Research

The chronology of culture: a comparative assessment of European Neolithic dating approaches

Katie Manning[1,*], Adrian Timpson[1,2], Sue Colledge[1], Enrico Crema[1], Kevan Edinborough[1], Tim Kerig[1] & Stephen Shennan[1]

Archaeologists have long sought appropriate ways to describe the duration and floruit of archaeological cultures in statistical terms. Thus far, chronological reasoning has been largely reliant on typological sequences. Using summed probability distributions, the authors here compare radiocarbon dates for a series of European Neolithic cultures with their generally accepted 'standard' date ranges and with the greater precision afforded by dendrochronology, where that is available. The resulting analysis gives a new and more accurate description of the duration and intensity of European Neolithic cultures.

Keywords: Neolithic, culture, chronologies, dendrochronology, radiocarbon dating, Bayesian analysis

Supplementary material is provided online at http://antiquity.ac.uk/projgall/manning342/

Introduction

The construction of cultural chronologies in archaeology remains a point of major ongoing debate. Although the increasing application of Bayesian analysis has dramatically improved the temporal control that can be achieved in archaeology by making it possible to integrate stratigraphic and other information to narrow down calibrated radiocarbon date distributions (Bayliss & Whittle 2007; Bronk Ramsey 2009; Whittle *et al.* 2011), such analyses are complex and time-consuming. At present, stratigraphically constrained Bayesian analyses of dates are not available in sufficient quantity to provide a basis for characterising regional and inter-regional patterns in the cultures of Neolithic Europe, and it is open to question whether this approach is applicable to such broad scales of archaeological enquiry. In any case, chronological reasoning, particularly in Neolithic Europe, is still largely reliant on typological sequences. Although attempts at providing absolute sequences for cultural chronologies are becoming increasingly common (e.g. Stadler *et al.* 2001; Raetzel-Fabian 2002; Furholt 2003; Reingruber & Thissen 2009; Denaire 2011), there have been few efforts to systematically compare the beginning and end dates assigned to cultures, and the date ranges indicated by all of the available radiocarbon evidence. Some obvious exceptions include, for example, the review by Włodarczak (2009) of the Corded Ware dating

[1] *Institute of Archaeology, University College London, 31–34 Gordon Square, London WC1H 0PY, UK*
[2] *Research Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK*
[*] *Author for correspondence (Email: k.manning@ucl.ac.uk)*

methods, and the use of multivariate statistics in conjunction with radiocarbon data as a means of refining typochronological classifications. In particular, correspondence analysis and principal component analysis have been used to identify chronologically meaningful aspects of regional typologies, e.g. in the Mittelelbe-Saale region in Germany (Müller *et al.* 2000; Czebreszuk & Müller 2001; Müller & Van Willigen 2001, Müller 2009), for Linearbandkeramik (LBK) settlements in the Netherlands (van de Velde 2012) and also for the Corded Ware (Ullrich 2008). All of these, however, are relatively localised studies of internal development.

In this paper, we approach the issue at a different scale, adopting a broader and therefore lower resolution view of the relationship between different dating approaches for the European Neolithic; we then go on to estimate the underlying shape of the intensity of a culture's presence through time. We investigate the temporal ranges of various cultural groups by analysing and comparing the estimates from three sources: the 'standard' date range, the radiocarbon date range, and the dendrochronological (dendro) date range, where available. We use the term 'standard' date range to mean the estimated start and end date for a culture as found in relevant current literature, which we consider to represent the best-informed archaeological knowledge on the subject (see online supplementary Table S1 for all 'standard' dates and the references used). The motivation behind our analysis came from attempting to investigate temporal patterns in subsistence data. Given that awareness of available radiocarbon dates forms an important part of that up-to-date archaeological knowledge, we were surprised to find a large difference in our results when using site phases that were dated using associated radiocarbon samples, compared to the current estimate of the 'standard' date range for that cultural phase. Therefore, in order to undertake a valid analysis of subsistence trends, or of any other archaeological phenomena for that matter, we must first investigate the congruency between the 'standard' date range and the contextually associated radiocarbon dates for a given culture.

It is worth pointing out at this stage that our interest in the temporal ranges of different cultural groups is not intended to promote the notion of archaeological culture as a monolithic and unchanging entity: cultures as 'bricks'. On the contrary, one of us has consistently argued the opposite (e.g. Shennan 1978, 1989), following Clarke's (1968) polythetic definition of archaeological cultures. We do not need to take the traditional view of cultures as monolithic blocks to explore their chronological patterns, which may in fact relate to specific cultural packages. Just as in genetics a contrast is made between the analysis of gene trees and population trees, both of them equally valid for different purposes, so it is equally appropriate in the case of transmitted cultural variation to work at different scales of resolution for different purposes. Moreover, we do not need to subscribe to the idea that our classifications 'carve nature at its joints' in Plato's famous phrase (*Phaedrus*, 265d–266a; Fowler 1925) to recognise that some categorisations are better than others. One key criterion is whether variation between the entities being analysed is significantly greater than that within them (see Shennan *et al.* 2014), regardless of the processes that resulted in this patterning, which often remain the focus of dispute (see Furholt 2014 for an example). In short, while archaeological 'cultures' may not be useful for investigating local-scale socio-economic dynamics, they have remained indispensable for characterising broad-scale spatial and temporal trends, such as those which prompted this analysis.

Method and dataset

All radiocarbon data were derived from the database collated by the EUROEVOL project. These data include a lab code, ^{14}C age and standard deviation (SD), latitude and longitude. Each ^{14}C sample was calibrated via the IntCal09 calibration curve (Heaton *et al.* 2009). In order to warrant inclusion in this analysis, the sample also had to have an associated culture, assigned either by the excavator, laboratory or post-excavation analyst. The resultant ^{14}C dataset for this study therefore consists of 5594 radiocarbon-dated samples from 1784 archaeological site-phases (see Shennan *et al.* 2013 for the applied definition of phase), from 71 Neolithic and Early Bronze Age cultures (see Figure 1 and online supplementary Table S1). In addition to this ^{14}C dataset, a second dendrochronological dataset comprising 350 samples were obtained from the Erziehungsdirektion des Kantons Bern website (Dendro n.d.).

Each ^{14}C sample can therefore be considered to have two properties: a radiocarbon calibrated date, and a 'standard' date, each with its own probability distribution. Clearly there is a degree of circularity between these two properties, since the cultural assignment of a sample may be influenced by its radiocarbon date, especially in the absence of associated diagnostic material; more importantly, the current 'standard' date ranges will have been strongly influenced by the radiocarbon dates themselves. Hence our surprise at the difference in results when using a site's cultural assignment as the basis for dating it as opposed to the site's own radiocarbon dates.

