ROLLING OUT REVOLUTION: USING RADIOCARBON DATING IN ARCHAEOLOGY

Alex Bayliss

English Heritage, 1 Waterhouse Square, 138-142 Holborn, London, EC1N 2ST, United Kingdom.
Email: Alex.Bayliss@english-heritage.org.uk.

ABSTRACT. Sixty years ago, the advent of radiocarbon dating rewrote archaeological chronologies around the world. Forty years ago, the advent of calibration signaled the death knell of the diffusionism that had been the mainstay of archaeological thought for a century. Since then, the revolution has continued, as the extent of calibration has been extended ever further back and as the range of material that can be dated has been expanded. Now a new revolution beckons, one that could allow archaeology to engage in historical debate and usher in an entirely new kind of (pre)history. This paper focuses on more than a decade of experience in utilizing Bayesian approaches routinely for the interpretation of 14C dates in English archaeology, discussing both the practicalities of implementing these methods and their potential for changing archaeological thinking.

INTRODUCTION

Archaeologists have become accustomed to radiocarbon revolutions. It seems that as soon as we are getting to grips with the implications of one upheaval, another scientific development initiates new turmoil in our thinking. The constant challenge has been to ensure that the discipline exploits these successive waves of technical advance as effectively as possible. Chronology is essential for all archaeologies. It is a measure of its importance that developments in 14C science have fundamentally changed not only our views of the past, but also our approaches to archaeology itself.

This paper focuses on my experience of rolling out the latest 14C revolution in England. The types of archaeological questions posed by English sites are in no way exceptional, although there are, of course, local idiosyncrasies in the way archaeology is funded and organized, and in the types of sites that need to be dated and samples that are available for dating. What is unusual is that in England this new revolution has been underway for nearly 15 yr.

English Heritage1 was quick to realize the potential of routine Bayesian modeling of archaeological chronologies (Bayliss and Bronk Ramsey 2004). This means that we now have a corpus of well over 500 archaeological sites that have been dated using this methodology. Considerable efforts have been made to introduce the technique to the wider archaeological community, and to extend its application into the commercial sector (beyond projects funded by English Heritage or universities). Our experience of the potential and limitations of this methodology in practice, and of the challenges of building Bayesian chronological models routinely, may therefore be of value for other archaeologies that have yet to experience the full onslaught of this latest revolution.

THE FIRST RADIOCARBON REVOLUTION

As soon as the principles of 14C dating had been established (Arnold and Libby 1949), samples from English archaeological sites were submitted for dating. The first was wood from the Mesolithic site at Star Carr, Yorkshire, soon followed by samples from the Neolithic lakeside settlement at Ehenside

---

1English Heritage is the Government Agency responsible for archaeology and the wider historic environment in England; it previously formed part of the Department of the Environment (1970–80), the Ministry of Public Buildings and Works (1962–70), and the Ministry of Works (1944–62).
Tarn, Cumberland, and Stonehenge, Wiltshire (C-353, 9844 ± 350 BP; C-462, 4964 ± 300 BP; C-602, 3798 ± 275 BP; Arnold and Libby 1951:113; Libby 1951:292).

It quickly became apparent that the accepted chronology for British, indeed European, prehistory, based on typological crossdating with historical chronologies in Egypt, Mesopotamia, and the eastern Mediterranean, was untenable (Renfrew 1973: chapter 3). The English Neolithic, for example, was at least a millennium older than had been suggested previously (Piggott 1949; Renfrew 1974). Cultures, which had formerly been assigned to no more than 500 yr, now spanned 1500 yr or more. The pace of cultural change was much slower than had been anticipated. Elsewhere, the impact of this revolution was even more startling; suddenly there was chronology where previously there had been none. For the first time, world prehistory could be attempted (Clark 1961).

In England, as elsewhere, this revolution gained momentum throughout the 1950s and 1960s, although $^{14}$C dates were costly and difficult to obtain. The Ministry of Works was not able to commission its first dates until 1961 (Figure 1), when the $^{14}$C dating laboratory at the National Physical Laboratory at Teddington, England, was established (Callow et al. 1963). In these circumstances, it took time to accumulate a large enough sample of dates for meaningful patterns to appear.

THE SECOND RADIOCARBON REVOLUTION

Just as archaeologists were rebuilding some form of chronology from the collapsed ruins left by the first $^{14}$C revolution, it became clear that $^{14}$C dates needed to be calibrated (Suess 1967). These calibrated dates were several hundred years older than the uncalibrated $^{14}$C measurements. On the surface, this just required another shift in chronology—the English Neolithic became earlier by another 800 yr. But actually something more fundamental was wrong. The calibrated chronology exploded the whole diffusionist framework of prehistory by demonstrating that many of the developments

---

2In this paper, $^{14}$C ages are cited as “BP,” dates “BC” or “AD” are estimated or actual dates on the calendar scale, dates “cal BC” or “cal AD” derive from calibrated $^{14}$C ages, and dates “cal BC” or “cal AD” in italics are posterior density estimates derived from Bayesian modeling.
that were supposedly carried to Europe from the eastern Mediterranean actually appeared in Europe first (Renfrew 1973: chapter 5). In England, nowhere was this more startling than at Stonehenge, which was now too early to be the product of influence from Mycenae.

