Journal of the Royal Statistical Society
SERIES C Applied Statistics

Appl. Statist. (2008)
57, *Part* 4, *pp.* 399–418

A simple monotone process with application to radiocarbon-dated depth chronologies

John Haslett and Andrew Parnell

Trinity College Dublin, Republic of Ireland

[Received July 2007. Revised February 2008]

Summary. We propose a new and simple continuous Markov monotone stochastic process and use it to make inference on a partially observed monotone stochastic process. The process is piecewise linear, based on additive independent gamma increments arriving in a Poisson fashion. An independent increments variation allows very simple conditional simulation of sample paths given known values of the process. We take advantage of a reparameterization involving the Tweedie distribution to provide efficient computation. The motivating problem is the establishment of a chronology for samples taken from lake sediment cores, i.e. the attribution of a set of dates to samples of the core given their depths, knowing that the age–depth relationship is monotone. The chronological information arises from radiocarbon (^{14}C) dating at a subset of depths. We use the process to model the stochastically varying rate of sedimentation.

Keywords: Compound Poisson–gamma distribution; Monotone processes; Radiocarbon dating; Tweedie distribution

1. Introduction

Continuous monotone stochastic processes arise in many areas of application: the study of accumulated wear in aspects of reliability (Heinricher and Stockbridge, 1993; Lam and Zhang, 2003), of exposure to disease in epidemiology (Lee *et al.*, 2004) and in the modelling of forest fires (Reed and McKelvey, 2002). Here we study the accumulation of lake sediment cores which can give an insight into past climates. We focus on modelling the age–depth relationship (i.e. the chronology) given partial information. We propose a very simple stochastic process with properties (continuous, monotonic and Markovian) that may make it attractive elsewhere.

The most familiar monotone processes are counting processes, of which the simplest is the Poisson process. Such processes, being piecewise constant, are not in themselves attractive for modelling smoothly increasing random functions. The process that is presented here may be interpreted as an integrated process $y(t) = \int_0^t u(s)\,\mathrm{d}s$; here $u(\cdot)$ denotes a positive piecewise constant process, in our case the sedimentation rate. Our process is thus piecewise linear. The very simple properties derive from an underlying Poisson process.

Monotone processes are also used as part of the technical apparatus of monotone Bayesian smoothing, where interest lies in smooth latent random functions as priors. These arise in applications as diverse as price effect data (e.g. Brezger and Steiner (2005)), leukaemia prognosis (Holmes and Heard, 2003), sleeping problems (Dunson, 2005), dose–response (Kong and Eubank, 2006) and growth curves (e.g. Ramsay (1998)). The use of Gaussian processes in Bayesian smoothing is now commonplace, but they are less natural for monotone smoothing: hence

Address for correspondence: Andrew Parnell, Department of Statistics, Lloyd Institute, Trinity College Dublin, Dublin 2, Republic of Ireland.
E-mail: Andrew.Parnell@tcd.ie

 0035–9254/08/57399

the need for simple alternatives. Smoothers that are defined via basis functions (such as splines) are random if the number and/or locations of the knots are random. The proposals here can be seen as using linear basis functions and a Poisson process to locate the knots. Crucially in this paper we show that it is possible to marginalize with respect to this process. Effectively, the prior for the knots is conjugate to the prescription for the basis functions. In our proposal there are no issues to be overcome by the unknown number of knots.

The methods that are discussed here may also be seen as an instance of another area of research—the study of stochastic processes given partial realizations. For diffusions this is a very active research area (e.g. Dellaportas *et al.* (2006)). The process that is proposed here is not a diffusion; it may in fact be seen as an example of a 'partially deterministic Markov process' (Davis, 1984). The generic challenge lies in computing the likelihood for the observed data and in generating random sample paths conditional on the observations. We see below that in this case both tasks are very simple. In particular, the simplicity of the likelihood derives from a close connection between the compound Poisson–gamma (CPG) and Tweedie distributions, which are themselves well studied, being members of the exponential dispersion family (Jørgensen, 1987). Despite this simplicity, the method seems to be robust.

In this paper our focus is on the depths $\{d_i; i = 1, \ldots, n\}$ that are taken from a core of lake sediment. Biological proxies in such cores provide reconstructions of past climates; see Haslett *et al.* (2006). The corresponding calendar ages θ_i are crucial in estimating the temporal nature of the reconstructions. Necessarily the underlying function $d(\theta)$ is continuous and monotone; older ages must lie at deeper sediments. However, these ages are only available indirectly via radiocarbon (^{14}C) dating. Formally, for $m < n$ of the samples with depths $\{d_j; j = 1, \ldots, m\}$, suitably qualified laboratory estimates $x_j \pm \tau_j$ are given in ^{14}C years before present (BP). The challenge is to determine the true calendar dates θ_i at all depths d_i, i.e. to provide a chronology, the set $\{\theta_i, d_i\}$, or more generally a random function $\theta(d)$. This function is the key contribution of the paper. Uncertainty arises both in the calibration of the radiocarbon ages and in stochastic interpolation for the $n - m$ depths at which we have no age information. The study of this uncertainty is our central focus. We approach this below by generating continuous random functions $\theta(d)$ that are consistent with the data.

In a context such as this, the following features seem desirable, in addition to monotonicity.

(a) Continuity; chronologies can be requested for any set of depths within a core. More generally, although any continuous process can be approximated by a discrete equivalent, it seems desirable that the procedure be well behaved in the limit.
(b) Minimal assumptions on smoothness; there is no reason *a priori* to exclude either almost flat or very steep sections, within the constraints of continuity, for great variations in sedimentation rate are observed in practice.
(c) Increased uncertainty away from ^{14}C-dated points, which can be very irregular, as we discuss in Section 4; in the interests of generality, it seems desirable that the procedure be an interpolator in the limiting case where the measurements are without error.

Additionally the model should be simple, with few parameters, facilitating theoretical analysis and efficient sampling.

