HUMANS UNDER CLIMATE FORCING:

HOW CLIMATE CHANGE SHAPED HUNTER-GATHERER POPULATION DYNAMICS IN EUROPE
30,000–4000 YEARS AGO

Miikka Tallavaara

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Arts of the University of Helsinki, for public examination in auditorium XII, University main building, on 19th December 2015, at 12 noon.

Helsinki 2015
Supervisors

Doc. Tuija Rankama
Department of Philosophy, History, Culture and Art Studies,
University of Helsinki, Finland

Prof. Mika Lavento
Department of Philosophy, History, Culture and Art Studies,
University of Helsinki, Finland

Reviewers

Prof. Robert L. Kelly
Department of Anthropology, University of Wyoming, USA

Prof. Stephen Shennan
UCL Institute of Archaeology, University College London, United Kingdom

Opponent

Prof. Robert L. Kelly
Department of Anthropology, University of Wyoming, USA

© Miikka Tallavaara (Summary paper)
© Elsevier Ltd (Paper I)
© Authors (Paper II)
© SAGE Publications (Paper III)
© Archaeopress (Paper IV)
© National Academy of Sciences of the United States of America (Paper V)

ISBN 978-951-51-1766-3 (pbk.)
ISBN 978-951-51-1767-0 (PDF)
http://ethesis.helsinki.fi

Unigrafia
Helsinki 2015
ABSTRACT

Knowledge of prehistoric human population dynamics and its drivers is important for the understanding of cultural and social evolution. Working within the human ecological framework, this study aims to contribute to that knowledge, by reconstructing hunter-gatherer population dynamics in Europe 30,000–4000 years ago and by exploring the role of climate change in population size fluctuations. Archaeological reconstructions of population dynamics in Pleistocene Europe and Holocene Finland are based on spatio-temporal distributions of archaeological radiocarbon dates, which are taken as a proxy of human activity in time and place. The reliability and validity of the population history reconstructions are evaluated by studying potentially biasing effects of research history, taphonomic loss of archaeological material, and radiocarbon calibration.

In addition to making use of these archaeological methods, this study aims to develop and evaluate systematic means to use ethnographic data and palaeoclimate model simulations to reconstruct prehistoric hunter-gatherer population dynamics. This climate envelope modelling approach is used to simulate changes in population size and range in Europe between 30,000 and 13,000 years ago, and also to a lesser extent in Holocene Finland.

The results suggest that archaeological reconstructions based on the distribution of radiocarbon dates are not determined or strongly affected by biases related to research history. Instead, the reconstructions appear to reflect a true demographic signal from the past. However, radiocarbon calibration introduces high-frequency variation in the reconstructions, which has to be taken into account before any demographic interpretations are made. Due to this non-demographic variation, the method may not be able to reliably detect short-term variation in past population size and it is thus currently better suited to tracking long-term trends in population history. The taphonomic loss of archaeological material can potentially have a strong impact on the distribution of archaeological radiocarbon dates, but the current methods of taphonomic correction may not sufficiently take into account the geographical variability in taphonomic factors. In the future, it is therefore important to further develop taphonomic correction methods.

The ability of the climate envelope model simulation of human population to replicate archaeological patterns indicates that this novel approach is suitable for studying long-term hunter-gatherer population dynamics. The method allows not only the exploration of relative changes in population size, but also the estimation of absolute population density and size. It may also be able to detect potential inadequacies in the geographical distribution of archaeological data.

In the Finnish data, the correlation between the archaeological population size reconstruction and palaeoclimatic data suggests that the climate was an
important driver of long-term hunter-gatherer population dynamics, and that population appears to have changed in equilibrium with climate. The impact of the climate on human population was mostly indirect, mediated by its impacts on environmental production and consequently on food availability. The important role of climate is also supported by the correspondence between archaeological population reconstruction and the climate envelope model simulation of past human population size, which assumes long-term population dynamics to be in equilibrium with the climate. This correspondence also suggests that the impact of the climate on terrestrially adapted hunter-gatherer population dynamics has remained relatively constant from the Late Pleistocene to the present.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all of those who have supported me over the seven (!) years for completing my PhD study. Firstly, I would like to thank my supervisors, docent Tuija Rankama and professor Mika Lavento of the University of Helsinki, who have offered me trust, guidance and encouragement during this work. Tuija has been a close and valuable mentor already since the beginning of my journey in archaeology, for which I’m very appreciative.

I would also like to thank my co-authors, professor Heikki Seppä, professor Miska Luoto, professor Heikki Järvinen, docent Markku Oinonen, Esa Hertell, Natalia Korhonen, Mikael A. Manninen and Petro Pesonen, without whom I would not been able to complete my thesis. Esa, Mikael and Petro are also thanked for their friendship throughout the years we have known each other. Christopher TenWolde is thanked for revising the language of my thesis, Wesa Perttola for so willingly advising me in the numerous MapInfo issues, and my room mates at the University of Helsinki, Sanna Kivimäki, Satu Koivisto, Teemu Mökkönen and Kati Salo for the joyful company and stimulating discussions in the office and elsewhere.

I’m grateful to my thesis reviewers, professor Robert L. Kelly and professor Stephen Shennan for their encouraging comments on my work. It has been a real honour to have my thesis examined by persons, whose work I so greatly admire.

I appreciate the financial support from Finnish Graduate School in Archaeology, Finnish Academy of Science and Letters, and Kone Foundation. This support made my dissertation possible. Especially, the interest and trust given to my research by the people at Kone Foundation was very encouraging at the early stages of this work.

Finally, I would like to thank my parents, Kari and Marja-Sisko, who have always supported me and my, perhaps a bit odd, choice of career. Most of all, however, I would like to thank my beloved wife Meri, because of your love, support, insightful discussions, and especially, for so patiently standing those numerous occasions when I have been mentally away from you.
## CONTENTS

Abstract ................................................................................................................... 3  
Acknowledgements ................................................................................................. 5  
Contents .................................................................................................................. 6  
List of original publications .................................................................................... 7  
Author’s contribution ............................................................................................ 8  

1 **Introduction** .................................................................................................. 10  
1.1 Human population dynamics and cultural change - a brief review .................. 11  
1.2 Ecological causes of population size changes .............................................. 16  
1.3 Reconstructing prehistoric population history ............................................. 17  
1.4 Aims of the thesis .......................................................................................... 22  

2 **Materials and methods** ............................................................................ 23  
2.1 Distribution of radiocarbon dates as a human population proxy .................. 23  
2.1.1 Evaluation methods for different biases .................................................... 28  
2.1.1.1 Research bias ...................................................................................... 28  
2.1.1.2 Taphonomic bias .................................................................................. 32  
2.1.1.3 Calibration bias .................................................................................. 32  
2.2 Modelling of prehistoric hunter-gatherer population dynamics ...................... 34  
2.2.1 Climate envelope modelling approach ....................................................... 35  
2.2.1.1 Ethnographic training data ................................................................... 38  
2.2.1.2 Statistical model .................................................................................. 39  
2.2.1.3 Human population range and density simulation ....................... 40  
2.3 Palaeoenvironmental data .............................................................................. 42  
2.3.1 Palaeoclimatic and palaeoecological proxy data ....................................... 42  
2.3.2 Simulated climate data .............................................................................. 44  

3 **Results** .................................................................................................... 45  
3.1 Evaluation of archaeological population proxy .......................................... 45  
3.2 Correspondence between Holocene climate change and long-term population dynamics ........................................................................................................ 48  
3.3 Impact of event-like climate change on human population ....................... 50  
3.4 Climate envelope model simulation of Ice Age population dynamics in Europe .......................................................... 51  

4 **Discussion** ................................................................................................ 53  
4.1 Methodological considerations .................................................................... 53  
4.2 Impacts of climate .......................................................................................... 57  
4.3 The differing dynamics of foragers and farmers ......................................... 61  

5 **Conclusions** .............................................................................................. 64  
References ........................................................................................................... 67  
Publications I–V
LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:


The publications are referred to in the text by their roman numerals.
AUTHOR’S CONTRIBUTION

I  MT initiated the study and MT, PP and MO designed it. PP and MO provided the archaeological data. MT analysed the data and wrote the paper with contributions from PP. All authors provided editorial comments and approved the final manuscript.

II All authors initiated and designed the study. PP and MO provided the data. MO analysed the data and wrote the paper with contributions from PP and MT. All authors provided editorial comments and approved the final manuscript.

III MT initiated the study and MT and HS designed it. MT provided the archaeological data and HS provided the palaeoecological data. MT analysed the data and wrote the paper with contributions from HS.

IV MT and PP initiated the study and MT, PP, MAM designed it. PP, MAM and EH provided the archaeological data. MT analysed the data with contributions from MAM. MT wrote the paper. All authors provided editorial comments and approved the final manuscript.

V MT and HS initiated the study and MT, HS and ML designed it. MT prepared the archaeological and ethnographic data. NK and HJ provided the model-based climate data. MT and ML built predictive models, analysed their predictive accuracy and predicted presence and density values. MT constructed ensemble forecasts of the models, analysed modelling results, compared the model to the archaeological data and interpreted the results. MT wrote the paper. All the authors provided editorial comments and approved the final version of the manuscript.
INTRODUCTION

How many people were there in prehistory? This is a common question asked by laymen to archaeologists, but archaeologists usually struggle to give a reasonable estimate. Yet, there is a pressing need to be able to reconstruct prehistoric population history, because the question is also scientifically important: recent efforts on model building and testing have reaffirmed the importance of human population dynamics in our socio-cultural and linguistic evolution (e.g. Bromham et al. 2015; Derex et al. 2013; Henrich 2004; Kempe and Mesoudi 2014; Kline and Boyd 2010; Powell et al. 2009; Shennan 2001). The importance of demography in social and cultural change can be felt in modern society as well. In Finland, for example, the sustainability gap in public finances caused by changes in population size and age structure is used as a reason to tear down the structures of the welfare state. In addition to cultural and societal change, human population size can potentially have impact on the destiny of other species as well: growing populations of anatomically and behaviourally modern humans have been partly responsible for past ecosystem changes such as the extinctions of Pleistocene megafauna (Lorenzen et al. 2011) and Neanderthal humans (Mellars and French 2011) and current human population growth is probably leading towards a state shift in the Earth’s biosphere in the near future (Barnosky et al. 2012).

If knowing how human population size has changed through time is important, it is equally important to know what factors have affected human population size, because these factors are the ultimate causes of the cultural and ecological changes triggered by human population size. This study addresses both of these questions: it is about reconstructing how human population size has varied throughout prehistory, and about exploring the causes of changes in population size. The main focus of the study is on hunter-gatherer populations in Europe between c. 30,000 and 4000 years ago. In addition to reconstructing long-term population dynamics, it aims to explore whether climate, which is one of the most important determinants of global ecosystems, has affected hunter-gatherer population dynamics. Such questions regarding demography and human-environment interactions were recently listed as part of the 25 grand challenges for archaeology (Kintigh et al. 2014a,b).

From the theoretical point of view, the approach of the study is ecological, because insights from human life history theory (Hill and Hurtado 1996), behavioural ecology (Kelly 2013) and macroecology (Burnside et al. 2012; Hamilton et al. 2012) provide the best tools to understand hunter-gatherer population dynamics.
1.1 HUMAN POPULATION DYNAMICS AND CULTURAL CHANGE - A BRIEF REVIEW

Shennan (2000) has argued that human population dynamics is the single most important factor in understanding cultural change. In archaeology and anthropology, human population size has indeed featured as an important explanatory variable from the beginning of the New Archaeology in the 1960s onwards. These early theories were influenced by Boserup’s (1965) ideas of the imbalance between a population and its resources being a cause for economic intensification and change. According to Boserup (1965), economic intensification is a result of human population growth that will increase population size to the point to which further population growth would cause food shortage. This stress or pressure would then lead to economic innovations and intensification that would allow for a new cycle of growth. Because such innovation and intensification usually implies increasing labour input relative to yields, it is assumed not to happen spontaneously without the pressure to innovate.

Although Boserup’s original population pressure theory considered, first and foremost, agricultural intensification, it was soon generalised to cover changes among hunter-gatherers as well. Binford (1968) and Flannery (1969) suggested that pressure in the marginal environments of the Near East forced hunter-gatherers to broaden their diet to include previously marginal food sources, leading eventually to the deliberate tending and cultivation of plant species. Cohen (1977) generalised even further by arguing that the slow but steady population growth and filling up of suitable environments during the Late Pleistocene and Early Holocene led eventually, if not totally in sync, to the global imbalance between the hunter-gatherer population and its resources. This prehistoric food crisis resulted in increased dietary breadth and the invention and adoption of agriculture. On a more local scale, Dumond (1972) provided example of population pressure-induced subsistence changes among Eskimo hunter-gatherers in Alaska.

In addition to subsistence changes, population pressure has been seen as a cause of social change and especially for the rise of hierarchies. An important concept here is circumscription, which was featured prominently for the first time in Carneiro’s (1970) theory of the origins of the state. According to Carneiro, population growth within an area of circumscribed agricultural land, set off by mountains, seas, or deserts, leads to competition between previously autonomous groups. As a result, the strongest group subjugates others, because in their circumscribed environment the groups do not have any real option to avoid dominance by voting with their feet. This creates hierarchies between people and groups, and as the process is repeated, it results in larger and larger political groups and, in some cases, in the formation of states. In addition to environmental circumscription, Carneiro (1970) highlights social or, rather, demographic circumscription that can occur even in areas where environmental circumscription is not a
problem. This means that as suitable habitats are filled up, groups are circumscribed by other groups and therefore their options to avoid the dominance-seeking group are restricted. The process of demographic circumscription can also be seen as a more general mechanism leading not just to state formation, but also to the rise of hierarchies and social and political inequality within and between groups. Quite early on, population induced infilling and circumscription was considered as an important factor behind social changes among hunter-gatherer populations as well (Cohen 1977, 1981).

Another important idea related to the rise of hierarchical social organisations is scalar stress, as introduced by Johnson (1982). Scalar stress forms when the number of decision making units, such as households, increases within the group, and the coordination of cooperation and information flow becomes increasingly difficult. This stress is relieved either through group fissioning or the development of hierarchical social organisation. According to Johnson (1982), egalitarian sequential, or horizontal, hierarchies form when smaller units (e.g. nuclear families) form new, higher level units (e.g. extended families) that are larger and fewer in number than previous units. However, the increasing group size and consequent scalar stress may bring the sequential hierarchy to its limit, and if group fissioning is not an option due to circumscription, a new form of non-egalitarian vertical hierarchy can arise to facilitate decision making and implementation. In such systems, certain individuals or groups are associated with leadership functions and statuses as well as with some degree of control over resources, which development, according to Johnson (1982), is required for the coordination and regulation of their utilisation. The combination of circumscription and scalar stress has been used to explain the rise of hierarchical hunter-gatherer societies in north-western North America, as well as worldwide (Ames 1985; Cohen 1985).

Many of these basic ideas are still very much alive in the current understanding of the role of human population size and density in behavioural change. However, due to the rise of the role of evolutionary theory in anthropology and archaeology since the early 1980s, interaction between the human population and its resources and forms of social organisation are nowadays more often explicitly analysed within an evolutionary ecological framework (e.g. Bayham et al. 2012; Hertell 2009; Janetski 1997; Jerardino 2010; Kennett et al. 2009; Prentiss et al. 2014; Stiner et al. 2000).

