Radiocarbon (2026), **68**, pp. 21–31
doi:10.1017/RDC.2024.56

Selected Papers from the 2nd Latin American Radiocarbon Conference, Mexico City, 4–8 Sept. 2023

CONFERENCE PAPER

Modelling chronologically ordered radiocarbon dates in R

Maarten Blaauw[1], Marco Aquino-López[2] and J. Andrés Christen[3]

[1 14]CHRONO Centre for Climate, the Environment and Chronology, School of Natural and Built Environment, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, United Kingdom, [2]Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, United Kingdom and [3]Centro de Investigación en Matemáticas CIMAT, Guanajuato 36023, Guanajuato, Mexico
Corresponding author: Maarten Blaauw; Email: maarten.blaauw@qub.ac.uk

Received: 22 November 2023; **Revised:** 12 January 2024; **Accepted:** 29 February 2024; **First published online:** 18 September 2024

Keywords: Bayesian age-models; radiocarbon calibration; chronological ordering; stratigraphy; wiggle-match dating

Abstract
Studies with multiple radiocarbon dates often contain useful information on the relative locations of the dated levels. Such information can be used to obtain robust, integrated site chronologies, with at times more precise ages than those of the individual dates, where outliers can be identified and downweighted, and where the ages of any undated levels can also be estimated. Examples include trees with radiocarbon dates separated by exactly known amounts of yearly tree-rings, or sedimentary sites where ages further down the stratigraphy can be assumed to be older than ages further up. Here we present *coffee*, an R package for Bayesian models that apply *c*hronological *o*rdering *f*or *f*ossils and *e*nvironmental *e*vents. *Coffee* runs natively within the popular and versatile R environment, with no need for importing or exporting data or code from other programs, and works with plain-text input files that are relatively easy to read and write. It thus provides a new, transparent and adaptable educational and research platform designed to make chronology building more accessible.

Introduction

Chronology-building is an essential part of much research in paleoecology, paleoclimate, geology and archeology. Although some sites benefit from many dates, other studies have to work with relatively few dates. Additionally, at times reversals or unexpectedly large dating scatter are encountered, and often there is a need to combine the dates with additional information in order to arrive at reliable, informed chronologies. Over the past three decades, Bayesian methods which formally combine existing knowledge with new data (Bronk Ramsey, 1995; Buck et al., 1996) have become very popular tools to produce such chronologies. Although these tools are mostly used by the radiocarbon community, Bayesian methods can also be applied to other types of dates (e.g., De Vleeschouwer and Parnell, 2014).

Current standalone software tools to build Bayesian chronologies include *OxCal*[1] (Bronk Ramsey, 1995, 2001), *BCal*[2] (Buck et al., 1999), and *ChronoModel*[3] (Lanos and Philippe, 2018). Besides these stand-alone tools, a range of Bayesian chronology-building packages is available within the Open Source statistical software environment *R*, including *rcarbon* (Crema and Bevan, 2021) to analyse the densities of dates over time, *Bchron* (Haslett and Parnell, 2008) and *Bacon* (Blaauw and Christen, 2011)

[1] https://c14.arch.ox.ac.uk/oxcal.html
[2] https://bcal.shef.ac.uk
[3] https://chronomodel.com/

© The Author(s), 2024. Published by Cambridge University Press on behalf of University of Arizona. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

for age-depth modelling of cores with dated depths, and *oxcAAR*[4] which provides an *R* interface to *OxCal*.

R has become a very popular data–handling toolkit in many research communities. Its approach to providing all code freely promises to be beneficial to science, because it enables higher transparency and reproducibility, a better long-term usability and more flexibility for analysis than closed-source and/or standalone alternatives. Given that around 20,000 user-contributed packages are currently available on *R*'s CRAN repository, most or all of a research project's data analysis can now be performed within a single, transparent and shareable environment. It is in this spirit that we present a new *R* package that provides two methods to combine dates with stratigraphical information. The *R* package is called *coffee* – *c*hronological *o*rdering *f*or *f*ossils and *e*nvironmental *e*vents.

Methods

Dates

Samples can be dated using historical information (e.g., when fossils are found of a species that was introduced to a region at a known historical time), or using radiometric dating where the ratio of certain radioactive isotopes is measured, with this ratio changing in a predictable manner over time. If the calendar age θ_i (in cal BP, calendar years before AD 1950) of a sample x_i is estimated by a measurement y_i with a degree of uncertainty σ_i, it is often assumed that this measurement is distributed normally around the (unknown) true age,

$$y_i \sim N(\theta_i, \sigma_i). \tag{1}$$

If the age of a sample is known exactly, then $y_i = \theta_i$, while in other cases we know that an event took place before a specific date, $\theta_i > \theta_{i-1}$, or after it, $\theta_i < \theta_{i+1}$.

