Bayesian Analysis (2011) **6**, Number 3, pp. 457–474

Flexible Paleoclimate Age-Depth Models Using an Autoregressive Gamma Process

Maarten Blaauw* and J. Andrés Christen†

Abstract. Radiocarbon dating is routinely used in paleoecology to build chronologies of lake and peat sediments, aiming at inferring a model that would relate the sediment depth with its age. We present a new approach for chronology building (called "Bacon") that has received enthusiastic attention by paleoecologists. Our methodology is based on controlling core accumulation rates using a gamma autoregressive semiparametric model with an arbitrary number of subdivisions along the sediment. Using prior knowledge about accumulation rates is crucial and informative priors are routinely used. Since many sediment cores are currently analyzed, using different data sets and prior distributions, a robust (adaptive) MCMC is very useful. We use the t-walk (Christen and Fox, 2010), a self adjusting, robust MCMC sampling algorithm, that works acceptably well in many situations. Outliers are also addressed using a recent approach that considers a Student-t model for radiocarbon data. Two examples are presented here, that of a peat core and a core from a lake, and our results are compared with other approaches.

Keywords: AR Gamma process; Radiocarbon; Paleoecology; Age-Depth models

Past climates and environments can be reconstructed from deposits such as ocean or lake sediments, ice sheets and peat bogs. Within a vertical sediment profile (core), measurements of microfossils, macrofossils, isotopes and other variables at a range of depths serve as proxy estimates or "proxies" of climate and environmental conditions when the sediment of those depths was deposited. It is crucial to establish reliable relationships between these depths and their ages. Age-depth relationships are used to study the evolution of climate/environmental proxies along sediment depth and therefore through time (e.g., Lowe and Walker 1997).

Age-depth models are constructed in various ways. For sediment depths containing organic matter, and for ages younger than c. 50,000 years, radiocarbon dating is often used to create an age-depth model. Cores are divided into slices and some of these are radiocarbon dated. A curve is fitted to the radiocarbon data and interpolated to obtain an age estimate for every depth of the core. The first restriction to be considered is that age should be increasing monotonically with depth, because sediment can never have accumulated backwards in time (extraordinary events leading to mixed or reversed sediments are, most of the time, noticeable in the stratigraphy and therefore such cores are ruled out from further analyses). Moreover, cores may have missing sections, leading to flat parts in the age depth models.

*School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, Belfast, UK, mailto:maarten.blaauw@qub.ac.uk

†Centro de Investigación en Matemáticas (CIMAT), Guanajuato, Mexico, http://www.cimat.mx/~jac

 DOI:10.1214/11-BA618

For peat cores the problem was studied in Blaauw and Christen (2005) using a piece-wise linear approach with few sections and considering constraints on accumulation rates. This approach has been successfully applied in many examples (e.g., Wohlfarth et al. 2006; Blaauw et al. 2007a; Chambers et al. 2007; Sillasoo et al. 2007; Blaauw et al. 2007b, 2010; Charman et al. 2009; Plunkett et al. 2009). This approach seems useful if high resolution dating is available (many core slices are dated) and the core is not very long (1 or 2 m). Otherwise the piece-wise linear model appears too restrictive and many sections are needed to be used to render the methodology more useful (Wohlfarth et al. 2008; Blaauw et al. 2010). Also, in piece-wise linear models with relatively long sections, the variability decreases in between section vertices (Yeloff et al. 2006; Blockley et al. 2007) (an axes or turning point effect, see Blaauw and Christen 2005, fig. 1(c)).

Haslett and Parnell (2008) address the age-depth chronology building in a new perspective. Based on a bivariate monotone Markov process (with gamma increments) they construct an increasing model with increments determined at random points provided by a renewal process. A linear interpolation is considered in between renewal points, therefore leading to a random number of sections. The process is also proven to be mean-square continuous, thus signifying a palatable and flexible model for age-depth relationships. Haslett and Parnell (2008) represent a workout improvement on the piece-wise linear approach of Blaauw and Christen (2005). However, in our perspective, Haslett and Parnell (2008) concentrate on the sound stochastic properties of their model and consider "minimal assumptions on smoothness" as a basic building block.

Inspired by the ideas of Haslett and Parnell (2008), we concentrate on modeling the accumulation process using a simple autoregressive time series in which smoothness is part of the controlling features, since we believe exactly the opposite: Prior input on the evolution and shape of accumulation rates (smoothness) is crucial in building realistic age-depth models. Because radiocarbon dates are costly (ca. 500USD per determination) and not every depth contains sufficient dateable carbon, not every depth of a core may be dated. Interpolating is the key issue here, and given the non-monotonic nature of the radiocarbon calibration curve (Reimer et al. 2009), it is common to find entire sections of a core where dates increasing in depth have overlapping calibrated distributions over several centimeters. A multitude of non-decreasing models may be imputed. However, available prior information on accumulation rates render many highly unlikely options from an ecological perspective, therefore leading to reduced uncertainty and more realistic modeling. See Section 2 for more details and discussion.

Other general purpose approaches to monotone regression could be attempted here, such as isotonic regression (e.g. Lavine and Mokus 1995; Cai and Dunsun 2007; Holmes and Heard 2003). However, radiocarbon dates are modeled in a non–standard fashion (a calibration is required, with added non-gaussian noise on the calendar scale) and it would be very important to have a means of entering the prior information available regarding the accumulation process just mentioned. It is not clear how an off-the-shelf monotonic regression approach can handle these issues and therefore tailor suited methodologies have been attempted, here and elsewhere (Christen et al. 1995; Blaauw et al. 2003; Blaauw and Christen 2005; Haslett and Parnell 2008; Bronk Ramsey 2008).

Here we concentrate on modeling accumulation rates to establish a coherent evolution of deposition along the core depth, and possibly including some sections where discontinuities are suspected (hiatuses). This is done via a gamma autoregressive process, where the corresponding integrated process represents the age-depth model. The resulting methodology can be used in high-density dated cores, where in some cases the piece-wise linear model with few sections of Blaauw and Christen (2005) is visually reproduced, although the age-depth model variance along the core is better estimated. Our methodology also produces reasonable interpolations in low-density dated cores, and since accumulation rates are inferred taking into account ecological prior information, only sensible alternatives are relevant, leading to better suited inferences. We show our methodology to work in cores from peat as well as from lake deposits (see Section 2). Our methodology uses MCMC (as well as Blaauw and Christen 2005; Haslett and Parnell 2008), but we use a self adjusting algorithm (Christen and Fox 2010) that could potentially enable its use by non-experts. Moreover, we use a parsimonious and novel approach to outliers in radiocarbon dating that has been shown to provide excellent results (Christen and Pérez 2009).

