USING SEDIMENTOLOGICAL PRIORS TO IMPROVE ^{14}C CALIBRATION OF BIOTURBATED SEDIMENT ARCHIVES

Bryan C Lougheed*

Department of Earth Sciences, Uppsala University, Sweden

ABSTRACT. Radiocarbon (^{14}C) dating is often carried out upon multi-specimen samples sourced from bioturbated sediment archives, such as deep-sea sediment. These samples are inherently heterogeneous in age, but existing ^{14}C calibration techniques were originally developed for age homogeneous material, such as archaeological artifacts or individual tree rings. A lack of information about age heterogeneity leads to a systematic underestimation of a sample's true age range, as well as the possible generation of significant age-depth artifacts during periods of the Earth's history coinciding with highly dynamic atmospheric Δ^{14}C. Here, a new calibration protocol is described that allows for the application of sedimentological priors describing sediment accumulation rate, bioturbation depth and temporally dynamic species abundance. This Bayesian approach produces a credible calibrated age distribution associated with a particular laboratory ^{14}C determination and its associated sedimentological priors, resulting in an improved calibration, especially in the case of low sediment accumulation rates typical of deep-sea sediment. A time-optimized computer script (*biocal*) for the new calibration protocol is also presented, thus allowing for rapid and automated application of the new calibration protocol. This new calibration protocol could be applied within existing age-depth modeling software packages to produce more accurate geochronologies for bioturbated sediment archives.

KEYWORDS: Bayesian, bioturbation, calibration, sediment.

INTRODUCTION

Radiocarbon (^{14}C) analysis is routinely used to determine the age of marine sediment archives up to ~50 ka in age, and has been fundamental in increasing our understanding of the spatio-temporal development of palaeoclimate during the last glacial and the Holocene. However, due to ^{14}C being a very rare radioisotope in the environment (approximately one in 10^{12} carbon dioxide molecules in the atmosphere is ^{14}CO$_2$), it is more difficult to measure than more common, stable carbon isotopes. From a practical standpoint, this rarity results in a requirement of relatively large sample sizes to attain a sufficient measurement signal using, e.g., accelerated mass spectrometry (AMS). In the case of, e.g., deep-sea sediment archives, many tens of single microfossil specimens from a discrete core depth are often pooled into a single sample for measurement. However, AMS will only report a mean ^{14}C activity (along with a measurement uncertainty), so any information about the ^{14}C heterogeneity of the sample will be lost.

Systematic bioturbation of deep-sea sediment causes discrete downcore intervals of deep-sea sediment to have an age distribution that is characterized by an exponential probability density function with a long tail towards older ages (Berger and Heath 1968). This age distribution is mainly governed by the sediment accumulation rate (SAR) and bioturbation depth (BD), the latter of which is typically around 10 cm (Trauth et al. 1997; Boudreau 1998). The presence of the aforementioned age distribution is supported by studies of, e.g., particle mixing, stable isotopes, ^{14}C, species abundance and tephras (Bramlette and Bradley 1942; Nayudu 1964; Ruddiman and Glover 1972; Peng et al. 1979; Hutson 1980; Pisias 1983; Schiffelbein 1984; Andree 1987; Bard et al. 1987; Wheatcroft 1992; Trauth
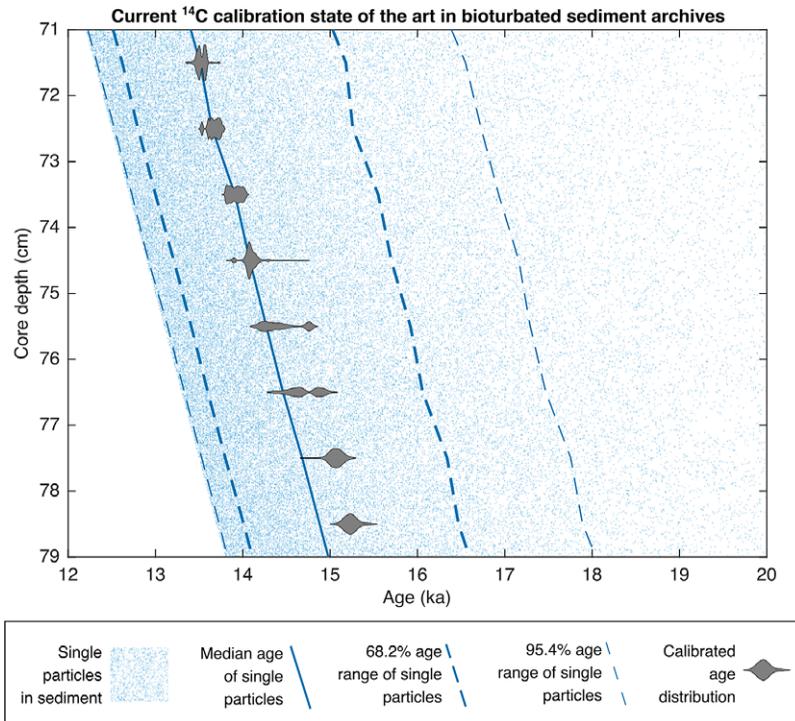
*Corresponding author. Email: bryan.lougheed@geo.uu.se

Figure 1   5 cm ka^{-1} sediment simulation of single particles using a global average BD of 10 cm (Trauth et al. 1997; Boudreau 1998) and a best-case 10^4 simulated particles per cm. Shown also is the discrete 1 cm depth median age, as well as the associated 68.2% and 95.4% age range. Also shown are the calibrated age distributions that would result if one were to use the existing state of the art to calibrate the mean ^{14}C activity resulting from all the particles contained in each 1 cm discrete depth. The single particle simulation is carried out using SEAMUS (Lougheed 2020), using the IntCal20 calibration curve (Reimer et al. 2020) and assuming no reservoir affect. Calibration is carried out using MatCal (Lougheed and Obrochta 2016).

et al. 1997; Henderiks et al. 2002; Löwemark and Grootes 2004; Sepulcre et al. 2017; Lougheed et al. 2018; Abbott et al. 2018; Missiaen et al. 2020; Dolman et al. 2021).

In the case of a wholly non-bioturbated sediment archive (such as laminated sediment retrieved from an anoxic environment), a SAR of 5 cm ka^{-1} can be assumed to have a temporal resolution of 1000/5 = 200 yr cm^{-1}. However, in the case of bioturbated sediment typical of the oxygenated deep-sea, the 1σ age value of a 1 cm slice of sediment with a typical deep-sea SAR of 5 cm ka^{-1} and BD of 10 cm can be approximated as 10/5×1000 = 2000 yr (Berger and Heath 1968). Somewhat counter-intuitively, that same bioturbated sediment archive will also exhibit a downcore increasing mean age of 200 yr cm^{-1}, which can deceptively mask the fact that the sediment is bioturbated. In essence, downcore increase in mean age is not the same concept as the discrete-depth age variance. This concept is visualized in Figure 1 by a 5 cm ka^{-1} single particle sediment simulation, where it can be seen that the discrete-depth median age of the single particle population increases by ~200 yr per cm, whereas the actual age range contained in each discrete 1 cm depth is actually much greater, and characterized by an exponential distribution with a long tail towards older ages (which can be seen in Figure 1 as the decreasing density of single

particles towards older ages). Hence, the true temporal resolution of the 5 cm ka^{-1} discrete-depth archive is much greater than 200 yr, and failing to consider this point risks leading to false high-precision in age-depth chronologies.

The current state of the art in palaeoclimate includes no information about bioturbation when ^{14}C calibrating multi-specimen samples retrieved from deep-sea (or lacustrine) sediment archives. In essence, the current state of the art considers only ^{14}C-centric priors (calibration curve and reservoir effect) when estimating (calibrating) true age. Such an approach incorrectly treats deep-sea sediment as having discrete age increments, similar to non-bioturbated archives such as tree rings, speleothems and/or varves. By ignoring bioturbation when calibrating ^{14}C measurements from sediment archives, one essentially assumes that a BD of 0 cm and SAR of 1000 cm ka^{-1}. Hence, the current lack of method for including correct sedimentological priors when applying the ^{14}C method to sediment archives can lead to an underestimation of the full age uncertainty. This underestimation is also illustrated in Figure 1, where the existing state of the art in ^{14}C dating and calibration is virtually applied to a simulated bioturbated sediment core created using the SEAMUS single foraminifera simulator (Lougheed 2020). In can further be seen in Figure 1 that the current calibration method can produce significant age-depth artifacts when applied to bioturbation, which is due to the mixing of single elements (e.g., foraminifera) from periods of past dynamic Δ^{14}C into the same discrete depth interval (Lougheed et al. 2020).

