

A BAYESIAN APPROACH TO THE ESTIMATION OF RADIOCARBON CALIBRATION CURVES: THE INTCAL09 METHODOLOGY

T J Heaton¹ • P G Blackwell • C E Buck

Department of Probability and Statistics, University of Sheffield, United Kingdom.

ABSTRACT. This article presents a new approach to the construction of radiocarbon calibration curves. The Bayesian methodology was developed specifically to facilitate construction of the 2009 updates to the internationally agreed ¹⁴C calibration curves known as IntCal09 and Marine09. The curve estimation approach taken uses Markov chain Monte Carlo sampling, specifically a Metropolis-within-Gibbs sampler, which offers improved flexibility and reliability over the approaches used in the past. In particular, the method allows accurate modeling of calibration data with ¹⁴C determinations that arise from material deposited over several consecutive calendar years and that exhibit complex uncertainty structures on their calendar date estimates (arising from methods such as wiggle-matching and varve counting).

1. INTRODUCTION

The original radiocarbon dating work of Willard Libby was based upon the assumption that the proportion of radioactive carbon (¹⁴C) had been constant throughout history. It was soon recognized that, in fact, the proportion of ¹⁴C has varied considerably over time. For accurate ¹⁴C dating, it is necessary to identify these fluctuations and adjust the dates obtained using Libby's technique (with the assumption of a constant ¹⁴C proportion) accordingly. However, the processes underlying these changes are complex and not sufficiently understood to permit a deterministic approach. Instead, a data-based calibration curve is required based upon a large set of calibration data with ¹⁴C determinations and, independently found, calendar date estimates.

Two key factors determine the reliability of such a ¹⁴C calibration curve. The first is the quality and quantity of the calibration data upon which it is based. Equally important, however, is the appropriateness of the methods used in getting from these data to an estimate of the curve. Within the ¹⁴C community, much work has been put in to the former, but the latter has only become the focus of real debate in recent years. This is especially significant since many of the calibration data sets possess complex structures that have not always been taken into account during curve estimation. Without formal modeling of such structures, not only is one not able to use the calibration records to their full potential but, more seriously, one may draw incorrect inference. In this article, building upon the IntCal04 (Reimer et al. 2004) methodology of Buck and Blackwell (2004) and Blackwell and Buck (2008), we set out a complete Bayesian approach to curve construction using Markov chain Monte Carlo (MCMC). Our statistical modeling framework permits clearer and more accurate modeling of the structures and uncertainties within calibration data and the curve-building process. Furthermore, it can provide samples of plausible curves over the entire calendar range of interest simultaneously as well as pointwise summary statistics. The additional information this gives can, in principle, be incorporated into the calibration process itself.

Calibration curve construction is based upon a series of calibration data having ¹⁴C determinations X_i together with calendar date estimates T_i that are found independently of the ¹⁴C determining process. Such data can have several complications. Often, ¹⁴C determinations are not performed upon a sample representing a single year but rather a block of material representing n_i consecutive years. Furthermore, the processes by which the calendar date estimates are obtained may be imprecise and also significantly dependent upon other data in the calibration database. Factors such as these will affect the resultant curve and should not be ignored in any curve-building methodology. Alternative

¹Corresponding author. Email: t.heaton@shef.ac.uk.

calibration curves, for example the CalPal-2007_{Hulu} curve of Weninger and Jöris (2008) or that of Fairbanks et al. (2005), do not, however, provide a rigorous framework to incorporate them.

The desire to account for such features provided motivation for the previous IntCal04 methodology (Buck and Blackwell 2004). Here, the calibration curve was modeled as a random walk with the data providing noisy observations of its value at a set of, mainly unknown, calendar dates. Posterior means and variances were then evaluated pointwise using subsets of the data. While this addressed many of the issues involved in modeling the data themselves, it did possess some drawbacks. By calculating points on the calibration curve separately, it was not possible to determine covariances between values at different calendar dates. Furthermore, there was a limit to the modeling flexibility, for example, the inclusion of known calendric ordering within data, which could be incorporated without some degree of approximation.

An MCMC-based approach offers a number of potential advantages over the methodology used to create IntCal04. In addition to significantly greater flexibility in modeling the structures within calibration data, it allows unknown parameters such as the random walk's variance to be estimated adaptively from the data. Also, by generating plausible samples, given the calibration data, of complete curves rather than considering the values at each calendar date separately, one is able to obtain posterior covariances between all points on the curve together with any other summary statistic desired. Blackwell and Buck (2008) illustrated how a curve might be constructed using such MCMC, although only for data with single-year ¹⁴C determinations and without some of the more complex uncertainty structures in the calendar age estimates that appear in the IntCal database. Here, we extend their approach so all data within the IntCal database can be used. In particular, we allow consideration of blocked data and a wider range of calendrical age modeling.

The layout of the article is as follows. In Section 2, we provide detail on the format of calibration data together with the random walk model for the curve itself. Section 3 presents the varying approaches to calendar age estimation used for the calibration data that need to be accurately modeled. Each of these results in differing structures for the uncertainty in the calendar age estimates, which we formalize mathematically. The Metropolis-within-Gibbs MCMC sampler used to update the curve paths and other unknown parameters is then set out in Section 4. Finally, in Section 5 we discuss some of the computational issues involved in the final production of the IntCal09 curve.