Fig. 1 indicates the nature of the problem. Separate radiocarbon dating of each of the m samples generates uncertain data (calendar ages—represented by horizontal bars, being the 95% highest density region (HDR) intervals) at various depths. A valid chronology is a continuous monotonic function; our model is defined on random piecewise linear functions. Given the data, our model permits the sampling of many possible piecewise linear sedimentation histories or, equivalently, chronologies. Each such chronology allows a deterministic reconstruction of the

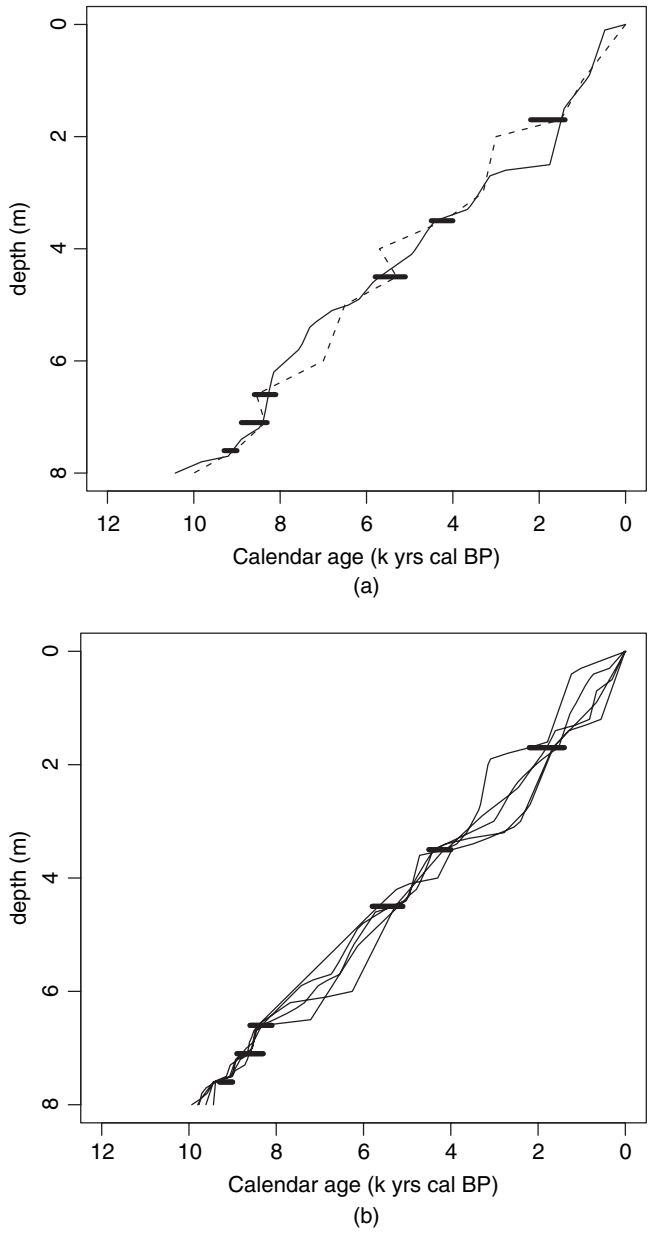

Fig. 1. (a) Sample data (▬, 95% HDR intervals for calibrated radiocarbon dates at various depths) with a satisfactory monotonic chronology (———) and an unsatisfactory non-monotonic chronology (-------) (at depths of around 4 and 6.5 m) and (b) five sampled satisfactory monotonic chronologies (———) fitted to sample data (▬)

calendar ages for all of the n; importantly, the samples are dated jointly. Given the data and the model, sampling many such histories provides many chronologies, allowing analysis of any aspect of the joint uncertainty that is associated with the entire chronology.

In Section 2 we review at a more technical level the options that are available for monotone smoothing. Section 3 provides details of the stochastic process that we propose, covering

theoretical aspects and setting out the algorithm. We present in Section 4 some more details of the motivating problem, reviewing previous literature on calibration. Section 5 presents applications. Section 6 indicates possibilities for extension.

2. Monotone processes and smoothing

The modelling of continuous monotonic processes $y(x)$ (equivalently $x(y)$) for $\{x, y\} \in \mathbb{R}^2$ arises in several fields. Our focus is on sampling realizations of such functions that are consistent with possibly noisy observations on $\{x_j, y_j; j = 1, \ldots, m\}$ but are interpolators in the absence of noise. A Bayesian formulation is natural. The key is the family of priors for the random functions.

One natural family of stochastic processes is those which are described as 'partially deterministic Markov processes' (Davis, 1984). Diffusions do not have monotonic sample paths and pure jump processes are not continuous. Several of the approaches to Bayesian smoothing can be seen as being based on (monotone) splines defined on possibly randomly located knots. In the present study the requirement of interpolation (in the limiting case) is a strong condition; it requires *a priori* that we have knots at and between data points. A wide discussion of the general issues of monotone smoothing may be found in for example Brezger and Lang (2006), in particular, the subselecting of monotone paths from the posterior or, almost equivalently, only proposing monotone paths. More recently, Wang and Dunson (2007) have considered the even wider issue of smoothing under stochastic constraints.

Holmes and Heard (2003) proposed a piecewise constant process for y, with a random number of knots at random locations; they used flat priors for this. Although they envisaged that the number of knots is less than the number of samples, there is nothing in principle to insist on this. Within disjoint sets that are defined by the knots, they envisaged that the y have a set-specific distribution. They imposed no explicit constraint on monotonicity, in the interests of computational simplicity; they simply subsampled monotone paths from the posterior. In the applied ^{14}C literature Bronk Ramsey (2007) has proposed a Poisson counting process which can be seen as having piecewise constant sample paths; it brings with it a very natural model for the number and location of knots, which has much in common with the process that we propose. However, such paths are clearly discontinuous, yielding in our context sedimentation rates that are either zero or infinite.