METHOD

The new calibration protocol presented here involves complementing the traditional ^{14}C priors (past Δ^{14}C from a calibration curve, reservoir effect) with sedimentological priors (SAR, BD and temporal changes in species abundance). This improved calibration protocol for sediment archives allows us to estimate an improved age distribution from the ^{14}C activity measurement carried out on a given bioturbated sample (Figure 2).

Establishing a Prior Distribution for Calendar Age

In order to calibrate ^{14}C activity measurements carried out upon heterogeneous samples retrieved from bioturbated sediment, the following sedimentological priors are defined:

s = estimated sediment accumulation rate (SAR), in cm yr^{-1}
m = bioturbation (mixing) depth (BD), in cm
k = the fraction of the analyzed microfossils that are fragmented (a value between 0 and 1)
a = time series of abundance of the analyzed species relative to itself (values between 0 and 1)

Both SAR and BD are considered here as a constant value, i.e., not as a time series of temporally variable values. These inputs are kept constant foremost to reduce computation time, and also because temporal changes in, e.g., SAR (the relationship between mean age and depth) are not known when an age-depth chronology has yet to be developed. In short, applying detailed information about temporal changes in SAR when the age-depth relationship of the sediment is not yet known would constitute circular thinking.

Prior information is often applied within Bayesian analysis to construct an expected prior probability distribution based on established understanding of physical processes. In this case, we use SAR and BD priors to construct a prior distribution of relative age for the sample being calibrated, based on theoretical understanding of the influence of bioturbation
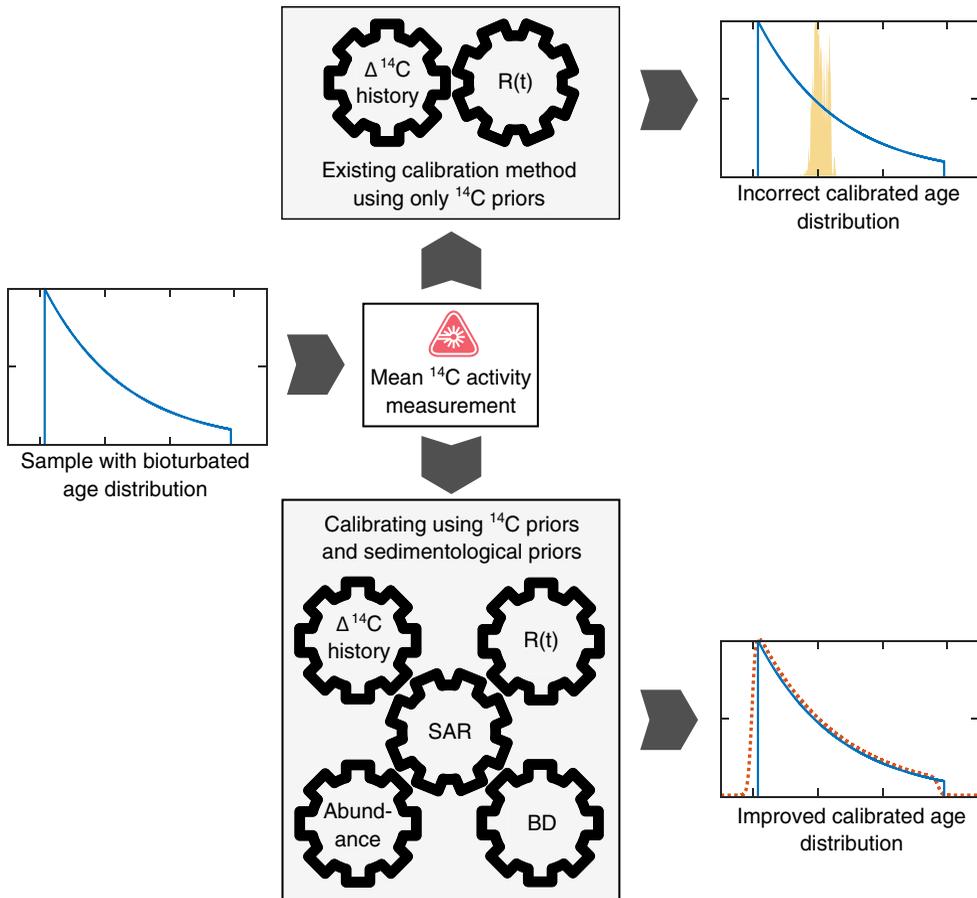
Figure 2   A flow chart demonstrating the principle of complementing ^{14}C priors with sedimentological priors to produce a more accurate calibrated age distribution for bioturbated sediment archives.

upon the age distribution of sediment. Following Berger and Heath (1968), the age distribution for a given depth of fully bioturbated sediment core can be represented by an exponential probability distribution, which can be considered the basis of the prior probability distribution for a sample's calibrated age:

$$p_{prior}(r_1, r_2, \ldots, r_n) = exp\left(\frac{-(r_1, r_2, \ldots, r_n)s}{m}\right) \qquad (1)$$

where r is the relative age (starting at 1 yr) within P_{prior}. The low-probability long tail of an exponential probability function continues to infinity, which obviously cannot be stored in computer memory. The prior distribution is therefore limited to the age equivalent value of five bioturbation depths, i.e., a relative age of $r_{limit} = 5m/s$, which is rounded to the nearest whole year.

When picking microfossils for ^{14}C analysis, palaeoceanographers generally prefer to pick whole and/or pristine specimens. The fragmented and/or dissolved microfossils that are not picked have been resident in the bioturbation depth for a longer time and have been

exposed to more bioturbation cycles, and as such represent the oldest fraction of the sample (Rubin and Suess 1955; Ericson et al. 1956; Emiliani and Milliman 1966; Barker et al. 2007). There is therefore a benefit in not picking the older, broken foraminifera, as it results in a more constrained age distribution (the long tail of the age distribution is shortened). Information regarding the fact that the oldest/broken foraminifera are not picked can be incorporated into the prior distribution. The estimated fraction of fragmented microfossils (k) can be related to the cumulative expression of Eq. (1):

$$1 - k = 1 - \exp\left(\frac{-rs}{m}\right) \tag{2}$$

Eq. (2) can be solved to attain $r(k)$, the threshold age for fragmented foraminifera:

$$r(k) = \frac{-m.\ln(k)}{s} \tag{3}$$

Regions of the prior probability distribution (p_{prior}) older than $r(k)$ can, therefore, be considered to consist of fragmented microfossils that are not picked by palaeoceanographers. When $r(k) < r_{limit}$, p_{prior} is truncated at the discrete relative age $r(k)$ to incorporate prior information from the picking process. When $r(k) \geq r_{limit}$, $r(k)$ is approximated to r_{limit}. All discrete probability values in p_{prior} are subsequently normalized such that they sum to 1.

Establishing a Distribution for [14]C Activity

Please note that, to avoid ambiguity, throughout this text the use of the term "age" refers exclusively to true/calibrated age, while ^{14}C activity is always referred to as ^{14}C activity, i.e., not as "[14]C age".

The new calibration protocol must incorporate the full uncertainty regarding ^{14}C activity, which includes uncertainties regarding the laboratory ^{14}C activity determination, the calibration curve ^{14}C activity, and the ^{14}C activity depletion as a result of the reservoir effect. These are expressed here as follows:

A_{det} = The laboratory ^{14}C activity determination of the sample (in ^{14}C yr BP).

σ_{det} = The measurement uncertainty associated with A_{det} (in ^{14}C yr).

$A_{cc}(t)$ = The ^{14}C activity (in ^{14}C yr BP) predicted by the calibration curve for a discrete age t.