2. A BAYESIAN MODEL AND DATA-ADAPTIVE CURVE

As in Buck and Blackwell (2004), let us define, at any calendar date θ , the value of the calibration curve to be $\mu(\theta)$. While the physical processes that determine $\mu(\cdot)$ are not well understood, it is still feasible to express prior beliefs about its form through a statistical model. Following Buck and Blackwell (2004), one such possibility is a random walk where the changes from one year to the next follow independent Gaussian distributions with known mean β and, possibly unknown, variance r^2 . As such, for any θ ,

$$\mu(\theta + 1)|\mu(\theta) \sim N(\mu(\theta) + \beta, r^2) \quad (1)$$

Here, and later, the notation $|$ indicates conditioning such that $\mu(\theta + 1)|\mu(\theta)$ denotes “the value of $\mu(\theta + 1)$ given the value of $\mu(\theta)$.” However, we also have calibration data that give us more information on the value of the true curve at past points in time. Using these calibration data, we can update the prior to narrow our belief about μ and arrive at a posterior estimate of the curve. Intuitively, in calculating this posterior we aim to find which, out of all possible random walks defined by Equation 1, are plausible given the observed data.

Suppose we have a set \mathcal{I} of such calibration data. In the simplest case, an individual calibration datum will have known calendar date θ^* and a ¹⁴C determination, which is a noisy observation of $\mu(\theta^*)$. To complicate matters, however, a calibration datum often does not consist of a ¹⁴C determination from a single year but rather a determination based upon material from a block of several successive years. Also, it may not be possible to determine their calendar age absolutely but merely to estimate it.

We consider the general case where a calibration datum i consists of material arising from n_i successive years of metabolization and has an estimate T_i of the most recent year of deposition along with a ¹⁴C determination X_i for the complete calibration datum. Our complete calibration data set thus consists of paired values $(T_i, X_i)_{i \in \mathcal{I}}$ where for each i ,

$$T_i = \theta_i + \varepsilon_i \tag{2}$$

$$X_i = \sum_{j=0}^{n_i-1} \mu(\theta_i + j) + \eta_i$$

Here, and elsewhere in the article, θ_i is the true value for the most recent year of deposition. It is assumed that the errors on the ¹⁴C determinations n_i are independently distributed $N(0, \sigma_i^2)$, but the errors on the calendar dates ε_i can, and do, take more complex forms.

3. UNCERTAINTY IN CALENDAR AGES: DATA SETS AND ERROR STRUCTURES

3.1. Importance of Modeling

As highlighted in Equation 2, estimation of the calendar date θ_i is required for many calibration data. Such estimates are often associated with considerable uncertainty. These uncertainties are important, not least since the methods by which calendar dates are estimated are varied and can create significant dependence in the values obtained. Some data may have calendar dates that are known accurately, while others cannot be estimated so precisely. Furthermore, the calendar age of 1 datum in a data set may be related to all or some of the others in the same sequence. The presence and form of this error will have an effect on the calibration curve we produce. As such, its proper treatment should be part of any reliable curve-building methodology. Recognition of calendar dating error was one of the significant steps forward in calibration curve construction that was offered by the IntCal04 model (Buck and Blackwell 2004) and continues to be so in the updated MCMC methodology presented here.

Below, we set out the various types of calendar dating methodology that have been used to obtain the data in the IntCal09 calibration database and the structures we have chosen to formalize. Sections 3.2.1 to 3.2.4 articulate structures used in IntCal04, while Section 3.3 describes a new structure developed for the IntCal09 update.

3.2. Types of Calendar Date Error

3.2.1. No Error

For many, more recent, calibration data the calendar age θ is genuinely known (or any error is sufficiently small that it can be considered known). In such cases, we observe

$$T_i = \theta_i$$

Please note, as explained in Section 2, that if the sample represents a block of material over a period of n_i successive years, then θ_i is the most recent year of deposition. Within the current update curve, all the terrestrial data, having been dated via dendrochronology, are treated as having known calendar ages.

3.2.2. Independent Error

Alternatively, if a datum within a specific data set has an estimate T_i for its calendar age θ_i derived independently from all other data, we can model the relationship between the two as

$$T_i = \theta_i + \varepsilon_i$$

where ε_i are independent $N(0, \sigma_i^2)$. Note that the variance σ^2 may be, and generally is, different for each datum. Such independence occurs if the data have been U/Th dated as is the case for the corals from Barbados, Tahiti, Muroroa, and New Guinea (Bard et al. 1990); the corals from Huon Peninsula, Papua New Guinea (Edwards et al. 1993); the corals from Vanuatu and Papua New Guinea (Burr et al. 1998); the corals from Araki, Barbados, and Kirimati (Fairbanks et al. 2005); and the corals from Vanuatu and Papua New Guinea (Cutler et al. 2004).

In addition, we model the foraminifera from Cariaco Basin non-varved sediments (Hughen et al. 2006) and Iberian Margin non-varved sediments (Bard et al. 2004) as having independent calendar dates. This is known to be incorrect, but the calendar dating methodology used on these 2 data sets was complex and it has not possible to quantify exact dependence. Investigation of the effect of varying degrees of dependence in these 2 data sets showed minimal effect on the resulting curve.

3.2.3. Wiggle-Match Error

Wiggle-match error occurs when the relative chronology of all data within a particular data set \mathcal{J} is known (for example, due to layer counting) but it is not known where the data set as a whole fits in to the absolute chronology. Such “floating” data sets need to be tied in to other data where the calendar ages are known. This is usually done by matching “wiggles” within the ^{14}C determinations to other absolutely dated records.

With a pure wiggle-matching approach, the only calendar dating error is common to all data within the data set—the error made in making the single tie to other records. Once this single tie is made, the ages of all data are determined absolutely. This can be formalized by modeling, for any 2 observations $i, j \in \mathcal{J}$,

$$T_i = \theta_i + \Omega$$

$$T_j = \theta_j + \Omega$$

For completeness, we place a $N(0, \rho^2)$ prior on the value of this wiggle-match error Ω .

