeastern United States. *Ecology* 64:76-88.
- Sanderson, E.W., Jaiteh, M., Levy, M.A., Redford, K.H., Wannebo, A.V., and Woolmer, G. 2002. The human footprint and the last of the wild. *Bioscience* 52: 891-904.
- Sands, R.D., and Leimbach, M. 2003 Modeling agriculture and land use in an integrated assessment framework. *Climatic Change* 56:185-210.
- Sapper, C. 1924 Die Zahl Und Die Volksdichte der Indianischen Bevolkerung in Amerika Vor Der Conquista und in Der Gegenwart. Proc. Of the 21st Internat. Congr. of Americanists Part 1. The Hague.
- Saucier, J.-P., J.-F. Bergeron, P. Grodin, and A. Robitaille. 1998. Les régions écologiques du Quebec méridional (3e version): un des éléments du système hiérarchique de classification écologique du territoire mis au point par le ministère des Ressources naturelle du Quebec. Supplément - L'Aubelle, février-mars 1998.
- Sauer, C.O. 1956. The agency of Man on the earth. In W. L. Thomas, Jr. (ed.), Man's Role in Changing the Face of the Earth, Vol. 1, pp. 49-69. The University of Chicago Press.
- Savelle, J.M., and Dyke, A.S. 2014. Paleoeskimo occupation history of Foxe Basin, Arctic Canada: Implications for the Core Area Model and Dorset origins. *American Antiquity* 79: 249-276.

- Savoie, L., and Richard, P.J.H. 1979. Paléophytogéographie de l'épisode de Saint-Narcisse dans la région de Sainte-Agathe, Quebec. *Géographie physique et Quaternaire* 33: 175-188.
- Scarry, C.M. 2003. Patterns of wild plant utilization in the Prehistoric eastern Woodlands. Pages 50-1047 in P.E. Minnis, editor. *People and Plants in Ancient Eastern North America*. Smithsonian Institution, Washington, USA.
- Scharf, E.A. 2010. Archaeology, land use, pollen and restoration in the Yazoo Basin (Mississippi, USA). *Vegetation History and Archaeobotany* 19: 159-175.
- Schurr, T.G., and Sherry, S.T. 2004. Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: Evolutionary and demographic evidence. *American Journal of Human Biology* 16(4): 420 – 439.
- Scott, D.W. 1979. On optimal and data-based histograms. *Biometrika*, 66: 605–610.
- Scott, D.W. 1992. *Multivariate Density Estimation: Theory, Practice, and Visualization* (John Wiley, New York).
- Seppä, H., and Bennett, K.D. 2003. Quaternary pollen analysis: recent progress in palaeoecology and palaeogeography. *Progress in Physical Geography* 27(4): 548-579.
- Shennan, S., Edinborough, K., 2007. Prehistoric population history: from the Late Glacial to the Late Neolithic in Central and Northern Europe. *Journal of Archaeological Science* 34: 1339-1345.
- Shennan, S., Downey, S.S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., Manning, K., and Thomas, M.G., 2013. Regional population collapse followed initial agricultural booms in mid-Holocene Europe. *Nature Communications* 4: 2486.
- Shimatani, K. 2002. Point processes for fine-scale spatial genetics and molecular ecology. *Biometrical Journal* 44(3): 325–352.
- Shuman, B., Bartlein, P., Logar, N., Newby, P., and Webb III, T. 2002. Parallel climate and vegetation responses to the early Holocene collapse of the Laurentide Ice Sheet. *Quaternary Science Reviews* 21: 1793-1805.
- Shuman, B., Newby, P., Huang, Y., and Webb, III, T. 2004. Evidence for the close climatic control of New England vegetation history. *Ecology* 85(5): 1297-1310.
- Shuman, B.N., Newby, P., and Donnelly, J.P. 2009. Abrupt climate change as an important agent of ecological change in the Northeast US throughout the past 15,000 years. *Quaternary Science Reviews* 28: 1693–1709.
- Silver, T. 1990. *A new face on the countryside. Indians, colonists and slaves in South Atlantic forests, 1500-1800*. Cambridge University Press, Cambridge, England.
- Silverman, B.W. 1986. *Density estimation for statistics and data analysis*. Chapman & Hall, London.
- Simek, J.F., Blankenship, S.A., Cressler, A., Douglas, J.C., Wallace, A., Weinand, D., and Welborn, H. 2012. The Prehistoric cave art and archaeology of Dunbar Cave, Montgomery County, Tennessee. *Journal of Cave and Karst Studies* 74(1): 19-32.
- Singarayer, J., Vldes, P.J., Friedlingstein, P., Nelson, S., and Beerling, D.J. 2011. Late Holocene methane rise caused by orbitally controlled increase in tropical sources. *Nature*, 470 (7332): 82-85.
- Sioui, G.E. 1999. *Huron Wendat: The Heritage of the Circle*. Michigan State University Press, USA.
- Smith, B.D. 2006. Eastern North America as an independent center of plant domestication. *Proceedings of the National Academy of Sciences USA* 103:12223-12228.

- Smith, B.D. 2007. Niche construction and the behavioural context of plant and animal domestication. *Evolutionary Anthropology* 16:188-199.
- Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., Rice, C., Scholes, B., Sirotenko, O., Howden, M., McAllister, T., Pan, G., Romanenkov, V., Schneider, U., Towprayoon, S., Wattenbach, M., and Smith, J. 2008. Greenhouse gas mitigation in agriculture. *Philosophical Transactions of the Royal Society B* 363: 789-813.
- Snow, D. 2010. *Archaeology of Native North America* (Prentice Hall, Boston), chapter 3.
- Soepboer, W., Sugita, S., and Lotter, A.F. 2010. Regional vegetation-cover changes on the Swiss Plateau during the past two millennia: A pollen-based reconstruction using the REVEALS model. *Quaternary Science Reviews* 29: 472-483.
- Speller, C.F., Kemp, B.M., Wyatt S.D., Monroe, C., Lipe, W.D., Arndt, U.M., and Yang, D.Y. 2010. Ancient mitochondrial DNA analysis reveals complexity of indigenous North American turkey domestication. *Proceedings of the National Academy of Sciences USA* 107(7): 2807-2812.
- Steckle, R.H., and Rose, J.C. 2005. *The backbone of history: health and nutrition in the western hemisphere*. Cambridge University Press, Cambridge, UK.
- Steele, J., and Politis, G. 2009. AMS <sup>14</sup>C dating of early human occupation of southern South America. *Journal of Archaeological Science* 36: 419-429.
- Steele, J. 2010. Radiocarbon dates as data: quantitative strategies for estimating colonization front speeds and event densities. *Journal of Archaeological Science* 37: 2017-2030.
- Steward, J.H. 1945. The changing American Indian. In: *The Science of Man in the World Crisis*, Ralph Linton, ed. New York, pp. 282-305.
- Steward, J.H. 1949. The native population of South America. In: *The Comparative Ethnology of South American Indians*. Handbook of South American Indians, 5, B. A. E. Bull., Washington, 143: 655-668.
- Stewart, O. 2002. *Forgotten Fires. Native Americans and the transient Wilderness*. University of Oklahoma Press, Norman, Oklahoma, USA.
- Strandberg, G., Brandefelt, J., Kjellström, E., and Smith, B. 2011. High-resolution regional simulation of the last Glacial Maximum in Europe. *Tellus* 63A, 107-125.
- Strandberg, G., Kjellström, E., Poska, A., Wagner, S., Gaillard, M.-J., Trondman, A.-K., Mauri, A., Davis, B.A.S., Kaplan, J.O., Birks, H.J.B., Bjune, A.E., Fyfe, R., Giesecke, T., Kalnina, L., Kangur, M., Van Der Knaap, W.O., Kokfelt, U., Kunes, P., Latalowa, M., Marquer, L., Mazier, F., Nielsen, A.B., Smith, B., Seppä, H., Sugita, S., 2014. Regional climate model simulations for Europe at 6 and 0.2 k BP: sensitivity to changes in anthropogenic deforestation. *Climates of the Past* 10: 661-680.
- Stuiver, M., Reimer, P.J. 1993. Extended <sup>14</sup>C data base and revised CALIB 3.0 <sup>14</sup>C age calibration program. In: Stuiver, M., Long, A., and Kra, R.S., eds., *Calibration 1993*. *Radiocarbon* 35: 215-230.
- Stuiver, M., Reimer, P.J., Bard, E., Beck, J.W., Burr, G.S., Hughen, K.A., Kromer, B., McCormac, G., Van Der Plicht, J., Spurk, M. 1998. IntCal98 radiocarbon age calibration, 24,000-0 cal BP. *Radiocarbon* 40: 1041-1083.
- Sugita, S. 1993. A model of pollen source area for an entire lake surface. *Quaternary Research* 39: 239-244.
- Sugita, S. 1994. Pollen representation of vegetation in quaternary sediments: Theory and method in patchy vegetation. *Journal of Ecology* 82(4): 881-897.

- Sugita, S., Gaillard, M.-J., Broström, A. 1999. Landscape openness and pollen records: a simulation approach. *Holocene* 9(4): 409-421.
- Sugita, S. 2007a. Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. *Holocene* 17(2): 229-241.
- Sugita, S. 2007b. Theory of quantitative reconstruction of vegetation II: All you need is LOVE. *Holocene* 17(2): 243-257.
- Sugita, S., Parshall, T., Calcote, R., and Walker, K. 2010. Testing the Landscape Reconstruction Algorithm for spatially explicit reconstruction of vegetation in northern Michigan and Wisconsin. *Quaternary Research* 74: 289-300.
- Surovell, T.A., and Brantingham, P.J. 2007. A note on the use of temporal frequency distributions in studies of prehistoric demography. *Journal of Archaeological Science* 34: 1868-1877.
- Surovell, T.A., Finely, J., Smith, G.M., Brantingham, P.J., and Kelly, R.L. 2009. Correcting temporal frequency distributions for taphonomic bias. *Journal of Archaeological Science* 36: 656-664.
- Surovell, T.A., Pelton, S.R., Anderson-Sprecher, R., and Myers, A.D. 2016. Test of Martin's overkill hypothesis using radiocarbon dates on extinct megafauna. *Proceedings of the National Academy of Sciences USA* 113(4): 886-891.
- Tallavaara, M., and Seppä, H. 2011. Did the mid-Holocene environmental changes cause the boom and bust of hunter-gatherer population size in eastern Fennoscandia? *Holocene* 22(2): 215-225.
- Tallavaara, M., Luoto, M., Korhonen, N., Järvinen, H., and Seppä, H., 2015. Human population dynamics in Europe over the Last Glacial Maximum. *Proceedings of the National Academy of Sciences USA* 112: 8232-8237.
- Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D.G., Mulligan, C.J., Bravi, C.M., Rickards, O., Martinez-Labarga, C., Khusnutdinova, E.K., Fedorova, S.A., Golubenko, M.A.V., Stepanov, V.A., Gubina, M.A., Zhadanov, S.I., Ossipova, L.P., Damba, L., Voevoda, M.I., Dipierri, J.E., Villems, R., and Malhi, R.S. 2007. Beringian Standstill and spread of Native American Founders. *PLoS ONE* 9: e829.
- Taylor, R. W., and Spurlock, B.D. 1982. The changing Ohio River naiad fauna: a comparison of early Indian middens with today. *The Nautilus* 96: 49-51.
- Thomas, E.R., Wolff, E.W., Mulvaney, R., Steffensen, J.P., Johnsen, S.J., Arrowsmith, C., White, J.W.C., Vaughn, B., and Popp, T. 2007. The 8.2 ka event from Greenland ice cores. *Quaternary Science Reviews* 26: 70-81.
- Thompson, R. S., Anderson, K.H., and Bartlein, P.J. 1999. *Atlas of Relations between Climatic Parameters and Distributions of Important Trees and Shrubs in North America*. US Geological Survey Professional Paper 1650A-B. Denver: USGS.
- Thornton, R., and Marsh-Thornton, J. 1981. Estimating prehistoric American Indian population size for United States: Implications of the nineteenth century population decline and Nadir. *American Journal of Physical Anthropology* 55: 47-53.
- Thornton, R. 1987. *American Indian Holocaust and Survival* (University of Oklahoma Press, Oklahoma), Pp 15-41.
- Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., Schindler, D., Schlesinger, W.H., Simberloff, D., and Swackhamer, D. 2001. Forecasting agriculturally driven global environmental change. *Science* 292: 281-284.

- Timpson, A., Colledge, S., Crema, E., Edinborough, K., Kerig, T., Manning, K., Thomas, M.G., and Shennan, S. 2014. Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. *Journal of Archaeological Science* 52: 549-557.
- Timpson, A., Manning, K., Shennan, S. 2015. Inferential mistakes in population proxies: A response to Torfing's "Neolithic population and summed probability distribution of  $^{14}\text{C}$  -dates". *Journal of Archaeological Science* 63: 199-202.
- Torfing, T. 2015. Neolithic population and summed probability distribution of  $^{14}\text{C}$  -dates. *Journal of Archaeological Science* 63: 193-198.
- Törnqvist, T.E., and Hijma, M.P. 2012. Links between early Holocene ice-sheet decay, sea-level rise and abrupt climate change. *Nature Geoscience* 5: 601–606.
- Trondman, A.-K., Gaillard, M.J., Mazier, F., Sugita, S., Fyfe, R., Nielsen, A.B., Twiddle, C., Barratt, P., Birks, H.J., Bjune, A.E., Björkman, L., Broström, A., Caseldine, C., David, R., Dodson, J., Dörfler, W., Fischer, E., van Geel, B., Giesecke, T., Hultberg, T., Kalnina, L., Kangur, M., van der Knaap, P., Koff, T., Kuneš, P., Lagerås, P., Latalowa, M., Lechterbeck, J., Leroyer, C., Leydet, M., Lindbladh, M., Marquer, L., Mitchell, F.J., Odgaard, B.V., Peglar, S.M., Persson, T., Poska, A., Rösch, M., Seppä, H., Veski, S., and Wick, L. 2015. Pollen-based quantitative reconstructions of Holocene regional vegetation cover (plant-functional types and land-cover types) in Europe suitable for climate modelling. *Global Change Biology* 21: 676-697.
- Trondman, A.-K., Gaillard, M.-J., Sugita, S., Björkman, L., Greisman, A., Hultberg, T., Lagerås, P., Lindbladh, M., and Mazier, F. 2016. Are pollen records from small sites appropriate for REVEALS model-based quantitative reconstructions of past regional vegetation? An empirical test in southern Sweden. *Vegetation History and Archaeobotany* 25: 131-151.
- Tull, A.J.T., Burr, G.W., and Hodgins, G.W.L., 2013. Radiocarbon dating, reservoir effects, and calibration. *Quaternary International* 299: 64-71.
- Tulowiecki, S.J., and Larsen, C.P.S. 2015. Native American impact on past forest composition inferred from species distribution models, Chautauqua County, New York. *Ecological Monographs* 85: 557-581.
- Turney, C.S.M., and Hobbs, D. 2006. ENSO influence on Holocene Aboriginal populations in Queensland, Australia. *Journal of Archaeological Science* 33: 1744-1748.
- Turney, C.S.M., Haberle, S., Fink, D., Kershaw, A.P., Barbetti, M., Barrows, T.T., Black, M., Cohen, T.J., Corrège, T., Hesse, P.P., Hua, Q., Johnston, R., Morgan, V., Moss, P., Nanson, G., van Ommen, T., Rule, S., Williams, N.J., Zhao, J.-X., D'Costa, S., Feng, Y.-X., Gagan, M., Mooney, S., and Xia, Q. 2006. Integration of ice-core, marine and terrestrial records for the Australian Last Glacial Maximum and Termination: a contribution from the OZ INTIMATE group. *Journal of Quaternary Science* 21(7): 751-761.
- Ubelaker, D. H. 1976a. The sources and methodology for Mooney's estimates of North American Indian populations. Chapter 8 in *Native Population of the Americas in 1492*. William M. Denevan, ed. University of Wisconsin Press, Madison,
- Ubelaker, D.H. 1976b. Prehistoric New World population size: Historical review and current appraisal of North American estimates. *American Journal of Physical Anthropology* 45(3): 661-665.
- Ubelaker, D.H. 1992. Patterns of demographic change in the Americas. *Human Biology* 64: 361-379.

- Ubelaker, D.H. 2006. Population size, contact to nadir. In: Ubelaker, D.H., ed., *Environment, Origins, and Population. Handbook of North American Indians*, vol. 3. Smithsonian Institution, Washington, pp. 694-701.
- Vale, T.R. 2002. The pre-European landscape of the United States: pristine or humanized? In T. R. Vale (Ed.), *Fire, Native Peoples and the Natural Landscape*, pp. 1–40. Washington DC: Island Press.
- Vermeersch, P.M. 2005. European population changes during Marine Isotope Stages 2 and 3. *Quaternary International* 137: 77-85.
- Vermeersch, P.M. 2011. The human occupation of the Benelux during the Younger Dryas. *Quaternary International* 242: 267-276.
- Vermeersch, P.M. 2015. Radiocarbon Palaeolithic Europe Database, Version 18. Available at: <http://ees.kuleuven.be/geography/projects/14c-palaeolithic/index.html>.
- Viau, A.E., Gajewski, K., Sawada, M.C., and Fines, P. 2006. Millennial-scale temperature variations in North America during the Holocene. *Journal of Geophysical Research* 111(D09102).
- Vitousek, P.M., Mooney, H.A., Lubchenco, J., and Melilo, J.M. 1997. Human domination of Earth's ecosystems. *Science* 277: 494-499.
- Vu, V.Q. 2011. Ggbiplot: A ggplot2 based biplot. R package version 0.55.
- Wackernagel, H. 2003. *Multivariate Geostatistics* (3rd ed.). Berlin: Springer.
- Waguespack, N.M. 2007. Why were still arguing about the Pleistocene occupation of the Americas. *Evolutionary Anthropology* 16: 63-74.
- Wand, M.P. and Jones, M.C. 1995. *Kernel Smoothing*. London: Chapman & Hall.
- Walker, M.J.C., Berkelhammer, M., Bjorck, S., Cwynar, L.C., Fisher, D.A., Long, A.J., Lowe, J.J., Newnham, R.M., Rasmussen, S.O., and Weiss, H., 2012. Formal subdivision of the Holocene Series/Epoch: a discussion paper by a Working Group of INTIMATE (Integration of ice-core, marine and terrestrial records) and the Subcommittee on Quaternary Stratigraphy (International Commission on Stratigraphy). *Journal of Quaternary Science* 27, 649e659.
- Wang, Y., and Herzschuh, U. 2011. Reassessment of Holocene vegetation change on the upper Tibetan Plateau using the pollen-based REVEALS model. *Review of Palaeobotany and Palynology* 168(1): 31–40.
- Wang, G., B. Knapp, S. Clark, and B. Mudder. 2013. The Silvics of *Castanea dentata* (Marsh.) Borkh., American Chestnut, Fagaceae (Beech Family). USDA Forest Service. General Technical Report SRS-173. Southern Research Station. Asheville, North Carolina.
- Wang, C., Lu, H., Zhang, J., Gu, Z., and He, K. 2014. Prehistoric demographic fluctuations in China inferred from radiocarbon data and their linkage with climate change over the past 50,000 years. *Quaternary Science Reviews* 98: 45-59.
- Warrick, G. 2008. *A population history of the Huron-Petun, A.D. 500-1600*. Cambridge University Press, Cambridge, England.
- Waters, M.R., and Stafford Jr., T.W. 2007. Redefining the age of Clovis: implications for the peopling of the Americas. *Science* 315: 1122-1126.
- Webb, T., III., Yoryos Yeracaris, G., and Richard, P. 1978. Mapped patterns in sediment samples of modern pollen from southeastern Canada and northeastern United States. *Géographie physique et Quaternaire* 32: 163-176.
- Webb, T. III., 1988. Eastern North America. Pages 155-192 in B. Huntley, and T. Webb III, editors. *Vegetation history*. Kluwer, Dordrecht, Netherlands.