For radiocarbon dates, we have a good understanding of how ^{14}C ages relate to cal BP years through the different calibration curves such as IntCal20 (Reimer et al., 2020), Marine20 (Heaton et al., 2020), SHCal20 (Hogg et al., 2020) or postbomb curves (Hua et al., 2022). Calibration curves essentially consist of three columns with the first a range of cal BP ages θ, followed by their associated ^{14}C ages $\mu(\cdot)$ and finally by their uncertainties $\sigma(\cdot)$. Then, radiocarbon ages are assumed to be normally distributed given the calibration curve's ^{14}C age belonging to the calendar year of interest:

$$y_i \sim N(\mu(\theta_i), \omega(\theta_i)), \tag{2}$$

where $\omega(\theta_i) = \sqrt{\sigma^2(\theta_i) + \sigma_i^2}$ combines the uncertainties of both the calibration curve and of the date (e.g., Christen and Litton, 1995).

Outliers

Analysis of material dated through multiple radiocarbon measurements often shows a scatter beyond what the individual lab errors would suggest (Christen and Pérez, 2009; Scott et al., 2018; Scott, 2023). This overdispersion can be modelled in several ways (Christen, 1994a,b; Bronk Ramsey, 2009b). The most common approach to modelling outliers is to enable a shift of the measured mean:

$$y_i' \sim N(\mu(\theta_i) + \psi_i \delta_i, \omega(\theta_i)), \tag{3}$$

where $\psi_i = (0, 1)$ is a 'flag' and $\delta_i = N\left(0, \beta\sigma_i^2\right)$ is a shift on the radiocarbon scale (often, β is set to 2; shifts on the cal BP scale can also be modelled). If a date is not outlying, ψ_i is set to 0 and the date is

[4] https://CRAN.R-project.org/package=oxcAAR

treated as usual. If the date is an outlier, ψ_i is set to 1 and a shift of its age is modelled. This prior outlier probability is often set to 5% (e.g., Blaauw and Christen, 2005; Haslett and Parnell, 2008; Bronk Ramsey, 2009b) although Haslett and Parnell (2008) model two types of outliers (moderate outliers with a 5% prior probability and more extreme outliers with a 1% prior probability). This 'flag and shift' approach requires several additional parameters to be modelled for each date in a site. Posterior outlier probabilities are then calculated as the proportion of times during modelling that the outlier flag had to be set to 1. It is thus usually not necessary or recommendable to manually remove outliers - they are downweighted automatically based on the other dates and the modelling constraints.

Another approach to overdispersion is to treat the uncertainty of each date not as a known single value but as a distribution (Christen and Pérez, 2009). For example, if a lab reports a date to have an uncertainty of, say, 20 years, the true uncertainty could be lower or higher than the reported value. By multiplying the lab error by a gamma distribution, in effect we are replacing the normal distribution of Eq. (2) with a Student-t distribution, with parameters $t.a$ and $t.b$. Depending on the parameter values used, the latter distribution will appear very similar to the normal distribution, but will have wider tails. *Bacon* (Blaauw and Christen, 2011) uses the Student-t distribution to model dates, as it makes the age-model robust to dating scatter without the need to model additional parameters (see also Blaauw et al., 2018). We will be using the Student-t distribution in the approach developed below. Also with this approach, usually there is no need to manually remove outlying dates.

Stratigraphy

The dates can be integrated with any information regarding the time spans between (dated or undated) sample positions. For example, if we have radiocarbon samples from a tree with n reliable yearly rings but no master dendrochronology to align it with, we can still precisely count the exact gaps, or time spans, between specific (sections of) tree-rings, and combine these with radiocarbon–dated rings to produce a chronology. If say 10 annual tree-rings were deposited between the older sample 1 and the younger sample 2, then we know that whatever the known or unknown age is of sample 1, it must be 10 years older than that of sample 2. In statistical terms (e.g., Christen, 1994b; Christen and Litton, 1995), this can be written as there being m dated positions in a tree, with each position x_i having an associated calendar age θ_i separated by known integer time-spans from other rings within the tree (here we start counting from the youngest, outermost ring):

$$\delta_i = \theta_i - \theta_{i-1} \in \mathbb{N} \tag{4}$$

Some rings will have one or more radiocarbon dates and some none. We will leave this implicit in the definition of the likelihood in Eq. (7) to avoid a more cumbersome notation. Other positions of interest, while not radiocarbon dated, may still require an age estimate. Given θ_0 (the age of the outermost ring), the age of any ring can then be calculated by adding up the gaps:

$$\theta_i = \theta_0 + \sum_{j=1}^{i} \delta_j. \tag{5}$$

In sites with less exact information on the time spans between levels, but where we can still safely assume a chronological ordering of the different levels based on the site's stratigraphy (e.g., Harris, 1989), often all we know is that between two stratigraphically separated levels x_i and x_{i-1}, the time span $\theta_i - \theta_{i-1} = \delta_i > 0$. Some layers will not have such information and then dates within said layer cannot safely be assumed to be in chronological order (we will call these layers 'blocks'). In some cases however, there could be additional information, say, if two samples are separated by a soil which is known from other sources to take a certain amount of time to form. These stratigraphical gaps could be modelled using a range of distributions, e.g., uniform: $\delta_i \sim U(a_i, b_i)$ (see Bronk Ramsey, 2000; Steier and Rom, 2000), normal: $\delta_i \sim N(\mu_i, \sigma_i)$, gamma: $\delta_i \sim \Gamma(\mu_i, \sigma_i)$, or even an exact gap: $\delta_i = \mu_i$. (For

sites such as bogs, lakes or marine deposits which can reliably be assumed to have accumulated approximately continuously over time, *age–depth* models such as *BChron* (Haslett and Parnell, 2008), *Bacon* (Blaauw and Christen, 2011) or *OxCal's P_Sequence* (Bronk Ramsey, 2009b) would be more appropriate.)

Combining dates and stratigraphy

Trees with their exactly known time-gaps between ^{14}C-dated rings are relatively easy to model. Combining Equations (2) and (5), we can calculate the likelihood of a calendar age for each ^{14}C-dated tree-ring $y_i \sim N\left(\mu(\theta_i), \omega_i^2\right)$ where, as before, $\theta_i = \theta_0 + \sum_{j=1}^{i} \delta_j$ and $\omega_i^2 = \sqrt{\sigma^2(\theta_i) + \sigma_i^2}$. For any chosen calendar year θ_0, the likelihood of each ^{14}C-dated ring is calculated as $p(x_i|\theta_i)$, and we can use the formula for the normal distribution:

$$p(x_i|\theta_i) = \frac{1}{\omega_i(\theta_i)\sqrt{2\pi}} \exp\left\{-\frac{(x_i - \mu(\theta_i))^2}{2\omega_i^2(\theta_i)}\right\}. \tag{6}$$

Here, the vertical bar means 'given', so, we calculate the probability of a chosen age x_i given the (unknown) true age θ_i (e.g., Christen and Litton, 1995). As outlined above, instead of the normal distribution, the Student-t distribution could also be used. Now, we calculate the product of all likelihoods,

$$p(\boldsymbol{x}|\boldsymbol{\theta}) = \prod_{i=1}^{m} p(x_i|\theta_i), \tag{7}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$.

Bayes' Theorem (Buck et al., 1996; Bronk Ramsey, 2009a) states that

$$p(\theta|x) \propto p(x|\theta)p(\theta). \tag{8}$$

In words, the posterior probability (our distribution of interest, here the age of a sample given the data) is proportional to the probability of the data given the age (see above), times the prior. This brings us to the prior information.

Prior information is any knowledge we have that could help interpreting the data — information that existed *prior to* obtaining or seeing the dating results. In the case of producing chronologies, this is the relative information between the dated levels, such as Eq. (5) for tree-rings, or any knowledge about the stratigraphical ordering of sites (e.g., Harris, 1989; Buck et al., 1991; Christen and Litton, 1995; Bayliss, 2009). A very basic prior could be to constrain the ages to be within a certain timeframe, e.g., if we know that a site is of Holocene age, then we do not need to consider pre-Holocene ages, and radiocarbon-derived ages should be within the ^{14}C dating limit (< 55,000 cal BP).

For trees with year-rings, the prior information could take either of two values (Christen and Litton, 1995):

$$p(\boldsymbol{\theta}) = \begin{cases} 1 & 0 < \theta_0 < \infty; \theta_i = \theta_0 + \sum_{j=1}^{i} \delta_j \forall i = 1, 2, ..., m \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

So, if a proposed modelling solution lies within a timeframe and has the correct exact time-gaps between all ^{14}C-dated rings, then all prior information is obeyed, and the prior becomes 1. If however any of these constraints are violated, the prior becomes 0.

For stratigraphical deposits, as outlined above, we usually assume chronological ordering (unless other information is available). For this, we measure the timespans δ_i between all neighboring ages θ_i and θ_{i-1}, check that they are positive (i.e., in chronological order or $\theta_{i-1} < \theta_i$), and calculate the corresponding

probability $p_i(\theta_i - \theta_{i-1})$, where p_i is some prior distribution for the (positive) length of time-span δ_i. The joint prior is calculated by multiplying all individual probabilities (i.e. *a priori* independent):

$$p(\boldsymbol{\theta}) = \begin{cases} \prod_{i=1}^{n} & p_i(\theta_i - \theta_{i-1}) \text{ for } \theta_0 < \theta_1 < \ldots < \theta_n \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