3.2.4. Varve Counting Error

Varve counting error relates to the error that accumulates as one counts annual layers of deposition, which are difficult to determine clearly. Consider an annually laminated object from which a series of data are to be taken. Furthermore, assume θ_{init} , the calendar date of the start of this layered sequence, is known and denoted by T_{init} with

$$T_{init} = \theta_{init}$$

Suppose, for the purpose of illustration, data are selected by trying to count back varves from the start of the laminated sequence until a certain number is achieved and then cutting a section from the sequence. The alternative possibility, where data are chosen before the number of varves between them is known, will not affect the structure that follows. Specifically, assume that the first calibration datum for inclusion in the database is found by trying to count back, say, 100 yr from the start. There may, however, have been errors in this counting and so the true calendar date of this datum will be $\theta_1 = \theta_{init} + 100 - \varepsilon_1$ where ε_1 represents how much we have over- or under-counted the real layers. Hence, the date we estimate

$$T_1 = T_{init} + 100 = \theta_1 + \varepsilon_1$$

Suppose then that our second datum is found by trying to count back another 30 yr during which we may make further counting errors ε_2 . The true date for this datum will thus be $\theta_2 = \theta_1 + 30 - \varepsilon_2$, whereas our date estimate will be

$$T_2 = \theta_2 + \varepsilon_1 + \varepsilon_2$$

In this way, our error accumulates as we count back along the layers within the object. At each new step, the accuracy of the calendar age measurement depends crucially on how accurate we were during the counting of each of the previous steps. If we miscounted on any earlier step, this error will accumulate into the calendar date accuracy of our new datum. We can represent such a structure by

$$T_i = \theta_i + \phi_i$$

where we define

$$\phi_i = \sum_{k=1}^i \varepsilon_k \text{ with } \varepsilon_k \sim N(0, \sigma_k^2)$$

3.2.5. Wiggle-Matching with Other Errors

Pure wiggle-matching in the form described in Section 3.2.3 is not feasible for several data sets. While an initial tie may be made to locate a data set \mathcal{J} within the absolute chronology, the relative chronology within that data set is often not known completely. If an independent dating method is used to find this relative chronology within the set, then we can model the error structure for data $i, j \in \mathcal{J}$ as

$$T_i = \theta_i + \Omega + v_i$$

$$T_j = \theta_j + \Omega + v_j$$

Here, the v_i and v_j relate to the additional error in finding the relative chronology and are independent from the wiggle-matching Ω .

Within our calibration data, the foraminifera from Cariaco Basin varved sediments (Hughen et al. 2004) are modeled as above with the initial use of a wiggle-match to locate the start of the data set in time followed by varve counting to determine the relative chronology within the set. As such, the additional error terms $v_i = \sum_{k=1}^i \varepsilon_k$, where each ε_k are independent $N(0, \sigma_k^2)$.

3.3. Ordering

In addition to the above error structures, several of the data sets are known to have ordering constraints within their calendar dates. In such a data set, although the precise dates of the constituent data are unknown, they must conform to a definite ordering. This may occur if, for example, the data come from sections of the same core that have been deposited through time. A datum further down the core must be older than a datum at the start of the core.

Within the IntCal09 calibration data, the foraminifera from Cariaco Basin varved sediments (Hughen et al. 2004), Cariaco Basin non-varved sediments (Hughen et al. 2006), and Iberian Margin non-varved sediments (Bard et al. 2004) are treated as possessing a known ordering.

4. A METROPOLIS-WITHIN-GIBBS SAMPLER

4.1. What Is a Gibbs Sampler?

There is not the space in this article to give a complete explanation of the Gibbs sampler, but for readers unfamiliar with the technique, we give a short review in the usual context of Bayesian inference (for more detail, see e.g. Casella and George 1992). Let us suppose that we have some, perhaps multidimensional, observed data \mathbf{Z} , which come from a distribution with a set of unknown parameters $\psi = (\psi_1, \dots, \psi_n)$. A Bayesian approach would place a prior $\pi(\psi)$ on these parameters, which represents the prior belief about their values. These beliefs would then be updated, in light of the data \mathbf{Z} , to obtain the posterior according to

$$p(\psi|\mathbf{Z}) = \frac{p(\mathbf{Z}|\psi)\pi(\psi)}{p(\mathbf{Z})}$$

Often, however, it is not possible to calculate the right-hand side of this equation directly since $p(\mathbf{Z}|\psi)$ is too complex. In such cases, the Gibbs sampler provides an alternative. It generates a sample $\psi^1, \dots, \psi^m \sim p(\psi|\mathbf{Z})$ indirectly using, often simpler, conditional distributions without ever having to calculate the full $p(\psi|\mathbf{Z})$. By taking a large enough sample, we can make inference on any characteristic of the posterior density we desire.

