ARTICLE IN PRESS

Quaternary Geochronology xxx (2010) 1–7

Contents lists available at ScienceDirect

Quaternary Geochronology

journal homepage: www.elsevier.com/locate/quageo

Research Paper

Methods and code for 'classical' age-modelling of radiocarbon sequences

Maarten Blaauw*

School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, United Kingdom

ARTICLE INFO

Article history:
Received 19 February 2009
Received in revised form
22 October 2009
Accepted 27 January 2010
Available online xxx

Keywords:
Radiocarbon age-modelling
Calibration
Software
Exploratory analysis

ABSTRACT

Age–depth models form the backbone of most palaeoenvironmental studies. However, procedures for constructing chronologies vary between studies, they are usually not explained sufficiently, and some are inadequate for handling calibrated radiocarbon dates. An alternative method based on importance sampling through calibrated dates is proposed. Dedicated R code is presented which works with calibrated radiocarbon as well as other dates, and provides a simple, systematic, transparent, documented and customizable alternative. The code automatically produces age–depth models, enabling exploration of the impacts of different assumptions (e.g., model type, hiatuses, age offsets, outliers, and extrapolation).

1. Introduction

Age–depth models are built to estimate the calendar ages of depths in a core from a deposit, based on limited numbers of dated depths and on assumptions as to how the deposit has accumulated between those dated depths. Given the intensive use of age–depth modelling, surprisingly few details have been published about the techniques involved, adequate software seems to be largely missing or under-used, and details of applied age–depth modelling assumptions are often missing in the literature.

A review of all papers published in 2008 in the journals Quaternary Geochronology, Quaternary Science Reviews, Journal of Quaternary Science and The Holocene, indicates 93 papers applying some kind of age–depth model through fossil proxy deposits (Table 1). Of these papers, all 82 that used ^{14}C dates provide details about their calibration procedures (e.g., which calibration curve and software were applied). However, often much less information is supplied about subsequent age–depth modelling (regularly this information can only be gleaned from graphs). Sixty papers do not tell which (if any) point estimates were used, 71 do not mention the software used to produce age–depth curves, 65 do not indicate whether age–depth uncertainties were calculated, and 18 do not specify the applied age–depth model. Eleven papers apply some form of Bayesian age–depth modelling, whereas the remainder use "classical" age–depth modelling such as linear interpolation (31) or linear/polynomial regression (18).

As will be discussed below, calibrated ^{14}C dates often have irregular distributions which cannot be properly reduced to single point estimates with symmetric error bars. However, most age-modelling routines do indeed simplify calibrated distributions to symmetric ones. Bennett (1994) and Bennett and Fuller (2002) discussed basic age–depth models based on linear interpolation, polynomial and spline regression, but based on uncalibrated radiocarbon dates. Proxy graphing software Tilia (Grimm, 1990) produces age–depth models, but *i*) does not calibrate dates, *ii*) does not take into account errors, and *iii*) does not provide age–depth graphs. Heegaard et al. (2005) applied mixed-effects modelling, but assumed that calibrated ^{14}C ages can be approximated by a simpler statistical distribution (e.g., Gaussian). It is likely that studies which do not report their age–depth modelling software often use regression or spline functions which generally cannot deal with irregular distributions. Whereas proxy graphic software psimpoll (Bennett, 1994) can handle calibrated ^{14}C dates, they need to be calibrated outside the software.

Although Bayesian methods can produce very reliable age–depth models (e.g., Blaauw and Christen, 2005; Bronk Ramsey, 2008), Table 1 shows that by far most studies continue to use more basic, 'classical' age-depth models (especially for sites with only a handful of dates). Here these classical approaches will be discussed, and a technique based on probability sampling methods will be proposed in order to take into account the multi-modal and asymmetric nature of calibrated ^{14}C dates. A more systematic and transparent approach to produce classical age–depth models is presented, based on free and open-source software.

* Tel.: +44 28 9097 3895.
E-mail address: maarten.blaauw@qub.ac.uk

1871-1014/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.quageo.2010.01.002

Table 1
Literature analysis of primary literature reporting age–depth models, published in 2008 in Quaternary Geochronology (2 papers), Quaternary Science Reviews (40), Quaternary Research (10), Journal of Quaternary Science (13) and The Holocene (28). Publications citing previously published age–depth models were not taken into account. As several papers applied a number of age–depth models and types of dates, the numbers do not always add up. 17 papers mentioned the removal of dates identified as outlying.

Dates	Point estimate	Model	Model error	Age-model software
^{14}C (82)	Not specified (60)	Linear interpolation (31)	Not specified (65)	Not specified (71)
Tephra (11)	Full distribution (13)	Not specified (18)	2 sd error (17)	Oxcal (6)
^{210}Pb/Cs (9)	Mid (5)	Linear regression (13)	1 sd error (6)	Bpeat (4)
U/Th (8)	Median (4)	Bayesian (11)		Mixed-effect (3)
OSL (5)	Intercept (3)	Linear regression (5)		Bchron (1)
Tuning (4)	Mean (1)	Spline (4)		psimpoll + BCal (1)
Varves (2)	Weighted mean (1)	Mixed-effect (3)		Other (2)
	Mid of most probable range (1)	CRS (2)		
		Other (2)		

2. Radiocarbon calibration

The importance of using calibrated ^{14}C dates cannot be overstated. Owing to past changes in atmospheric ^{14}C concentration, the relationship between ^{14}C and calendar age is far from linear, thus requiring a calibration curve (Reimer et al., 2009). Whereas an uncalibrated ^{14}C date can be assumed to have symmetrical errors (normal distribution), calibration will often result in widened, asymmetrical and multi-peaked calendar age uncertainties (Fig. 1; all calendar ages are expressed in years before AD 1950 or cal BP, unless mentioned otherwise). These asymmetrical distributions can cause problems when applying classical interpolation/regression techniques, as they often assume symmetrically/normally distributed errors.