- Webb, T. III., P Bartlein, S. Harrison, and K.H. Anderson. 1993. Vegetation, lake levels and climate in eastern North America for the past 18000 years. Pages 415-467 in Wright, H.E. Jr., J. Kutzbach, T. Webb III, W.F. Ruddiman, F.A Street-Perrott, and P.J. Bartlein, editors. *Global Climates since the last Glacial Maximum*. University of Minnesota Press, Minneapolis, Minnesota, USA.
- Webb, T. III., Anderson, K.H., Bartlein, P.J., and Webb, R.S. 1998. Late Quaternary climate change in eastern North America: a comparison of pollen-derived estimates with climate model results. *Quaternary Science Reviews* 17(6-7): 587-606.
- Wendland, W.M. and R.A. Bryson. 1974. Dating climatic episodes of the Holocene. *Quaternary Research* 4: 9-24.
- Weninger, B., Jöris, O., Danzeglocke, U., 2013. CalPal: Cologne Radiocarbon CALibration & PALaeoclimate Research Package. Europe Database.
- West, D., Crawford, M., and Savinetsky, A.B. 2007. Genetics, prehistory and the colonisation of the Aleutian Islands. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 98: 47-57.
- Western, D. 2001. Human modified ecosystems and future evolution. *PNAS* 98: 5458-5465.
- Whitehead, D.R. 1979. Late-glacial and postglacial vegetational history of the Berkshires, western Massachusetts. *Quaternary Research* 12: 333-342.
- Whitehead, D.R., and Sheehan, M.C. 1985. Holocene vegetational changes in the Tombigbee River valley, eastern Mississippi. *American Midland Naturalist* 113: 122-137.
- Whitmore, J., Gajewski, K., Sawada, M., Williams, J.W., Minckley, T., Shuman, B., Bartlein, P.J., Webb III, T., Vial, A.E., Shafer, S., Anderson, P., and Brubaker, L. 2005. North American Modern Pollen Database for Multi-scale Paleoecological and Paleoclimatic Applications. *Quaternary Science Reviews* 24: 1828-1848.
- Whitney, G.G. 1994. *From coastal wilderness to fruited plain*. Cambridge University Press, Cambridge, England.
- Wickham, H. 2009. *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York.
- Wickham, H., and Chang, W. 2016. *Devtools: Tools to make developing R packages easier*. R package version 1.12.0.
- Wilcox, W. F. 1931 Increase in the population of the earth and of the continents since 1650. In: *International Migrations, Volume 2. Interpretations*, W. F. Wilcox, ed. National Bureau of Economic Research, Inc., New York, pp. 33- 82.
- Williams, M. 1989a. Deforestation: past and present. *Progress in Human Geography* 13: 176-208.
- Williams, M. 1989b. *Americans and their Forests: A Historical Geography*. New York: Cambridge University Press.
- Williams, J.W., Shuman, B.N., Webb III, T., Bartlein, P.J., and Leduc, P.L. 2004. Late-Quaternary vegetation dynamics in North America: scaling from taxa to biomes. *Ecological Monographs* 74: 309-324.
- Williams, J.W., and Shuman, B.N. 2008. Obtaining accurate and precise environmental reconstructions from the modern analog technique and North American surface pollen dataset. *Quaternary Science Reviews* 27: 669-687.
- Williams, A.N., Smith, M.A., Turney, C.S.M., and Cupper, M.L. 2008a. AustArch1: a database of <sup>14</sup>C and luminescence ages from archaeological sites in the Australian arid zone. *Australian Archaeology* 66: 99.

- Williams, A., Santoro, C.M., Smith, M.A., and Latorre, C. 2008b. The impact of ENSO in the Atacama Desert and Australian arid zone: exploratory time-series analysis of archaeological records. *Chungara* 40, 245-259.
- Williams, A.N., Ulm, S., Goodwin, I.D., and Smith, M. 2010. Hunter-gatherer response to late Holocene climatic variability in northern and central Australia. *Journal of Quaternary Science* 25: 831-838.
- Williams A.N. 2012. The use of summed radiocarbon probability distributions in archaeology: a review of methods. *Journal of Archaeological Science* 39: 578-289.
- Williams, A.N. 2013. A new population curve for prehistoric Australia. *Proceedings of the Royal Society B Biological Sciences* 280(1761): 20130486.
- Williams, A.N., and Smith, M. 2013. AustArch3: a database of  $^{14}\text{C}$  and luminescence ages from archaeological sites in southern Australia. *Australian Archaeology* 76: 102.
- Williams, A.N., Ulm, S., Cook, A.R., Langley, M.C., and Collard, M. 2013. Human refugia in Australia during the last glacial maximum and terminal Pleistocene: a geospatial analysis of the 25-12ka Australian archaeological record. *Journal of Archaeological Science* 40: 4612-4625.
- Williams, A.N., Ulm, S., Smith, M., and Reid, J. 2014. AustArch: a database of  $^{14}\text{C}$  and non- $^{14}\text{C}$  ages from archaeological sites in Australia – composition, compilation and review (Data Paper). *Internet Archaeology* 36. Available at <http://dx.doi.org/10.11141/ia.36.6>.
- Williams, A.N., Veth, P., Steffen, W., Ulm, S., Turney, C.S.M., Reeves, J.M., Phipps, S.J., and Smith, M. 2015a. A continental narrative: human settlement patterns and Australian climate change over the last 35,000 years. *Quaternary Science Reviews* 123: 91-112.
- Williams, A.N., Mooney, S.D., Sisson, S.A., and Marlon, J. 2015b. Exploring the relationship between Aboriginal population indices and fire in Australia over the last 20,000 years. *Palaeogeography, Palaeoclimatology, Paleoecology* 423: 49-57.
- Williams, A.N., Ulm, S., Turney, C.S., Rohde, D., and White, G. 2015c. Holocene demographic changes and the emergence of complex societies in Prehistoric Australia. *PLoS ONE* 10, e0128661.
- Wolff, E.W. 2014. Ice sheets and the Anthropocene. In: Waters, C.N., Zalasiewicz, J., Williams, M., Ellis, M.A., Snelling, A. (Eds.), *A Stratigraphical Basis for the Anthropocene*, Geological Society, London, Special Publications 395, pp. 255e263.
- Woods, R. 2015. From revolution to convention: the past, present and future of radiocarbon dating. *Journal of Archaeological Science* 56: 61-72.
- Wright, Jr., H.E., and Bartlein, P.J. 1993. Reflections on COHMAP. *Holocene* 3(1): 89-92.
- Wycoff, W.W. 1991. Black Walnut on Iroquoian landscapes. *Northeast Indian Quarterly* 8, 4-17.
- Xu, Q., Cao, X., Tian, F., Zhang, S., Li, Y., Li, M., Li, J., Liu, Y., and Liang, J. 2014. Relative pollen productivities of typical steppe species in northern China and their potential in past vegetation reconstruction. *Science China: Earth Sciences* 57(6): 1254-1266.
- Xu, Q., Zhang, S., Gaillard, M.-J., Li, M., Cao, X., Tian, F., and Li, F. 2016. Studies of modern pollen assemblages for pollen dispersal- deposition- preservation process understanding and for pollen-based reconstructions of past vegetation, climate, and human impact: A review based on case studies in China. *Quaternary Science Reviews* 149: 151-166.
- Zahid, H.J., Robinson, E., and Kelly, R.L. 2016. Agriculture, population growth, and statistical analysis of the radiocarbon record. *Proceedings of the National Academy of Sciences USA* 113(4): 931-935.

- Zalasiewicz, J., Waters, C.N., Wolfe, A., Barnosky, A.D., Cearreta, A., Edgeworth, M., Ellis, E.C., Fairchild, I.J., Gradstein, F.M., Grinevald, J., Haff, P., Head, M.J., 12, Ivar do Sul, J.A., Jeandel, C., Leinfelder, R., McNeill, J.R., Oreskes, N., Poirier, C., Revkin, A., de Richter, D., Steffen, W., Summerhayes, C., Syvitski, J.P.M., Vidas, D., Wapreisch, M., Wing, S., and Williams, M. 2017. Making the case for a formal Anthropocene Epoch: an analysis of ongoing critiques. *Newsletters on Stratigraphy* 50(2): 205-226.
- Zelinsky, W. 1973. *The Cultural Geography of the United States*. First Edition. Prentice Hall.
- Zettler, F.W. 2015. *The biohistory of Alachua County, Florida*. Pineapple Press Inc: Sarasota, Florida.

## Appendix 1 (Supplementary Information for Chapter 3)

The methodology for chapter 3 comprised of 5 major steps which include the calculation of the 1) Radiocarbon Frequency Population Estimates (RFPEs), 2) taphonomic correction, 3) sampling intensity, 4) edge correction, and 5) a comparison with ethnohistorical data. RFPEs are computed as a ratio between preliminary RFPEs and sampling intensity. Similar techniques based on kernel density estimation are used in steps 1 and 3 but they differ in two important ways. In step 1 (RFPE calculation), multiple dates at a single site within a given time interval are considered individually, and only cultural dates are used. In step 3 (sampling intensity calculation), only geographically distinct sites are considered such that multiple dates at the same location are considered as one, and all cultural and paleoenvironmental dates of all ages in the CARD are used.

### *Kernel Density Estimation*

The kernel density technique is one of the most common non-parametric methods used for quantifying the spatial behaviour of a point pattern of sites (Silverman, 1986). Surrounding each site is an individual kernel function which can be interpreted as a “hat” with a specific radius  $r$  (or bandwidth) which sits on top of the site. The considered site has an influence on the intensities at all locations within a circular area with radius  $r$  centered on the site. In our study, the radius was defined as 600 km (**Fig. A1.5**), a value close to the optimal radius (595 km) calculated using Scott’s rule (Scott, 1992). Choosing a higher or lower radius would lead to an over- or under-smoothing of the maps (**Fig. A1.1**). Each site is considered individually using the following Epanechnikov kernel function (Scott, 1992):

$$K(u) = \frac{2}{\pi} (1 - u^T u) I_{\{u^T u \leq 1\}} \quad [13]$$

where  $u=(u1, u2)$  is a point on the map. This kernel is transformed to have the radius  $r$  and additionally scaled by two values: 1) the weight  $w$  (mentioned in the Methods section) for computing the RFPE (not for sampling intensity), and 2) the edge correction  $e$  (see below). For each possible location on the map, the values of the scaled kernels which cover this location are summed to produce the intensity function  $\lambda(\cdot)$ , i.e.,

$$\lambda(u) = \frac{1}{r^2} \sum_{i=1}^n w_i e_i K\left(\frac{x_i - u}{r}\right) \quad [14]$$

where  $u=(u1, u2)$  is again a point on the map. For the computation of the sampling intensity,  $x_1, \dots, x_n$  denote the locations of all geographically distinct CARD sites and  $w_i$  is always 1. When computing the RFPEs,  $x_1, \dots, x_n$  are the sites with dates in the corresponding interval ( $\pm 400$  years), where multiple dates from one site are summed individually. Possible modifications of kernel density estimation are discussed below.

## *Taphonomic Bias*

The taphonomic bias inherent in the CARD, due to long term loss of samples through time in the fossil record, is addressed by computing the mean age of each interval and using:

$$n_t = 5.726442 \times 10^6 (t + 2176.4)^{-1.3925309} \quad [15]$$

where  $n_t$  is an indicator for the number of  $^{14}\text{C}$  dates in the CARD surviving from time  $t$ , or the mean age of the interval in question. This taphonomic loss rate was developed based on ice core and geologic volcanic activity records and operates on a removal-through-time basis (Surovell et al., 2009). We assume that the archaeological remains used to derive the CARD data share similar degradation characteristics as these geological contexts (Peros et al., 2010; Williams et al., 2010; Tallavaara and Seppä, 2011; Williams, 2012) and thus, that  $^{14}\text{C}$  dates are subject to the same rate of taphonomic loss. Based on this, the resulting value will be higher for more recent intervals and lower for older ones. The whole map is globally divided by the value computed from the taphonomic formula (Surovell et al., 2009); values of RFPEs from older intervals are too low due to taphonomic loss and are divided by a lower value than the earlier intervals (Peros et al., 2010). We are aware that the application of this correction may be suboptimal in some areas where taphonomic loss is probably lower (e.g., the Arctic where permafrost effectively preserves archaeological materials) or in later time intervals (e.g., between the present and 0.5 ka when archaeological interest declines and historical information is used instead of  $^{14}\text{C}$  dates for dating), and may lead to decreased RFPEs. This could potentially be addressed via regional analyses but this was beyond the scope of this study.

## *Edge Correction*

Edge corrections are necessary since RFPEs would always decrease towards the coast (since dates near the coast have less of an influence). The circular area around a site with radius  $r$  can potentially intersect with an area where no population is possible, for example water or ice. If this is the case, then this location would have a lower capacity for human population and thus a lower influence on the RFPE function. In other words, a kernel has a volume within which the population can occur, if half of the circular area around the site occurs over water then the volume available for population is reduced by half. Ideally, every individual kernel will have an equal influence on the intensity. Therefore, the numerical integral of the kernel function over land is calculated, and the value of the kernel at each influenced location is divided by this numerical integral (denoted as  $e$  above) so that each corrected kernel has an equal influence on intensity (i.e., kernels that are at least partially over water or ice are divided such that they have the same influence as those completely over land). If half of the kernel is over water, the volume over land is only half of the other kernels over land, so this is scaled by two to allow the part of the kernel over land to have the same influence as any other kernel over land. When computing RFPEs, water and ice are taken into account for the edge correction, whereas for the sampling intensity only water is considered.

### *Estimating Population Density at Contact: Comparison to Results from Ubelaker (1992)*

The same techniques used to produce the sampling intensity and RFPE maps are used to produce an estimate of population size following European contact (**Fig. 3.3a**). This is compared to ethnohistorical documentation of North American tribal area population estimates (**Fig. 3.3b**) (Ubelaker, 1976b; 1992), which is largely based on a compilation of population estimates made by multiple authors and early explorers (Mooney, 1928). The maps (**Fig. 3.3**) have been rescaled and generalized to correspond to early representations of tribal areas (Ubelaker, 1992). In this case, data for Canada are a very generalized, and the estimates for “Greenland” are used in this study for northern Canada.

### *Shape and Bandwidth of the Kernel Density Estimator*

A variety of kernel functions exist for use in spatio-temporal analyses. The Epanechnikov kernel, which is used in this analysis, and the Gaussian kernel are two methods that are frequently used. The latter, however, has an unbounded support, i.e., dates would have an influence on the whole continent. The Epanechnikov kernel compliments our approach of assuming each date influences an area within a certain radius of its location. It is generally accepted that the choice of the kernel function has only a minor effect on the resulting estimate and its statistical properties. Another reason for choosing the Epanechnikov kernel is that this kernel minimizes the approximate mean integrated square error (AMISE) of the kernel density estimator, which is a desirable statistical property (Silverman, 1986).

The choice of the bandwidth is a more complex consideration. One option is to use an adaptive bandwidth, i.e., a bandwidth which is not globally-fixed but varies across space. There are two ways of achieving this. In the nearest-neighbor approach, the bandwidth is chosen based on the location  $u$  at which the intensity is computed, i.e., we would use  $r(u)$  instead of  $r$  in formula [2]. In the variable kernel approach, the bandwidth depends on the current data site, which implies that  $r(x_i)$  is written instead of  $r$  and the fraction is shifted inside the sum in formula [2]. Both approaches are quite similar with one conceptual difference. In the variable kernel approach all sites can have a different radius, but these radii remain fixed for all locations  $u$  at which the intensity is estimated. In the nearest-neighbor approach, the radius assigned to a certain data site can vary, if different locations  $u_1$  and  $u_2$  are considered. These approaches to bandwidth-choice contradict our fundamental logic that each date should have the same influence on a surrounding area with a spatially and temporally fixed radius. The bandwidths for all data sites could be determined adaptively based on all sampling sites and these bandwidths could be used for all estimations. But then, these bandwidths are not adapted to the total number of dates in different time intervals. For these reasons we do not use an adaptive bandwidth approach. To illustrate, RFPE maps created based on the nearest-neighbour and variable kernel approaches are provided (**Fig. A1.6**).

A second option is to algorithmically determine a globally-fixed bandwidth, in which case likelihood cross validation is most frequently used. However, this method is sensitive to outliers

and to multiple data points, and is inefficient when large datasets are considered making it a poor choice in this study also.

Instead, we opt to set the bandwidth to a fixed value of 600 km following Scott's rule (Scott, 1992). Our goal was to use the same bandwidth for the sampling intensity function and the RFPEs for all time intervals. This is necessary to maintain comparability through time, to avoid odd "jumps" between different time intervals and to avoid numerical anomalies when computing RFPEs as ratios. Thus, a bandwidth is computed for each time interval separately using Scott's rule (see **Fig. A1.5** for a histogram of obtained values). A bandwidth close to the mean value (595 km) seems to be an adequate choice for all time intervals and the sampling intensity.

### *Sensitivity Analysis*

A sensitivity test was performed to determine the effects of varying the kernel function radius. At first, RFPEs computed with a globally fixed radius greater or less than 600 km (**Fig. A1.1c**) are considered. A radius of 500 km or smaller is not large enough to capture the overall pattern since often too few sites occur within such a small area (**Fig. A1.1a&b**). When a radius of 800 km or more is used, the overall signal remains but the maps are too smoothed to discern regional patterns (**Fig. A1.1d-f**). It is possible that a radius of 1,000 to 1,250 km would be ideal for a global scale study, but we recommend *ca.* 600 km for a continental scale analysis. Likewise, a radius smaller than 500 km would be more suitable for a regional study.

Furthermore, we study the effect of choosing an adaptive bandwidth according to the generalized nearest-neighbor and variable kernel approaches. The concept is similar in both cases; the radius  $r(u)$  for each location  $u$  in the generalized nearest-neighbor approach as well as the radius  $r(x_i)$  for each site  $x_i$  in the variable kernel approach are chosen to be inversely proportional to the local density of all sampling sites. These radii are used for computing the sampling intensity and the RFPEs for all time intervals. Note, however, that these radii are not inversely proportional to the site density of the different time intervals anymore, which contradicts the purpose of choosing an adaptive bandwidth. Results show similar population estimates for both approaches (**Fig. A1.6**). The maps reveal largely-smoothed patterns in regions with a low site density (central Canada) but under-smoothing occurs in more frequently sampled regions, e.g., the southeast or Wyoming. In general, however, both approaches lead to the same population events, although the spatial scale and size of single events can vary slightly. Furthermore, both sequences of maps show the same patterns as in the fixed bandwidth approach, which reinforces the results obtained in this paper.

### *Quality of the CARD: Removal of "Anomalous" Data*

The CARD contains over 35,000 archaeological dates collected by thousands of submitting authors over many years. Although carefully checked, radiocarbon dates are subject to well-known sources of error, for example, a date could be contaminated with younger organic matter or be older due to the incorporation of older carbon. The database could potentially include

incorrect dates and 1,419 dates (of the 29,609 which we used to create the RFPE maps) are currently flagged as “anomalous” in the CARD. A rigorous, non-subjective method of removing potentially erroneous data from the CARD based on descriptive, geographical, cultural and other information needs to be developed. Until this can be done, however, we suggest research based on the CARD include all available data, “anomalous” or otherwise (Hellerstein, 2008). This avoids selectively removing dates based on preconceived notions of paleodemographic trends and eliminates the risk of over-cleaning the database. In this study, all available CARD data were included for the purpose of a) mapping the raw data at a continental scale, a task that has not yet been done, and b) identifying aspects of the data which may lead to new insight into the human history of North America. Anomalous data in a local context may not be anomalous in a regional or continental context.

To determine to what extent the dates described as “anomalous” in the CARD influence the resulting RFPE patterns, we recalculated RFPEs between 0.5 and 13 ka excluding the flagged dates. These dates (n=1,419 or 5% of the original data points) originated from sites spanning the entire continent (**Fig. A1.7**) resembling a random sample of the data in terms of spatial distribution, but being a “worst case scenario”, as they have a greater potential of influencing the resulting patterns than a true random sample. It is of interest to note that the date appearing below the glacial ice in the central Arctic was not described as anomalous and is therefore retained in this analysis. The majority of the sites in **Figure A1.7** were associated with multiple “anomalous” dates most often contributed by a single author per site. We would emphasize that these data may appear “anomalous” in the cultural context of the original study but that a) they may still provide information about population numbers, and b) they may not be “anomalous” when studied in the context of other data at regional to continental scales.

The recalculation of RFPEs excluding “anomalous” dates resulted in a series of maps nearly identical to the original maps shown in **Figure A1.2**. We provide a comparison of the recalculated RFPE maps with the original maps in **Figure A1.3**. The overall patterns remain the same throughout the entire study period, at continental to regional scales. The timing of settlement of the High Arctic, Banks and Victoria Islands, the Aleutian Islands, the western continental coastline and all other major centers of population visible in the maps remain the same. The single exception to this is Newfoundland where the RFPEs first begin to increase from 0 to 0.17 around 11.8 ka in the original maps, and 10.5 ka in the recalculated maps. Differences are never greater than one order of magnitude, are greater in earlier time intervals, and mostly occur along the coasts. Very small differences are observed (**Fig. A1.3**) in northern Alaska (12.5-13 ka), Newfoundland (11-11.5 ka), the Kivalliq region (7-7.5 ka) and Louisiana (7-7.5 and 4.5-5 ka). In each of these cases, the smoothing indicates a slightly earlier occupation in the series of maps based on all available data (i.e., when “anomalies” are included). After 2 ka, visual differences no longer exist.

