Upper Palaeolithic population histories of Southwestern France: a comparison of the demographic signatures of $^{14}$C date distributions and archaeological site counts

Jennifer C. French $^a$,*, Christina Collins $^b$

$^a$ McDonald Institute for Archaeological Research, Department of Archaeology and Anthropology, University of Cambridge, Downing Street, Cambridge, CB2 3ER, UK
$^b$ School of Archaeology, Classics and Egyptology, University of Liverpool, 12-14 Abercromby Square, Liverpool, L69 7WZ, UK

ARTICLE INFO

Article history:
Received 13 September 2014
Received in revised form 31 December 2014
Accepted 2 January 2015
Available online 14 January 2015

Keywords:
Prehistoric demography
Temporal frequency distributions
Upper Palaeolithic
Southwestern France

ABSTRACT

Radiocarbon date frequency distributions and archaeological site counts are two popular proxies used to investigate prehistoric demography, following the assumption that variations in these data reflect fluctuations in the relative size and distribution of past populations. However, the two approaches are rarely applied to the same data-set and their applicability is heavily conditioned by the archaeological record in question, particularly research histories, agendas, and funding availability. In this paper we use both types of data to examine the population history of the Upper Palaeolithic hunter-gatherers (~40 000–12 000 cal BP) of Southwestern France, comparing the demographic signatures generated.

Both proxies produce similar signatures across the Upper Palaeolithic sequence of the region, strengthening the interpretation of relative demographic changes as the cause of the pattern. In particular, a marked population decline is seen in both datasets during the Late Gravettian (~28 000 cal BP), as well as a population increase in the Late Solutrean (~25 000 cal BP) supporting the notion that the region acted as a population refugium during the Last Glacial Maximum. Where the two proxies diverge in the demographic signatures they produce, the radiocarbon date distribution shows peaks compared to troughs in site counts; the opposite pattern expected given taphonomic issues surrounding cultural carbon. Despite differences in chronological resolution and sampling bias, our data suggest that the two proxies can be considered broadly equivalent; a finding which warrants the investigation of prehistoric demography in regions where either extensive survey data or radiometric dating programmes are unavailable. While some preliminary observations are made, the impact of changing mobility on diachronic patterns seen in both proxies remains, however, difficult to assess.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Two popular archaeological approaches to investigating past population histories are examining temporal frequency distributions of the number of archaeological sites and radiocarbon ($^{14}$C) dates. Both methods are based on a common premise: that the frequency of the archaeological data serves as a proxy for past human occupation and that variations in these data reflect fluctuations in the relative size and distribution of past populations. Examples from the archaeological literature are abundant, using data from a range of geographical and chronological contexts at both regional and continental scales (e.g. Armit et al., 2013; Bocquet-Appel and Demars, 2000a, 2000b; Bocquet-Appel et al., 2009; Collins 2012; Gamble et al., 2004, 2005; Hinz et al., 2012; Holdaway et al., 2005; Kelly et al., 2013; Martínez et al., 2013; Mellars and French, 2011; Petraglia et al., 2009; Shennan and Edinborough, 2007; Shennan et al., 2013; Straus, 2011; Straus et al., 2000a; Wicks and Mithen, 2014; Williams, 2013; Williams et al., 2010) and frequently focusing on the relationship between climatic variability and human demographic fluctuations (Munoz et al., 2010; Kelly et al., 2013; Schmidt et al., 2012; Tallavaara and Seppä, 2011).

Despite both methods sharing the same underlying assumption of the relationship between the archaeological proxy material and past population histories, the validity of this shared assumption
remains largely untested. Following both previous studies which compare the two methods (Tallavaara et al., 2010), and calls for a multi-proxy approach to aid the verification and interpretation of palaeodemographic trends (Crombé and Robinson, in press; French, in press; Williams, 2012), we present here a comparison of the demographic signatures of \(^{14}\)C date distributions and archaeological site counts using the case study of the hunter–gatherer populations of the Upper Palaeolithic (~40 000–12 000 cal BP) of Southwestern France.

2. Background: using site counts and \(^{14}\)C date distributions to reconstruct prehistoric demographic patterns

Implicit in the use of both site counts and \(^{14}\)C date distributions (‘dates as data’; Rick, 1987) for demographic reconstruction is the assumption that for a given period the density of the proxy data is roughly proportional to the population, and the correlation between the density and distribution of archaeological material and past population is linear and uniform throughout (Bocquet-Appel et al., 2005 cf. Peros et al., 2010). Both approaches similarly assume that the density and distribution of archaeological material and the correlation between the density and distribution of archaeological material and past population is random and linear throughout (Bocquet-Appel et al., 2005 cf. Peros et al., 2010). Both approaches similarly assume ‘cumulative archaeological pressure’ (that the intensity of archaeological research was approximately uniform across the region under study until the completion of the database used) and ‘uniform archaeological pressure’ (that the observed frequency of the proxy data is a non-biased estimate of the original density) (Bocquet-Appel et al., 2009). Potential problems with these hypotheses relate to research biases and taphonomic biases respectively.