1 The model

We have a series of radiocarbon determinations $y_j \pm \sigma_j; j = 1, 2, \ldots, m$ taken along a (peat, lake, etc.) core at depths d_j. A semiparametric model is proposed to establish a relationship between the (unknown) age of peat and depth, d,

$$G(d, \theta, x) = \theta + \sum_{j=1}^{i} x_j \Delta c + x_{i+1}(d - c_i);$$

where $c_i \leq d < c_{i+1}, i < K$, and $c_0 < c_1 < \cdots < c_K$ are depths uniformly spaced along the peat core with difference Δc and $x = (x_1, x_2, \ldots, x_K)$. That is, in our model the core is divided into K equally spaced sections and x_j is the accumulation rate of section j. (The actual physical slicing of the core may or may not differ from the one presented here for our age-depth modeling purposes.)

Radiocarbon dating is an analytic technique for dating organic matter younger than c. 50,000 years. The radioactive carbon isotope ^{14}C is produced in the upper atmosphere and $^{14}CO_2$ becomes part of the biosphere. When organisms die, their organic matter is subtracted from the biosphere and its ^{14}C content decreases gradually owing to radioactive decay. By measuring and comparing the ^{14}C content of such organic matter with a modern standard, an approximate date is obtained in the form of a mean "radiocarbon age" and a standard measuring error, $y \pm \sigma$. However, the atmosphere's ^{14}C content has not remained constant through time and calibration curves $\mu(\cdot)$ have been constructed to translate calendar (true) years into radiocarbon years (Reimer et al. 2009). We assume that $E[y_j \mid d_j, x] = \mu(G(d_j, \theta, x))$. The specific model for the radiocarbon determinations y_j will be explained below.

A non-Gaussian autoregressive model is proposed for the sediment accumulation rates $x_j = w x_{j+1} + (1-w)\alpha_j$, where $w \in [0, 1]$ and $\alpha_j \sim Gamma(a_\alpha, b_\alpha)$ iid, with a_α and

b_α known, representing the *a priori* information available on accumulation rates. That is, the accumulation rate at depth c_j is a weighted average of the accumulation rate of the previous (below) depth c_{j+1} and an independent noise α_j. The value of w provides the "memory" or coherence in accumulation rates along the core, from independent $w = 0$ to fixed $w = 1$. This in turn provides a degree of smoothness, avoiding *a priori* extreme variations or extremely low or high accumulation rates. Since $x_j > 0$, the resulting age-depth model is increasing. A prior distribution for the intercept θ, $f(\theta)$, is also assumed.

Note that the x_j process may be written as $x_j = wx_{j+1} + z_j$, where $z_j \sim Gamma(a_\alpha, b_\alpha/(1-w))$ are iid "innovations". It may be seen that this process (given w, a_α and b_α) is stationary and that the stationary distribution corresponds to a self-decomposable (SD) random variable X, with the property $X =^d wX + Z$, with $Z \sim Gamma(a_\alpha, b_\alpha/(1-w))$(Mena and Walker 2005). It is also known that such stationary distribution is infinitely divisible and unimodal (Mena and Walker 2004). This process can also be seen as a discrete approximation of the continuous time Ornstein–Uhlenbeck process with gamma innovations or OU-Γ. Barndorff-Nielsen and Shephard (2003) study these types of processes and properties of the corresponding integrated process that would represent here the actual age-depth model. Much of the literature of non-Gaussian autoregressive processes has concentrated on controlling the resulting stationary distribution and thus finding the required distribution for the innovations (e.g. Mena and Walker 2005). We model the actual core accumulation process itself, and therefore the distribution for z_j is fixed (to be elicited as part of the prior information) and, as we will explain (see Section 2), this proves to be crucial in establishing age-depth models. This constitutes our key modeling feature, since it forms a better basis for eliciting informed priors, as opposed to having purely instrumental modeling parameters with no clear physical meaning to control the behaviour of possible age-depth models.

Barndorff-Nielsen and Shephard (2003) calculate the Laplace transform of the stationary distribution of the OU-Γ process, however, we have not been able to find an analytic version of the corresponding stationary distribution (given w, a_α and b_α) for our model. Barndorff-Nielsen and Shephard (2001) explain that a square integrable stationary OU process has an autocorrelation function that decreases exponentially with the lag. By first principles one can calculate $E(x_n x_{n-k})$ and noting that (by the SD property) $E(x_n)(1-w) = E(Z)$ one obtains the autocorrelation between x_n and x_{n-k}

$$\rho(k) = w^k,$$

which is independent of the distribution for z_j. Regarding the prior distribution for w, note that this depends on Δc, that is, how separated in depth x_j and x_{j+1} are. In order to have a standardized prior distribution for the "memory" in our model, independent of the particular Δc chosen, we note that $c(d) = \rho(d/\Delta c) = w^{\frac{d}{\Delta c}}$ represents the autocorrelation at a lag corresponding to a depth of d. Let d_s be a fixed depth and let $R = c(d_s)(\in [0,1])$ be the correlation of accumulation rates of any two sections of the core separated by d_s cm. What remains is to elicit the prior distribution of R, and we assume it to be $Beta(a_w, b_w)$. It is clear that the corresponding *a priori* density for

w is

$$f(w) = \frac{d_s w^{\frac{d_s}{\Delta c}-1}}{\Delta c} f_R(w^{\frac{d_s}{\Delta c}}),$$

where $f_R(\cdot)$ is the beta prior density of R with parameters (a_w, b_w). We let $d_s = 1$ cm in the following. This will be the prior distribution of w to be used in our model.

It is common in paleoecology to find stratigraphical indications for abrupt changes, such as a shift from lake to bog deposits, or a discontinuity in accumulated deposits. Since different types of deposits could have accumulated with unique accumulation rates or "memory", in our model specific prior settings can be applied to distinct sections of a core. Similarly, in peat bogs fires may destroy part of the accumulated peat, after which the bog continues to accumulate. Similar processes happen also in lakes (i.e., lowstands) and other deposits. This possibly creates a missing section, or a hiatus in the core, and would be seen as a jump in the age-depth model. At depths where a suspected hiatus occurred, a hiatus location may be introduced in our model as follows.