Standard ^{14}C calibration software generally applies 'probabilistic calibration' to provide calendar age estimates of ^{14}C dates (Stuiver and Reimer, 1989, 1993; van der Plicht and Mook, 1989). In this method, a ^{14}C date with reported mean and error $y \pm sd$, is assumed to have a Gaussian distribution on the ^{14}C age scale; $y \sim N(y, sd)$. Whereas fluctuations in the calibration curve often cause single ^{14}C ages to correspond to multiple calendar ages, every calendar age will always have only one single corresponding ^{14}C age. Therefore, to calibrate a ^{14}C date, we aim to find the probability of that date belonging to a given calendar year. For a range of calendar years θ, the corresponding ^{14}C age of the calibration curve, $\mu(\theta)$, is looked up and compared with the ^{14}C measurement (see Christen, 1994):

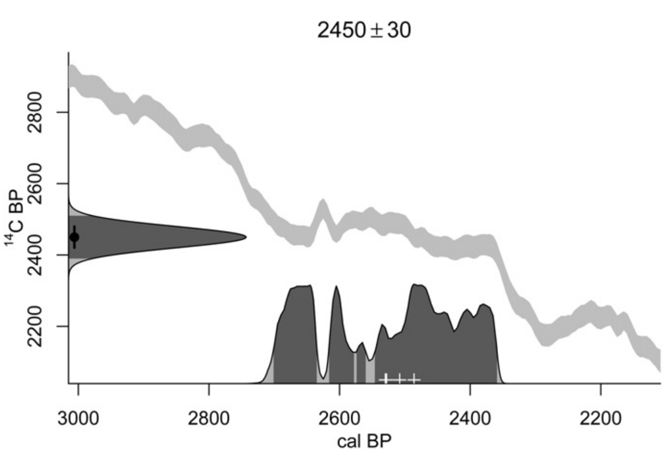

Fig. 1. Calibration of a ^{14}C date of 2450 ± 30 ^{14}C BP at 2 standard deviation. Highest posterior density ranges (dark grey) are 2359–2546 (probability 57.6%), 2560–2574 (2.5%), 2578–2616 (10.6%) and 2635–2701 cal BP (24.7%). Axes of the probability distributions are of arbitrary heights. From left to right, white crosses indicate the date's midpoint, weighted mean, median and mode, respectively. Multiple calendar ages intercept 2450 ^{14}C BP; the chosen intercept had the highest probability when taking the error on the calibration curve into account.

$$y|\theta \sim N\left(\mu(\theta), \sigma^2\right) \quad (1)$$

where | means 'given', and σ is a combination of the measurement uncertainty and that of the calibration curve at the calendar age, $\sigma(\theta)$, $\sigma = \sqrt{sd^2 + \sigma(\theta)^2}$. Usually the total probability over a large range of calendar years is normalised to 100%. Because the resolution of most calibration curves is decadal, calibration at yearly intervals requires (usually linear) interpolation of $\mu(\theta)$ and $\sigma(\theta)$.

The currently internationally accepted calibration curves are IntCal09 (northern hemisphere atmospheric; Reimer et al., 2009), Marine09 (for marine dates; Reimer et al., 2009), and SHCal04 (southern hemisphere atmospheric; McCormac et al., 2004). Alternative curves exist, e.g., for dating organic material deposited after nuclear weapons testing (Hua and Barbetti, 2004; Levin et al., 2008).

In case any of the ^{14}C dates in a sequence is suspected to be prone to a systematic age-offset (e.g., owing to a lake hard-water effect, or to an additional local reservoir effect for marine data; see Reimer and Reimer, 2001, http://calib.org/marine/), this should be corrected for before calibrating. If an uncertainty estimate of the reservoir effect is available, the error of the ^{14}C date should be increased accordingly ($sd = \sqrt{sd_{res}^2 + sd_m^2}$, where sd_{res} is the error of the reservoir effect).

Calibrated distributions are usually reduced to 1 or 2 sd calibrated ranges, taking into account their asymmetric and multi-peaked shapes. In standard calibration software such as CALIB (Stuiver et al., 2005), these ranges are obtained by calculating the highest posterior density (hpd) ranges: *i*) probability distributions are normalised to 100% (Eq. (1)), *ii*) the calendar years are ranked according to their probabilities, *iii*) those calendar ages with a cumulative summed probability at or below the 1 sd (c. 68%) or 2 sd (c. 95%) threshold are retained, and *iv*) the extremes and probabilities of any sub-ranges within these calendar ages are found (Fig. 1).

The probabilities within hpd ranges are either reported as ratios of the entire calibrated distribution (where the hpd ranges add up to say 68% or 95%; OxCal, Bronk Ramsey, 2008), or as ratios of the hpd region only (where the ranges are normalised to 100%, e.g., CALIB, Stuiver et al., 2005). Calibrated ranges at 1 sd will obviously result in narrower confidence intervals, and thus a perceived higher precision, than at 2 sd. However, given the often asymmetric and multi-modal nature of calibrated distributions, the probability that the 'true' calendar date lies outside the 1 sd ranges is considerable (c. 32%). Therefore the use of non-normalised 2 sd calibrated ranges is preferable.