Biases in research intensity and history can affect not only the number of sites known in a region, but also the type, with cave and rock-shelter sites often better represented than open-air sites due to their visual prominence (Aubry, 2013:200; Bocquet-Appel and Demars, 2000b; 557–558). Some chronologically diagnostic artefact types are more visually distinctive than others. This enhanced visibility may ‘inflate’ the number of sites for that period; a concern which Straus et al. (2000a, 2000b) have voiced with regard to the striking ‘Solutrean points’ in their studies of demography and numbers of sites in Iberia. Archaeological budgets often dictate the extent to which radiocarbon dating is used in a regional context (Bocquet-Appel et al., 2009) and \(^{14}\)C date distributions can be affected by such biases as: 1) the type of sites from which the date originates, with caves and rock-shelters better dated than open-air sites due to a frequent lack of organic material in the latter (van Andel et al., 2003); 2) where in an archaeological sequence the date comes from, with a frequent bias towards earlier levels in order to establish initial occupation (Bird and Frankel, 1991:3), and; 3) when the site excavation took place, with older excavations often less well-dated than recent ones.

The impact of taphonomic bias on the demographic signature generated through the use of both approaches has been examined most prominently by Surovell and colleagues (Surovell and Brantingham, 2007; Surovell et al., 2009) who demonstrated that the positive curvilinear distributions regularly observed in these proxies in archaeological contexts also occur in natural palaeontological and geological contexts (see also Ballenger and Mabry, 2011; Johnson and Brook, 2011). Their findings suggest that any results should be treated with caution and could instead be interpreted in terms of increased taphonomic loss with time-depth, unless some form of ‘correction curve’ is applied (Surovell et al., 2009). In particular, they suggest using the underlying geological formations of the study region, and/or the distribution of cave and rock-shelter sites as a check against the frequency distributions of open-air sites and any radiocarbon dates which come from them (e.g. Kelly et al., 2013). However, taphonomic loss has been rejected as a catch-all explanation for many \(^{14}\)C date distribution patterns (Riede, 2009), and it has been proposed that taphonomic biases could be offset by other biasing factors such as the over- and under-representation of dates within a sequence as a result of differing research priorities (Peros et al., 2010).

Past human behaviours — whether dependent or independent of population fluctuations — could also impact on \(^{14}\)C date distributions and site counts. For example, changes in human burial practices or subsistence strategies can affect the amount of cultural carbon entering the archaeological record, and changing patterns of mobility and land-use strategies would impact on the number of sites created. While both methods are not without their controversies, among their advocates the consensus is that both \(^{14}\)C dates and site counts are appropriate palaeodemographic proxies for long-term relative population fluctuations, with increased confidence being placed in the validity of any trends identified for regions with long and intensive histories of archaeological research and/or for which a large number of reliable \(^{14}\)C dates are available (Shennan and Edinborough, 2007:1340).

The relative merits of each approach have been summarised by Bocquet-Appel et al. (2005), with archaeological site counts providing better geographical sampling but with lower chronological resolution, while \(^{14}\)C date distributions provide much higher chronological resolution but with less control over sampling quality. In particular, the use of site counts suffers from the increased effect of what is termed the ‘contemporaneity problem’; the difficulty of assessing whether sites that belong to the same general archaeological period were occupied simultaneously or sequentially (Schacht, 1984). Nonetheless, the use of site counts is the best available palaeodemographic proxy in regions which lack extensive archaeological excavation and for which only survey data are available. Thus, the choice of either method is largely constrained by the archaeological record of the study region, rather than representing a deliberate decision to prioritise either chronological or geographical resolution in the study of demographic patterns. The archaeological record of the Upper Palaeolithic of Southwestern France is ideally suited as a case-study to compare the demographic signature generated from both types of data. Evidence of continuous human occupation in the late Pleistocene has resulted in a long history of excavation and survey limiting the impact of past research biases, and providing a large database which permits the valid application of both approaches.

3. Materials and methods

3.1. The Upper Palaeolithic of Southwestern France

The area of Southwestern France examined centres on the modern département of Dordogne, and incorporates the surrounding départements of Charente-Maritime, Charente, Corrèze, Lot, Lot-et-Garonne, and Gironde, spanning approximately 1.5° in latitude, from 44°3’N in Lot-et-Garonne to 45°7’N in Charente-Maritime (Fig. 1). The study area covers ~50 000 km²; an area large enough to accommodate the range in ethnographically documented hunter–gatherer occupation areas and annual ranges (see Binford, 1983:110; Kelly, 2013) — which we assume are useful estimates when investigating prehistoric hunter–gatherer populations — and which permits an appropriate scale for the understanding of regional demographic processes. The available area for occupation was undoubtedly larger during the Upper Palaeolithic and has been reduced due to subsequent global changes in sea level (Lambeck et al., 2002). The dominant geological features of the region are the caves and rock-shelters which are especially abundant in the Dordogne (Tixier, 2009:12), and from which the majority of archaeological evidence from the Upper Palaeolithic has been recovered.
The Upper Palaeolithic occupation of the region dates from ~40,000–12,000 cal BP. Human occupation occurred against a background of dramatic and often rapid global climatic changes, including the millennial-scale D-O cycles (Dansgaard et al., 1993; Grootes et al., 1993; Svensson et al., 2008) and the less frequent, but more severe, Heinrich events (Andrews, 1998; Hemming, 2004) identified in Greenland ice core and ocean core sediments respectively. Data from a range of terrestrial and marine climatic proxy...
records from the region indicate that these global climatic changes had local climatic and environmental effects (Ampel et al., 2008, 2010; Bertran, 2005; Bertran et al., 2008, 2013; Daniau et al., 2009; Genty et al., 2003, 2010; Naughton et al., 2007, 2009; Sánchez-Góñi et al., 2008; Wainer et al., 2009; Wohlfarth et al., 2008), although these may have been subject to time-lags (Blaauw et al., 2010).