Let $h_1 > h_2 > \ldots > h_H$ be depths at which hiatuses are suspected to have occurred. The approach here is that the autoregressive accumulation rates process loses memory after passing a hiatus, that is if $c_{k-1} \leq h_l < c_k$ then $x_k \sim Gamma(a_H, b_H \Delta c)$, where $Gamma(a_H, b_H)$ is the prior information provided on the jump length for a hiatus (as in Blaauw and Christen 2005, fig. 3(b)). Afterwards, $x_{k-1} = w x_k + (1 - w)\alpha_{k-1}$ as before.

1.1 Outliers and the radiocarbon data model

Our model for the radiocarbon determinations y_j is based on Christen and Pérez (2009). While the traditional Normal model may be used for comparison purposes, that is, $y_j \mid d_j, x \sim N\{\mu(G(d_j, \theta, x)), \sigma_j^2 + \sigma^2(G(d_j, \theta, x)\}$ (where $\sigma^2(\cdot)$ is the variance in the calibration curve, readily available in for example the IntCal09 calibration curve), we use the generalized robust Student-t model of Christen and Pérez (2009), where determination j has the likelihood

$$\left[b + \frac{\{y_j - \mu(\theta_j)\}^2}{2\omega_j^2(\theta_j)} \right]^{-\left(a + \frac{1}{2}\right)},$$

where $\theta_j = G(d_j, \theta, x)$ and $\omega_j^2(\theta) = \sigma_j^2 + \sigma^2(\theta)$. We follow the recommendation of Christen and Pérez (2009) setting $a = 3$ and $b = 4$. This model has as limiting distribution (for $b = a + 1$ and a large) the traditional Normal model and, more importantly, is robust to the presence of outliers.

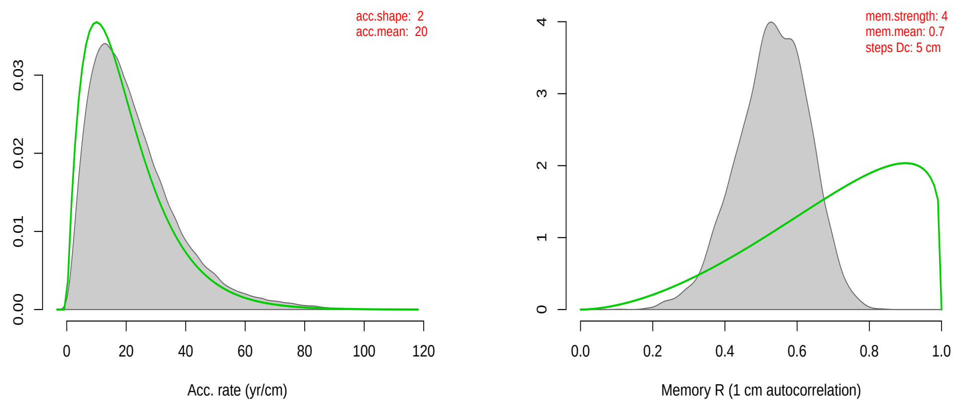
Blaauw and Christen (2005) use the approach of Christen (1994) of "shift outliers" to identify outliers in radiocarbon data and consequently build robust age-depth models. A shift δ_j is introduced in the Normal model to detect outliers $N(\mu(G(d_j, \theta, x) - \phi_j \delta_j), \sigma_j^2 + \sigma^2(\theta))$ where ϕ_j is a Bernoulli (indicator) variable. Using MCMC, the probability $P(\phi_j = 1|Data)$ is calculated and represents the posterior probability that determination j is an outlier (that is, requires a shift on the radiocarbon scale in order to be properly explained with the rest of the data, given the model at hand). Recently this approach has been extended to consider other types of shifts (e.g., on the calendar scale, etc. Bronk Ramsey 2009b). However, Christen and Pérez (2009) explain that most likely the Normal model itself is wrong, since assuming the reported error σ_j to be *known* is simply not true (indeed, σ_j is part of the radiocarbon measurement process as much as the mean radiocarbon age y_j is). By not assuming σ_j as known, the Student-t model is obtained. This model is robust to the presence of outliers and is capable of explaining the otherwise long standing "unexplained scatter" (FiriGroup 2003) in the most controlled and recent radiocarbon interlaboratory studies. We have obtained excellent results using this model (see Section 2), with the added feature that it is far more parsimonious than the former model of Christen (1994), which requires two extra parameters per determination. Moreover, besides robustness, this more parsimonious model contributes to greater speed and ease of convergence in our MCMC.

1.2 A self-adjusting MCMC

A good deal of techniques are available nowadays for numerically evaluating the posterior calculations needed in a Bayesian analysis. In particular, simulating from the posterior distribution using MCMC is, theoretically, straightforward even in high dimensional settings, typically via the Metropolis-Hastings (M-H) algorithm (see for example Gamerman and Lopes 2006). However, tuning the resulting MCMC is not a trivial task, is commonly only achieved partially and requires a lot of input from an experienced statistician on a case-by-case basis. If the statistical methodology is expected to be used by non-statisticians, this part is critical and numerical methods are needed that are robust to different prior distributions and data inputs. Substantial current research on MCMC runs in this direction, developing adaptive algorithms that learn about the problem at hand to automatically tune the MCMC simulation (see Andrieu and Thoms 2008, for a review). However, many additional regularity conditions are required and commonly some additional input is needed (e.g. the expected number of modes in the posterior, etc.). Christen and Fox (2010) present a self-adjusting MCMC algorithm (dubbed "the t-walk") that requires no tuning and has been shown to provide good results in many cases of up to 400 dimensions. In some cases it even rivals ideally tuned MCMCs (reaching theoretical optimality). The t-walk is not adaptive and therefore does not require new restricting conditions since it is a M-H algorithm (that runs in the product space of the original parameter space). The only needed input is the -log of the posterior and two initial points.

A program (in C++) is used to calculate $-\log f(\theta, x, w | y_1, y_2, \ldots, y_m)$. The other input required by the t-walk are two initial points. We simulate w_0 and w'_0 from the prior distribution of w, and given w_0 and w'_0 we simulate the autoregressive series to obtain x_0 and x'_0. θ_0 and θ'_0 are provided by the user. This creates the two initial "city park" values (θ_0, x_0, w_0) and (θ'_0, x'_0, w'_0).