Non-^{14}C dating information, such as a tephra layer of known or estimated calendar age, a historically anchored pollen event, or a modern age for the top of a sequence (e.g., the year of sampling), can provide additional dating points for age-depth modelling. Whenever available, it is advisable to include such data, with an estimate of their chronological uncertainty.

Please cite this article in press as: Blaauw, M., Methods and code for 'classical' age-modelling of radiocarbon sequences, Quaternary Geochronology (2010), doi:10.1016/j.quageo.2010.01.002

3. Age–depth modelling

Having estimated the age distributions of the individual dated depths, the next step is to provide age estimates for all depths in a sequence. This is done by modelling the accumulation of the sequence through time. The researcher will have to decide on the most likely type of age–depth model. For example, a deposit in a stable environment will probably have accumulated with fewer events of hiatuses or accumulation rate changes than one from a more variable environment, and should thus probably be modelled using a smooth age-depth model. The very common age–depth model of linear interpolation between the dated levels (Table 1), assumes that abrupt changes in accumulation rate took place exactly at the dated depths. Although this assumption is often likely to be wrong, linear interpolation often produces what look like plausible age–depth models (Bennett, 1994). In some sequences one or more hiatuses appear present, requiring models through subsets of the data. Additionally, at times age–depth models need to be extrapolated beyond the dated levels (necessarily introducing a large degree of age-model uncertainty for the undated parts of the sequence).

Calibrated distributions and ranges of ^{14}C dates, and resulting uncertainties in age–depth models, often span several centuries. Even so, users often require a single 'best' age–depth model in order to plot proxy values against calendar age. Drawing straight, polynomial, or exponential curves through age–depth data with symmetrical errors is straightforward and can be done automatically in spreadsheets and age-modelling software such as psimpoll (Bennett, 2008), the age–depth code of Heegaard et al. (2005), and Tilia (Grimm, 1990). Similarly, confidence intervals of such interpolation or regression curves can be easily calculated using standard techniques.

However, ^{14}C ages are notorious for their highly asymmetric calibrated distributions, and reducing them to single points is problematic (Telford et al., 2004b; Michczyński, 2007). To take the date of Fig. 1 as example, its midpoint of the 95% hpd ranges lies at 2530 cal BP, a year with rather low calibrated probability. Also other point estimates such as the median (2508 cal BP) or the weighted mean (2528 cal BP) fall at unlikely calendar ages. The year with the highest probability (mode or intercept; 2486 cal BP) is an inadvisable point estimate, owing to *i*) the possibility of multiple intersecting calendar years and *ii*) its sensitivity to small variations in the calibration curve (Telford et al., 2004b). Moreover, the error ranges of all these point estimates (except the midpoint) will often be asymmetric. Therefore classical statistical regression techniques

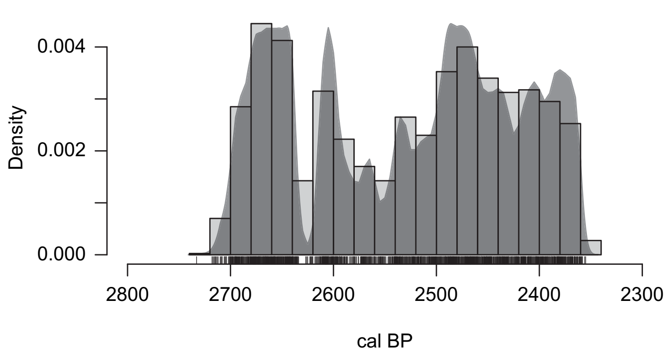

Fig. 2. Two thousand point calendar year estimates (short lines at bottom graph) were sampled from the calibrated distribution of the date of Fig. 1 (dark grey distribution). As explained in the text, calendar years with higher probabilities were sampled more frequently. A histogram of the sampled ages (light grey bars) closely resembles the original calibrated distribution.

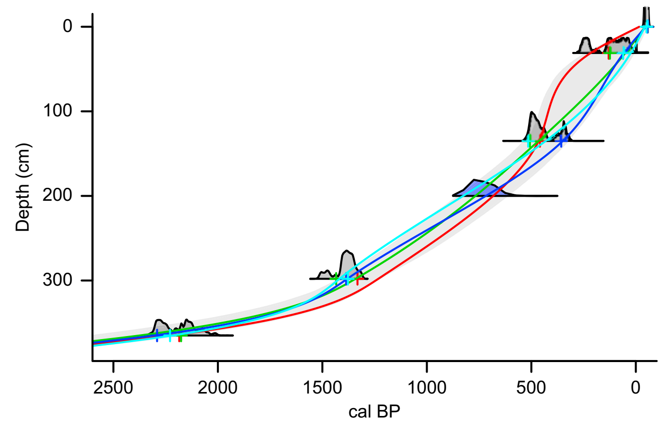

Fig. 3. Schematic drawing of an iterative sampling process in age–depth modelling. Four point estimates (coloured crosses) are sampled from calibrated distributions (grey histograms), and error-weighted smooth splines drawn through the crosses (each colour corresponds to an iteration process). Note that the splines do not necessarily cross each individual age point estimate. Splines were weighted by the calibrated probabilities of the sampled calendar ages. The grey envelope shows the final 95% confidence intervals based on 1000 iterations. The age distribution of all splines at 200 cm depth is indicated by the blue histogram. The shown data are a selection from core Quilichao (Berrio et al., 2002; Fig. 4). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

which depend on symmetric error distributions should not be applied to ^{14}C age–depth modelling.