Similarly, each dendro sample can be considered to have two properties: a dendro date and the 'standard' date of the culture to which it was assigned. Therefore, we can think of these relationships from the perspective of estimating the date range of a culture, by comparing its ^{14}C date range, its 'standard' date range, and (where available for a handful of cultures) its dendro date range. Our objective in what follows is thus to establish how much agreement there is between these date range estimates, and discuss the implications of those findings, not least for the analysis of other archaeological patterns. Breunig (1987) had a similar programme, although at a larger spatial scale, and of course at that time far less radiocarbon information was available.

Our study area encompasses central and north-western Europe (Figure 1) and covers the Late Mesolithic through to the Early Bronze Age. The term 'culture' refers here to the archaeological cultures as single entities, so, for example, the Cortaillod is treated as a whole rather than being divided into the sub-phases of the Cortaillod i.e. ancien, classique and tardif. Three of our cultures—Bell Beaker, Corded Ware and Linearbandkeramik—are multi-regional, 'culture groups' in Clarke's (1968) terms.

The dataset included 5452 samples from 1748 site-phases, which had been assigned to only one of these 71 cultures, while 142 samples (2.5%) from 36 phases had been assigned to two (or in one case three) of these cultures (see online Table S1 for a list of all cultures). This second type of cultural assignment may have three interpretations: firstly that the deposits indicate overlap between two different cultures; secondly, that they belong to a transitional phase between two cultures; and thirdly that they have an equal probability of belonging to either cultural range. It is likely that all three interpretations were used on different occasions; nevertheless, we have adopted the third interpretation for all cases. This provides

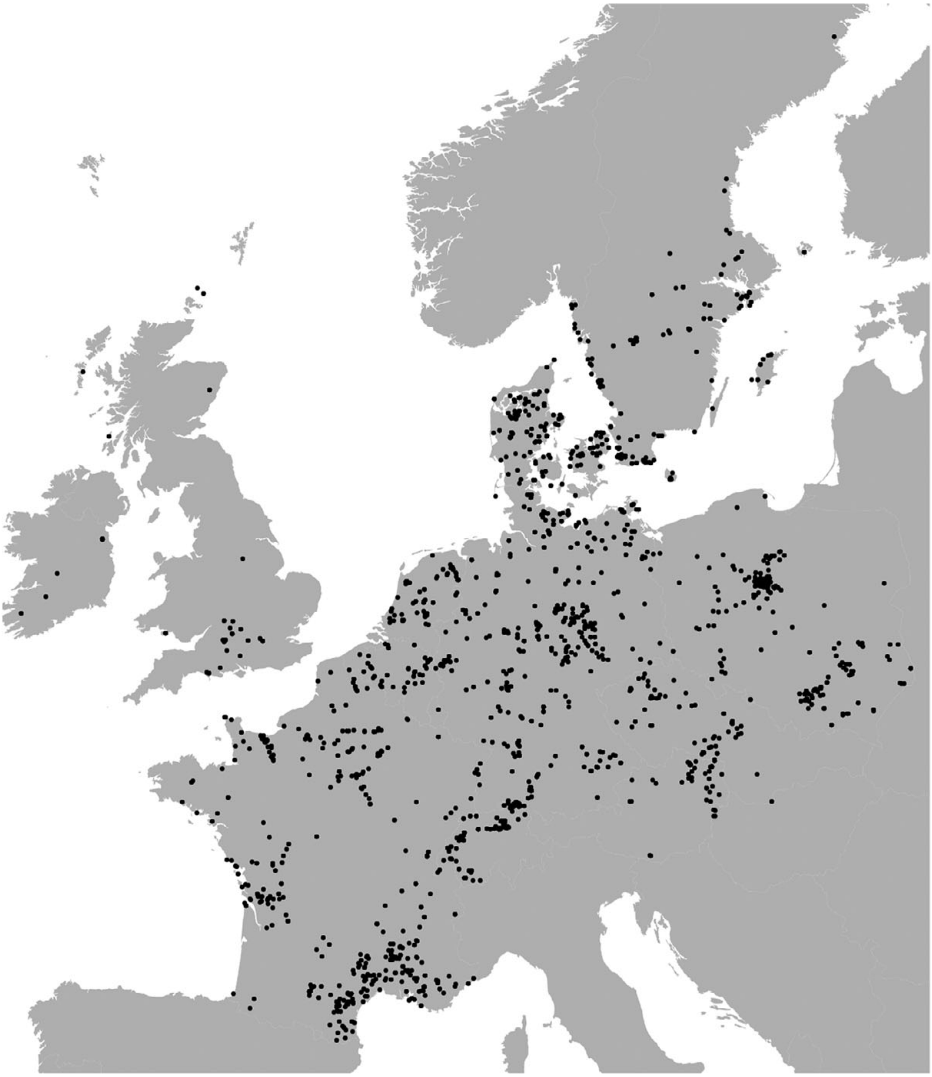

Figure 1. Map of north-western Europe showing the site locations from which the ^{14}C samples were taken (black dots).

a conservative approach with the effect of erring on the side of slightly greater uncertainty in the date range.

First, we investigate the broad-scale relationship between the radiocarbon date and the 'standard' cultural date for all samples combined (Figure 2). Both before and after calibration the exact radiocarbon date of a sample remains elusive. Instead we have a distribution, which describes the probability of each possible date. Similarly, the true 'cultural' date could be any date within a culture's range, which is assumed to be a uniform distribution. We generate 1000 random realisations, such that for each realisation every sample has a random point estimate: both a discrete radiocarbon date and a discrete 'standard' date obtained from their respective probability distributions. For each realisation the correlation between the two estimates (using Pearson's R) and the gradient of a best-fitted linear model were calculated. The linear model used Deming regression since both variables have an error distribution. By repeating the random sampling 1000 times we build up a distribution of possible values, from which confidence intervals (CIs) can be estimated. These confidence intervals describe the amount of variance caused by the uncertainty in the probability distributions.

Having established the overall correlation between our date range estimates, we turned our attention to assessing the relationship between the different estimates for each culture separately. We selected individual cultures that met a minimum sampling criterion to ensure good geographic and temporal representation, such that a minimum of 18 phases per culture were required, leaving us with the 22 (31%) best-represented cultures (Table 1), each of which comprised between 18 and 329 site-phases (mean = 64.4), and between 34 and 1030 samples (mean = 195), and constituted the majority of the data, comprising 4281 ^{14}C dates (79%), and 1416 phases (81%).