The "energy" function $U(\theta, x, w|y) = -\log f(\theta, x, w|y)$ (that is -log of the posterior



(a) (b)

Figure 1: Prior (green) and posterior (grey) distributions of accumulation rate (a) and memory R (b) for core MSB2K. For interpretation of the references to color in this figure and other figures, the reader is referred to the web version of this article.

distribution, besides perhaps a normalizing adding constant, Liu 2001, p. 80) is

$$\begin{aligned} U(\theta, w, x|y) \quad = \quad & \sum_{j=1}^{m} \left(a + \frac{1}{2}\right) \log\left(b + \frac{\{y_j - \mu(\theta_j)\}^2}{2\omega_j^2(\theta_j)}\right) + \\ & - \log f(\theta) + \\ & \frac{d_s}{\Delta c}(1 - a_w)\log(w) + (1 - b_w)\log\left(1 - w^{\frac{d_s}{\Delta c}}\right) + \\ & \sum_{k=1}^{K} U_\alpha(k, x), \end{aligned}$$

where as above $\theta_j = G(d_j, \theta, x)$ and $\omega_j^2(\theta_j) = \sigma_j^2 + \sigma^2(\theta_j)$ and $U_\alpha(k, x)$ is $-\log$ of the prior for α_k, that is

$$U_\alpha(k, x) = \begin{cases} (1 - a_H)\log(\alpha_k) + b_H\alpha_k, & \alpha_k = x_k, \quad \text{where} \quad c_{k-1} \le h_l < c_k \\ (1 - a_\alpha)\log(\alpha_k) + b_\alpha\alpha_k, & \alpha_k = \frac{x_k - w x_{k+1}}{1 - w}, \quad \text{otherwise.} \end{cases}$$

One added feature of the t-walk is that for many examples it maintains the Integrated Autocorrelation Time (IAT, Geyer 1992) bounded around $30n$, where n is the dimension of the posterior; in this case $n = K + 2$. This means that to obtain a semi-independent sample one would need to resample ("thin") the MCMC output every $30n$ iterations;

to be on the safe side we subsample every $100n$ iterations, assuming that $IAT/n \leq 100$. The burn-in is established by checking the time series plot of $-U(\theta^{(t)}, x^{(t)}, w^{(t)}|y)$ (log-posterior), and commonly takes fewer than 20,000n iterations. Therefore to obtain a final sample size of T (after burn-in and thinning) we require a sample size of $20{,}000n + 100nT$. These guidelines facilitate convergence checking and have proven to be accurate and useful for several age models we have analyzed. These figures might seem extreme resulting in many millions of iterations needed, but the performance of the t-walk has been shown to be consistent for more than 90 cores analyzed and results are *linear* on the dimension $n = K + 2$. This is a remarkable and very useful property of this MCMC, and would potentially enable its use by non-experts. Fortunately, our optimized C++ implementation may run 10^7 iterations in c 1 minute, making these types of runs perfectly feasible, as will be shown by some particular examples in the next Section. In our implementation, to save disk space, we save only accepted iterations. As an estimate, only one in every $20n$ iterations is accepted, resulting in the final numbers for the burn-in and thinning being typically slightly larger than 20,000n and $100n$, respectively.

2 Examples

2.1 The MSB2K peat core

MSB2K consists of a 100 cm thick vertical section from the raised bog peat deposit Meerstalblok in Eastern Netherlands (Blaauw et al. 2003). The core contains 40 ^{14}C dates of selected above-ground plant remains, and was analyzed every cm for about 60 different fossil proxies (e.g., pollen, leaves and seeds of fossilized plant remains) that can inform us about past local to regional environmental conditions during the mid Holocene (Blaauw et al. 2003). From previous research we expected that peat from this region has accumulated at c. 5–50 year per cm (mean c. 20) (Blaauw et al. 2003; Blaauw and Christen 2005). Therefore we set our prior for the accumulation rate as a gamma distribution with shape 2 and mean 20 (that is $a_\alpha = 2$ and $b_\alpha = 1/10$, see Figure 1a). We did not expect that environmental conditions would have changed drastically during the period of interest and therefore the prior for the variability of accumulation rate was set at relatively low levels ("high memory"; steps every $\Delta c = 5$ cm, $a_w = 7$ and $b_w = 3$, that is R has a mean correlation of 0.7; see Figure 1b). In c. 3 minutes on a dual 3.16 GHz processor running Linux, 23.46 million iterations were run, sub-sampling every 2,300 iterations with a burn-in of 20,000 iterations for a selected final sample size of 10,000. The final log-posterior time series plot was stable (not shown here; see the Supporting Material[1]). As explained in Section 1.2, we only save accepted iterations to save disk space, and the final numbers for a particular run vary slightly from those given here.

Posterior distributions were obtained for the accumulation rates and memory (Figure 1), as well as for the age-depth relationship (Figure 2). The posterior for the memory

[1] http://www.cimat.mx/~jac/BaconSupportingMaterial.pdf

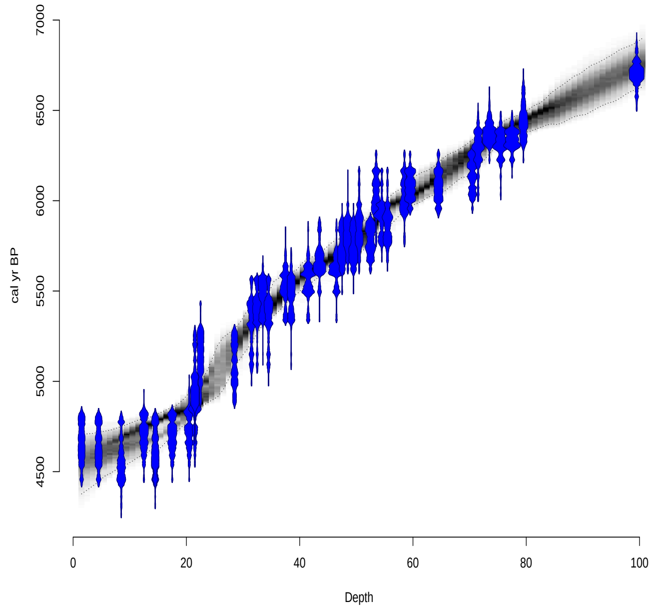


Figure 2: Posterior age–depth model of core MSB2K (grey), overlaying the calibrated distributions of the individual dates (blue). Grey dots indicate the model's 95% probability intervals.

