A BAYESIAN APPROACH TO THE USE OF $^{14}$C DATES IN THE ESTIMATION OF THE AGE OF PEAT

J. A. CHRISTEN,1 R. S. CLYMO2 and C. D. LITTON3

ABSTRACT. Peatlands occupy a significant portion of the land surface of the Earth and form a large carbon store. Most peat-forming systems have two layers. The upper layer, the acrotelm, fixes carbon by photosynthesis, loses some of it by decay and passes the remainder on to the lower layer, the catotelm. In the catotelm, decay continues at a slower rate. Mathematical models of the growth of the catotelm have been proposed which relate the cumulative mass of peat above a particular depth to calendar age of the peat at that depth. We demonstrate how $^{14}$C dating and the Bayesian approach to data analysis can be used to make inferences about the relation between calendar age and cumulative mass, and to estimate the accumulation and decay rates.

INTRODUCTION

Peatlands cover ca. 3% of the land surface of the Earth. They may be as much as 10–15 m deep, averaging 2 m, and they form a large store of carbon. Peat accumulates and decays through time and therefore a complex relation between depth and age can be expected. From chemical and biological analyses of peat at successive depths, one can learn about environmental and climatic change through time (see Warner, Tolonen and Tolonen 1991).

Most peat-forming systems consist of two layers (Ingram 1978). The upper layer, the acrotelm, fixes carbon by photosynthesis, loses much of it by decay and passes the remainder on to the lower layer, the catotelm, which is below the summer water table. The depth of the acrotelm is usually 5–50 cm, whereas the depth of the catotelm can be as much as 1000–1500 cm, although 200 cm is usual. Until recently, scientists believed that there was little or no decay in the catotelm but recent research is beginning to overturn this notion (Clymo 1984, 1991).

Several mathematical models of the growth of the catotelm, based upon differing interpretations of physical and chemical processes, have been proposed (see Clymo 1992). These models do not relate age directly to depth but instead relate the age of peat at a particular depth to the cumulative dry mass of material above that depth (but still within the catotelm). However, as Clymo (1984: 619) remarks, it is not necessary to know where the present surface of the catotelm is; an arbitrary datum may be taken at any depth within the catotelm, and values of the age and mass expressed relative to that datum.

Consider an arbitrary fixed datum at depth $d_0$ below the surface of the catotelm and let $\theta_0$ denote the unknown calendar age, measured in years BP, of the peat at that depth. Consider peat at depth $d$ ($>d_0$). Let its calendar age, measured in years BP, be $\theta$ ($>\theta_0$) and let $M$ be the cumulative dry mass (in g cm$^{-2}$) of material between depths $d_0$ and $d$. Let $p$ be the rate at which dry mass is added (on an areal basis) to the catotelm, $\alpha$ be the proportional rate of decay in the catotelm and $a$ be a constant. Clymo (1992) gives the following three possible models based upon differing assumptions for the form of $\alpha$

$$M = \frac{p}{a} (1 - e^{-a(\theta - \theta_0)})$$

(1)

1Instituto de Matemáticas, Unidad Morelia, UNAM Nicolás Romero 150, Centro 58000 Morelia, Michoacán, Mexico
2School of Biological Sciences, Queen Mary and Westfield College, London, E1 4NS, England
3Department of Mathematics, University of Nottingham, University Park, Nottingham, NG7 2RD, England

\[ M = \frac{P}{a} \log_e (1 + a (\theta - \theta_0)) \]  

and

\[ M = \frac{P}{a} (\sqrt{1 + 2a (\theta - \theta_0)} - 1). \]

For model (1) \( \alpha = a \); for model (2) \( \alpha = a \nu(\theta) \); and for model (3) \( \alpha = a \nu^2(\theta) \) where \( \nu(\theta) \) is the proportion of the initial mass remaining after time \( \theta - \theta_0 \). In practice, \( a \) is small, and if we consider a short time period then all three models reduce to \( M = p (\theta - \theta_0) \). That is, if we are considering a relatively short time span of, say, a few hundred years, the amount of decay in the catotelm is negligible, and the cumulative mass is therefore directly proportional to the relative age. It is only for periods of time greater than several millennia that these more complex models are necessary.