Neelon and Dunson (2004) proposed (as we do below) a piecewise linear process. They adopted a normal distribution truncated at zero for the slopes, which they discussed in terms of unrestricted latent slopes and implemented in a natural fashion by never accepting sets of slopes that contain negative values. Indeed they went further by allowing point masses at zero. They envisaged a large number of knots at prespecified equally spaced locations; point masses at zero effectively thin the posterior knot process. Knots at data points will lead to interpolation if that is desired. One of the strengths of their procedure is that it facilitates the testing of hypotheses about 'flat patches'. However, this procedure is not well behaved in the limit, being in this case an arbitrarily large number of knots, for arbitrarily large sets of unconstrained slopes will contain no negative slopes with vanishingly small probability.

In the context of ^{14}C dating, the work of Blaauw and Christen (2005) can be seen as a very special case of the Neelon and Dunson (2004) general approach, involving piecewise linear basis functions. However, they have unresolved issues on the very small number of knots that they recommended and their process does not lend itself naturally to interpolation. In Section 4.3 we further discuss their procedure and that of Bronk Ramsey (2007).

Our proposal is to use a simply defined stochastic process in continuous x, with attractive continuity and Markov properties, based on a random number of piecewise linear segments of

random length. Crucially, we show that it is simple to marginalize with respect to both. We adopt a simple gamma model for independent and identically distributed data for the total increment associated with such segments and show that the process can be structured to have mean-square continuous sample paths. Although the process is strictly monotone, it does naturally produce almost flat patches. It may therefore have possibilities for wide applicability.

3. A piecewise linear Markov monotone stochastic process

We first introduce our procedure in the context of a bivariate renewal process $(\tilde{x}, \tilde{y}) = \{(\tilde{x}_i, \tilde{y}_i); i = 1, 2, \dots\}$. This facilitates the specification in Section 3.1 of a three-parameter (λ, α, β) *piecewise linear* CPG process $x(y)$ which is continuous with respect to x and Markov with respect to y. In Section 3.2 we establish connections to the closely related Tweedie distribution. We consider some theoretical properties of interpolation in Section 3.3 showing that it is natural to fix α. We conclude by considering inference in Section 3.4 and, in Section 3.5, the robustness of the model to departures from assumptions.

3.1. Process construction

We consider here the generation of the piecewise linear random sample paths in Fig. 1. Initially we consider a piecewise linear process that is defined on a bivariate gamma renewal process. This facilitates the specification of a piecewise linear CPG process; we find it most convenient to describe this as a function $x(y)$. We propose the following constructive algorithm based on a bivariate monotone renewal process (Hunter, 1974) with increments from gamma distributions as below. We then consider the continuity of such processes and specialization to a Poisson process.

Firstly, consider a process that is constructed as follows.

(a) Define $\tilde{x}_0 = \tilde{y}_0 = 0$, and create pairs $(\tilde{x}_i, \tilde{y}_i) = \{(\tilde{x}_{i-1}, \tilde{y}_{i-1}) + (r_{x_i}, r_{y_i}); i = 1, 2, 3, \dots\}$ where the (r_{x_i}, r_{y_i}) are independent and identically distributed gamma(α_x, β_x) and gamma(α_y, β_y).

(b) Use linear interpolation to create line segments.

Now, $x(y) = \Sigma_{i=1}^{N(y)} r_{x_i} + s_{y,i+1}(y - \tilde{y}_i)$ is a piecewise linear process that is indexed by continuous label y where $s_{y,i} = r_{x_i}/r_{y_i}$ and $N(y)$ is a counting process in the range $(0, y)$. Furthermore, pairs of differences $(\tilde{x}_k - \tilde{x}_j, \tilde{y}_k - \tilde{y}_j)$, with $j < k$, are independent yet can still be written as a sum of gamma random variables with respect to $N(y_j - y_k)$. The set of renewal points has an independent increments property.

Consideration of the mean-square continuity of $x(y)$ with respect to x is facilitated by thinking of the random function as $y(x)$; in this notation the slopes are $s_{x,i} = r_{y_i}/r_{x_i}$ and are of course independent and identically distributed. Continuity is assured if $\lim_{h \to 0}[\mathrm{var}\{y(x+h) - y(x)\}]/h^2 = 0$ (Stein, 1999). Conditioning on $\tilde{x}(x) = \arg\max(\tilde{x}_j, j = 1, \dots, m; \tilde{x}_j < x)$, the most recent renewal point before x, the probability of another renewal point in $(x, x+h)$ is $g_x\{x - \tilde{x}(x)\}h$, for small h, where g is the hazard rate of gamma(α_x, β_x). It follows, after unconditioning, that $\mathrm{var}\{y(x+h) - y(x)\} = h^2 \,\mathrm{var}(s_x) + O(h^3)$ for $\tilde{x}_{i-1} < x \leqslant \tilde{x}_i$. Thus, provided that $\mathrm{var}(s_x)$ is defined, the process is mean square continuous. Now s_x is the product $r_y(1/r_x)$ of independent gamma and inverse gamma random variables. However, the variance of the latter is not defined unless $\alpha_x > 2$, which becomes a necessary and sufficient condition for mean-square continuity.

With unit shape parameters (α_x, α_y) the renewal processes are of course Poisson processes. The piecewise linear process $x(y)$, with $\alpha_y = 1$ and $\alpha_x > 2$, is thus Markov with respect to y and mean square continuous with respect to x. We thus adopt more conventional notation,

parameterizing the Poisson process by λ and the gamma increments by (α, β) and referring to the process as a piecewise linear CPG, with parameters (λ, α, β). Subsequently we shall see that it is possible to fix α.

Note that the slope process, which we may denote as $s(y)$, is piecewise constant and Markov. The number $N(y)$ of *changes* in slope within an interval of length y follows a Poisson distribution. The process $x(y)$ is thus an integrated Markov process.