The “anomalous” dates in question were further investigated by choosing 3 sites and comparing the “anomalous” dates from those sites with all non-anomalous dates within a 25 km radius. **Figure A1.8a** suggests removing the “anomalous” dates could potentially show no human

presence during the mid-Holocene, which may be the case, or may be simply due to lack of samples. The time of the first arrival in the region would be the same. Removing these dates could also indicate people arrived later in a particular region (**Fig. A1.8b**), although the impact of their removal on the maps is small (**Fig. A1.3**). In many cases, the removal of “anomalous” dates would produce no difference in our interpretation of population numbers through time (**Fig. A1.8c**). Overall, the number of questionable dates is low compared to accepted dates, their removal has little effect on the general patterns shown in the RFPE maps, and this favours their inclusion in spatio-temporal analyses such as this one. The high degree of similarity between RFPE maps based on all available data and non-anomalous data implies the CARD in its current form is both spatially and temporally representative of paleodemographic trends across North America.

### *Completeness of the CARD: Randomly Removing 50% of the dates*

Whereas the section above identifies a worst-case scenario in which the RFPE maps are produced following the removal of dates which have been flagged as “anomalous”, here the database size is reduced to test whether or not the CARD is spatially representative and has a sufficiently large size for all time intervals. To do this we performed a Bernoulli experiment with a success probability of one half for each date and removed all dates showing success, which resulted in 16,894 (around 50%) of the original dates. The RFPEs are then recomputed based on the reduced dataset. **Figure A1.4** indicates that, despite the smaller number of dates, the spatial patterns remain largely the same, even during the earlier intervals when there are comparatively few dates. Some regional differences are visible, for example a population center is missing in the ice-free corridor between 11 and 11.5 ka, and Florida shows almost no human presence between 9.5 and 10 ka.

Based on our multiple sensitivity analyses, we propose that the CARD is sufficiently complete for continental-scale spatio-temporal paleodemographic analyses, that individual dates (and even large numbers of dates) do not have a major influence on RFPEs, and non-systematic errors (e.g., “anomalous” dates) have only a minor influence as well.

### *Pooled versus Summing Methods*

Recent studies have investigated the use of pooled mean dates for individual site phases (when multiple dates are found at a single site) as opposed to the basic summing method (Buchanan et al., 2008; Lesure, 2008; Steele, 2010). The question is: are multiple dates due to 1) the original researcher simply attempting to obtain replicates, in which case it would bias the population estimates, or 2) the relative size of the site, and the duration and amount of human activity there, in which case all dates should be considered individually. Although we judge this issue in need of further attention, it was not the purpose of this study to test this, and we refer readers to a prior investigation of this (Shennan and Edinborough, 2007). Shennan and Edinborough show small-scale differences in population fluctuations and a shuffling of the peak population period when summed probabilities and pooled means are compared but conclude that the use of a pooling

method in lieu of a summing method does not lead to a more straightforward interpretation of  $^{14}\text{C}$  dates.

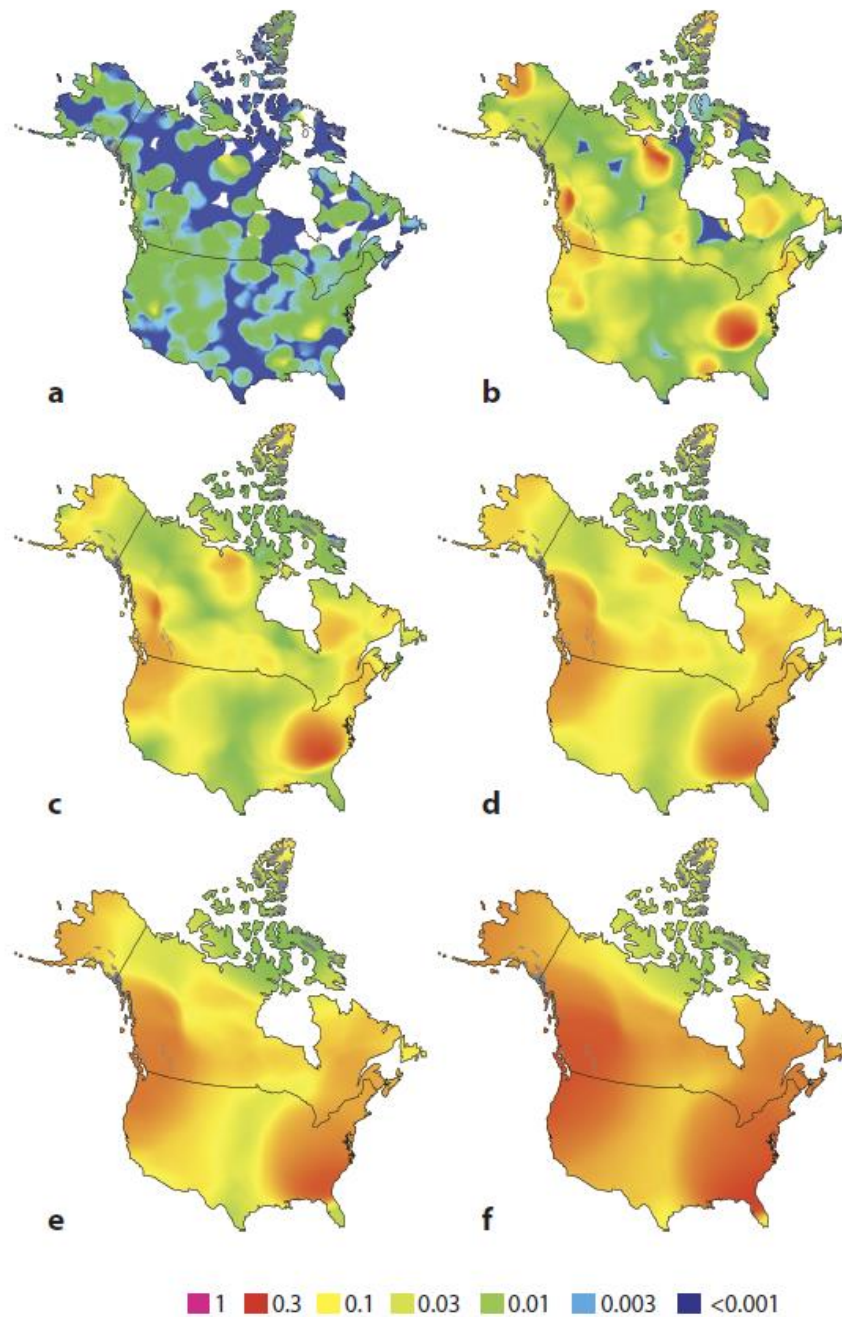
### *Calibration Curves: IntCal04 versus IntCal13*

In this study, we use a previously calibrated version of the CARD based on CALIB v.5.0.2 and IntCal04 (Munoz et al., 2010; Peros et al., 2010; Gajewski et al., 2011) since recalibrating the dates using IntCal13 would make no difference to our results. To demonstrate how little of an effect the changes made between IntCal04 and IntCal13 have on the calibration results (Bronk Ramsey et al., 2013; Reimer et al., 2013) we chose a random sample of 200 dates which span the Holocene, and calibrated these using IntCal13 and the northern hemisphere calibration curve (removing marine samples which require a marine correction;  $n=8$ ) and using 2 standard deviations. We compared these to the IntCal04 calibrated data we use here (**Fig. A1.9**). The two calibrated series are identical except for two points which are from freshwater molluscs, and even these two points differ by much less than our sampling interval.

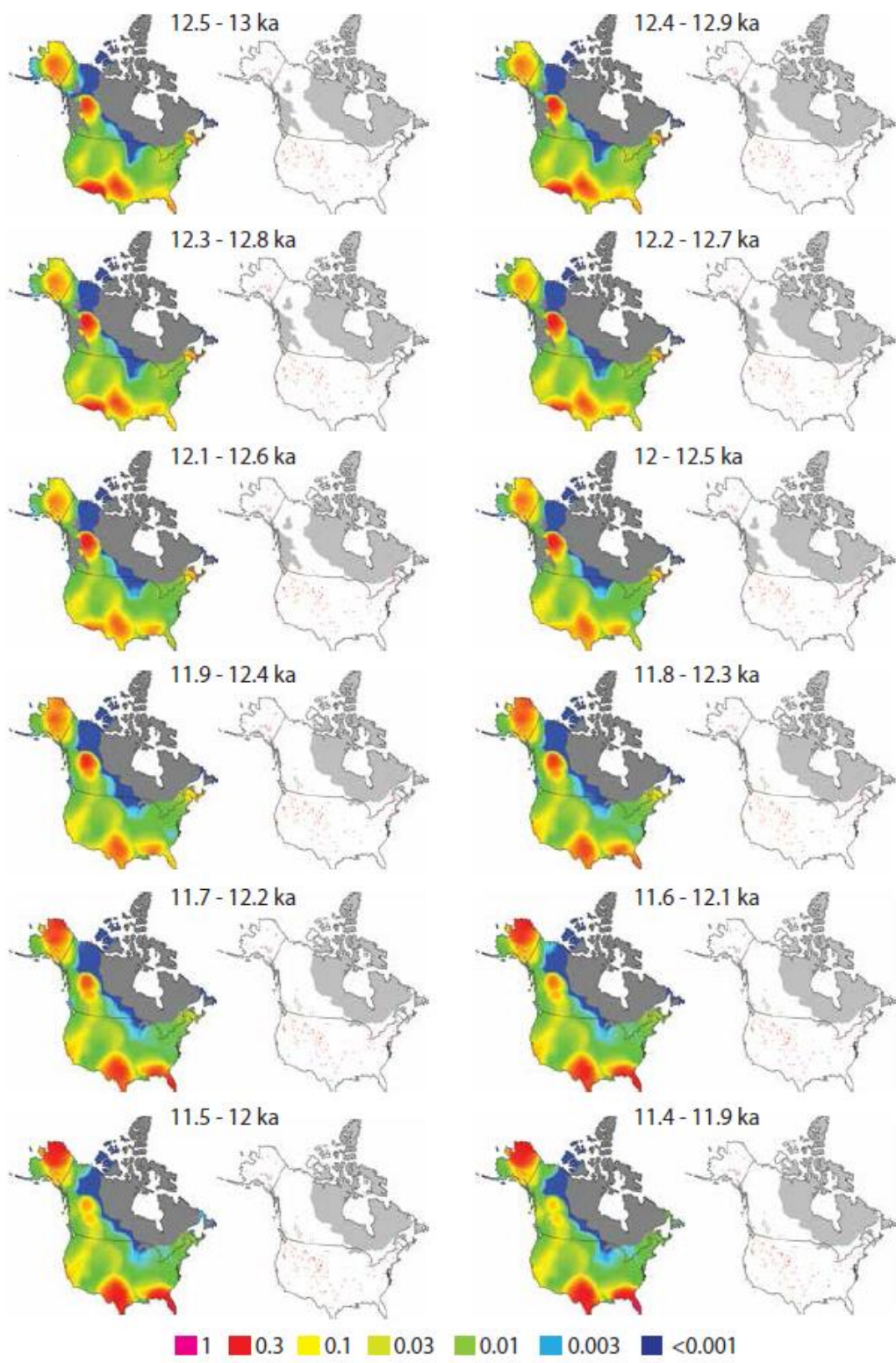
### *Robustness of Results*

When analyzing RFPEs the question about robustness of inferred results arises. In the previous paragraphs we showed that modifications of the method have only a minor influence on the results. For the broad population patterns, the effect of varying the kernel function is just as small as that of applying an adaptive smoothing parameter. The main regions of population activity stay the same; only the sizes of different population events may vary. A larger effect is obvious if the global bandwidth is varied. An effect on the scale of population events is then visible, whereas affected regions still stay the same.

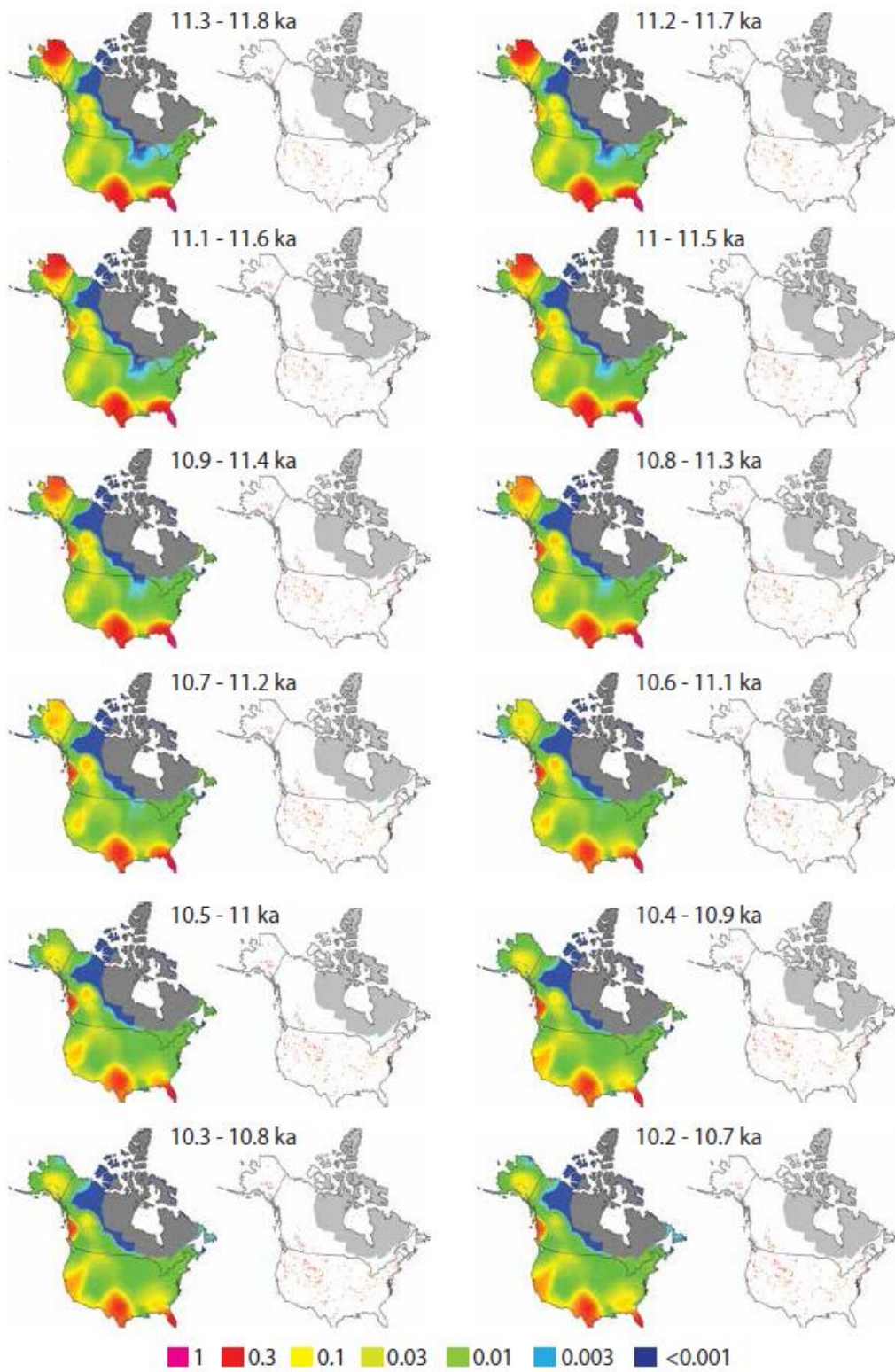
Additionally, we showed that the method is robust in dealing with outliers and anomalous data. If 5% of the dates are considered to be “anomalous” and dropped, results are almost identical. Even if the CARD is reduced randomly by 50%, our results remain valid – a conclusion which is remarkable in particular for earlier time intervals. This shows the true significance and robustness of the RFPEs. Naturally, the next step would be to further validate our results by using a statistical test for significance of estimated RFPEs (Shennan et al., 2013). However, performing such a test would require a stochastic spatio-temporal model for RFPEs; this is a non-trivial issue and beyond the scope of this paper. The development of an asymptotic test would require extensive theoretical consideration, which is not possible in this context. Bootstrap methods would be appropriate for performing statistical tests, but cannot be applied here due to computational (time) restrictions. The development and implementation of a model-based statistical test will be subject to future research.



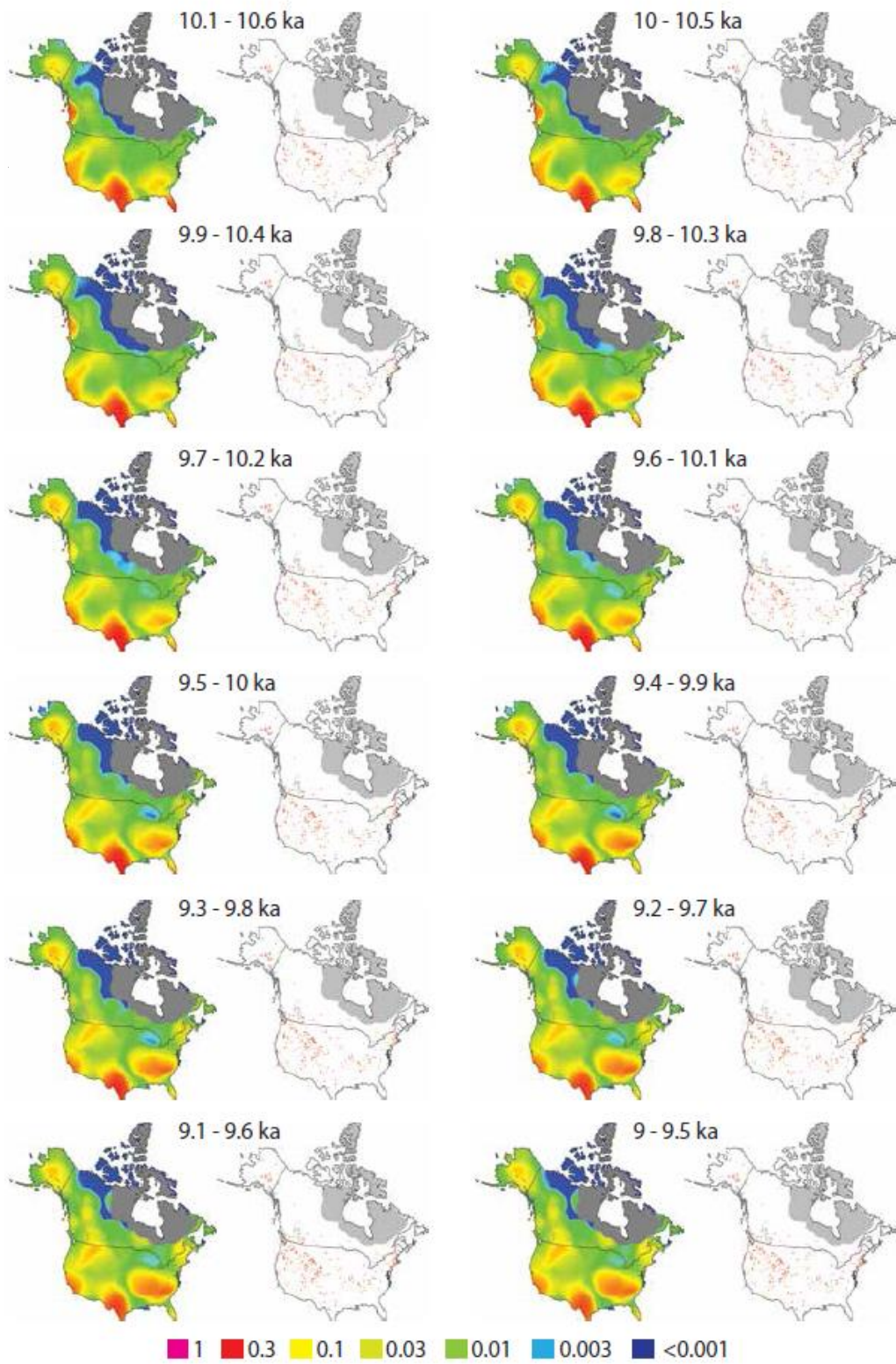
**Figure A1.1** Maps showing the effects of choosing different kernel function radii on RFPes between 3 and 3.5 ka in relation to the (c) 600 km radius used in this study. When a radius less than 500 km is used the continental-scale pattern is lost (a: 200km, b: 400km). When the radius exceeds 800 km local changes are over-smoothed (d: 840km, e: 1,000km, f: 1,250km).