Undermining the culture-historical approach, which had been the basis of prehistory for a century, left a vacuum. The “good objective chronology” provided by \(^{14}C\) dates (Renfrew 1973:109) demanded explanation. This second \(^{14}C\) revolution did not create the “New Archaeology,” but it contributed to a climate in which processual archaeology could flourish (Trigger 2006: chapters 7–8).

In England, the second \(^{14}C\) revolution began in the early 1970s and continues today. As tree-ring calibration has been extended ever further back in time (Suess 1967; Ralph et al. 1973; Pearson et al. 1986; Kromer and Becker 1993; Stuiver et al. 1998; Reimer et al. 2004), more and more archaeologists have had to “cope” with calibration (Pearson 1987). The establishment of an internationally-agreed series of calibration curves and the availability of user-friendly calibration software have been critical in allowing calibration to become routine. Although it did take some time for the need for calibration to be appreciated, nowadays almost all archaeologists working in the Holocene in England will calibrate as a matter of course. This will no doubt extend into Paleolithic archaeology as calibration is extended further back in time (van der Plicht et al. 2004).

THE CONTINUING RADIOCARBON REVOLUTION

Part of the success of the second \(^{14}C\) revolution can be attributed not just to advances in calibration itself, but, from the early 1970s, to the increasing availability of dates. \(^{14}C\) measurements were still expensive, but as the number of laboratories increased, it became possible for archaeologists to send samples for dating on a commercial basis. This expansion was linked to the increasing use of liquid scintillation spectrometry, but also to the expansion of “rescue” archaeology undertaken in advance of construction projects (Rahtz 1974). For example, in England the number of samples submitted by the Ministry of Public Buildings and Works/Department of the Environment increased from a handful each year during the 1960s to an average of more than 100 during the 1970s (Figure 1).

As more samples were processed, the selection of appropriate material for dating became an everyday issue; and here English archaeologists must accept some blame. Despite the clear, practical, and accessible guidance provided by Hans Waterbolk (1971), many samples with age-at-death offsets and which were not directly associated with the archaeological event of interest were submitted for dating. Although the sheer quantity of material required for conventional dating undoubtedly led to the submission of samples simply because they were big enough, this is not the whole story. In many cases, more suitable samples were available, or samples could have been rehabilitated, just by the identification and isolation of fragments of charcoal from short-lived species.

In the mid-1980s, dating by accelerator mass spectrometry (AMS) became available to archaeologists (Dennell 1987), although it was only in the early 1990s that AMS could be obtained at an equivalent price and precision to conventional dating (Figure 1). In some areas of archaeology, the impact of AMS has been revolutionary (Taylor 1995). The ability to date tiny samples has enabled questions to be addressed that were previously beyond the reach of \(^{14}C\) (either because the datable material was too small or too precious). Our understanding of early plant domestication, the spread of anatomically modern humans, and the population of the New World has been transformed (Harris et al. 1987). But for most archaeologies, including those of England, AMS has expanded the remit of the second \(^{14}C\) revolution rather than forming a revolution in its own right. Archaeology is still (post)processual.
This is not to deny the impact of AMS, which has allowed archaeologists to submit a much greater range of material for dating. Now, almost every site produces many thousands of potentially datable samples. The concept of a “radiocarbon sample,” specially excavated and recorded, painstakingly wrapped in swathes of tin foil, is outmoded. Every charred seed or fragment of charcoal recovered from flotation and every bone or pot sherd recovered from the finds tray is a potential sample. New types of material, cremated bone and residues on pottery (Lanting et al. 2001; Hedges et al. 1992; Berstan et al. 2008), can now be dated. So, our choice has never been greater, but again we must admit failure. Generally, short-lived material is now isolated for dating and increasingly single-entity samples (Ashmore 1999) are the rule rather than the exception, but routinely still far too little attention is paid to the association between the sample, the context from which it was recovered, and the archaeological event that our dating targets (Lanting and van der Plicht 1994; Prendergast 2000).

In England, the availability of AMS coincided with a significant increase in the number of samples dated. Partially, this is because over the past 20 yr the cost of measurements has remained stable (and so in real terms has gone down), but largely this is a product of a change in the planning regime in the early 1990s, which meant that developers paid for the investigation of archaeological remains that were destroyed during construction projects. The explosion in commercial archaeology that followed has been reflected in the number of 14C dates obtained from English sites (for example, in Greater London, Figure 2).