The sampler works by specifying an initial value for the set of parameters $\psi^0 = (\psi_1^0, \dots, \psi_n^0)$ and then iteratively creating a sequence of new values $\psi^1, \psi^2, \dots, \psi^k, \dots$. Each new value in the sequence is found by cycling through an algorithm whereby each individual variable ψ_i in turn is updated by sampling from its conditional distribution given the current values of all other variables:

- Start with $\psi^j = (\psi_1^j, \dots, \psi_n^j)$,
 - 1) Sample $\psi_1^{j+1} \sim p(\psi_1|\mathbf{Z}, \psi_2^j, \dots, \psi_n^j)$
 - 2) Sample $\psi_2^{j+1} \sim p(\psi_2|\mathbf{Z}, \psi_1^{j+1}, \psi_3^j, \dots, \psi_n^j)$
 - ⋮
 - i) Sample $\psi_i^{j+1} \sim p(\psi_i|\mathbf{Z}, \psi_1^{j+1}, \dots, \psi_{i-1}^{j+1}, \psi_{i+1}^j, \dots, \psi_n^j)$
 - ⋮
 - n) Sample $\psi_n^{j+1} \sim p(\psi_n|\mathbf{Z}, \psi_1^{j+1}, \dots, \psi_{n-1}^{j+1})$
- Set $\psi^{j+1} = (\psi_1^{j+1}, \dots, \psi_n^{j+1})$.

For some large k , under reasonably general conditions, the values from the k^{th} iteration ψ^k onwards will then be, effectively, samples from the posterior $p(\psi|\mathbf{Z})$.

We use a slight modification of the above approach known as Metropolis-within-Gibbs, which gives more flexibility in the iterative algorithm. This methodology is identical to that described above but for some of the individual ψ_i^{j+1} update steps, instead of sampling directly from the conditional distributions given all the other variables, we sample ψ_i^{j+1} from a different distribution, which is easier to calculate. This proposed value ψ_i^{j+1} is then either rejected or accepted with the appropriate probability.

In using such a Gibbs-based approach, the parameters of interest can include the value of the curve at several preselected calendar dates. As such, each time we iterate through the sampler we can find posterior values of the curve at a set of points simultaneously. We are not required to consider the curve independently at each point but instead can determine characteristics about how values at multiple calendar dates vary jointly.

4.2. Variables Relating to the Data

Our calibration data provide us with a set of observed variables:

- $\mathcal{T} = \{T_i\}_{i \in \mathcal{I}}$ —the estimated calendar dates of the data. If the calibration datum is blocked, this value refers to the most recent (start) date of the block.
- $\mathcal{X} = \{X_i\}_{i \in \mathcal{I}}$ —the estimated ^{14}C determinations of the (possibly blocked) data.

We also have the following variables which we wish to learn about and will be updated in turn within our sampler:

- $\Theta^{\text{Data}} = \{\theta_i\}_{i \in \mathcal{I}}$ —the true calendar dates of the data. Again, if the calibration datum is blocked, this will refer to the most recent date of the block.
- $\mathcal{M}^{\text{Data}} = \{\mu_i\}_{i \in \mathcal{I}}$ —the value of the calibration curve at the dates Θ^{Data} .
- Ω —wobble-match error in a data set. We only present the case of a single wobble-match error that occurs in a varve counted data set as required for the IntCal09 curve. The extension to multiple independent wobble-match errors possibly combined with other error formats is feasible within the MCMC framework.
- r^2 —variance of the Gaussian random walk.

In addition, we must also specify the drift of the Gaussian random walk β in Equation 1. While in principle this can also be treated as an unknown random variable, in this update it was decided, as in the IntCal04 curve, to treat it as fixed with a value of 1. As explained in the IntCal04 methodology (Buck and Blackwell 2004), this is based upon the expectation that the value of the calibration curve changes by approximately 1 ^{14}C year per calendar year. The authors of that paper also found that the resultant curve is very insensitive to the specific value of β used, a more precise value based upon current knowledge of ^{14}C decay rates making very little difference to the curve.

4.3. Addition of a Grid

One of the theoretical advantages of MCMC is that calibration of a new datum with ^{14}C determination X_0 does not require evaluation of the curve on a fixed grid. Indeed, by reducing the curve to a regularly spaced series of posterior means and variances, a huge amount of information is lost. Blackwell and Buck (2008) showed this can have a noticeable effect on calibration. A variety of alternative approaches that do make full use of the additional information provided by complete realizations of the curve are possible. See Blackwell and Buck (2008) for further details.

In practice, however, the ^{14}C calibration programs currently in use are not able to utilize such extra information and do require the curve estimates take the form of pointwise means and variances on a

fine-scale, regular grid. In order to produce output in the form required, we introduce a grid into our sampler, which is updated along with the other variables.

Let us suppose that the calendar dates chosen for our output grid are $\Theta^{Grid} = (\theta_1^{Grid}, \dots, \theta_m^{Grid})$.

The values of the calibration curve at these dates will then be $\mathcal{M}^{Grid} = (\mu_1^{Grid}, \dots, \mu_m^{Grid})$. Our MCMC sampler estimates the curve at both these known grid dates and the unknown dates of the calibration data, i.e.

$$\Theta = (\Theta^{Data}, \Theta^{Grid})$$

where the values will be

$$\mathcal{M} = (\mathcal{M}^{Data}, \mathcal{M}^{Grid})$$

Note that once a grid is added, the observed data, i.e. \mathcal{X}^{Data} and \mathcal{T}^{Data} , will be of a smaller dimension than Θ and \mathcal{M} .