The prior gap probabilities p_i depend on both the gap length and on the chosen distributions. Most commonly used in stratigraphies are uniform priors where any gap length between certain limits has the same prior probability. Since Jones and Nicholls (2002) show that uniform prior densities favour a span of 2δ over that of δ by a factor of c. 2^{M-1}, the uniform distributions are weighted by the amount of events within the time-span (see Bronk Ramsey, 2000; Steier and Rom, 2000; Jones and Nicholls, 2002):

$$p_i(\delta_i) = \begin{cases} \frac{1}{(R-\delta_i)} \frac{1}{(\delta_i)^{M-1}}, & \text{if } 0 < \delta_i < R \\ 0 \text{ otherwise,} \end{cases} \tag{11}$$

where $R = P - A$ is the total possible time span based on the upper and lower bounds, and M is the number of phases within the gap. If the gap length is assumed to be normally distributed as $\mu_i \pm \sigma_i$, we use

$$p_i(\delta_i|\mu_i, \sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(\delta_i - \mu_i)^2}{2\sigma_i^2(\delta_i)}\right\}. \tag{12}$$

A Gamma distribution ensures that $\delta_i > 0$:

$$p_i(\delta_i|\alpha_i, \beta_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \delta_i^{\alpha_i - 1} e^{-\beta_i \delta_i} \tag{13}$$

with mean $\mu_i = \frac{\alpha_i}{\beta_i}$ and variance $\sigma_i^2 = \frac{\alpha_i}{\beta_i^2}$. For exact gaps, we use a Dirac's delta:

$$p_i(\delta_i) = \begin{cases} 1 & \delta_i = \mu_i \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

Implementation

Let $\theta_{0,j}$ be a calendar age assigned to the outermost, youngest ring x_0 of a ^{14}C-dated tree. Since this youngest ring now has an assigned age, all of the ^{14}C-dated levels $x_1, x_2, \ldots, x_m$ too have their corresponding ages $\theta_1, \theta_2, \ldots \theta_m$ (after Eq. 5, which ensures the correct exact time gaps), and these ages are used to calculate the 'fit' of the ^{14}C ages (Eq. 6) given the age $\theta_{0,j}$. The product (Eq. 7) gives the 'fit' of the entire ^{14}C-dated tree. Now the same calculations are repeated for a sufficiently wide and dense range of calendar ages $\theta_{0,j}$, providing a 'wiggle-matched' age distribution for the ^{14}C-dated tree (Christen and Litton, 1995; Bronk Ramsey et al., 2001).

For sedimentary sites, we resort to Markov Chain Monte Carlo (MCMC) (Robert and Casella, 2004; Christen and Fox, 2010). The process starts with a simulated set of ages through Monte Carlo simulation of parameters from the prior distribution. Namely, initially an age $\theta_{0,j}$ is assigned to the topmost, youngest level x_0. This age is sampled from a uniform prior distribution with very wide boundaries (e.g., across the entire range of ^{14}C dating). Then, an age gap $\delta_{i,j}$ is simulated between levels x_0 and x_1 (this will provide a simulated age $\theta_{1,j} = \theta_{0,j} + \delta_{1,j}$ for x_1). This process is repeated for all levels in the site, providing gap and age estimates for all levels. For most levels, the gaps will be uniform with $\delta_{i,j} > 0$ to ensure chronological ordering, but levels within 'blocks' will allow for $\delta_{i,j} \leq 0$, and other gaps will need sampling from a normal or gamma distribution or require an exact gap. This series of ages and gaps then provides the first of $j = 1 : k$ iterations of modelled years $\theta_{0:n,k}$ — but many more will be needed.

Table 1. *A selection of the Pazyryk cultural complex (Kuzmin et al., 2004) dates that were used to produce Fig. 1. Only the first 10 lines are shown. The first column shows labels, the second and third show the uncalibrated radiocarbon ages and their (1σ) uncertainties. The fourth column indicates the rings (in this case, the midpoints of multiple rings), starting with the youngest, outermost ring and working down backwards in time toward the date of the oldest, innermost ring. The fifth column shows which calibration curve is to be used for each date: cc = 1 for IntCal20 (Reimer et al., 2020), cc = 2 for Marine20 (Heaton et al., 2020), cc = 3 for SHCal20 (Hogg et al., 2020) or cc = 4 for a tailor-made curve. Dates that are already on the cal BP scale get cc = 0. Commas are printed to highlight the formatting as a .csv file*

Lab ID,	C14 age,	error,	ring,	cc
U-35(rings 341–350),	2310,	43,	345.5,	1
U-34(rings 331–340),	2260,	38,	335.5,	1
U-33(rings 321–330),	2230,	34,	325.5,	1
U-32(rings 311–320),	2205,	34,	315.5,	1
U-31(rings 301–310),	2280,	36,	305.5,	1
U-30(rings 291–300),	2280,	46,	295.5,	1
U-29(rings 281–290),	2260,	40,	285.5,	1
U-28(rings 271–280),	2270,	40,	275.5,	1
U-27(rings 261–270),	2340,	41,	265.5,	1
U-26(rings 251–260),	2450,	41,	255.5,	1