We demonstrate here the use of \(^{14}\text{C} \) dating and the Bayesian approach to data analysis to estimate the age \( \theta \), corresponding to a given cumulative dry mass \( M \). To illustrate our approach we use model (1), but we can also apply our methods to the other models. To estimate the age \( \theta \), corresponding to the cumulative mass \( M \), we take a peat core and \(^{14}\text{C} \) date samples from various depths and also measure the cumulative mass corresponding to those depths. The primary objective is to use model (1) and the \(^{14}\text{C} \) dates to predict the age (on the calendar scale) of the peat at different depths. A secondary objective is to obtain reliable estimates of the parameters \( p \) and \( a \).

The relation between the cumulative mass and the calendar dates is complex, as is the relation between the \(^{14}\text{C} \) dates and the calendar time scale. The latter relation is non-monotonic because of the “wiggly-nature” of the high-precision calibration curve (Pearson and Stuiver 1986; Stuiver and Pearson 1986). Clymo et al. (1990) attempted to fit models of peat growth using Pearson's (1986) wiggle-matching techniques, but to do so, had to make some assumptions about the values of the unknown parameters. Clymo (1992) and Warner, Clymo and Tolonen (1993) used the function minimization procedures of Nelder and Mead (1965) but were unable to include any uncertainty regarding the calendar dates corresponding to the \(^{14}\text{C} \) dates. In this paper, we include in the analysis all the relevant uncertainties.

**METHODS AND DISCUSSION**

**Outline of the Bayesian Methodology**

We use the Bayesian statistical framework, which is fundamentally different from the traditional approach adopted by most researchers. Litton and Buck (1994) review its application to archaeological problems and Christen (1994) demonstrates its usefulness for the interpretation of \(^{14}\text{C} \) dates. We do not intend to give a full exposition of the Bayesian approach; for more details, see Buck, Cavanagh and Litton (1996), who describe the Bayesian methodology and its application to the analysis and interpretation of archaeological data. For readers who may be unfamiliar with the approach, we give a brief review here.

Researchers develop statistical models that represent, in a mathematical manner, the problem at hand. Within these models, a relation is posited between the data, denoted by \( x = (x_1, ..., x_n) \), and the unknown parameters, denoted by \( \psi = (\psi_1, ..., \psi_k) \), is posited. Researchers following the Bayesian approach use probabilities to express the uncertainty about the parameters \( \psi \), before and after seeing the data \( x \). The Bayesian framework offers a simple way of computing the posterior uncertainty from the prior uncertainty and the data via the model.
In broad terms, the Bayesian method has the following three components:

1. The prior probability density function, denoted by \( g(\psi) \). This may be read as “how much belief do I attach to possible values of the unknown parameters before (prior to) observing the data?”. A relatively large value of \( g(\psi) \) for a particular value of \( \psi \) indicates that this value is thought to be relatively likely; a relatively small value, that it is relatively unlikely.

2. The likelihood, denoted by \( l(\psi ; x) \). This may be read as “how much are particular values of the parameters supported by the data?”. If the value of the likelihood at a particular value of the parameters, say \( \psi_1 \), is relatively large then this indicates that the data strongly support \( \psi_1 \). If at another value of \( \psi \), say \( \psi_2 \), the likelihood is relatively small, then this suggests that the support from data for \( \psi_2 \) is relatively small. In practice, the likelihood function is determined from the probability distribution of \( x \) conditional on \( \psi \), which we denote by \( f(x | \psi) \).

3. The posterior density function, denoted by \( h(\psi | x) \). This may be read as “how much belief do I attach to possible values of the unknown parameters after (posterior to) observing the data?”. A relatively large value of \( h(\psi | x) \) indicates that the corresponding value of \( \psi \) is, a posteriori, relatively likely; a relatively small value, that it is relatively unlikely.