$\sigma_{cc}(t)$ = The uncertainty (in ^{14}C yr) associated with $A_{cc}(t)$.

$R(t)$ = The predicted ^{14}C activity depletion (in ^{14}C yr) of A_{det} relative to the calibration curve at discrete age t, due to a local reservoir effect (Stuiver et al. 1986). $R(t)$ can be substituted with $\Delta R(t)$ in the case of a marine calibration curve.

$\sigma_R(t)$ = The uncertainty (in ^{14}C yr) associated with $R(t)$ (or $\Delta R(t)$).

Activity depletion due to $R(t)$ is considered here by incorporating it into the calibration curve ^{14}C activity. This approach to handling $R(t)$ allows, if desired, for temporally dynamic $R(t)$ to be correctly incorporated (Waelbroeck et al. 2019). The calibration curve is adjusted as follows, for each discrete calendar age t:

$$A_{ccR}(t) = A_{cc}(t) + R(t) \tag{4}$$

Uncertainties pertaining to calibration curve ^{14}C activity and the ^{14}C reservoir effect ($\sigma_{cc}(t)$ and $\sigma_R(t)$) are both Gaussian, so they can be easily propagated into one term, for each discrete calendar age t:

$$\sigma_{ccR}(t) = \sqrt{(\sigma_{cc}^2(t) + \sigma_R^2(t))} \tag{5}$$

Before proceeding, all of the aforementioned ^{14}C-related values are first converted into F^{14}C space to facilitate more accurate calculations that take isotope mass balance into account, which is especially relevant in the case of wide range of ^{14}C activity (Erlenkeuser 1980; Bronk Ramsey 2008; Keigwin and Guilderson 2009), such as is the case with bioturbated sediment archives.

A sequence of probabilities can describe the closeness of a sequence of ^{14}C activities predicted for all discrete ages t (represented as T) available within the calibration curve (i.e., $A_{ccR}(T)$), to a single ^{14}C activity predicted by the calibration curve for a discrete age t (i.e., $A_{ccR}(t)$). This closeness, which includes a quantification of calibration curve and reservoir effect uncertainties, can be evaluated using a normal distribution for each instance of t, summing through all n values available in T to give the total relative ^{14}C probability for each t:

$$p_{14C}(T|t) = \sum_{T_1}^{T_n} \left(\frac{1}{\sigma_{ccR}(t)\sqrt{(2\pi)}} \exp\left(\frac{-(A_{ccR}(T) - A_{ccR}(t))^2}{2\sigma_{ccR}^2(t)} \right) \right) \tag{6}$$

The Prior Calibration Process

The prior calibration process involves moving the p_{prior} distribution along a sliding window of calendar ages and each time computing the the hypothetical laboratory mean ^{14}C activity determination (h_{det}) that would result from each p_{prior} placed at a sliding window starting at each t:

$$h_{det}(t) = \sum_{r=1}^{r(k)} \left(A_{ccR}(t + r - 1) \cdot p_{14C}(T|t + r - 1) \cdot p_{prior}(r) \cdot a(t + r - 1) \right) \tag{7}$$

Subsequently, it is possible to evaluate the single probability value of each $h_{det}(t)$ as a function of its closeness to the normal distribution of the sample's observed laboratory determination $A_{det} \pm \sigma_{det}$:

$$p_{h_{det}}(t) = \frac{1}{\sigma_{det}(t)\sqrt{(2\pi)}} \exp\left(\frac{-(h_{det}(t) - A_{det})^2}{2\sigma_{det}^2} \right) \tag{8}$$

For each sliding window placed at each t, a vector of calibrated age probabilities is calculated, corresponding to each discrete age in the sliding window:

$$p_{cal}(t) = p_{h_{det}}(t) \cdot \left(p_{prior}(r, r+1, \ldots, r(k)) \odot a(t, t+1, \ldots, t + r(k) - 1) \right) \tag{9}$$

Subsequently, each $p_{cal}(t)$ is sorted into a large matrix, referred to here as $M_{cal}(T)$:

$$M_{cal}(T) = \begin{bmatrix} p_{cal}(t_1)_1 & p_{cal}(t_1)_2 & \ldots & p_{cal}(t_1)_n & 0 & 0 \\ 0 & p_{cal}(t_2)_1 & p_{cal}(t_2)_2 & \ldots & p_{cal}(t_2)_n & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & p_{cal}(t_n)_1 & p_{cal}(t_n)_2 & \ldots & p_{cal}(t_n)_n \end{bmatrix} \quad (10)$$

The final credible calibrated probability distribution corresponding to all ages T can be calculated simply by summing all rows in $M_{cal}(T)$:

$$p_{cal}(T) = \sum_{i=1}^{n} M_{cal}(T)_{ij} \quad (11)$$

All elements in the resulting vector $p_{cal}(T)$ are subsequently normalized such that they sum to 1.

Script for Automated Calibration (*biocal*)

Here, a fully documented Matlab function (*biocal.m*) is provided for automated calculation of the calibration protocol outlined in this study, with full compatibility in Octave. Other programming language versions of the script (e.g., Python, Julia, R) are forthcoming and will be uploaded to the same software repository upon completion. The *biocal* script takes full advantage of computer memory to carry out calculations using vectorized programming, thus resulting in a time-optimized routine. In the calibration protocol described in the previous section, it is assumed that it is possible to calculate P_{prior} sliding windows along the the entire history covered by the calibration curve. However, as it would be computationally prohibitive to calibrate for the entire history of the calibration curve, *biocal* restricts its P_{prior} sliding window calculations to an interval of the calibration curve covering a 3σ distance in each direction from the laboratory ^{14}C determination, with added padding to accommodate a long tail of P_{prior} sitting at +3σ distance. In future, when computer memory and processor power increases by another order of magnitude, it will be possible to compute sliding windows across the entire calibration curve, assuming that would ever be deemed necessary. For now, as long as the tails of the final calibrated probability distribution gradually fall to very small values near to zero, we can know that a sufficient interval of the calibration curve has been considered.

The calculation time and memory usage for *biocal* increases with decreasing SAR, increasing BD, increasing ^{14}C measurement uncertainty, increasing calibration curve uncertainty and increasing reservoir effect uncertainty. Testing using Matlab 2020a on a Linux system with an Intel i7-9700 CPU resulted in the following times and memory usage: a Younger Dryas aged sample with SAR of 4 cm ka^{-1} and BD of 10 cm required 1.7 s calculation time and 2GB memory; the same sample, but with a SAR of 20 cm/ka^{-1}, required 0.2 s to calculate and used 100 MB of memory.