(1 cm correlation, $R = w^{\frac{d_s}{\Delta c}}; d_s = 1$ cm) indicates that our prior belief of a high correlation between peat accumulation at a distance of 1 cm (approximately equivalent to 20 years) seems not very accurate. The substantially lower correlation (compare the prior and posterior in Figure 1(b)) possibly suggests a higher than expected, sub-centennial, climate variability. On the other hand, the average distribution of all accumulation rates x_j's is quite similar to the (prior) distribution used for the innovations (α_j's). This means that the prior is sufficiently strong and the likelihood (data) does not provide further information on accumulation rates.

The radiocarbon data do not provide sufficient information on accumulation rates, especially within sparsely dated parts of the core and/or for flat parts of the radiocarbon calibration curve. Note for example the chronology between 0 and 20 cm core depth in Figure 2. Resulting from a flat part of the calibration curve, the data is simply flat in that part of the core. Any sort of non-decreasing age depth model could be fitted to this part of the core and, indeed, reasonable models are considered by the use of a properly elicited prior, bounding physically possible accumulation rates (for peat, a different prior is used for the lake core presented in the next Section). Other current approaches to age-depth modeling, to be discussed in the next Section, do not control accumulation rates explicitly and may in fact overestimate or underestimate the age-depth variability without the proper weight of a prior; see Figure 5.

Alternative age-depth models for core MSB2K have previously been obtained by Blaauw et al. (2003) and Blaauw and Christen (2005). The latter paper used an approach comparable to the one presented in this paper, albeit with much fewer sections, applying a hiatus and using outlier analysis. The age–depth model by (Blaauw and Christen 2005) was based on two piece-wise linear sections and as such appears much more constrained and rigid than the age–depth model with 21 sections presented here. The model of (Blaauw and Christen 2005) was quite constrained in the middle of both linear sections, which is probably an artefact of the model (Blockley et al. 2007). Our present model seems to capture well the increased uncertainties at sections which are dated at lower density (e.g., between 100 and 80 cm depth).

The importance of taking into account constraints on accumulation rate is obvious from looking at the model behaviour between c. 20 and 0 cm depth (Figure 2). In this section, a plateau in the calibration curve causes nearly constant ages, and many different age-depth curves could be drawn. However, it can be seen that only some age-depth models are likely given the priors of accumulation rate and its variability.

Blaauw and Christen (2005) inferred a hiatus between c. 5500 and 5000 cal BP for core MSB2K, and provided some ecological reasons as to why such a hiatus could have occurred (e.g., a so-called "bog burst" where floods eroded part of the peat deposit). In our present age-depth model for the site however, a gradual change in accumulation rate is inferred during the afore-mentioned period. If a hiatus is forced at 22 cm depth (choosing as prior for the hiatus length a gamma distribution with shape 1 and mean 200), no discernible hiatus is produced (data not shown), indicating that a hiatus–free model would be the most parsimonious choice for this core.

2.2 The RLGH3 lake core

The relatively high precision chronology for core MSB2K is partly due to its unusually high dating resolution. For cores with a lower ratio of radiocarbon dates per centimeter, greater chronological variability is to be expected. In order to investigate this, we assessed a core from lake sediment that was dated at lower resolution. Generally, lake bottoms gradually accumulate organic and inorganic matter derived from within and outside the lake system. Core RLGH3 consists of Holocene lake sediments from the Round Loch of Glenhead, Scotland (Jones et al. 1989; Stevenson et al. 1990; Harkness et al. 1997). The age–depth models shown here are based on 20 radiocarbon dates and we included the prior of θ_0 as a $N(-35, 10)$, cal BP, for the surface sediment (the core was sampled around AD 1985). The sediment being from a lake, we expected lower accumulation rates with less "memory" than for peat core MSB2K. Therefore, we set the priors for accumulation rate as a gamma distribution with shape 2 and mean 50 (yr/cm) and for the accumulation variability a beta distribution with "strength" 20 and mean 0.1 (that is $a_\alpha = 2$, $b_\alpha = 1/25$, $a_w = 2$, $b_w = 18$), and $\Delta c = 5$ cm.

Our age–depth model for core RLGH3 reconstructs a generally smooth accumulation

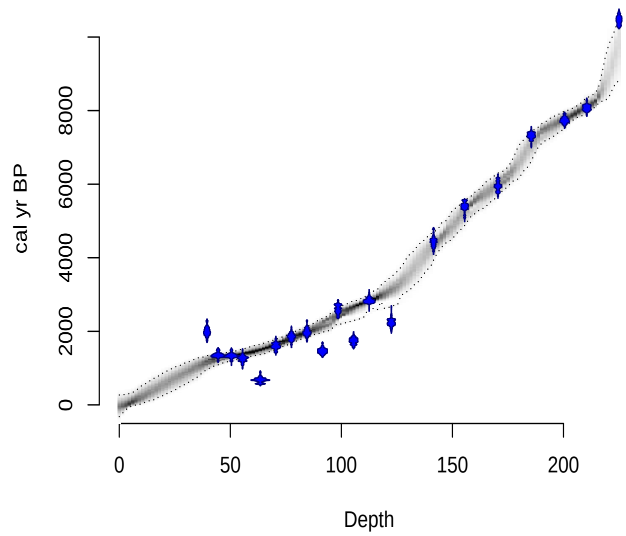
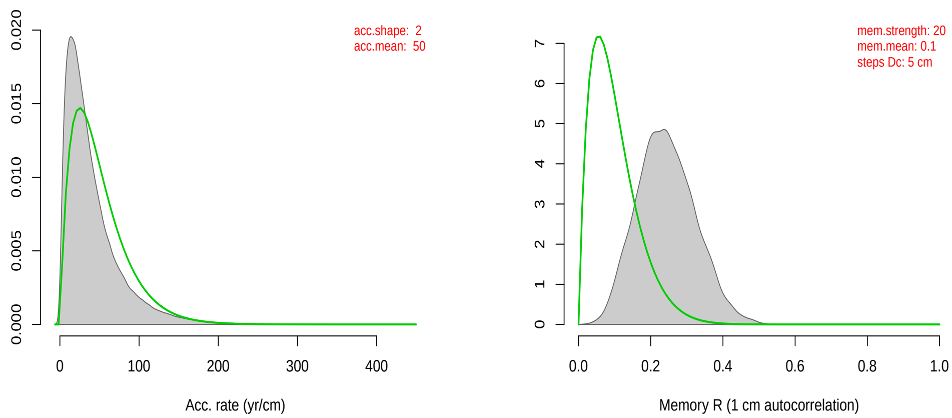