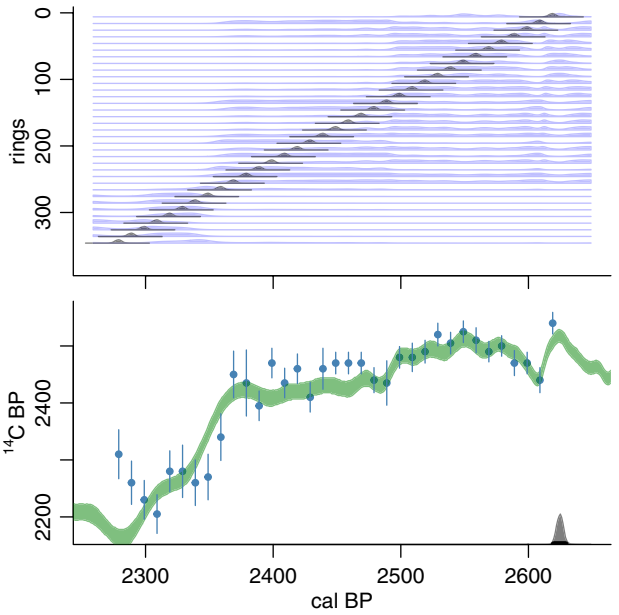

Figure 1. *Wiggle-match dating of a tree from the Pazyryk cultural complex (Kuzmin et al., 2004). Blue distributions on top panel show the unmodelled calibrated distributions for each of the* ^{14}C*-dated rings. Grey histograms show the 'wiggle-matched' age distributions for each ring. Bottom panel show the fit of the uncalibrated* ^{14}C *dates (blue) against the IntCal20 calibration curve (Reimer et al., 2020, green), and the age distribution for the oldest ring of the tree. The mean offset of the dates from the calibration curve is 0.85 standard deviations, ranging from 0 (date 24) to 3.23 (date 1).*

Table 2. *Simulated stratigraphical sections for Fig. 2 (top 6 rows) and Fig. 3 (bottom 13 rows). An undated position is indicated by cc = 10; also shown are an exact (cc = 11) and normal (cc = 12) gap. A block of 4 unordered dates is indicated with repeated entries for the position (5). Commas are printed to highlight the formatting as a .csv file*

Lab ID,	C14 age,	error,	position,	cc
sim1,	3946,	39,	1,	1
sim1,	3724,	36,	2,	1
undated,	,	,	3,	10
sim3,	4258,	46,	4,	1
sim4,	4151,	43,	5,	1
sim5,	4441,	49,	6,	1
sim1,	3946,	35,	1,	1
sim2,	3924,	34,	2,	1
exactgap,	20,	,	2.5,	11
sim3,	3834,	37,	3,	1
normalgap,	100,	10,	3.5,	12
sim4,	4014,	39,	4,	1
block1.1,	4200,	41,	5,	1
block1.2,	4338,	41,	5,	1
block1.3,	4003,	47,	5,	1
block1.4,	4250,	45,	5,	1
sim5,	4220,	38,	6,	1
sim6,	4124,	42,	7,	1
sim7,	4121,	45,	8,	1

From this initial set, randomly one of the ages will be chosen and changed somehow. This becomes a 'proposal'. For this proposal, the 'energy' will be calculated as the product of the likelihoods of the dates given the assigned ages (according to Eq. 6), times the product of all priors (the gaps, according to Eq. 11-13 as required). The proposal will also be checked to ensure that it still obeys all constraints (e.g., no reversals). Now, this proposed iteration may either be accepted as the next iteration, or it may be rejected (in which case, the previous iteration is copied to be the next iteration) using a Metropolis-Hastings acceptance probability (Robert and Casella, 2004).

Over time, the MCMC process provably approximates the underlying distribution of all parameters involved (Robert and Casella, 2004) — that is, the ages and gaps. Since each iteration depends on the previous iteration, there is a large degree of 'memory' between neighbouring iterations. To remove this memory, after running a few hundreds of thousands to millions of iterations, 'thinning' is carried out, storing only 1 every say few dozen iterations (depending on the number of parameters that are estimated). Additionally, since the first iteration is simulated from the prior, it represents sub-optimal 'ballpark' values, and commonly the first dozens of MCMC iterations often fit poorly – therefore, this 'burn-in' will have to be removed as well (by default the first 100 iterations). It is important to check that after removing the burn-in and thinning, the time-series of iterations (e.g., any of the parameters, or, the energy), does not exhibit an initial clear jump (too short burn-in) or clear auto-correlation. Several implementations of the MCMC approach are available, and here we use the t-walk (Christen and Fox, 2010) as it runs natively within *R* and is flexible enough to work with many parameters without the need for expert tuning.