GROUND-TRUTH EVALUATION

Evaluating Calibration Using Sedimentological Priors

Here, a test is carried out to determine if the calibration protocol incorporating sedimentological priors results in an improved calibration process (i.e., a better estimation
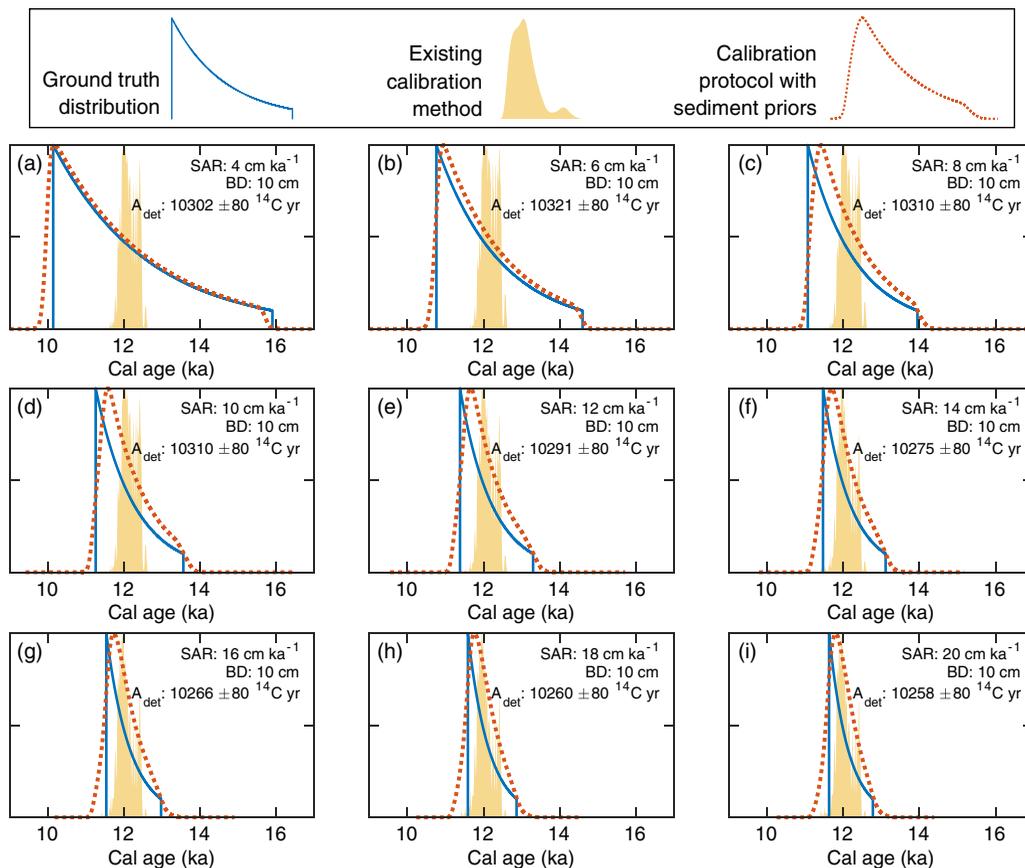
Figure 3   Comparing the new ^{14}C calibration protocol to the existing ^{14}C calibration method, in the case of samples with a mean age of 12 ka, constant species abundance and various sedimentological prior scenarios. Shown in all panels: the ground-truth age distribution (solid blue line); the age distribution estimated using the new ^{14}C calibration protocol with sedimentological priors (dashed orange line); the age distribution estimated using the traditional ^{14}C calibration method (filled yellow area). A_{det} is the expected mean ^{14}C activity determination resulting from the ground-truth age distribution according to IntCal20. The following scenarios are considered as sedimentological priors: Panel A: SAR 4 cm ka^{-1}, BD 10 cm; Panel B: SAR 6 cm ka^{-1}, BD 10 cm; Panel C: SAR 8 cm ka^{-1}, BD 10 cm; Panel D: SAR 10 cm ka^{-1}, BD 10 cm; Panel E: SAR 12 cm ka^{-1}, BD 10 cm; Panel F: SAR 14 cm ka^{-1}, BD 10 cm; Panel G: SAR 16 cm ka^{-1}, BD 10 cm; Panel H: SAR 18 cm ka^{-1}, BD 10 cm; Panel I: SAR 20 cm ka^{-1}, BD 10 cm. (Please see electronic version for color figures.)

of the true age distribution of the measured sample) for a number of SAR scenarios and using a globally representative BD of 10 cm. First, the established understanding of bioturbation's effect upon age-depth models on geological timescales (Berger and Heath 1968; Guinasso and Schink 1975; Peng et al. 1979; Trauth et al. 1997; Trauth 1998; Dolman and Laepple 2018; Lougheed 2020) is used to calculate the associated annualized age distribution that would be expected for a discrete-depth, 1 cm sediment sample. For all scenarios, the mean value of the age distribution is set at 12 ka, and it is assumed that the oldest 10% of the foraminifera are broken foraminifera that are not picked and, therefore, not included in the distribution. These age distributions represent the ground-truth age distribution of our virtual sample (represented as solid blue lines in Figure 3), the target age distribution that a calibrated age distribution can be judged against.

Subsequently, we can carry out a "virtual AMS analysis" upon the ground-truth distribution by using the IntCal20 (Reimer et al. 2020) calibration curve to determine the mean ^{14}C activity that could be expected, in a best-case scenario, to result from the aforementioned age distribution. For simplicity's sake, no reservoir effect is included in this demonstration, and it is assumed that the mean ^{14}C activity reported by IntCal20 perfectly represents the ^{14}C activity recorded by the sediment archive, with linear interpolation applied to IntCal20 where necessary to achieve annual resolution.

Assuming an appropriate ^{14}C measurement uncertainty of ±80 ^{14}C yr, the mean ^{14}C activity can then be calibrated in two ways, which can subsequently be compared to each other: (1) using IntCal20 and Matcal 3.1 (Lougheed and Obrochta 2016) to carry out the existing, standard ^{14}C calibration procedure following, e.g., Bronk Ramsey (2008), shown in Figure 3 as filled yellow areas; (2) using the aforementioned *biocal* in combination with IntCal20, supplemented by the SAR and BD priors associated with each scenario, to carry out the new calibration protocol outlined in this study, which is represented in Figure 3 as broken orange lines.

As could be expected, the calibration protocol using sedimentological priors outperforms the standard calibration procedure in estimating the ground-truth age distribution, as shown in Figure 3 for a number of SAR scenarios ranging between 4 and 20 cm ka^{-1}, with a BD of 10 cm and constant temporal species abundance. In such use case scenarios, using the calibration protocol with sedimentological priors demonstrably leads to a more accurate calibrated age distribution, which would be ideal for improving age-depth modeling of low SAR sediment archives.

In Figure 4, we repeat the same SAR scenarios as previously, but in the case of a much older ground-truth scenario (mean age of 32 ka), whereby Gaussian uncertainties associated with both the sample ^{14}C activity (±300 ^{14}C yr assumed here) and the ^{14}C calibration curve are both markedly increased. In Figure 4(e–i), it can be seen that these larger uncertainties, when combined with increasing SAR, lead to the sedimentological priors becoming overwhelmed by the Gaussian ^{14}C uncertainties and, consequently, the calibrated age distribution determined by the procedure starts to approach a normal distribution. In these use case scenarios, the new calibration protocol using sedimentological priors does not necessarily offer any advantage over the traditional calibration method.

