Journal of Archaeological Science 87 (2017) 59–72

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

Comparing archaeological proxies for long-term population patterns: An example from central Italy

Alessio Palmisano*, Andrew Bevan, Stephen Shennan

Institute of Archaeology, University College London, United Kingdom

ARTICLE INFO

Article history:
Received 25 July 2017
Received in revised form
27 September 2017
Accepted 2 October 2017

Keywords:
Demography
Summed probability distributions
Radiocarbon
Aoristic methods
Mediterranean
Archaeological survey
Settlement patterns

ABSTRACT

Raw counts of archaeological sites, estimates of changing settlement size and summed radiocarbon probability distributions have all become popular ways to investigate long-term regional trends in human population. Nevertheless, these three archaeological proxies have rarely been compared. This paper therefore explores the strengths and weaknesses of different kinds of archaeological evidence for population patterns, as well as how they address related issues such as taphonomic loss, chronological uncertainty and uneven sampling. Our overall substantive goal is to reconstruct demographic fluctuations in central Italy from the Late Mesolithic to the fall of the Roman Empire (7500 BC-AD 500), and with this in mind, we bring to bear an unusually detailed and extensive dataset of published central Italian archaeological surveys, consisting of some 10,971 occupation phases at 7383 different sites. The comparative results demonstrate reassuring consistency in the suggested demographic patterns, and where such patterns diverge across different proxies (e.g. Late Bronze Age/Iron Age) they often do so in useful ways that suggest changes in population structure such as site nucleation or dispersal.

1. Introduction

Over the past decade or so, there has been renewed archaeological interest in demographic reconstruction, in close step with other trends, such as the growing popularity of both cultural evolutionary and human ecological frameworks. The role of population size as a driver of cultural change was perhaps first emphasised by social anthropologists (Naroll, 1956; Carneiro, 1962) and then adopted by archaeologists to explain long-term variation in subsistence-settlement systems (Binford, 1968; Sanders and Price, 1968; Shennan, 2000, 2001) or shifts in sociopolitical complexity (Feinman and Neitzel, 1984; Feinman, 2011). More recent studies, in Europe for example, have stressed the upward-impact of Neolithic economies on local population densities (Shennan and Edinborough, 2007; Shennan, 2009; Shennan et al., 2013) from the often-lower population levels present when hunter-gatherers were active in the same region. Such discussions also feed into ongoing debate about whether agricultural innovation and intensification typically develops in response to population growth or vice versa (Boserup, 1965; Cohen, 1977; Netting, 1993; Peregrine, 2004), while a range of separate research continues to emphasise how population growth in a given landscape has typically run in step with increasingly substantial cultural modifications, often in a clearly coupled human demographic-ecological system (see Butlin and Roberts, 1995; Allen, 2001; Mercuri et al., 2002; Fyfe et al., 2010; Walsh, 2013; Langgut et al., 2016; Wigand and McCallum, 2017).

With this wider background in mind, it is clear that successful characterisation of human population fluctuations over the *longue durée* (and assessment of the causes of these fluctuations) is pivotal for how we understand cultural and environmental change. While genetic (both modern and ancient) or palaeodemographic (osteological) estimates of changing population size are also important (e.g. Bocquet-Appel, 2002; Cassidy et al., 2016), the most popular archaeological proxies for investigating regional demographies over the long-run have been data on counts of archaeological sites, sometimes with accompanying estimates of changing settlement size, and the summed probability distributions of radiocarbon dates (hereafter SPD). The first two have a longer archaeological pedigree in being used to estimate population across many different regional contexts (Sanders, 1965; Adams, 1965, 1981; Wright and Johnson, 1975; Sanders et al., 1979). More recently, over the past two decades, SPDs of archaeological (i.e. anthropogenic) radiocarbon dates have also become popular especially for

* Corresponding author.
E-mail address: a.palmisano@ucl.ac.uk (A. Palmisano).

https://doi.org/10.1016/j.jas.2017.10.001
0305-4403/

inferring population in prehistoric periods (Rick, 1987; Shennan and Edinborough, 2007; Bocquet-Appel et al., 2009; Shennan et al., 2013; Downey et al., 2014; Timpson et al., 2014; Balsera et al., 2015; Crema et al., 2016) and for assessing demographic responses to climate change (Weninger et al., 2009; Williams et al., 2010; Maher et al., 2011; Woodbridge et al., 2014; Flohr et al., 2016). Nevertheless, these proxies are rarely compared directly. Building on previous work (e.g. Tallavaara et al., 2010; French, 2015; French and Collins, 2015; Demjan and Dreslerov, 2016), we advocate greater use of multiple lines of demographic evidence and here present a comparison of radiocarbon SPDs and various modelled treatments of settlement counts and sizes for central Italy from the Late Mesolithic (7500 BC) to the fall of the Roman Empire (AD 500).