3.2. Marginalizing over N(y)

We address here a central issue: it is possible to marginalize with respect to $N(y)$. We focus on the embedded CPG renewal process $\tilde{x}(y) = \Sigma_{i=1}^{N(y)+1} r_{x_i}$ and exploit the similarity with $x_0(y) = \Sigma_{i=1}^{N(y)} r_{x_i}$, a Poisson sum of independent and identically distributed gamma random variables. This latter follows a Tweedie distribution (Tweedie, 1984), which is alternatively known as the CPG distribution, some of whose properties we describe below. Clearly these are very similar for large λ, which is of particular interest to us; below we expand on this connection.

The density of $\tilde{x}(y)$, marginalized with respect to $N(y)$, can be expressed simply as

$$f(x; \lambda y, \alpha, \beta) = \exp(-\beta x) \exp(-\lambda y) \sum_{n=0}^{\infty} \frac{\beta^{(n+1)\alpha}}{\Gamma\{(n+1)\alpha\}} x^{(n+1)\alpha - 1} \frac{(\lambda y)^n}{n!}. \tag{1}$$

By contrast the CPG distribution may be written as

$$f_0(x; \lambda y, \alpha, \beta) = \exp(-\beta x) \exp(-\lambda y) \sum_{n=1}^{\infty} \frac{\beta^{n\alpha}}{\Gamma(n\alpha)} x^{n\alpha - 1} \frac{(\lambda y)^n}{n!}, \tag{2}$$

with an atom at zero corresponding to $N(y) = 0$. Although the infinite sum is not available analytically, very fast routines are available for the evaluation of $f_0(x)$; see Dunn and Smyth (2005). But we note that

$$f(x; \lambda y, \alpha, \beta) = f_0(x; \lambda y, \alpha, \beta) + \frac{\partial f_0(x; \lambda y, \alpha, \beta)}{\partial(\lambda y)} \tag{3}$$

is easily computable by using a numerical approximation for the derivative.

Interestingly, the conventional Tweedie parameterization (e.g. Kaas (2001))

$$\lambda y = \frac{\mu^{2-p}}{\psi(2-p)}, \qquad \alpha = \frac{2-p}{p-1}, \qquad \frac{1}{\beta} = \psi(p-1)\mu^{p-1} \tag{4}$$

shows that it is a member of the exponential dispersion family (Jørgensen, 1987). It is then easy to show that $\mathbb{E}[x_0(y)] = \lambda y \alpha/\beta = \mu$ and that $\mathbb{E}[\tilde{x}(y)] = \mu + \alpha/\beta \approx \mu$ for large λ. Similarly $\mathrm{var}\{x_0(y)\} = \lambda y \alpha(\alpha+1)/\beta^2 = \psi\mu^p$ and $\mathrm{var}\{\tilde{x}(y)\} = \psi\mu^p + \alpha/\beta^2 \approx \psi\mu^p$ for large λ.

3.3. The conditioned process

We consider here the uncertainty in $x(y)$, given the parameters and realizations $\{(x_i, y_i); i = 1, 2, \ldots, s\}$. Our interest here is in the properties of the process as an interpolator. We shall see that for large λ the uncertainty is a very weak function of α and that fixing $\alpha = 4$ is natural. For simplicity we shall confine ourselves to the case where the realizations are in fact a subset of the renewal points or may be well approximated as such. We may thus use the independent increments version of the process. This approximation is not unnatural when λ is large.

In this case, the increments $\tilde{x}_i - \tilde{x}_{i'}$ are a sum of $N(\tilde{y}_i - \tilde{y}_{i'}) + 1$ independent gamma random variables, where $N(y) \sim \mathrm{Poisson}(\lambda y)$, given the parameters. It is thus sufficient to consider the

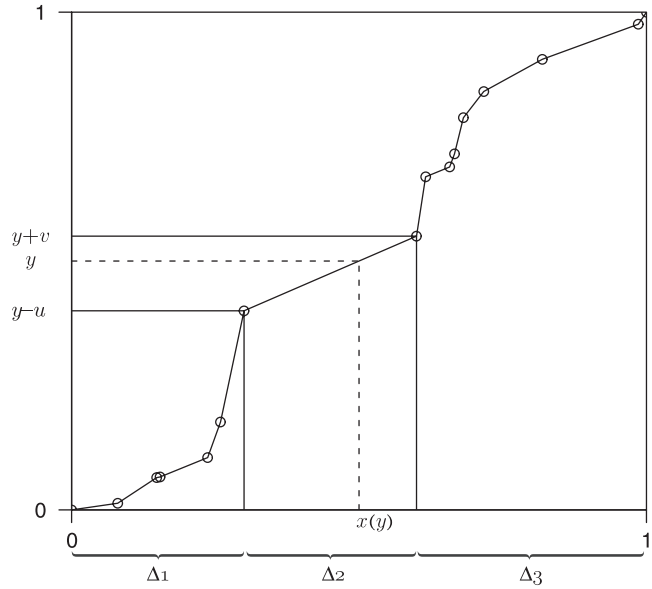

Fig. 2. Schematic diagram of the set-up for the calculation of var$\{x(y)\}$ for the conditioned CPG process

uncertainty in $x(y)$ given the realizations immediately above and below y for, conditional on these, the other realizations are irrelevant. (This is in fact the key element of the approximation, which is valid even when λ is not very large.) For simplicity, we scale this such that the given realizations lie at opposite corners of the unit square, the scaling being equivalent to setting $\beta = 1$; Fig. 2. We focus on var$\{x(\frac{1}{2})|x(0) = 0, x(1) = 1\}$, dropping explicit reference to the conditioning for simplicity of notation.

Such a process may be constructively defined:

(a) draw N points from a Poisson distribution with parameter λ;
(b) draw independent multivariate samples of length $N + 1$ from the Dirichlet distributions with multivariate parameters **1** and α**1** for the y- and x-axes respectively (**1** denoting a vector of 1s of length $N + 1$);
(c) form partial sums;
(d) plot on the unit square with linear interpolation.

We use here the close connection between the gamma and Dirichlet distributions.