Additionally, it is also possible to revisit the 5 cm ka^{-1} scenario from Figure 1, where it was shown that the traditional calibration method would misrepresent the age distribution of bioturbated (deep-sea) sediment. The new calibration protocol using sedimentological priors is applied to the same simulated sediment core (Figure 5), resulting in a much-improved calibration, whereby the 95.4% age interval predicted by the new calibration protocol provides an almost complete overlap with the actual 95.4% age interval of the single particle population. There remain some minor age-depth artifacts which result from single particles during periods of highly dynamic Δ^{14}C (e.g., the last deglaciation) being mixed into the same discrete depths. This is an unavoidable fact of ^{14}C dating of bioturbated sediment records, so researchers should remain vigilant when interpreting apparent SAR changes during periods of highly dynamic Δ^{14}C. However, when one uses the new calibration protocol detailed here, the relative effect of these age-depth artifacts is reduced due to the much more realistic and wider calibrated age confidence intervals.
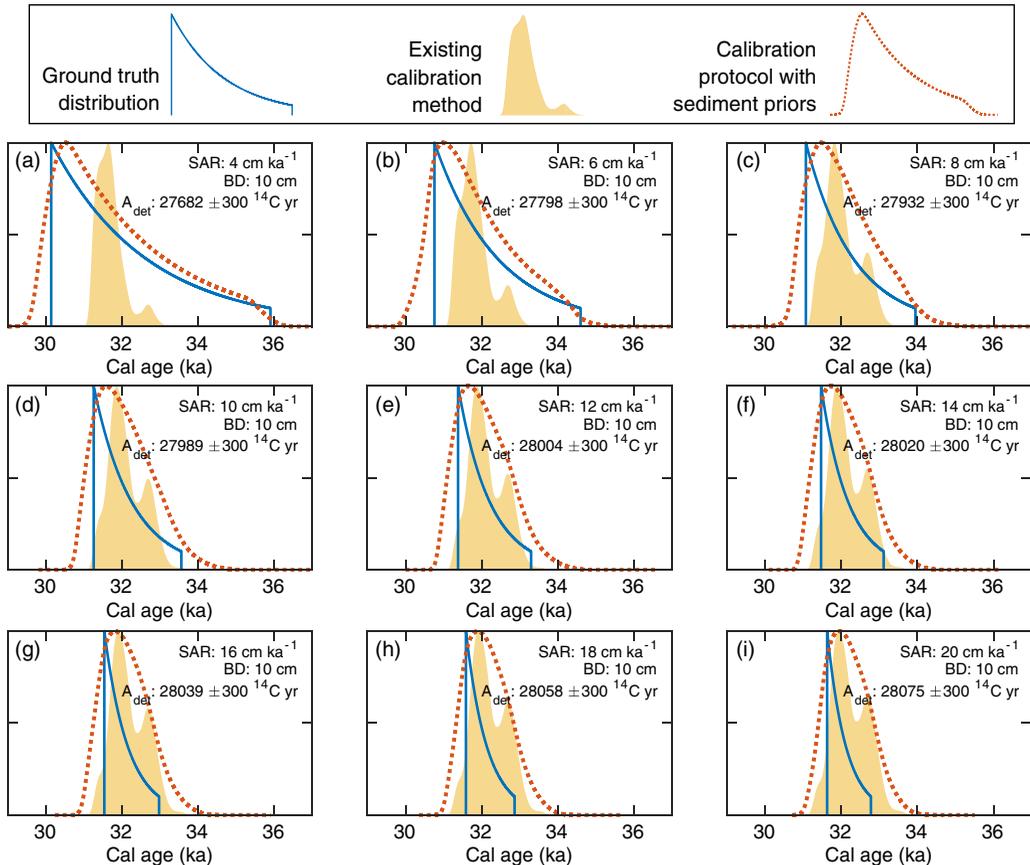
Figure 4   Comparing the new ^{14}C calibration protocol to the existing ^{14}C calibration method, in the case of samples with a mean age of 32 ka, constant species abundance and various sedimentological prior scenarios. Shown in all panels: the ground-truth age distribution (solid blue line); the age distribution estimated using the new ^{14}C calibration protocol with sedimentological priors (broken orange line); the age distribution estimated using the traditional ^{14}C calibration method (filled yellow area). A_{det} is the expected mean ^{14}C activity determination resulting from the ground-truth age distribution according to IntCal20. The following scenarios are considered as sedimentological priors: Panel A: SAR 4 cm ka^{-1}, BD 10 cm; Panel B: SAR 6 cm ka^{-1}, BD 10 cm; Panel C: SAR 8 cm ka^{-1}, BD 10 cm; Panel D: SAR 10 cm ka^{-1}, BD 10 cm; Panel E: SAR 12 cm ka^{-1}, BD 10 cm; Panel F: SAR 14 cm ka^{-1}, BD 10 cm; Panel G: SAR 16 cm ka^{-1}, BD 10 cm; Panel H: SAR 18 cm ka^{-1}, BD 10 cm; Panel I: SAR 20 cm ka^{-1}, BD 10 cm.

Evaluating Calibration Using Sedimentological and Abundance Priors

Temporal changes in species abundance (e.g., of foraminifera) will affect the shape of the species' age distribution for a given discrete depth. Here, a sine wave with a wavelength of 2000 yr is used, purely for demonstrational purposes, as a theoretical temporal abundance function (Figure 6). In Figure 7, the same SAR scenarios as in Figure 3 are analyzed, but this time with the application of the abundance aspect. Firstly, the aforementioned sinusoidal temporal abundance function is applied to the ground truth distribution. Subsequently, the same abundance function is used as an additional prior input when running *biocal*, to complement the sedimentological priors. The results in Figure 4 demonstrate how known information about temporal changes in species abundance can be used to produce better informed calibrated age estimations for bioturbated sediment
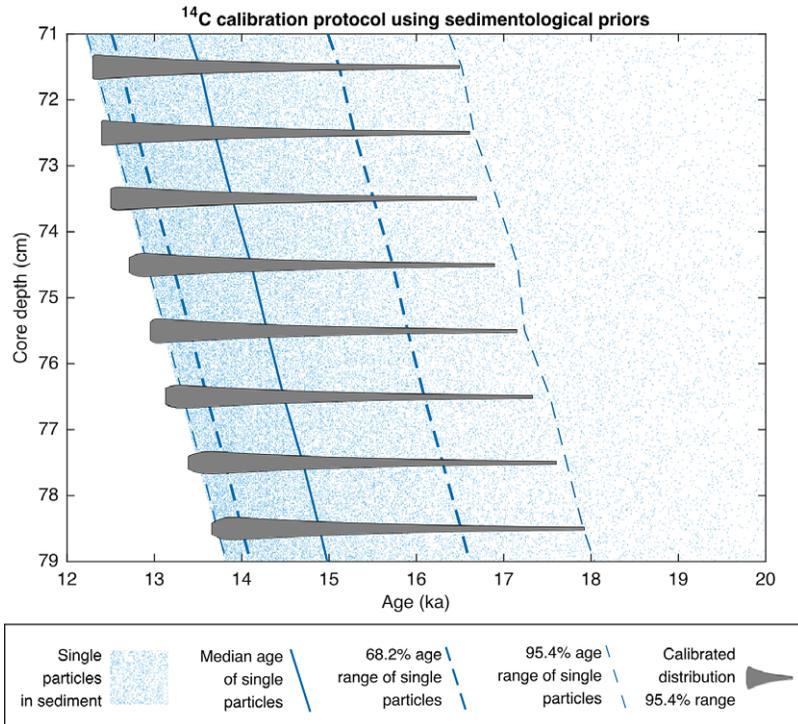
Figure 5 5 cm ka^{-1} sediment simulation of single particles using a global average BD of 10 cm (Trauth et al. 1997; Boudreau 1998) and best-case 10^4 particles per cm. Shown also is the discrete 1 cm depth median age, as well as the associated 68.2% and 95.4% age range. Also shown are the calibrated age distributions that would result if one were to use the new calibration protocol outlined in this manuscript to calibrate the mean ^{14}C activity resulting from all the particles contained in each 1 cm discrete depth. Specifically, the *biocal* routine is applied to the mean ^{14}C age of each discrete depth, with a SAR prior of 5 cm ka^{-1} and a BD prior of 10 cm. The single particle simulation is carried out as in Figure 1.
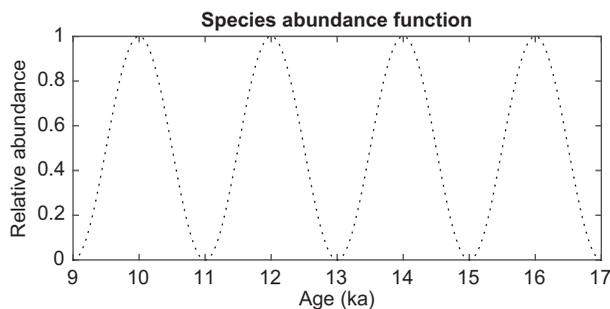

Figure 6 Visualization of the theoretical species abundance function used in this study to demonstrate the incorporation of prior information about species abundance in the ^{14}C calibration protocol developed in this study. The abundance function is implemented as a sine wave with a wavelength of 2000 yr.