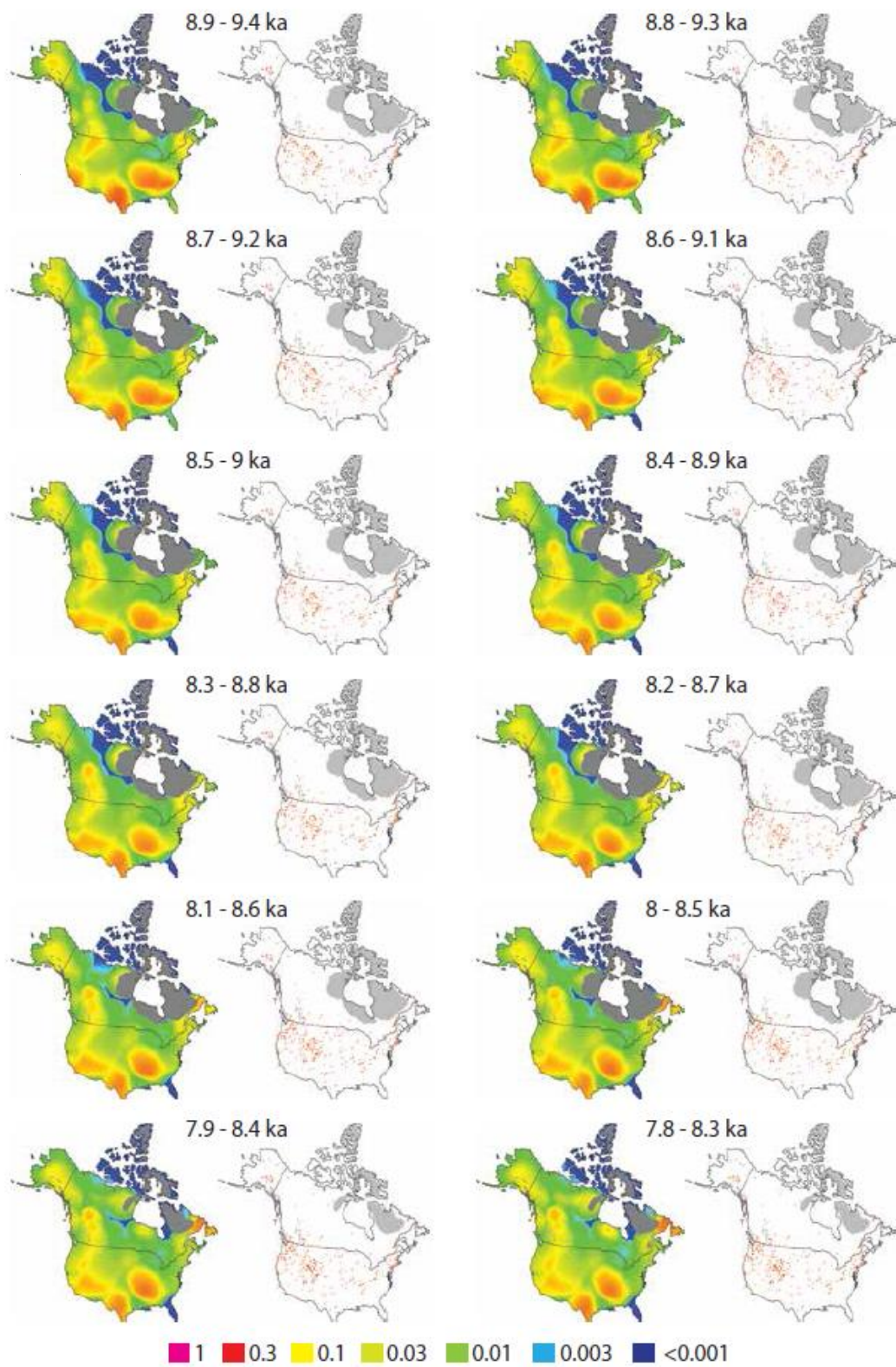
**Figure A1.2** Radiocarbon Frequency Population Estimate (RFPE) maps for the past 13 ka. RFPEs and the distribution of the associated CARD dates between 0.5 and 13 ka. Grey = ice.



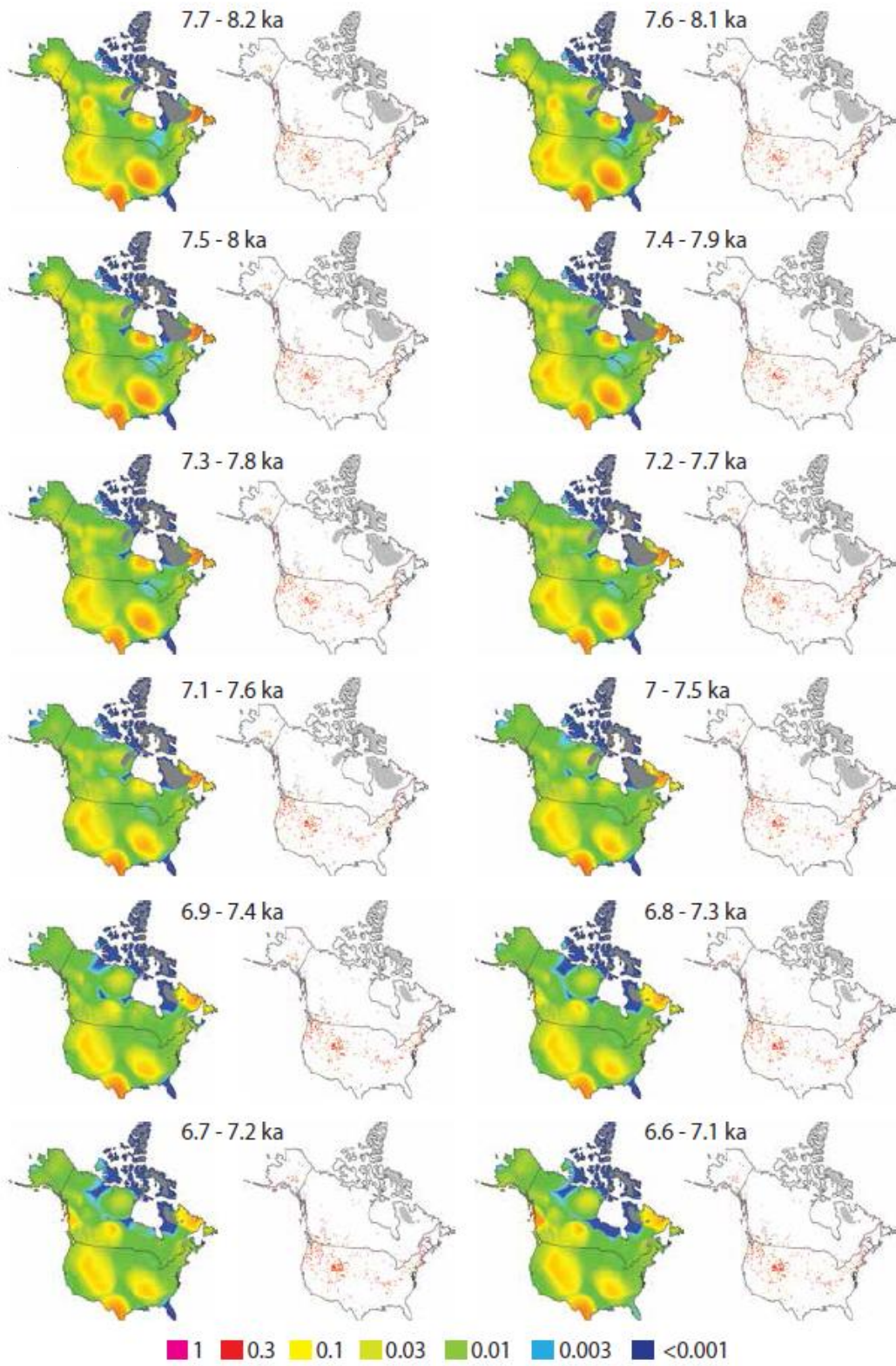
**Figure A1.2** Continued.



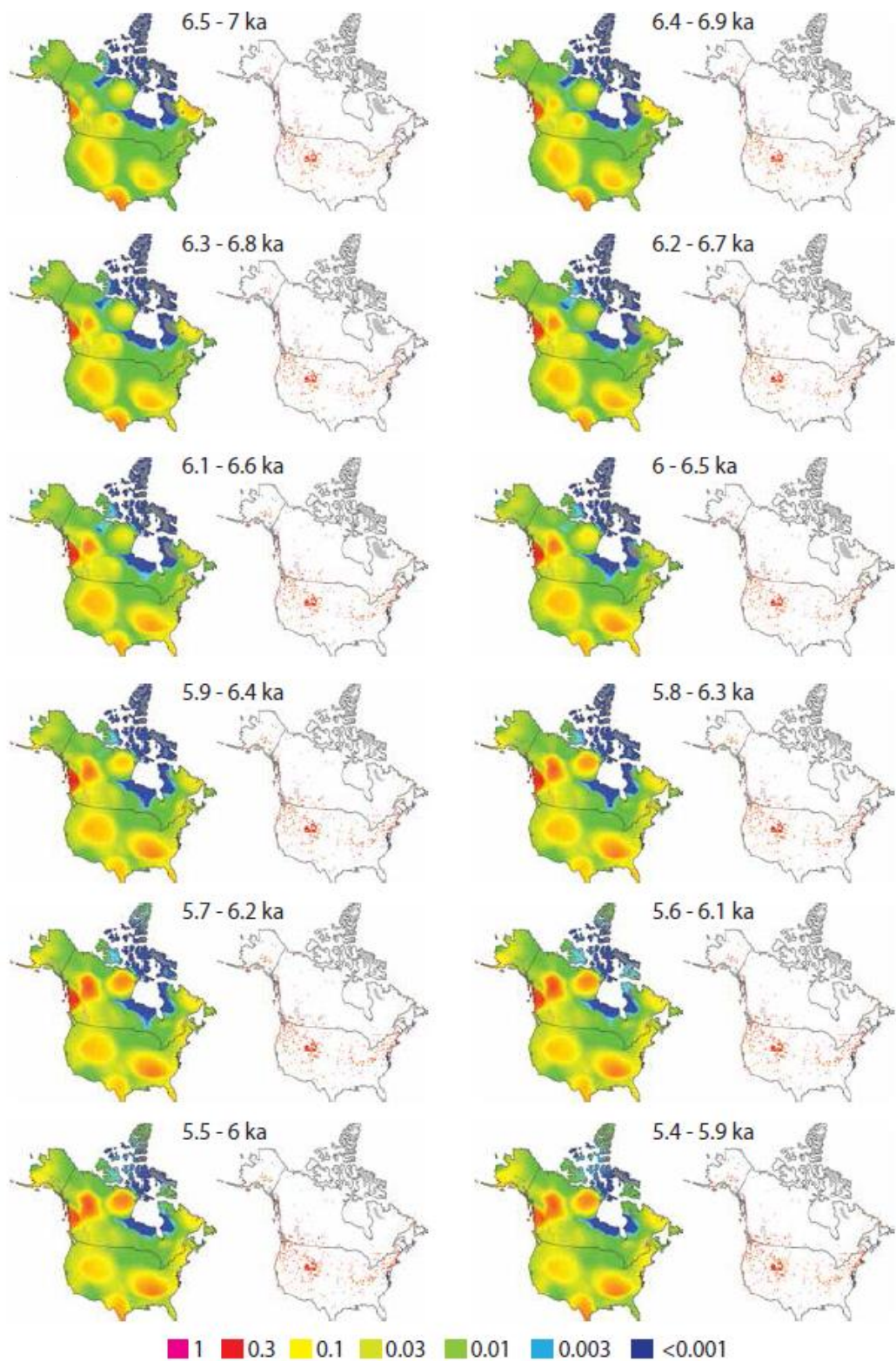
**Figure A1.2** Continued.



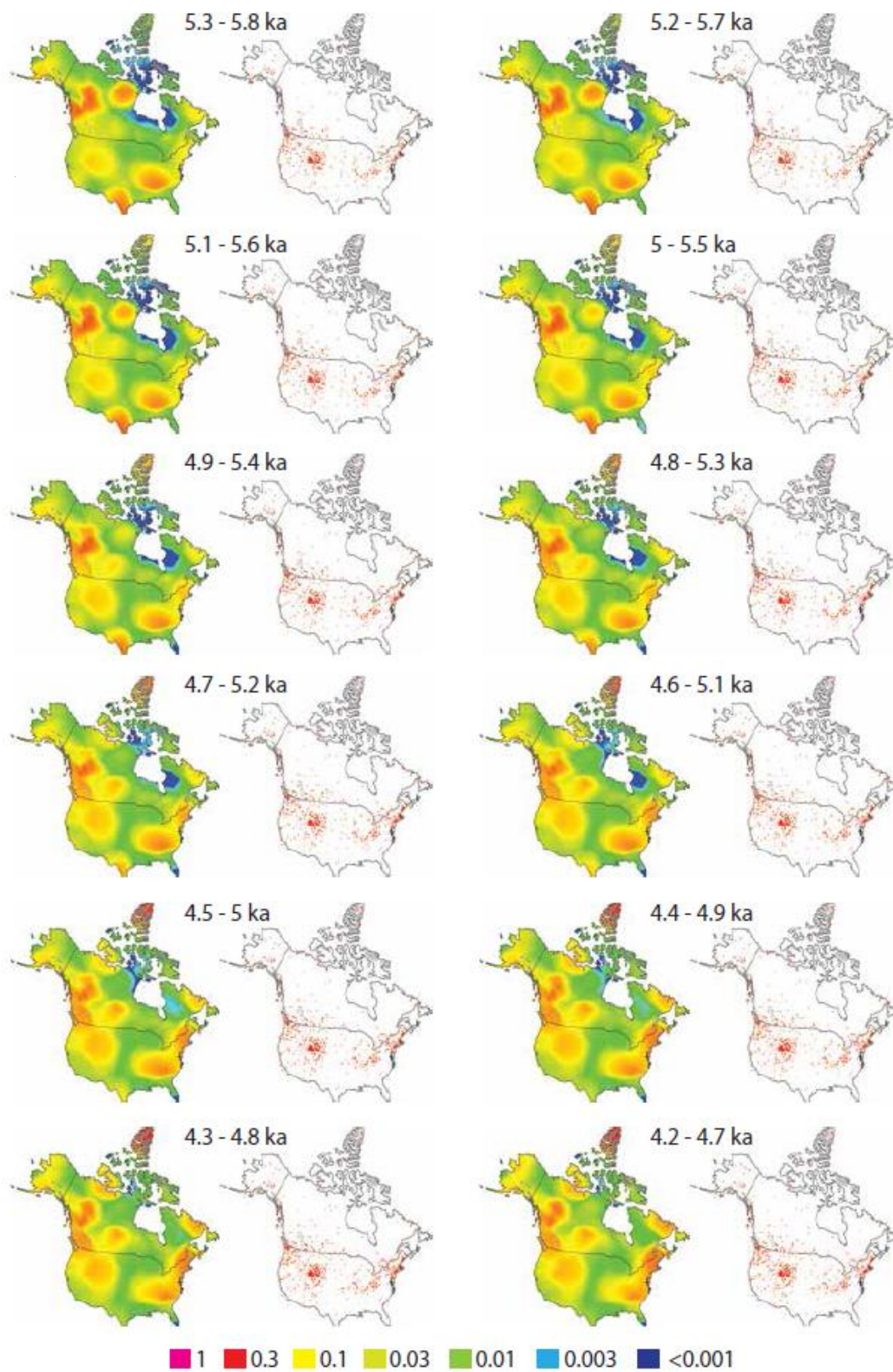
**Figure A1.2** Continued.



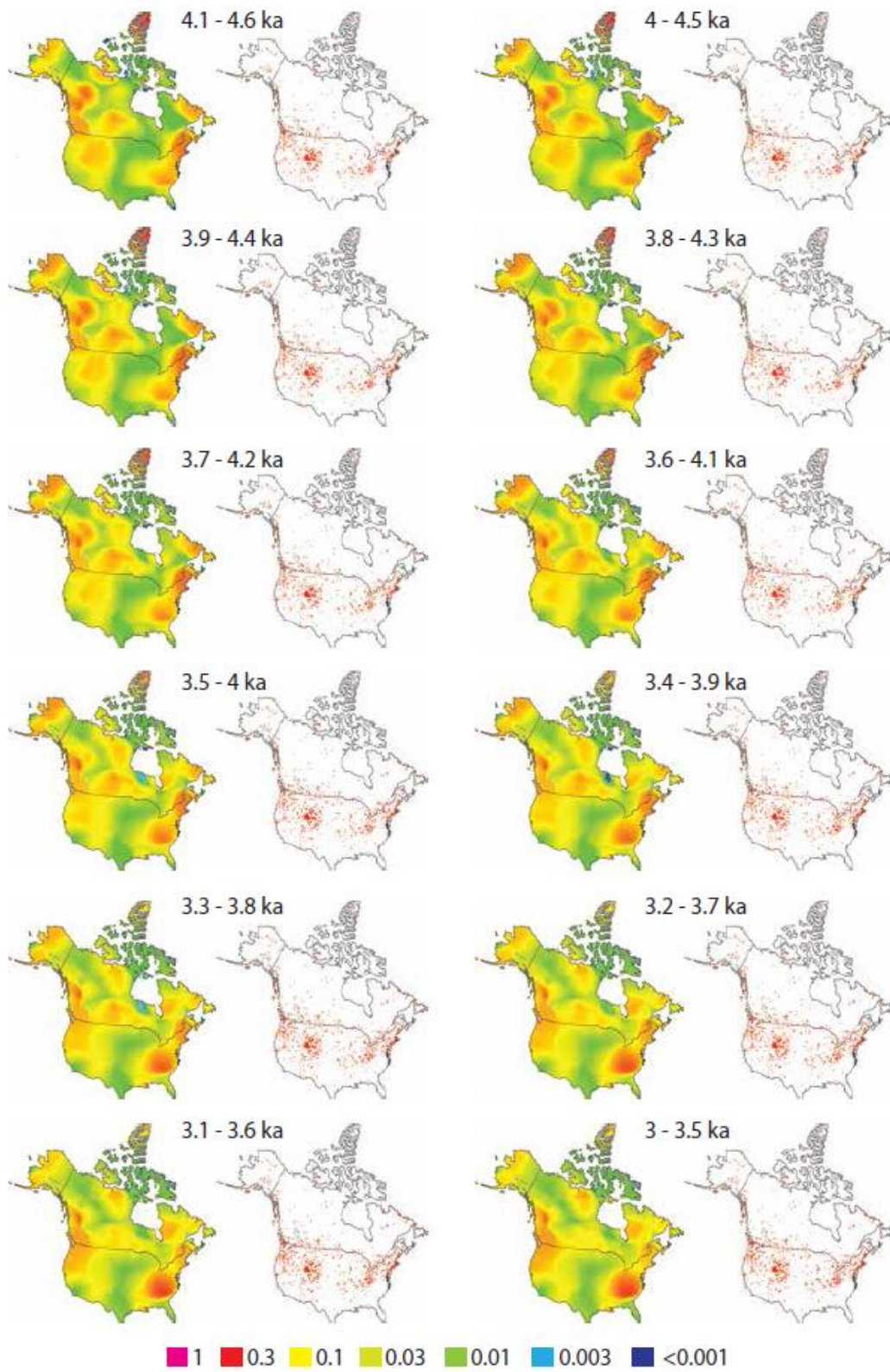
**Figure A1.2** Continued.