The properties of the sample paths are dictated entirely by the parameters α and λ, which control the lengths of the $N + 1$ renewal intervals on the x-axis and the number of renewal points respectively. Note that $y = \frac{1}{2}$ selects at random one of the renewal intervals. Let $y - u$ and $y + v$ denote the end points of this interval on the vertical axis. The point y thus partitions the axis into three segments: $(0, y - u)$, $(y - u, y + v)$ and $(y + v, 1)$. Similarly, it partitions the x-axis; denote the lengths of these three segments on this axis by $(\Delta_1, \Delta_2, \Delta_3)$. Suppose that the first and last of these comprise M and $N - M$ renewal intervals, where $0 \leqslant M \leqslant N$. It follows that $(\Delta_1, \Delta_2, \Delta_3) \sim \text{Dirichlet}\{M\alpha, \alpha, (N - M)\alpha\}$. We have $N \sim \text{Poisson}(\lambda)$ and, with λ sufficiently large, $M|N \sim \text{binomial}(N, y)$. We also define $z = u/(u + v)$ and write $x(y) = \Delta_1 + z\Delta_2$. A schematic view of this set-up in shown in Fig. 2.

We can now write

$$\mathbb{E}[x^2|N, M, z] = \mathbb{E}[\Delta_1^2|N, M] + z^2\,\mathbb{E}[\Delta_2^2|N, M] + 2z\,\mathbb{E}[\Delta_1\Delta_2|N, M] \tag{5}$$

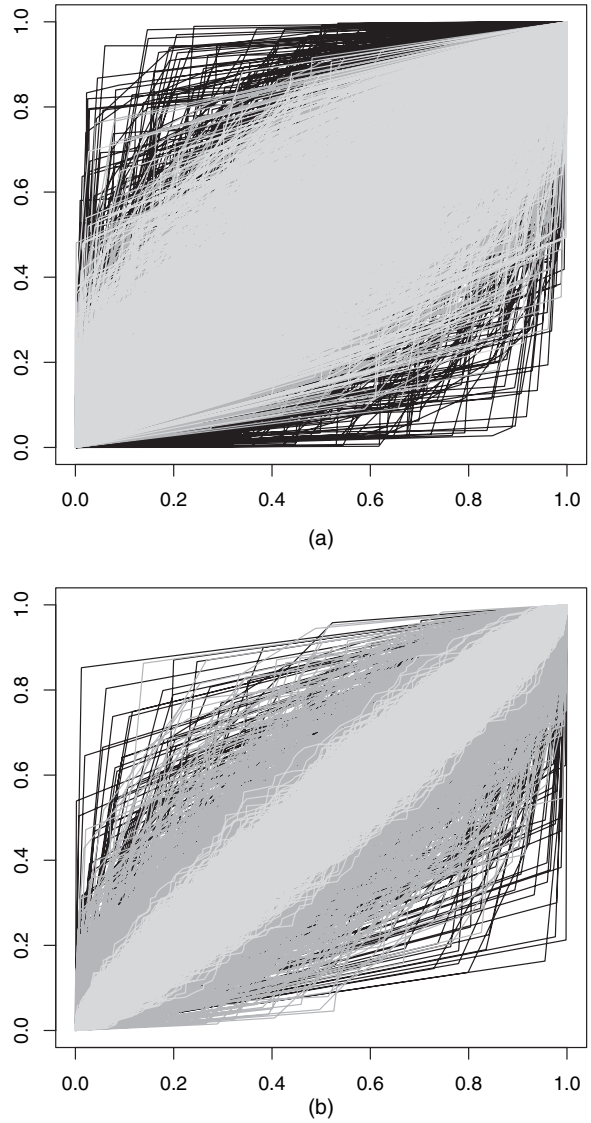

Fig. 3. Simulated sample paths from (a) varying α with fixed λ and (b) varying λ with fixed α: the lighter greys represent increased values of the varying parameter

$$= \frac{M^2\alpha + M(1+2z\alpha) + z^2(1+\alpha)}{(N+1)(N\alpha+\alpha+1)}, \tag{6}$$

using standard results on the moments of the Dirichlet distribution. With λ large we also have $z \sim \text{uniform}(0,1)$ so that

$$\mathbb{E}[x^2|N] \approx \frac{1}{4} - \frac{\alpha+1}{6(N+1)} + \frac{(2\alpha+3)(\alpha+1)}{12(N\alpha+\alpha+1)}. \tag{7}$$

Since $\alpha > 2$, we also have $\lambda\alpha$ large; thus we can approximate $1/(N\alpha+\alpha+1)$ by $1/(N\alpha+\alpha)$ to give (after taking expectations over N)