Four of our selected cultures also have associated dendro dates (Corded Ware n = 73, Cortaillod n = 66, Horgen n = 66 and Pfyn n = 55), which generally provide greater precision for an individual sample than a typical radiocarbon date (mean dendro error = 6.9 years per sample, compared to mean ^{14}C error of 74.7 ^{14}C years). Dendrochronology can therefore shed further light on the date range of each culture, so long as comparisons are made within the relevant region (see below in regard to the Corded Ware Culture in Switzerland).

For each of these latter cultures, the dendro date range is reported in Table 1 using the earliest and latest date, and are plotted in Figure 3 with each sample as a vertical red bar with a width equal to the sample's error. Similarly, the 'standard' dates are reported as a range between earliest and latest, and plotted as a uniform distribution, represented by a black bar in Figures 3 and 4. The radiocarbon date ranges are reported in Table 1 as the two-tailed 95% interval of a summed probability distribution (SPD), while the full probability distribution (between 8000–0 BC) is plotted in Figures 3 and 4. Each SPD was constructed in three stages: first, raw radiocarbon dates (mean and SD) were calibrated (see Shennan *et al.* 2013 for details of calibration); second, the calibrated probability distributions from every sample in the same phase were summed and normalised for unity; and third, these distributions from every phase in the same culture were summed and normalised for unity. The rationale behind this approach is that if random sampling is assumed, the law of large numbers predicts that as the number of samples increases, the sample distribution becomes increasingly similar to the true distribution. By summing the individual samples' probability distributions, the new probability distribution contains the combined knowledge of each sample, equally weighted, as each sample was initially considered an equally fair possibility of the date of the event. This applies at the level of both the phase and the culture.

In order to assess the level of agreement between the radiocarbon dates and the 'standard' dates, we calculated the proportion of the SPD mass that fell outside the 'standard' cultural date range. This value is reported in Table 1 as the proportion that fell prior to the start of the 'standard' date range (older), the proportion that fell subsequent to the 'standard' end date (younger) and the total proportion.

Finally, we assessed how the intensity or 'floruit' of an individual culture varied through time, and the underlying shape of this change. We Z-transformed the SPD for each of the selected 22 cultures, to ensure they were on the same scale (Figure 5). This involved obtaining the mean date for the culture's summed probabilities and the standard deviation of this distribution and then expressing the date values not on the original year scale but as numbers of standard deviations away from the mean, thus eliminating the effect created by the fact that some cultures lasted longer than others. The 22 Z-transformed SPDs were weighted by number of phases per culture, then summed and normalised to unity (Figure 5, blue line), since we can expect the 'shape' of those cultures with more phases to be more representative of the true culture shape. Further, for comparison we summed and normalised the 22 Z-transformed SPDs with an equal weighting (Figure 5, red line), to provide some indication of the efficacy of our minimum inclusion criterion of 18 phases per culture.

Results

Overall correlation

Figure 2 illustrates 10 realisations (out of the total 1000 realisations that were computed) for each of our 5594 radiocarbon samples plotted in grey. Note the 'blocky' nature of the scatterplot in the horizontal range, due to the 'standard' date being sampled from each culture's wide uniform probability distribution. The vertical range, meanwhile, is scattered more evenly since each radiocarbon date was sampled from its own relatively narrow probability distribution. Clearly, when all samples and all cultures are included, there is a strong correlation between the two dating methods. (R = 0.859, 95% CI = 0.855–0.862), with almost three quarters of the variance in either estimate explained by the other (R squared = 0.737, 95% CI = 0.730–0.744). The gradient of the best-fitted linear model plotted in blue is 1.038 (95% CI = 1.030–1.045), which is to be expected as both the radiocarbon date range and the 'standard' date range are two different estimates of the same quantity (the true date range of the culture).

Comparing individual culture date ranges

Having established an overall strong relationship between the probability distributions of all radiocarbon dates and 'standard' culture dates, the question remains as to whether certain cultures correlate better with their associated radiocarbon dates than others. Figure 3 plots the 'standard', radiocarbon and dendro date ranges for the four cultures with associated dendro samples, and Figure 4 plots the 'standard' and radiocarbon date ranges for the remaining 18 selected cultures. Generally, there is good agreement between the dendro date ranges and the 'standard' date ranges, which is perhaps not surprising due to the high level of precision on dendro dates and the excellent preservation of the wetland sites where these cultures are predominantly found. The exception to this is the Corded Ware, which has a much wider 'standard' range than the dendro range, due to the fact that the Corded Ware is found over a much wider geographic area than the typical wetlands required for dendrochronology, and lake shore settlement ceased *c.* 2450 cal BC (Billamboz & Köninger 2008).