Figure 2 Sources of funding for 14C measurements undertaken on samples from Greater London (1959–2004)

THE THIRD RADIOCARBON REVOLUTION

This was the state of 14C dating in English archaeology in the early 1990s when the first papers outlining a new approach for interpreting 14C dates appeared (Naylor and Smith 1988; Buck et al. 1991, 1992, 1994a,b, 1996; Christen 1994). The technical details of Bayesian chronological modeling will be not be covered here, as they are discussed in another paper in this volume (Bronk Ramsey, this issue). What is of relevance for this narrative is that the potential of the approach for constructing site-specific chronologies was immediately appreciated.
The vast majority of archaeological excavations in England, then and now, are development-led. Many of these excavations are on sites with extensive and complex stratigraphic relationships; single-context recording and the construction of Harris matrices are standard professional practice (Harris 1989; Spence 1993). Consequently, a very large proportion of sites already had significant quantities of relative dating information that could be used as “informative prior beliefs” in Bayesian models and, because of AMS, they almost all also had a wealth of datable samples.

The last requirement for rolling out Bayesian chronological modeling routinely was fulfilled in 1993, when the first software that could be used by archaeologists to construct models was issued (OxCal v1.3; Bronk Ramsey 1994). This was swiftly followed by further versions of OxCal (Bronk Ramsey 1995, 1998, 2001, 2008, this issue), BCal (Buck et al. 1999), and DateLab (Jones and Nicholls 1999).

INTRODUCING THE CONCEPTS

Most people have trouble with statistics, which produces “instant mental paralysis in many otherwise competent archaeologists” (Fletcher and Lock 1991: viii). It is therefore much more effective to introduce the concepts of Bayesian modeling to an archaeological audience with words, pictures, and examples rather than equations (Figure 3). Analogy to archaeological methods and theory can be helpful—be it the hermeneutic spiral of post-processualism (Hodder 1992) or the hypothetico-deductive method of processualism (Binford and Binford 1968). Alternatively, the simple logic that a chronology that relies on all the dating information available—\( ^{14} \text{C} \) dates, stratigraphy, coins, typology, etc.—is bound to be more reliable than one that relies only on a single strand of information may make the principles involved clear.

At this stage, an example, such as that shown in Figure 4, is usually helpful. The first point to highlight from this graph is that, if the concept of “the radiocarbon sample” became outmoded with the advent of AMS, then so now is that of “the radiocarbon date.” Renfrew’s “good objective chronology” is now contaminated by our archaeological opinions—the “posterior density estimates” shown in black in the graphs are not just based on independent scientific evidence, but also on the “prior beliefs” that have been included in the model. So, the onus is squarely on us as archaeologists to ensure that those beliefs are correct.

The idea that posterior density estimates are simply more precise date estimates that include archaeological information as well as the calibrated \( ^{14} \text{C} \) date is fairly simple. So, for example, at Wayland’s Smithy I, \( KIA-27625 \) suggests that individual ws15 died and was buried in 3600–3550 cal BC (84%...
Figure 4  Probability distributions of dates from Wayland’s Smithy (see Whittle et al. 2007a for full details of this model). Each distribution represents the relative probability that an event occurs at a particular time. For each 14C date, 2 distributions have been plotted: one in outline that is the result of simple 14C calibration, and a solid one based on the chronological model used; the “event” associated with, for example, ‘KIA-27625’ is the growth of the person whose bones were dated. The other distributions correspond to aspects of the model. For example, the distribution ‘start Wayland’s Smithy I’ is the posterior density estimate for the first burial activity on the site. The large square brackets down the left-hand side and the OxCal keywords define the overall model exactly.

Probability) or 3545–3525 cal BC (11% probability; Figure 4). Slightly more difficult to grasp is the idea that a posterior density estimate does not have to equate with a 14C sample. So, start Wayland’s Smithy I is our estimate for when the first burial was placed in that monument, in 3610–3550 cal BC (83% probability) or 3545–3525 cal BC (12% probability; Figure 4). This estimate is derived from all the 14C dates from the site, as well as all the archaeological information we have included in the model.
Moving on to the components of this model, the easiest part of a Bayesian chronological model for most archaeologists to grasp is the need to constrain the calibration of $^{14}$C dates on the basis of relative dating known from stratigraphic sequences. Stratigraphy is bread-and-butter to the profession, and is probably the most familiar source of chronological information for archaeologists. In the burial chamber at Wayland's Smithy I, we have a series of articulated, or partially articulated, human skeletons whose relative date of burial can be inferred from their relative positions in the tomb and expressed as a stratigraphic matrix (Figure 5; and see also Whittle et al. 2007a: Figure 2).