With the rise of the evolutionary paradigm, the understanding of how population size can affect cultural change has diversified. Cultural transmission models, originally adapted from population genetics, help us understand how and why frequencies of different cultural variants change through time (Boyd and Richerson 1988). These new models and insights allow us to analyse cultural evolution in a way roughly analogous to genetic evolution. Shennan (2001) and Henrich (2004) have built models to analyse
the role of population size in cultural transmission and the creation of cultural variation. Their models indicate that large pools of interacting individuals can create and maintain adaptive skills more effectively, and are also capable of faster cumulative cultural evolution than small populations. A decrease in population size may, in turn, result in a loss of complex cultural traits. Thus, the effects of population size on cultural variation would be roughly similar to its effects on genetic variation (Frankham 1996). Shennan (2001) and Powell et al. (2009) link the punctuated appearance of modern behavioural traits between 100,000 and 50,000 years ago to changes in human population size and density, whereas Henrich (2004) has suggested that the gradual loss of cultural traits observed in the archaeological and ethnological records of Tasmania was caused by a reduction in population size. This reduction was a result of the separation and isolation of Tasmanian populations from the larger Australian metapopulation at the end of the Pleistocene (Henrich 2004). These models have been tested in experimental laboratory studies, which found strong support for the idea that the size of the pool of interacting individuals has an impact on cultural evolution (Derex et al. 2013; Kempe and Mesoudi 2014; Muthukrishna et al. 2014). In addition, a test in the natural context using ethnohistorical data from Oceania found that population size best explains the complexity of fishing technologies among the island-living populations (Kline and Boyd 2010).

From early on, theories that postulate an important role for human population size in behavioural and cultural change have been criticised. In order to evoke technological and social innovations, it has been thought that population-resource-imbalance has to be relatively severe and represent a long-lasting state, but this important assumption has been questioned (Cowgill 1975; Cowgill and Wilmsen 1975). This state of affairs is also problematic from the population ecological point of view as the relationship between a population and its resources are usually assumed to be in equilibrium, even though this can also mean stable oscillation (Hanski et al. 1998). Population pressure arguments have been found especially troubling with respect to hunter-gatherers who, according to a popular idea, are argued to actively regulate their population size well below carrying capacity. This active regulation, for which infanticide is the most effective means, is seen as a major reason for the near zero population growth assumed to characterise most of the Pleistocene (Cowgill 1975; Cowgill and Wilmsen 1975; Hassan 1981: 144).

However, assumption of active regulation is problematic. Firstly, as the tendency to maximise (inclusive) fitness has been under strong selection throughout the evolution of all species, it is theoretically highly unlikely that humans would voluntarily restrict their reproductive behaviour for the common good (Shennan 2002: 102–117). Secondly and more importantly, there is no empirical evidence that birth control mechanisms, such as infanticide, would have actively been used to control the population size, or that these would in fact have had an inhibitive effect on population size.
Indeed, quite the opposite may be the case. Smith and Smith (1994) analysed Inuit sex-ratio data that has been used to argue for infanticide and they indeed found evidence for some preferential female infanticide. Although Smith and Smith (1994) were not able to directly test the population regulation hypothesis, they found it unconvincing both theoretically and in the light of the available data. Instead, Smith and Smith (1994) suggested that preferential female infanticide among the Inuits is best explained by the differential payback sons and daughters are able to contribute to their parents’ inclusive fitness in the socio-ecological context of the Inuits. In addition, Lee (1972) has noted that the workload of !Kung San mothers increases significantly if the interbirth interval is shorter than four years, because shorter intervals require mothers to carry much greater loads of baby and food on her foraging trips. Thus, the interbirth interval of four years, by whatever means achieved, is linked to mother’s personal wellbeing and endurance, not necessarily to any intentional attempts to regulate population size. Blurton Jones (1986) has further argued that the four years interbirth interval appears to maximise the reproductive success of !Kung women, most likely because the strain caused by a shorter interval could lead to increased offspring mortality (but see Hill and Hurtado 1996: 380–385). Thus, by increasing mothers reproductive success, birth control among !Kung does not constrain population growth, but can rather accelerate it.

Due to the above mentioned reasons, the common idea of an extremely low or even zero population growth rate during most of human prehistory is also problematic. Ethnographic data indicate no hunter-gatherer groups that would have had a zero growth rate (Hill and Hurtado 1996: 471; Pennington 2001) and it has been suggested that human life history has evolved to be much more “r-selected” (population size is governed by maximum reproductive capacity) than the life histories of other great apes, also allowing much more rapid population growth (Hill and Hurtado 1996: 472). If this potential for rapid growth characterises hunter-gatherers, one has to assume that periods of growth were frequently followed by crashes in order to explain the inferred slow or zero long-term population growth during the Pleistocene (Boone 2002; Hill and Hurtado 1996: 471).

Even if humans, like other species, have evolved a tendency to maximise their reproductive success, and thus are not likely to constrain their reproductive behaviour for the common good, population pressure arguments may still appear problematic. If, due to the density-dependent factors, population tends to be in equilibrium with its resources or oscillate around some “quasi equilibrium”, a key question becomes: under what circumstances would the imbalance between a human population and its resources be sufficiently high to trigger economic innovations such as agriculture? A stable population at the equilibrium would hardly indicate population pressure, otherwise it would not be in equilibrium. However, Winterhalder et al. (1988; Winterhalder and Lu 1997) and Belovsky (1988) suggest that a common type of dynamics among hunter-gatherers would be
stable limit cycle, where the population and its dietary breadth oscillates around some mean with the wavelength varying between 50 to 100 years. This kind of density-dependent oscillation results from the interplay between the density of the human population and its resources. Human population grows fast when resources are abundant, but at the same time the density of the resources start to decline. The declining abundance of resources slows the growth of human population and eventually turns the growth into decline. As human population declines, resources recover, which eventually allows new human population growth, and so forth. At the peaks of the cycle there is, indeed, population pressure and the widest diet, but the pressure starts to be relieved rather quickly, within less than 50 years. This appears to be a short time for technological innovations to occur, especially as the archaeological record usually indicates rather gradual economic changes (see also Bettinger et al. 2010).

Despite the potential problems related to population-resources-imbalance arguments, it is still possible that under favourable conditions of resource abundance, population growth can lead to environmental and/or demographic circumscription, causing the above-mentioned problems in the coordination of co-operation and dominance avoidance, and consequent changes in social organisation, even without an imbalance between a population and its resources. However, Hayden (2001) has questioned the role of population size in the rise of non-egalitarian social organisation as well. According to Hayden, population size and unequal social relations can be correlated, but they are not causally related, because they are both affected by resource abundance. Thus, high population size or density would not be a necessary condition for inequalities to arise and the correlation is only spurious. Instead, the surplus hypothesis argues that non-egalitarian social organisation would arise as a result of the action of dominance-seeking individuals who use abundant resources for their own advantage by creating alliances, gaining prestige, and creating debts (Hayden 2001). The same surplus can lead to population increase.

Also, the theories that propose an important role for population size in cultural transmission have been criticised on theoretical (Querbes et al. 2014; Read 2008) and empirical (Collard et al. 2005, 2011, 2013a,b; Read 2006, 2008, 2012) grounds. Collard et al. have argued that in the ethnographic hunter-gatherer data, variation in toolkit complexity is not explained by population size or mobility, but subsistence-related risk, which is measured by effective temperature (Collard et al. 2005, 2011, 2013a,b). Perhaps the most vocal critic of the role of population size in cultural variation has been Read (Read 2006, 2008, 2012). Using ethnographic data, he has argued that hunter-gatherer toolkit complexity is explained by the combined effects of growing season length and residential mobility, not by population size (Read 2006, 2008, 2012). According to Read, this implies that variation in toolkit complexity is “driven by the response of a hunter-gatherer group to ecological constrains through its mode of resource procurement” (Read...
2008: 620). Later, Read (2012) has generalised his critique and argued that the variation in fishing technology in Oceania is also better explained by subsistence-related risk rather than population size (but see the discussion in the supplementary material of Read 2012; see also Henrich 2006).

All the above-mentioned strongly indicates that population size and density have a potentially very important role in cultural change, but that their relative importance is far from resolved. This means that the role of demography remains an important and interesting research topic, and that such research will likely generate new insights into human cultural and biological variability in time and space.

1.2 ECOLOGICAL CAUSES OF POPULATION SIZE CHANGES

To fully explain cultural change in cases where population size or density turns out to be an important factor, one has to search for factors that affect population size. From the human ecological perspective, the study of these factors is also an important aim unto itself. Changes in population size are determined by a complex interplay between the number of births and deaths and in and out migration. Current theoretical understanding of human demography is best formulated in life history theory. One of its fundamental assumptions is that because our decision making facility (brains) has evolved through natural selection, fitness maximisation is expected to be the goal (conscious or unconscious) of decision makers. This theory predicts that resource availability would affect the severity of so called life history trade-offs, such as the one between the maximum number of offspring that can be produced and the maximum number that can reach the stage of being successful parents themselves (Hill and Hurtado 1996: 18–35; Shennan 2009). Increasing the amount of resources would decrease these severities and eventually lead to population expansion as a result of increased fertility and higher offspring survival. For humans, this link between resource availability and reproductive success seems to hold, at least in pre-industrial settings, where the amount of resources available to parents is indeed shown to correlate with their fertility and with the survival of their offspring (Hill and Hurtado 1996: 293–320; Pettay et al. 2007; Rickard et al. 2010). The availability of resources often dictates migration-related decisions as well. At the population level, Baumhoff (1963) was able to show that the density of important food resources was predicative of hunter-gatherer population size in California.

The issue can also be considered from a more macroecological perspective. Early on, Bartholomew and Birdsell (1953) argued that on a long-term basis the population size of human hunter-gatherers is in thermodynamic equilibrium with the trophic levels below it, because the energy for growth and reproduction for hunter-gatherers comes from these
lower trophic levels. Thus, hunter-gatherer population depends on the production of primary producers, either directly or via higher-order producers. In terrestrial environments, primary production is mainly controlled by temperature and hydrology, i.e., climate: as the climate changes, primary production changes as well. Therefore, it is reasonable to assume that, on time scales typically studied by archaeologists, mean hunter-gatherer population size varies with climate. As climate has a strong influence on the patterning of plant and animal species distributions and on ecosystems in general (Anderson-Teixeira and Vitousek 2012; Currie 1991; Holdridge 1967; Stephenson 1990), it is difficult to understand why things would be different with humans, who are themselves part of the ecosystem and depend on other species. It has already been convincingly argued, for example, that hunter-gatherer mobility patterns and territorial requirements are linked to climatic factors (Binford 1980; Hamilton et al. 2012; Kelly 2013: 88–94). It is useful to keep in mind that the relationship between climate and other aspects of an ecosystem may not be linear, because the rates of any biochemical reactions increase exponentially with temperature (Anderson-Teixeira and Vitousek 2012).

However, several archaeological studies have not found a strong link between climate and hunter-gatherer population size or cultural change (e.g. Fiedel and Kuzmin 2007; Gamble et al. 2005; Gamble 2005; Jochim 2012; Meltzer and Bar-Yosef 2012). The idea that human population history and thus, at least in some cases, cultural change would have been controlled by environmental factors can also be ideologically hard to accept and is easily deemed as environmental determinism, which is considered inherently wrong. Culture, and especially technology, is often thought to liberate humans from such environmental constrains. During the heydays of post-modern archaeology in the 1980s and 1990s, all kinds of ecological approaches were dismissed because they were seen to impose modern capitalistic ideology on prehistoric people (Johnson 1999: 146–148). However, from the scientific point of view, it is important to study how integral a part of the ecosystem humans actually are, rather than to assume it a priori.

1.3 RECONSTRUCTING PREHISTORIC POPULATION HISTORY

One of the most important challenges in the study of prehistoric population dynamics is how to reconstruct past population trajectory. The ability to reliably reconstruct population histories is an obvious prerequisite for the study of the causes and possible effects of changes in population size and density. Problems with these reconstructions left the early theories potentially open for severe criticism. As Renfrew pointed out already in 1973, in order to avoid circularity, one needs evidence of population change that is
independent of the type of economies or social organisations that one seeks to explain with population size (Renfrew 1973). Even today some studies use measures such as changes in dietary breadth or prey species morphology and population structure to trace changes in human population size and density (Klein and Steele 2013; Stiner et al. 2000; Stute et al. 2009). Such arguments require that these changes are determined only by human population size or density, an assumption that is hardly justified.

In 1953, Bartholomew and Birdsell stated that “Some archaeologist have hoped to reconstruct preagricultural population figures from studying the temporal and spatial distribution of sites, but the inescapable sampling errors in this approach render it unreliable. We suggest that an analysis of the energy relationships and efficiency of the techniques for obtaining food offer a promising approach” (Bartholomew and Birdsell 1953). Their suggestion has indeed formed one approach of reconstructing past population figures in archaeology, either by relating known human energy requirements to the energy available for human consumption in particular environments, or by projecting into the past the densities of ethnographically known populations who are assumed to live in environmentally and technologically analogous conditions to their prehistoric counterparts. The former alternative provides estimates of the maximum density or size (density × area) of human population a given area is able to support (e.g. Casteel 1972; Hassan 1981). However, it requires highly detailed information on past environments, which is still difficult to obtain. It also makes far-reaching assumptions about the harvestable biomass suitable for human consumption in different environments. Furthermore, it is difficult to know how to relate maximum population to actual realised population (Casteel 1972; Kelly 2013: 184–185).

Ethnographic analogy overcomes some of these problems by providing estimates of realised population densities in certain environmental conditions. The use of analogies, nevertheless, requires some consideration. For example, Eller et al. (2009) used a single ethnographically informed density estimate of 0.4 individuals per 100km² to estimate the global Middle Pleistocene human census population size. It is without saying that this is a totally unreasonable approach. Even when the correspondence between the environmental conditions of ethnographic and prehistoric populations is justified (e.g. Gräslund 1974), one cannot be sure how representative the population density of the chosen ethnographic case(s) is for the given environmental conditions: at the time of the ethnographic documentation, the density of the hunter-gatherer population could have been very different from the long-term mean due to its naturally oscillating dynamics.

In general, the use of ethnographic data in estimating prehistoric population density and size has been rather unsystematic. The availability of global databases of ethnographically documented hunter-gatherers (Binford 2001; Kelly 2013) now allows a more systematic approach. Fitting a mean function of population density over chosen environmental gradients yields
the most likely estimate of population density in given environmental conditions. Such an approach takes into account the fact that individual ethnographically documented groups could have been at different stages of their population cycles at the time of the ethnographic documentation. Actually, this approach is no longer based on ethnographic analogy. Instead, through inductive model building, it informs us how chosen environmental variables affect human population density. This information or theory can then be used to estimate past population densities, if the past states of environmental variables are known (see also Binford 2001: 447–464).

Despite the scepticism of Bartholomew and Birdsell (1953), the temporal and spatial distribution of archaeological material has become the most commonly used measure for prehistoric population size. This might be simply a result of more and more archaeological research being done: as the archaeological dataset grows, sampling error becomes less severe. The number of sites (sometimes together with size of the sites) through time was already used in the late 1960s and early 1970s as a proxy, i.e. an indirect measure, of population size (e.g. Flannery 1969; Hole et al. 1969; Renfrew 1972; Sanders 1972). These data were usually based on surveys. Survey data can be problematic, however, because they do not necessarily yield temporally diagnostic artefacts, and even when present the diagnostic artefacts often do not cover the whole period the site was in use. In areas where postglacial rebound has been strong, it is possible to date sites using shore displacement chronology independently of diagnostic artefacts. The temporal distribution of sites dated using shore displacement has also been used as a proxy for human population size (Siiriäinen 1981), but obviously its geographical applicability is rather restricted.