Archaeologically, the Upper Palaeolithic of Southwestern France is divided into five main successive chrono-typological periods or industries: Aurignacian, Gravettian, Solutrean, Magdalenian, and Azilian (Table 1). The Chatelperronian, which immediately precedes the Aurignacian in the region, is frequently considered to belong to the Upper Palaeolithic, although as the consensus is that this industry was manufactured by Neanderthals, rather than *Homo sapiens* (Bailey et al., 2009; Mellars, 2005; Pelegrin and Soressi, 2007; cf. Bar-Yosef and Bordes, 2010), we have excluded it from our study.

These five phases show a clear chronological and stratigraphic succession in the region, originally identified through the presence of diagnostic lithic ‘type-fossils’ and later confirmed by radiometric dates. These phases are divided into sub-phases, again based on the presence of diagnostic lithic (or bone) types, although ambiguities are present at certain stages of the sequence due to such factors as conflicts between absolute dates and stratigraphy, the validity of the extrapolation of sequences from key sites to the wider region, and questions surrounding the chronologically diagnostic nature of certain type fossils (e.g. Aubry and Almeida, 2013; Bon, 2002; Ducasse, 2012). In view of these ambiguities we adopted simplified sub-divisions of each of the five periods for this study (Table 1; see Supplementary Material for details).

### 3.2. 14C date distributions

Radiocarbon dates for the study region were collected from the literature (Table S1, Supplementary Material). The majority of radiocarbon dates from Southwestern France are produced from bone samples, although some charcoal samples were also included. Dates were calibrated using OxCal 4.2 (Bronk Ramsey, 2009a) and the IntCal13 calibration curve (Reimer et al., 2013). Where possible, Bayesian models were built for sites (Table S2, Supplementary Material). These models utilize the prior information available to archaeologists (e.g. stratigraphic sequences). This information can be formalized and included in the model to reduce the distribution of the calibrated date and to identify outliers, which are then transformed into more realistic date distributions based on stratigraphic data. We utilized the boundary and phase constructions within OxCal (Bronk Ramsey, 2005), with dates from phases at individual sites grouped together, and separated by boundaries. Outlier analysis was also conducted, using a uniform prior probability of 0.05 for each date, corresponding to a 1 in 20 probability of each radiocarbon date being an outlier (Bronk Ramsey, 2009a). Any outlying dates identified were down-weighted in the analysis (Bronk Ramsey, 2009b), eliminating the need for their exclusion (see Supplementary Material for further details).

Models were built for sites that featured multiple radiocarbon dates in good stratigraphic sequence, but not for sites yielding only one or two radiocarbon dates, or where the stratigraphy was unreliable. Several studies have suggested criteria for the inclusion or exclusion of radiocarbon dates according to chronometric hygiene, proposing that *a priori* unreliable dates (whether due to contamination/laboratory error or deriving from a questionable context) should be excluded from analyses (Pettitt et al., 2003; Spriggs, 1989; Waterbolk, 1971). However, the use of site-specific Bayesian models with outlier analysis should identify any unreliable dates and this will be reflected in the posterior distribution obtained, with outlying dates down-weighted in the analysis (see Figs. S1, S2 Supplementary Material). We therefore felt it unnecessary to exclude dates which might be regarded as outliers *a priori*, as these outlying dates will be ‘corrected’ in the Bayesian models. We then compared the distributions obtained from the modelled and unmodelled dates to help to understand the effect (if any) of the inclusion of dates from sites that could not be modelled (Fig. S3 Supplementary Material). For example, Fig. 2 shows a sequence of dates from the site of Les PeyruGES (Lot), constructed in OxCal 4.2 using the boundary and phase constructions. The dates from phases 6 and 7 clearly lie outside of the sequence. These original calibrated,

<table>
<thead>
<tr>
<th>Period</th>
<th>Approx. dates (cal BP, IntCal13)</th>
<th>Length of period (1000 years)</th>
<th>Traditional French systematic sub-phases</th>
<th>Sub-stages used here</th>
<th>Dates of sub-stages (cal BP, IntCal13)</th>
<th>Length of sub-stages (1000 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aurignacian</strong></td>
<td>39,500–34,000</td>
<td>5.5</td>
<td>Aurignacian I</td>
<td>Early Aurignacian</td>
<td>39,500–36,000</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aurignacian II</td>
<td>Late Aurignacian</td>
<td>36,000–34,000</td>
<td>2.0</td>
</tr>
<tr>
<td><strong>Gravettian</strong></td>
<td>34,000–26,100</td>
<td>7.9</td>
<td>Perigordian IV</td>
<td>Early Gravettian</td>
<td>34,000–31,500</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Perigordian Va</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Perigordian Vb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Perigordian Vc/Noaillian</td>
<td>Middle Gravettian</td>
<td>31,500–29,500</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Perigordian VI</td>
<td>Late Gravettian</td>
<td>29,500–26,100</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Perigordian VII/Proto-Magdalenian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Solutrean</strong></td>
<td>26,100–24,600</td>
<td>1.5</td>
<td>Proto-Solutrean</td>
<td>Early/Middle Solutrean</td>
<td>26,100–25,500</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower Solutrean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Middle Solutrean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Upper Solutrean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Magdalenian</strong></td>
<td>24,600–15,500</td>
<td>9.1</td>
<td>Final Solutrean</td>
<td>Late Solutrean</td>
<td>25,500–24,600</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Badegoulian/Magdalenian 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Magdalenian I</td>
<td>Early Magdalenian/Badegoulian</td>
<td>24,600–21,800</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Magdalenian II</td>
<td>Middle Magdalenian</td>
<td>21,800–18,200</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Magdalenian III</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Magdalenian IV</td>
<td>Upper Magdalenian</td>
<td>18,200–17,000</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Magdalenian V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Azilian</strong></td>
<td>15,500–11,500</td>
<td>4</td>
<td>None</td>
<td>Final Magdalenian</td>
<td>17,000–15,500</td>
<td>1.5</td>
</tr>
</tbody>
</table>