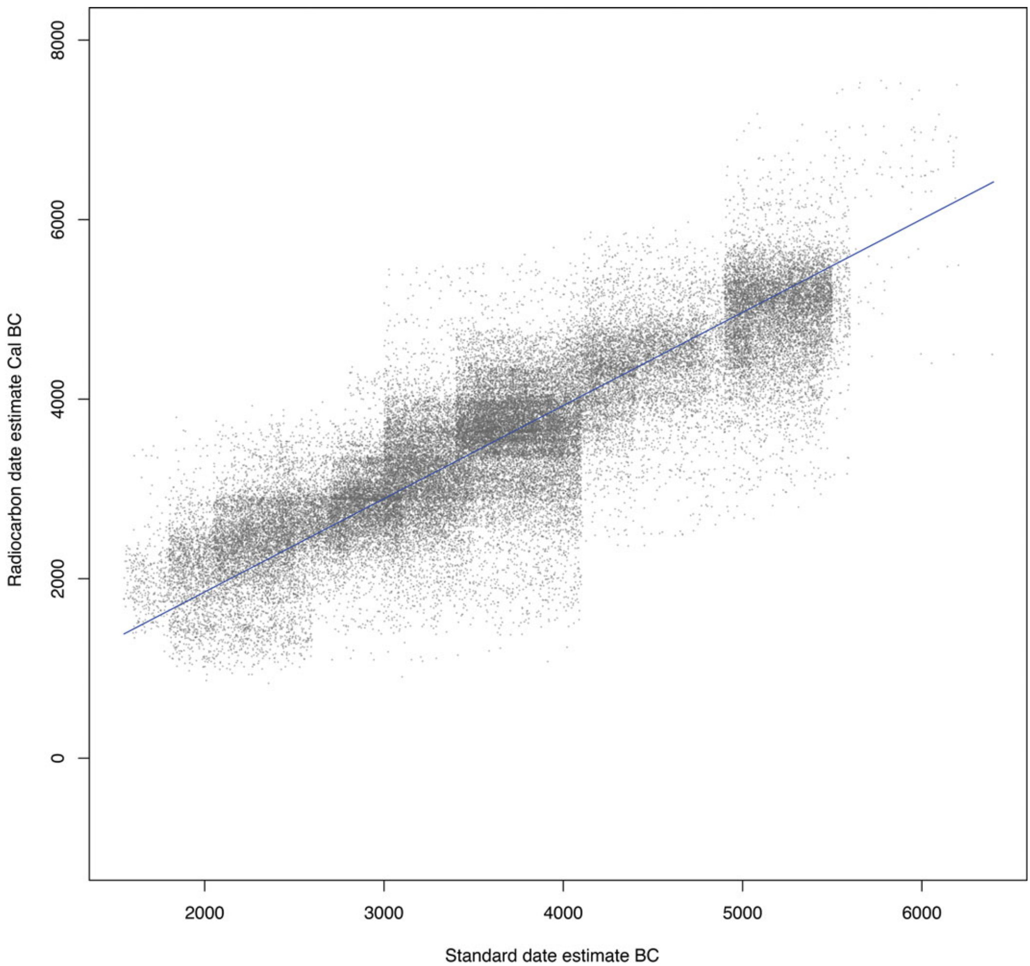

Figure 2. Ten random realisations of each of the 5594 samples' radiocarbon dates plotted against their respective 'standard' date (grey dots), with the line of best fit for all 1000 realisations (blue).

In contrast, the radiocarbon distributions are substantially wider than the 'standard' date ranges in all cases. Table 1 reports the proportion of the radiocarbon probability mass that exceeds the 'standard' range, revealing a massive spread of values from 14.8% (Pitted Ware) to 75.1% (Veraza) with a mean of 38.7%. Such a broad range suggests that some cultures may have more reliable chronological sequences than others. On a more general note, however, it is clear that the level of agreement between radiocarbon dates and the conventional date ranges for European Neolithic cultures remains quite poor. Overall the radiocarbon ranges tend to be skewed towards being younger than the 'standard' date ranges; on average 23% of the radiocarbon probability mass is younger than the 'standard' date range and 15% is older (22% and 12% respectively when weighted by sample size). This is rather surprising in light of our inclusion of charcoal dates that might lead to 'old wood' effects.

Other differences in the nature of this discrepancy for each culture are evident. Some, for example Veraza and Iwienska, indicate an overall shift in the peak of radiocarbon probability

Table 1. The 22 best-represented cultures with their start and end dates estimated using the three different methods, and the mean and SD using the improved normally distributed model. The last three columns estimate the difference between 'standard' and radiocarbon date range.

			Radiocarbon (BC)		Standard (BC)		Dendrochronology (BC)		Proportion of radiocarbon probability mass outside the standard range		
Culture	Phases	Samples	Start	End	Start	End	Start	End	Older	Younger	Total
Bell Beaker	113	228	3040	1607	2500	1800			0.264	0.046	0.310
Western Cardial	57	151	6574	3849	5600	4950			0.157	0.349	0.506
Chasseen	75	213	4663	2952	4400	3400			0.134	0.115	0.249
Corded Ware	189	341	3251	1901	2800	2050	2751	2418	0.171	0.062	0.233
Cortaillod	33	111	4503	2963	4100	3400	3895	3517	0.201	0.147	0.348
Ertebølle	46	360	5709	3478	5450	4100			0.077	0.206	0.283
Globular Amphora	60	145	3411	2009	3100	2700			0.138	0.363	0.501
Horgen	26	130	3540	2525	3500	2750	3482	2765	0.038	0.131	0.169
Iwienska	23	131	2195	1082	2600	1800			0.000	0.613	0.613
Lengyel	39	147	5080	3540	4800	4000			0.115	0.183	0.298
Linearbandkeramik	139	535	5719	4127	5500	4900			0.072	0.313	0.385
Michelsberg	45	126	4529	3158	4250	3500			0.150	0.106	0.256
Néolithique final	42	98	3473	2040	2900	2150			0.228	0.042	0.270
Néolithique moyen II	24	96	4534	2973	3950	3500			0.556	0.083	0.639
Peu Richard	18	56	3488	2483	3300	2900			0.173	0.268	0.441
Pfyn	19	92	4251	3138	3850	3400	3919	3507	0.359	0.151	0.510
Pitted Ware	51	130	3637	2257	3400	2350			0.100	0.048	0.148
Rossen	20	50	5018	4045	4750	4400			0.169	0.178	0.347
Seine-Oise-Marne	26	39	3730	1777	3500	2700			0.130	0.495	0.625
Trichterbecher	329	1030	4323	1793	4100	3000			0.049	0.261	0.310
Veraza	20	38	3622	1989	3400	2850			0.069	0.682	0.751
Wartberg	22	34	3488	2205	3500	2800			0.014	0.302	0.316

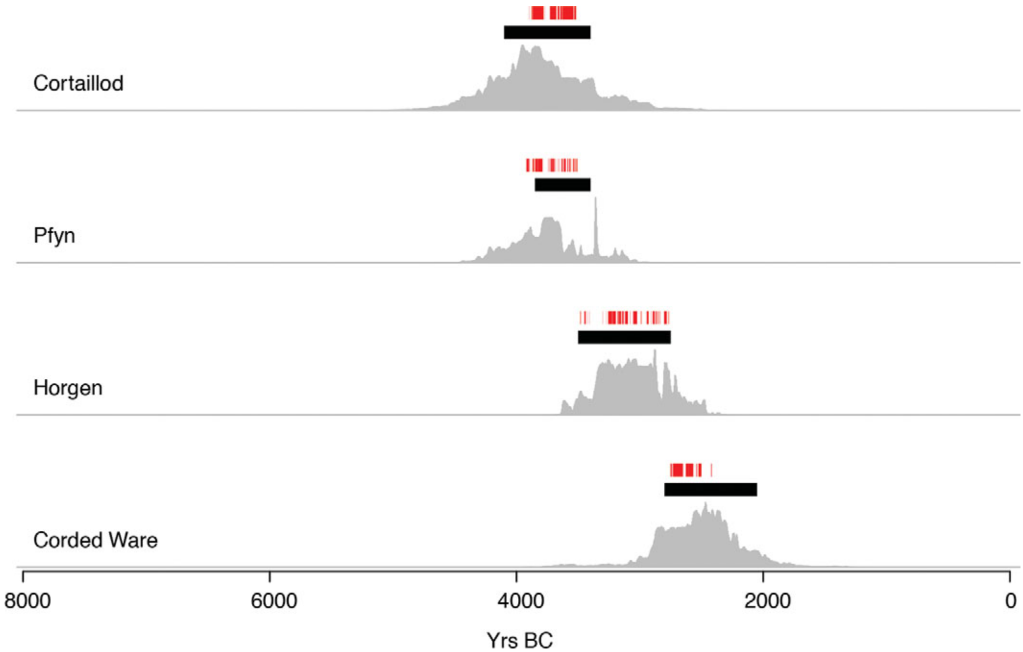

Figure 3. Comparative date ranges for the four cultures with associated dendro samples arranged in chronological sequence. The dendro date range is shown in red, radiocarbon summed probability distribution (SPD) in grey and the 'standard' date range in black.