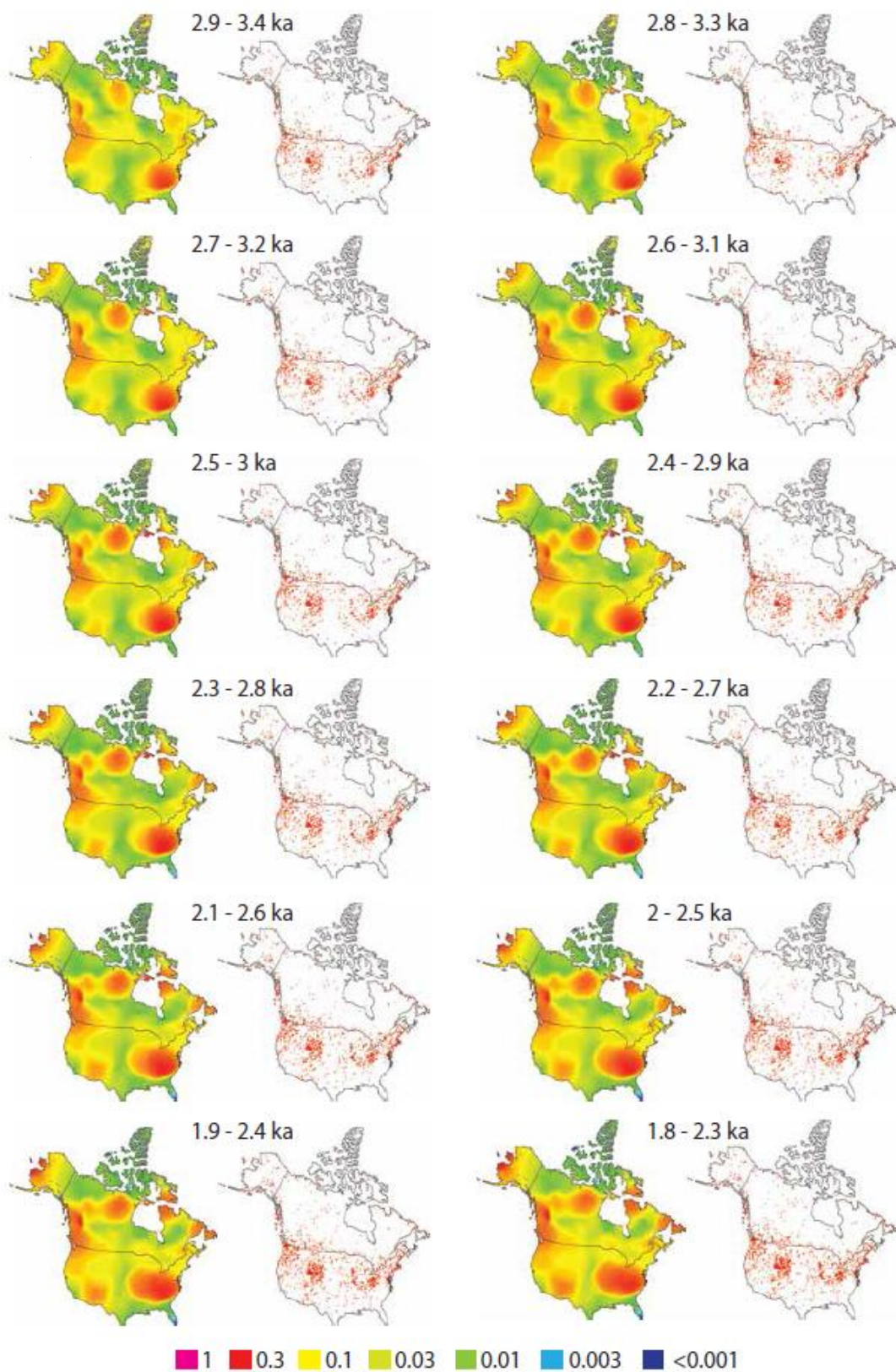
**Figure A1.2** Continued.



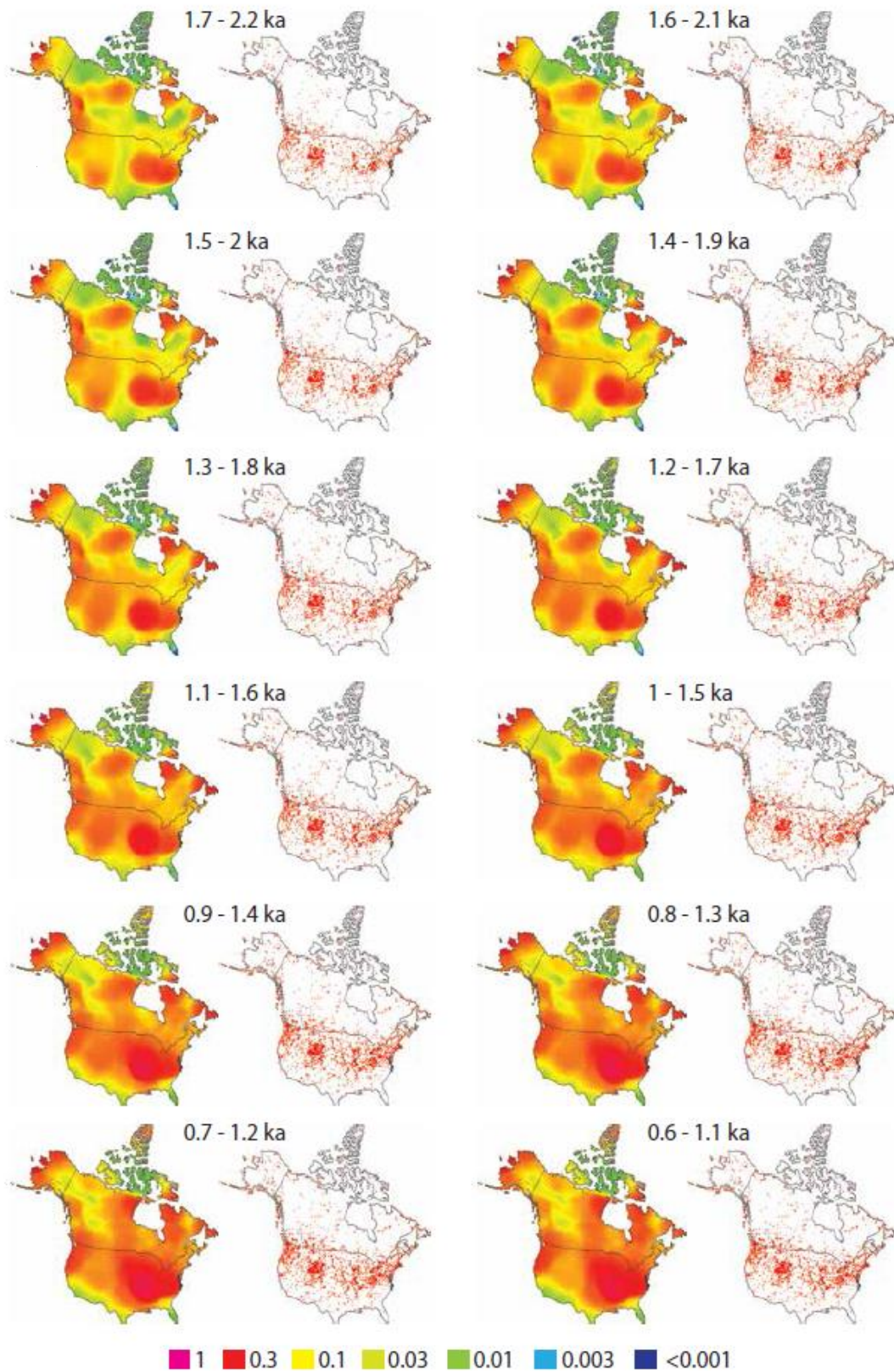
**Figure A1.2** Continued.



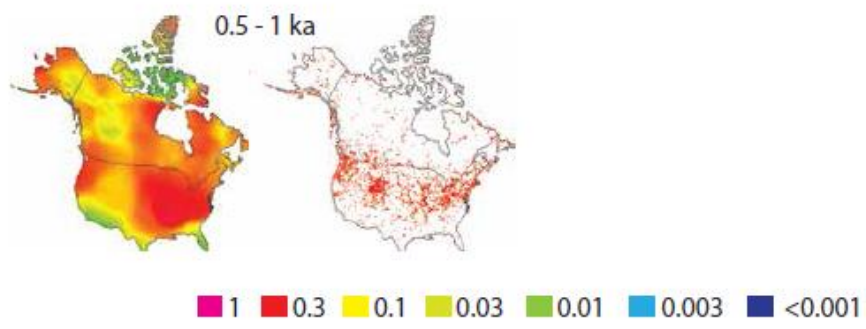
**Figure A1.2** Continued.



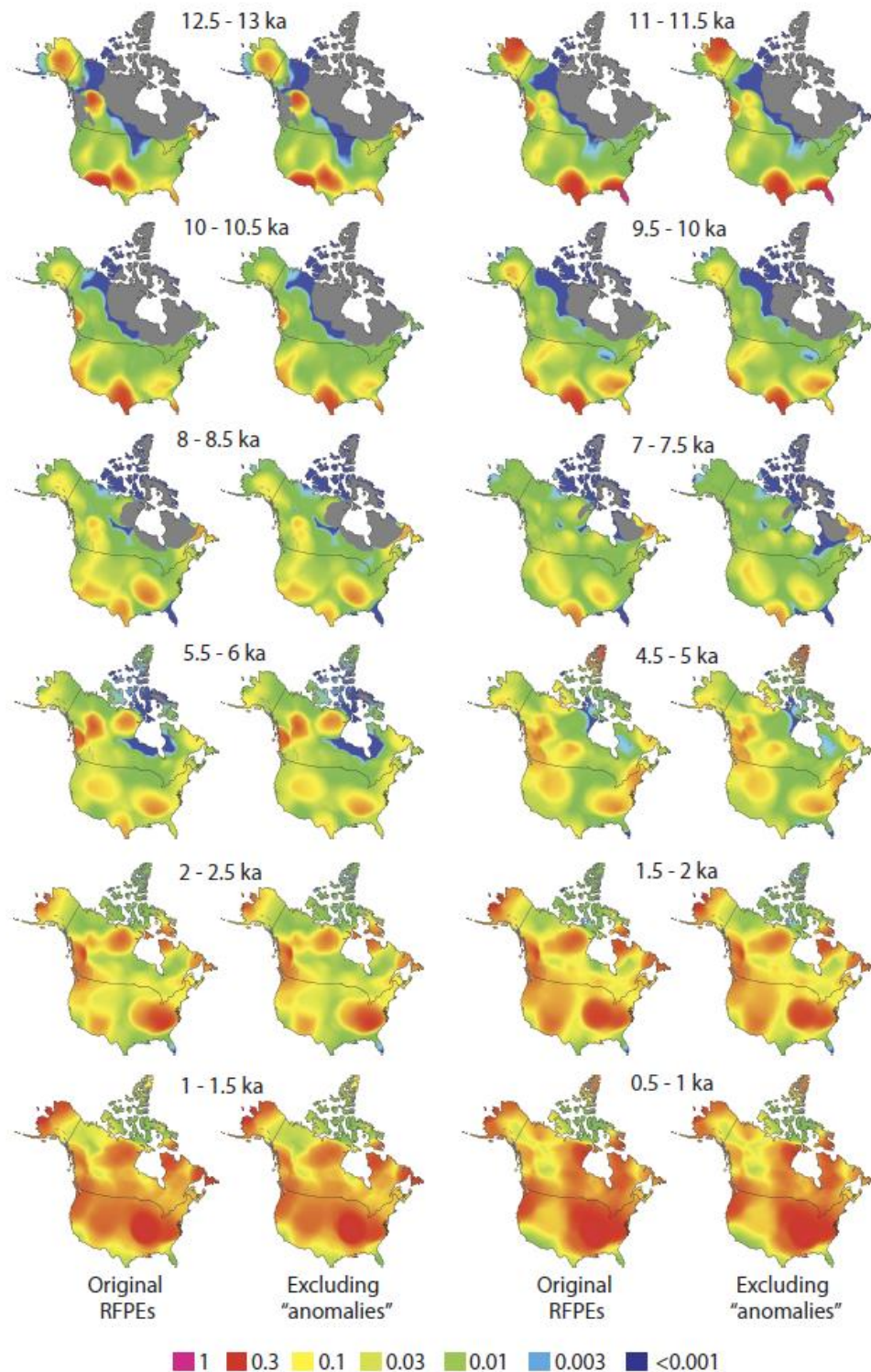
**Figure A1.2** Continued.



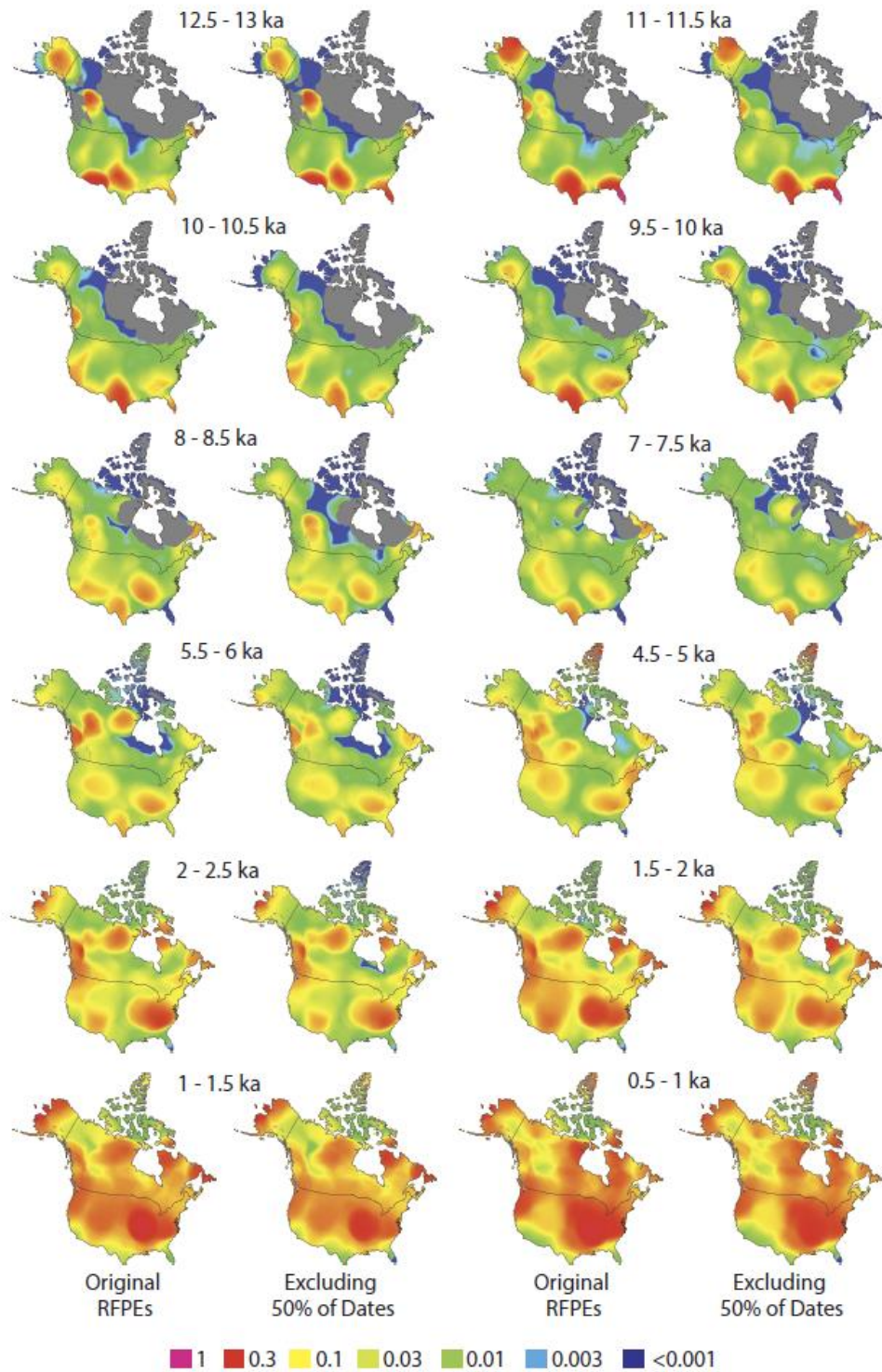
**Figure A1.2** Continued.



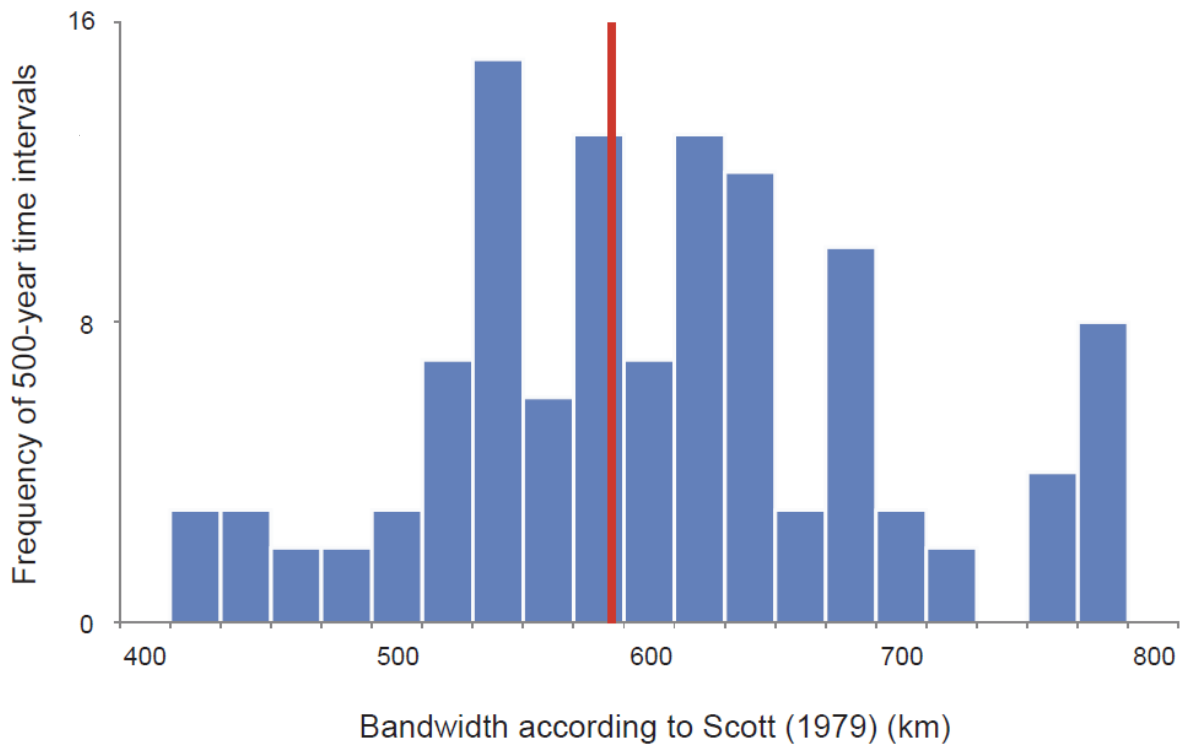
**Figure A1.2** Continued.



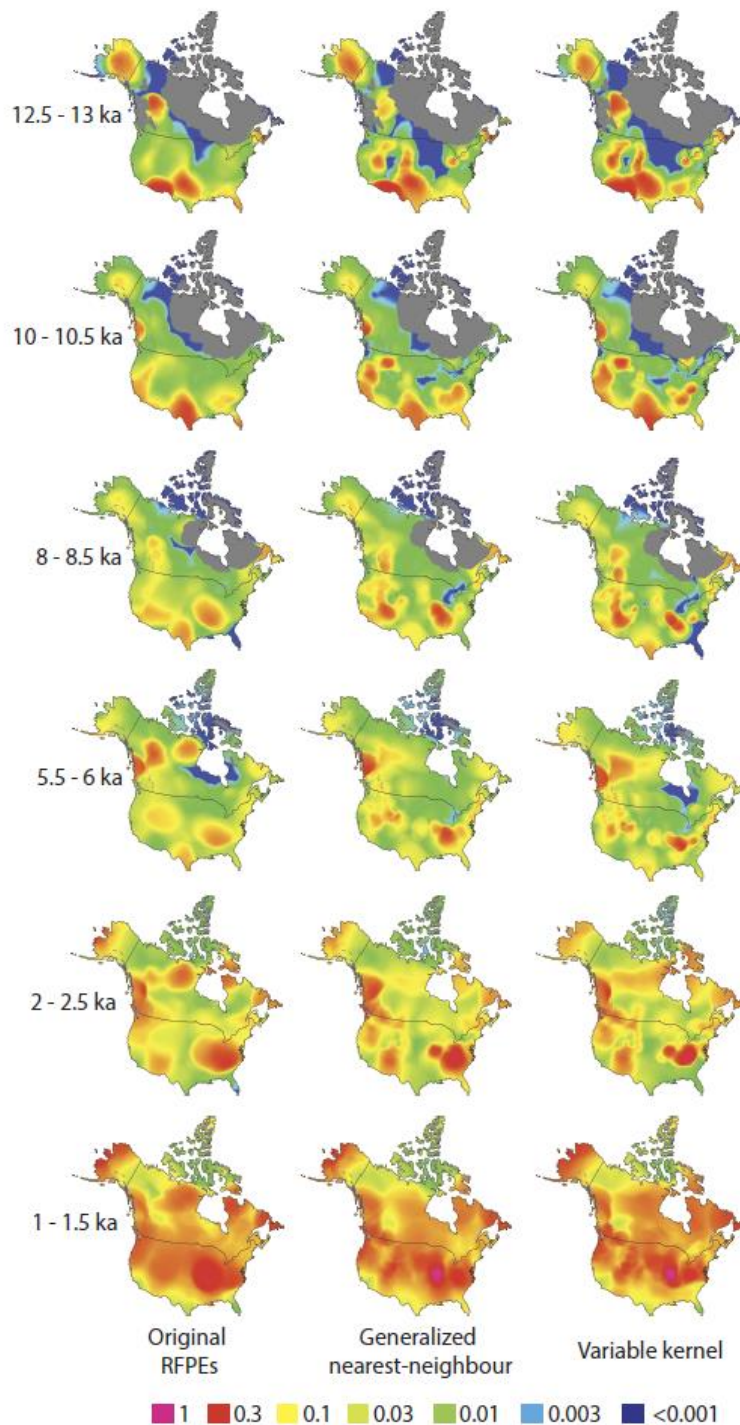
**Figure A1.3** Comparison of the original RFPEs shown in [Figure 3.2](#) with recalculated RFPEs based on the exclusion of CARD dates described as “anomalous”. The original RFPEs are calculated using all available data in the CARD whereas the recalculated RFPEs are based on a subset of the data with 1,419 fewer dates.