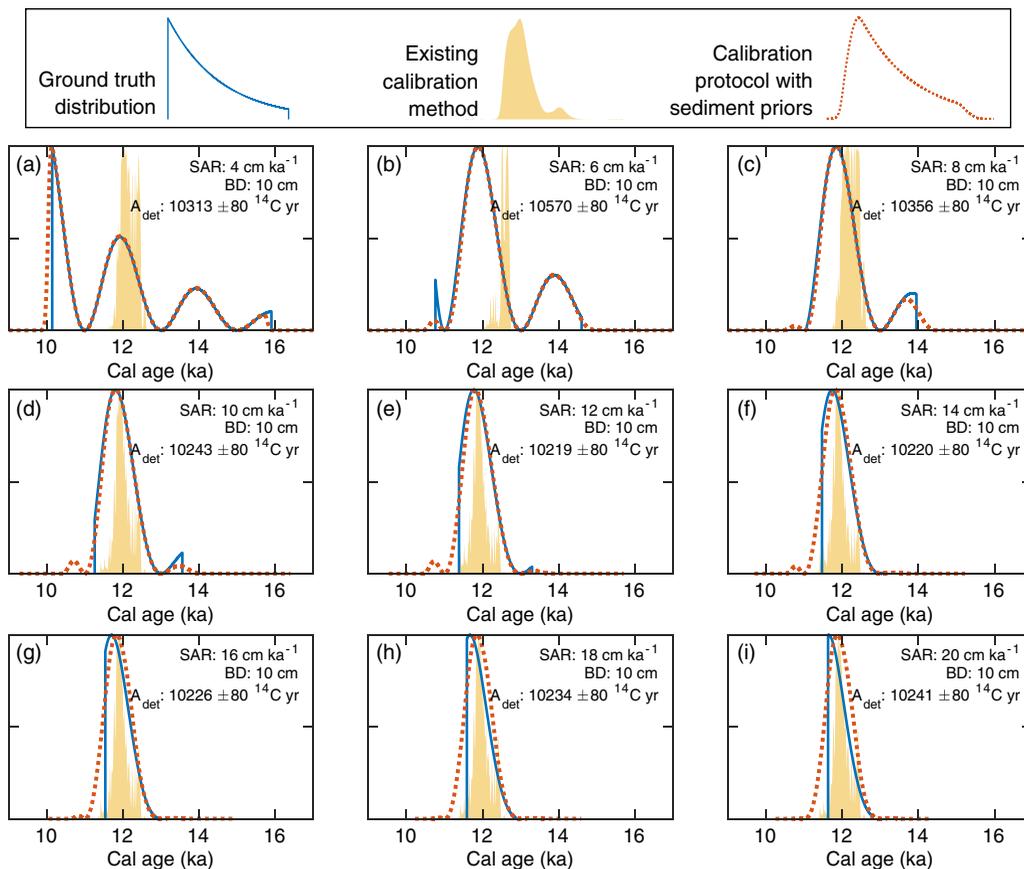Figure 7   Comparing the new ^{14}C calibration protocol to the existing ^{14}C calibration method, in the case of samples with a mean age of 12 ka, temporally dynamic species abundance and various sedimentological prior scenarios. Shown in all panels: the ground-truth age distribution (solid blue line); the age distribution estimated using the new ^{14}C calibration protocol with sedimentological priors (broken orange line); the age distribution estimated using the traditional ^{14}C calibration method (filled yellow area). A_{det} is the expected mean ^{14}C activity determination resulting from the ground-truth age distribution according to IntCal20. The following scenarios are considered as sedimentological priors: Panel A: SAR 4 cm ka^{-1}, BD 10 cm; Panel B: SAR 6 cm ka^{-1}, BD 10 cm; Panel C: SAR 8 cm ka^{-1}, BD 10 cm; Panel D: SAR 10 cm ka^{-1}, BD 10 cm; Panel E: SAR 12 cm ka^{-1}, BD 10 cm; Panel F: SAR 14 cm ka^{-1}, BD 10 cm; Panel G: SAR 16 cm ka^{-1}, BD 10 cm; Panel H: SAR 18 cm ka^{-1}, BD 10 cm; Panel I: SAR 20 cm ka^{-1}, BD 10 cm.

archives. In Figure 8, the 2000-year wavelength abundance function is also applied to in the case of an older ground-truth distribution, demonstrating that abundance priors can also be used as a tool to better constrain ^{14}C analysis of older samples that have greater uncertainty.

ADVICE FOR DETERMINING PRIOR VALUES

In order to carry out the calibration protocol detailed here, prior values for SAR, BD, fraction broken foraminifera, temporal species abundance and temporal reservoir effect are required. A first order estimate for the sediment accumulation rate can be ascertained by examining the general relationship between age-depth determinations (including ^{14}C-derived age estimates based on existing calibration methods without sedimentological priors). This approach does
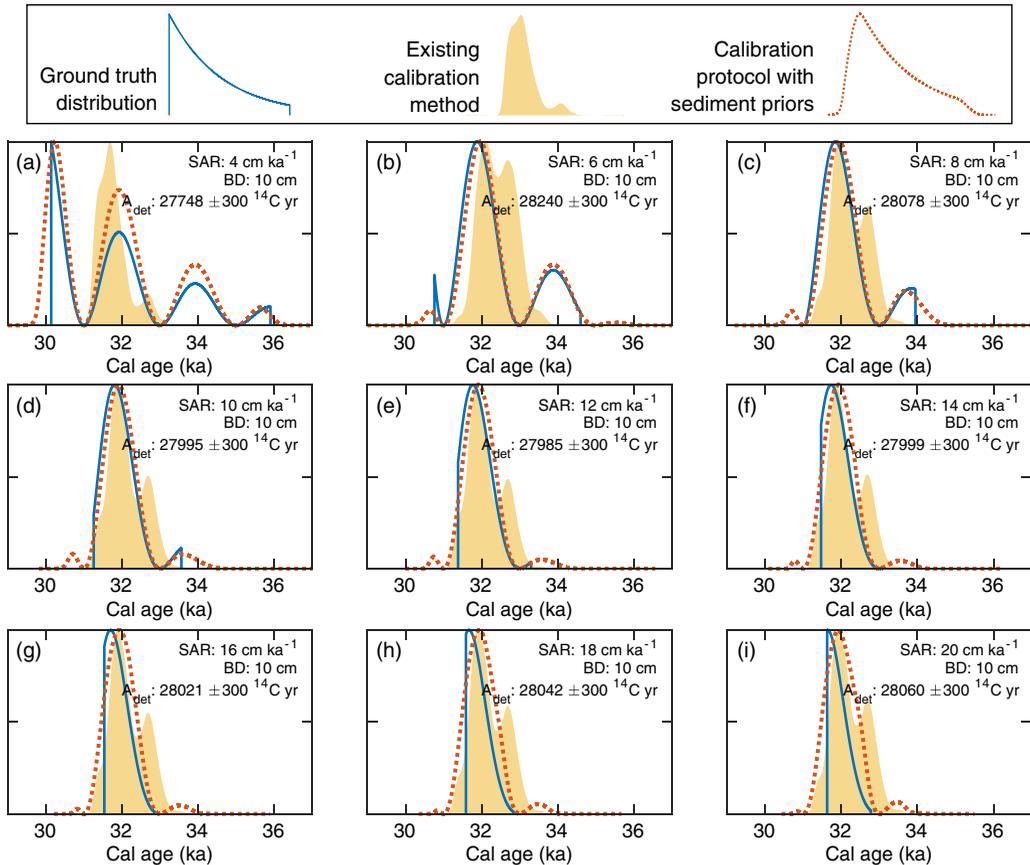
Figure 8 Comparing the new ^{14}C calibration protocol to the existing ^{14}C calibration method, in the case of samples with a mean age of 32 ka, temporally dynamic species abundance and various sedimentological prior scenarios. Shown in all panels: the ground-truth age distribution (solid blue line); the age distribution estimated using the new ^{14}C calibration protocol with sedimentological priors (broken orange line); the age distribution estimated using the traditional ^{14}C calibration method (filled yellow area). A$_{det}$ is the expected mean ^{14}C activity determination resulting from the ground-truth age distribution according to IntCal20. The following scenarios are considered as sedimentological priors: Panel A: SAR 4 cm ka^{-1}, BD 10 cm; Panel B: SAR 6 cm ka^{-1}, BD 10 cm; Panel C: SAR 8 cm ka^{-1}, BD 10 cm; Panel D: SAR 10 cm ka^{-1}, BD 10 cm; Panel E: SAR 12 cm ka^{-1}, BD 10 cm; Panel F: SAR 14 cm ka^{-1}, BD 10 cm; Panel G: SAR 16 cm ka^{-1}, BD 10 cm; Panel H: SAR 18 cm ka^{-1}, BD 10 cm; Panel I: SAR 20 cm ka^{-1}, BD 10 cm.

represent a *Catch-22* situation, however: we need an approximate indication of the age-depth relationship to determine the SAR prior, but the combination of SAR and Δ^{14}C history can influence the ^{14}C age distribution shape for a particular sediment interval, and hence apparent age, of the sediment archive. It would be prudent, therefore, to test a number of realistic SAR priors and examine the consequences for geochronological interpretation.