At this point, it is absolutely critical to emphasize the fundamental importance of the taphonomy of the dated samples. How did the datable material get into the deposit from which it was recovered? Stratigraphy provides relative dating between deposits. $^{14}$C dating does not date deposits. It dates organic material. Consequently, the relative sequence of the deposits cannot be used to constrain the calibration of the $^{14}$C dates unless the samples died and were deposited at the time when the context in which they were found formed. This is, of course, never known but may be interpreted with varying degrees of certainty by the archaeologist (Figure 6). Short-life, single-entity samples are not enough. All samples must be presumed to be residual (older than the deposit in which they were found) unless there is evidence, or at least a convincing argument, that they were fresh when deposited. Guilty until proven innocent!

---

**Figure 5** Harris matrix of stratigraphic relationships between articulated/partially articulated skeletons from Wayland’s Smithy I.

---

**Figure 6** The relationship between interpretation, an archaeological context, and the material recovered from it.
This consideration of taphonomy is so vital because the stratigraphic information included in a model is extremely “informative.” This means that it has a considerable effect on the outputs of our model, and that our estimated chronology will be incorrect if the order between dated samples incorporated into the model is wrong. Occasionally stratigraphy may be resolved incorrectly, occasionally samples may get mislabeled or contaminated, but incorrect interpretation of the taphonomy of the dated material is by far the most common source of error in this situation.

The use of stratigraphy as informative prior information is easy to grasp in principle, if difficult to implement successfully in practice. The need to impose a statistical distribution on a group of \(^{14}\)C dates to counteract the inevitable statistical scatter—which will make them appear to start earlier, end later, and endure for longer than was actually the case—is mysterious to most archaeologists (see e.g. Brown 2007; Sheridan 2007). Indeed, when Bayesian methods were first introduced to archaeology, it took some time for the importance of this type of prior information to be appreciated (Bayliss 1995; Bayliss and Harry 1997; Steier and Rom 2000).

One strategy to demonstrate the need for such “uninformative” prior information is to simulate a series of calibrated \(^{14}\)C dates from a suite of samples of known age. Figure 7, for example, shows 25 calibrated \(^{14}\)C dates simulated from a uniform series of calendar dates running from 6799 to 6775 BC. When archaeologists are asked to estimate from this graph when the activity started, when it ended, and for how long it continued, they will almost all get it wrong. There is a strong tendency to

Figure 7 Calibrated \(^{14}\)C dates (Stuiver and Reimer 1993; Reimer et al. 2004) from a fictitious settlement. The \(^{14}\)C ages have been simulated from samples that actually date between 6799 and 6775 BC.
suggest from such graphs that activity started earlier, ended later, and went on for longer than it did in reality.\textsuperscript{4} This is because \textsuperscript{14}C dating is a probabilistic process and estimated values scatter around the actual values. Visual inspection of graphs of calibrated dates is misleading because no account is taken of this scatter.

If, however, we construct a Bayesian model which simply incorporates the information that the dates derive from a coherent (uniformly distributed) phase of archaeological activity (Figure 8), accurate date estimates are produced, even when no stratigraphic information is included in the model. In this example, the model suggests that the activity started in 6955–6750 cal BC (95% probability; start; Figure 8) and ended in 6815–6680 cal BC (95% probability; end; Figure 8)—in both cases comfortably including the actual dates of 6799 BC and 6775 BC, respectively. Other examples of this type of simulation are given in Bayliss et al. (2007a, 2008).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure8.png}
\caption{Probability distributions of dates from a fictitious settlement, incorporating a uniform distribution for the occupation of the settlement (informative prior information from stratigraphy is not included in this model). The format is identical to that of Figure 4. The simulated dates are those shown in Figure 7 (6799–6775 BC). The large square brackets down the left-hand side along with the OxCal keywords define the overall model exactly.}
\end{figure}

\textsuperscript{4}An audience of archaeologists which met at Çatalhöyük, Turkey in August 2008 were asked to estimate the dates of establishment and abandonment, and duration of use of this settlement from the simulation shown in Figure 7. The vast majority of the audience got the answers importantly wrong—91% estimated the start date inaccurately, all estimated the end date inaccurately, and everybody significantly overestimated the duration (see also Bayliss et al. 2007a: footnote 5)!
The need for a statistical distribution to be imposed on a group of dates is difficult to grasp, even when the alternative of not imposing a distribution demonstrably leads to incorrect interpretation. This supposedly “uninformative” prior information can actually constrain the calibrated $^{14}$C dates considerably (Figure 8), especially in cases where the duration of past activity was short in comparison to the range of the calibrated dates (as is often the case). In these circumstances, the imposition of any type of distribution, although essential if accurate date estimates are to be produced, makes most archaeologists uncomfortable. Fortunately, the use of a uniform distribution has proven very forgiving in practice, at least for site-based models (Bayliss et al. 2007a: Figures 15–17). Now that a range of distributions are available (Bronk Ramsey, this issue; Karlsberg 2006), the task of determining which distribution should be chosen for a particular application will be more complex. It remains to be seen which approaches will prove robust in everyday usage.