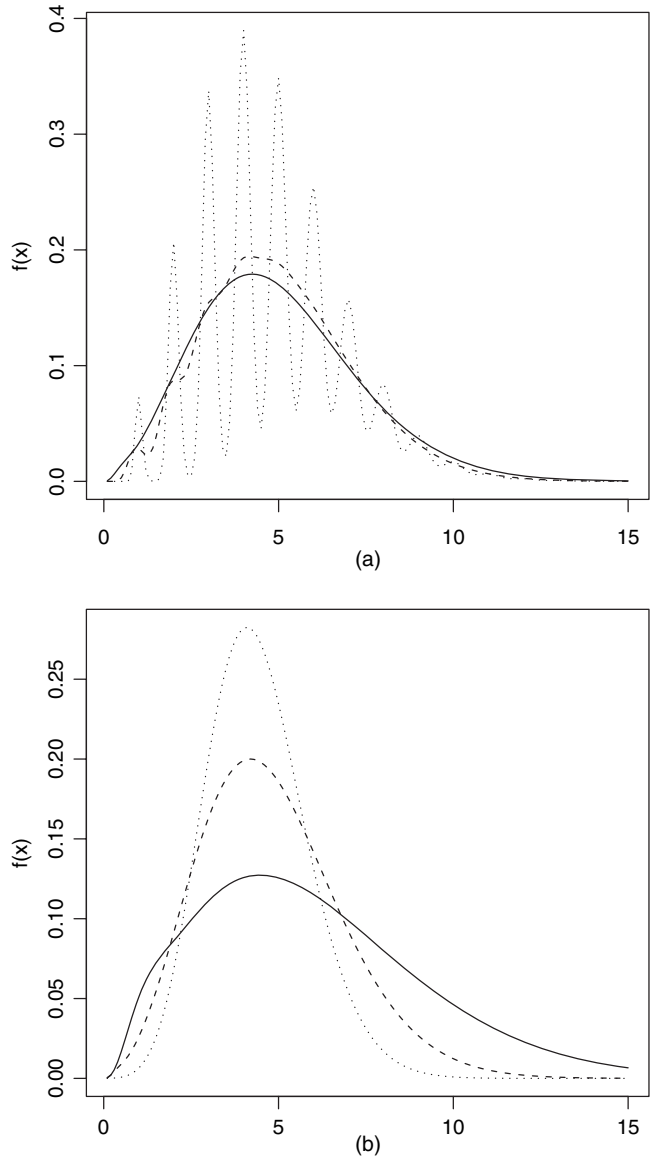

Fig. 4. Density functions for the CPG distribution (defined in equation (1)) with varying parameter values: (a) ———, $\alpha = 4$, $\lambda = 4$ and $\beta = 4$; - - - - - - -, $\alpha = 15$, $\lambda = 4$ and $\beta = 15$; · · · · · · ·, $\alpha = 100$, $\lambda = 4$ and $\beta = 100$; (b) ———, $\alpha = 4$, $\lambda = 2$ and $\beta = 2$; - - - - - - -, $\alpha = 4$, $\lambda = 5$ and $\beta = 5$; · · · · · · ·, $\alpha = 4$, $\lambda = 10$ and $\beta = 10$

$$\mathrm{var}\{x(\tfrac{1}{2})\} \approx \frac{\{1-\exp(-\lambda)\}(1+\alpha)}{4\alpha\lambda} \approx \frac{\psi\mu^{p-2}}{4}. \tag{8}$$

Thus, in the limit for large λ, the conditioned process is *exactly* the line $y = x$. (Clearly this is also so for the less interesting very small λ.) Furthermore the process is relatively insensitive to α. Fig. 3 shows the effect of varying α and λ on the variability of the sample paths.

It therefore seems natural, in situations where the focus is on the conditioned process, to simplify the model even further by fixing α at an appropriate large value for given μ. But, for large

α, $f\{x(y); \lambda y, \alpha, \beta)\}$ is multimodal; Fig. 4. Indeed, in the limit, $f(x; \lambda y, \alpha, \beta)$ tends to the degenerate scaled Poisson distribution, confined to discrete quanta $n\mu/\lambda y$ for integer n. Although in our application all large values of α lead to the same inferences, it seems more natural to set α as large as possible while $f\{x(y); \lambda, \alpha, \beta)\}$ is unimodal. Dunn and Smyth (2005) suggested that $\alpha = 4$ is a conservative limit and this is what we adopt.

3.4. Model fitting

Our interest here is with inference given observations. We adopt the independent increments process, treating the data as observations on a subset $(\tilde{x}_s, \tilde{y}_s)$ of the renewal points. Subsequently in Section 4 we consider having noisy observations on $\tilde{x}_s$; our task will then include inference on the $\tilde{x}_s$, given a model for the noise. We firstly consider inference on the parameters λ and β. We subsequently consider the predictive distribution of y-values corresponding to any x given the data, using the conditioned process above. For this we are not restricted to the renewal points of the process.

We write $\Omega_j = \tilde{x}_s(y_{s,j}) - \tilde{x}_s(y_{s,j-1})$, and $\omega_j = y_{s,j} - y_{s,j-1}$, so that

$$\Omega_j = \sum_{i=1}^{N(\omega_j)+1} G_i, \tag{9}$$

with $G_i \sim \text{gamma}(\alpha, \beta)$ and $N(\omega_j) \sim \text{Poisson}(\lambda\omega_j)$. The posterior for λ and β is obtained from

$$\pi(\lambda, \beta | \boldsymbol{\Omega}, \boldsymbol{\omega}) \propto \pi(\boldsymbol{\Omega} | \alpha, \beta, \lambda, \boldsymbol{\omega})\, \pi(\lambda, \beta), \tag{10}$$

with $\{\boldsymbol{\Omega}, \boldsymbol{\omega}\}$ denoting the complete set of $\{\Omega_j, \omega_j\}$ pairs. The likelihood, $\pi(\boldsymbol{\Omega} | \alpha, \beta, \lambda, \boldsymbol{\omega}) = \Pi_{j=1}^{m-1}(\Omega_j | \alpha, \beta, \lambda, \omega_j)$, is now of the form that is given in equation (1), with $\Omega_j = x$, and $\lambda y = \lambda\omega_j$. Observe that it is not necessary to condition on the unobserved renewal points.

The predictive distribution of $x(y)$ given $\tilde{y}_s$ and $\tilde{x}_s$ is most easily described constructively as, for each segment, repeatedly drawing from $\pi(\lambda, \beta | \boldsymbol{\Omega}, \boldsymbol{\omega})$ and, for each such draw, repeatedly drawing from the conditioned process, as outlined in Section 3.3. More formally, we have

$$\begin{aligned} \pi\{x(y) | \tilde{x}_s, \tilde{y}_s\} &= \int \pi\{x(y), \beta, \lambda | \tilde{x}_s, \tilde{y}_s\}\, \mathrm{d}\beta\, \mathrm{d}\lambda \\ &= \int \pi\{x(y) | \beta, \lambda, \tilde{x}_s, \tilde{y}_s\}\, \pi(\beta, \lambda | \tilde{x}_s, \tilde{y}_s)\, \mathrm{d}\beta\, \mathrm{d}\lambda, \end{aligned} \tag{11}$$

which is the predictive density marginalized over the posterior.

3.5. Model validation

We initially test our model by generating underlying sample paths from a process, which is not necessarily that considered in this section, with varying parameter values. We then pass a subset of $\{x, y\}$ points from the process to the model. In each of many realizations of the process we construct 95% highest posterior densities corresponding to four fixed values of y. We report the coverage rates. In particular we consider four scenarios.