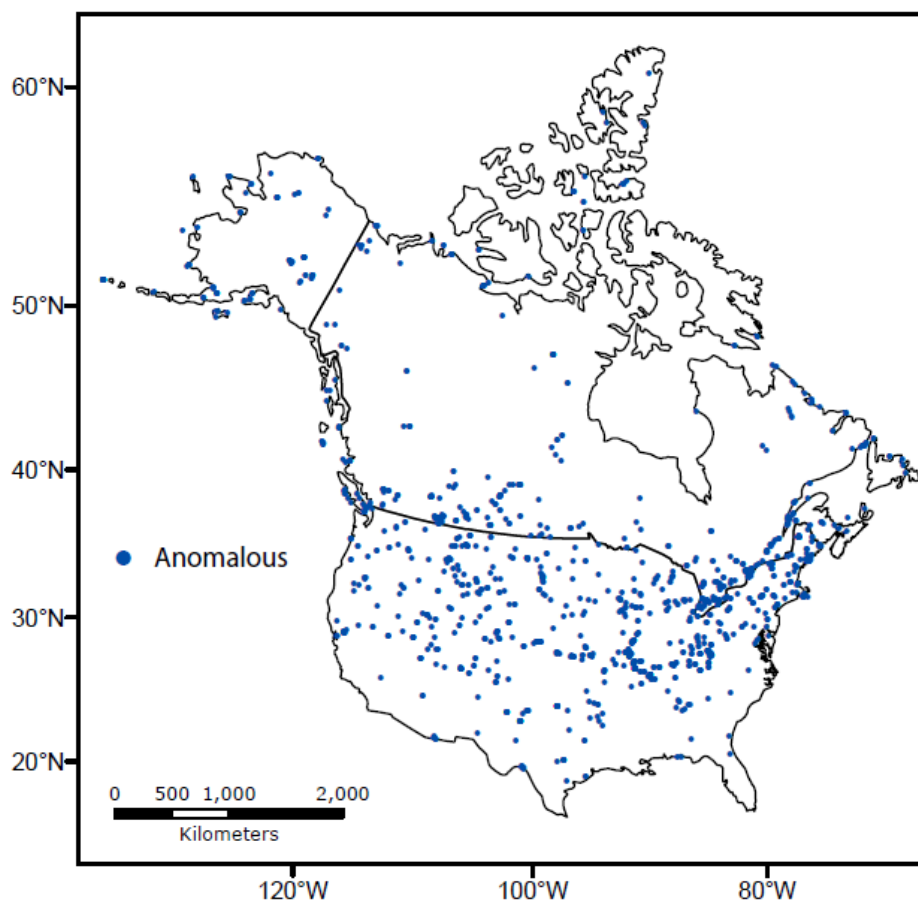
**Figure A1.4** Comparison of the original RFPEs shown in [Figure 3.2](#) with recalculated RFPEs based on random exclusion of 50% of all CARD dates. The original RFPEs are calculated using all available data in the CARD whereas the recalculated RFPEs are based on a randomly-selected subset of the data with 16,894 fewer dates.



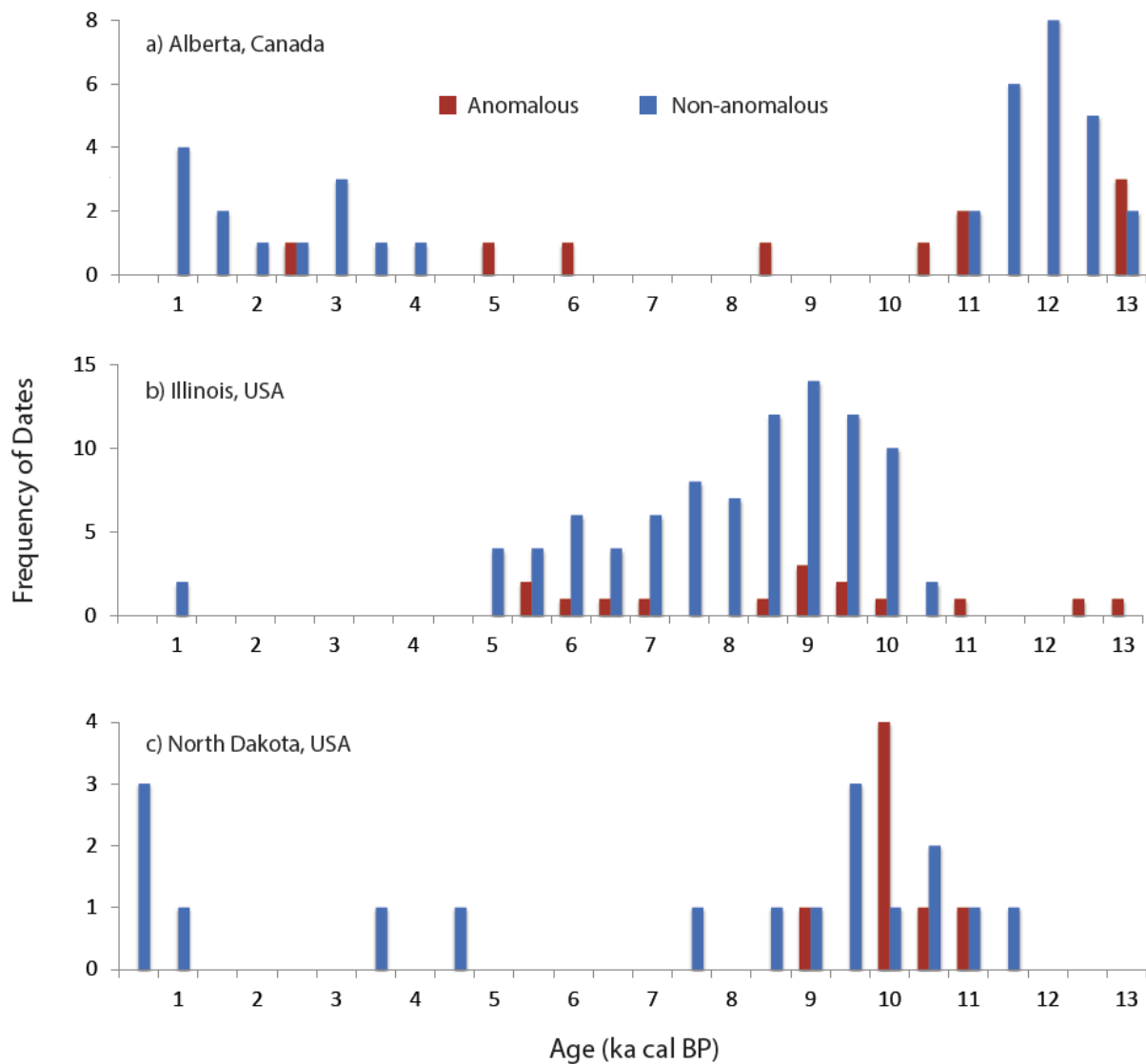
**Figure A1.5** Histogram of automatically-calculated bandwidths based on Scott's rule (Scott, 1992). Optimal bandwidths calculated for all 500-year time intervals. Choosing the mean (595 km indicated by the red line) appears to be an adequate choice for all time intervals.



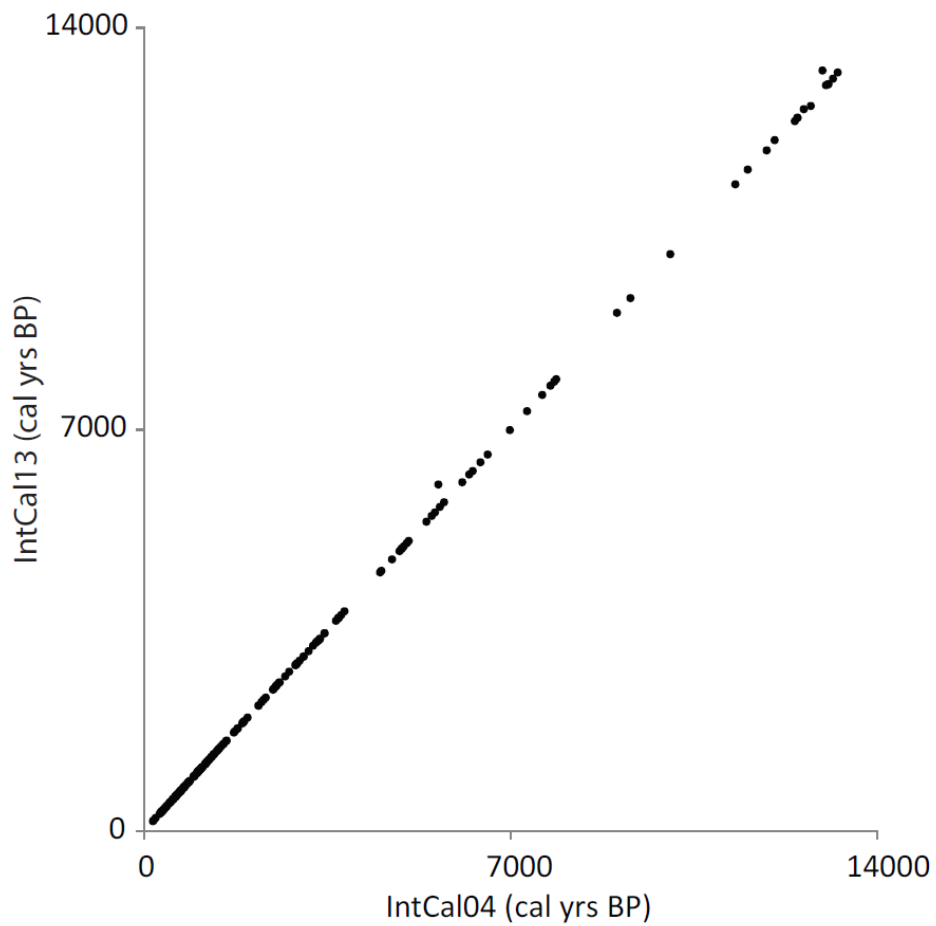
**Figure A1.6.** Maps showing the effects of varying the bandwidth when a nearest-neighbour bandwidth selection approach and a variable kernel approach are used. “Original RFPEs” refers to the Epanechnikov kernel function used in this study. Varying the bandwidth spatially changes the intensity and scale of population activity but locations of activity and major settlements events generally remain the same.



**Figure A1.7** Locations of the archaeological sites in the CARD with <sup>14</sup>C dates described as “anomalous”. Anomalous descriptions in the CARD are based on the interpretations of the submitting authors and are typically a result of being younger or older than expected.



**Figure A1.8** Histograms of “anomalous” dates from 3 sites and non-anomalous dates within a 25 km radius of each site. In some instances, the inclusion of “anomalous” dates alters the duration of human occupation (a, b) which is observed in the smoothing of the RFPE maps. In contrast, at some sites the distributions of “anomalous” and correct dates are identical (c) resulting in little to no effect on the RFPE patterns.



**Figure A1.9.** Calibrations using IntCal04 and IntCal13. Randomly selected dates from the CARD were calibrated using IntCal04 and IntCal13 to show the scant difference when using one or the other to calibrate dates from the Holocene.

## Appendix 2 (Video for Chapter 3)

Video available at <http://www.lpc.uottawa.ca/members/chaput/Animations.html>

**Video A2.1** The spatiotemporal distribution of Holocene populations in North America based on archaeological radiocarbon dates.

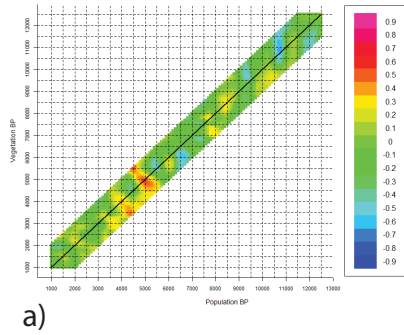
### Appendix 3 (Supplementary Information for Chapter 5)

**Figure A3.1** The following are heat maps summarizing the spatial correlations of maps of the 10 taxa, and their cross-correlation with population. Each page is dedicated to one taxon, and each graph is the stacked summary of 121 spatial correlation or cross-correlation functions. The taxa included are, in order: *Picea*, *Pinus*, *Quercus*, *Castanea*, *Juglans*, *Carya*, *Acer*, *Fagus*, *Populus*, *Poaceae*.

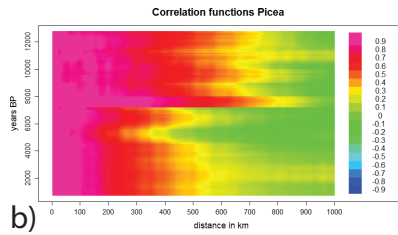
- a) Summary of cross-correlations between difference maps; as in [Figure 5.3](#)
- b) Correlation functions of the maps of taxon abundance
- c) Cross-correlations of maps of taxon abundance and population estimates
- d) Correlation functions of 500-year differences in taxon abundance, and cross-correlation of these maps with population at
- e) zero lag
- f) lag of +100 years
- g) lag of -100 years
- h) lag of +300 years
- i) lag of -300 years
- j) lag of +500 years
- k) lag of -500 years
- l) lag of +1000 years
- m) lag of -1000 years

A positive lag means that vegetation change corresponds to a more recent 500-year change interval than change in population and a negative lag means the opposite.

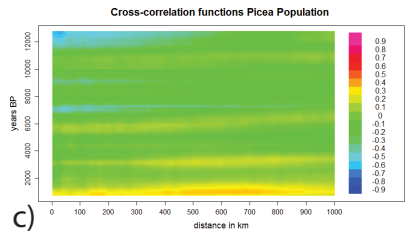
# Picea



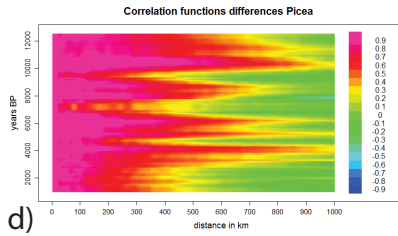
a)



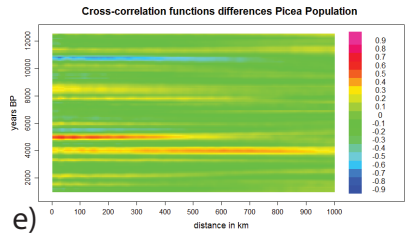
b)



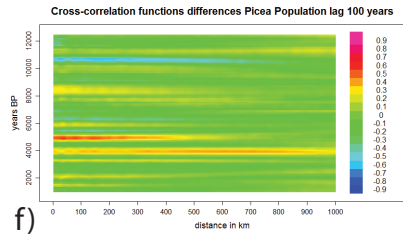
c)



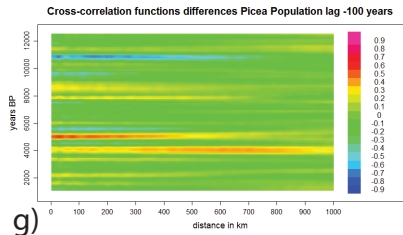
d)



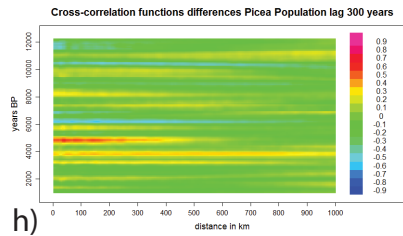
e)



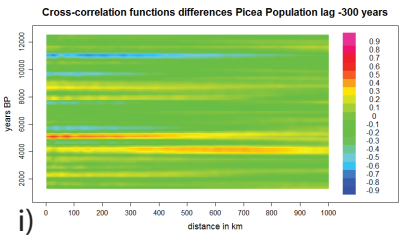
f)



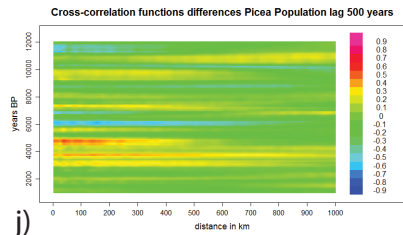
g)



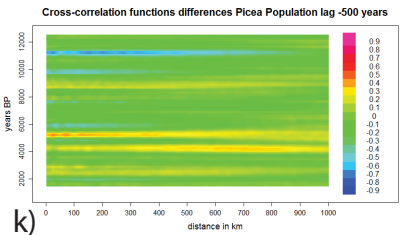
h)



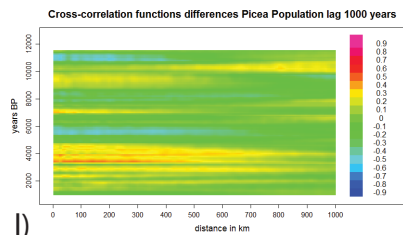
i)



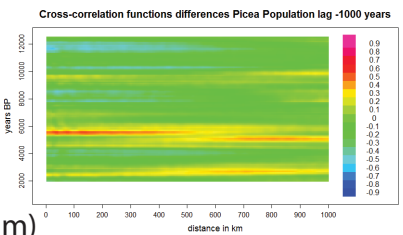
j)



k)

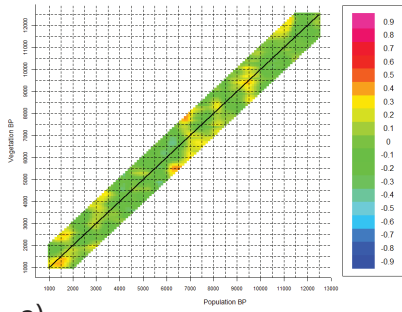


l)

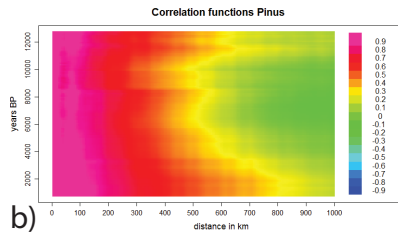


m)

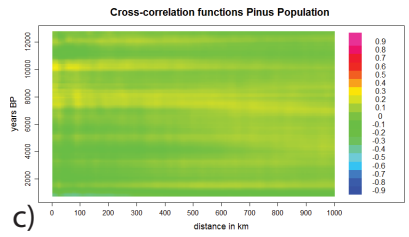
# Pinus



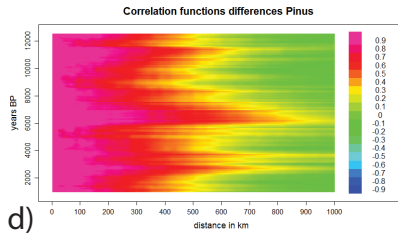
a)



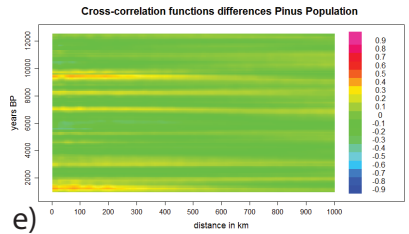
b)



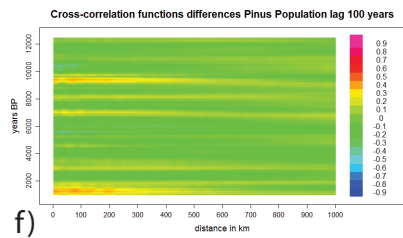
c)