To evaluate the effectiveness of our model in enhancing the accuracy of dating, either in dendrochronology or stratigraphy, a 'fit' measure is calculated for each date, by finding the percentage of modelled ages that fall within any of the highest posterior density (hpd) ranges of the unmodelled

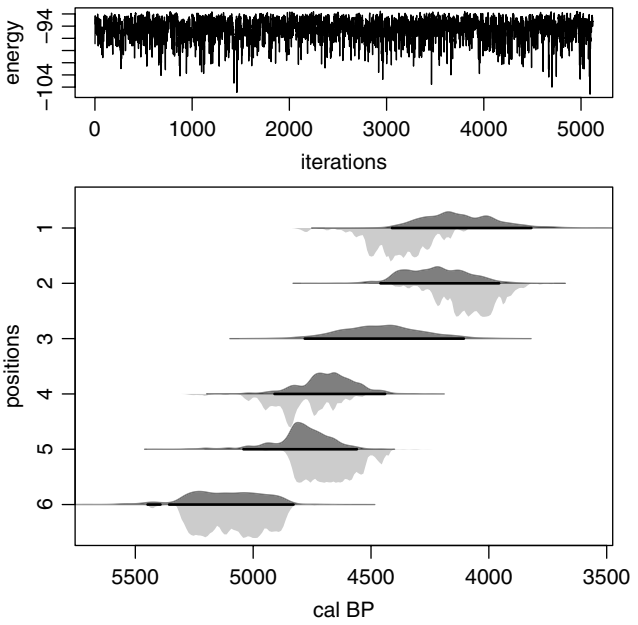

Figure 2. *A simulated stratigraphy with 5 chronologically-ordered dated positions, and one undated level which is constrained by the ages of the second and third dates (top 6 rows of Table 2). Dark-grey 'swimming elephants' or 'volcanic arc islands' show the modelled ages taking into account chronological ordering, light-grey 'reflections' show the individually calibrated ages. Note that the undated level has no reflection. This run of 400,000 iterations took around 4 minutes on a 7-year old laptop. The top panel shows the 'energy' of the 5,121 remaining MCMC iterations, with the pattern indicating a well-mixed run.*

date. By default, this is done on the 95% hpd ranges. For the tree, a goodness-of-fit is also calculated as the standardized offset (z-score) between the dates and the calibration curve's ^{14}C ages (at the assigned 'best' calendar age for each date).

Case studies

Tree

Kuzmin et al. (2004) ^{14}C-dated 35 blocks of a tree at the Pazyryk cultural complex in the Altai Mountains, Russia (Table 1). Each of the 35 blocks of rings covered 10 annual rings worth of wood, and the midpoints of neighboring blocks were separated by 10 rings. Ideally, one would sample a single year worth of wood, but this was not possible with the techniques available at the time. Re-modelling with *coffee* using IntCal20 (Reimer et al., 2020) results in a best estimate for the start of the tree growth at 2625 cal BP, with a 9-year 95% range (Fig. 1). This is a much higher precision than that of the individual calibrated dates, some of which span several centuries. Note also the scatter of the ^{14}C dates around the calibration curve – although the mean standardized offset is only 0.85 (so, less than 1σ), three dates (8% of the total) are offset from the calibration curve by $> 2\sigma$.

Stratigraphical sections

A stratigraphy was constructed by simulating an ordered site with 5 dates (the top 6 rows of Table 2; Fig. 2), accompanied by an age estimate for one undated level which is constrained by the ages of its

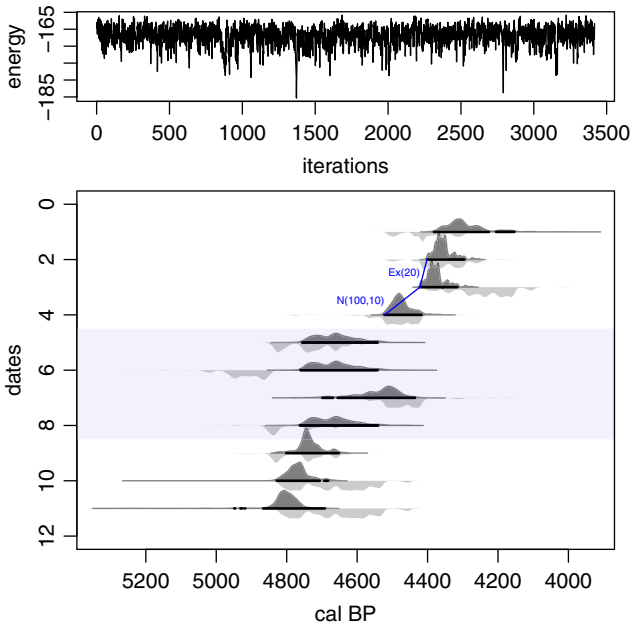

Figure 3. *A more complex simulated stratigraphy, including a 'block' (highlighted in blue) within which the dates cannot be assumed to be chronologically ordered but where the block itself is modelled to be older than the dates above it and younger than the dates below it. Also included are an exact gap of 20 years (Ex*(20)*) and one that is assumed to be normally distributed (N*(100, 10)*) (see Table 2). This run of 2 million iterations took c. 27 minutes – the remaining 3,417 MCMC iterations (top panel) show some minor areas with structure, suggesting that a longer run might be advisable. For further details, see the caption of Fig. 2.*