Figure 3: Posterior age–depth model of core RLGH3 (grey), overlaying the calibrated distributions of the individual dates (blue). Grey dots indicate the model's 95% probability intervals.

history, with some fluctuations in accumulation rate at sub-millennial scale (Figure 3). The posteriors for accumulation rate and its variability are comparable to their priors, although the posterior indicates more memory than assumed *a priori*. As in the previous example of the peat core, the prior used for accumulation rates is sufficiently strong and data cannot provide further information. Little more can be learned from radiocarbon data than what we already know about sediment accumulation in this lake core. What constitutes a reasonable (non-decreasing) age-depth model between 0 and 50 cm, for example? This information is not contained within the radiocarbon data and must be an external input, indeed formally provided by our prior for accumulation rates.

There is a clear disagreement between the dates c. 130 to 40 cm depth, which can be attributed to erosion of older organic material from the catchment (Jones et al. 1989). Our model seems unbothered by the discording dates and effectively bypasses them. As detailed in Section 1.1, our method uses a Student-t model to calibrate the dates, instead of the usual normal model. The wider tails of our calibration model reduce the need for detecting and removing outliers.

For comparison purposes, we ran the RLGH3 data through BChron (Parnell et al.



(a) (b)

Figure 4: Prior (green) and posterior (grey) distributions of accumulation rate (a) and memory R (b) for core RLGH3.

2008) and the P Sequence of OxCal (Bronk Ramsey 2008) (Figure 5). The P Sequence models sediment accumulation through sampling from Poisson distributions (here with parameter k set at 2 and interpolation at every 5 cm). As discussed in Section **??**, BChron models the sediment accumulation through a monotonic model with increments determined at random points provided by a renewal process. BChron nor OxCal's P sequence take into account prior information on accumulation rates, and while the variation in accumulation rates can be set in the P sequence, this cannot be done within BChron. In both cases, the final variability in the age-depth models is in fact a result of some parameter settings but owing to a lack of a workable interpretation of the corresponding priors in terms of the accumulation rates (or other physical properties of the core) there is little room to justify the actual variance depicted in the resulting models.

BChron exhibits wider intervals than ours while P sequence narrower (using "$k = 2$" in the OxCal settings), as may be seen from Figures 3 and 5. However using our approach, a prior for accumulation rates is required and becomes crucial in explaining the age-depth model variability, as explained above. Indeed, wider or narrower confidence intervals may indeed be obtained by varying the prior for accumulation rates, α_j's (see the Supporting Material[2] for further discussion regarding the sensitivity of the prior selection for the α_j's). A well elicited informative prior will therefore enable us to downweight unreasonable models, which will lead to our variance estimates along the age-depth model. As opposed, for BChron and the P sequence, what is a "reasonable"

[2] http://www.cimat.mx/~jac/BaconSupportingMaterial.pdf

model depends simply on instrumental priors, which leads to unjustifiable (low or high) variance estimates.

Moreover, both models show only the 95% confidence intervals, while our method provides the entire posterior distribution as grey–scales (Figure 3). Owing to the spread of dates between c. 130 and 40 cm depth, both BChron and OxCal were only able to produce age–depth models when including many additional parameters for outlier analysis. Priors for outliers were a Student-t distribution with 5 degrees of freedom for OxCal, while for BChron prior outlier probabilities for "standard" outliers were 5%, and 0.1% for "extreme" outliers. BChron took c. 17 min to run. OxCal's P sequence ran at speeds comparable to that of our method (2 min.), however it took much trial-and-error to find parameter distributions that resulted in successful, converged P sequence runs. From the cores that we have tested, our routine seems to be more robust, producing stable age-depth models for a wide range of sites.

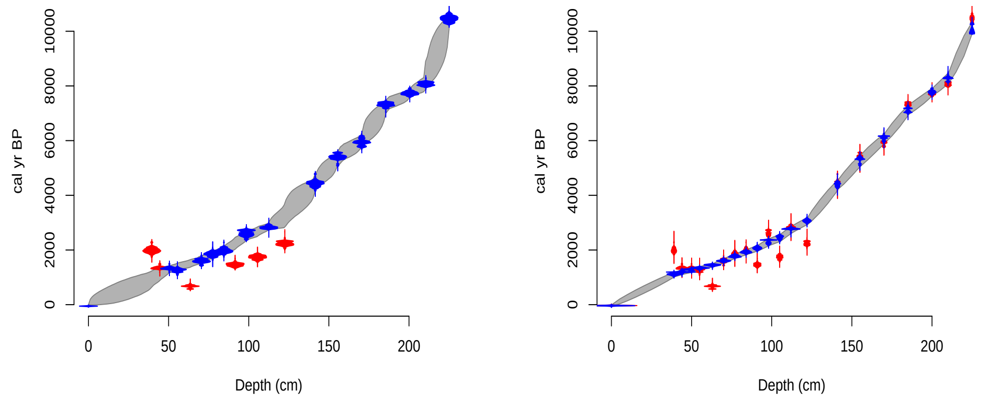
BChron as well as Oxcal use conventional Metropolis Hastings MCMC with shift outlier analysis, producing hugely parametrized models. Moreover, conventional MCMC may become extremely inefficient when the same setting is applied to different posterior distributions (Christen and Fox 2010). OxCal has been developed over nearly 20 years and many tailor–made features in its MCMC have been added (see Bronk Ramsey 2009a, and analysis details online[3]) such as multiple variable moves and 'a small adaptive element BChron in that the range of possible update positions is narrowed once reasonable convergence has been achieved' (C. Bronk Ramsey, personal communication).