MODELING IN ACTION

The reflexive approach to implementing Bayesian chronological modeling in archaeological projects that has been adopted by English Heritage over the past decade is summarized in Figure 9.

The first step is to construct a simulation model for the chronology of the site. As modeling is a formal, explicit process, this means that the objectives of the dating program have to be clearly defined. Are you interested in the actual dates between which a site was occupied? Or is the duration of its use more interesting? Or perhaps, the date when occupation ended is well known from other sources, and the dating program needs to be designed more to determine when the site was established.

The components of the simulation model are those of any model. First, the available “informative” prior beliefs have to be established. A list of all datable material that is demonstrably not residual in the context from which it was recovered forms a pool of potential $^{14}$C samples. The site matrix is needed, to allow relative dating sequences between potential samples to be established. Then, the structural phasing may be needed, for example, to determine which deposits relate to the use of which building so that meaningful groups of samples relating to periods of activity (uniformly distributed phases/groups in modeling terms) can be identified. Once the prior beliefs of the model have been established, $^{14}$C dates can be simulated from the pool of suitable datable material. Errors on the measurements are estimated from those recently obtained by the selected laboratory on similar material of similar age. In this process, the actual date of the site has to be fed into the model,
although in practice it is usually possible to give reasonable estimates of this parameter, even if the site is only known to be, for example, “middle Bronze Age.” Multiple models can be run for different actual ages and for different sampling strategies to see which approach might be most effective. In cases where the date of activity is really entirely unknown, a small number of “range finder” dates are useful to inform the simulation.

Probably the most important outcome of these simulations is to identify which archaeological problems cannot be resolved to a useful resolution with the available samples, stratigraphy, and phasing. This prevents dates being obtained which simply confirm what the archaeologist already knew—“the site is middle Bronze Age.” Resources can be concentrated to ensure that targeted questions are satisfactorily resolved. Hard decisions may have to be made. For example, at Lodge Farm, St Osyth, Essex (Germany 2007), it was decide to concentrate dating resources on the Neolithic causewayed enclosure and Bronze Age funerary activity on the site, and not to attempt to date the middle Iron Age settlement using $^{14}C$ at all. This ensured that the earlier periods of activity could be dated effectively (e.g. Figure 10, and see Hamilton et al. 2007a: Figures 64, 65).

After all this effort, the minimum number of samples that might be able to achieve the archaeological objectives should be submitted for dating (assuming that the site falls on the most helpful piece of calibration curve possible). When the results are reported, these replace the simulated dates in the model, and, of course, gremlins appear! Sometimes, something will have gone wrong with a $^{14}C$
measurement in the laboratory, but almost always it will be the assessment of the taphonomy of the
dated sample (and therefore its relative chronology within the dated sequence) that will be in error.
Once these problems have been identified and remodeled in an appropriate way, further simulated
dates are added to the existing dates and the cycle repeats (Figure 9). Ideally, it repeats until adding
more simulated dates does not materially improve the precision of the chronology produced by the
model. In practice, however, usually either there is no more money for more samples or the post-
evacuation timetable cannot accommodate further rounds of sampling.

This process is time-consuming; frequently as much staff time is spent in selecting samples and run-
ing simulations as is spent in analysis and publication of the final suite of results. However,
projects where the samples are selected around the model, rather than where the model is grafted
onto an existing series of dates, have consistently provided much more precise chronologies and
been much more cost-effective.

Finally, a model is built with all the results and described for publication. Frequently, it is useful to
provide a series of alternative models, perhaps exploring different readings of the archaeological
sequence or different statistical assumptions. These alternative models are termed sensitivity analy-
ses and provide an indication of the reliability of the preferred model (see Bayliss and Whittle 2007
for examples of sites where multiple models are reported).

THE POTENTIAL IN PRACTICE

Since 1993, when Bayesian modeling was first implemented by its Scientific Dating Team, English
Heritage has submitted over 4500 samples for 14C dating. Of these, well over 3000 (73%) have been
selected and interpreted within a Bayesian statistical framework (Figure 11).

![Figure 11 Number of 14C measurements funded by English
Heritage since Bayesian modeling became routinely available
for the selection and interpretation of 14C dates (1993–2007).](https://doi.org/10.1017/S0033822200033750)
Bayesian modeling is nearly always useful for sites which are excavated, where 90% of submitted samples are selected and interpreted in this way. In contrast, only 17% of samples submitted for landscape survey projects are subjected to formal modeling. This is because many samples submitted during such projects provide range finder dates. For example, the top and bottom of a sediment column may be dated to determine the date span of a paleoenvironmental record, or a fish-trap in the intertidal zone may be dated to determine whether it is of prehistoric or medieval date. In such circumstances, simple calibration suffices. Where Bayesian models have been constructed for such projects, they have usually been designed to provide chronologies for geomorphological models of river terrace development (Chiverrell et al. 2008).