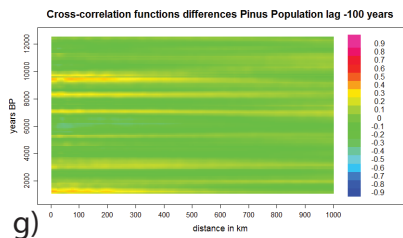
d)



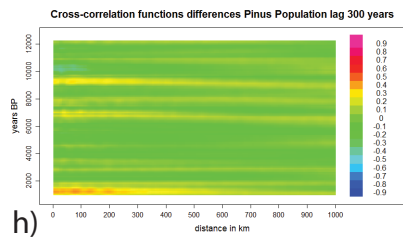
e)



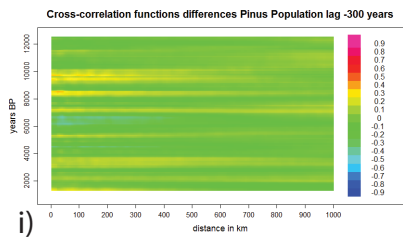
f)



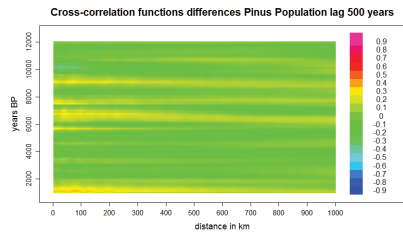
g)



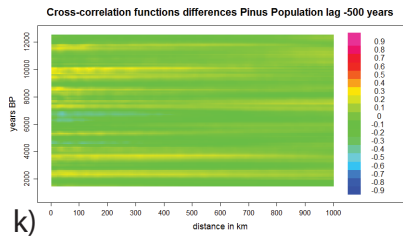
h)



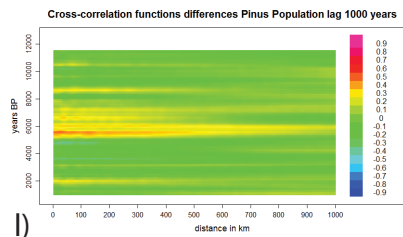
i)



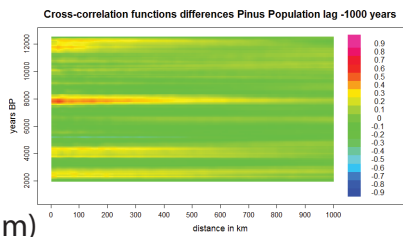
j)



k)

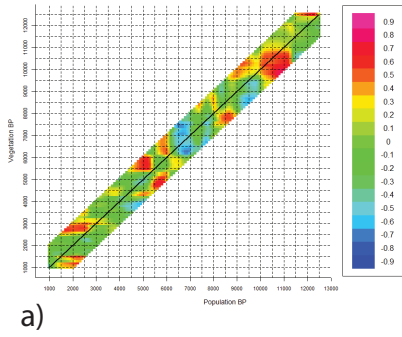


l)

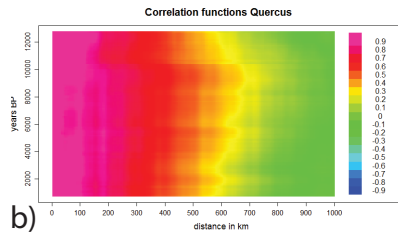


m)

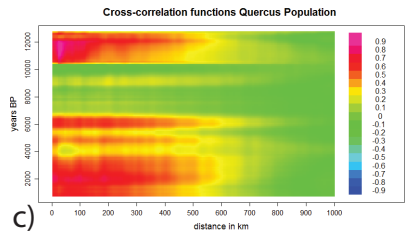
# Quercus



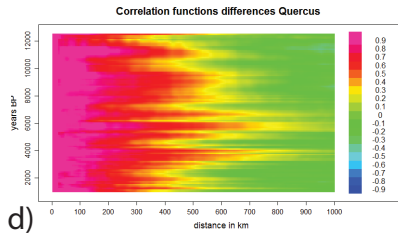
a)



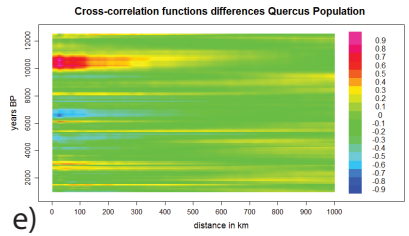
b)



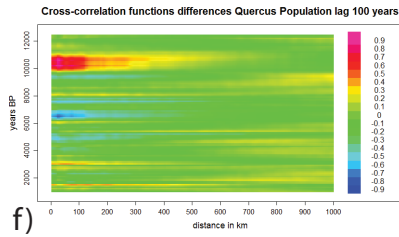
c)



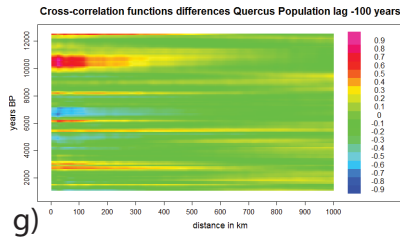
d)



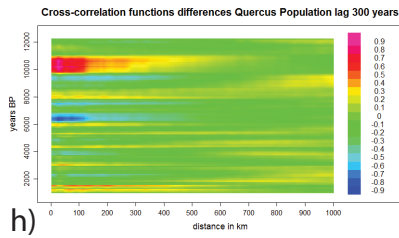
e)



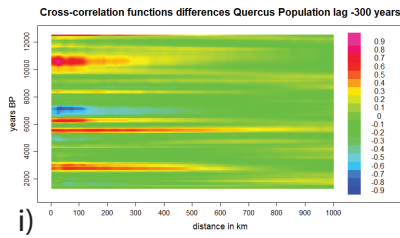
f)



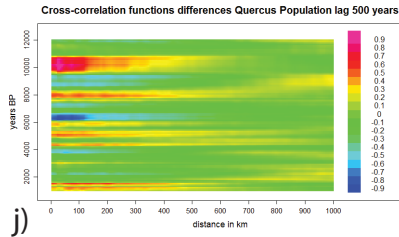
g)



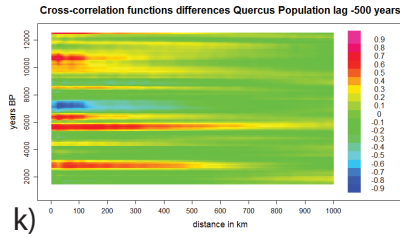
h)



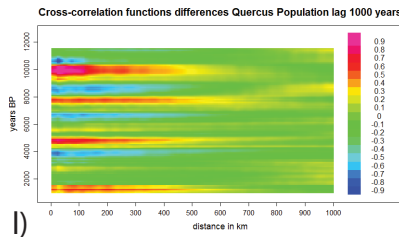
i)



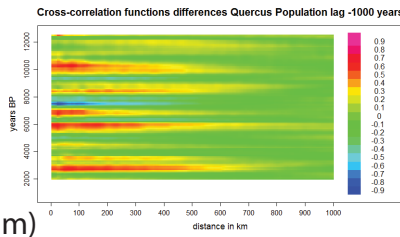
j)



k)

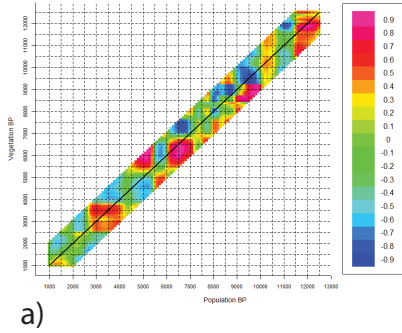


l)

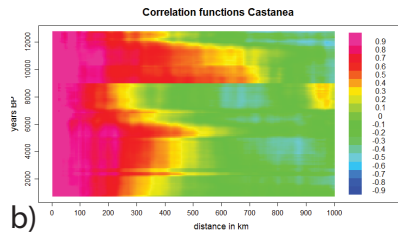


m)

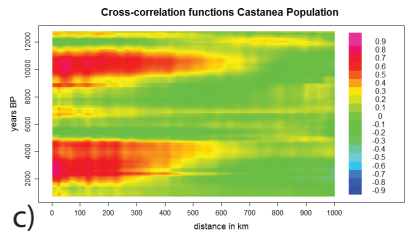
# Castanea



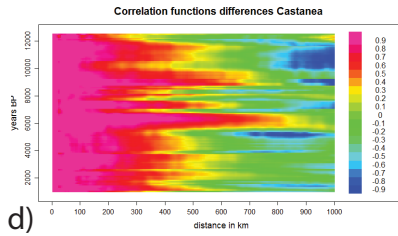
a)



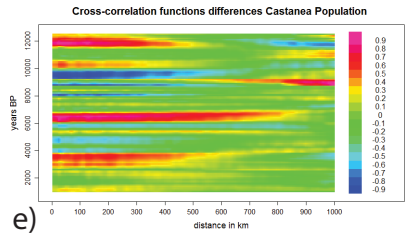
b)



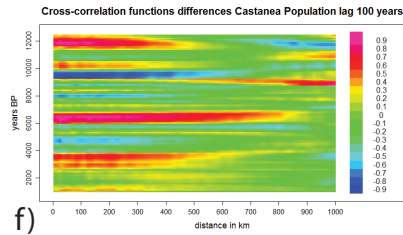
c)



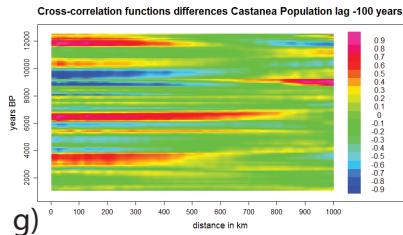
d)



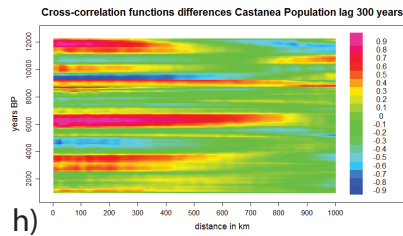
e)



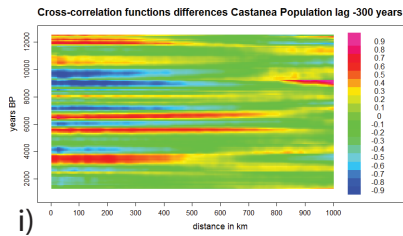
f)



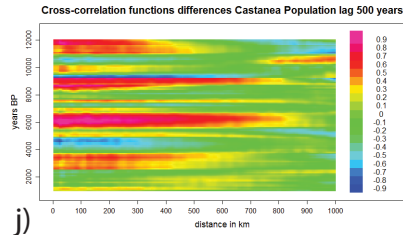
g)



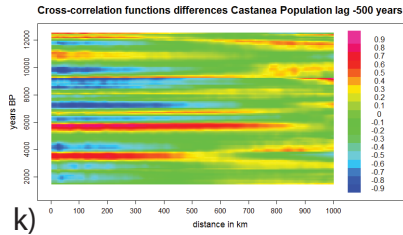
h)



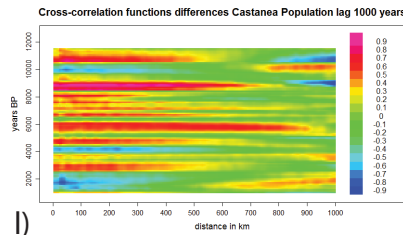
i)



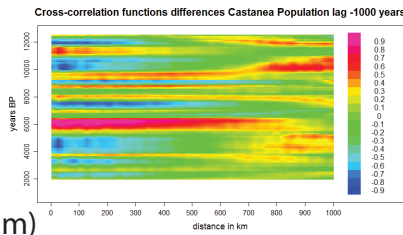
j)



k)

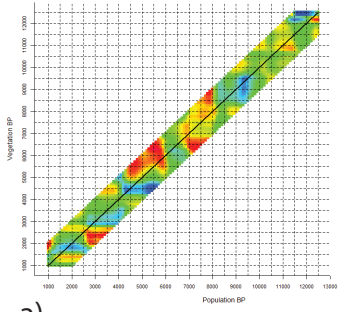


l)

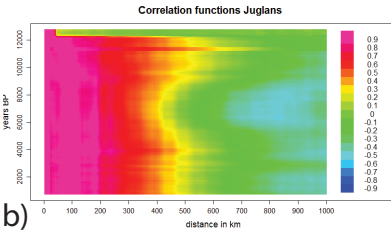


m)

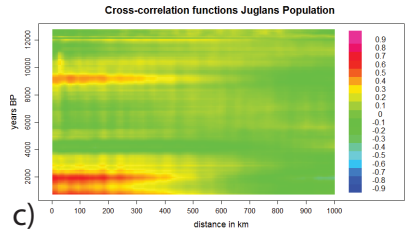
# Juglans



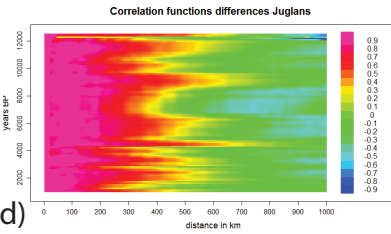
a)



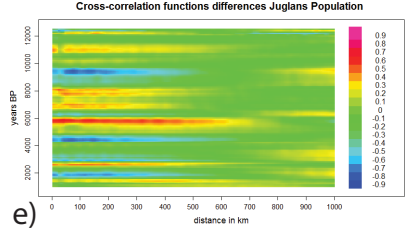
b)



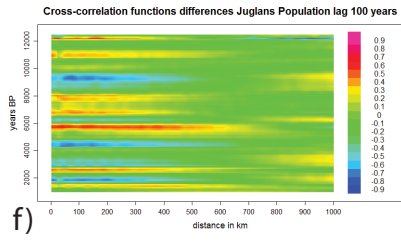
c)



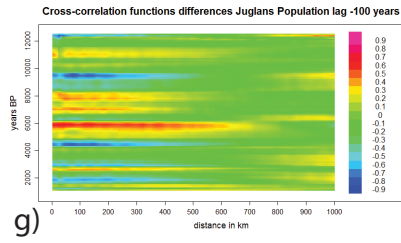
d)



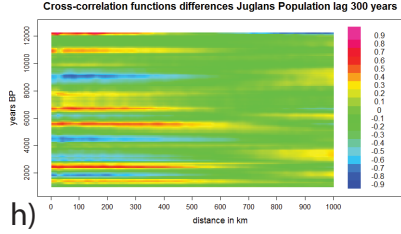
e)



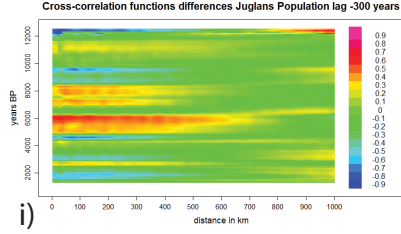
f)



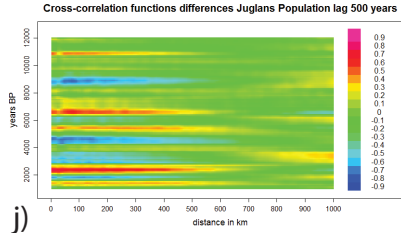
g)



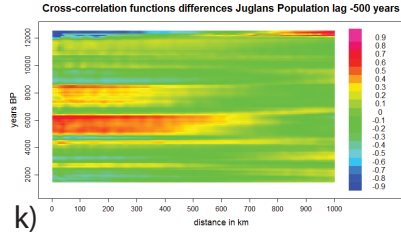
h)



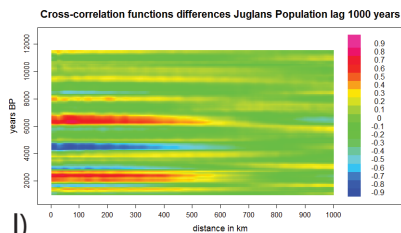
i)



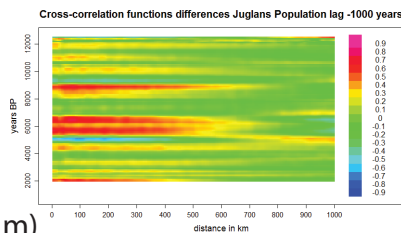
j)



k)

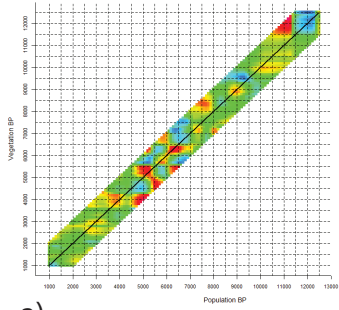


l)

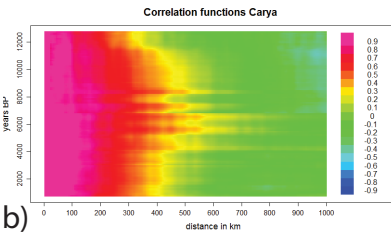


m)

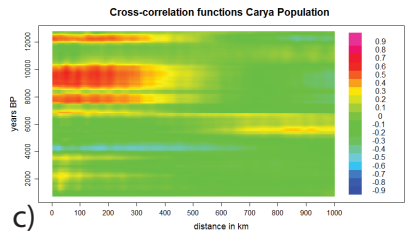
# Carya



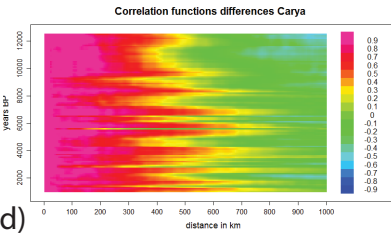
a)



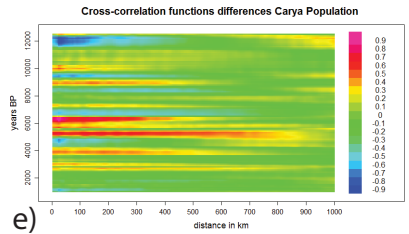
b)



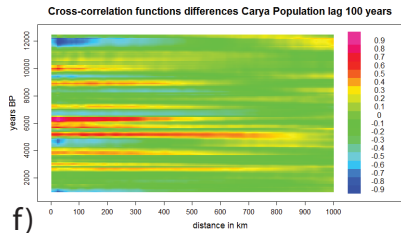
c)



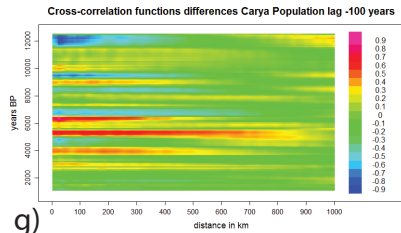
d)



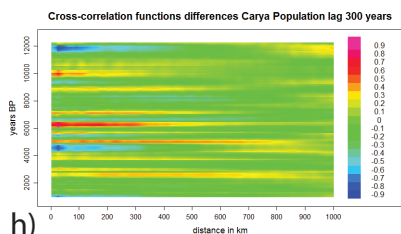
e)



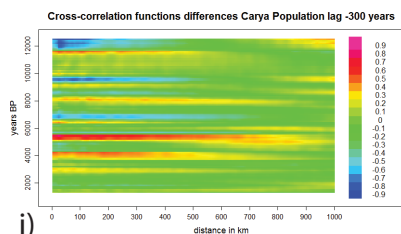
f)



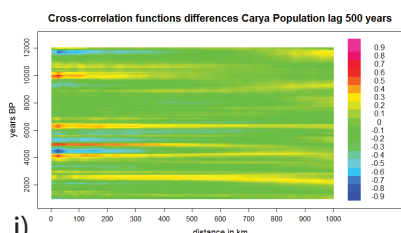
g)



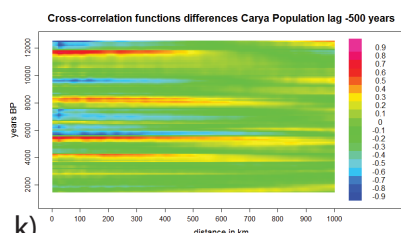
h)



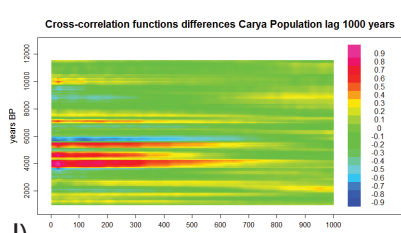
i)



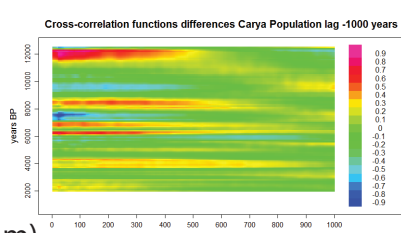
j)



k)