*Table 1* Chronological scheme for the Upper Palaeolithic adopted for this study, indicating chronological ranges, duration of each techno-complex and sub-phases and the relationship between the framework used here and traditional French Upper Palaeolithic systematics. The lack of available radiocarbon dates for the Middle Solutrean means that we were unable to estimate the duration of this sub-stage of the techno-complex, and it has been combined with the Early Solutrean in this analysis. See Supplementary Material for further information on how date ranges were obtained.
outlying, dates have been down-weighted, and new model dates constructed on the basis of dates from elsewhere in the sequence. In this manner, outlying dates can be altered into more realistic dates, which is preferable to merely excluding samples. This is particularly true for demographic studies, where we are chiefly interested in the frequency of anthropogenic carbon in the archaeological record, and the exclusion of dates that could otherwise be modelled in this way would erode the emerging picture of demographic patterns and potentially reinforce some of the research biases discussed in Section 2.

After calibration, the raw data were collated from OxCal, in the form of a probability distribution for each date, with a probability density for every five-year interval of the range covered. In total, 502 radiocarbon dates were included in the distribution. This exceeds the minimum sample requirement set out by Williams (2012). To overcome research bias, these probability distributions were normalized through averaging within sites to prevent the over-representation of well-dated sites, as described by Grove (2011). Essentially, the multiple radiocarbon dates from each site were collapsed into a single distribution, ensuring that no single site was over-represented in the final analysis. These distributions were then summed across the entire region and study period. Although frequently used to correct radiocarbon temporal frequency distributions for taphonomic loss, the curve of Surovell et al. (2009) is not applicable to this data-set as all of the dates derive from cave/rock-shelter sites, rather than the open-air sites for which the curve was devised. While not denying that some destruction of cave and rock-shelter deposits has occurred, at present no similar curve exists to correct for this loss, and these deposits are generally better protected than those at open-air sites.

One common criticism is that the primary visible signal in radiocarbon date summed probability distributions is that of the calibration curve (Blackwell and Buck, 2003; Chiverrell et al., 2011 cf. Miller and Gingerich, 2013). To control for the shape of the calibration curve, a simulated dataset was produced with radiocarbon dates uniformly distributed across the study period. The simulated dataset was produced using the R_simulate function in OxCal 4.2 (Bronk Ramsey, 2005), running the simulation for 1000 iterations. The R_simulate function simulates a radiocarbon date for a calendar date entered, and we used it to produce a series of radiocarbon dates that are uniform in calendar time. The summed probability distribution produced through this method should represent the distribution that would be expected if the regional population had remained static across the duration of the Upper Palaeolithic and if no taphonomic loss occurred (cf. Shennan et al., 2013). Comparison of the two distributions allows us to observe demographic events, rather than the effects of the calibration curve. To assess further whether the calibration curve was affecting the shape of the distribution for this data set, we compared the distribution produced using the IntCal13 calibration curve with that produced using the earlier IntCal09 curve (Reimer et al., 2009).

3.3. Archaeological site counts

Site counts were collected through a literature review, compiled into a database which includes all the reported sites in the study region which are dated to any of the five phases of the Upper Palaeolithic discussed in Section 3.1 (Table S3, Supplementary Material). 542 sites (865 occupations, as many of the sites were occupied in multiple periods) were included. A ‘site’ was defined as any location where at least one lithic artefact chronologically diagnostic of any of the five study periods was present. We do not assume that all the sites included in the database represented permanent, long-term, habitation sites, nor that all the sites documented for each phase were occupied simultaneously; the figures generated represent averages across the periods in question. Each period was divided further into the sub-stages presented in Table 1. To account for the different lengths of each period and sub-stages, site counts were standardised by converting these into estimates of the number of sites/1000 years, assuming the approximate lengths given in Table 1 and rounded to the nearest whole number. To compensate for any potential taphonomic loss relative to time depth, the correction curve of Surovell et al. (2009) was applied to the open-air site counts, taking the chronological midpoint of each period as the correction factor.

4. Results

4.1. $^{14}$C date distributions

Fig. 3 depicts the summed probability distribution of radiocarbon dates for the region, based on all dates both modelled and unmodelled. We compared this with a simulated dataset which assumes that the regional population was static and that no taphonomic loss occurred. While the ‘real’ and simulated datasets share some features, the simulated dataset conforms to a uniform distribution, while the real dataset does not. The two distributions are clearly very different, which we interpret as indicating that population density fluctuated throughout the Upper Palaeolithic in Southwestern France. However, the possibility remains that such factors as regional dating strategy are impacting the distribution; a factor we consider in more detail in Section 5.1.
The radiocarbon date distribution shows peaks in activity (indicating an increase in population) at 30,000 cal BP, 27,000 cal BP, 25,000 cal BP, 20,000 cal BP, 18,000 cal BP, 15,000 cal BP and 13,000 cal BP. These peaks are punctuated by troughs (indicating population decrease), including deep troughs between 28,000 cal BP and 26,000 cal BP, separated by a small peak. The calibration curve used (IntCal09 or IntCal13) has a limited effect on the probability distribution produced, with the observable differences present relating to the relative magnitude of fluctuations, rather than the overall pattern of the distribution (Fig. 4). However a peak present at ~32,000 cal BP in the IntCal09 distribution is absent in that produced by IntCal13; a discrepancy for which we can think of no obvious explanation.