It is possible to use an approximate prior for BD using an estimate based on globally representative values (generally between 8 and 12 cm) (Trauth et al. 1997; Boudreau 1998). One could also directly estimate for the sediment archive itself based on ^{14}C investigations of the core top (Peng et al. 1979; Trauth et al. 1997; Henderiks et al. 2002), or by using ^{14}C measurements on single foraminifera (Lougheed et al. 2018) or, more accessibly,

by measuring ^{14}C on a number of samples with low numbers of foraminifera and using a statistical analysis of the sample variation to infer downcore bioturbation depth (Dolman et al. 2021).

The fraction of unpicked, fragmented microfossils can be estimated by simply investigating the sample material (Le and Shackleton 1992). There is a risk, however, that the very oldest microfossils of the original population are completely dissolved and are therefore no longer present in the sample material as broken material (Ruddiman and Heezen 1967), which could affect assumptions regarding the p_{prior} age distribution. In any case, one can take into account the susceptibility of a particular species to breakage (Boltovskoy 1991; Boltovskoy and Totah 1992) in combination with knowledge of bottom water chemistry (Ruddiman and Heezen 1967; Parker and Berger 1971), as well as the average residence time in the bioturbation zone, itself a function of SAR and BD (Lougheed et al. 2020).

Additional challenges are associated with determining temporal changes in species abundance, seeing as the abundance record estimated from the depth domain (i.e., the downcore, discrete-depth record) is itself modified by bioturbation (Lougheed 2020), and therefore does not reflect the original species abundance signal in the time domain. Species abundance in the time domain, which is called for in the calibration protocol outlined here, could be based on an estimate from, e.g., a transient palaeoclimate model run linked to an ecological model (Lombard et al. 2011; Morard et al. 2013; Roche et al. 2018; Metcalfe et al. 2020), although estimating relative temporal abundance of a species using such an approach remains a challenging task. Temporal reconstructions of abundance represent an inherent difficulty for the interpretation not just of ^{14}C chronological data, but downcore, multi-specimen microfossil records in general (Hutson 1980; Boyle 1984; Bard 2001; Löwemark and Grootes 2004; Löwemark et al. 2008; Lougheed 2020). If one is simply not aware of the temporal abundance history at a site, a suitable approach could involve applying multiple plausible abundance scenarios when calibrating ^{14}C dates using the calibration protocol outlined here and examining if the spread of calibrated age outcomes significantly affects the geochronological interpretation. Such an approach is similar to the current state of the art, when one might reasonably experiment with multiple reservoir effect scenarios or calibration curve versions.

CONCLUSION

Current ^{14}C calibration workflows for sediment archives do not incorporate information about sedimentological processes such as SAR and BD, meaning that current ^{14}C-based geochronologies systematically underestimate the total age range of a multi-specimen sample, and potentially also contain age-depth artifacts. By taking into account sedimentological processes in addition to ^{14}C uncertainties, a more credible calibrated age distribution can be ascertained using the protocol outlined here. This new calibration protocol offers most improvement in the case of lower SAR typical of deep-sea sediment archives. It should be noted, however, that SAR itself can influence the age distribution (and hence ^{14}C activity distribution) of a sample, but in order to determine the SAR prior accurately one needs to know the approximate age-depth relationship of the sediment. This *Catch-22* type situation inherently limits high-temporal resolution geochronological analysis of deep-sea sediment, so an exploratory approach involving a range of plausible scenarios could help understand consequences for geochronological interpretation and allow researchers to test the effect of their assumptions. Such an approach can be facilitated by

the computerized implementation (*biocal*) of the calibration protocol presented here, allowing for many scenarios to be rapidly explored. This time-efficient, vectorized computer script could be ported to and included in existing geochronological software packages typically applied to sediment archives (Bronk Ramsey 1995; Haslett and Parnell 2008; Parnell et al. 2008; Blaauw 2010; Blaauw and Christen 2011; Lougheed and Obrochta 2019), thus leading to improved age-depth chronologies, and ultimately improving the accuracy of geochronological interpretation of sediment archives.

ACKNOWLEDGMENTS

SOFTWARE AVAILABILITY

The latest version of *biocal* can be downloaded from https://github.com/bryanlougheed/biocal/ and release versions are permanently archived at https://doi.org/10.5281/zenodo.5787164.

REFERENCES

Abbott PM, Griggs AJ, Bourne AJ, Davies SM. 2018. Tracing marine cryptotephras in the North Atlantic during the last glacial period: protocols for identification, characterisation and evaluating depositional controls. Marine Geology 401:81–97. doi: 10.1016/j.margeo.2018.04.008.

Andree M. 1987. The impact of bioturbation on AMS ^{14}C dates on handpicked foraminifera: a statistical model. Radiocarbon 29(2):169–175. doi: 10.1017/S0033822200056927.

Bard E. 2001. Paleoceanographic implications of the difference in deep-sea sediment mixing between large and fine particles. Paleoceanography 16(3):235–239.

Bard E, Arnold M, Duprat J, Moyes J, Duplessy JC. 1987. Reconstruction of the last deglaciation: Deconvolved records of $\delta^{18}O$ profiles, micro-paleontological variations and accelerator mass spectrometric ^{14}C dating. Climate Dynamics 1(2):101–112.

Barker S, Broecker W, Clark E, Hajdas I. 2007. Radiocarbon age offsets of foraminifera resulting from differential dissolution and fragmentation within the sedimentary bioturbated zone. Paleoceanography 22(2):PA2205. doi: 10.1029/2006PA001354.

Berger WH, Heath GR. 1968. Vertical mixing in pelagic sediments. Journal of Marine Research 26:134–143.

Blaauw M. 2010. Methods and code for 'classical' age-modelling of radiocarbon sequences. Quaternary Geochronology 5(5):512–518.

Blaauw M, Christen JA. 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. Bayesian Analysis. 6:457–474. doi: 10.1214/11-BA618.

Boltovskoy E. 1991. On the destruction of foraminiferal tests (laboratory experiments). Révue de Micropaléontologie 34(1):12–25.

Boltovskoy E, Totah V. 1992. Preservation index and preservation potential of some foraminiferal species. Journal of Foraminiferal Research 22(3):267–273. doi: 10.2113/gsjfr.22.3.267.

Boudreau BP. 1998. Mean mixed depth of sediments: The wherefore and the why. Limnology and Oceanography 43(3):524–526. doi: 10.4319/lo.1998.43.3.0524.

Boyle EA. 1984. Sampling statistic limitations on benthic foraminifera chemical and isotopic data. Marine Geology 58(1–2):213–224. doi: 10.1016/0025-3227(84)90124-5.

Bramlette M, Bradley W. 1942. Geology and biology of North Atlantic deep-sea cores. Part 1. Lithology and geologic interpretations. Prof. Pap. U.S. Geological Survey 196 A:1–34.

Bronk Ramsey C. 1995. Radiocarbon calibration and analysis of stratigraphy; the OxCal program. Radiocarbon 37(2):425–430.

Bronk Ramsey C. 2008. Radiocarbon dating: revolutions in understanding. Archaeometry 50(2):249–275.

Dolman AM, Groeneveld J, Mollenhauer G, Ho SL, Laepple T. 2021. Estimating bioturbation from replicated small-sample radiocarbon ages. Paleoceanography and Paleoclimatology 36(7):e2020PA004142. doi: 10.1029/2020PA004142.

Dolman AM, Laepple T. 2018. Sedproxy: a forward model for sediment archived climate proxies.

Climate of the Past Discussions 1–31. doi: 10.5194/cp-2018-13.

Emiliani C, Milliman JD. 1966. Deep-sea sediments and their geological record. Earth-Science Reviews 1(2–3):105–132. doi: 10.1016/0012-8252(66)90002-X.

Ericson DB, Broecker WS, Kulp JL, Wollin G. 1956. Late-Pleistocene climates and deep-sea sediments. Science. 124(3218):385–389. doi: 10.1126/science.124.3218.385.

Erlenkeuser H. 1980. ^{14}C age and vertical mixing of deep-sea sediments. Earth and Planetary Science Letters 47(3):319–326. doi: 10.1016/0012-821X(80)90018-7.