![Diagram](image_url)

**Figure 1.** The use of the distribution of $^{14}$C dates as a proxy for relative human population size is based on the simple idea that there should be a positive correlation between the number of people and the amount of waste, i.e. archaeological material, they produce. Large populations leave more sites, hearths, tools, refuse pits, etc. behind them than smaller populations.
Rick (1987) was the first to measure changes in prehistoric population size with the temporal distribution of archaeological radiocarbon dates (\(^{14}\text{C}\) dates), thus avoiding the problems related to the lack of sufficiently high-resolution artefact typologies. The use of the distribution of \(^{14}\text{C}\) dates as a proxy is based on the same idea as the use of distribution of sites: It is assumed that there is a positive correlation between the number of people and the amount of dateable contexts (sites, hearths, refuse pits, graves, etc.) they produce (Fig. 1). Contrary e.g. to ethnographic analogy, this kind of proxy can only give relative, not absolute, estimates of population size; i.e., it can tell when population size was larger and when smaller, but it cannot assign any actual population size (numbers) to the estimate.

Since Rick’s (1987) pioneer study, the method has been increasingly used, especially during the last ten years, which can be seen in the development of citations to Rick’s article (Fig. 2). Figure 2 also shows the number of citations to some other influential papers. The reasons for the recent popularity of the method may, again, relate to the increased availability of radiocarbon dates, not least as a result of some large projects (Van Andel et al. 2003; Gamble et al. 2004, 2005), and to the increased theoretical interest in population size as an explanatory variable in cultural change.

![Graph](image)

**Figure 2.** The development of citations to six influential studies that use distributions of \(^{14}\text{C}\) dates as a human population proxy. Citation data from the ISI Web of Science.

Despite its current popularity, the method is not without its problems. These relate to the factors that can disturb the link between the number of prehistoric people and the strength of their archaeological signal as measured by the distribution of \(^{14}\text{C}\) dates. These biases include, for example, variable research interests that affect what is dated, taphonomic loss of
archaeological material through time, the different rates of archaeological material produced by mobile and sedentary people, differential cultural practices dealing with organic (dateable) materials, and the biasing effects of $^{14}$C calibration. Rick (1987) was already aware of these biases, and carefully evaluated their potential impact on the data. With the new research boom, the amount of studies evaluating the method and developing means to filter out the impacts of different biases, as well as developing the method in general, has increased (e.g. Ballenger and Mabry 2011; Bamforth and Grund 2012; Brown 2015; Downey et al. 2014; French and Collins 2015; Grove 2011; Peros et al. 2010; Rhode et al. 2014; Shennan et al. 2013; Steele 2010; Surovell et al. 2009; Surovell and Brantingham 2007; Timpson et al. 2014; Williams 2012). It is likely that the awareness of these biases and the eagerness to develop the method are the reasons why a couple of attempts to completely dismiss the method have not been very successful (Contreras and Meadows 2014; Mökkönen 2014). A large number of potential biases (Mökkönen 2014) nevertheless ensures that there is a constant need to evaluate and develop the method.

Bartholomew and Birdsell (1953) were hardly able to envision the most recent developments in reconstructing prehistoric population dynamics, namely the ones based on the analysis of variation in human DNA. Skyline-plot methods (Atkinson et al. 2008, 2009) and pairwise or multiple sequentially Markovian coalescent analyses (Li and Durbin 2011; Schiffels and Durbin 2014) use coalescent theory to infer changes in effective population size from genetic data. However, effective population size does not have a straightforward relationship with the actual census population size (Hawks 2008). In addition, these methods depend on estimates of DNA mutation rate and molecular clock calibrations, which are still debated (Fu et al. 2013; Scally and Durbin 2012) and imprecise, leading to poor temporal resolution. Furthermore, at least skyline-plot methods may not be able to detect demographic histories that go beyond a major reduction in population size (Grant et al. 2012). This is evident in the Bayesian skyline-plots of marine species that show flat demographic curves prior to the Last Glacial Maximum (LGM), when a distinct population decline is assumed to have occurred (Grant et al. 2012). Similar flat patterns characterise skyline-plot curves of pre-LGM human populations (Atkinson et al. 2008; Zheng et al. 2012). All these problems make it difficult to meaningfully compare DNA-based population reconstructions with the records of cultural and environmental changes. Until significant improvements in the DNA-based method, reconstructions based on archaeological and ethnographic data must be considered more reliable.
1.4 AIMS OF THE THESIS

In his book Genes, Memes and Human History, Shennan (2002: 112) defines the tasks of archaeology of population as follows: 1) to characterise regional population patterns through time, 2) to identify the factors affecting these, and 3) to examine the impact of population size and density on other aspects of human activity and social institutions. This thesis contributes to the first two of the suggested tasks. By building on the understanding and methodological developments described above, and by taking into account the open questions, this thesis aims to:

1. use archaeological radiocarbon dates to reconstruct the prehistoric population history of eastern Fennoscandia between 11,000 and 1000 cal BP;
2. evaluate the reconstruction by comparing it with other, spatially and temporally more restricted population proxies and by studying the effects of research history, taphonomic loss, and $^{14}$C calibration on the distribution of radiocarbon dates;
3. develop a systematic approach of utilising ethnographic data to reconstruct hunter-gatherer population history within the climate envelope modelling approach, and to use this approach together with model-based climate data to reconstruct long-term population dynamics in Europe between 30,000 and 13,000 cal BP;
4. study the effect of climate and climate-related environmental factors on human population dynamics. This is achieved by:
   a) comparing the reconstruction of Fennoscandian population history with the local palaeoclimatic and palaeoenvironmental records of long-term and event-like environmental changes that would have been relevant in terms of human population dynamics;
   b) comparing the climate envelope model simulation of past human population size, which assumes long-term population dynamics to be in equilibrium with the climate, with the archaeological population proxy in order to see how realistic the simulation is.

The main focus of the thesis will be on hunter-gatherer populations. This allows to look at extremely long-term dynamics and, thus, to see possible recurrent features in the data. The study areas in Holocene eastern Fennoscandia (mainly Finland) and Late Pleistocene Europe are well suited to these questions: Pleistocene humans in Europe were purely hunter-gatherers, and eastern Fennoscandia contains one of the longest Holocene hunter-gatherer records in Europe and the best available palaeoclimatic and palaeoenvironmental data. This arrangement also allows to study whether the impact of the climate on human populations has been different under glacial and interglacial climate regimes.
2 MATERIALS AND METHODS

2.1 DISTRIBUTION OF RADIOCARBON DATES AS A HUMAN POPULATION PROXY

The archaeological radiocarbon analyses performed at the Dating Laboratory of the Finnish Museum of Natural History (University of Helsinki) and its predecessors form the backbone (c. 80%) of the radiocarbon dataset that is used in reconstructing population history in eastern Fennoscandia. The dataset was also extended to cover, as thoroughly as possible, the published archaeological radiocarbon measured elsewhere. In addition, the data also contain unpublished dates that were released for use by many Dating Laboratory customers. At the time of the first publications (papers I and II), the whole database consisted of almost 2600 individual $^{14}$C dates, extending from the earliest colonization of the area (c. 11,000 cal BP) to the modern era. When compiling the database, it was decided not to make any a priori exclusive selections of dates. This means that the dataset also includes dates that potentially have no link to human activity due to erroneous sample selection in the field. However, such erroneous selections are expected to form a clear minority within the dataset, and their role is merely random, not systematic. Furthermore, exclusion of dates that do not seem to match the archaeological material of the dated context can bias the real signal of human activity. For example, the Hossanmäki site in Lohja (southern Finland) yielded an archaeological assemblage that can be dated on typological and technological grounds exclusively to the Stone Age (Pesonen and Tallavaara 2006). However, all the radiocarbon dates from burnt bone and hearths showed much younger ages, extending from Pre-Roman Iron Age to the Medieval Period (Pesonen and Tallavaara 2006). The exclusion of these dates because of the lack of an apparent link between the dates and archaeological assemblage, would have led to a biased result in terms of the depiction of human activity in the region (see also Shennan et al. 2013).

In paper V, the archaeological proxy of the European Palaeolithic population is based on 3718 $^{14}$C dates from 895 sites. Most of the dates are extracted from the INQUA Radiocarbon Palaeolithic Europe Database v12 (Vermeersch 2005). The dataset has been expanded by including dates from recent publications. The dataset was critically evaluated using the information given in the INQUA database. Dates were excluded according to several criteria: a) all dates that were qualified as unreliable or contaminated, b) dates without coordinates or laboratory reference, c) duplicate dates, d) dates with standard errors greater than 5% of the mean $^{14}$C age, e) dates from gytta, humus, peat, soil or soil organics, organic sediment, humic acid fraction of the sediment, and fossil timber, f) dates of marine origin, such as shell, marine shell, and molluscs, g) dates without a clear link to human
activity, such as terminus ante and post quem, surface, above, up from, top, below, and beneath the cultural layer(s), minimum or maximum age of the layer, and beyond site, and h) dates of cave bear (Ursus spelaeus), that may have no link to human activity. In some cases, coordinates or even ages were corrected according to the original publication of the date.

The whole Fennoscandian dataset and its suitability for reconstructing population history was evaluated in paper II, whereas smaller subsets of the data were used in papers I, III, and IV. To study the prehistoric population history in eastern Fennoscandia, only dates that are older than 800 \(^{14}\)C years were used. The fact that differing research emphases may cause some sites, or certain chronological phases of a particular site, to have significantly more dates relative to other sites or phases, can distort population-related conclusions based on date distributions (Gamble et al. 2005). Shennan and Edinborough (2007) tackled this problem by calculating a pooled mean of all the dates from a single chronological phase and constructing the population curve using these mean dates. Because there are few stratigraphically distinct phases in eastern Fennoscandia, a slightly different procedure was used in this study (papers I, III). First, each site’s dates were arranged in an ascending order. Then the first cluster of dates that fell within an arbitrary interval of 200 radiocarbon years was combined. After that, it was continued towards the oldest date and combined the second cluster of dates falling within the interval, and so on. For example, if a site originally had the following six dates: 4000 bp ± 80, 4120 bp ± 50, 4200 bp ± 95, 4500 bp ± 80, 5100 bp ± 45 and 5250 bp ± 50, the first three and the last two dates would have been combined. Combinations of dates were carried out using the weighted average method in CalPal calibration program (Weninger and Jöris 2004). After the combination, the following three dates would represent the site in a combined database: 4105 bp ± 39, 4500 bp ± 80 and 5167 bp ± 33. When all sites were handled this way, the combined database contained 1160 individual dates. A similar approach to reduce the effect of site phases with multiple dates has now been used in other studies as well (Shennan et al. 2013).

When studying the earliest colonization of eastern Fennoscandia, a slightly different approach was followed (paper IV). Because the studied time period was narrower, and because the number of Mesolithic \(^{14}\)C dates between 10,900–8500 cal BP is relatively small (N=107), it was decided to use the dates without combining them. Since the focus in paper IV is on the impacts of a short-lived climate event, it was necessary to use archaeological \(^{14}\)C dates that would accurately date human activity on site. Therefore, if a site yielded charcoal and burnt bone dates, the latter were prioritised in the reconstruction of spatio-temporal population patterns due to their smaller own-age effect (see paper II).

However, it is well known that there is an offset between marine and atmospheric carbon reservoirs that affects radiocarbon dates made on marine samples, such as seal bones. There is considerable spatial and
temporal variation in the marine reservoir effect values in different parts of the world, and the values for the different stages of the Baltic Sea are currently unknown or only partly understood (e.g. Eriksson 2004; Hedenström and Possnert 2001; Lindqvist and Possnert 1999; Olsson 2006). On the other hand, research on the origin of the carbon source in burnt bone indicates transfer of CO$_2$ from the fuel material into the bone during combustion (Hüls et al. 2010; Olsen et al. 2013; Strydonck et al. 2010; Zazzo et al. 2011). This result has obvious implications for the dating of burnt bone samples, because it suggests that the marine reservoir effect is counterbalanced to some degree by the transfer of CO$_2$ from wood used as a fuel. At the same time, terrestrial bones may yield dates that are too old due to the “old wood effect” and thus decrease the age difference between charcoal and bone deriving from the same context. However, unless the exchange of bone carbonate with CO$_2$ is 100%, burnt bone should give more accurate dates of human activity than charcoal. Several examples where calcined bone and charcoal from the same context have been dated show a clear age correspondence between these two sample materials, so that bone dates are systematically slightly younger than charcoal dates (e.g. Lanting et al. 2001; Manninen and Tallavaara 2011; Olsen et al. 2008).

The Fennoscandian radiocarbon database was also geographically divided into three subsets. The border between the southern and central areas approximates a cultural border that is observable at some points of prehistory, and especially in the ethnographic material (Carpelan 1999; Sarmela 1994; Vuorela 1976) as well as in the genetic data (Hedman et al. 2004; Lappalainen et al. 2006; Salmela et al. 2008). The border between the central and northern areas, on the other hand, is more arbitrary, but it nevertheless approximates the southern limit of the historically known territories of Sami societies (e.g. Näkkäläjärvi 2003: 115). Paper III used only the combined dates from the southern and central regions (N=869).

To construct a proxy curve for the relative human population size, radiocarbon dates were calibrated and individual probability distributions were summed using CalPal (paper I) and OxCal (paper III) calibration programs (Bronk Ramsey 2009; Weninger and Jöris 2004) and IntCal04 and IntCal09 calibration curves (Reimer et al. 2004, 2009). This summing produces a summed probability distribution (SPD) of calibrated dates. This method is still widely used in studies of prehistoric population size (Collard et al. 2010; French and Collins 2015; Gamble et al. 2005; Hinz et al. 2012; Kelly et al. 2013; Riede 2009; Shennan et al. 2013; Shennan and Edinborough 2007; Smith et al. 2008; Timpson et al. 2014; Wang et al. 2014; Williams 2013). However, the summing of individual probability distributions causes problems if standard errors of $^{14}$C measurements vary, which is usually the case. Measurements with a smaller standard error show more sharply peaked (higher kurtosis) calibrated probability distribution than measurements with greater standard error (see also Bamforth and Grund 2012). This is because the same probability mass is distributed within
Materials and methods

a narrower range. Thus, the summing of dates with small and large errors creates a curve that indicates higher relative population size during periods of small-error-dates, and lower population size during periods of large-error-dates, even if the number of dates is equal during both periods. There is also the possibility of systematic bias as standard errors tend to grow with age, so that Holocene and Late Pleistocene patterns in SPDs would appear different only because of systematic change in standard errors, not because of different population dynamics.

This problem is illustrated in Figure 3, which shows the SPDs of 21 evenly spaced (50 ${}^{14}$C yrs) simulated radiocarbon dates. In the first set of dates (blue), the standard error is 80 years for each date. In the second set (red), ${}^{14}$C ages are exactly the same as in the first set, but four dates have a standard error of 40 years (Table 1). These four dates create a very distinct peak in the red distribution, even though the number and spacing of ${}^{14}$C dates is the same as in the grey distribution. Thus, SPD may not be the best way to illustrate temporal distributions of radiocarbon dates.

Figure 3. The changes in standard errors affect the shape of summed probability distributions (SPD), but not the shape of temporal frequency distributions. Narrow standard errors create peaks in the SPD. In the temporal frequency distribution of calibrated median dates such peaks do not show up. The data used in the figure is shown in Table 1.

In addition to SPD, Figure 3 shows the kernel density curve of the distribution of calibrated median dates of the red set. This curve correctly depicts the distribution without any peak, therefore avoiding the problem related to varying standard errors. Unlike SPD, the distribution of calibrated medians, weighted averages or modes produces a distribution of dates based on temporal frequency, even when it is illustrated as a kernel density plot. It is thus a better representation of a population proxy, as a higher frequency (or density) really means a higher number of dates, which can be interpreted as a higher relative population size. In SPD, high probability density can just indicate dates with smaller standard errors.
Table 1. The dates used to create Figure 3.