### 4.2. Archaeological site counts

Numbers of archaeological sites/1000 years vary greatly across the Upper Palaeolithic of Southwestern France. Peaks in number of archaeological sites are seen in the Middle Gravettian, Solutrean and Upper/Final Magdalenian. Dips in the number of archaeological sites are seen in the Early Gravettian, Badegoulian, Middle Magdalenian and Azilian, with the lowest number seen in the Late Gravettian (Fig. 5).

To correct for taphonomic loss the correction curve of Surovell et al. (2009) was applied to both the open-air sites and those for which the designation of sites as either sheltered (cave/rock-shelter) or open were unknown. For each period the gross number of sheltered sites, the gross number of open-air sites, and the corrected value for open-air sites were scaled between 0 and 1 by dividing each value by the maximum observed value in that category to create a relative temporal frequency distribution of numbers of different types of sites (Fig. 6). The shape of the distributions of corrected and un-corrected open-air sites are very similar, suggesting that taphonomic loss has not greatly affected the known number of open-air sites in the region. However, while some similarities are present (most noticeably the decreases seen in both site types ~33,000 cal BP), the distributions of open-air and sheltered sites are different throughout the Upper Palaeolithic, even when open-air sites are corrected for taphonomic bias. Key differences include; 1) a peak in the frequency of open-air sites ~23,000 cal BP (Badegoulian) compared with a dip in the frequency of sheltered sites, and; 2) a decrease in the frequency of open-air sites in the latter part of the sequence (20,000–16,000 cal BP), contrasting with an increase in the relative frequency of sheltered sites. These differences are problematic for the interpretation of relative demographic trends from these data. One possible explanation for this divergence is that the study region is outside of the geographical region of applicability of Surovell et al.’s (2009) correction curve. Periods of clear divergence between the frequency distributions of sheltered and open-air sites could also reflect the deliberate and preferential selection by Upper Palaeolithic populations of either type of site.

**Fig. 4.** Comparison of the ‘real’ and simulated probability distributions of radiocarbon dates. The solid line depicts the summed probability distribution for the region, based on calibrated radiocarbon dates (n = 502), both modelled and unmodelled. The broken line is the simulated dataset (n = 502).

**Fig. 5.** Numbers of archaeological sites in Southwestern France across the Upper Palaeolithic displayed as sites/1000 years of each techno-complex. EA = Early Aurignacian; LA = Late Aurignacian; EG = Early Gravettian; MG = Middle Gravettian; LG = Late Gravettian; E/M S = Early/Middle Solutrean; LS = Late Solutrean; B = Badegoulian; MM = Middle Magdalenian; UM = Upper Magdalenian; FM = Final Magdalenian; AZ = Azilian.
4.3. A comparison of the two approaches

Despite the different chronological resolutions of the two proxies, the overall patterns are broadly similar, and while the $^{14}$C date distribution shows fluctuations within chrono-typological stages, several points of convergence are found between the two datasets (Fig. 7). Most prominently, both $^{14}$C dates distributions and site counts document a deep trough during the Late Gravettian (~28 000 cal BP). Similar agreement between the two proxies are found in peaks at ~29 500 cal BP (Middle/Late Gravettian interface), ~25 000 cal BP (Late Solutrean), and ~18 000 cal BP (Upper Magdalenian). Noticeable differences in the distributions produced by the two datasets are found at ~20 000 cal BP (Middle Magdalenian) and ~13 000 cal BP (Azilian). In both cases of noticeable divergence the $^{14}$C date distribution is dominated by peaks, where dips are present in the numbers of archaeological sites.

5. Discussion

There are fluctuations in both $^{14}$C date distributions and numbers of archaeological sites during the Upper Palaeolithic in South-Western France. While we cannot entirely discount the role of other factors, the pattern produced suggests that neither time-transgressive taphonomic bias, nor effects of the calibration curve are impacting the overall distribution of either proxy. The lack of $^{14}$C dates from open-air locales is a potential biasing factor which we discuss in further detail in Section 5.1. The data from both proxies indicate a population increase in the (late) Middle Gravettian; a hypothesis previously postulated by David (1973, 1985), David and Bricker (1987). However, the peak seen in the site counts is more muted than that seen in the $^{14}$C date distribution and it is possible that this peak is enhanced by the drop-off in $^{14}$C contamination ~30 000 cal BP. It is noted elsewhere that the potential effects of contamination will increase with the age of a sample (Petit et al., 2003); very old samples will contain low amounts of $^{14}$C, and therefore the introduction of just a small amount of modern carbon can have dramatic effects on the perceived age of the sample. This effect could possibly be influencing the shape of the radiocarbon distribution in the Aurignacian and Early Gravettian, which is noticeably lacking in significant trends compared to the later stages of the Upper Palaeolithic distribution.

The two proxies produce broadly similar distributions; a finding also obtained by both Williams (2012) and Tallavaara et al. (2010) in earlier methodological comparisons. We consider the close correspondence between the distributions of the two types of data to enhance the reliability of the general trends noted above, and to strengthen the interpretation of relative demographic change as the cause of the observed pattern. In particular, while previous studies have questioned the reliability of many earlier conventional (as opposed to accelerator) radiocarbon ages from the Upper Palaeolithic of Southwestern France (e.g. Mellars et al., 1987), the general correspondence between the two proxies suggests that our use of outlier analysis in Bayesian models has identified and down-weighted any questionable dates, and that these are not significantly distorting the shape of the overall temporal frequency distribution (see Supplementary Material).