4.4. An Iteration

One iteration of our Gibbs sampler consists of a complete cycle through the following set of steps which are each explained further in Sections 4.5, 4.6, and 4.7, respectively:

- Step 1. **Update calibration curve \mathcal{M} and Θ given other variables i.e.** $\mathcal{M}, \theta | \mathcal{T}, \mathcal{X}, \Omega, r^2$. Performed pairwise using a Metropolis Hastings algorithm where, for each pair j , a new set of values (μ_j, θ_j) is proposed. In-depth explanation of this step can be found in Blackwell and Buck (2008). We use the same algorithm, although minor modifications, explained in Section 4.5, are required to account for the blocking that is observed within our calibration data. Small changes are also required in the proposal of a new calendar date θ_j if the error relating to that date includes varve counting.
- Step 2. **Update wiggle-match error Ω given other variables i.e.** $\Omega | \mathcal{M}, \Theta, \mathcal{T}, \mathcal{X}, r^2$. For any particular wiggle-matched data set, we can update the value of that wiggle-match by considering the difference between the observed calendar dates \mathcal{T} and the true calendar dates Θ . In the general case, this will have a multivariate normal distribution. However, in the case where the wiggle-match is combined with varve counting, as in the IntCal09 database, it is sufficient to consider the difference between recorded and true calendar date for the first sample, which allows a simplification to be made.
- Step 3. **Update walk variance r^2 given other variables i.e.** $r^2 | \mathcal{M}, \Theta, \mathcal{T}, \mathcal{X}, \Omega$. The increment in the calibration curve between any 2 neighboring calendar dates θ_i and θ_j will have a likelihood dependent upon the random walk variance. Conditioning on the observed values of these increments allows sampling from the posterior for r^2 .

Each iteration of our MCMC will provide a value for the true calendar dates θ_i^{Data} of each calibration datum and the calibration curve at that date $\mu_i^{Data} = \mu(\theta_i^{Data})$. These can be considered as draws from the posterior of the curve on a random rather than fixed grid. In addition, we will have data-adaptive estimates for the wiggle-match error and random walk variance. Finally, we have values of the curve at the pre-specified fixed dates θ^{Grid} , which can be reduced to the means and variances required.

4.5. Updating the Calibration Curve (μ_j, θ_j)

4.5.1. Generalized Metropolis-Hastings Step

Referring to Blackwell and Buck (2008) for details, the general update of a pair (μ_j, θ_j) takes the following form:

- a) Propose a new θ'_j from a symmetric distribution centered at current value θ_j . A discretized normal distribution is used since the calendar dates of calibration data are reported on an annual scale.
- b) Accept with probability $p = \min \{HR, 1\}$ where the Hastings ratio

$$HR = \frac{p(\mathcal{X}|\theta'_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)p(\theta'_j|\mathcal{T}, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r)}{p(\mathcal{X}|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)p(\theta_j|\mathcal{T}, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r)}$$

and $\mathcal{Y}_{(i)}$ denotes a set \mathcal{Y} with element y_i removed.

- c) If accepted, sample new μ_j from distribution $\mu_j|\theta'_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \mathcal{X}, \Omega, r, \beta$.

This is modified, however, in the case of pairs where the calendar date θ_j is known absolutely, i.e. the grid values and terrestrial data. For such pairs, steps a) and b) are unnecessary. Instead, we simply automatically accept a new value for μ_j sampled from $\mu_j|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \mathcal{X}, \Omega, r, \beta$.

4.5.2. The Case of Blocking

The introduction of blocking within the data does not modify the overall approach to the sampler but does affect the computational simplifications that one could otherwise make. Consider the initial term in both the numerator and denominator of the Hastings ratio. In the presence of blocking, it is not possible to reduce the likelihood of the complete data from $p(\mathcal{X}|\theta'_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)$ to that of the single determination x_j . While initially this appears to make the problem computationally prohibitive, savings can still be found. Due to the independent increment property of the Gaussian random walk, the likelihood can be split into a product of independent multivariate likelihoods only one of which will depend upon θ'_j . Let \mathcal{X}'_j denote those ¹⁴C determinations x_i whose likelihood is affected if calibration datum j has a true calendar date of θ'_j . Precisely which these are will depend upon the overlap of the relative data blocks into the region containing θ'_j . Specifically, we can write

$$p(\mathcal{X}|\theta'_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta) = p(\mathcal{X}'_j|\theta'_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta) \times \prod_{\mathcal{K} \in \mathcal{P}'} p(\{x_l\}_{l \in \mathcal{K}}|\mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)$$

where \mathcal{P}' denotes the partition of all the observations, excluding \mathcal{X}'_j , into the independent multivariate likelihoods if calibration datum j , with determination X_j , has a true calendar age of θ'_j and the value of the curve is known at $\Theta_{(j)}$.

Similarly, the denominator $p(\mathcal{X}|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)$ can be decomposed. Let \mathcal{X}_j denote those ¹⁴C determinations x_i whose likelihood is affected if calibration datum j has a true calendar date of θ_j . Here,

$$p(\mathcal{X}|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta) = p(\mathcal{X}_j|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta) \times \prod_{\mathcal{K} \in \mathcal{P}} p(\{x_l\}_{l \in \mathcal{K}}|\mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)$$

where \mathcal{P} denotes the partition of all the observations, excluding \mathcal{X}_j , into the independent multivariate likelihoods if calibration datum j , with determination X_j , has a true calendar age of θ_j and the value of the curve is known at $\Theta_{(j)}$. Again, each term on the right-hand side has the form of a multivariate normal.

Many of the terms subsumed into the 2 product terms above are identical and so can be canceled within the Hastings ratio. In fact, the only terms that do not are those relating to the union of the ^{14}C determinations $\mathcal{X}'_j \cup \mathcal{X}_j$ so that

$$\frac{p(\mathcal{X}|\theta'_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)}{p(\mathcal{X}|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)} = \frac{p(\mathcal{X}'_j \cup \mathcal{X}_j|\theta'_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)}{p(\mathcal{X}'_j \cup \mathcal{X}_j|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r, \beta)}$$

This greatly reduces the calculation required.