<table>
<thead>
<tr>
<th></th>
<th>Blue set</th>
<th></th>
<th>Red set</th>
<th></th>
<th>Black set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Std</td>
<td>14C Age</td>
<td>Std</td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td>80</td>
<td>5000</td>
<td>80</td>
<td></td>
<td>5746</td>
</tr>
<tr>
<td>5050</td>
<td>80</td>
<td>5050</td>
<td>80</td>
<td></td>
<td>5797</td>
</tr>
<tr>
<td>5100</td>
<td>80</td>
<td>5100</td>
<td>80</td>
<td></td>
<td>5830</td>
</tr>
<tr>
<td>5150</td>
<td>80</td>
<td>5150</td>
<td>80</td>
<td></td>
<td>5903</td>
</tr>
<tr>
<td>5200</td>
<td>80</td>
<td>5200</td>
<td>80</td>
<td></td>
<td>5973</td>
</tr>
<tr>
<td>5250</td>
<td>80</td>
<td>5250</td>
<td>80</td>
<td></td>
<td>6043</td>
</tr>
<tr>
<td>5300</td>
<td>80</td>
<td>5300</td>
<td>80</td>
<td></td>
<td>6087</td>
</tr>
<tr>
<td>5350</td>
<td>80</td>
<td>5350</td>
<td>40</td>
<td></td>
<td>6133</td>
</tr>
<tr>
<td>5400</td>
<td>80</td>
<td>5400</td>
<td>40</td>
<td></td>
<td>6221</td>
</tr>
<tr>
<td>5450</td>
<td>80</td>
<td>5450</td>
<td>40</td>
<td></td>
<td>6248</td>
</tr>
<tr>
<td>5500</td>
<td>80</td>
<td>5500</td>
<td>40</td>
<td></td>
<td>6300</td>
</tr>
<tr>
<td>5550</td>
<td>80</td>
<td>5550</td>
<td>80</td>
<td></td>
<td>6352</td>
</tr>
<tr>
<td>5600</td>
<td>80</td>
<td>5600</td>
<td>80</td>
<td></td>
<td>6390</td>
</tr>
<tr>
<td>5650</td>
<td>80</td>
<td>5650</td>
<td>80</td>
<td></td>
<td>6441</td>
</tr>
<tr>
<td>5700</td>
<td>80</td>
<td>5700</td>
<td>80</td>
<td></td>
<td>6497</td>
</tr>
<tr>
<td>5750</td>
<td>80</td>
<td>5750</td>
<td>80</td>
<td></td>
<td>6551</td>
</tr>
<tr>
<td>5800</td>
<td>80</td>
<td>5800</td>
<td>80</td>
<td></td>
<td>6601</td>
</tr>
<tr>
<td>5850</td>
<td>80</td>
<td>5850</td>
<td>80</td>
<td></td>
<td>6663</td>
</tr>
<tr>
<td>5900</td>
<td>80</td>
<td>5900</td>
<td>80</td>
<td></td>
<td>6726</td>
</tr>
<tr>
<td>5950</td>
<td>80</td>
<td>5950</td>
<td>80</td>
<td></td>
<td>6787</td>
</tr>
<tr>
<td>6000</td>
<td>80</td>
<td>6000</td>
<td>80</td>
<td></td>
<td>6847</td>
</tr>
</tbody>
</table>

In papers IV and V, the population proxy is based on the distribution of calibrated median dates, and in this summary paper all shown 14C-based population reconstructions are based on median dates calibrated using the IntCal13 calibration curve (Reimer et al. 2013) and clam calibration algorithm (Blaauw 2010) in R statistical software (R Development Core Team 2014).

In papers IV and V, the spatial distribution of 14C dates is used to make inferences about the range of human population. The idea is that the calibrated median age reflects human presence at a certain place at a certain time. In paper V, dates were binned using 1000 year bins. When evaluating spatial distributions of dates, one has to keep in mind that such distributions are relatively prone to research biases (see below).
2.1.1 EVALUATION METHODS FOR DIFFERENT BIASES

There are several factors that may bias interpretations based on the temporal frequency distributions of $^{14}$C dates. The potential impacts of these factors must be considered before variation in the temporal frequency distribution can be interpreted in terms of past population history. The most notable biasing factors are the effects of varying research interests, the taphonomic loss of dateable archaeological material, and the effect of calibration of $^{14}$C dates.

2.1.1.1 Research bias

One of the most important assumptions in the use of temporal frequency distribution of archaeological material as a population proxy is that the known sites represent a random sample of all the existing sites in relation to their age. In Finland, this situation is more tenable, since most surveys are carried out by cultural heritage management agencies, mainly the National Board of Antiquities (NBA). Usually cultural heritage management does not have a scientific interest or focus on certain periods. The same applies to excavations as well, which usually are rescue excavations at sites determined by present land use rather than by scientific questions. In the case of radiocarbon dates, one has to assume further that the researchers have also made the radiocarbon determinations randomly in relation to their age. This is not as plausible as the previous assumption, since it is likely that varying research emphases will have an effect here. The biasing effect of research interests is strongest in small samples, whereas in larger samples the effect of different research interests may cancel each other out. Therefore, the large European-wide dataset of glacial $^{14}$C date used in paper V is not evaluated against research biases. This kind of evaluation would also have been difficult, due to the lack of sufficient information easily available for those dates.

It is likely that the spatial distribution of dates is more biased than the temporal distribution. This is because archaeological activity is mainly determined by land use intensity: areas with higher intensity would show a higher density of archaeological sites. However, this kind of spatial bias should not affect temporal distributions if archaeological data are a sufficiently random sample of the temporal dimension of the true archaeological signal. This is demonstrated in Figure 4. Otherwise, possible spatial biases are not evaluated in this study.

The possible biasing effects of research interests are studied by comparing the temporal frequency distribution of $^{14}$C dates from eastern Fennoscandia to other independent archaeological population proxies, by studying how the distribution might have changed throughout the history of radiocarbon dating in Finland, and by studying how the distribution might vary between different date submitter classes.
Ceramic site frequency index
The frequency of sites dated typologically using ceramic finds is used as an alternative measure of population size. From c. 7200 cal BP onwards, Finnish Stone Age and Bronze Age (c. 3500–2500 cal BP) sites are most easily recognised via pottery finds, which form the basis for the chronological and geographical division of the archaeological cultures. It should be noted that the beginning of pottery use in hunter–gatherer populations did not indicate a shift towards productive economies. This means that pottery was adopted and used in Mesolithic conditions. In Finland, this era is often referred to as “Subneolithic,” contrary to the true farming Neolithic.

Figure 4. Simulated example showing how modern land use intensity could bias the spatial, but not temporal distribution of the archaeological sample. In the left panel of A, the true signal is uniformly distributed over the total area. Because of modern land use, the spatial distribution of the archaeological sample is systematically biased and does not reflect the true signal. However, despite this spatial bias, the temporal distribution in the archaeological sample reflects the bimodal distribution of the true signal. The figure is based on simulated data that have three variables: age, longitude (x), and latitude (y). In the horizontal panel A, coordinates follow a uniform distribution (within the defined area) and in B, y follows normal and x follows gamma distribution. In A and B, age follows the temporal distribution shown in the left panel (true signal panel). The left panel shows the complete data. For the right panel (archaeological signal), the data was sampled. In A, the sample was biased so that lower latitudes (smaller y) were more likely to be included in the sample. In B, the data were randomly sampled.
The ceramic database was constructed by Pesonen during the late 1990s for educational purposes (Pesonen 1999). The database is a compilation of information from several sources, the most important being the collections of the National Museum of Finland. Originally, the database contained c. 6000 catalogue numbers, including Neolithic, Early Metal Age, and Iron Age ceramics. The database was later modified so that each site has as many entries as it has recognizable ceramic types. This made it possible to calculate the number of sites containing each ceramic type. As the Iron Age collections were not studied, and the number of sites is based only on the published data, it is obvious that the database cannot be used as a proxy for the Iron Age population levels. Thus, the frequency distribution is presented only for its Subneolithic–Bronze Age portion.

Because most ceramic types overlap both chronologically and geographically, the number of sites could not be used directly in the alternative proxy. Therefore, the number of sites having a certain ceramic type was divided by the length of the period of use (in hundred years) of this type. In this way, an average number of sites per century for each ceramic period was derived. These average numbers were then summed together for each century. The proxy, the Ceramic Site Frequency Index (henceforth CSF-Index), is formed from these calculations, and is coarser than the distribution of radiocarbon dates. For example, because of the averaging procedure, it does not take into account changes in frequency that might have occurred within a certain ceramic period.

**Other alternative proxies**

In addition to the CSF-index, the eastern Fennoscandian $^{14}$C date-based proxy is compared to other published archaeological and non-archaeological indicators of human population size: Siiriäinen (1981) presented a distribution of coastal sites dated using shore displacement chronology. He assumed that the distribution (number of sites per 100 years) reflects changes in human population size. Siiriäinen’s (1981) curve of site frequency is presented here so that the time boundaries of different cultural periods are updated according to the current knowledge.

Recently, Sundell (2014; Sundell et al. 2014) used the number of stone artefacts and number of stone artefact types in a given period as a human population proxy. Here, the artefact frequency and number of types in a given period are divided by the length of the period to account for the varying durations of different cultural periods.

In his study about the beginning of agriculture in Finland, Hertell (2009) used the percentage of seal bones in coastal assemblages as a proxy for human population density. This is based on the fact that in the ethnographic record there is a positive correlation between hunter-gatherer population density and the proportion of aquatic resources in the diet (e.g. Binford 2001; Kelly 2013). In this summary paper the same idea is used, but with a larger dataset (osteological archives compiled by Pirkko Ukkonen and Kristiina
Mannermaa at the Finnish Museum of Natural History and unpublished osteological reports at the National Board of Antiquities) and a different periodisation: Early Mesolithic (11,000–8500 cal BP), Late Mesolithic (8500–7200 cal BP), Early Neolithic (7200–6000 cal BP), Middle Neolithic (6000–5400 cal BP), and Late Neolithic (5400–3500 cal BP). Assemblages were dated using information from the ceramic and \(^{14}\)C date databases described above. Here the mean percentage of marine mammal bone fragments per period is used as a proxy (see also Tallavaara et al. 2014).

Genetics provide data on population history that is totally independent of archaeological data and methods. In this study, archaeological proxies are compared to the timing of the genetic bottleneck inferred from Finnish genetic data, and dated using a molecular clock (Sajantila et al. 1996). To put it simply, a genetic (or population) bottleneck means that before and after the bottleneck, the effective population size has been larger than at the bottleneck (for more about prehistoric population bottlenecks see Sundell 2014; Sundell et al. 2014).

**Historical changes in research interest**

The data produced by the Dating Laboratory of the Finnish Museum of Natural History and its predecessors (N=1979) includes the most complete information on, for instance, sample submitters and submission dates within the total data set, thus allowing to study the variation through time and between different date submitter classes.

Radiocarbon dating in Finland started during the late 1960s. To study the possible changes in research interests through time, uncalibrated \(^{14}\)C dates were divided according to the decade during which the sample was submitted to the laboratory (1968–1979, 1980–1989, 1990-1999, 2000–2008). If research interests or other systematic effects have changes over the years, the shape of the date distribution should vary over the decades.

**Variation between sample submitter classes**

The same subset of the data was used to study the possible differences between different sample submitter classes, by classifying the data into four mutually inclusive classes. These are: a) Individual NBA-submitted dates not belonging to any larger date series. These are assumed to form the least biased distribution, because in their studies NBA should not have major a priori preferences set over any archaeological period. b) The dates submitted by the 16 most active submitters. c) The remaining 127 sample submitters (there are altogether 143 different sample submitters). d) The dates submitted by the Early in the North project, which was the largest individual project that had submitted dates prior to 2010.

If there is a marked research bias affecting the temporal frequency distribution of \(^{14}\)C dates, this should show up as a variation between the above-mentioned classes.
2.1.1.2 Taphonomic bias

Surovell et al. (2009) argue that younger findings are over-represented relative to older findings in the archaeological record due to the time-dependent influence of destructive processes such as erosion and weathering. Similar time-dependent loss-processes seem to affect geological and palaeontological data (Surovell and Brantingham 2007), as well as historical coin records (Peros et al. 2010).

Surovell et al. (2009) propose a model of taphonomic bias that is based on the observation of the deviation between terrestrial and ice core records of volcanic activity. Terrestrial record shows a clear time-dependent decrease in the number radiocarbon dates associated with volcanic deposits, whereas ice core records of volcanism do not show such a time-dependent pattern. Since the ice core record does not suffer from taphonomic loss due to constant accumulation, this difference between records suggests that the long-term trend in the terrestrial record is governed by taphonomic loss of volcanic deposits.

The model can be used to make a taphonomical correction to the temporal frequency distribution of the dates. The model gives the number of sedimentary deposits preserved from the given time. Dividing the number of archaeological dates from this time by the number of surviving deposits gives the taphonomically corrected, unstandardized value for the number of dates.

In paper II, a slightly modified version of this model was used, based only on the Holocene dates. Williams (2012) published yet another modified model of the taphonomic bias. His version of the model corrects the increasing deviation in residuals in the most recent and oldest time intervals, and unstable variance that changed over time. Williams (2012) also suggests that the original and his version of the model over-corrects the number of dates older than 25,000 cal BP (see also Williams 2013). Despite this possible over-correction, Williams’ model was used in paper V to correct the temporal frequency distribution of dates between 30,000 and 13,000 cal BP. In this summary paper, Williams’ model is also used to correct the distribution of dates from eastern Fennoscandia.

2.1.1.3 Calibration bias

Because the rate of radiocarbon production is not constant, the calibration curve that transfers $^{14}$C ages to calendar ages is not linear. This, in turn, has some problematic consequences: when a uniform distribution of calendar dates is converted to $^{14}$C ages and recalibrated after adding some constant or random standard error, the resulting distribution of calibrated dates (SPD or distribution of calibrated median dates) is far from uniform (Fig 5). Instead, calibration creates some packing of the dates and, consequently, also less dense parts in the distribution, even though the true distribution of calendar ages is uniform. It is suggested that sharp declines in the calibration curve
create distinct peaks in the probability distributions, whereas plateaus in the curve even out any peaks (Guilderson et al. 2005; Michczyński and Michczyńska 2006). However, as Figure 5 shows, the correlation between the shape of the calibration curve and resulting SPD or distribution of calibrated medians is not perfect (see also Brown 2015).

Nevertheless, all of this indicates that calibration procedure can create artificial peaks and troughs in the date distribution, which can mistakenly be interpreted in demographic terms. This is problematic, especially when the focus is on short-lived population events, such as the ones possibly related to rapid climate events. The fact that at least some of the rapid cold events coincide with the anomalies in the atmospheric $^{14}$C content (Hughen et al. 2000) makes this even more problematic. Recently, Shennan et al. (2013) suggested a computationally heavy simulation approach to evaluate the impact of calibration bias on SPD. In paper IV, a simpler, albeit less effective and accurate, method is used to evaluate calibration bias when studying the impact of the so-called 10,300 cal BP cold climate event on the early pioneer population in eastern Fennoscandia. The calibration bias is studied first by creating an artificial, evenly spaced (10 $^{14}$C yrs) set of dates with identical standard errors (80 yrs) covering the same span as the archaeological data. After calibrating the simulated set, it is possible to compare the peaks and
troughs of the simulated distribution of median dates with the peaks and troughs of the archaeological date distribution, and to determine which peaks and troughs in the archaeological distribution are merely calibration anomalies.