Whereas the use of Bayesian chronological modeling in archaeological survey is likely to remain relatively restricted, it is probable that the number of environmental sequences that are subject to Bayesian analysis will increase over the coming years. At present, 48% of samples submitted from such sites by English Heritage are part of Bayesian models. Although the statistics of applying Bayesian modeling to sedimentary deposits have been available for some time (Christen et al. 1995), its full implementation has been restricted because Bayesian age-depth modeling has only become routinely available recently with the advent of OxCal v4 (Bronk Ramsey 2008) and BPeat (Blaauw and Christen 2005). Although some attempt has been made to model environmental sequences using more basic models (Bayliss 2005; Wiltshire and Bayliss 2006; Boomer et al. 2007), the acknowledged limitations of these have restricted their use. With the routine availability of more sophisticated and more appropriate models, the chronological modeling of environmental deposits is likely to become more widely adopted over the coming years (Blockley et al. 2007). It should be particularly noted, however, that the accuracy of such chronologies is crucially dependent on the derivation of the dated material in the sediment (which is often problematic).

Archaeological excavation in England is overwhelmingly undertaken on sites dating from the Holocene, and so almost all the chronological models that have been built over the past 15 yr are also on sites of this period. Less than 4% of the \(^{14}\)C measurements included in the Archaeological Site Index to Radiocarbon Dates from Great Britain and Ireland are on Palaeolithic material (Figure 12).

![Figure 12](https://doi.org/10.1017/S0033822200033750)
Although, since the advent of AMS, $^{14}$C dating has been fundamental to our understanding of the period, material suitable for $^{14}$C dating is often scarce, sites may lie beyond the limit of the technique, or other scientific dating methods may be more appropriate. In the Paleolithic, synthetic studies modeling particular archaeological issues (Blackwell and Buck 2003; Blockley et al. 2004) are more common than site-specific applications.

Bayesian modeling has, however, proven useful on sites of all periods within the Holocene. As yet, there have been comparatively few applications on Mesolithic sites, although this seems to be more a product of the rarity of datable deposits than an indication that archaeology of this period is unsuitable for the technique. The hut at Howick, Northumberland produced a series of dates that produced a coherent chronology spanning a century or two in the earlier 8th millennium cal BC (Figure 13; Bayliss et al. 2007b), and chronologies of similar resolution have been produced for human remains from Aveline’s Hole, Somerset (Marshall and van der Plicht 2005) and for occupation at Star Carr, Yorkshire (Dark et al. 2006).

It is probably in the earlier Neolithic that Bayesian modeling has so far made its greatest impact on English archaeology. A series of studies are providing dating at a generational scale for a sample of the long barrows (Bayliss and Whittle 2007), causewayed enclosures (Germany 2007; Allen and Bayliss 2008; Mercer and Healy 2008; Whittle et al. 2008), and cursus monuments (Barclay and Bayliss 1999; Bayliss et al. 2003a), which lie at the heart of our understanding of this period. For the first time, this body of evidence has allowed comparison between sites used at the same time (Figure 14; Whittle et al. 2007b). The sequence of sites and the tempo of change can now be considered, from sites and material that were previous undifferentiated somewhere within the earlier 4th millennium cal BC (Whittle et al. 2008). Furthermore, the first attempts to formally model the date when Neolithic practices appeared in England are being undertaken (Bayliss et al. 2008), to place this mass of monumentality within the context of everyday life.

The impact of Bayesian chronologies on our understanding of the early Neolithic of England may just be beginning, but it stands in stark contrast to the position in research into the later Neolithic. Following the pioneering use of Bayesian methods in the dating of Stonehenge (Allen and Bayliss 1995; Bayliss et al. 1997), little progress was made for more than a decade. The comparatively precise dating of this monument stood as an isolated beacon in a fuzzy sea of uncertainty. Nothing more graphically illustrates that chronology is sterile—just a number—if it does not allow comparison. It is only recently that dating of other late Neolithic sites nearby has begun to allow informed discussion of the context of Stonehenge and the pace of change at the end of the 3rd millennium cal BC, and to discuss alternative models for the chronology of Stonehenge itself (Bayliss et al. 2007e; Parker Pearson et al. 2007). There is as yet little beyond Wessex.

We also lack a critical mass of applications in the Bronze Age for the full impact of Bayesian chronologies to be realized in this period. Wiggle-matching of floating tree-ring series has allowed a number of sites to be precisely dated (Bayliss et al. 1999, 2003b; Bayliss and Pryor 2001; Masefield et al. 2004), yet this really has led only to wider discussion of the impact of the introduction of sewn-plank boat technology into northwestern Europe (van der Noort et al. 1999; Wright et al. 2001). Other sites and monuments have been precisely dated (Bayliss 2002a; Bayliss et al. 2002; Hamilton et al. 2007a, 2008; Marshall et al. 2008a), including even those that fall on the calibration plateau at the end of the 3rd millennium cal BC (Crowson and Bayliss 1999; Marshall et al. 2008b), although not as yet in sufficient numbers to allow more than site-specific discussion.