The prior, the likelihood and the posterior are connected by Bayes’ theorem, which may be expressed in the form

\[
h(\psi | x) \propto l(\psi ; x) \cdot g(\psi).
\]  

Here, \( g(\psi) \) represents the prior information about \( \psi \). We then observe the data \( x \) and Bayes’ theorem tells us that the posterior information, represented by \( h(\psi | x) \), is proportional to the prior \( g(\psi) \) times the likelihood \( l(\psi ; x) \). By the result of this operation, we have combined two sources of information, that provided by the prior and that provided by the data. Note that before using \( h(\psi | x) \) to make inferences, the constant of proportionality in (4) must be determined so that \( h(\psi | x) \) integrates to 1.

Suffice it to say that within the Bayesian framework, the fitting of complex models is relatively straightforward, although we may need to use sophisticated numerical techniques to make inferences about the parameters of interest. An important feature of the Bayesian approach is that one can incorporate into the analysis, coherently and logically, relevant prior knowledge about the parameters. Of course, one must clearly state the use of such prior information so that the assumptions are clear to the reader.

**Applying the Bayesian Methodology**

To apply the Bayesian methodology, we must develop a model, expressed in terms of parameters, that relates the \( ^{14}C \) dates to the cumulative dry mass data. Then we need to determine the likelihood, establish the prior information about the parameters and convert this into a prior density. Next, we apply Bayes’ theorem to calculate the posterior density of the parameters, which we then use to make inferences.

There are two components to our peat model: the first relates the cumulative mass to the calendar age and the second relates the calendar age to the \( ^{14}C \) age.

**Modeling the Cumulative Mass – Calendar Age Relation**

Suppose we have a series of \( n \) \( ^{14}C \) dates \( x_1 \pm \sigma_1, x_2 \pm \sigma_2, \ldots, x_n \pm \sigma_n \) from peat samples taken at successive depths \( d_1 < d_2 < \ldots < d_n \) corresponding to calendar years \( \theta_1, \theta_2, \ldots, \theta_n \). Recall that \( d_n \) and \( \theta_0 \) represent the reference (datum) depth and the corresponding calendar age of the peat at that depth. Because peat age increases with depth, the associated calendar years for the \( ^{14}C \) dates (the calendar
years marking the death of the organic materials contained in the peat samples) must be ordered so that \( \theta_0 < \theta_1 < \theta_2 < \ldots < \theta_n \).

Let \( m_i \) be the cumulative dry mass at depth \( d_i \), then according to model (1), \( m_i \) and \( \theta_i \) are related by the equation

\[
m_i = \frac{p}{a} (1 - e^{-a(\theta_i - \theta_0)}).
\]  

(5)

Rearranging, \( \theta_i \) can be expressed as

\[
\theta_i = \theta_0 - a^{-1} \log_e (1 - p^{-1}a m_i).
\]

(6)

Thus the calendar age \( \theta_i \) can be expressed in terms of the unknown parameters \( \psi = (\theta_0, p, a) \).

**Modeling the Calendar Date - \(^{14}\)C Date Relation**

Suppose we have a \(^{14}\)C date, denoted by \( x = \sigma \), where \( x \) is the estimated \(^{14}\)C age and \( \sigma \) is the corresponding standard deviation reported by the laboratory. Let \( \theta \) represent the unknown calendar date of the corresponding event. Here we view \( x \) as a realization of the random variable

\[
X \sim N(\mu(\theta), \sigma^2 + \sigma^2(\theta))
\]

(7)

where \( \mu(\theta) \) denotes the calibration curve and \( \sigma^2(\theta) \) represents the uncertainty in the calibration curve itself as \( \mu(\theta) \) is not known exactly.

We use a high-precision calibration curve in piecewise linear form to approximate \( \mu(\theta) \), taking

\[
\mu(\theta) = \begin{cases} 
  c_0 + b_0 \theta & (\theta \leq t_0) \\
  c_k + b_k \theta & (t_{k-1} < \theta \leq t_k, \ k = 1, 2, \ldots, K) \\
  c_K + b_K \theta & (t_K < \theta)
\end{cases}
\]

(8)

where the \( t_k \) are the knots of the calibration curve, \( K + 1 \) is the number of knots used, and the \( b_i \) and \( c_i \) are known constants. Then we have

\[
X \sim N(c_0 + b_0 \theta, \omega^2(\theta)) \quad (\theta \leq t_0)
\]