Our findings largely corroborate previous continental-scale research into the Upper Palaeolithic population histories of Europe (Bocquet-Appel and Demars, 2000b; Bocquet-Appel et al., 2005; Gamble et al., 2004, 2005). Overall, relative peaks and troughs do not show any consistent correlations with prevailing climatic conditions. Comparison of the high resolution $^{14}$C date probability distribution with the NGRIP ice core (used as a proxy for global temperature change) however, reveals that population peaks only occur during cold stages (Fig. 8). While these peaks do not occur during every cold stage, or last for their duration, a cross-correlation analysis of the NGRIP curve and the summed $^{14}$C probability distribution does show a weak negative correlation (coefficient = −0.224 (0 time-lag)). This contrasts with the data for ethnographic hunter-gatherers which show a positive correlation between temperature (through its impact of animal and plant biomass) and hunter-gatherer regional population densities (Binford, 2001; Layton and O’Hara, 2010; Marlowe, 2005). One possible explanation for this discrepancy is that the $^{14}$C dates are not measuring population fluctuations, but instead reflect fluctuations in mobility (e.g. Naudinot et al., in press; Niekus, 2005/2006). Increased mobility (which would appear as peaks in the $^{14}$C date distribution) is a common buffering response amongst
hunter–gatherers in the face of resource uncertainty and decreased productivity induced by climatic cooling (Binford, 1980; Grove, 2009).

The peak seen in both proxies during the Late Solutrean (~25 000 cal BP) supports the widely-cited notion of Southwestern France as part of a Franco-Cantabrian refugium during the Last Glacial Maximum [LGM] (Achilli et al., 2004; Jochim, 1987). The agreement between the two proxies suggests that the aforementioned enhanced visibility of the Solutrean points used as chronological diagnostic type-fossils for this period (Straus et al., 2000a, 2000b) are not artificially inflating the Late Solutrean site count estimate (as an isolated Solutrean lithic find-spot identified through the clearly visible type-fossils would be included in the site count analysis, but would not generate any 14C dates). In this refugium model, the increased population signature seen is likely the result of population movement from central and northern France, rather than representing in situ population growth (Demars, 1996). The region would then have ‘emptied out’ in the warmer Badegoulian/Early Magdalenian as resettlement of more northerly or adjacent high-altitude areas of France occurred (see, for example site distribution maps in Bocquet-Appel and Demars (2000b) and Demars (1996, 2002).

While increased mobility cannot be discounted as a possible explanation, given the role of the study region as a refugium identified during this cold phase, we suggest that the overall trend of a negative correlation between temperature and population seen in the 14C probability distribution indicates that Southwestern France may have acted as a refugium for hunter–gatherer populations during other cold stages of the Upper Palaeolithic (Bertran et al., 2013; 2274). For example, the peak seen in the 14C distribution ~15 000 cal BP and the subsequent trough ~14 000 cal BP reflects (to a lesser degree) the pattern seen during the LGM. The peak ~15 000 cal BP may have been caused by population retraction into the South-Western France refugium zone during the Greenland Stadial 2a cold stage with the relative trough ~14 000 cal BP attributed to the recolonisation of adjacent areas and population movement out of the region brought about by higher temperatures and climatic amelioration with the onset of Greenland Interstadial 1e (Bølling/Meierdorf) (Wolff et al., 2010). However, the broad chronological resolution offered by site distribution maps (Schacht, 1984) and the site counts analysis presented here makes testing this hypothesis difficult, particularly, when as in this instance, the two dates being compared fall within the same general techno-complex.

5.1. Differences between the two proxies

While the overall pattern generated by the two types of data is similar, as discussed in Section 4.3, there are some noticeable points of divergence found at ~20 000 cal BP and ~13 000 cal BP. Building on the discussion in Section 2, we offer some observations on three key contrasts between 14C date distributions and archaeological site counts which could potentially account for these differences.

5.1.1. Chronological resolution

As discussed previously, a fundamental difference between the two methods used is the much higher chronological resolution and control permitted by 14C date distributions, in contrast to the coarse resolution offered by archaeological site counts. The standardisation of site counts into numbers per 1000 years further diminishes the resolution; taking an average number of sites, and ignoring any (likely) changes in frequency that may have occurred throughout a given phase. As such, nuanced chronological variations in the number of archaeological sites may be obscured, which could explain the discrepancies in the patterns observed between the two proxies. For example, while the site counts indicate a
population decrease during the Late Gravettian, the enhanced resolution of the $^{14}$C date distribution shows considerable fluctuations within this phase with the population decrease ~28 000 cal BP flanked by increases at the interface with the earlier Middle Gravettian ~29 500 cal BP and ~27 000 cal BP.

Although the differences in scales of resolution can never be entirely rectified, as a test, we reduced the chronological resolution of the $^{14}$C date distribution, producing a moving-sum histogram with 1000 year bins to permit a more like-for-like comparison with the site count distribution (Fig. 9). To produce the histogram, the modal values of each radiocarbon posterior probability distribution were collated and tallied into 1000-year time bins. Where technocomplexes of the Upper Palaeolithic covered more than one 1000 year bin, fluctuations were seen within phases.