Estimating past population has generally been considered a problematic goal by most archaeologists, but the past couple of decades has seen a slow resurgence of interest in reconstructing demographic variables. Population estimates build on the assumption that an observable density of archaeological evidence over time and across a study region is somehow proportional to population despite the presence of certain archaeological biases (Drennan et al., 2015, 11). Put simply, the bigger the population, the stronger the signal in the archaeological record (e.g. the higher the density of pottery sherds, stone tools, site counts, radiocarbon dates, etc.). Hence, the first step in modelling population dynamics over the long-term is to identify those archaeological materials that might provide the most reliable indirect measures of population, and exclude those more strongly affected by other factors. Furthermore, it is usually assumed that such indicators do not offer good evidence for *absolute* numbers of people in the past, but rather offer an idea of *relative* intensities of population and proportional change through time (Tallavaara et al., 2010, 252; Drennan et al., 2015, 12). In this work, we use three types of archaeological data as proxies for estimating population fluctuations over the long run: 1) Settlement data including site counts; 2) summed estimated settlement sizes, effectively a weighted version of site counts; and 3) SPDs of radiocarbon dates. Two main potential issues common to three lines of evidence relate to the presence of both research and taphonomic biases, which can negatively affect the density and visibility of the archaeological signal known in a given region. For example, all archaeological periods are not necessarily equally represented in either settlement data or radiocarbon date lists, due to a series of factors: 1) the research priorities of different archaeological excavations and surveys resulting in specific periods being better investigated than others; 2) variation in the field methods adopted; and 3) the enhanced visibility of particular diagnostic artefacts that are easier to detect and collect. In addition, the archaeological record has been shaped by a wide variety of natural and cultural taphonomic processes (e.g. agriculture, erosion, alluviation, post-depositional deposits, human and animal excavations, wind deflation, etc.; cf. Roper, 1976, 372; Hirth, 1978, 125; Ammermann, 1985, 33; Gregg et al., 1991; Brantingham et al., 2007). Several studies have argued that a broad gradient exists in which there is increasing taphonomic loss with increasing time depth, or put another way, a higher level of destruction of earlier archaeological deposits (Surovell and Brantingham, 2007; Surovell et al., 2009), leaving them underrepresented when compared with the more recent deposits.

Turning more specifically to settlement evidence, a “site count” approach to population inference is typically based on the assumption that the overall number of sites is representative of population across space and time, but such counts can of course be biased by the intensity of archaeological surveys carried out in a given region (Plog et al., 1978; Cherry, 1983), by the ease with which a given site type can be observed and discovered archaeologically, etc. In addition, it is sometimes difficult to distinguish settlements from other kinds of site (e.g. cemeteries, specialized ritual sites, temporary agricultural or hunting installations), and, even if we can do so, to decide what kinds of site should be part of the counting exercise. A further issue is that we struggle to date the creation, duration and abandonment of sites and, without the support of stratigraphic data and/or calibrated radiocarbon dates, a given site's profile of occupational intensity through time can be only established by rough assessment of the stylistic chronologies of artefacts recovered from it. We are similarly left uncertain about the relative permanence vs seasonality of site use or about whether there is exact contemporaneity among multiple sites across a wider landscape. The spatial structure and size hierarchies of settlement sites are a further key variable that is often poorly understood. For example, a simple site count rarely does justice to changing population levels where a settlement system exhibits a move towards growing concentration of people in a few larger centres and we have to be able to observe large, contiguous spatial regions of settlement to understand how such a nucleation process plays out. Paying attention to estimates (from survey and excavation) of site size is therefore a useful addition to site counts, and typically rests on the assumption that the number of inhabitants is somehow proportional to the area of a settlement. Nonetheless, this correlation is neither likely to scale in a linear way (e.g. larger cities are often also more densely packed, albeit with less inhabited, functionally specialist zones as well) nor to be universally consistent across different regions of the world (Drennan et al., 2015, 20–25).

Turning to radiocarbon dates, large lists of archaeological radiocarbon dates can be calibrated and counted up (summed in the manner of a histogram) as a proxy for population, based on the assumption that the more people living in a given region, the more the archaeological deposits, the more organic materials, and the more radiocarbon samples collected and dated (Rick, 1987). Although this approach has been widely used by archaeologists for estimating population fluctuations for the Paleolithic and the Neolithic, it faces several challenges, in addition to the general ones discussed above, which may undermine its validity (Williams, 2012; Contreras and Meadows, 2014; Torfing, 2015). First, radiocarbon samples are often strategically collected for dating stratigraphic sequences within a site and, therefore, are not a random sample of human activity in every phase. Second, both the instrumental error associated with each date and the radiocarbon calibration curve have effects on the shape of each calibrated date's probability distribution and hence on the SPD of all summed calibrated dates (Michczyński and Michczyńska, 2006, 4; Williams, 2012, 581–584; Weninger et al., 2015). Third, research budgets can determine the extent to which radiocarbon samples are collected and used in an archaeological excavation, so some regions are richer in collected dates than others. Finally, certain chronological periods are more likely to be sampled than others: if datable coins, documents or fine-ware pottery exist, for instance, there is typically greater reliance on these forms of chronological evidence and less interest in paying for expensive radiocarbon dates.

Although the SPD of radiocarbon dates, site counts, and estimated settlement sizes have been widely used as proxies for population, the above limitations point to a need for cross-comparison among them where possible to strengthen our overall interpretation of demographic trends through time. The resolutions of these different kinds of evidence vary as well: an SPD of radiocarbon dates usually provides better chronological resolution, but typically less geographical coverage and control over sampling quality, when compared with site counts and estimated site sizes, but the latter are usually time-sliced to a much coarser level of resolution. For sites, there is a further imbalance between the kinds of evidence produced by extensive and methodical archaeological excavation in a given region, versus use of archaeological surface survey data. In