(9)

\[
X \sim N(c_k + b_k \theta, \omega^2(\theta)) \quad (t_{k-1} < \theta \leq t_k, \ k = 1, 2, \ldots, K)
\]

(10)

\[
X \sim N(c_K + b_K \theta, \omega^2(\theta)) \quad (t_K < \theta)
\]

(11)

where

\[
\omega^2(\theta) = \sigma^2 + \sigma^2(\theta).
\]

(12)

Thus, the probability density function of the random variable \( X_i \), representing the \( i \)th \(^{14}\)C date, conditional on \( \psi \) is

\[
f_i(x_i|\psi) \propto \omega_i^{-1}(\theta_i) \exp \left\{ -\frac{(x_i - \mu(\theta_i))^2}{2\omega_i^2(\theta_i)} \right\}
\]

(13)

where \( \theta_i \) is given by (6) and

\[
\omega_i^2(\theta_i) = \sigma_i^2 + \sigma^2(\theta_i).
\]

(14)
**A Bayesian Approach to the Use of $^{14}$C Dates**

**The Likelihood**

Noting that, conditional on $\psi$, the $^{14}$C dates are independent, the likelihood can be expressed as

$$l(\psi; x) = \prod_{i=1}^{n} f_i(x_i|\psi) \propto \prod_{i=1}^{n} \omega_i^{-1}(\theta_i) \exp \{-(x_i - \mu(\theta_i))^2/2\omega_i^2(\theta_i)\}.$$  \hspace{1cm} (15)

where $\theta_i$ is given by (6) and $\omega_i^2(\theta_i)$ by (14).

**The Prior for $\theta_0$, $p$ and $a$**

In general, $\theta_0$ will be independent of $p$ and $a$, so that the joint prior density of $\psi$ can be expressed as $g(\psi) = g_1(\theta_0) g_2(p, a)$ where $g_1(\theta_0)$ is the prior density for $\theta_0$ and $g_2(p, a)$ is the joint prior density for $p$ and $a$. It is much more difficult to make a statement about whether or not $p$ and $a$ are independent. Both are affected by numerous common factors such as temperature, availability of water and plant species, but little is known about the exact botanical, physical and chemical mechanisms involved. These two parameters are probably correlated but we do not know whether the correlation is positive or negative. In addition, when incorporated into particular models, the actual values may result in numerical correlations with no biological basis.

**The Posterior**

Let $g(\psi)$ be the joint prior density function of $\psi$; then by Bayes' theorem, the joint posterior density of $\psi$ is given by

$$h(\psi | x) \propto l(\psi; x) g(\psi).$$  \hspace{1cm} (16)

From this joint posterior density function, by using numerical integration techniques, one can calculate the marginal posterior densities of $\theta_0$, $a$ and $p$. Also, for a given value of the cumulative mass, $m_0$, at depth, $d_0$, by using (6) one can calculate the posterior density of the corresponding calendar date, $\theta_0$. Further, from the joint and marginal posterior densities, one can calculate appropriate summary statistics, such as posterior means, posterior standard deviations and 95% intervals.

**Example 1**

Warner, Tolonen and Tolonen (1991) report the results of a study of peatland formation and development at Point Escuminac, New Brunswick, Canada. From a 532-cm peat core, the authors collected for $^{14}$C analysis 22 samples from different depths in the catotelm. They used the $^{14}$C dates to estimate the ages, on the $^{14}$C time scale, of pollen and macrofossils found in the core and to understand the development of the bog and the surrounding environment. Warner, Clymo and Tolonen (1993) used the 16 dates that (in 1992) were within the time period of the calibration curve to fit model (1) using Nelder and Mead's (1965) function minimization procedure. They estimated the dates on the calendar scale as well as the peat accumulation and decay parameters, $p$ and $a$. We will re-analyze the data, given in Table 1, using the Bayesian framework.