neighbouring positions. Note that several of the individually-calibrated dates show reversals. However, after taking into account the prior information that the ages of the positions should be in chronological order (i.e., the lowermost date should be the oldest, the fourth date should be younger than the fifth, but older than the third, and so on), the modelled, posterior distributions are indeed in chronological order. For this simulation, 400,000 MCMC iterations were run. After thinning and removal of the burn-in, 5,121 iterations remained. The mean fit between the unmodelled and modelled ages was 72.46%, ranging from 47.2% for date 1 to 87.01% for the bottommost date.

Instead of modelling the undated position as done in Table 2, we could also have modelled its age using the MCMC output — for each iteration taking the age estimates of the second and third date and modelling a random age from a uniform distribution bounded by the two age estimates.

Fig. 3 shows a more complex simulated stratigraphy, with a block of dates (within which no chronological ordering is assumed), as well as two levels indicating either exact or normally–distributed gaps. This more complex example was run using 2 million MCMC iterations, storing 1 in every 45 iterations. The mean fit between the unmodelled and modelled dates was 78.3%, ranging from 0% (date 6) to 100% (date 2).

Discussion

The R package presented here uses methods developed over decades (Buck et al., 1991; Christen, 1994a,b; Bronk Ramsey, 1995, 2001; Christen and Litton, 1995; Christen and Pérez, 2009; Christen and Fox, 2010). However, its availability within the *R* environment makes it a versatile and flexible new tool

to producing chronologies for trees and stratigraphies, given the possibility to link with R code written by users. For example, the MCMC iterations of the ages and gaps could be used as input for further analysis using other *R* packages. Note that *oxcAAR* can import *OxCal* output into *R*, but that *R* package still depends on running *OxCal* outside of *R*, whereas all calculations of *coffee* are done within *R*.

A potential advantage of *coffee* is that it uses relatively basic commands (e.g., `require(coffee); rings(); strat()`) and that the .csv files are straightforward to write and read (much like Bacon's .csv files, which have become popular for age–depth modelling of sedimentary deposits). Its *R*-based operation could enable its usage for non-interactive, batch-based chronology-building, e.g., for databases. A disadvantage of running the MCMC iterations within *R* is that it is relatively slow, taking minutes to hours to run. We are planning to implement the MCMC run as faster *c++* code in the future, if there is interest in this from the community.

Further information about *coffee* including versions, development history, a tutorial ('vignette'), news about features and bug fixes is available at https://CRAN.R-project.org/package=coffee.

Acknowledgements. We thank the organisers of CLARa2 2023 for giving us the opportunity to present the *coffee* R package at a training workshop during the conference, and thank the referee for their helpful comments which enhanced this work.

Competing interests. The authors declare that they have no competing interests.

References

Bayliss A (2009) Rolling out revolution: using radiocarbon dating in archaeology. *Radiocarbon* **51**, 123–147.

Blaauw M and Christen JA (2005) Radiocarbon peat chronologies and environmental change. *Journal of the Royal Statistical Society Series C-Applied Statistics* **54**, 805–816.

Blaauw M and Christen JA (2011) Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Analysis* **6**, 457–474.

Blaauw M, Christen JA, Bennett KD and Reimer PJ (2018) Double the dates and go for Bayes – impacts of model choice, dating density and quality on chronologies. *Quaternary Science Reviews* **188**, 58–66.

Bronk Ramsey C (1995) Radiocarbon calibration and analysis of stratigraphy: the OxCal program. *Radiocarbon* **37**, 425–430.

Bronk Ramsey C (2000) Comment on 'the use of Bayesian statistics for ^{14}C dates of chronologically ordered samples: a critical analysis'. *Radiocarbon* **42**, 199–202.

Bronk Ramsey C (2001) Development of the radiocarbon program OxCal. *Radiocarbon* **43**, 355–363.

Bronk Ramsey C (2008) Deposition models for chronological records. *Quaternary Science Reviews* **27**, 42–60.

Bronk Ramsey C (2009a) Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360.

Bronk Ramsey C (2009b) Dealing with outliers and offsets in radiocarbon dating. *Radiocarbon* **51**, 1023–1045.

Bronk Ramsey C, van der Plicht J and Weninger B (2001) 'Wiggle matching' radiocarbon dates. *Radiocarbon* **43**, 381–389.

Buck CE, Cavanagh WG and Litton CD (1996) *Bayesian Approach to Interpreting Archaeological Data*. Hoboken, NJ: Wiley.