3 Discussion

The main goal for age–depth models is to produce age estimates for all depths in a core, and most classic approaches do just that by fitting a curve through the dated points (Blaauw 2010). Instead, the method presented here aims to produce more environmentally realistic age–depth models by reconstructing the underlying processes, i.e. the accumulation/sedimentation process itself. We think of our model as a likely simulation of the deposition process of lakes, bogs or other types of sediment, in which the accumulation process is influenced by environmental conditions. Gradual – or at times abrupt – environmental changes will force responses in the deposition processes, causing accumulation rates to change somewhat from previous rates.

It seems that the P sequence in OxCal (Bronk Ramsey 2008) was constructed with comparable ideas in mind, in that sedimentation is seen as grains settling in a column over time (modelled using a Poisson distribution). BChron (Parnell et al. 2008) produces a random number of sections which accumulate with random positive increments, and as such is perhaps somewhat further away from an "environmentally inspired" model (e.g., varibility in accumulation rate cannot be controlled, leading to perhaps overly inflated confidence intervals). Bpeat assumes piece–wise linear accumulation of peat deposits, based on environmental data that suggest rather constant peat accumulation

[3] http://c14.arch.ox.ac.uk/oxcalhelp/hlp_analysis_detail.html



(a) (b)

Figure 5: Output from alternative age-modelling routines BChron (a, Parnell et al. 2008) and OxCal (b, Bronk Ramsey 2008). Dark grey areas show the 95% confidence intervals of the models. For BChron, red distributions show rejected (outlying) dates, blue distributions show accepted dates. For OxCal, red distributions show individually calibrated dates without taking the other dates or accumulation model into account, while blue distributions show posterior distributions after age–depth modelling and outlier analysis.

over limited time intervals (Belyea and Clymo 2001; Blaauw and Christen 2005). However, linear accumulation might be an overly optimistic assumption, resulting in too constrained confidence intervals especially for longer sections or cores dated at lower resolution (Yeloff et al. 2006; Blockley et al. 2007). On the other hand, BChron tends to overestimate variability in–between dated depths, since it considers unreasonably large or low accumulation rates. Both BChron and the P sequence of OxCal tend to follow isolated radiocarbon dates (see the date at the bottom of the core at c. 225 cm in Figure 5). Indeed, the crucial use of well elicited prior information on accumulation rates leads to more reasonable and consistent models, in the examples presented here and several others already analyzed.

Our methodology is programmed in C++ with a user interface in R and is available in Linux, Mac and Windows platforms (currently in beta test, available for download[4]). The software is called "Bacon" (a name partly inspired by how specific prior information will produce smooth "floppy" or "crispy" *B*ayesian *ac*cumulati*on* models). Bacon has already been tested in more than 90 (lake and peat) cores in a handful of research projects (Charman et al. 2011; de Vleeschouwer et al. 2010). Certainly, we believe it

[4]`http://chrono.qub.ac.uk/blaauw/bacon.html`

provides a useful alternative to building age–depth models in paleoecology.

References

Andrieu, C. and Thoms, J. (2008). "A tutorial on adaptive MCMC." *Statistics and Computing*, 18(4): 343–373. 462

Barndorff-Nielsen, O. E. and Shephard, N. (2001). "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion)." *Journal of the Royal Statistical Society, Series B*, 63: 167–241. 460

— (2003). "Integrated OU processes and non-Gaussian OU-based stochastic volatility models." *Scandinavian Journal of Statistics*, 30: 277–295. 460

Belyea, L. W. and Clymo, R. S. (2001). "Feedback control of the rate of peat formation." *Proceedings of the Royal Society of London. Series B, Biological sciences*, 268: 1315–1321. 470

Blaauw, M. (2010). "Methods and code for 'classical' age-modelling of radiocarbon sequences." *Quaternary Geochronology*, 5: 512–518.
URL http://dx.doi.org/10.1016/j.quageo.2010.01.002 469

Blaauw, M., Bakker, R., Christen, J. A., Hall, V. A., and van der Plicht, J. (2007a). "A Bayesian framework for age modeling of radiocarbon-dated peat deposits: case studies from the Netherlands." *Radiocarbon*, 49(2): 357–367. 458

Blaauw, M. and Christen, J. A. (2005). "Radiocarbon peat chronologies and environmental change." *Applied Statistics*, 54(5): 805–816. 458, 459, 461, 464, 466, 470

Blaauw, M., Christen, J. A., Ampel, L., Veres, D., Hughen, K., and Preusser, F. (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. 458

Blaauw, M., Christen, J. A., Mauquoy, D., van der Plicht, J., and Bennett, K. D. (2007b). "Testing the timing of radiocarbon-dated events between proxy archives." *The Holocene*, 17: 283–288. 458

Blaauw, M., Heuvelink, G. B. M., Mauquoy, D., van der Plicht, J., and van Geel, B. (2003). "A numerical approach to 14C wiggle-match dating of organic deposits: best fits and confidence intervals." *Quaternary Science Reviews*, 22(14): 1485–1500. 458, 464, 466

Blockley, S. P. E., Blaauw, M., Bronk Ramsey, C., and van der Plicht, J. (2007). "Assessing uncertainties in age modelling sedimentary records in the Lateglacial and early Holocene." *Quaternary Science Reviews*, 26: 1915–1926. 458, 466, 470

Bronk Ramsey, C. (2008). "Deposition models for chronological records." *Quaternary Science Reviews*, 27(1-2): 42–60. 458, 468, 469, 470

— (2009a). "Bayesian analysis of radiocarbon dates." *Radiocarbon*, 51(1): 337–360. 469

— (2009b). "Dealing with outliers and offsets in radiocarbon dating." *Radiocarbon*, 51(3): 1023–1045. 462

Cai, B. and Dunsun, D. B. (2007). "Bayesian Multivariate Isotonic Regression Splines: Applications to Carcinogenicity Studies." *Journal of the American Statistical Association*, 102(480): 1158–1171. 458

Chambers, F. M., Mauquoy, D., Brain, S. A., Blaauw, M., and Daniell, J. R. G. (2007). "Globally synchronous climate change 2800 years ago: proxy data from peat in South America." *Earth and Planetary Science Letters*, 253: 439–444. 458

Charman, D., Barber, K., Blaauw, M., Langdon, P., Mauquoy, D., Daley, T., Hughes, P., and Karofeld, E. (2009). "Climate drivers for peatland palaeoclimate records." *Quaternary Science Reviews*, 28: 1811–1819. 458