**Specification of the Prior Density**

As mentioned above, we know very little *a priori* about $p$ and $a$ or how they relate. First, we have no prior information as to whether $p$ and $a$ correlate, and if they do, whether positively or negatively. Therefore, we assume their independence and express the joint prior density of $p$ and $a$ as the product of their marginal densities, *i.e.*,

$$g_2(p, a) = g_p(p) g_a(a).$$  \hspace{1cm} (17)
Second, Clymo (1984) suggests that the most likely (modal) values for the northern hemisphere are $0.005 \text{ g cm}^{-2} \text{ yr}^{-1}$ and $0.0002 \text{ yr}^{-1}$, respectively. To reflect the large uncertainty about these values, wide margins are stated in terms of multiplicative factors of $10^{-1}$ and 10. That is, it is believed that $p$ lies between 0.0005 and 0.05, and the range of $a$ is 0.00002 to 0.002. Because small values of both parameters are thought to be more likely than large ones, we model $\log_{10} p$ and $\log_{10} a$ as being normally distributed. That is, we assume that

$$\log_{10} p \sim N(\log_{10} 0.005, 0.25) \quad (18)$$

$$\log_{10} a \sim N(\log_{10} 0.0002, 0.25) . \quad (19)$$

Therefore, the prior probability density functions of $p$ and $a$ are given by

$$g_p (p) = \frac{2}{p (\log_{10} 10) \sqrt{2\pi}} \exp \left\{ -2 (\log_{10} p - \log_{10} 0.005)^2 \right\}$$

and

$$g_a (a) = \frac{2}{a (\log_{10} 10) \sqrt{2\pi}} \exp \left\{ -2 (\log_{10} a + \log_{10} (-0.0002))^2 \right\}$$

respectively.

Finally, we have no information about the date, $\theta_o$, of the material at the datum level except that it must lie within, say, the last 6000 yr. Other than this, we cannot say that one year is more likely than another. Thus, a priori, we assume that $\theta_o$ has a uniform distribution over the interval 0 to 6000, i.e.,

$$g_1 (\theta_o) = \frac{1}{6000} \quad \text{for} \quad 0 < \theta_o < 6000 . \quad (22)$$

This completes our specification of the prior density.
Interpreting the Posterior Density

Using Bayes’ theorem (see Equation (4)), we calculated the joint posterior distribution of $\theta_0$, $a$ and $p$. The marginal posterior density of $\theta_0$ is given by

$$h_0(\theta_0|x) \propto \int_{0}^{\infty} \int_{0}^{\infty} h(\theta_0, a, p|x) \, dp \, da . \tag{23}$$

This integral cannot be evaluated analytically, nor can the constant of proportionality, so numerical methods are used (see Buck, Cavanagh and Litton: Chap. 8). Likewise for the marginal posterior densities of $p$ and $a$. Figure 1 gives these densities and shows that some of the three distributions are

![Diagram of posterior distributions](image)

Fig. 1. Marginal posterior densities of the calendar date at the datum level, $\theta_0$, the proportional decay rate, $a$, and the rate of addition $p$, using peat growth model $\theta_i = \theta_0 + a^{-1} \log \left(1 - p^{-1} a m_i\right)$ and the data from Point Escuminac, New Brunswick, Canada. The dashed lines show the values, with one standard error bars, estimated by ad hoc function minimization techniques before the present work (Clymo 1992).
not symmetric. The modal (most likely) values of \( \theta_0 \), \( p \) and \( a \) are ca. 220 cal BP, 0.022 g cm\(^{-2}\) yr\(^{-1}\) and 0.00015 yr, respectively. The 95\% highest posterior density (HPD) interval (the shortest interval with posterior probability of 0.95) for \( \theta_0 \) is 130–300 cal BP and the 95\% HPD intervals for \( p \) and \( a \) are 0.021–0.024 g cm\(^{-2}\) yr\(^{-1}\) and 0.00013–0.00017 yr\(^{-1}\), respectively.

We observe how the age of peat varies with the cumulative mass. We can obtain the posterior density of the age of peat at any desired cumulative mass: Figure 2 gives examples of this. Figure 2 shows a wider range of possible values at both ends of the peat core. For example, at a cumulative mass of 40 g cm\(^{-2}\), the range is ca. 60 yr, whereas at a cumulative mass of 100 g cm\(^{-2}\), the range is >300 yr. This was expected because we have more information from the middle section of the core than from either end.