Analogously, sampling a new value for μ_j is altered by the presence of blocking with the distribution $\mu_j|\theta_j, \mathcal{M}_{(j)}, \Theta_{(j)}, \mathcal{X}, \Omega, r, \beta$ affected by multiple ^{14}C determinations rather than the single x_j . As above, the likelihood of the x s can be decomposed into independent multivariate distributions, with most not depending upon the value of μ_j , before applying Bayes' theorem as in Blackwell and Buck (2008).

4.5.3. Hastings Ratio and Calendar Date Uncertainty

Let us now turn to the other term in the acceptance probability, $p(\theta_j|\mathcal{T}, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r)$. It is here that the dependence structure within the calendar date error plays a crucial role as it will significantly affect this value and hence the curve. We describe the 3 possibilities of independent, wiggle-matched and varved error set out in Sections 3.2.2, 3.2.3, and 3.2.4, respectively.

If the calendar dating process has only independent error, $p(\theta_j|\mathcal{T}, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r)$ reduces to the probability given the observed calendar date of that individual calibration datum, i.e. $p(\theta_j|t_j) \sim \mathcal{N}(0, \sigma_j^2)$. This is the case for the majority of the data used in the calibration curve.

On the other hand, for data that have been wiggle-matched, the likelihood of θ_j will depend not just upon t_j but on every data point in that set. This suggests complexities in the calculation of $p(\theta'_j|\mathcal{T}, \mathcal{M}_{(j)}, \Theta_{(j)}, \Omega, r)$. However, the value of the wiggle-match Ω completely encapsulates all wiggle-match dependence and, conditional on its value, we need not concern ourselves with what form this structure takes. As a consequence, we can reduce our consideration to that of incorporating any additional error resulting from dating processes after the wiggle-match, for example, that caused by later varve counting.

In any varve-counted data, the simplification used for the independent case is not possible since error in dating 1 calibration datum will affect the observed calendar dates for other data in that set. We must consider the joint likelihood of the data that have been varve counted. This requires further calculation.

4.5.4 Acceptance of θ' in Presence of Varve Counting Error Φ

Suppose we have a sequence of data points that are varve counted, i.e.

$$T_i = \theta_i + \phi_i$$

where $\phi_i = \sum_{k=1}^i \varepsilon_k$ and each ε_k is independent $N(0, \sigma_j^2)$. Having proposed a new value θ' for a calibration datum i with the above structure, in accordance with the Metropolis step above, to determine acceptance we are required to calculate

$$p(\theta_j | \mathcal{T}, \mathcal{M}_{(j)}, \Theta_{(j)}, r) = p(\theta_j | \mathcal{T}, \Theta_{(j)})$$

Given \mathcal{T} and $\Theta_{(j)}$, we know the errors ϕ_{j-1} and ϕ_{j+1} and hence

$$\begin{aligned} p(\theta_j | \mathcal{T}, \Theta_{(j)}) &= p(\varepsilon_j = T_j - \theta_j - \phi_{j-1} | \varepsilon_j + \varepsilon_{j+1} = \phi_{j+1} - \phi_{j-1}) \\ &\propto p(\varepsilon_j = T_j - \theta_j - \phi_{j-1}) p(\varepsilon_{j+1} = \phi_{j+1} + \theta_j - T_j) \end{aligned}$$

As indicated in Section 3.2.4, for the implementation in IntCal09, varve counting occurs in the presence of wiggle-matching and so $T_i = \theta_i + \phi_i + \Omega$. However, given the value of this wiggle-match parameter, the structure simplifies to that described above simply by replacing each θ_i with $\theta_i^* = \theta_i + \Omega$

4.6. Updating Wiggle-Match Error Ω

Data i and j are related by a wiggle-match error Ω if the error structure of the calendar dates satisfies the relation

$$T_i = \theta_i + \Omega + \phi_i$$

$$T_j = \theta_j + \Omega + \phi_j$$

For the purposes of IntCal09, we need only consider the case when the additional error ϕ takes the form of varve counting and hence, in this paper, we only present the update step in this instance, which is found to have an attractive simplification. The more general case where the ϕ s are assumed to have a multivariate normal distribution with mean 0 and covariance Σ can also be solved by considering the sufficient statistic $\mathcal{T} - \Theta$.

In the case that ϕ has the varve counting structure, we can decompose the likelihood as

$$\begin{aligned} p(\mathcal{T} | \Theta, \Omega) &= p(\varepsilon_1 = T_1 - \theta_1 - \Omega) p(\varepsilon_2 = (T_2 - T_1) - (\theta_2 - \theta_1)) \\ &= \exp\left[-\frac{1}{2} \frac{(T_1 - \theta_1 - \Omega)^2}{\sigma_1^2}\right] h(\mathcal{T}, \Theta) \end{aligned}$$

for some known function $h(\cdot)$. Using a $N(0, \rho^2)$ prior for the wiggle-match Ω and applying Bayes' theorem, we find

$$\begin{aligned} p(\Omega | \mathcal{T}, \Theta) &\propto \pi(\Omega) p(\mathcal{T} | \Theta, \Omega) \\ &\propto \exp\left[-\frac{1}{2} \left\{ \frac{\Omega^2}{\rho^2} + \frac{(T_1 - \theta_1 - \Omega)^2}{\sigma_1^2} \right\}\right] \end{aligned}$$

In other words, the posterior distribution $\Omega | \mathcal{T}, \Theta$ depends only upon the error in the first observation of the varve-counted sequence. Hence to update Ω given the other variables, we sample from $\Omega | \mathcal{T}, \Theta \sim N(\mu', \sigma'^2)$ where

$$\sigma'^2 = \frac{1}{\rho^{-2} + \sigma_1^{-2}} \quad \mu' = \sigma'^2 \frac{T_1 - \theta_1}{\sigma_1^2}$$