At the very start of the 1st millennium cal BC, at the transition from the Middle to the Late Bronze Age in England, however, the impact of precise Bayesian chronologies is becoming apparent. The
Figure 13: Probability distributions of dates from the Mesolithic structure at Howick, Northumberland, during site phases 1A, 1B, 2, and 3 (see Bayliss et al. 2007b for full details of this model). The format is identical to that of Figure 4. The large square brackets down the left-hand side along with the OxCal keywords define the overall model exactly.
slope of the calibration curve in this period allows precise dating, even on sites with limited vertical stratigraphy and relatively limited numbers of \(^{14}\)C dates (Figure 15). This is allowing questions such as the pace of the transition from Middle to Late Bronze Age material culture and the longevity of settlement to be assessed (Bayliss et al., forthcoming). It also allows English archaeological evidence to be discussed within the protohistoric framework of archaeology from continental Europe (Needham et al. 1998; Bayliss et al. 2006).
Bayesian modeling has rehabilitated 14C dating in the English Iron Age. In southern England, calibrated 14C dates simply do not provide the precision necessary to date archaeological activity in this period to a resolution that allows comparison with artifact-based chronologies. At present, no Bayesian applications have tackled the challenges of the major plateau in the calibration curve in the earlier 1st millennium cal BC, although such limitations do not apply to sites dating to the middle and late Iron Age. Here, useful chronologies can be constructed (Figure 16), and there are an increasing number of site-based studies (Buck and Litton 1995; Hey et al. 1998; Bayliss et al. 2003c, 2005; Marshall and Hamilton 2007). These are particularly important in providing independent dating for artifact, particularly pottery, typologies. At Sutton Common, Yorkshire, Bayesian modeling of an environmental sequence has allowed this to be correlated with a neighboring waterlogged archaeological site dated by 14C and dendrochronology (Gearey et al., forthcoming). Where the necessary stratigraphic sequences and suitable samples exist, the potential of this approach is such that it has been recommended that “multiple radiocarbon dating must become routine for Iron Age sites” (Haselgrove et al. 2001:5).

Figure 16 Probability distributions of dates from Conderton Camp, Worcestershire (see Bayliss et al. 2005 for full details of this model). The format is identical to that of Figure 4. The large square brackets down the left-hand side along with the OxCal keywords define the overall model exactly.
By contrast, studies in the Roman period remain extremely rare as there is a perception that artifact-based dating is more precise (and less expensive!). Useful dating has been provided for the (in fact not uncommon) sites where artifactual evidence is sparse (Bayliss 2002b; Marshall 2004), but Bayesian chronologies have really made their impact in the succeeding sub-Roman and early medieval periods. Precise dating has been provided for sites in western England that are largely aceramic in the centuries following the collapse of Roman Britain (Bayliss 1995; Bayliss and Harry 1997; Bayliss and Bronk Ramsey 2007; Hamilton et al. 2008, 2009), and for sites further east where few artifacts were recovered (Bayliss et al. 1996; Bayliss and Hey 2004). It has proven possible to correlate paleoenvironmental evidence with nearby archaeological sites (Bayliss 2005; Wiltshire and Bayliss 2006), and even provide dating that is precise enough to be useful in areas with good artifact-based dating (Figure 17; Scull and Bayliss 1999). Overall, Bayesian modeling enables $^{14}$C dating to provide chronologies of sufficient resolution to allow archaeology to engage in historical debate in this period.

![Figure 17](https://doi.org/10.1017/S0033822200033750)

**Figure 17** Probability distributions of dates from Buttermarket, Ipswich, Suffolk (see Scull and Bayliss 1999 for full details of this model). The format is identical to that of Figure 4. The large square brackets down the left-hand side along with the OxCal keywords define the overall model exactly.

Increasingly, Bayesian chronologies are also allowing $^{14}$C dating to play a part in medieval archaeology—a period when archaeology itself has been characterized as “an expensive way of telling us what we already know” (Sawyer 1976)! Here the challenge is to provide chronologies that are both precise and accurate enough to compare with documentary dating and dendrochronology. This reinforces the importance of otherwise subtle technical issues, such as small marine or freshwater components in human diet, which may not have a substantive effect on less precise chronologies but may
be significant for Bayesian applications (see Bayliss et al. 2007a for further discussion of these issues). So far, models have been concentrated in cemetery studies, where relative dating from stratigraphy is often good, but artifacts are often scarce (Bayliss et al. 2004, 2007f; Sidell et al. 2007). The first steps are also being taken to utilize 14C dating with the deep stratigraphy of middle and late Saxon towns to provide chronologies for the establishment and development of towns such as Oxford (Marshall 2003), and Bayesian modeling is also allowing chronologies to be produced for medieval and post-medieval standing buildings that are unsuitable for dendrochronology (Figure 18; Bayliss 2007). Although the laboratory accuracy required for such applications is challenging, the posterior density estimates produced are consistently in agreement with independent tree-ring dating (Hamilton et al. 2007b; Tyers et al. 2009).