Charman, D., Beilman, D. W., Blaauw, M., Booth, B., Brewer, S., Chambers, F. M., Christen, J. A., Gallego-Sala, A., Harrison, S. P., Hughes, P. D. M., Jackson, S. T., Korhola, A., Mauquoy, D., Mitchell, F. J. G., Prentice, I. C., van der Linden, M., de Vleeschouwer, F., Yu, Z. C., Bauer, I. E., Garneau, M., Karofeld, E., Le Roux, G., Moschen, R., Nichols, J. E., Phadtare, N. R., Rausch, N., Swindles, G. T., Väliranta, M., van Bellen, S., and van Geel, B. (2011). "Carbon–cycle implications of climate–driven changes in peat accumulation during the last millennium." *Nature*, (Revision). 470

Christen, J., Clymo, D., and Litton, C. (1995). "A Bayesian approach to the use of C14 dates in the estimation of the age of peat." *Radiocarbon*, 37: 431–442. 458

Christen, J. A. (1994). "Summarising a set of radiocarbon determinations: a robust approach." *Applied Statistics*, 43(3): 489–503. 461, 462

Christen, J. A. and Fox, C. (2010). "A general purpose sampling algorithm for continuous distributions (the t-walk)." *Bayesian Analisis*, 4(2): 263–282. 459, 462, 469

Christen, J. A. and Pérez, S. (2009). "A new robust statistical model for radiocarbon data." *Radiocarbon*, 51(3): 1047–1059. 459, 461, 462

de Vleeschouwer, F., Pazdur, A., Luthers, C., Streel, M., Mauquoy, D., Wastiaux, C., Le Roux, G., Moschen, R., Blaauw, M., Pawlyta, J., Sikorski, J., and Piotrowska, N. (2010). "A millennial record of environmental change recorded in peat deposits from the Misten Bog (east Belgium)." *Quaternary International*, (Submitted). 470

FiriGroup (2003). "The Third International Radiocarbon Intercomparison (Tiri) and the Fourth International Radiocarbon (Fir) – 1999-2002 – Results, analysis and conclusions." *Radiocarbon*, 45(2): 135–328. 462

Gamerman, D. and Lopes, H. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. London: Chapman and Hall. 462

Geyer, C. J. (1992). "Practical Markov Chain Monte Carlo." *Statistical Science*, 43: 179–189. 463

Harkness, D., Miller, B. F., and Tipping, R. M. (1997). "NERC radiocarbon measurements 1977-1988." *Quaternary Science Reviews*, 16: 925–927. 466

Haslett, J. and Parnell, A. (2008). "A simple monotone process with application to radiocarbon-dated depth chronologies." *Applied Statistics*, 57(4): 399–418. 458, 459

Holmes, C. C. and Heard, N. A. (2003). "Generalized monotonic regression using random change points." *Statistics in Medicine*, 22: 623–838. 458

Jones, V. J., Stevenson, A. C., and Battarbee, R. W. (1989). "Acidification of lakes in Galloway, southwest Scotland: a diatom and pollen study of the post-glacial history of the Round Loch of Glenhead." *Journal of Ecology*, 77: 1–23. 466, 467

Lavine, M. and Mokus, A. (1995). "A nonparametric Bayes method for isotonic regression." *Journal of Statistical Planning and Inference*, 46: 235–248. 458

Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. New York: Springer. 463

Lowe, J. J. and Walker, M. J. C. (1997). *Reconstructing Quaternary Environments (2nd edition)*. Prentice Hall. 457

Mena, R. H. and Walker, S. G. (2004). "A density function connected with a non–negative self-decomposable random variable." *Journal of Statistical Computation and Simulation*, 74(10): 765–775. 460

— (2005). "Stationary autoregressive models via a Bayesian nonparametric approach." *Journal of Time Series Analysis*, 26: 789–805. 460

Parnell, A., Haslett, J., Allen, J., Buck, C., and 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. 467, 469, 470

Plunkett, G., Whitehouse, N. J., Charman, D. J., Blaauw, M., Kelly, E., and Mulhall, I. (2009). "A multi-proxy palaeoenvironmental investigation of the findspot of an Iron Age bog body from Oldcroghan, Co. Offaly, Ireland." *Journal of Archaeological Science*, 36: 265–277. 458

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., and Weyhenmeyer, C. E. (2009). "IntCal09 and Marine09 radiocarbon age calibration curves, 0–50,000 years cal BP." *Radiocarbon*, 51: 1111–1150. 458, 459

Sillasoo, Ü., Mauquoy, D., Blundell, A., Charman, D., Blaauw, M., Daniell, J. G. R., Toms, P., Newberry, J., Chambers, F. M., and Karofeld, E. (2007). "Peat multi-proxy from Männikjärve bog as indicators of Late Holocene climate changes in Estonia." *Boreas*, 36: 20–37. 458

Stevenson, A. C., Jones, V. J., and Battarbee, R. W. (1990). "The cause of peat erosion: a palaeolimnological approach." *New Phytologist*, 114: 727–735. 466

Wohlfarth, B., Blaauw, M., Davies, S. M., Andersson, M., Wastegård, S., Hormes, A., and Possnert, G. (2006). "Constraining the age of Lateglacial and early Holocene pollen zones and tephra horizons in southern Sweden with Bayesian probability methods." *Journal of Quaternary Science*, 21: 321–334. 458

Wohlfarth, B., Veres, D., Ampel, L., Lacourse, T., Blaauw, M., Preusser, F., Andrieu-Ponel, V., Kéravis, D., Lallier-Vergès, E., Björck, S., Davies, S. M., de Beaulieu, J. L., Risberg, J., Hormes, A., Kasper, H. U., Possnert, G., Reille, M., Thouveny, N., and Zander, A. (2008). "Rapid ecosystem response to abrupt climate changes during the last glacial period in Western Europe, 40–16 kyr BP." *Geology*, 36: 407–410. 458

Yeloff, D., Bennett, K. D., Blaauw, M., Mauquoy, D., Sillasoo, Ü., van der Plicht, J., and 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. 458, 470

Acknowledgments

Part of this research was completed while JAC was visiting Queen's University Belfast. This paper benefited from a workshop on Bayesian age–depth modeling in San Miguel de Allende (Mexico), hosted by CIMAT and partly funded by PAGES and INQUA (through the INTREPID project). The authors thank the reviewers of *Bayesian Analysis* that help to improve the presentation of the paper.