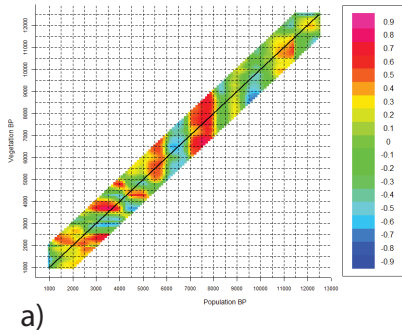


l)

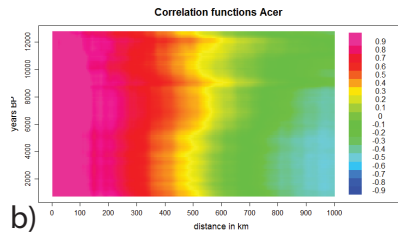


m)

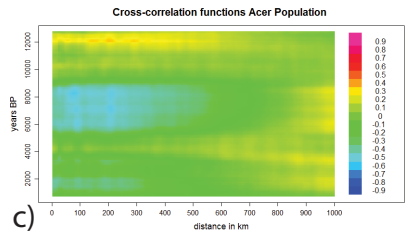
# Acer



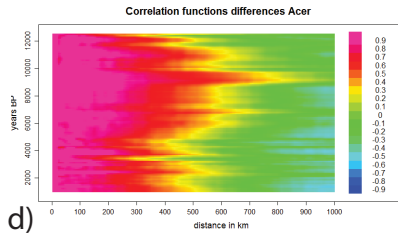
a)



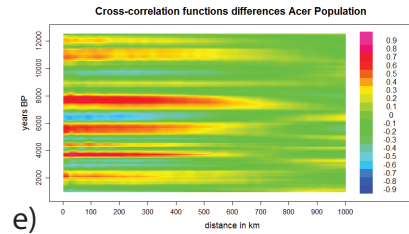
b)



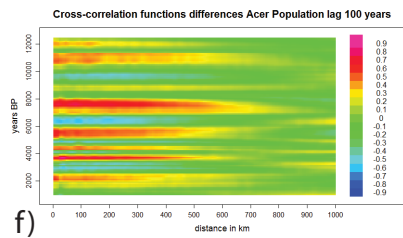
c)



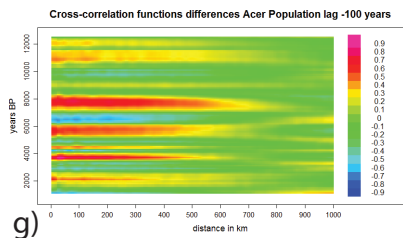
d)



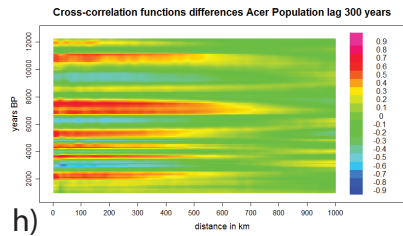
e)



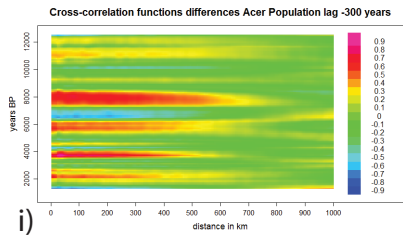
f)



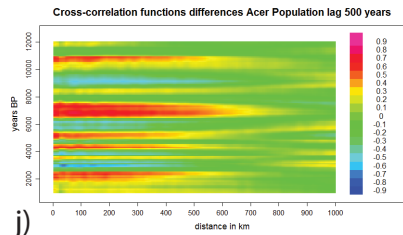
g)



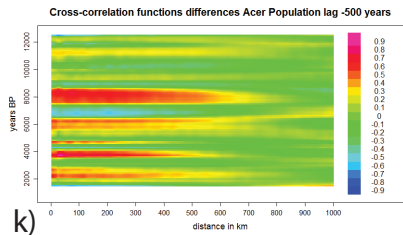
h)



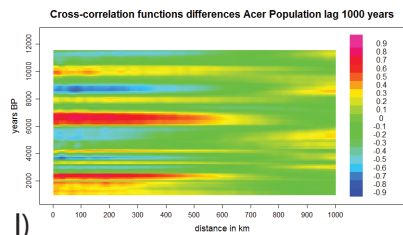
i)



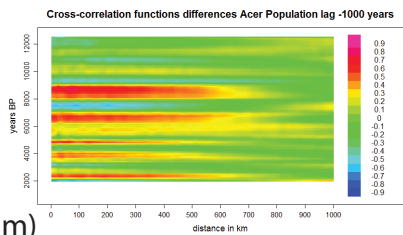
j)



k)

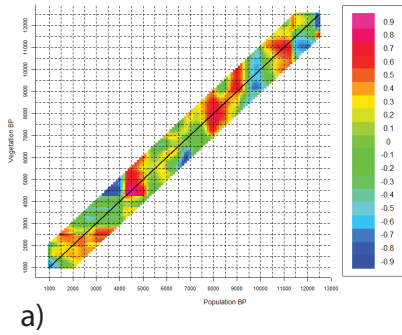


l)

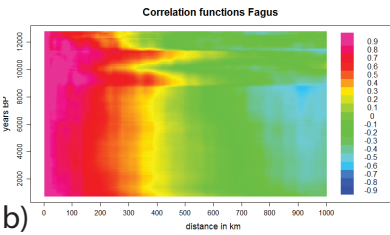


m)

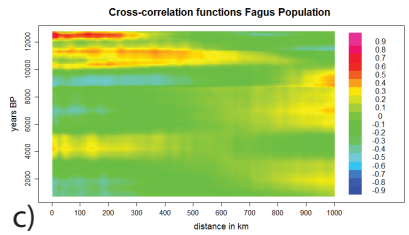
# Fagus



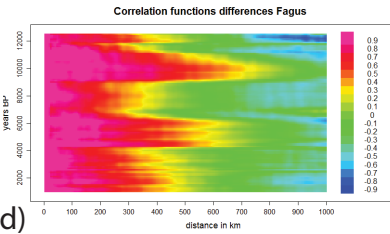
a)



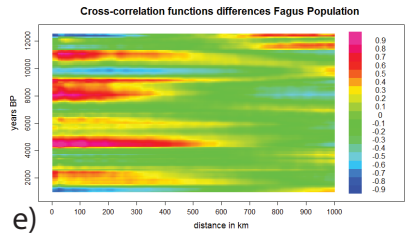
b)



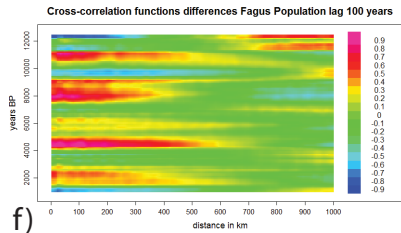
c)



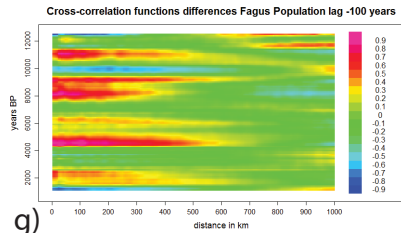
d)



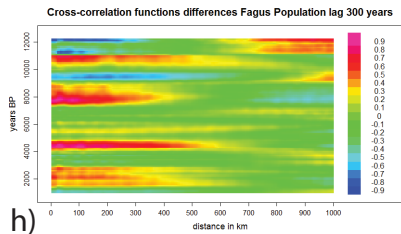
e)



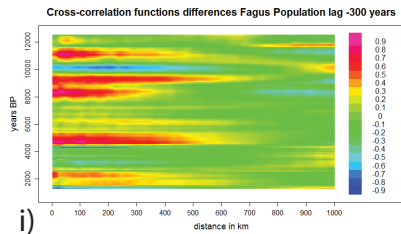
f)



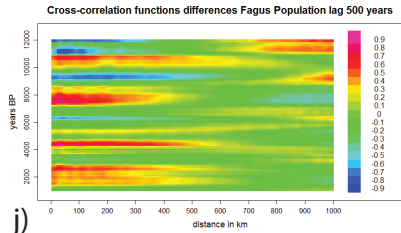
g)



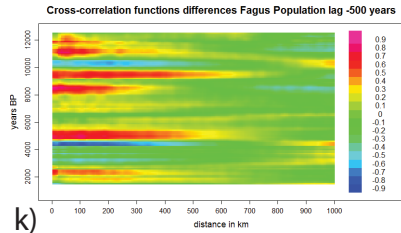
h)



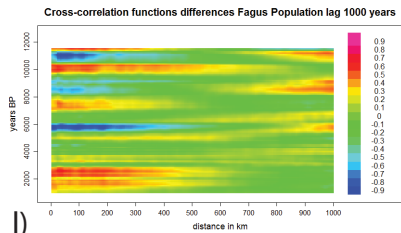
i)



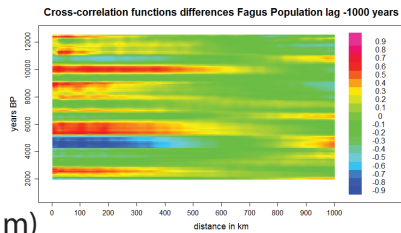
j)



k)

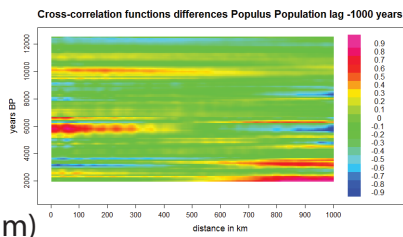
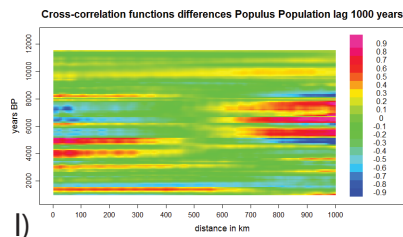
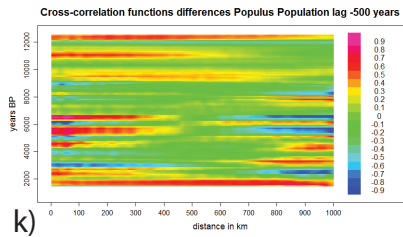
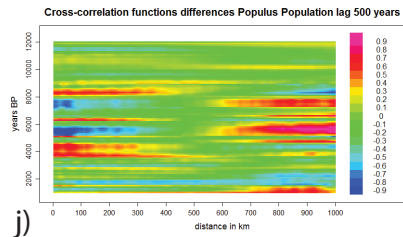
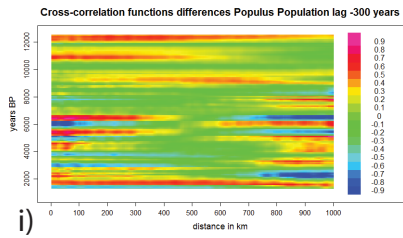
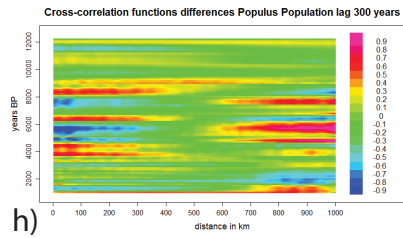
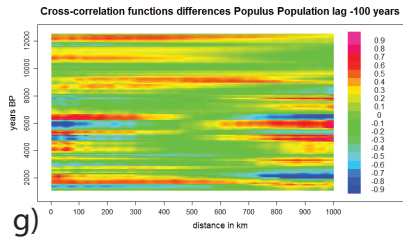
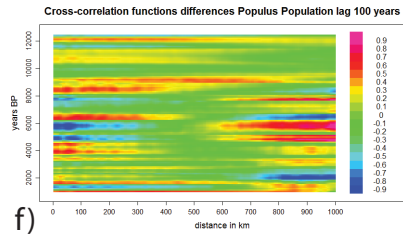
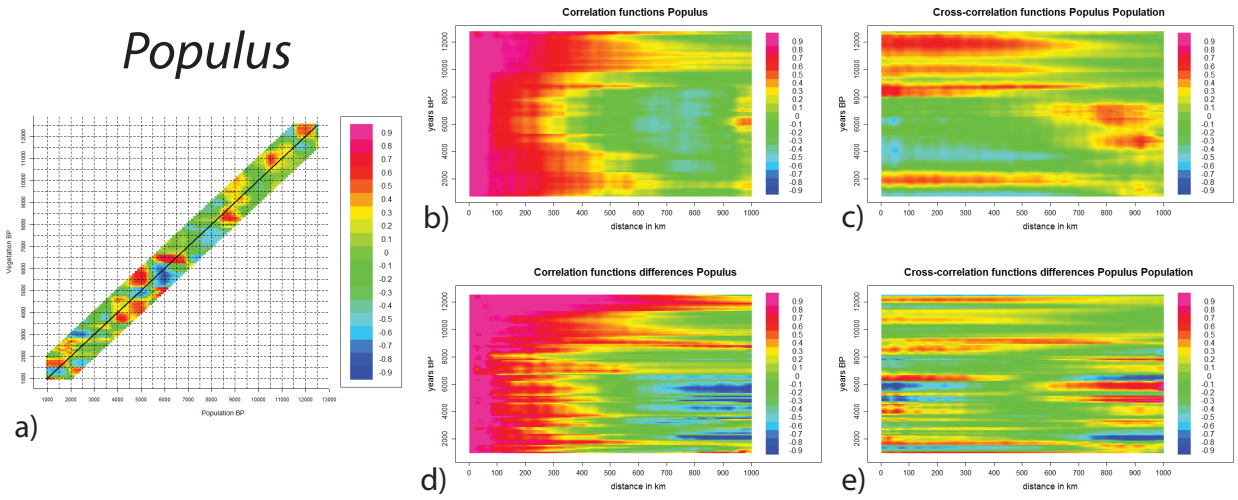


l)

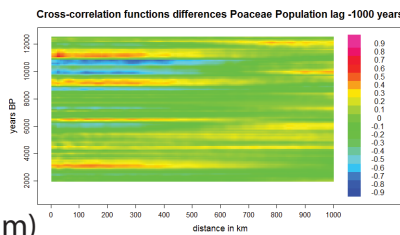
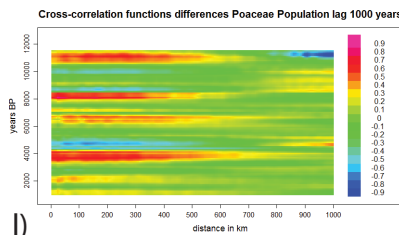
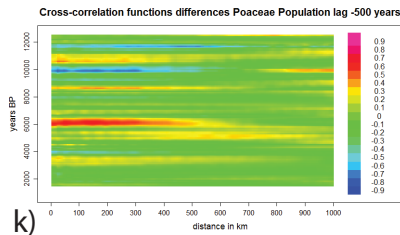
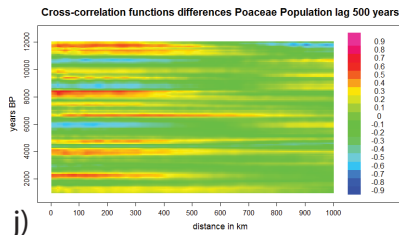
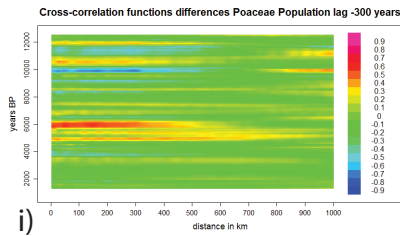
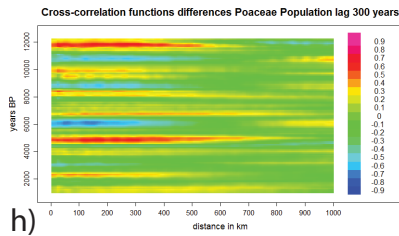
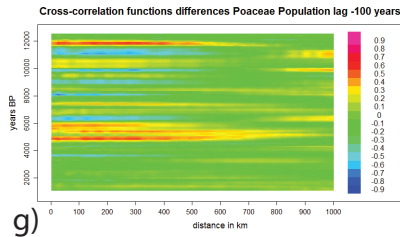
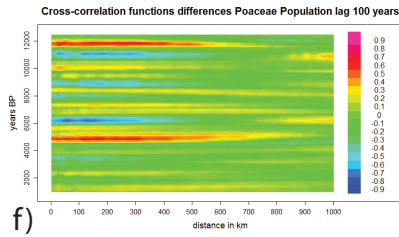
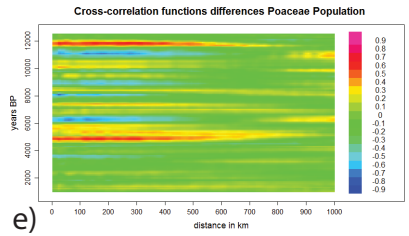
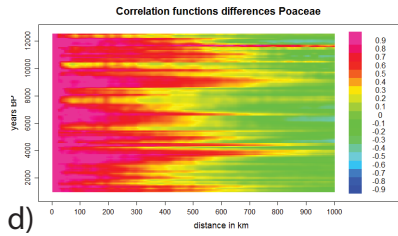
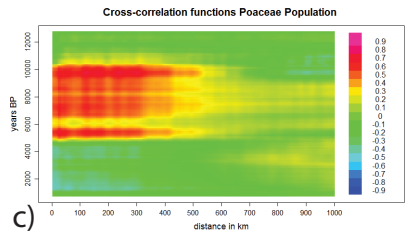
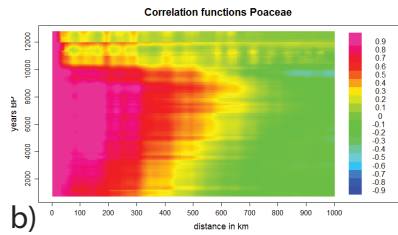
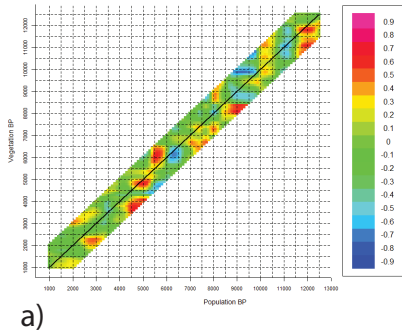


m)

# Populus



# Poaceae



## Appendix 4 (Videos for Chapter 5)

Videos available at <http://www.lpc.uottawa.ca/members/chaput/Animations.html>

**Video A4.1** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Acer*

**Video A4.2** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Carya*

**Video A4.3** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Castanea*

**Video A4.4** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Fagus*

**Video A4.5** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Juglans*

**Video A4.6** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Picea*

**Video A4.7** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Pinus*

**Video A4.8** Pollen abundances through time with data for both the entire continent and the area restricted to the range of Poaceae

**Video A4.9** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Populus*

**Video A4.10** Pollen abundances through time with data for both the entire continent and the area restricted to the range of *Quercus*

**Video A4.11** Estimated pollen percentages for *Acer* with population estimates within the taxon range

**Video A4.12** Estimated pollen percentages for *Carya* with population estimates within the taxon range

**Video A4.13** Estimated pollen percentages for *Castanea* with population estimates within the taxon range

**Video A4.14** Estimated pollen percentages for *Fagus* with population estimates within the taxon range

**Video A4.15** Estimated pollen percentages for *Juglans* with population estimates within the taxon range

**Video A4.16** Estimated pollen percentages for *Picea* with population estimates within the taxon range

**Video A4.17** Estimated pollen percentages for *Pinus* with population estimates within the taxon range

**Video A4.18** Estimated pollen percentages for Poaceae with population estimates within the taxon range

**Video A4.19** Estimated pollen percentages for *Populus* with population estimates within the taxon range

**Video A4.20** Estimated pollen percentages for *Quercus* with population estimates within the taxon range

**Video A4.21** Population and pollen changes (*Acer*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.22** Population and pollen changes (*Carya*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.23** Population and pollen changes (*Castanea*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.24** Population and pollen changes (*Fagus*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.25** Population and pollen changes (*Juglans*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.26** Population and pollen changes (*Picea*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.27** Population and pollen changes (*Pinus*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.28** Population and pollen changes (Poaceae) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.29** Population and pollen changes (*Populus*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years

**Video A4.30** Population and pollen changes (*Quercus*) that correspond to intervals that are shifted by the temporal lags 0, +100, +200, ..., +1000 years, and -100, -200, ..., -1000 years