A proposed alternative to the above point estimates is to use the entire calibrated distribution of the dates. This can be done by repeated random sampling of the calibrated distributions of the dates, each time calculating an age–depth model through the sampled ages. If a large number of calendar ages is drawn from a calibrated distribution $f(\theta)$, with the probability of an age being sampled proportional to its height of the probability distribution at that age (see Eq. (1)), a histogram of these age estimates will form a close approximation of the original calibrated distribution (Fig. 2). This technique is known as importance, bootstrap or Monte Carlo sampling; Bayesian age–depth modelling makes use of comparable though more sophisticated iterative sampling methods.

Similar to estimating calibrated distributions by importance sampling, for a core with multiple dates we can sample ages from each of its calibrated distributions ($f_n(\theta)$ for each of n dated depths), thus resulting in point age estimates for all dated depths (Fig. 3). A curve can then be drawn through these age–depth point estimates, resulting in a calendar age estimate for any depth d (dated or undated). Repeated sampling and subsequent age–depth modelling will thus result in multiple age estimates for depth d, and histograms from a sufficiently large amount of iterations will provide us with a calendar age distribution of this depth, $f_d(\theta)$. The explicit assumption made here is that although the true calendar ages of a core's dated depths will never be known, the type of age–depth curve applied through these unknown points (e.g., a smooth spline) will be a reliable approximation of the unknown true accumulation history of the site. Through repeated sampling of the calibrated distributions, each time calculating its age–depth model, we will thus build up an approximation of the site's accumulation history, based on the dated depths, their uncertainties, and a chosen type of age-model. This procedure can be run in psimpoll (Bennett, 2008) by feeding it age distributions calculated by, e.g., BCal (Buck et al., 1999; for an application see Yeloff et al., 2006). A comparable bootstrap procedure was reported by Higuera et al. (2009).

Heegaard et al. (2005) using mixed-effect modelling to add an extra error term to dated levels, with the aim to include an

Please cite this article in press as: Blaauw, M., Methods and code for 'classical' age-modelling of radiocarbon sequences, Quaternary Geochronology (2010), doi:10.1016/j.quageo.2010.01.002

uncertainty as to whether the reported date of a dated level is a true representation of its real age (e.g., see Scott, 2007 for the variability between repeated ^{14}C measurements on a single piece of wood). However, Heegaard et al. (2005) did not use calibrated distributions and instead assumed that these can be approximated by symmetrical distributions. For calibrated distributions, the "Heegaard" added error could be simulated by sampling say 1000 ^{14}C ages from a reported ^{14}C age ($y \pm sd$), assuming a normal distribution for the ^{14}C measurement. Then each sampled ^{14}C age is calibrated, and a single calendar year is sampled from its calibrated distribution (Eq. (1)). As with the sampling method described above, a histogram of the mixed-effect sampled calendar ages will form an estimate of the date's calibrated distribution. As expected, mixed-effect age-models generally obtain wider confidence limits than models without the added error (data not shown).

Even if all uncertainties have been taken into account, users often need a single 'best' age–depth model in order to produce plots of the core's analysed proxies against age (however, see Blaauw et al., 2007 for plots which include chronological uncertainties of proxies). As explained above, there are considerable problems with age-depth models through point estimates based on the midpoints, (weighted) mean, median or mode of calibrated dates. Even more, the considerable scatter between closely spaced dates, as well as between dates and the applied model, indicates that the accuracy of ^{14}C measurements should not be overestimated (see Scott, 2007 for the variable outcome of repeated ^{14}C dating of a single sample). As an alternative to drawing 'best' age–depth models through point estimates from individual dates, I propose to use the point age–depth estimates derived from the age–depth model itself, as it is based on all dating information and contains an interpretation of a likely accumulation history for a site. A robust point estimate for any depth in a core could then be provided by the weighted mean of all age-model iterations at that depth.

4. Alternative age-modelling software

Here software is presented which performs the methods detailed above simultaneously and in an automated and documented manner. The approach can be used to test the results of different age-model choices, and to systematically and rapidly produce publication-quality graphs in addition to files containing information about the settings, results, and quality of the obtained age-model. The code produces non-Bayesian, 'classical' age–depth models, and therefore is called *clam*. It can also be used to calibrate single ^{14}C dates.

Clam has been written in the open-source statistical environment R (R Development Core Team, 2010). R can be downloaded freely at http://www.r-project.org for Windows, Mac and Linux (current version 2.10.1, c. 25 MB). The software makes use of a command-window, which may not appear very user-friendly at first sight. However, for *clam* the number of commands to be typed is limited. Moreover, this approach enables a highly customizable environment. R is relatively easy to learn and is being used for a growing number of applications in palaeo-research (Telford et al., 2004b; Blaauw and Christen, 2005; Heegaard et al., 2005; Blaauw et al., 2007, 2010; Haslett and Parnell, 2008). R is similar to S and *clam* could thus also be used in S-PLUS (Insightful Corp.). *Clam* comes with a manual which contains information for installing and using the code (see Section 8).