In this summary paper, calibration bias is studied differently from the approach used in paper IV, following a slightly simplified version of the Shennan et al. methodology (2013; see also Brown 2015): first, calendar dates are sampled (N=N of archaeological sample) from the distribution that follows the trend in archaeological date distribution (loess smooth with the span = 1). Such distribution takes into account the long term trend in the number of dates in the archaeological data, but leaves out all short-term fluctuations. Next, calendar dates are transferred to $^{14}$C age scale and calibrated after adding measurement error. These steps are repeated 200–1000 times to account for the sampling variation in calendar date sampling and in calendar-age-to-$^{14}$C-age-transfer. Finally, the 5th-to-95th percentile range is calculated. This range covers 90% of the simulations, excluding the highest and lowest 5% of the values in a given time bin. Such range, illustrated e.g. by a polygon, represents an archaeological signal of the truly monotonic growth pattern, which does not contain any major fluctuations. In other words, it indicates how the truly smooth distribution would show up in archaeological data if long-term growth rate, sampling, and calibration would have been the only factors affecting the shape of the distribution. Negative and positive deviations between this null model and the observed archaeological pattern are so rare, if the null model holds that they can be considered as statistically significant differences. On the other hand, if the simulated and archaeological distributions show corresponding short-term fluctuations, archaeological fluctuations may not represent a true demographic signal, but only bias related to calibration and/or sampling.

All these steps are performed using R statistical software (R Development Core Team 2014). Calibration is performed using the functions of clam algorithm in R (Blaauw 2010).

### 2.2 MODELLING OF PREHISTORIC HUNTER-GATHERER POPULATION DYNAMICS

As mentioned in Chapter 1.3, archaeological data can provide information mainly on the relative changes in past population size, and if quantitative estimates of population size or density are needed one must turn to other approaches. Modelling and simulations can provide one solution to this problem. At the same time, by providing a complementary line of evidence, they contribute to the fuller and more valid picture of prehistoric population patterns.

The tradition of predictive modelling of site locations is especially strong in archaeology, but modelling was also used to study socio-cultural processes
already in the 1970s, and during recent decades this has become even more important as computational power has increased (Lake 2014). Today, modelling and simulations have a strong foothold especially in the study of human-environment interaction (e.g. Crabtree and Kohler 2012).

### 2.2.1 CLIMATE ENVELOPE MODELLING APPROACH

The modelling approach used in this study (paper V, and in a much smaller and simpler scale in paper III) is a kind of a predictive modelling. It is based on the simple idea of using information on how climate affects population densities of ethnographically known hunter-gatherers to predict or hindcast (or retrodict) prehistoric hunter-gatherer population densities. Basically, the only information this approach requires about the past is the information about past climate conditions, which is possible to acquire by using palaeoclimate model data.

As shown in chapter 1.3, the idea of using ethnographic and/or environmental information to infer prehistoric population densities is not novel, but ethnographic data has only rarely been put to systematic use regarding this issue. Instead, the fields of biogeography and ecology have developed highly sophisticated quantitative methods of using information from the present to model future or past changes in species range and abundance. These climate envelope models, or niche models, are increasingly used in ecology to predict forthcoming climate-related changes in species distribution (Araújo and Peterson 2012; Pearson and Dawson 2003; Virkkala et al. 2013). They are also used in paleobiology to complement incomplete and coarse fossil-based species distributions, to provide hypotheses regarding past communities and their dynamics, and to assess the determinants of organism distributions (Nogués-Bravo et al. 2008; Svenning et al. 2011).

The aims of climate envelope modelling in palaeobiology (Svenning et al. 2011) translate easily to archaeology, where such modelling can provide:

a. quantitative and high resolution predictions of past human distributions
b. testable ecological hypotheses regarding human populations and their dynamics
c. quantitative assessment of the climatic determinants of human distribution

In addition to demography, the same approach can be used to model other aspects of human behaviour that potentially have link to climate, such as mobility, subsistence and technology.

In series of publications, Banks et al. (2008a,b, 2009, 2011, 2013) apply what they call eco-cultural niche modelling that aims to elucidate the influence of environmental factors on social and technical systems. Their approach is closely related to ecological niche modelling, but it differs from the modelling strategy used in this study. When modelling human range
during the Last Glacial Maximum in Europe, Banks et al. (2008b) use the spatial distribution of archaeological sites as a basis for their model, which they then validate by comparing it to the distribution of archaeological sites in a test set. This might introduce a risk of circularity, as the training and test sets are both sampled from the same population of archaeological sites. In this study (paper V), the model of human population size and range dynamics is built on ethnographic data, and is thus independent of the archaeological record, which allows the evaluation of the model against archaeological population proxy. This must be considered as a clear advantage.

Climate envelope models belong to the class of species distribution or niche models. According to classic definition, the fundamental ecological niche of a species comprises those environmental conditions within which a species can survive and grow (Hutchinson 1957). Climate envelopes, for their part, constitute the climatic component of the fundamental niche, and can therefore also be defined as a climatic niche (Pearson and Dawson 2003). Figure 6 illustrates the idea of the climate envelope by showing the envelope of human hunter-gatherers in terms of four climate variables.

![Figure 6](image.png)

**Figure 6.** Climatic envelope or niche of hunter-gatherers in terms of four climate variables. The size of the bubble is related to population density. The figure is based on Binford’s (2001) data.

It shows the climate conditions where ethnographic hunter-gatherers are present, and how abundant they are under different conditions. Climate envelope modelling uses that kind of information on how climate affects the current distribution of a species in order to hindcast past (or predict future) changes in the range and density of the species under the selected climate change scenario.
Climate envelope modelling requires (Araújo and Peterson 2012; Nogués-Bravo 2009; Pearson and Dawson 2003; Svenning et al. 2011):

1. Training or calibration data that includes information on current species distribution and current climatic conditions.
2. A statistical model, or several models, that are fitted to the training data (a model algorithm is trained using the training data). This model predicts species distribution as a function of chosen climate variables. It is a calibration model that transfers the values of climate predictors to population density and presence/absence, when past (or future) species distribution is simulated.
3. Climate data that describes past (or future) climate conditions. This climate data is input data to the statistical model, when past (or future) species distribution is simulated.
4. Independent validation data against which the realism of the simulation is evaluated. Not always available, but when the focus is on the past, archaeological or fossil records can provide such data.

These requirements for the model used in this study are described in more detail below.

All the species distribution models, including climate envelope models, assume an equilibrium between species distribution and environment. This means that species occurs everywhere the climate (or the environment in general) is suitable, and is absent from unsuitable areas. It is well known that this assumption can easily be violated, as a species’ range and abundance can be in disequilibrium with the environment because of dispersal limitations and/or interactions with other species (Nogués-Bravo 2009; Svenning et al. 2011).

It is clear, for example, that the geographical distribution of ethnographically documented hunter-gatherers does not reflect the area that is suitable for hunter-gatherers, because large areas previously occupied by foragers are dominated by agricultural populations from the Mid-Holocene onwards. However, as Binford (2001: 130–159) has shown, an ethnographic sample of hunter-gatherers reflects their niche variability reasonably well, because there is no bias regarding the plant community types or plant productivity despite the clear geographical bias: there are ethnographic hunter-gatherer populations living in niches that, in some parts of the world, are now occupied by agricultural populations.

An even more important assumption for the studies, where information on a present niche is used to simulate the past or future range and abundance of a species, is niche stability through time. Any considerable shift in a species’ niche will reduce model transferability across time (Nogués-Bravo 2009; Svenning et al. 2011). Through technological evolution, humans have been especially prone to expand their niche by adapting to new environments and increasing the productivity of already occupied areas. This poses a clear challenge to bioclimatic envelope modelling of prehistoric humans. However, in hindcasting studies, it is possible to compare model simulations to the
independent archaeological (or palaeobiological) record. Significant differences between simulated and archaeological distributions would indicate niche shift, provided that the archaeological record is not a completely biased sample of the true distribution.

### 2.2.1.1 Ethnographic training data

In the model construction process, the ethnographic dataset (N=339) of hunter-gatherer populations compiled by Binford (2001) was used to extract calibration data to train the statistical models. Among other things, this dataset includes information on population density, resource use, mobility, and climatic and other environmental conditions of hunter-gatherer populations. In paper III, a subset of the data containing hunter-gatherer populations from arctic to temperate zones were used as training data (N=198).

Because isotope studies of human bone collagen indicate that the Pleistocene hunter-gatherers obtained, at most, 30% of their dietary protein from aquatic resources (Drucker and Bocherens 2004; Richards et al. 2005), only populations whose main livelihood comes from terrestrial resources (SUBSP≠3) were selected for the training data in paper V, where focus is on the Ice Age human population. Increased use of aquatic resources is an example of niche expansion that occurred among some hunter-gatherer populations during the Holocene. Aquatic resources are essential for groups living in polar environments, and they enable generally higher population densities among hunter-gatherers (Fig. 7). In addition to the exclusion of aquatic hunter-gatherers, cases whose subsistence is based on mutualistic relations with non-hunter-gatherers (SUBPOP=X) and mounted hunter-gatherers (SYSTATE3=1) were excluded from the training data. Furthermore, to keep the simulated population densities conservative, populations that either move into and out of a central location that is maintained for more than one year, or are completely sedentary (GRPPAT=2) were excluded. Such groups usually live under high population densities. This exclusion means that the simulation assumes that the Pleistocene human populations in Europe were residentially fully mobile, an assumption commonly held by archaeologists.

Altogether, the training data includes information on 127 hunter-gatherer populations. Because this dataset gives information only on environments where the hunter-gatherers have existed in recent historical times, 120 pseudo-absence data points to the climate space, where terrestrially adapted hunter-gatherers have not existed during the historical period (e.g. extremely cold or extremely hot and dry), were added to enhance the performance of the statistical models (paper V). The climate data for these points were obtained from the WorldClim database (Hijmans et al. 2005). Addition of pseudo-absence data to presence-only data is a standard procedure in ecological modelling (Barbet-Massin et al. 2012; Phillips et al. 2009).
Figure 7. The adaptation to aquatic resources is an important factor that has enabled the colonisation of the coldest environments, and also the support of higher population densities. Data from Binford (2001).

Of the available climate variables, mean annual temperature (CMAT) was selected as a predictor of population density (DENSITY) in paper III focusing on the Holocene population history in eastern Fennoscandia. In paper V, potential evapotranspiration (PET), water balance (WAB), and mean temperature of the coldest month (MCM) were selected as predictors of the density and presence (DENSITY>0) / absence (DENSITY=0) of the Ice Age human population in Europe. PET and WAB were selected because they exert a strong influence on ecosystem productivity and species richness, and therefore also on food availability for human foragers, whereas MCM affects winter mortality (Currie 1991; Franklin 2010; Stephenson 1990). PET and MCM values are directly available from the ethnographic dataset. WAB-values were calculated as the difference between annual precipitation and PET.

2.2.1.2 Statistical model

In paper III, the climate envelope modelling approach was used in a simplified way. A calibration model was constructed simply by fitting a generalized linear model (McCullagh and Nelder 1989) with gamma error distribution and log link function to predict population density with annual mean temperature.

In paper V, significantly more sophisticated methodology was implemented. To model the distribution and density of the human population, two frameworks were used: One predicting the range (presence/absence) of the human population and the other predicting
population density. The human population occurrence was modelled as a binary response variable, and density as a continuous response variable. Because different modelling algorithms can give different predictions, six alternative state-of-the-art techniques were used to relate human presence/absence and density with the explanatory climatic variables. These techniques are generalized linear modelling (McCullagh and Nelder 1989), generalized additive modelling (Hastie and Tibshirani 1990), support vector machines (Cortes and Vapnik 1995; Drucker et al. 1997), classification tree analysis (Breiman et al. 1984), random forest (Breiman 2001), and generalized boosting methods (Elith et al. 2008). The final simulation of Ice Age population dynamics was achieved by averaging the resulting six predictions (see below).

In paper V, the models' predictive accuracy was assessed using cross validation. This means that the ethnographic training data was randomly split into training (70%) and validation (30%) subsets 100 times. In each round, a model was built using the training split and its ability to correctly predict population density and presence/absence was tested using the validation split. The predictive power of the binary models was determined by testing the accuracy of predictions made for the validation dataset (i.e. data that were not used in model development), calculating the area under the curve of a receiver operating characteristic plot (AUC) (Fielding and Bell 1997) and the true skill statistic (TSS) (Allouche et al. 2006). For density models, mean $R^2$ between predicted and true values were calculated.

2.2.1.3 Human population range and density simulation

In paper III, the aim of the simple modelling exercise was to produce a null model or simulation that would indicate how population density would have varied throughout the Holocene if it was determined by annual mean temperature alone. This was achieved by inputting the annual mean temperature values of a pollen-based temperature reconstruction (see below) into a calibration model that transferred temperature values into hunter-gatherer population densities. This simulation, which only includes a temporal component, was compared to an archaeological population proxy to determine if other factors than annual mean temperature would also have influenced long-term hunter-gatherer population dynamics, i.e., whether there are clear differences between simulated and archaeological patterns.

In paper V, the range of the human population for every 1000 years between 30,000 and 13,000 cal BP was simulated by predicting the presence/absence of humans for every $0.375^\circ$ by $0.25^\circ$ cell containing land area in Ice Age Europe. Eustatic changes in the sea level and the consequent changes in the land area of Europe were taken into account by adjusting the sea level according to the global sea level change curve (Peltier and Fairbanks 2006). The spatial and temporal resolution of this simulation is determined by the resolution of the climate model (see below) that was used to obtain the
values of climate predictors for the Ice Age Europe. The range simulation was
done by using the above-mentioned six calibration model algorithms and
climate predictor values derived from the climate simulation. The climate
simulation-based monthly average temperature and annual precipitation
values were used to calculate PET and WAB values. WAB was calculated as
the difference between annual precipitation and potential
evapotranspiration. PET was calculated as:

\[ \text{PET} = 58.93 \times T_{\text{above } 0 \degree C} \]

where \( T_{\text{above } 0 \degree C} \) is the average temperature of the months with positive
temperatures (Holdridge 1967).

Because different calibration model algorithms usually give slightly different
predictions, their results were averaged by using ensemble averaging
methods (Araújo and New 2007). For binary models, a majority vote was
used, which is an ensemble forecasting method that assigns a presence
prediction only when more than half of the models (i.e. >3 in this case)
predicts a presence. Next, population density was predicted for every 0.375°
by 0.25° cell inside the modelled range. For density models, to average the
results based on different algorithms, their median was calculated for each
cell.

To calculate the human population size in Europe for every 1000 years,
the land area of each cell was first calculated. Here, the systematic north-
south areal change of the 0.375° by 0.25° cells was taken into account. Next,
the predicted population density (median of six predictions) of the cell was
multiplied by the land area of the cell, and these values were summed to get
the total population size.

To evaluate the uncertainty of the population size estimates, the whole
process from calibration model fitting to the calculation of population size
was repeated 500 times, each time using a random sample (70%) of the
training/calibration data. This enabled the calculation of confidence limits
for the simulated population size estimates. The set of modelling techniques
and climate data was held constant throughout the process.

To evaluate the results of the bioclimatic envelope model simulation, the
simulated range and size of the human population was compared to an
archaeological population proxy based on archaeological \(^{14}\text{C}\) dates (see
above). The temporal comparison between the simulated population size and
the relative population size given by archaeological proxy was done
statistically, by calculating correlation of these data between 30,000 and
13,000 cal BP. Spatial comparison was done only by eyeballing. These
comparisons indicate whether the climate envelope of hunter-gatherers has
remained the same from the Pleistocene to the Holocene, and whether the
climate has an impact on hunter-gatherer population dynamics.
2.3 PALAEOENVIRONMENTAL DATA

As has been seen, palaeoclimatic reconstructions play a critical role in this study. However, because the aim in this study is not to produce new knowledge on past climate or environments and because reconstructions are made by experts in palaeoclimatology and climate modelling, these data and methods are only briefly described here.