density, suggesting the 'standard' range may be slightly too early or too late. Others, for example Seine-Oise-Marne, have a generally broader radiocarbon distribution, indicating a potential underestimation of the 'standard' date range. The specifics of some of these cultural groups are discussed below in relation to the potential underlying causes of chronological discrepancies.

The 'temporal shape' of Neolithic cultures

In addition to comparing different dating methods, the technique of summing probability distributions also allowed us to investigate the broad-scale 'temporal shape' of Neolithic cultures. Care should be taken when interpreting the distribution of individual cultures across relatively narrow ranges since the radiocarbon calibration process often generates spurious wiggles on the short-term scale (below around 200 years). By Z-transforming each SPD we can compare the shape of each culture's radiocarbon distribution on a like-for-like scale. This transformation also had the benefit of causing these wiggles to shift and change in size, removing their constructive interference and instead reducing them to random background noise in the overall shape.

Figure 5 shows SPDs after Z-score transformation for each of our 22 cultures. The red line shows the mean of all cultures, while the blue line shows the mean weighted by the number of phases per culture. Both lines show remarkable similarity, suggesting that our inclusion criterion of a minimum 18 phases was successful in ensuring each SPD was well

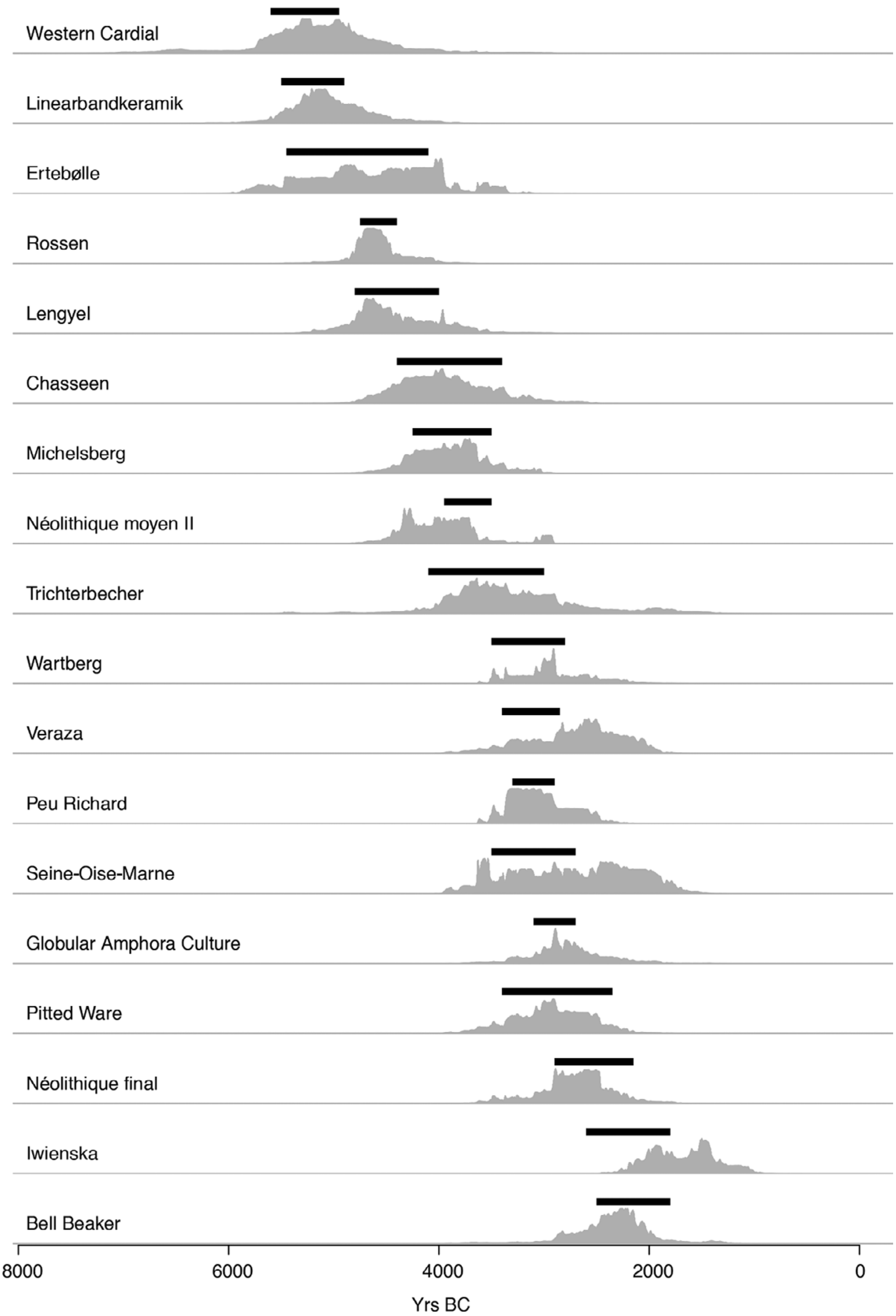

Figure 4. Comparative date ranges for the remaining 18 best-represented cultures, arranged in chronological sequence according to the mean of the 'standard' date range.