Figure 18 Probability distributions of dates for the felling of timbers from the north wing of Baguley Hall, Greater Manchester (see Hamilton et al. 2007b for full details of this model). The posterior density estimate for Felling/Construction, cal AD 1450–1470 (at 95% probability), can be compared with the felling date of 2 timbers from this wing determined by dendrochronology, AD 1443–1478 (95% confidence). The format is identical to that of Figure 4.

**COME THE REVOLUTION?!**

So, we now have a methodology that allows 14C dating to produce accurate chronologies routinely to a resolution of less than a century—often to within the span of a human lifetime and sometimes down to a generation or a few decades. But it is hard work. It requires rigorous attention to detail, and demands much of both archaeologists and 14C laboratories in terms of selecting and measuring samples accurately. Is it worth the effort involved? What is at stake?

In today’s multivocal archaeology, perhaps the time for revolution is past—there is no single paradigm at the root of the subject, no unilinear evolutionary view of the development of the discipline, and so perhaps nothing to overturn (Trigger 2006:6–9). Maybe in its maturity this revolution will change not so much how people choose to do archaeology, but more the questions that they can address within whichever theoretical framework they prefer. As David Clarke (1973:10) remarked more than a generation ago, “the consequences arising from the introduction of new methodologies are of far greater significance than the new introductions themselves.”
For example, over recent decades there has been an emphasis on the long-term in archaeology (Hodder 1987). The time-depth of the archaeological record is undoubtedly one of the strengths of the discipline, but is this emphasis not also predicated by the fact that most of our existing chronologies have a resolution that only allows us to look at the long-term, because our chronology is “smeared” over centuries of uncertainty (Baillie 1991; Dean 1993:60)? But now we are beginning to see that prehistoric societies were much more dynamic than we have previously appreciated: change happens more quickly, occupation is less enduring (Whittle et al. 2007b). In the English early Neolithic, our new corpus of Bayesian chronologies is beginning to allow us to look for medium-term structures and Braudel’s chains of events. We can consider causality at the scale of individual human agency (Dobres and Robb 2000) and attempt to untangle medium-term from long-term structure, and perhaps even consider the relationship between agency and structure for particular episodes in the past. Prehistorians need to remember that “the span of a human life is in itself a significant durée” (Le Goff 1989:405).

Scholars working in historic and protohistoric periods know this (Foxhall 2000). Here the outcome of routine Bayesian chronology will be to allow archaeological evidence to engage with a greater range of historical debates. If we can provide the chronological resolution, the longer timescale of archaeology can flesh out the bones of earlier historic periods and illuminate their origins. In any case, our chronologies can bring new forms of evidence into these debates—environmental evidence from pollen cores, sites, and typological series dated so that the pace and sequence of environmental and cultural change can be mapped and related.

Above all, it must be remembered that a date, be it ever so precise, is just a number. Dates only become meaningful when they allow comparison between different events in the past and measure the lapse of time between them. The experience of English archaeology shows that, 15 yr after the first applications, it is only now that we are beginning to accumulate a sufficient corpus of well-dated sites for their impact to be apparent. If we want these new archaeologies, then revolution has to become routine.

ACKNOWLEDGMENTS

Revolution is hard work! The work embodied in this paper would not have been possible without the enthusiasm and dedication of my past and present colleagues in the Scientific Dating Team of English Heritage: Peter Marshall, Derek Hamilton, John Meadows, Jane Sidell, and Isabelle Parsons; and of the English Heritage dendrochronologist based at Sheffield University, Cathy Tyers. We have been privileged to collaborate with a significant proportion of the archaeology sector working in England over the past 15 yr—and our thanks go to all the excavators who have given generously of their time and welcomed us into their projects, even if the process was to them sometimes painful and slightly baffling! Without the statisticians and programmers who have developed the methods and programs that allow us to roll-out this revolution, Christopher Bronk Ramsey and Caitlin Buck in particular, nothing would have been possible. Finally, thanks go to Christopher Bronk Ramsey, Gordon Cook, Gerry McCormac, and Johannes van der Plicht, and their teams at Oxford, Glasgow, Belfast, and Groningen, respectively, who have labored over many years to produce the accurate 14C ages that form the bedrock of our models.
REFERENCES


A Bayliss


Stuiver M, Reimer PJ, Bard E, Beck JW, Burr GS,