The aim was to select climatic and ecological variables that would exert a strong influence on ecosystem productivity, and thus on food availability for hunter-gatherers, who depend on natural plant and animal resources for their survival. In addition, factors that have a more direct impact on human well-being, such as winter temperature, play an important role in hunter-gatherer population dynamics.

2.3.1 PALAEOCLIMATIC AND PALAEOECOLOGICAL PROXY DATA

In eastern Fennoscandia, archaeological population proxies are compared to palaeoenvironmental proxies, of which pollen-based temperature reconstructions are the most important. These are based on using modern pollen-climate calibration sets to create transfer functions that can be applied to obtain quantitative reconstructions of past temperature or, in some cases, precipitation. The idea is thus similar to the human population model described above: to use the link between the composition of modern pollen assemblages and climate to infer past climate. Here, the only information required from the past is pollen composition obtained from sediment records.

There are many aspects in quantitative climate reconstructions that can affect the reconstruction output, such as the selection of the transfer function techniques and the design of the calibration set used for constructing the transfer function. These aspects have been discussed by Juggins & Birks (2012). In paper III, the technique used in the reconstructions is weighted-averaging partial least squares (WA-PLS), one of the most commonly used reconstruction techniques in quantitative palaeoclimatology. Juggins & Birks (2012) discuss the many advantages of this technique. One problem with this method is that the absolute reconstructed values of climate parameters can be highly sensitive to the characteristics of the calibration dataset (Salonen et al. 2013b). However, the reconstructed patterns appear to be robust irrespective of the calibration data selection, which enables the correction of possibly biased absolute values by fixing the curve with modern observed values (Salonen et al. 2013b). Furthermore, in this study, the sensitivity of the absolute reconstructed values to the calibration set design is not a critical problem, since the focus is on the patterns, i.e. on the relative changes of climate parameters (papers I, III, IV).

In papers I and IV, temperature reconstructions from single high resolution pollen records are used (Heikkilä and Seppä 2003; Seppä et al.
2002), whereas in paper III, the mean annual temperature reconstruction is based on the combination (stacking) of four pollen records from southern Finland (Heikkilä and Seppä 2003; Ojala et al. 2008; Sarmaja-Korjonen and Seppä 2007; Seppä et al. 2009b).

In addition to temperature, proxies for hydroclimate, the intensity of the primary production season, and aquatic productivity are used in paper III. The hydroclimatic conditions of eastern Fennoscandia are tracked with cellulose-inferred oxygen isotope (δ^{18}O) values from one lake sediment core (Heikkilä et al. 2010). The oxygen isotope values of the lake sediment core reflect partly the isotopic composition of the meteoric water, and partly the general water volume of the lake, which is controlled by the outcome of precipitation and evapotranspiration. As summer evaporation causes enrichment of the heavier δ^{18}O in relation to lighter δ^{18}O, increased δ^{18}O values indicate warmer and drier conditions. The annual accumulation of organic matter in the annually laminated Lake Nautajärvi reflects the autochthonous primary production of the lake, while at the same time, and perhaps more importantly, it can be used as a proxy for the warmth and length of the primary production season (Ojala and Alenius 2005; Ojala et al. 2008). Another proxy used for aquatic production in paper III is the nitrogen content of the lake sediment (Heikkilä et al. 2010), which, however, is rather lake-specific and does not reflect any general pattern in the productivity of aquatic ecosystems, and therefore is not shown in this summary paper.

The Baltic Sea has been an important resource environment for humans throughout the Holocene. In addition to temperature, its productivity is affected by the inflow of saline North Atlantic waters (Möllmann et al. 2000; Ojaveer et al. 1998; Sohlenius et al. 2001; Tuovinen et al. 2008; Vuorinen et al. 2009). After the beginning of the latest brackish water Litorina stage at around 8500 cal BP, the inflow has been controlled by meteorological conditions, freshwater input, and the change in the inlet cross-section area at the Danish Straits (Gustafsson and Westman 2002), which is the most important non-climatic factor. In this study, it is assumed that the highest productivity of the Baltic Sea occurs when high salinity and temperature values coincide. In paper I, the highest salinity levels are indicated according to several sources (Emeis et al. 2003; Tuovinen et al. 2008; Westman et al. 1999; Westman and Sohlenius 1999), while in paper III and in this summary paper the salinity of the sea is tracked with the continuous proxy data given in Gustafsson and Westman (2002).

Changes in the forest ecosystem from tundra to birch forest (paper IV), and from temperate mixed forest to spruce dominated boreal forest (paper III), are tracked with several pollen records forming relevant south-north (Bondestam et al. 1994; Salonen et al. 2013a; Seppä 1996; Seppä et al. 2002; Seppä and Hammarlund 2000; Subetto et al. 2002) and east-west (Alenius and Laakso 2006; Heikkilä and Seppä 2003; Ojala and Alenius 2005; Ojala et al. 2008) transects over the research area.
2.3.2 SIMULATED CLIMATE DATA

In paper V, monthly average temperature and annual precipitation values for Europe were generated using a full last glacial cycle simulation (126000 years ago until the present-day) with the CLIMBER-2-SICOPOLIS model system (Ganopolski et al. 2010), which simulates climate at a temporal resolution of 1000 years. CLIMBER-2 is a so-called Earth system model of intermediate complexity (EMIC). It includes model components for the atmosphere, ocean, sea ice, land surface, terrestrial vegetation, and ice sheets (Petoukhov et al. 2000). The reason for using an EMIC instead of an Earth system model of full complexity (ESM), or general circulation models (GCM), is that the time-period is too long for GCM simulations with adequate spatial resolution. Therefore, downscaling the low resolution simulation results produced by the climate model is necessary regardless of the choice of the simulation model. Climate data was first downscaled statistically to the resolution of 1.5° (longitude) by 0.75° (latitude) (Korhonen et al. 2014). The temperature data were then re-gridded to 0.375° by 0.250°. During the re-gridding process, monthly temperature values were lapse by the pseudo adiabatic lapse rate (6.4 °C/km) to account for differences in average elevation between the fine-scale and coarse-scale grids (Vajda and Venäläinen 2003).

Although climate models provide more detailed information on past climates than palaeoclimatic records, they are only models, and as with all models they are not perfect representations of reality. Different climate models would give at least slightly different results. One specific problem with the current model is that it cannot simulate high-frequency climate variations. Therefore, for example, some of the Late Pleistocene cold events, such as Heinrich Stadial 1, do not show up in the model data.
3 RESULTS

3.1 EVALUATION OF ARCHAEOLOGICAL POPULATION PROXY

Figure 8 shows that the patterns in the temporal distribution of $^{14}$C dates are remarkably consistent between different date submitter classes and throughout the history of $^{14}$C dating in Finland. All curves indicate low Early Holocene frequencies, a distinct period of growth and decline between 6000 and 4000 $^{14}$C years ago and a new rise in date frequencies starting from 3000 $^{14}$C years ago. For the Early to Mid-Holocene, the clearest deviance is produced by the dates of the Early in the North project, because their highest frequencies are reached slightly earlier than in other curves. The spatial coverage of these dates is, however, rather restricted (northern Finland) when compared to the other date submitter classes, and the small differences may reflect regional peculiarities in the temporal distribution of archaeological material. These can relate to true differences in prehistoric population history, or the research situation in sparsely populated northern Finland. Nevertheless, the Early in the North dates also show the same general Mid-Holocene pattern of growth and decline as other curves.

Figure 9 shows further that, for the Stone Age (11,000—3500 cal BP), the pattern in date frequency (Fig. 9 A) is similar to other spatially and/or temporally more restricted archaeological population proxies (Fig. 9 D-H): all the records show the same Mid-Holocene rise and decline culminating slightly after 6000 cal BP. Furthermore, genetic data indicate a Late Stone Age population bottleneck (Fig. 9 I) occurring around 3900 cal BP (Sajantila et al. 1996), i.e. at the same time with the lowest post Mid-Holocene peak levels in date distribution. Based on the characteristics of the archaeological data, Lavento (2001: 175–177) has also suggested that population size declined, and may even have gone locally extinct at the end of the Late Stone Age.

All this strongly suggests that research biases do not have a marked effect on the shape of the temporal frequency distribution of $^{14}$C dates and that the date distribution truthfully reflects the distribution of known archaeological material. As suggested in Chapter 2.1.1.1, it is possible to assume that the distribution of the known archaeological material reflects the real distribution of archaeological material when a dataset is large and/or most of the research is done by cultural heritage management without any strong research interests, as is the case in Finland.

It is also obvious that the Mid-Holocene part shows significantly more, and the Late Stone Age-Early Bronze Age significantly less, dates than the null model of monotonic growth pattern would suggest (Fig. 9 B). Thus, the
observed long-term pattern is not influenced by calibration and/or (random) sample variation.

Based on the evidence presented so far, it is possible to assume that the temporal distribution of $^{14}$C dates reflects the strength of true archaeological signal and therefore the size of the human population that produced the archaeological material. However, when the possible taphonomic bias is taken into account, the pattern changes (Fig. 9 C). In the taphonomically corrected distribution, the Late Holocene population remains smaller than during the Mid-Holocene population peak. This may have interesting implications regarding the cultural history of Finland (see Discussion). Nevertheless, taphonomic correction does not markedly change the pattern for Early to Mid-Holocene hunter-gatherers, which are the main focus of this study: the distinct period of Mid-Holocene growth and decline remains unchanged. In contrast, when dealing with the Pleistocene data in paper V, taphonomic correction exerted a strong influence on the shape of the date distribution and thus also on the results of the paper (see paper V).
Figure 9. Comparison of $^{14}$C date-based population proxy to other proxies, and evaluation against calibration and taphonomic bias. A: Human population proxy based on the temporal frequency distribution of calibrated median $^{14}$C dates from Finland (with 100 years bins). B: Population proxy evaluated against calibration (and sample) bias. Red areas indicate significantly higher and blue areas significantly lower frequency of dates than the null model of monotonic growth pattern (dotted line and grey polygon) would suggest. Archaeological realisation of the null model (grey polygon) is based on 1000 simulations, and it shows the range that covers 90% of the simulations. C: Taphonomically corrected distribution of dates. D: Ceramic Site Frequency Index that shows the distribution of sites dated typologically using ceramic finds (Paper I). E: Distribution of the number of coastal sites (Siiräinen 1981), adjusted for the length of each period (Site Frequency Index). F: Distribution of the absolute number of stone artefacts assigned to a given period (Sundell et al. 2014), adjusted for the length of each period (Artefact Frequency Index). G: Distribution of the number of different artefact types associated with each period (Sundell et al. 2014), adjusted for the length of each period (Type Index). H: Percentage of marine mammal bone fragments in coastal archaeological bone assemblages. I: Timing of the genetic bottleneck inferred from Finnish genetic data (Sajantila et al. 1996).
3.2 CORRESPONDENCE BETWEEN HOLOCENE CLIMATE CHANGE AND LONG-TERM POPULATION DYNAMICS

Figure 10 shows that the outlines of the \(^{14}\text{C}\) date-based population proxy and different indicators of climate and ecosystem productivity are correlated between 11,000 cal BP and 3500 cal BP in Fennoscandia. Similarly to human population, temperature as reflected in pollen based reconstruction and lacustrine \(\delta^{18}\text{O}\) data, warmth and length of primary production season, as well as Baltic Sea productivity, all show a rise, culminating at or slightly after 6000 cal BP, and a subsequent decline. Similarity between the population proxy and the simulated population density suggests further that annual mean temperature alone is an important variable affecting hunter-gatherer population dynamics. Despite the overall similarity, these two patterns differ in the relative strength of the population peak, which is more pronounced in the proxy data than in the model. This difference suggests that the annual mean temperature does not solely explain the variability in hunter-gatherer population size. Instead, the pattern of the warmth and length of the primary production season (Fig. 10 C) is highly similar to that of the human population, with its clear peak at 6000 cal BP. Together with the salinity of the Baltic Sea, this indicates that the productivity of the terrestrial and aquatic ecosystems, and consequently food availability for human hunter-gatherers, probably was especially high between 6000 and 5000 cal BP.

Figure 10 also explores the possible link between hunter-gatherer population size and major changes in the forest composition in Finland. Palaeoecological research has shown that a forest composition change occurred during the Holocene in Fennoscandia, as Norway spruce gradually invaded and expanded in the region from the east. Norway spruce eventually became the dominant species by outgrowing and suppressing the temperate deciduous taxa, particularly lime, that were important components of the forests prior to the spruce expansion (Seppä et al. 2009a). It appears that there is an interesting correlation with the timing of the forest ecosystem change and population size in Finland, as the population decline coincides with the rise of the spruce and consequent rise of the boreal forest ecosystem.

However, the positive correlation between human population size and environmental factors seems to vanish after c. 3500 cal BP, as population size starts to grow again despite the declining temperature and productivity. This new rise correlates with the intensification of farming economies that started from the Bronze Age onwards. There are recent claims of much earlier cultivation (Alenius et al. 2013), but these are based on single pollen grains in pollen records and cannot be taken as a conclusive evidence of a productive economy. The firmest pollen evidence for the earliest cultivation in the southern part of the research area dates nevertheless to c. 4000 cal BP, whereas the first macrofossil evidence is c. 500 years younger (Vuorela 1998,
The earliest signs of dairy farming date slightly earlier to c. 4500 cal BP (Cramp et al. 2014).

Figure 10. Human population proxy compared to palaeoclimatic and palaeoenvironmental proxies in eastern Fennoscandia. A: $^{14}$C date-based human population proxy together with the simulation that shows how population density should have varied if it was solely affected by annual mean temperature. B: Pollen-based reconstruction of annual mean temperature (loess smoothed). C: Organic varve thickness (loess smoothed) in the annually laminated Lake Nautajärvi (Ojala et al. 2008). Varve thickness reflects the warmth and length of the primary production during summer. The Late-Holocene rise is linked to human activity, not to climate (Ojala and Alenius 2005; Ojala et al. 2008). D: Oxygen isotope ($\delta^{18}$O) values (loess smoothed) of the sediment of Lake Saarikko (Heikkilä et al. 2010). $\delta^{18}$O tracks changes in hydroclimate. High values indicate warm and dry conditions. E: Changes in the forest composition indicated by changes in the pollen percentages of temperate tree taxa and Norway spruce (Alenius and Laakso 2006; Heikkilä and Seppä 2003; Ojala and Alenius 2005; Ojala et al. 2008; Seppä et al. 2009a). F: Salinity of the Baltic Sea (Gustafsson and Westman 2002). Together with temperature, salinity affects the productivity of the sea. The figure also highlights the Mid-Holocene period, when the productivity is assumed to be high and when the hunter-gatherer population size has also been highest.
The first scanty evidence for farming is followed by the increased number of sample sites with positive Cerealia pollen indications dating to the Bronze Age, while the beginning of more intense agriculture dates to the Late Bronze Age, and especially to the Early Iron Age (c. 2500 cal BP onwards). This intensification is indicated by the continuous presence of Cerealia pollen in pollen diagrams, as well as by the increased number of cultivars and increased importance of cattle (Asplund 2008; Edgren 1999; Vuorela 1998). In the inland, signs of cultivation increase during the Late Bronze Age/Early Iron Age, even though the beginning of continuous cultivation is dated well into the historical period (Taavitsainen et al. 1998, 2007). The expansion of agricultural economies observed at the beginning of the Iron Age is also a more general phenomenon around the Baltic Sea, which interestingly occurs in the context of long-term climatic cooling (Asplund 2008: 280).