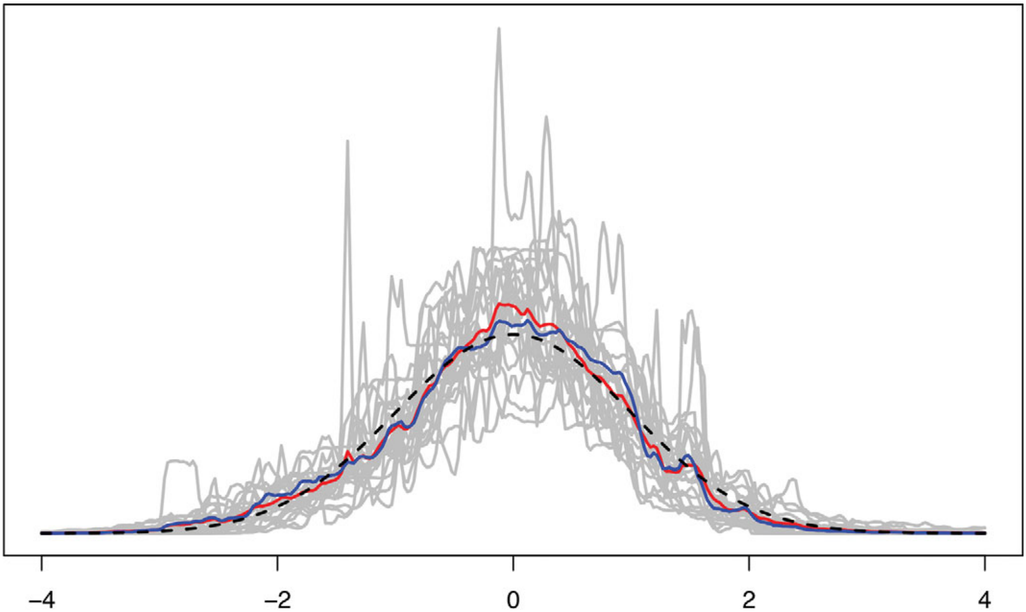

Figure 5. The 22 grey lines represent each culture's SPD (Z-transformed); the red line shows the mean of all 22 cultures' SPD; the blue line shows the mean weighted by number of phases per culture; and the dashed black line shows a standard normal distribution.

representative of the culture's true underlying distribution. The dashed black line shows a standard normal distribution (mean = 0, SD = 1), which describes the shape of any normal distribution after Z-score transformation. The fit is extremely compelling.

The principal advantage of this normal distribution is its increased explanatory power at no extra cost of parameters. Currently, the 'standard' date range uses two parameters: a start and end date, while a normal distribution also only requires two parameters: the mean and SD. For example, currently the 'standard' date range for the Linearbandkeramik is described using start = 5500 BC and end = 4900 BC, which tells us little about the actual distribution of the available evidence. In contrast, a normal distribution more accurately describes its shape, including its central tendency as well as duration, using mean = 5088 BC and SD = 310 yrs. Hence, the additional information that is gained, at no extra cost, allows us to begin exploring the statistical properties of cultural chronologies in a way that is not possible with the current start and end date ranges. Figure 6 shows examples of the fitted normal distribution with its mean and SD for five selected cultures. The mean and SD are reported for all 22 cultures in Table 2, and are plotted in Figure S1 in the supplementary information.

Discussion

It may be fair to assume that the additional comparative dates provided by dendrochronology would improve the overall reliability of a culture's chronology. Although this does appear to be the case for the Horgen and Corded Ware (only 17% and 23% of ^{14}C probability mass

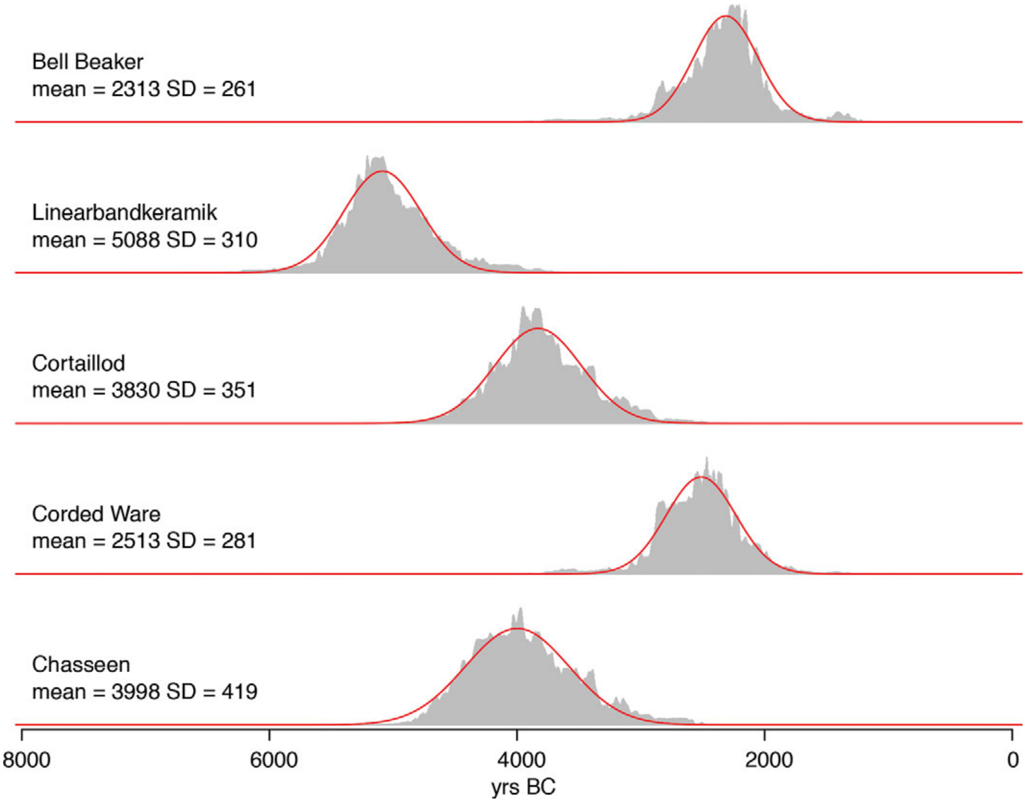

Figure 6. Examples of summed probability distributions for 5 of the 22 cultures in grey. The red line shows the best-fitted normal distribution.

outside their respective 'standard' dates), it does not apply to all cultures; for example, there is a discordance of 51% in the Pfyn data. This is in contrast to cultures such as the Pitted Ware, Chasseen and Michelsberg, which have no associated dendro dates, and yet show relatively good agreement between the radiocarbon and 'standard' date ranges (15%, 25% and 26% discordance respectively). This may be due to recent efforts aimed specifically at refining cultural chronologies such as the Michelsberg and its successive cultures (e.g. Raetzel-Fabian 2000, 2002; Geschwinde & Raetzel-Fabian 2009), or because of the archaeological visibility of particular cultural traits, e.g. the enclosures of the Michelsberg or the Chasseen, and the distinctive material and economic characteristics of the Pitted Ware. In contrast, the cultures with the highest disagreement values, such as the Veraza, Néolithique moyen II and Seine-Oise-Marne, are comparatively ill-defined and with relatively small sample sizes (38, 96 and 39 respectively).