**Example 2**

Clymo *et al.* (1990) consider a peat core from Ellergower Moss, Galloway, southwest Scotland. From the ca. 50-cm-long core at 2-cm intervals below the summer water-table, that is, from the top of the catotelm, 11 samples were \(^{14}\)C-dated. Table 2 contains details of the depths and corresponding dates. Presumably, this peat core represents only a short period of time, say a few hundred years, so the simplified model (not involving \( a \))

\[
m_i = p (\theta_i - \theta_0)
\]

(24)

is applicable. We take the same priors for \( p \) and \( \theta_0 \), as in the previous example.

**Interpretation of the Posterior Density**

Using numerical integration, we obtained the posterior densities of \( \theta_0 \) and \( p \). The posterior distribution of \( \theta_0 \) is flat—values between 50 and 200 cal BP are approximately equally likely. In contrast, the
distribution of $p$ has a mode at just over 0.005 g cm$^{-2}$ yr$^{-1}$ but has a long right-hand tail up to ca. 0.018 g cm$^{-2}$ yr$^{-1}$. The 95% HPD intervals for these parameters are 20–210 cal BP and 0.0045–0.0171 g cm$^{-2}$ yr$^{-1}$, respectively. Figure 3 shows the gray-scale representation of the posterior den-
sities of the age of peat at every 0.01 g cm\(^{-2}\) of cumulative mass. The two dark regions correspond to two unconnected regions of high posterior probability. The results are not so clear-cut as for the data from Point Escuminac. For example, for a cumulative mass of 1.0 g cm\(^{-2}\) the distribution is unimodal, with likely dates ranging from 200 to 290 cal BP, whereas at greater depths the distribution splits in two. For a cumulative mass of 1.5 g cm\(^{-2}\), the range is 270 to 380 cal BP, with modes at ca. 290 and 350, but with dates ca. 310–320 being relatively unlikely.

In this example, given any values for \(p\) and \(\theta_0\), we can calculate the value of \(\theta_i\) for specified \(m_i\) using the equation \(\theta_i = \theta_0 + m_i p^{-1}\). Using these \(\theta_i\)s, we can then “position” the corresponding dates \(x_i \pm \sigma_i\) against the calibration curve, a procedure known as “wiggle-matching”. In Figure 4 we plot the wiggle-match obtained using the modal values (\(\theta_0 = 61\), \(p = 0.0054\)) of the joint posterior density of \(\theta_0\) and \(p\). Apart from one date, the data appear to follow the calibration curve reasonably well, although the flat posterior distribution of \(\theta_0\) indicates that within the range 50 to 200 cal BP, the match is almost equally good at many other positions. It is unclear whether this is an inherent defect of wiggle-matching techniques or because of the temporal variation of \(p\), or a combination of both.

**Fig. 4** \(^{14}\)C dates from the peat core from Ellergower Moss with the relevant part of the calibration curve. \(* = \^{14}\)C age \(x_i\); \(\rightarrow\) \(x_i - \sigma_i\) to \(x_i + \sigma_i\). Here we used the modal values of the joint posterior distribution of \(\theta_0\) and \(p\), taking \(\theta_0 = 61\) yr and \(p = 0.0054\) g cm\(^{-2}\) yr\(^{-1}\).

**CONCLUSION**

We have described the use of Bayesian statistical techniques to estimate the age of \(^{14}\)C-dated peat cores using a simple peat growth model. Using the Bayesian approach, we included in a single analysis the \(^{14}\)C calibration curve, the available \(^{14}\)C dates, a specific peat growth model and prior information about the parameters of the model. We combined all these factors within the Bayesian framework to estimate the age of peat corresponding to a particular cumulative mass.
We used two examples to obtain, in terms of the posterior, distributions of the age of peat at any cumulative mass desired, and calculated highest posterior density intervals. We also obtained reliable estimates of the rates of accumulation and of decay, p and a, respectively. Apart from the immediate advantages of working within a coherent and consistent framework for statistical inference, we are able to include all the relevant uncertainties into the analysis and our techniques provide clear, visual and interpretable results.

ACKNOWLEDGMENTS

J. A. C. was supported by the Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México. We thank F. Oldfield and B. G. Warner for permission to use unpublished data.

REFERENCES