### 3.3 IMPACT OF EVENT-LIKE CLIMATE CHANGE ON HUMAN POPULATION

Figure 11 shows that the so-called 10,300 cal BP cold event coincides with the sharp decline in the number of $^{14}$C dates. This suggests that even short-lived climate events could have had a profound effect on the small pioneer population of hunter-gatherers. However, when the distribution of archaeological $^{14}$C dates is compared to the simulated distribution (grey polygon), it appears that the latter has a similar gap in the distribution. In general, the archaeological population proxy corresponds well with the simulation, i.e. the null model that assumes smooth growth over the Early Holocene. Instead of being a real demographic signal, the gap in the distribution of archaeological $^{14}$C dates between 10,300 and 10,100 cal BP may thus only be a calibration artefact. Although in paper IV it is argued that the cold event at 10,300 cal BP had a clear impact on human population size in Finland, this conclusion now seems premature. It is too early to conclude whether the cold event had any demographic effect or not, despite the fact that the decline in the date distribution coincides with the event.
Early Holocene population dynamics compared to climate variability and evaluated against calibration bias. A: NGRIP Oxygen isotope data (loess smoothed) from Greenland (Rasmussen et al. 2006). B: Pollen-based July temperature reconstruction from Ifjord northern Norway (Seppä et al. 2002). Ifjord temperature data shows at least three cold events (11,400, 10,800 and 10,300 cal BP) that coincide with the cold events in the Greenland data. C: $^{14}$C date-based human population proxy shown as a kernel density plot (black curve). Dark grey curve shows the null model of smooth growth and the grey polygon shows its archaeological realisation. The polygon is based on 1000 simulations, and shows a range that covers 90% of the simulations. Comparison shows that the archaeological proxy follows the model of smooth growth. The figure also shows the period of reduced vegetation cover in north-western Russia, linked to the cold event at 10,300 cal BP (Subetto et al. 2002), and two of the episodes of increased freshwater flux from the melting Laurentide Ice Sheet into the North Atlantic (R1 and R2), which are assumed to have caused the climatic cooling (Clark et al. 2001).

3.4 CLIMATE ENVELOPE MODEL SIMULATION OF ICE AGE POPULATION DYNAMICS IN EUROPE

Figure 12 shows that the temporal patterns in the simulated population size (Fig. 12 A) and relative population size estimates based on archaeological $^{14}$C data (Fig. 12 D) are remarkably consistent ($r_p = 0.84, p<0.00002$), which supports the validity of the simulation. Because the simulation is based only on three climate variables, this correspondence between the simulation and the archaeological record also indicates that climate has been an important factor driving the long-term population dynamics of humans. Both the simulation and archaeological proxy show relatively high late Marine Isotope Stage 3 population size levels, a decline towards the LGM minimum and a
Results

rapid growth during the Late Glacial. The simulation suggests that the human population size in Europe was about 330,000 at 30,000 cal BP, and 130,000 during its minimum at 23,000 cal BP and 410,000 during the Greenland interstadial 1 (Fig. 12 A).

Figure 12. Climate envelope model simulation of human population dynamics between 30,000 and 13,000 years ago. A: Simulated human population size in Europe. Error bars show the resampling-based confidence limits (95%). B: Simulated mean density in the inhabited area of Europe. Error bars show the resampling-based confidence limits (95%). C: Changes in the percentage of potentially inhabited land area in Europe. D: Archaeological population size proxy based on the taphonomically corrected number of dates. E: Percentage of time the area has potentially been inhabited between 30,000 and 13,000 years ago. F: Mean population density (people/100km$^2$) between 30,000 and 13,000 years ago.

The simulated spatial pattern of human population indicates a population contraction starting in line with the known ice sheet expansion at 27,000 cal BP (Fig. 12 C). The post-LGM recolonization of the continent started from 19,000 cal BP onwards (Fig. 12 C). The simulation suggests that the continuously suitable and inhabited area between 30,000 and 13,000 cal BP covered 36% of the European land area even during the coldest LGM (Fig. 12 C), stretching to the north of the Alps (Fig. 12 E), a result also supported by an emerging archaeological picture (Terberger and Street 2002). In addition, the simulation shows a persistent southwest-northeast gradient of decreasing population densities, with the densest populations throughout the LGM in the Iberian Peninsula and the Mediterranean region (Fig. 12 F).
4 DISCUSSION

4.1 METHODOLOGICAL CONSIDERATIONS

The evaluation of the population proxy described above shows that dating-related research biases do not determine or strongly influence the temporal distribution of $^{14}$C dates in eastern Fennoscandia. Larger samples, such as the one used in paper V, are even less likely affected by these biases. In the Fennoscandian data, the correlation with other proxies indicates that the archaeological population proxy based on $^{14}$C dates is a reliable indicator of the temporal distribution of known archaeological material instead of being biased by e.g. research interests. Elsewhere, Williams (2012) and French and Collins (2015) have compared SPDs with temporal frequency distributions of archaeological sites in Australia for the last 40,000 years and in the Upper Palaeolithic southwestern France respectively. Both studies found clear positive correlation between proxies. Furthermore, Downey et al. (2014) analysed proportions of immature skeletons from Mesolithic and Neolithic cemeteries in Europe and found evidence for an increase in the population growth rate after the adoption of agriculture. This pattern matches the picture obtained from the SPDs. All these comparisons in different geographical and chronological contexts provide support for the general validity of temporal distributions of $^{14}$C dates as a proxy for human population size.

However, Mökkönen (2011, 2014) has recently strongly criticised the archaeological population proxies. The most important part of his criticism is that the temporal distribution of the archaeological material reflects changes in the settlement pattern and, especially, in the current visibility of archaeological sites, not in the demography. According to Mökkönen (2011, 2014), the Late Stone Age decline in the frequency of $^{14}$C dates (and other proxies) in the Finnish material is a result of the decline in the number of highly visible pit house sites, a shift towards smaller and thus less visible pottery, and reduced residential mobility, which together lead to a smaller number of identified sites. Should this argument hold, one has to assume that after the Stone Age, residential mobility and the visibility of archaeological dwelling sites and pottery increases as the number of $^{14}$C dates increases again. Yet, it is known that after the Stone Age the number of pit house sites does not increase (Pesonen 2002), pottery becomes less visible (Ikäheimo 2002) and the settlement pattern likely becomes more sedentary due to the increasing influence of agriculture on the economy. Furthermore, the pattern in the temporal distribution of $^{14}$C dates has remained the same throughout the history of radiocarbon dating in Finland. The pattern was the same in the 1970s and 1980s as it is today (Fig. 2). It is also the same in Siiriäinen’s (1981) curve, which is based on even older data gathered before
1969. Thus, the boom in the research of house pit sites in the 1990s did not have any profound influence on the shape of the distributions of archaeological material and $^{14}$C dates.

It thus seems that Mökkönen's argument is invalid, and will remain so until it is shown that there is indeed such a strong correlation between the current visibility of archaeological remains and the frequency of $^{14}$C dates (or known sites) in Finland. In addition, genetic data on the population bottleneck and the simple population simulation shown in Figures 9 and 10, are independent of the archaeological record and yet provide clear support for the temporal distributions of different archaeological proxies. The fact that several archaeological and non-archaeological records in the Finnish material show the same pattern indicates that they are tracking a real demographic signal from the past (see also Tallavaara et al. 2014). The above mentioned comparison of two independent proxies by Downey et al. (2014) further suggests that this holds for other datasets as well.

Recently, Contreras and Meadows (2014) have also criticised the use of $^{14}$C date-based population proxies by arguing that even when there is a perfect correspondence between population size and the quantity of dateable material produced at any given time, these methods cannot trace the true population history. However, they fall into a logical fallacy of straw man argumentation when they show that archaeological methods cannot accurately reflect the two historical population events (< 200 yrs long) that they had chosen for comparison: the use of $^{14}$C dates for population reconstruction is not based on the argument that $^{14}$C date-based proxies would trace all demographic events. Furthermore, the simulation approach of Contreras and Meadows (2014) lacks the essential component of sample variation, which is another flaw in their study. In addition, comparison of the distribution of time points (true calendar ages) and the distribution of probability distributions is like comparing apples and oranges. Summed probability distributions can smooth out true variation (and create artificial ones), and consequently conceal small-scale fluctuations in the archaeological signal. The use of point estimates such as calibrated median dates may solve this problem. All in all, the attempt of Contreras and Meadows (2014) fails to provide a relevant critique of $^{14}$C date-based population proxies (see also Timpson et al. 2014).

However, the method is clearly not without its drawbacks. As shown here and elsewhere (e.g. Bamforth and Grund 2012; Brown 2015), short-term variation in the temporal distribution of $^{14}$C dates is strongly influenced by $^{14}$C date calibration. This is probably the biggest weakness of the method, because it hampers archaeologists reliably detecting short-lived demographic events. The simulation approach developed by Shennan et al. (2013), Timpson et al. (2014) and Brown (2015) provides a tool to evaluate the impact of calibration and sample variation on the distribution of dates, but it does not remove the problem. In this summary paper, the simulation approach revealed that despite remarkable fluctuations, the Early Holocene
archaeological pattern in Finland corresponds to the pattern simulated under the null hypothesis assuming smooth growth over the Early Holocene. Thus, simulations can be considered as an essential tool when the focus is on high-frequency fluctuations in population proxy. Simulation allows a case by case evaluation of whether the observed pattern really deviates from the pattern simulated under the null model of smooth demographic signal.

In this summary paper, the simulation approach also showed that the longer-term fluctuations – the Mid-Holocene boom and bust – are not calibration artefacts, but represent real fluctuations in the archaeological signal. The same pattern is observed all over northern Europe, where the observed pattern also deviates significantly from the null model (Shennan et al. 2013).

On the other hand, taphonomic bias appears to have a strong impact on the long-term trend in the distribution of $^{14}$C dates. The taphonomic correction changes the shape of the archaeological population proxy both in Finland and in Palaeolithic Europe. In the Finnish data, the taphonomically corrected proxy indicates that Mid-Holocene population levels would have been higher than Late Holocene levels. This may appear counter-intuitive. However, during the Mid-Holocene, environmental conditions were more suitable for the hunter-gatherers (see below) who populated the whole area. From the late Mid-Holocene onwards environmental conditions started to deteriorate due to the cooling and the gradual replacement of mixed forest by coniferous forest (see below). The environmental carrying capacity for hunter-gatherers would thus have been lower during the Late Holocene than during the Mid-Holocene. The intensification of the agricultural economy from the Bronze Age–Early Iron Age onwards induced a new population growth, but during the Iron Age, the area inhabited by the agricultural population was still relatively restricted. Consequently, the total population size in the area may have remained smaller than during the Mid-Holocene, when environmental conditions were optimal for the whole population. Thus, the taphonomically corrected proxy may not be so flawed a reflection of the true population history as it may initially seem.

Nevertheless, the model of taphonomic loss and the derived correction is a simplification of a rather complex phenomenon. It is clear that there is regional variation in taphonomic loss processes. On the continental scale, for example, the interplay of climatic and topographic factors result in higher erosion rate in the Mediterranean region than in central and northern Europe (García-Ruiz et al. 2013; Poesen and Hooke 1997; Vanmaercke et al. 2011). This probably also leads to a more severe loss or disturbance of archaeological contexts in the Mediterranean region than elsewhere in Europe, which could be useful to take into account when correcting European-wide datasets, such as the one used in paper V. The ideal situation, in general, would be to devise a correction factor for each regional dataset, based on local circumstances. In some cases this may work, but in most of the cases it is practically impossible.
Thus, one may argue that temporal frequency distributions should not be routinely corrected because it applies a theoretical correction that may not be valid in all situations or because it may over correct the number of very old dates (Williams 2012). However, since the time-dependent loss of material seems to be a very real process affecting the temporal frequency distributions of different types of data, it has to be taken into account somehow. So far, the model first formulated by Surovell et al. (2009) and later modified by Williams (2012) is the best tool available to tackle this bias.

Although $^{14}$C date-based population proxies are shown to be valid indicators of prehistoric population size, one may still ask, whether it would be possible to use some other archaeological proxies to avoid all the potential biases related to $^{14}$C date data. Of all the archaeological population proxies, $^{14}$C date-based proxy is the most comprehensive, because its use is not bound to periods or areas where good typological or shore displacement chronologies are available. In optimal circumstances, its resolution is also much higher than what can be achieved by typological dating. However, comparisons between different proxies are always useful as a means to evaluate the validity of observed patterns.

Even if archaeological proxies provide a reliable method for studying long-term population dynamics, archaeological data is still sparse and often ambiguous. As suggested in Chapter 2.1.1.1., this holds especially for spatial distributions. One major factor here is the Wallacean shortfall-like effect of incomplete information on species distribution. While the Wallacean shortfall is true for current plant and animal species, paleontological and archaeological records provide an obviously even more incomplete and coarse reflection of true ranges (Svenning et al. 2011), as one is moving further and further back in time. Therefore, the study of prehistoric population dynamics demands multiple lines of evidence. It is shown here that by using only three climate predictors within climate envelope or niche modelling approach, it is possible to simulate Pleistocene population pattern that corresponds with the archaeological pattern. This indicates that such a modelling approach provides an efficient tool that is a necessary complement to archaeological and DNA-based methods. Furthermore, with this method, it is possible to estimate population dynamics even for the areas without a sufficient archaeological record to create a reliable proxy. However, in such cases, results must be considered more tentative as model-data comparisons are not possible.

As with all other methods of prehistory, climate envelope modelling is also an imperfect one: the resolution of the simulation depends on the resolution of the climate model and, more importantly, the accuracy of the simulated density and range estimates depends on the ability of the climate model to correctly simulate past climate. While climate models are becoming better due to the intensive development, other critical issues remain to be solved. One is the assumption of niche stability. The further back in time one is hindcasting population dynamics, the more likely the potentially
confounding effect of non-analogous environments becomes. For example, during the glacial periods of the Pleistocene, atmospheric CO$_2$ levels were 80 ppm lower than during the Holocene, and almost 200 ppm lower than at present (Monnin et al. 2001). This may have had a major impact on plant productivity and ecosystem functioning (Gerhart and Ward 2010). It is thus possible that similar climatic conditions would have affected humans differently during the Pleistocene than during the Holocene. However, the similarity between the simulated pattern and the observed archaeological pattern indicates that this may not be a serious problem in this case. Of course, the correlation validates only the simulated pattern, i.e. the relative, not absolute, population sizes. Estimates of absolute population size will always remain somewhat speculative no matter what method is used to estimate them.

4.2 IMPACTS OF CLIMATE

The correlations between archaeological and several environmental proxies, and between climate envelope model simulations and the archaeological proxies, support the idea that changes in climate and environment have a strong impact on hunter-gatherer population dynamics: hunter-gatherer population size appears to change in sync with climate (see also Kelly et al. 2013; Munoz et al. 2010). Because climate envelope model simulations are based on information derived from modern or recent historical hunter-gatherer populations, the ability of models to replicate archaeological patterns indicates further that the effect of climate has remained relatively constant from the last glacial to the present. Millennia of cultural evolution have not fundamentally changed the constraints on terrestrially adapted hunter-gatherer populations posed by the climate.