Guinasso NL, Schink DR. 1975. Quantitative estimates of biological mixing rates in abyssal sediments. Journal of Geophysical Research 80(21):3032–3043. doi: 10.1029/JC080i021p03032.

Haslett J, Parnell A. 2008. A simple monotone process with application to radiocarbon-dated depth chronologies. Journal of the Royal Statistical Society: Series C (Applied Statistics) 57(4):399–418. doi: 10.1111/j.1467-9876.2008.00623.x.

Henderiks J, Freudenthal T, Meggers H, Nave S, Abrantes F, Bollmann J, Thierstein HR. 2002. Glacial–interglacial variability of particle accumulation in the Canary Basin: a time-slice approach. Deep Sea Research Part II: Topical Studies in Oceanography 49(17):3675–3705. doi: 10.1016/S0967-0645(02)00102-9.

Hutson WH. 1980. Bioturbation of deep-sea sediments: oxygen isotopes and stratigraphic uncertainty. Geology 8(3):127–130. doi: 10.1130/0091-7613(1980)8<127:BODSOI>2.0.CO;2.

Keigwin LD, Guilderson TP. 2009. Bioturbation artifacts in zero-age sediments. Paleoceanography 24(4):PA4212. doi: 10.1029/2008PA001727.

Le J, Shackleton NJ. 1992. Carbonate dissolution fluctuations in the western Equatorial Pacific during the Late Quaternary. Paleoceanography 7(1):21–42. doi: 10.1029/91PA02854.

Lombard F, Labeyrie L, Michel E, Bopp L, Cortijo E, Retailleau S, Howa H, Jorissen F. 2011. Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach. Biogeosciences 8(4):853–873. doi: 10.5194/bg-8-853-2011.

Lougheed BC. 2020. SEAMUS (v1.20): a Δ^{14}C-enabled, single-specimen sediment accumulation simulator. Geoscientific Model Development 13(1):155–168. doi: 10.5194/gmd-13-155-2020.

Lougheed BC, Ascough P, Dolman AM, Löwemark L, Metcalfe B. 2020. Re-evaluating ^{14}C dating accuracy in deep-sea sediment archives. Geochronology 2(1):17–31. doi: 10.5194/gchron-2-17-2020.

Lougheed BC, Metcalfe B, Ninnemann US, Wacker L. 2018. Moving beyond the age-depth model paradigm in deep sea palaeoclimate

archives: dual radiocarbon and stable isotope analysis on single foraminifera. Climate of the Past 14:515–526. doi: 10.5194/cp-2017-119.

Lougheed BC, Obrochta SP. 2016. MatCal: open source Bayesian ^{14}C age calibration in Matlab. Journal of Open Research Software 4:e42. doi: 10.5334/jors.130.

Lougheed BC, Obrochta SP. 2019. A rapid, deterministic age-depth modelling routine for geological sequences with inherent depth uncertainty. Paleoceanography and Paleoclimatology 34:122–133. doi: 10.1029/2018PA003457.

Löwemark L, Grootes PM. 2004. Large age differences between planktic foraminifers caused by abundance variations and Zoophycos bioturbation. Paleoceanography 19(2):PA2001. doi: 10.1029/2003PA000949.

Löwemark L, Konstantinou KI, Steinke S. 2008. Bias in foraminiferal multispecies reconstructions of paleohydrographic conditions caused by foraminiferal abundance variations and bio-turbational mixing: a model approach. Marine Geology 256(1–4):101–106. doi: 10.1016/j.margeo.2008.10.005.

Metcalfe B, Lougheed BC, Waelbroeck C, Roche DM. 2020. A proxy modelling approach to assess the potential of extracting ENSO signal from tropical Pacific planktonic foraminifera. Climate of the Past 16(3):885–910. doi: 10.5194/cp-16-885-2020.

Missiaen L, Wacker L, Lougheed BC, Skinner L, Hajdas I, Nouet J, Pichat S, Waelbroeck C. 2020. Radiocarbon dating of small-sized foraminifer samples: insights into marine sediment mixing. Radiocarbon 62(2):313–333. doi: 10.1017/RDC.2020.13.

Morard R, Quillévéré F, Escarguel G, de Garidel-Thoron T, de Vargas C, Kucera M. 2013. Ecological modeling of the temperature dependence of cryptic species of planktonic foraminifera in the Southern Hemisphere. Palaeogeography, Palaeoclimatology, Palaeoecology 391:13–33. doi: 10.1016/j.palaeo.2013.05.011.

Nayudu YR. 1964. Volcanic ash deposits in the Gulf of Alaska and problems of correlation of deep-sea ash deposits. Marine Geology 1(3):194–212. doi: 10.1016/0025-3227(64)90058-1.

Parker FL, Berger WH. 1971. Faunal and solution patterns of planktonic foraminifera in surface sediments of the South Pacific. Deep Sea Research and Oceanographic Abstracts 18(1): 73–107. doi: 10.1016/0011-7471(71)90017-9.

Parnell AC, Haslett J, Allen JRM, Buck CE, Huntley B. 2008. A flexible approach to assessing synchroneity of past events using Bayesian reconstructions of sedimentation history. Quaternary Science Reviews 27(19–20): 1872–1885. doi: 10.1016/j.quascirev.2008.07.009.

Peng T-H, Broecker WS, Berger WH. 1979. Rates of benthic mixing in deep-sea sediment as

determined by radioactive tracers. Quaternary Research 11(1):141–149.

Pisias NG. 1983. Geologic time series from deep-sea sediments: time scales and distortion by bioturbation. Marine Geology 51(1–2):99–113.

Reimer PJ, Austin WEN, Bard E, Bayliss A, Blackwell PG, Ramsey CB, Butzin M, Cheng H, Edwards RL, Friedrich M, et al. 2020. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). Radiocarbon: 1–33. doi: 10.1017/RDC.2020.41.

Roche DM, Waelbroeck C, Metcalfe B, Caley T. 2018. FAME (v1.0): a simple module to simulate the effect of planktonic foraminifer species-specific habitat on their oxygen isotopic content. Geoscientific Model Development 11(9):3587–3603. doi: 10.5194/gmd-11-3587-2018.

Rubin M, Suess HE. 1955. U.S. Geological Survey Radiocarbon Dates 11. Science 121:481–488.

Ruddiman WF, Glover LK. 1972. Vertical mixing of ice-rafted volcanic ash in North Atlantic sediments. Geological Society of America Bulletin 83(9):2817–2836.

Ruddiman WF, Heezen BC. 1967. Differential solution of planktonic foraminifera. Deep Sea Research and Oceanographic Abstracts 14(6):801–808. doi: 10.1016/S0011-7471(67)80016-0.

Schiffelbein P. 1984. Effect of benthic mixing on the information content of deep-sea stratigraphical signals. Nature 311(5987):651. doi: 10.1038/311651a0.

Sepulcre S, Durand N, Bard E. 2017. Large ^{14}C age offsets between the fine fraction and coexisting planktonic foraminifera in shallow Caribbean sediments. Quaternary Geochronology 38:61–74. doi: 10.1016/j.quageo.2016.12.002.

Stuiver M, Pearson GW, Braziunas T. 1986. Radiocarbon age calibration of marine samples back to 9000 cal yr BP. Radiocarbon 28(2B):980–1021.

Trauth MH. 1998. TURBO: a dynamic-probabilistic simulation to study the effects of bioturbation on paleoceanographic time series. Computers & Geosciences 24(5):433–441. doi: 10.1016/S0098-3004(98)00019-3.

Trauth MH, Sarnthein M, Arnold M. 1997. Bioturbational mixing depth and carbon flux at the seafloor. Paleoceanography 12(3):517–526.

Waelbroeck C, Lougheed BC, Vazquez Riveiros N, Missiaen L, Pedro J, Dokken T, Hajdas I, Wacker L, Abbott P, Dumoulin J-P, et al. 2019. Consistently dated Atlantic sediment cores over the last 40 thousand years. Scientific Data 6(1):165. doi: 10.1038/s41597-019-0173-8.

Wheatcroft RA. 1992. Experimental tests for particle size-dependent bioturbation in the deep ocean. Limnology and Oceanography 37(1):90–104.