In short, *clam* calibrates the ^{14}C dates (calendar scale dates such as tephras of known calendar age can be used as well), constructs an age-model based on decisions made by the user, and produces graphs and text files. Provided calibration curves are IntCal09, Marine09 (both Reimer et al., 2009) and SHCal04 (McCormac et al., 2004), while other curves such as post-bomb curves can be added. After correcting for any reservoir effect, all ^{14}C dates of a sequence are calibrated individually through calculating their calibrated distributions for every calendar year as in Eq. (1) (default in cal BP; cal BC/AD as alternative). In case a ^{14}C date is suspected to be outlying, it can be removed manually and is then not taken into account in the age-model. Any ^{14}C dates which fall outside the ranges of the calibration curve will not be used, while dates which fall partly beyond its ranges will be truncated (in both cases a warning will be given). Highest posterior density probabilities are not normalised (this differs from e.g., CALIB; see Section 2). Depending on the core and chosen settings, calculations take up to c. 10–20 s on a modern PC.

Age–depth models are constructed using the calibrated distributions described above. Several common types of age-models are provided here; *i*) linear interpolation between the dated levels, *ii*) linear or higher order polynomial regression, and *iii*) cubic, smoothed or locally weighted splines (the latter two can optionally be weighted by the inverse and squared dating errors). Sometimes the stratigraphy or the ^{14}C dates indicate the presence of hiatuses (Blaauw and Christen, 2005); *clam* comes with an option to include one or more hiatuses. Extrapolation beyond the ^{14}C dated levels is inadvisable but can be done if required.

Confidence intervals for the undated levels are calculated using the Monte Carlo approach detailed above: *i*) From the calibrated age distribution of every dated depth a calendar age is sampled (with the probability of a year being chosen proportional to its calibrated probability $p(\theta)$ as calculated from Eq. (1)); *ii*) an age-model is drawn through these points, and *iii*) for every depth (default 1 cm steps) the resulting calendar age is stored. Per default, curves with age-reversals are rejected with a warning. In step *i*, an extra error term can be included akin to the mixed-effect models described in Heegaard et al. (2005; here the size of the extra error is taken to be equal to a date's reported error). The sampling process is repeated often (default 1000 times, which suffices for most initial runs; use e.g., 10 times more iterations for final runs), resulting in as many calendar age estimates for every depth. For every depth, its calendar age point estimates are binned, ranked, normalised and summed, after which the range of those binned calendar ages which fall within e.g., 68% (1σ) or 95% (2σ) forms an estimate of the confidence interval for every individual depth. As the calculations are based on random iterations, a slightly different confidence interval will be estimated for every run.

To find the single 'best' age–depth model from all iterated age–depth models, *clam* calculates for every depth the weighted mean of all sampled calendar ages (default), or alternatively, the midpoint of the highest posterior density range of all calendar ages (default at 95% confidence intervals). In order to facilitate comparisons with commonly used although inadvisable point estimates (Table 1), *clam* can provide age–depth models based on *i*) the modes or intercepts (highest $p(\theta)$ of Eq. (1)), *ii*) the medians, *iii*) the weighted means, or *iv*) the midpoints of the highest posterior densities. The product of the probabilities $p(\theta)$ (see Eq. (1)) of the calendar age estimates of the dated levels provides us with a measure of the model's goodness-of-fit. If age-reversals occur in the chosen age-model, a warning is given.

5. Case studies

Telford et al. (2004a) tested the accuracy of age–depth modelling techniques by modelling the ^{14}C dates of a sediment sequence dated independently through varve counting. Here a slightly different approach is taken, simulating the accumulation of sediment over time and its subsequent ^{14}C dating and age–depth modelling (Fig. 4). Sediment starts to accumulate at a rate of 20 yr/cm for the upper depth. Each depth further down has an accumulation rate differing slightly from its overlying depth, by sampling from a normal distribution with the previous accumulation rate as mean. The standard

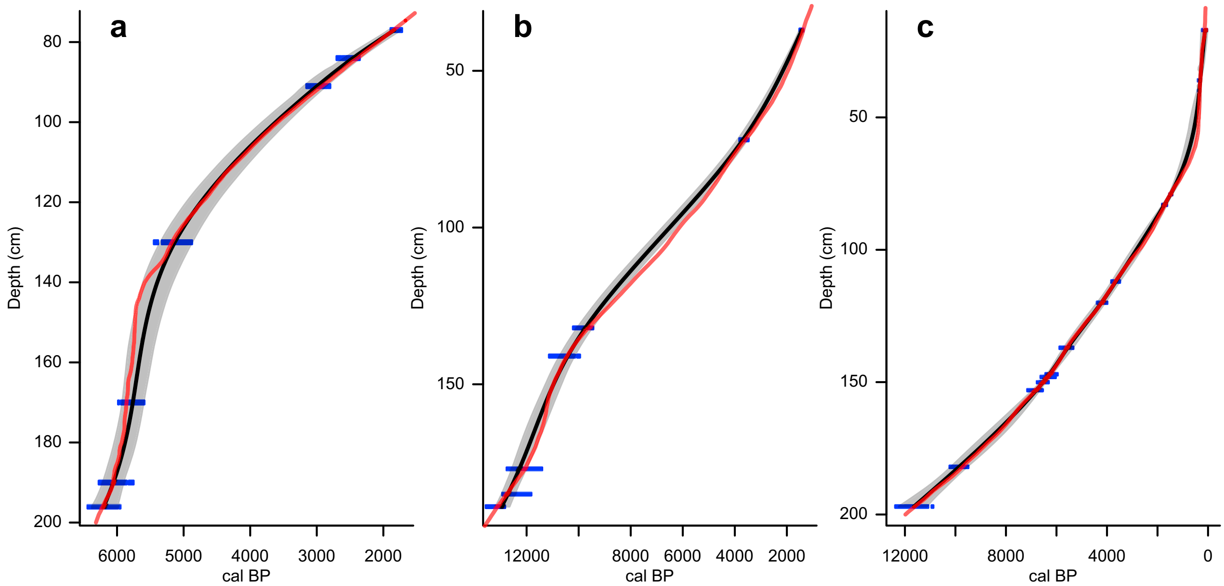

Fig. 4. Simulated accumulation (red curve; see text) and ^{14}C dating (blue blocks show 95% hpd ranges) of a sediment record, followed by *clam* age–depth modelling using a smooth spline (error-weighted with smoothness set at the default of 0.3; grey envelopes show 95% confidence intervals). The first two simulations used 7 dates, while the last one has 14 dated levels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