However, these results are mostly valid for long-term trends, and do not provide unambiguous evidence that short-lived climate events (less than c. 200 yrs long) would have impacted archaeologically visible population history. The Fennoscandian archaeological record indicates changes that may be related to the Early Holocene cold events, as Manninen (2014) suggests, but the impact on population size is not clear, because possible short-term population fluctuations are masked by variation induced by the $^{14}$C calibration procedure. This does not mean that climate events would not have had an impact on hunter-gatherer population dynamics: Riede (2009) and Wicks and Mithen (2014) have suggested that the cold events of the late glacial and Early Holocene would indeed have influenced hunter-gatherer populations. However, they have not included extensive resampling-based simulations in their studies to evaluate the possible calibration bias. The problems related to the detection of the impacts of rapid climate events are amplified by the fact that some of the cold events co-occur with shifts in atmospheric $^{14}$C content, as both may be linked to abrupt shifts in ocean
Discussion

circulation and deep ocean ventilation (Hughen et al. 2000). This means that the cold events would show up as troughs in a $^{14}$C date-based population proxy, even when population decline had not truly occurred. The problem is further illustrated in Figure 13. It shows how difficult it is to accurately radiocarbon date human activity that occurred during such events and, consequently, how difficult it is to determine the intensity of human action during the event.

![Figure 13: Simulated example showing how difficult it is to accurately date past human activity occurring during, e.g., the 10,300 cal BP cold event. Seven true calendar ages between 10,300 and 10,000 cal BP were transformed into calibrated radiocarbon ages using the simulation approach described above. Calibrated radiocarbon ages were simulated 1000 times for each calendar age. A: Distributions of simulated calibrated medians for each calendar age (in each distribution N=1000). The figure shows that practically only one of the calibrated radiocarbon ages falls within the period between 10,300 and 10,000 cal BP, despite the fact that all calendar ages, i.e. occurrences of human action, are within the period. The vertical bars show the means of each distribution. B: With the calibrated probability distributions the gap is not equally serious, only because probability distributions are so wide. Calibration nevertheless packs the youngest four dates on top of each other. The mean of simulated radiocarbon ages for each calendar age was used in the calibration, with a standard error of 80 years.](image)

In the eastern Fennoscandian data, correlation between long-term human population dynamics and temperature and, especially, productivity proxy strongly suggests that the impact of climate on human population was mostly indirect, mediated by its impacts on environmental production and consequently on food availability. Successful climate envelope model
simulation of the glacial human population in Europe, using potential evapotranspiration and water balance as two of the predictor variables, supports this idea. In this context, the negative correlation between effective precipitation and population size in eastern Fennoscandian data may appear counterintuitive at the outset. However, in the northern latitudes, temperature is generally the most important factor affecting primary productivity (e.g. Zheng et al. 2004), and it seems unlikely that water availability would have limited production during the Holocene. Instead, in the cold and dry glacial climate, the availability of water played an important role, especially in the southern and eastern parts of Europe.

In addition to the warmth and length of the growing season, climate affects ecosystem composition that may also have an effect on food abundance. Although not necessarily in complete equilibrium with the climate (Seppä et al. 2009a), the forest ecosystem change in eastern Fennoscandia correlates with climate change, which makes it difficult to determine whether it was the climate-controlled productivity as such, or changes in the forest composition and structure that had a more profound effect on food abundance, and consequently on hunter-gatherer population size. As shown in paper III, there is nevertheless a difference in animal biomass between coniferous and mixed forests in modern data: the temperate mixed forest ecosystem sustains a higher animal biomass than the boreal forest. This is evident for the whole animal biomass as well as for single species, such as moose. Although these biomass differences partly relate to climate through differences in growing season length, it is still likely that the differences in forest structure and soil biochemistry also exert a strong influence. Coniferous species, especially spruce, effectively shade out herbaceous, shrub and immature tree growth in the understory, thus reducing the abundance of important food sources, e.g. for ungulates (Mellars 1975). The negative effect of shading on the undergrowth is further enforced by acidic litter, which decreases soil fertility (e.g. Kohtsik et al. 2001; Levula et al. 2003; Wallrup et al. 2006). It is thus reasonable to assume that the ecosystem change from species-rich temperate mixed forests to typical boreal forests would have had a negative effect on hunter-gatherer food availability in eastern Fennoscandia.

In addition to its influence on food abundance, climate can have direct somatic effects that affect human survival. Even in modern Finland, mortality is higher during cold months than during warm months (Hassi and Ikäheimo 2013). Inhaled cold air, cooling of the body surface, and cold stress induced by lowering the core body temperature cause physiological responses leading to increased mortality risk (Barnett et al. 2007; Conlon et al. 2011; Mäkinen et al. 2009; Mourtzoukou and Falagas 2007). Cold environments are also more hazardous, and the risk of dying of hypothermia is obviously much higher than in temperate environments.

Because of the complex and partly collinear relations between the forcing variables, it is not possible to fully quantify the relative importance of the
different environmental factors affecting past hunter-gatherer population dynamics. The climate envelope modelling approach would allow such quantification by evaluating the predictive ability of multiple sets of predictor variables. However, such evaluations have not been done in this study. It is nevertheless clear that there were several interwoven environmental factors forcing population size in the same direction. Cultural factors cannot be totally ruled out, but archaeological data e.g. in eastern Fennoscandia do not indicate technological changes that would explain variation in population size. Adoption of pottery occurs c. 7200 cal BP, coinciding with the beginning of population growth, but pottery continues to be in use also during the late Mid-Holocene population bust. Thus far it is possible to attribute the main causality of long-term hunter-gatherer population fluctuations to climate related changes in food availability.

It is argued here that the long-term changes in hunter-gatherer population size tracked by archaeological population proxies reflect long-term changes in environmental productivity, i.e., carrying capacity. This implies that the changes that we see in population size are density independent, i.e., they do not depend on the density or size of the human population. Density dependent fluctuations occur when the density of the population affects its own growth rate (Hanski et al. 1998). For example, as predator population grows, prey population declines, which slows down the growth of the predator population and eventually turns it negative. As the predator population declines, resources recover, which will eventually enable the new growth of the predator population as well. The question is, can we see density dependent fluctuations in archaeological population proxies? The simulation study by Belovsky (1988) suggests that density dependent fluctuations are common in hunter-gatherer populations (see also Winterhalder et al. 1988; Winterhalder and Lu 1997). In his simulation, the amplitude of these fluctuations varies according to the productivity of the environment, so that the most severe fluctuations occur in the most productive environments. The length of the cycle (high density-low density-high density) is c. 100 years in Belovsky’s (1988) simulation. The historical data on the number of Saami tax payers from Finnish Lapland (Enbuske 2008: 294) supports this figure, although it is not a direct measure of population size. Such a length for the cycle is clearly too short to be visible in an average $^{14}$C dates-based archaeological population proxy, especially as much of the short-term variation in the proxy can be attributed to the $^{14}$C calibration.

It is thus more likely that archaeological population proxy usually reflects changes in the mean population size (Fig 14). This could explain why population growth rates calculated from archaeological population proxies are clearly smaller than growth rates based on ethnographic or historical data (Kelly et al. 2013; Peros et al. 2010): ethnographic data reflects the true current population growth rate ($r$), whereas archaeological population proxy reflects changes in the long-term mean population size, which is controlled
by environmental productivity. High-frequency variation inferred from the ethnographic data and simulations, and low-frequency variation visible in the archaeological proxies and climate envelope model simulation, indicate that hunter-gatherer populations were not stationary, neither in the long- nor short-term.

![Figure 14](image.png)

**Figure 14.** Simulated example showing how hunter-gatherer population size (black curve) would fluctuate through time due to density dependent factors. Population size oscillates around its mean (grey curve), which changes as environmental productivity changes, for example, due to climate changes. The resolution of archaeological population proxies is not usually sufficiently high to detect the density dependent high frequency fluctuation. It is more likely that archaeological proxies reflect changes in the mean population size.

### 4.3 THE DIFFERING DYNAMICS OF FORAGERS AND FARMERS

Although the main focus of this study is on hunter-gatherers, the results from Finland also enable comparisons between the population dynamics of foragers and farmers. As mentioned above, the correlation between population size and environmental factors ceases from about c. 3500 cal BP onwards, most likely due to the intensification of the farming economy. The firmest pollen and macrofossil evidence for the earliest cultivation dates to 4000–3500 cal BP in southern and 3500–3400 cal BP in central and eastern Finland (Taavitsainen et al. 1998, 2007; Vuorela 1998, 1999; Vuorela and Lempiäinen 1988), while the oldest dated bone of domesticated animals dates to c. 4000 cal BP (Bläuer and Kantanen 2013) and the oldest evidence for dairy economy to c. 4500 cal BP (Cramp et al. 2014).

During the Early Metal Age, indications of farming increase in palynological and archaeological records around the Baltic Sea, suggesting the intensification of agricultural economy in the area (Asplund 2008;
This indicates that agriculture thrived throughout the cooling late Holocene as well as during some of the colder anomalies of this cooling trend, such as the Little Ice Age of the 16th century (Lagerås 2007). Similarly, Munoz et al. (2010) noted that the rise in the agricultural population in the northeastern USA during the Late Woodland period appears to have occurred during the period of climatic cooling.

It therefore seems that the adoption of agriculture weakened the climatic control over the long-term human population dynamics. Whereas hunter-gatherer population size changes in equilibrium with climate, agricultural population can grow independently of climate trends. Since agricultural populations are not dependent only on natural resources, their food availability is not controlled solely by environmental productivity. This allows the population to grow even if environmental productivity is decreasing.

Short-term, inter-annual or -decadal variability is nevertheless known to have affected crop yields and consequently the size of historical farming populations (Eckstein et al. 1984; Pitkänen 1993), and it is thus likely that crop failures and consequent famines occurred during prehistory as well. However, historical grain figure data for southern Finland shows only a weak spatial synchrony, which indicates that crop failures and possible population effects were usually geographically restricted (Holopainen and Helama 2009). This would have permitted population growth over a larger spatial scale despite frequent but spatially non-synchronous local declines. Thus, even when in a marginal economic position, agriculture seemed to offer a buffer against periodic climate induced shortages, which allowed for more sustainable long-term population growth (Boone 2002).

To act as a buffer against climate-related periodic shortages implies that the availability of agricultural products reduces the mortality of agricultural populations relative to non-agricultural populations during these shortages. Recently, Helle et al. (2014) have shown, however, that the slightly higher population growth rate of farmers as compared to foragers in historical northern Finland was related to differences in fecundity. In fact, they do not find clear population level mortality differences between farmers and non-farmers during their study period (AD 1740-1900). Helle et al. (2014) suggest that the differences in fecundity may not relate to dietary differences, since diets of farmers and non-farmers in northern Finland would have been roughly similar. However, this may not be entirely correct: in addition to wild resources and dairy products, grain products already played a role in the diets of the earliest farmers of the late 17th century, and certainly from the 1750s onwards (Kehusmaa and Onnela 1995: 147–183). A wider resource base and consequently better food availability would have allowed a better nutritional status for farmers in general, and also during lean years. This likely contributed to their higher reproductive success as well.
Observations made from the Finnish archaeological and historical data – that the long-term population dynamics of farmers are independent from climate and that their growth rates are higher than the growth rates of non-farmers living in similar environments – illustrate nicely the competitive or evolutionary superiority of farmers over foragers even in agriculturally relatively marginal environments: farming has replaced and will replace foraging in all but agriculturally unsuitable environments by sheer outnumbering (see also Richerson et al. 2001).
5 CONCLUSIONS

Based on the results described and discussed above, it is argued that:

1. Archaeological human population proxy based on the temporal frequency distribution of $^{14}$C dates is a reliable method for the Finnish data. Together with positive results of evaluations in different periods and geographical areas, this supports the general validity of the method.
2. Instead of using the summed probability distribution of calibrated $^{14}$C dates as a population proxy, it might be better to use the distribution of calibrated median, mean, or mode $^{14}$C dates, since such distribution better facilitates the demographic interpretation of the data.
3. Due to the bias produced by calibration, the method has to be used cautiously when studying short term fluctuations. High-frequency variation may not relate to demographic signal, but to features in the calibration curve.
4. Therefore, the bias associated with the calibration of the dates has to be taken into account. This is especially pertinent to studies that focus on short-lived demographic events. The calibration bias can be evaluated by comparing observed data to the distribution of simulated data randomly re-sampled from the null model.
5. Due to the calibration bias and other issues, the temporal resolution of the $^{14}$C dates-based method is usually not sufficiently high to track density dependent fluctuation in human population size. Instead, the method tracks changes in the long-term mean population size.
6. The climate envelope modelling approach can be a reliable method for estimating long-term hunter-gatherer population dynamics, and even absolute population densities and sizes. Although it is not without its limitations, the approach provides a necessary complement to archaeological methods. The method can also be used when archaeological data is not available, and it can reveal potential inadequacies in the distribution of archaeological data.
7. The climate explains over 60% of the variation in hunter-gatherer population densities, demonstrating that climate alone is a major factor affecting long-term hunter-gatherer population dynamics.
8. The impact of the climate on terrestrially adapted hunter-gatherer population dynamics has remained relatively constant from the Late Pleistocene to present.
9. Climate influences hunter-gatherer food abundance, and consequently population density mainly by affecting the length and warmth of the growing season (i.e. environmental productivity) and vegetation. Climate also has an effect on wintering conditions such as winter mortality.
10. Climate controls the mean population size or density (or the carrying capacity) around which the true population size oscillates because of density dependent factors. This high-frequency density dependent oscillation is not possible to track with the methods used in this study.

11. Unlike the clear effects of climate on the long-term population dynamics of hunter-gatherers, the Early Holocene short-lived cold events did not have a reliably detectable effect on human population size in eastern Fennoscandia.

12. The adoption of a farming economy weakens the climate forcing on long-term human population dynamics.

In the future, one of the greatest challenges in the study of past population history relates to the reconstruction of short-lived demographic events and high-frequency variation in population size when calibration-induced variation is present. Automated methods for filtering out the calibration bias have already been developed (Timpson et al. 2014). These methods can remove false positive demographic signals from the data, but they do not help detect true positive signals that are masked by the resolution of radiocarbon dating. It may turn out that the accuracy of $^{14}$C date-based population proxies will never be sufficiently high to truly detect short-term demographic variation.

There is also a need to further develop the taphonomic correction procedure. The process of taphonomic loss is real and must be taken into account, but the application of a single correction function may over- or under-correct archaeological temporal frequency distributions. It might be possible to improve the correction method by taking the geographic variation in erosion rates into account. So far, the correction method has been applied to temporal distributions, but it would be worthwhile to develop the means to taphonomically correct spatial distributions as well.

Because population proxies based on the distribution of archaeological $^{14}$C dates are best suited to regional or even continental scales of analysis, large $^{14}$C date datasets are needed. Therefore, the practical and economical application, evaluation, and development of the method require access to online databases containing sufficiently detailed contextual and technical information on the dates. Such databases already exist or are under construction, but many geographical areas and periods are still not covered by these databases. There is, thus, an urgent need for new databases, as well as discussion about the database conventions and best practises (see also Kintigh et al. 2015).

Climate envelope or niche modelling tools are well established in biogeography and ecology. Their application to human populations may still require evaluation, and the tuning of methods that are best suited to the analysis of hunter-gatherers as relatively small training sets and, in some cases, complex non-linear relationships between variables pose problems that are not so common when the focus is on other species.
Conclusions

In addition to methodological issues, there is definitely still a need for the deeper analysis of the relative importance of different climatic and other environmental factors on hunter-gatherer population dynamics using both ethnographic snap-shot and archaeological long-term data.

Despite these challenges, this study has hopefully shown the usefulness of the described approaches. The model simulations are powerful tools to analyse past dynamics, but they are not perfect representations of the past reality. Archaeological data will always have the primary role in the study of prehistory, but neither do they provide perfect, unbiased representations of past reality. Therefore, we need multiple lines of evidence and multiple approaches in the study of prehistory, and especially in the study of prehistoric population dynamics.
REFERENCES


References


References


References


References


References


References


82