4.7. Updating r^2 —Random Walk Variance

At each step of the algorithm, we sample particular points $(\mu_i, \theta_i)_{i \in \mathcal{I}}$ on the random walk's path. Let us suppose that these have been ordered so that

$$\theta_i < \theta_j \quad \text{if and only if} \quad i < j.$$

Any identical calendar date values can be ignored since the values of μ at these dates will also be identical. To update the random walk variance, it is sufficient to just consider the changes in the sample curve values $\mu_j - \mu_{j-1}$ between these dates $(\theta_j - \theta_{j-1})$. Specifically,

$$(\mu_j - \mu_{j-1})_{2 \leq j \leq n} \sim N(\beta(\theta_j - \theta_{j-1}), r^2(\theta_j - \theta_{j-1}))$$

To update the variance, we place an Inverse Gamma prior on $r^2 \sim IG(\delta, \gamma)$ so, having observed

$$\left(\frac{\mu_j - \mu_{j-1}}{\sqrt{\theta_j - \theta_{j-1}}} \right)_{2 \leq j \leq n} \sim N(\beta \sqrt{\theta_j - \theta_{j-1}}, r^2)$$

we have a posterior $r^2 | \mathcal{M}, \Theta \sim IG(\delta^*, \gamma^*)$ where

$$\delta^* = \delta + \frac{n-1}{2} \quad \text{and} \quad \gamma^* = \gamma + \sum_{i=2}^n \frac{[(\mu_j - \mu_{j-1}) - \beta(\theta_j - \theta_{j-1})]^2}{2(\theta_j - \theta_{j-1})}$$

5. ADDITIONAL CONSIDERATIONS

5.1. Choice of Grid

For the purposes of IntCal09, a grid was selected with intervals of 10 yr for the range 12–15 cal kyr BP, 20 yr for 15–25 cal kyr BP, 50 yr for 25–40 cal kyr BP, and 100 yr for 40–50 cal kyr BP. The mean and variance for the posterior calibration curve was evaluated pointwise at each of dates to produce the final curve and confidence bands.

5.2. Ordering of Observations

Three of the calibration data sets (Bard et al. 2004; Hughen et al. 2004, 2006) possessed known ordering among their calendar dates. Such ordering was maintained by automatic rejection of any value θ'_j , as proposed in Step a) of the sampler, which violated these constraints. While this may lead to a higher rejection rate within the Metropolis step than optimal, it was not considered a major factor in computational speed.

5.3. Discretizing the Calendar Dates

All the data we received had calendar dates recorded to the nearest year. Indeed, many calibration data were calendar dated via counting layers that were deposited annually. It was felt appropriate that the methodology should work on the same scale and not consider dates on a finer timescale. Accordingly, a discretized normal was used for the proposal distribution of θ' in Step a) of our sampler.

5.4. Data with No Calendar Age Uncertainty

The calendar ages \mathcal{T} for tree-ring calibration data have no error associated with them and so it is not necessary to propose new values for their true calendar dates Θ . In fact, since we are not directly interested in the value of the ^{14}C curve at these dates, there is no need to perform Step a) of the sampler for any datum i with no calendar age uncertainty. We must, however, keep the ^{14}C determination X_i with associated date θ_i in the sampler for when we update all other variables.

5.5. Splitting the Curve

For the IntCal09 curve update, it was decided that the estimate should not change for the period 0–12,000 cal yr BP from that published in 2004 (Reimer et al. 2004). As a result, the statistical work reported here focused on providing estimates for IntCal09 in the period 12,000–50,000 cal yr BP.

Between these dates, the IntCal09 curve was created with 2 distinct sections: the first, from 12,000–12,540 cal yr BP, based solely on tree-ring data while the later section from 12,540–50,000 was constructed using only data from a marine environment with no tree rings present. Exploratory analysis indicated significant differences between the features of these 2 sources of data and suggested it was inappropriate to model them identically. The curve as evidenced by the data from a marine environment had considerably increased variability from that suggested by the tree-ring data. It was therefore decided that curves should be created for each section independently subject to the condition of continuity at the juncture. This allowed separate values of the random walk variance r^2 to be selected for regions based upon tree-ring data and those based upon marine data, which would better account for the 2 distinct data sources.

For estimating the 12,000–12,540 cal yr BP, tree-ring-based portion of the curve, the random walk variance was fixed to be 8^2 for the reasons set out in Buck and Blackwell (2004). Conversely, in the older, marine-data based portion of the curve r^2 was allowed to vary and adapt to these marine calibration data. In this section, a fairly uninformative prior was chosen for the variance r^2 centered around 8^2 . The posterior obtained for r^2 had a mean around 15.5^2 .

5.6. Convergence

A quantitative assessment of convergence was difficult due to the large number of variables that are updated in the sampler. Instead, we used a heuristic approach whereby, having let the sampler run for a burn-in period, independent subsamples of the later output were selected and the corresponding calibration curve estimate for each was plotted. No visible differences were observed, thus indicating that the Markov chain had reached and adequately explored its equilibrium distribution.

The 12,540–50,000 cal yr BP (marine-data based) segment of the IntCal09 curve was estimated using an MCMC run that had a burn-in period of 100,000 iterations with between 150,000 and 500,000 further iterations being used to produce the curve dependent upon the region of interest. For the 12,000–12,540 cal yr BP portion of the curve (for which the data structures are simpler), the first 50,000 iterations were discarded as a burn-in period with a further 100,000 iterations used to produce the curve itself.