Taking the discrepancy between the two proxies at ~20 000 cal BP as an example, the peak seen in the summed probability distribution is more moderate in the moving sum histogram, with the ~20 000 cal BP bin representing a dip relative to the preceding bin. At both this specific point, and more generally across the Upper Palaeolithic sequence, the reduced resolution of the moving sum histogram produces a distribution more similar in shape to that produced via site counts. This suggests that methodological differences between the summed probability distribution and the moving sum histogram may be affecting the resultant demographic signature. It is important to remember that while the product of a calibrated $^{14}$C date is a probability distribution, the sample dated actually originated in a single calendar year. The distribution obtained through calibration can be misleading in this respect and this may account for the differences observed between the summed probability distribution and the moving sum histogram; accumulations of many ‘tails’ of distributions will produce a more complex pattern to that observed in the histogram. The histogram is both simpler, and arguably more realistic, as it reduces the probability distribution to a series of individual ‘events’ (the modal value of the distribution, instead of the entire distribution).

5.1.2. Taphonomy and material quantified

A further key difference between the two data-sets is the type of material being measured and quantified. The radiocarbon dates used in the analysis, have, by definition, been obtained from organic material, while site counts have been determined by the presence of lithic type-fossils. Although the highly calcareous limestone deposits of the study region provide excellent conditions for the survival of organic remains, a priori we would expect greater preservation of the durable lithic artefacts which form the database for the site counts analysis, than of the organic material which provide the dates for the $^{14}$C distribution. Furthermore, while open-air sites constitute ~35% of the sites included in the site counts analysis, very few of these yielded organic remains and none of the $^{14}$C dates included in the distribution originated from open-air sites. Thus, both generally and with regard to the current dataset, the demographic signature generated from site counts defined by the presence of inorganic material is likely to be inflated when compared to that produced using $^{14}$C date distributions. However, this distinction between the types of material being quantified cannot explain the differences seen between the two distributions presented here, as in both instances of major divergence peaks are seen in the $^{14}$C dates, with comparative troughs in the numbers of archaeological sites; the opposite pattern to that expected if the differential destruction of the two types of material is evoked as a major causative factor in the discrepancies between the two distributions.

5.1.3. Sampling and dating strategy

Stark differences are also present in the sampling methods used in each analysis. All radiocarbon dates came from well-stratified sheltered sites, which were probably habitation sites. In contrast,
the site counts analysis a priori covered a greater number of sites, and likely a greater range, than large habitation sites as well as short-term occupations and isolated find-spots on the landscape. It is questionable whether including non-habitation sites is appropriate for assessing past population changes (Bird and Frankel, 1991), and it is possible that the definition of a ‘site’ used in this study has artificially enhanced the demographic signature across the Upper Palaeolithic sequence. While this is difficult to assess, again we stress that in both instances of discrepancy between the demographic signatures produced using the two methods, the $^{14}$C distributions show an increased signature relative to the numbers of archaeological sites.

Almost half (48.6%) of the radiocarbon dates included in the analysis come from just 8 sites (Abri Pataud, Combe-Saunière, Cuzou de Vers, Le Flageolet I, Roc-de-Combe, La Ferrassie, Pégourie, Les Peyrugues) out of a total of 542 Upper Palaeolithic sites identified for the region through the site counts analysis. As described in the methodology, to eliminate bias from over-representation, dates from these 8 sites were normalized before summing across the region, as advocated by Grove (2011). Although an individual site may have produced multiple radiocarbon dates, it is reduced to a single probability distribution. Using this method, no single site will skew the regional summed probability distribution.

Fig. 10 shows the summed probability distribution of modelled radiocarbon dates from these 8 well-dated sites. The distribution produced solely from these sites differs from the regional distribution (Fig. 3) indicating that these sites are not over-represented in the final analysis, and illustrating how the method of first averaging within sites ensures that no single site can skew the distribution. It is also interesting to note the lack of peaks in the Late Upper Palaeolithic in the distribution obtained from these 8 key sites (compared to the overall distribution shown in Fig. 3) suggesting the tendency of earlier Upper Palaeolithic sites, such as the Abri Pataud, to be the focus of intensive dating programmes. In the absence of a catch-all explanation for the points of divergence in population histories seen between the two proxies throughout the Upper Palaeolithic sequence, we conclude that the differences must result from the specifics of the archaeological record at the chronological points in question. In particular, we suggest that the contrast between the two proxies −20 000 cal BP (Middle Magdalenian) can be accounted for by the lack of unequivocally chronologically diagnostic Middle Magdalenian lithic types (White, 1982, 1985:64–68) and the tradition of this sub-stage of the Magdalenian functioning largely as a residual category when assemblages could not be attributed to either the earlier or later stages of the Magdalenian based on bone tool typologies (Koetje, 1987:6; White, 1987:226). This could potentially have artificially depressed the demographic signature produced by the site counts analysis compared to the $^{14}$C date distribution.

5.2. Mobility-based explanations for the patterns seen

Finally, we should mention the possibility that changes in hunter—gatherer behaviour could impact on the quantities of either of the two proxies or provide an alternative, non-demographic explanation for the frequency distributions described above. For hunter—gatherer groups, the most frequently cited alternative explanation for chronological variation in site and radiocarbon date frequency distributions is a change in mobility strategy and land-use patterns, with an increased signature being interpreted as representing periods of higher mobility (e.g. Attenbrow, 2004; Naudinot et al., in press; Niekus, 2005/2006; Tallavaara et al., 2010:253). This could refer either to differences in overall group mobility strategy which impact how often groups move home bases and the number of non-residential sites they generate (e.g. Binford’s (1980) continuum of ‘logistic’ and ‘residential’ mobility), or seasonal/annual population aggregations and dispersals (‘fusion—fission’; Aureli et al., 2008; Layton and O’Hara, 2010).