(a) The data are generated from a CPG process with $\alpha = 4$ (the value at which we fix α in our model). The data are a subset of the renewal points of the process.
(b) The data are generated from a CPG process with $\alpha \neq 4$. The data are a subset of the renewal points of the process.
(c) The data are generated from a CPG process with $\alpha \neq 4$. The data are not renewal points of the process. Instead, they are generated uniformly across the y-axis range of the true sample path.

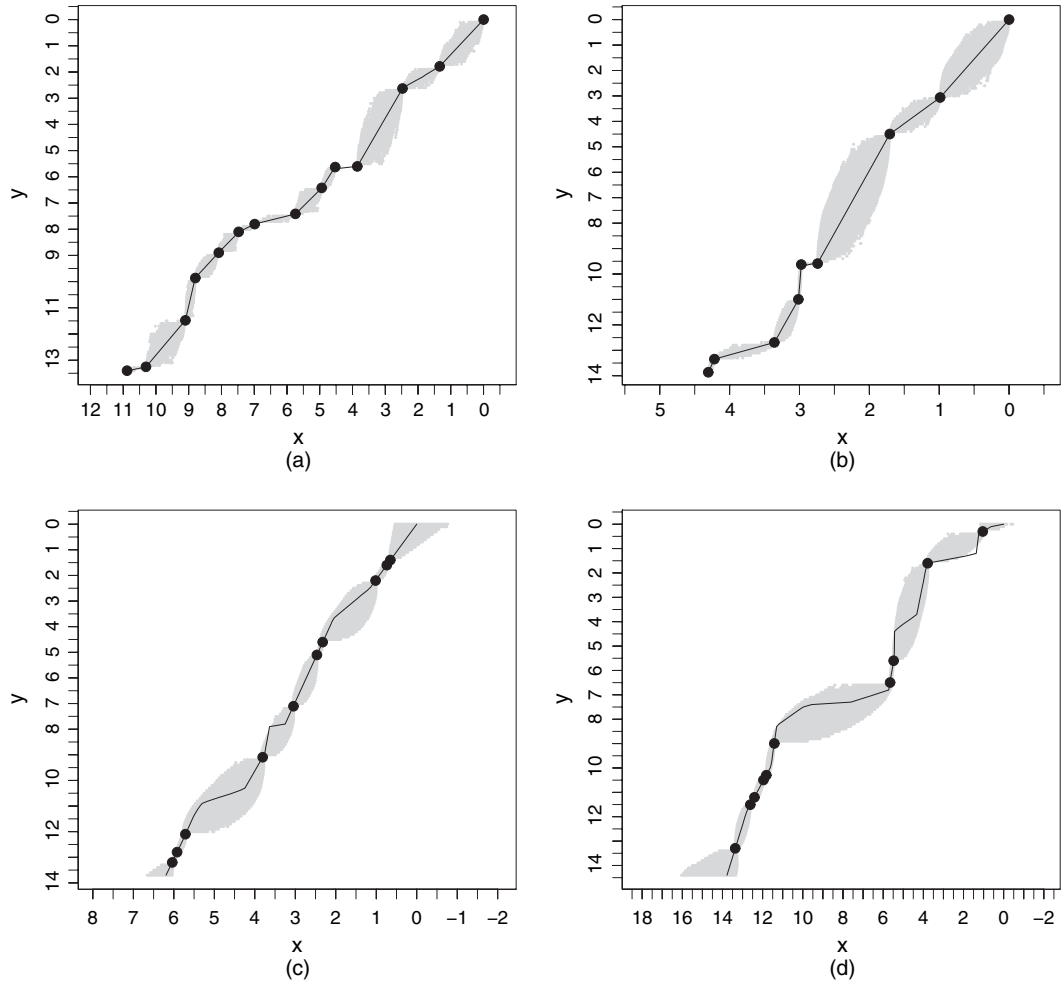

Fig. 5. Cross-validation sample paths for four different scenarios (▬, 95% HDR; ——, true path; •, data): (a) scenario (a); (b) scenario (b); (c) scenario (c); (d) scenario (d)

(d) The data are generated from a different monotone process. The data are not renewal points of the process. We use a bivariate truncated Gaussian renewal process and a bivariate log-Gaussian renewal process.

Our resultant sample paths are encouraging. In Fig. 5 we present the 95% highest posterior density intervals for a sample of each of the possible scenarios. In each case, the 'true' sample path is inside the bounds of our error intervals (with one small exception in scenario (d)). We repeat the sampling over 200 realizations (over four different y-values) creating a non-independent sample of 800. Variations of those scenarios outlined above give the coverage rates that are shown in Table 1. We note that these results include some situations where the model must extrapolate beyond the range of the given data. The scenarios show that the model fits well to various departures from the assumptions, with the worst results seen for processes that are generated when α is small; recall that the process is discontinuous when $\alpha \leqslant 2$. When $\alpha = 1$ the process is in fact a (linearized) bivariate Poisson process. Almost vertical segments have the

Table 1. Percentage coverage of 95% HDR in repeated generation of chronologies

Scenario	*% within 95% HDR*
(a) $\alpha = 4$	95.6
(b) $\alpha = 50$	97.0
$\alpha = 1$	87.6
(c) $\alpha = 50$	92.5
$\alpha = 1$	84.1
(d) Truncated Gaussian	96.7
Log-Gaussian	93.3

same probability as almost horizontal segments. Such a process is discontinuous; in fact the distribution of slopes has neither a mean nor a variance. We propose this as a realistic 'worst case', remarking that strong reasons would be needed to include discontinuity within the prior.