deviation of this distribution dictates the variability in accumulation rate (here set at 7). The cumulative sum of all simulated accumulation rates then forms the accumulation history of the core. Next, a number of depths are sampled randomly from all depths (using a uniform distribution), each of which stems from a known "true" calendar year. For each of these depths, the ^{14}C year corresponding to its calendar year is found in the IntCal09 calibration curve (Reimer et al., 2009). After adding random noise to this ^{14}C age by sampling from a normal distribution with the ^{14}C age as mean and 30 year as standard deviation, this value serves as a 'de-calibrated' ^{14}C age (Telford et al., 2004a). Finally *clam* calculates an age-depth model through the ^{14}C dates and their depths (Fig. 4).

Generally, the age–depth models produced by *clam* seem to provide reasonable approximations of the true (simulated) accumulation history. Occasionally the true curve lies outside the calculated confidence intervals (Fig. 4b), but these deviations tend to occur at sections of low dating density. With more dates, the age–depth models tend to follow the true history more closely (Fig. 4c).

The effects of different age-modelling choices can be checked from a case-study; core Quilichao from a lake in Colombia (Berrio et al., 2002). This core with 9 ^{14}C dates, and an assumed age of AD 2000 ± 5 for its surface, could be age-modelled assuming linear interpolation together with extrapolation beyond the dated levels to 800 cm (Fig. 5a). Between 485 and 445 cm depth, a large jump in ^{14}C ages appears, suggesting a hiatus. In Fig. 5b, the age-model is linear regression, assuming a hiatus at 477 cm depth (the level where pollen spectra change abruptly; Berrio et al., 2002). In Fig. 5c, the same hiatus is assumed, and the model is based on a smooth spline. Fig. 5d, a smooth spline is drawn through all dates except the one at 445 cm which is considered an outlier.

The age-models in this example were chosen rather arbitrarily and for illustration purposes only (although the reader is reminded that at least some age–depth models in the recent literature appear to have been chosen in a similarly arbitrary way; see Table 1). Users should apply those settings which are deemed adequate for the archive in question (such as deciding between linear interpolation and polynomial or spline regression as the most likely model given the site's deposition regime and/or stratigraphy).

6. Discussion

The software described here does not pretend to offer the latest advances in ^{14}C modelling. Rather it attempts to provide an easy, automated, transparent, documented and adaptable environment for producing age-models from ^{14}C sequences. Indeed, given the current suboptimal approaches and lack of information in most published age–depth models (Table 1), it is hoped that *clam* could aid in standardising and enhancing future 'classical' age–depth models.

Other software exists for producing age-models of ^{14}C dated sequences. In psimpoll (Bennett, 2008) one can choose from a large number of age-models, but ^{14}C dates are left uncalibrated or need to be calibrated outside the program (e.g., in BCal; Buck et al., 1999), and hiatuses cannot be included. Tilia (Grimm, 1990) is popular for age–depth model construction, but cannot handle hiatuses or asymmetric dates, and does not take into account the errors of dates. Heegaard et al. (2005) provide R code for ^{14}C age-modelling based on mixed-effect regression, but there calibrated ^{14}C dates are assumed to have a normal distribution on the calendar scale, hiatuses are not taken into account, and only one type of age-model can be chosen. Blaauw and Christen (2005) published Bayesian software Bpeat for 'wiggle-matching' ^{14}C dates based on the age-model of piece-wise linear accumulation; other age-models are not included in their program. OxCal (Bronk Ramsey, 2008) provides a number of age–depth models based on Bayesian statistics, and recently Bayesian software based on a monotonous accumulation model has been published (Haslett and Parnell, 2008).

The Bayesian age-modelling routines mentioned in the previous paragraph (Blaauw and Christen, 2005; Bronk Ramsey, 2008; Haslett and Parnell, 2008) use advanced, robust and flexible numerical methods, and should be the preferred approach for future age–depth models (Blockley et al., 2007). However, we have seen that many studies continue to apply 'classical' age–depth modelling routines, and existing Bayesian age–depth modelling methods might not add much extra value to low-resolution dated sites. Therefore, the intended use of *clam* is to rapidly, transparently and systematically produce age-models, while exploring the impact of different modelling choices (e.g., type of age-model, presence or