CONCLUSION

In this paper, we have presented a flexible, yet rigorous framework for the construction of ^{14}C calibration curves given suitable quality calibration data. Our methodology builds upon the random walk approach of Buck and Blackwell (2004) that was used for IntCal04 (Reimer et al. 2004) and

fully implements the MCMC curve construction ideas initially presented in Blackwell and Buck (2008). It uses these MCMC techniques to sample plausible curves based upon a random walk prior updated in light of the calibration data themselves. It offers advantages over alternative construction techniques in the improved flexibility to incorporate the complexities inherent in such data, for example, blocking and the various uncertainty structures in the calendrical age modeling. Furthermore, it allows the user to investigate the values of the curve at multiple calendar dates simultaneously and hence determine information about covariance between points which can, in principle, be used for later calibration. Our approach is used in the creation of the current IntCal09 curve (Reimer et al., this issue).

ACKNOWLEDGMENTS

The authors would like to thank both M Scott and S Pérez-Elizalde for their helpful comments on a draft of this paper, which have led to significant improvements.

REFERENCES

- Bard E, Hamelin B, Fairbanks RG, Zindler A. 1990. Calibration of the ^{14}C timescale over the past 30,000 years using mass spectrometric U-Th ages from Barbados corals. *Nature* 345(6274):405–10.
- Bard E, Rostek F, Ménot-Combes G. 2004. A better radiocarbon clock. *Science* 303(5655):178–9.
- Blackwell PG, Buck CE. 2008. Estimating radiocarbon calibration curves. *Bayesian Analysis* 3(2):225–68.
- Buck CE, Blackwell PG. 2004. Formal statistical models for estimating radiocarbon calibration curves. *Radiocarbon* 46(3):1093–102.
- Burr GS, Beck JW, Taylor FW, Recy J, Edwards RL, Cabioch G, Corrège T, Donahue DJ, O'Malley JM. 1998. A high-resolution radiocarbon calibration between 11,700 and 12,400 calendar years BP derived from ^{230}Th ages of corals from Espiritu Santo Island, Vanuatu. *Radiocarbon* 40(3):1093–105.
- Casella G, George EI. 1992. Explaining the Gibbs sampler. *The American Statistician* 46(3):167–74.
- Cutler KB, Gray SC, Burr GS, Edwards RL, Taylor FW, Cabioch G, Beck JW, Cheng H, Moore J. 2004. Radiocarbon calibration to 50 kyr BP with paired ^{14}C and ^{230}Th dating of corals from Vanuatu and Papua New Guinea. *Radiocarbon* 46(3):1127–60.
- Edwards RL, Beck JW, Burr GS, Donahue DJ, Chappell JMA, Druffel ERM, Taylor FW. 1993. A large drop in atmospheric $^{14}\text{C}/^{12}\text{C}$ and reduced melting in the Younger Dryas with ^{230}Th ages of corals. *Science* 260(5110):962–8.
- Fairbanks RG, Mortlock RA, Chiu T-C, Cao L, Kaplan A, Guilderson TP, Fairbanks TW, Bloom A, Grootes PM, Nadeau M-J. 2005. Radiocarbon calibration curve spanning 0 to 50,000 years BP based on paired $^{230}\text{Th}/^{234}\text{U}/^{238}\text{U}$ and ^{14}C dates on pristine corals. *Quaternary Science Reviews* 24(16–17):1781–96.
- Hughen KA, Southon JR, Bertrand CJH, Frantz B, Zermeno P. 2004. Cariaco Basin calibration update: revisions to calendar and ^{14}C chronologies for core PL07-58PC. *Radiocarbon* 46(3):1161–87.
- Hughen K, Southon J, Bertrand C, Turnbull J. 2006. Marine-derived ^{14}C calibration and activity record for the past 50,000 years updated from the Cariaco Basin. *Quaternary Science Reviews* 25(3–4):3216–27.
- Reimer PJ, Baillie MGL, Bard E, Bayliss A, Beck JW, Bertrand CJH, Blackwell PG, Buck CE, Burr GS, Cutler KB, Damon PE, Edwards RL, Fairbanks RG, Friedrich M, Guilderson TP, Hogg AG, Hughen KA, Kromer B, McCormac G, Manning S, Bronk Ramsey C, Reimer RW, Remmele S, Southon JR, Stuiver M, Talamo S, Taylor FW, van der Plicht J, Weyhenmeyer CE. 2004. IntCal04 terrestrial radiocarbon age calibration, 0–26 cal kyr BP. *Radiocarbon* 46(3):1029–58.
- Reimer PJ, Baillie MGL, Bard E, Bayliss A, Beck JW, Blackwell PG, Bronk Ramsey C, Buck CE, Burr GS, Edwards RL, Friedrich M, Grootes PM, Guilderson TP, Hajdas I, Heaton TJ, Hogg AG, Hughen KA, Kaiser KF, Kromer B, McCormac FG, Manning SW, Reimer RW, Richards DA, Southon JR, Talamo S, Turney CSM, van der Plicht J, Weyhenmeyer CE. IntCal09 and Marine09 radiocarbon age calibration curves, 0–50,000 years cal BP. *Radiocarbon*, this issue.
- Weninger B, Jöris O. 2008. A ^{14}C age calibration curve for the last 60 ka: the Greenland-HULU U-Th timescale and its impact on understanding the Middle to Upper Paleolithic transition in Western Eurasia. *Journal of Human Evolution* 55(5):772–81.