These results raise a number of important issues relating not only to chronological reasoning, but also to the characterisation of Neolithic cultures. On a very basic level, the amount of agreement between different dating methods at the broad level of analysis we have used seems to come down to economy and scale. The more established a culture

Table 2. The proposed mean and SD based on a normal distribution for all 22 cultures.

	Normal distribution (BC)	
Culture	Mean	SD
Bell Beaker	2313	261
Western Cardial	5126	448
Chasseen	3998	419
Corded Ware	2513	281
Cortaillod	3830	351
Ertebølle	4583	624
Globular Amphora	2806	282
Horgen	3066	276
Iwienska	1681	340
Lengyel	4499	362
Linearbandkeramik	5088	310
Michelsberg	3911	328
Néolithique final	2675	280
Néolithique moyen II	4031	317
Peu Richard	3097	268
Pfyn	3779	265
Pitted Ware	2911	362
Rossen	4620	156
Seine-Oise-Marne	2702	669
Trichterbecher	3432	481
Veraza	2625	426
Wartberg	2968	322

is, and to some degree the more widespread it is, the more agreement there is between estimates, implying greater accuracy. However, this is not simply attributable to sample size, and there is no clear correlation between sample size and the agreement between the two estimates. Trichterbecher has by far the largest ^{14}C sample size (1030), yet suffers 31% discordance between estimates, whilst Pitted Ware has fewer samples by an order of magnitude (130), yet has a smaller discordance of 15%. The Linearbandkeramik enjoys the second highest sample size of 535 but fares even worse, with a discordance between estimates of 39%. Despite the relatively broad geographic spread of the Cardial culture there is high discordance between estimates (51%). Hence there is no clear-cut cause for discrepancy between the different date estimates, and instead each culture needs to be considered individually in light of its own historical trajectory, environmental setting and history of research.

A more interesting result to have arisen from this analysis is the comparative shape of the SPD of each culture's ^{14}C date range estimate. The start and end dates of the 'standard' range provide us only with a simple uniform distribution, and while this may tell us something about the duration of a cultural phenomenon, it tells us nothing about its temporal nature, i.e. its floruit or intensity (Ottaway 1973; Aitchison *et al.* 1991).

It is improbable that the material associated with a given culture suddenly comes into existence maintaining a constant prevalence and then suddenly ceases. Our analysis from combining all 22 SPDs suggests that a normal distribution provides a better description of the fundamental underlying shape of a culture's rise and fall, and, therefore, the use of a mean and standard deviation would be more informative than the current start and end dates, which are so often used in the archaeological literature. Indeed, in some circumstances it may also be useful to supplement reporting the mean and SD of a cultural range by reporting the full ^{14}C probability distribution, in order to incorporate fine-grained information about the temporal variation in the intensity of the episode. This requires caution as finer resolution may introduce undesirable artefacts from sampling error and 'interference' effects from wiggles in the calibration curve (Blockley *et al.* 2000). As such, this would only be useful where large unbiased samples are available and when the culture spans a long time period. Nevertheless, this study provides empirical support that a Gaussian model, and perhaps other unimodal distributions (Lee & Bronk Ramsey 2012), are superior to the standard uniform distribution, both as a prior for Bayesian models of archaeological chronologies (Bronk Ramsey 2009) and as a simple summary description of the cultural time-span.

This improved underlying model of the expected temporal dynamics of a regional culture is of value for a range of different analyses. Bayesian analysis of ^{14}C dates aggregated at the level of culture is reliant on the selection of a prior distribution. In the absence of other information a uniform prior is usually adopted; instead, a normally distributed prior is both more reasonable and empirically supported. Analysis requiring Monte Carlo simulation of cultural dates (e.g. Crema 2013) would also benefit greatly from this improved model.

Finally, it is worth pointing out that the characteristic normal distribution we have identified bears a striking resemblance to the so-called 'battleship curves' produced when frequency seriations are carried out on individual artefact types that are chronologically sensitive. When cultures are taken as entities they seem to mirror this effect. In essence, the number of dated events that archaeologists are prepared to label, for example as Horgen or Michelsberg, starts small, increases to a peak and then declines again. The pattern could arise because of the waxing and waning popularity of temporally correlated styles across a geographical region. Another possibility, supported by demographic proxies in some cases (Shennan *et al.* 2013), is that they reflect fluctuations in local populations; at some periods there are simply more people in the region, so the number of dated events characterised by the styles of the period is also bound to be greater. Of course, these two possibilities are not mutually exclusive, and in some cases it could be that new cultural innovations themselves result in periods of population increase. These questions remain open for subsequent analysis.

Acknowledgements

We are grateful to all those who provided radiocarbon data, and to Johannes Müller and Felix Riede for their constructive review comments. This research was funded by the European Research Council Advanced Grant number 249390 to Stephen Shennan for the EUROEVOL Project. Information on the EUROEVOL Project is available at http://www.ucl.ac.uk/euroevol/

References

AITCHISON, T., B. OTTAWAY & A.S. AL-RUZAIZA. 1991. Summarizing a group of ^{14}C dates on the historical time scale: with a worked example from the Late Neolithic of Bavaria. *Antiquity* 65: 108–16.

BAYLISS, A.L. & A.W.R. WHITTLE (ed.). 2007. *Histories of the dead: building chronologies for five southern British long barrows* (*Cambridge Archaeological Journal* 17, Supplement 1). Cambridge: Cambridge University Press.

BILLAMBOZ, A. & J. KÖNINGER. 2008. Dendroarchäologische Untersuchungen zur Besiedlungs- und Landschaftsentwicklung im Neolithikum des westlichen Bodenseegebietes, in W. Dörfler, J. Müller & A. Wesse (ed.) *Umwelt—Wirtschaft—Siedlungen im dritten vorchristlichen Jahrtausend Mitteleuropas und Südskandinaviens*: 317–34. Neumünster: Wachholtz.

BLOCKLEY, S.P.E., R.E. DONAHUE & A.M. POLLARD. 2000. Radiocarbon calibration and Late Glacial occupation in north-west Europe. *Antiquity* 74: 112–21.

BREUNIG, P. 1987. *^{14}C-Chronologie des Vorderasiatischen, Suedost- und Mitteleuropaeischen Neolithikums*. Köln: Boehlau.

BRONK RAMSEY, C. 2009. Bayesian analysis of radiocarbon dates. *Radiocarbon* 51: 337–60.

CLARKE, D.L. 1968. *Analytical archaeology*. London: Methuen.