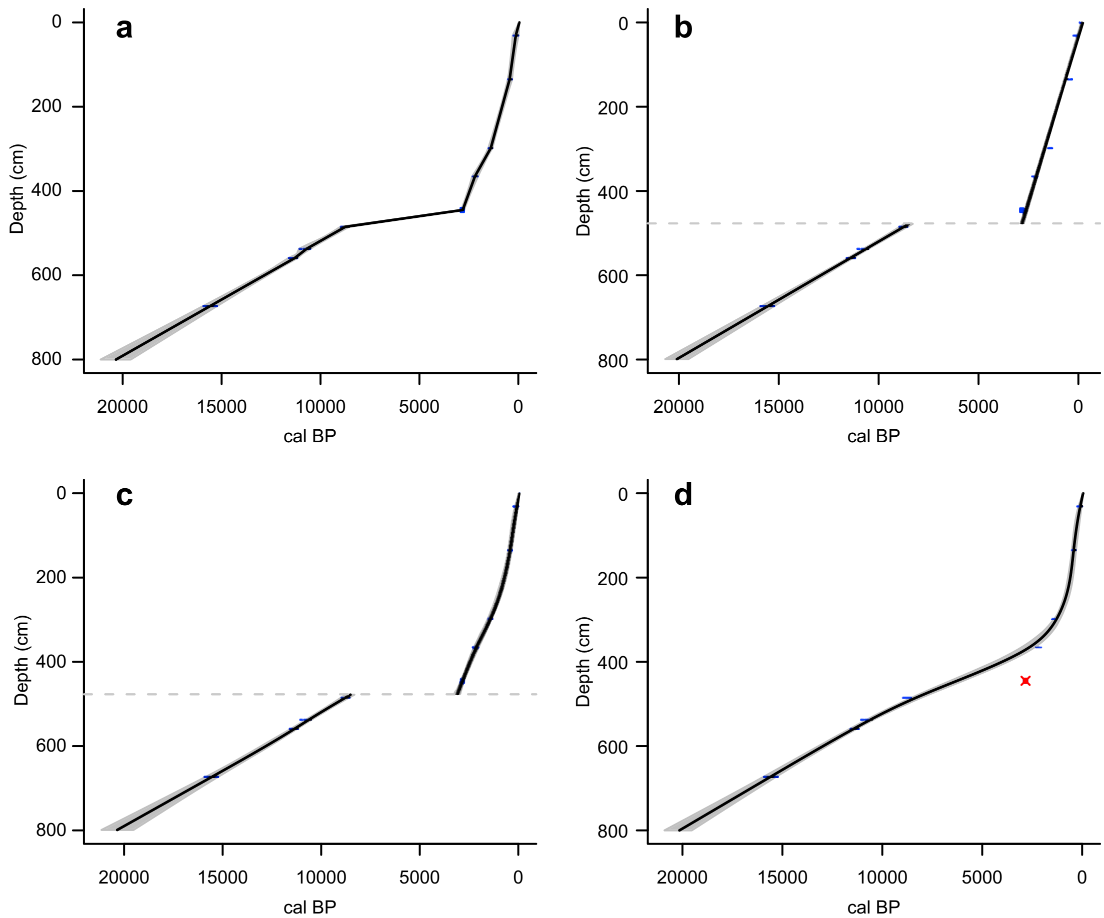

Fig. 5. Age-models of core Quilichao (Berrio et al., 2002). The age-models are extrapolated down to 800 cm. *a*, age-model based on linear interpolation between the dates levels. Between 485 and 445 cm depth, a large jump in ^{14}C ages appears, suggesting a hiatus. *b*, linear regression age-model, assuming a hiatus (horizontal dashed grey line) at 477 cm depth. *c*, the same hiatus is assumed, now using an error-weighted smooth spline (default smoothing parameter 0.3) as age-model. The core surface was assumed to be of recent age (AD 2000 ± 5). *d*, as in c but without hiatus and regarding the date at 445 cm date as outlying. Colours as in Fig. 4.

absence of a hiatus, reservoir effect, outliers, or anchor points). As a next step in the age-modelling process, more sophisticated approaches relevant for the core in question should be considered (Blockley et al., 2007).

It is vital to note the dangers of reducing calibrated ^{14}C dates to single-year point estimates and subsequent drawing of a single age–depth curve through these points. Point age estimates of calibrated ^{14}C dates and undated levels should not be regarded as the truth (e.g., see Telford et al., 2004b; Michczyński, 2007; Blaauw et al., 2007; Scott, 2007). Rather much attention should be paid to the estimated uncertainties for both the calibrated dates and the entire sequence.

The R code of *clam* can be understood relatively easily. Indeed, users are invited to adapt the code to their preferences, such as the figure settings. The open nature of the code enables the interested user to open the 'black box' of age-modelling, following in detail how ^{14}C dates can be calibrated and age-models derived. With this in mind, *clam* is also of use for educational purposes.

Acknowledgements

The ideas for *clam* were born during a stay at the Centro de Investigación en Matemáticas (CIMAT), Mexico. I thank Andrés Christen for introducing me to R, Juan Carlos Berrio for providing the Quilichao data, and Bas van Geel, Hans van der Plicht, Dmitri Mauquoy, Paula and Ron Reimer, Keith Bennett, Andrés Christen and Maytee Segura for stimulating discussions about age-modelling. Heather Binney and many more researchers and students are thanked for reporting bugs and desired features while testing *clam* on their cores.

Appendix. Supplementary information

Supplementary information associated with this article can be found, in the online version, at doi:10.1016/j.quageo.2010.01.002 and http://chrono.qub.ac.uk/blaauw/.

References

Bennett, K.D., 1994. Confidence intervals for age estimates and deposition times in late-Quaternary sediment sequences. The Holocene 4, 337–348.

Bennett, K.D., Fuller, J.L., 2002. Determining the age of the mid-Holocene *Tsuga canadensis* (hemlock) decline, eastern North America. The Holocene 12, 421–429.

Bennett, K.D., 2008. Psimpoll v. 4.263. http://chrono.qub.ac.uk/psimpoll/psimpoll.html (last accessed 17.02.2009.).

