G Model
ANCENE 86 No. of Pages 10

ARTICLE IN PRESS

Anthropocene xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Anthropocene

journal homepage: www.elsevier.com/locate/ancene

Radiocarbon dates as estimates of ancient human population size

Michelle A. Chaput*, Konrad Gajewski

Laboratory for Paleoclimatology and Climatology, Department of Geography, Environment and Geomatics, University of Ottawa, 60 University Private, Ottawa, Ontario, K1N 6N5, Canada

ARTICLE INFO

Article history:
Received 8 September 2015
Received in revised form 26 October 2015
Accepted 28 October 2015
Available online xxx

Keywords:
Anthropocene
Radiocarbon dates as data
Frequency distribution
Holocene
Paleodemography
HYDE 3.1

ABSTRACT

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

1. Introduction

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

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

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

In parallel to this model-based approach is an empirical approach which reconstructs past conditions using fossil or ethnographic data, or a combination of both. A large amount of information about human technological adaptations to past environments is provided in archaeological and ethnographic

* Corresponding author.
E-mail address: mchap036@uottawa.ca (M.A. Chaput).

http://dx.doi.org/10.1016/j.ancene.2015.10.002
2213-3054/

records, and a long history of archaeological investigations and paleoenvironmental studies from around the world has documented Holocene human–environment interactions. Approximate and qualitative estimates of population sizes in the Americas have been attempted through various means including theoretical calculation and ethnographic studies (e.g., Johnson, 2014; Johnson et al., 2015; Newson, 1996), but have arrived at greatly different estimates (Haynes et al., 2007; Waters and Stafford, 2007), and in any event require verification. Data from networks of archaeological sites enable the mapping and analysis of the magnitude of human impacts on the landscape, and also the potential effects of climate or environmental change on local human activities or population size.

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

If we are to arrive at more definite estimates of ancient population numbers, a necessary input for research on human–environment interactions in the Anthropocene, we must combine the results from current data-based and model-based studies. In this paper, we discuss the methodology of obtaining estimates of the spatiotemporal population distribution of ancient humans using ^{14}C databases. We assess the current state of regional ^{14}C datasets which can serve as the base for paleodemographic estimates and evaluate their potential for producing empirical estimates of past population density. Our comparisons of the most complete ^{14}C databases from North America and Australia with HYDE 3.1 population estimates (Klein Goldewijk et al., 2010, 2011) indicate that ^{14}C data can effectively be used to evaluate model-based estimates of past populations. Thus, they can serve as inputs to studies quantifying human-induced paleoenvironmental change or evaluating the early anthropogenic hypothesis.

2. Archaeological radiocarbon data at a continental scale

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

2.1. Radiocarbon background and methods

The statistical analysis of compilations of ^{14}C dates has a long history in the fields of Quaternary geology and paleoclimatology (e.g., Dyke, 2005; Gajewski et al., 2006; Wendland and Bryson, 1974). In archaeology, "dates as data" is the common name for the practice of using ^{14}C dates as a form of data from which spatiotemporal paleodemographic change may be inferred. It was first used in archaeology by Rick (1987) who suggested that an accumulation of culturally-associated dates in any given area should be related to the degree of human presence on the landscape, if inherent biases and sources of error are accounted for.

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

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

2.2. Interpreting radiocarbon datasets

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

1. Quality issues: an initial assessment of data quality and reliability should be done prior to any in-depth analysis. For example, a verification that dates have been logically assigned to the correct cultural phase based on its age range may be necessary (Gajewski et al., 2011). Duplicate laboratory identification numbers and large standard deviations should also be investigated. Flohr et al. (2015) describe a well-documented series of steps that could be followed. More generally, Woods (2015) explains the need for consistent publication practices when reporting ^{14}C results.
2. Sample size: as with most point data, a representative sample must first be obtained. It is unclear what constitutes a representative sample of ^{14}C dates, but in general a larger number of dates is preferable. Williams (2012) suggested that at least 500 dates are needed to confidently discuss dominant population trends. However, Peros et al. (2010) showed that even a small ($\sim$0.0001%) sample can be representative of the source population. In another study at a continental scale of North America, Chaput et al. (2015) randomly removed 50% (n = 16,894) of their data and this did not alter their conclusions related to paleodemography, suggesting sample size requirements may change on a case by case basis. In any event, the statistical field of sampling theory provides a guide to this question (e.g., Cochran, 1963).
3. Sampling bias: in some instances, sampling bias issues can be difficult to address. Some areas are studied more, or are more accessible, which would lead to more dates from that region. It is often the case that the sampling strategies used by archaeologists are designed with a particular region, site or culture in