CREMA, E. 2013. Cycles of change in Jomon settlement: a case study from eastern Tokyo Bay. *Antiquity* 87: 1169–81.

CZEBRESZUK, J. & J. MÜLLER (ed.). 2001. *Die absolute Chronologie in Mitteleuropa 3000–2000 v. Chr.* (Studien zur Archäologie in Ostmitteleuropa 1). Rahden: Marie Leidorf.

DENAIRE, A. 2011. Chronologie absolue de la séquence Hinkelstein-Grossgartach-Roessen-Bischheim dans le sud de la plaine du Rhin supérieur et le nord de la Franche-Comté à la lumière des dernières données, in A. Denaire, C. Jeunesse & P. Lefranc (ed.) *Nécropole et enceintes danubiennes du V^{e} millénaire dans le Nord-Est de la France et le Sud-Ouest de l'Allemagne*: 9–30. Strasbourg: Maison Interuniversitaire des Sciences de l'Homme-Alsace, Université de Strasbourg.

Dendro. n.d. Dendro- und C14-Daten der Schweiz. Available at http://www.erz.be.ch/erz/de/index/kultur/archaeologie/daten/dendro-_und_c14-daten.html (accessed 29 July 2014).

FOWLER, H.N. (trans.). 1925. *Plato in twelve volumes.* London: William Heinemann.

FURHOLT, M. 2003. *Die absolutchronologische Datierung der Schnurkeramik in Mitteleuropa und Südskandinavien* (Universitätsforschungen zur Prähistorischen Archäologie 101). Bonn: R. Habelt.

– 2014. Upending a 'totality': re-evaluating Corded Ware variability in Late Neolithic Europe. *Proceedings of the Prehistoric Society.* http://dx.doi.org/10.1017/ppr.2013.20

GESCHWINDE, M. & D. RAETZEL-FABIAN. 2009. *EWBSL: Eine Fallstudie zu den jungneolithischen Erdwerken am Nordrand der Mittelgebirge* (Beiträge zur Archäologie in Niedersachsen 14). Rahden: Marie Leidorf.

HEATON, T.J., P.G. BLACKWELL & C.E. BUCK. 2009. A Bayesian approach to the estimation of radiocarbon calibration curves: the IntCal09 methodology. *Radiocarbon* 51: 1151–64.

LEE, S. & C. BRONK RAMSEY. 2012. Development and application of the trapezoidal model for archaeological chronologies. *Radiocarbon* 54: 107–22. http://dx.doi.org/10.2458/azu_js_rc.v54i1.12397

MÜLLER, J. 2009. Dating the Neolithic: methodological premises and absolute chronology. *Radiocarbon* 51: 721–36.

MÜLLER, J. & J. VAN WILLIGEN. 2001. New radiocarbon evidence for European Bell Beakers and the consequences for the diffusion of the Bell Beaker phenomenon, in F. Nicolis (ed.) *Bell Beakers today*: 59–80. Trento: All'Insegna del Giglio.

MÜLLER, J., C. BECKER, H. BRUCHHAUS, E. KAISER, A. NEUBERT, S. PICHLER & M. ZABEL. 2000. Radiokarbonchronologie—Keramiktechnologie-Osteologie-Anthropologie-Raumanalysen. Beiträge zum Neolithikum und zur Frühbronzezeit im Mittel-Elbe- Saale-Gebiet. *Bericht der Römisch-Germanischen Kommission* 80: 25–211.

OTTAWAY, B. 1973. Dispersion diagrams: a new approach to the display of carbon-14 dates. *Archaeometry* 15: 5–12. http://dx.doi.org/10.1111/j.1475-4754.1973.tb00073.x

RAETZEL-FABIAN, D. 2000. *Calden.Erdwerk und Bestattungsplätze des Jungneolithikums. Architektur—Ritual—Chronologie*. Bonn: Habelt.

– 2002. Absolute chronology and cultural development of the Wartberg culture in Germany. Available at: http://www.jungsteinsite.uni-kiel.de/pdf/2002_2_fabian.pdf (accessed 29 July 2014).

REINGRUBER, A. & L. THISSEN. 2009. Depending on ^{14}C data: chronological frameworks in the Neolithic and Chalcolithic of southeastern Europe. *Radiocarbon* 51: 751–70.

SHENNAN, S.J. 1978. Archaeological 'cultures': an empirical investigation, in I. Hodder (ed.) *The spatial organisation of culture*: 113–39. London: Duckworth.

– 1989. Introduction: archaeological approaches to cultural identity, in S.J. Shennan (ed.) *Archaeological approaches to cultural identity*: 1–32. London: Unwin Hyman.

SHENNAN, S., S.S. DOWNEY, A. TIMPSON, K. EDINBOROUGH, S. COLLEDGE, T. KERIG, K. MANNING & M.G. THOMAS. 2013. Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nature Communications* 4: 2486. http://dx.doi.org/10.1038/ncomms3486

SHENNAN, S.J., E. CREMA & T. KERIG. 2014. Isolation-by-distance, homophily, and 'core' vs. 'package' cultural evolution models in Neolithic Europe. *Evolution and Human Behaviour.* http://dx.doi.org/10.1016/j.evolhumbehav.2014.09.006

STADLER, P., S. DRAXLER, H. FRIESINGER, W. KUTSCHERA, P. STEIER & E.M. WILD. 2001. Absolute chronology for early civilisations in Austria and Central Europe using ^{14}C dating with accelerator mass spectrometry with special results for the absolute chronology of Baden culture, in *Studia Danubiana Series Symposia II: Cernavodă III-Boleráz*: 541–62. Bucharest: Institutul Român de Tracologie.

ULLRICH, M. 2008. *Endneolithische Siedlungskeramik aus Ergersheim, Mittelfranken: Untersuchungen zur Chronologie von Schnurkeramik- und Glockenbechern am Rhein, Main und Neckar*. Bonn: Habelt.

VAN DE VELDE, P. 2012. Chronology of the Dutch Neolithic Bandkeramik culture: a new attempt, in C. Bakels & H. Kamermans (ed.) *The end of our fifth decade* (Analecta Praehistorica Leidensia 43/44): 293–305. Leiden: Faculty of Archaeology, Leiden University.

WHITTLE, A.W.R., F.M. HEALY & A. BAYLISS. 2011. *Gathering time: dating the Early Neolithic enclosures of southern Britain and Ireland*. Oxford: Oxbow.

WŁODARCZAK, P. 2009. Radiocarbon and dendrochronological dates of the Corded Ware culture. *Radiocarbon* 51: 737–49.

Received: 13 December 2013; Accepted: 15 April 2014; Revised: 20 June 2014