Buck CE, Kenworthy JB, Litton CD and Smith AFM (1991) Combining archaeological and radiocarbon information - a Bayesian approach to calibration. *Antiquity* **65**, 808–821.

Buck CE, Christen JA and James GN (1999) BCal: an on–line Bayesian radiocarbon calibration tool. *Internet Archaeology* **7**. doi:10.11141/ia.7.1

Christen JA (1994a) *Bayesian Interpretation of Radiocarbon Results*. Nottingham: University of Nottingham.

Christen JA (1994b) Summarizing a set of radiocarbon determinations - a robust approach. *Applied Statistics-Journal of the Royal Statistical Society Series C* **43**, 489–503.

Christen JA and Litton CD (1995) A Bayesian-approach to wiggle-matching. *Journal of Archaeological Science* **22**, 719–725.

Christen JA and Pérez ES (2009) A new robust statistical model for radiocarbon data. *Radiocarbon* **51**, 1047–1059.

Christen JA and Fox S (2010) A general purpose sampling algorithm for continuous distributions (the t-walk). *Bayesian Analysis* **5**, 263–282.

Crema ER and Bevan A (2021) Inference from large sets of radiocarbon dates: software and methods. *Radiocarbon* **63**, 23–39.

De Vleeschouwer D and Parnell AC (2014) Reducing time-scale uncertainty for the Devonian by integrating astrochronology and Bayesian statistics. *Geology* **42**, 491–494.

Harris EC (1989) *Principles of Archaeological Stratigraphy*. London: Academic Press.

Haslett J and Parnell A (2008) A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society Series C* **57**, 399–441.

Heaton T, Köhler P, Butzin M, Bard E, Reimer R, Austin W, Bronk Ramsey C, Grootes P, Hughen K, Kromer B, Reimer P, Adkins J, Burke A, Cook M, Olsen J and Skinner L (2020) Marine20 - the marine radiocarbon age calibration curve (0–55,000 cal BP). *Radiocarbon* **62**, 779–820.

Hogg A Heaton T, Hua Q, Palmer J, Turney C, Southon J, Bayliss A, Blackwell P, Boswijk G, Bronk Ramsey C, Petchey F, Reimer P, Reimer R and Wacker L (2020) SHCal20 Southern Hemisphere calibration, 0–55,000 years cal BP. *Radiocarbon* **62**, 759–778.

Hua Q, Turnbull JC, Santos GM, Rakowski AZ, Ancapichún S, De Pol–Holz R, Hammer S, Lehman SJ, Levin I, Miller JB, Palmer JG and Turney CSM (2022) Atmospheric radiocarbon for the period 1950–2019. *Radiocarbon* **64**, 723–745.

Jones M and Nicholls G (2002) New radiocarbon calibration software. *Radiocarbon* **44**, 663–674.

Kuzmin Y, Slusarenko I, Hajdas I, Bonani G and Christen JA (2004) The comparison of ^{14}C wiggle-matching results for the 'floating' tree-ring chronology of the Ulandryk-4 burial ground (Altai Mountains, Siberia). *Radiocarbon* **46**, 943–948.

Lanos Ph and Philippe A (2018) Event date model: a robust Bayesian tool for chronology building. *Communications for Statistical Applications and Methods* **25**, 131–157.

Reimer PJ, Austin W, Bard E, Bayliss A, Blackwell P, Bronk Ramsey C, Butzin M, Cheng H, Edwards R, Friedrich M, Grootes P, Guilderson T, Hajdas I, Heaton T, Hogg A, Hughen K, Kromer B, Manning S, Muscheler R, Palmer J, Pearson C, van der Plicht J, Reimer R, Richards D, Scott EM, Southon J, Turney C, Wacker L, Adolphi F, Büntgen U, Capano M, Fahrni S, Fogtmann-Schulz A, Friedrich R, Köhler P, Kudsk S, Miyake F, Olsen J, Reinig F, Sakamoto M, Sookdeo A and Talamo S (2020) The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757.

Robert C and Casella G (2004) *Monte Carlo Statistical Methods. Springer Texts in Statistics*. New York: Springer.

Scott EM, Naysmith P and Cook GT (2018) Why do we need ^{14}C inter-comparisons?: The Glasgow–^{14}C inter-comparison series, a reflection over 30 years. *Quaternary Geochronology* **43**, 72–82.

Scott EM (2023) Uncertainty and error in radiocarbon dating. In: *Encyclopedia of Quaternary Science*, 3rd Edn. Amsterdam: Elsevier.

Steier P and Rom W (2000) The use of Bayesian statistics for C-14 dates of chronologically ordered samples: a critical analysis. *Radiocarbon* **42**, 183–198.

Cite this article: Blaauw M, Aquino-López M, and Christen JA (2026). Modelling chronologically ordered radiocarbon dates in R. *Radiocarbon* **68**, 21–31. https://doi.org/10.1017/RDC.2024.56