A full study of potential differences in mobility strategy across the Upper Palaeolithic in Southwestern France was beyond the scope of this research, and, although they exist for earlier periods (e.g. Delanges and Rendu, 2011), there are no previous diachronic studies of mobility for the Upper Palaeolithic in the region. The study of some of the characteristics of the sites included in the analysis (and from which the dates derive) can provide insights into whether changing mobility is greatly impacting either distribution (French, 2013, submitted for publication). As shown in Fig. 6, there are clear periods of divergence in the relative frequency of open-air and sheltered sites across the Upper Palaeolithic in the region, even when the number of open-air sites has been ‘corrected’ using the curve of Surovell et al., 2009. This suggests that differences in hunter—gatherer settlement patterns with regard to the relative use (and subsequent preservation and archaeological discovery) of sheltered to open-air sites, could account for the differences in the number of sites documented at various parts of the sequence.

Data on other characteristics of the sites included can be used to look at key stages of the Upper Palaeolithic sequence. For example, the extreme trough seen in both the $^{14}$C date distribution and the number of archaeological sites during the Late Gravettian (~28 000 cal BP) could be attributed not to population decrease, but decreased population mobility relative to the Middle Gravettian. While a crude measure, as documented in French (2013) (as a percentage of the total number of sites with available data for each period) there are more Late Gravettian than Middle Gravettian sites with large lithic assemblages (defined as ≥1000 retouched tools) (38% and 31% respectively), potentially indicating more intensive or longer-term occupation of sites in the Late Gravettian compared to the Middle Gravettian. However, this small difference is unlikely to account for the extreme trough seen in both of our proxies during the Late Gravettian.

While the study of some of the variables discussed above can provide us with insights into alternative explanations for variation in archaeological proxies for demography, we stress that these explanations are not mutually exclusive, and that changes in
Using $^{14}$C date frequency distributions and archaeological site counts as proxies for changes in relative population size, we have documented fluctuations in hunter–gatherer population levels across the Upper Palaeolithic in Southwestern France. The patterns produced are similar between the two proxies, with specific points of interest across the sequence including a marked population decline in the Late Gravettian (~28 000 cal BP) and population peaks at the Middle/Late Gravettian interface (~29 500 cal BP), Late Solutrean (~25 000 cal BP) and Upper Magdalenian (18 000 cal BP). The Late Solutrean peak supports the long-held notion of this region as a population refugium during the Last Glacial Maximum. The negative correlation documented between temperature (as seen in the NGRIP curve) and the $^{14}$C date distribution raises the question as to whether Southwestern France acted as a refugium during other cold stages of the Upper Palaeolithic.

Overall, the data suggest that taphonomic bias and excavation and dating strategies have had little effect on the distributions produced. Similarly, the calibration curve used to create the radiocarbon date distribution had no significant impact on the overall shape of the distribution, although a peak present ~32 000 cal BP in the curve produced by IntCal09 is noticeably absent in that produced by IntCal13. The similarities between the distributions generated by the two approaches strengthen our interpretation of relative demographic change as the cause of the observed pattern. While some points of noticeable divergence are present we suggest that these cannot be explained with recourse to the general differences between the two approaches, and are probably related to the specifics of the archaeological record at the points in question. However, two major points of uncertainty remain. Firstly, within the limits of this study, it is difficult to assess the extent to which changes in hunter–gatherer mobility across the Upper Palaeolithic in Southwestern France affected (either or both) of the proxies studied, how these varied across the period, and as a result, affected the demographic signatures presented here. Secondly, it is unclear whether the broadly comparative results generated by the $^{14}$C date distributions and archaeological site counts are specific to our case study. While previous comparisons are rare, similar correspondences reported for datasets as diverse as Holocene Australia (Williams, 2012) and Mesolithic/Neolithic Fennoscandia (Tallavaara et al., 2010) suggest that the two proxies may be robust equivalents, and that the absence of extensive $^{14}$C dating or widespread archaeological survey in a region need not inhibit at least preliminary investigation of prehistoric population histories.

Acknowledgements

This research was funded by an AHRC Doctoral Studentship and a Research Fellowship at Peterhouse, University of Cambridge (JCF) and a University of Sheffield Doctoral Studentship (CC). Thanks to Andrew Chamberlain, Robert Foley, Matt Grove, Robert Kelly, Paul Mellars, Stephen Shennan, James Steele, and Todd Surovell for useful discussions on the ideas covered in this paper and/or comments on earlier drafts. We also thank three anonymous reviewers whose comments greatly improved the manuscript. All errors and omissions are our own.
Ducasse, S., 2012. What is left of the Badegoulian
Holdaway, S., Fanning, P., Shiner, J., 2005. Absence of evidence or evidence of
Gamble, C., Davies, W., Pettitt, P., Richards, M., 2004. Climate change and evolving
French, J.C (n.d. submitted for publication). The Demography of the Upper Palae-
Demars, P.-Y., 2002. Changements climatiques et occupation de l’espace. Les der-
Genty, D., Blamart, D., Ouahdi, R., Gilmour, M., Baker, A., Jouzel, J., Van-Exter, S.E.,
Layton, R., O’Hara, S., 2010. Human social evolution: a comparison of hunter-