Berrio, J.C., Hooghiemstra, H., Marchant, R., Rangel, O., 2002. Late-glacial and Holocene history of the dry forest area in the south Colombian Cauca Valley. Journal of Quaternary Science 17, 667–682.

Blaauw, M., Christen, J.A., 2005. Radiocarbon peat chronologies and environmental change. Applied Statistics 54, 805–816.

Blaauw, M., Christen, J.A., Mauquoy, D., van der Plicht, J., Bennett, K.D., 2007. Testing the timing of radiocarbon-dated events between proxy archives. The Holocene 17, 283–288.

Blaauw, M., Wohlfarth, B., Christen, J.A., Ampel, L., Veres, D., Hughen, K.A., Preusser, F., Svensson, A., 2010. Were last glacial climate events simultaneous between Greenland and France? A quantitative comparison using non-tuned chronologies. Journal of Quaternary Science 25, 387–394.

Blockley, S.P.E., Blaauw, M., Bronk Ramsey, C., van der Plicht, J., 2007. Assessing uncertainties in age modelling sedimentary records in the Lateglacial and Early Holocene. Quaternary Science Reviews 26, 1915–1926.

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

Buck, C.E., Christen, J.A., James, G.N., 1999. BCal: an online Bayesian radiocarbon calibration tool. Internet Archaeology 7. http://intarch.ac.uk/journal/issue7/buck/.

Christen, J.A., 1994. Bayesian Interpretation of Radiocarbon Results. PhD thesis, University of Nottingham, 267 pp.

Grimm, E.C., 1990. TILIA and TILIA*GRAPH. PC spreadsheet and graphics software for pollen data. INQUA Working Group on Data-Handling Methods Newsletter 4, 5–7.

Haslett, J., Parnell, A., 2008. A simple monotone process with application to radiocarbon-dated depth chronologies. Applied Statistics 57, 1–20.

Heegaard, E., Birks, H.J.B., Telford, R.J., 2005. Relationships between calibrated ages and depth in stratigraphical sequences: an estimation procedure by mixed-effect regression. The Holocene 15, 1–7.

Higuera, P.E., Brubaker, L.B., Anderson, P.M., Hu, F.S., Brown, T.A., 2009. Vegetation mediated the impacts of postglacial climate change on fire regimes in the south-central Brooks Range, Alaska. Ecological Monographs 79, 201–219.

Hua, Q., Barbetti, M., 2004. Review of tropospheric bomb ^{14}C data for carbon cycle modeling and age calibration purposes. Radiocarbon 46, 1273–1298.

Levin, I., Hammer, S., Kromer, B., Meinhardt, F., 2008. Radiocarbon observations in atmospheric CO_2: determining fossil fuel CO_2 over Europe using Jungfraujoch observations as background. Science of the Total Environment 391, 211–216.

McCormac, F.G., Hogg, A.G., Blackwell, P.G., Buck, C.E., Higham, T.F.G., Reimer, P.J., 2004. SHCal04 southern hemisphere calibration, 0–11.0 cal kyr BP. Radiocarbon 46, 1087–1092.

Michczyński, A., 2007. Is it possible to find a good point estimate of a calibrated radiocarbon date? Radiocarbon 49, 393–401.

van der Plicht, J., Mook, W.G., 1989. Calibration of radiocarbon ages by computer. Radiocarbon 31, 805–816.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. http://www.R-project.org (last accessed 01.02.2010.).

Reimer, P.J., Reimer, R.W., 2001. A marine reservoir correction database and on-line interface. Radiocarbon 43, 461–463.

Reimer, P.J., Baillie, M.G.L., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Buck, C.E., Burr, G.S., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Hajdas, I., Heaton, T.J., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., McCormac, F.G., Manning, S.W., Reimer, R.W., Richards, D.A., Southon, J.R., Talamo, S., Turney, C.S.M., van der Plicht, J., Weyhenmeyer, C.E., 2009. IntCal09 and Marine09 radiocarbon age calibration curves, 0–50,000 years cal BP. Radiocarbon 51, 1111–1150.

Scott, E.M., 2007. Radiocarbon dating: sources of error. In: Elias, S.A. (Ed.), Encyclopedia of Quaternary Science. Elsevier, Oxford, pp. 2918–2923.

Stuiver, M., Reimer, P.J., 1989. Histograms obtained from computerized radiocarbon age calibration. Radiocarbon 31, 817–823.

Stuiver, M., Reimer, P.J., 1993. Extended ^{14}C database and revised CALIB 3.0 ^{14}C age calibration program. Radiocarbon 35, 215–230.

Stuiver, M., Reimer, P.J., Reimer, R.W., 2005. CALIB 5.0.2html. http://calib.org (last accessed 22.10.2009.).

Telford, R.J., Heegaard, E., Birks, H.J.B., 2004a. All age-depth models are wrong: but how badly? Quaternary Science Reviews 23, 1–5.

Telford, R.J., Heegaard, E., Birks, H.J.B., 2004b. The intercept is a poor estimate of a calibrated radiocarbon age. The Holocene 14, 296–298.

Yeloff, D., Bennett, K.D., Blaauw, M., Mauquoy, D., Sillasoo, Ü., van der Plicht, J., van Geel, B., 2006. High precision C-14 dating of Holocene peat deposits: a comparison of Bayesian calibration and wiggle-matching approaches. Quaternary Geochronology 1, 222–235.

Please cite this article in press as: Blaauw, M., Methods and code for 'classical' age-modelling of radiocarbon sequences, Quaternary Geochronology (2010), doi:10.1016/j.quageo.2010.01.002