4. Radiocarbon dating and chronology modelling

In this section, we discuss the creation of calendar ages from given radiocarbon determinations. The process involves the technique of radiocarbon calibration via a given calibration curve which describes the relationship between calendar and radiocarbon years. In Section 4.1, we review the standard calibration method for creating calendar dates $\{\theta_j, j = 1, \ldots, m\}$. The technique can be naturally extended to the scenario where multiple dates are required, thus providing the opportunity for a joint prior distribution over the set of dates. It is this joint prior to which we attach our piecewise linear CPG process. In Section 4.2, we discuss techniques for dealing with dates which may, for reasons that are not connected with sedimentation, violate the monotonicity assumption. We conclude in Section 4.3 with a review of previous attempts at chronology modelling.

4.1. Radiocarbon dating

The radiocarbon age of a sample t_j is reported in radiocarbon years (written as ^{14}C years). The non-linear (and non-monotonic) transformation to calendar age θ_j (given in calendar years BP, present taken conventionally to be 1950) uses the published calibration curve $r(\theta_j)$ (Reimer *et al.*, 2005). The sources of uncertainty include

(a) the calibration curve, $r(\theta)$ itself, modelled as $r(\theta) \sim N\{\mu(\theta), \sigma^2(\theta)\}$ (the functions $\mu(\cdot)$ and $\sigma(\cdot)$ are known),
(b) laboratory errors, modelled with Gaussian uncertainty $t|\theta, \tau \sim N\{r(\theta), \tau^2\}$ and typically reported as $t \pm \tau$ (see Bowman (1999) and Aitken (1994)),
(c) reservoir effects where samples are believed to have come from non-standard sources of carbon (Stuiver and Pollach, 1977) and
(d) other errors, possibly induced by the laboratory treatment of the sample, or contamination of the sample itself (see Scott (2003)).

In the Bayesian format, given the model and radiocarbon data $t_j \pm \tau_j$, we obtain a posterior calendar age distribution via

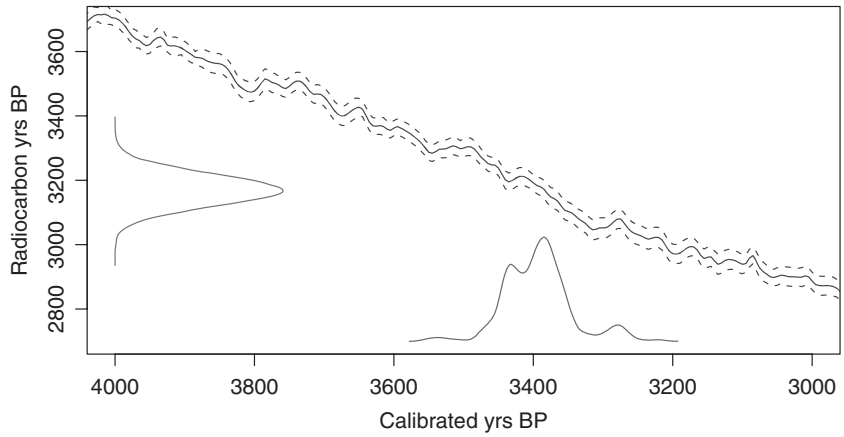

Fig. 6. Process of radiocarbon calibration

$$\pi(\theta_j|t_j,\mu,\tau_j,\sigma) \propto \pi(t_j|\theta_j,\mu,\tau_j,\sigma)\pi(\theta_j) \tag{12}$$

where $\pi(\theta_j)$ is a prior distribution on the unknown calendar age for the jth sample; note that the likelihood can be marginalized over $r(\cdot)$. For more detail, see, for example, Buck *et al.* (1996).

An example of radiocarbon calibration for an individual age is shown in Fig. 6. Here, the radiocarbon age 3180 ± 50 is shown in its original Gaussian format on the y-axis. This is transformed via the calibration curve (with dotted error bands) into the calendar age density on the x-axis. The multimodality in the calendar age distribution can be seen to arise directly from the non-monotonic nature of the calibration curve. Note also that the original modal age of 3180 ^{14}C years BP is transformed to just under 3400 calendar years BP.

In our case we have multiple samples and observe $\{(t_j \pm \tau_j; d_j); j = 1, \ldots, m\}$. We now write the calendar age as a function of depth to give $\theta(d_j)$. Interest focuses on the joint prior for the continuous monotone process $\theta(d)$, for which we adopt the piecewise linear CPG process. In particular we adopt the CPG distribution as the prior for the observed (depth, calendar age) increments; the independent increments assumption leads to a product likelihood. We allocate depth to the y-axis and time (being age) to the x-axis; this ensures continuity of sample paths with respect to time. Thus the rate process may now be thought of as the inverse of a sedimentation rate (i.e. years per unit depth). Its stochastic variation is piecewise constant, the number of changes in a given depth interval being Poisson distributed with depth durations following an exponential distribution. The sedimentation rates themselves are independent and identically distributed (being the ratio of exponential and gamma random variables). Their distribution is skewed; instances of very low rates are more common than are instances of very high rates. Of course, the rates and their durations are by construction not independent.

4.2. Outliers

Scott (2003) suggested that around 5% of all radiocarbon dates are outliers. Specifically, she suggested that the radiocarbon age uncertainty in some dates is an underestimate. Below we extend the standard model.

The standard approach is Christen (1994); $t_j|\theta_j,\sigma,\phi_j,\delta_j \sim N\{r(\theta_j)+\phi_j\delta_j,\sigma^2\}$. Here, ϕ_j is a binary flag parameter which identifies potential outlying dates, and δ_j is a shift parameter with prior distribution $\delta_j \sim N(0, b\sigma^2)$. Equivalently, t follows a distribution which is a binary mixture

of normal distributions with variances σ^2 and $(b+1)\sigma^2$. As in Blaauw and Christen (2005) we adopt $b=2$, and the simple prior distribution $\phi_j \sim \text{binomial}(1, 0.05)$, interpreting the posterior probability that $\phi_j = 1$ as the probability that the date in question is an outlier.

We allow a further type of outlier which allows the model to ignore completely certain dates which severely violate the monotonicity assumption. We thus use $t_j|\theta_j, \sigma, \phi_{1j}, \delta_{1j}, \phi_{2j}, \delta_{2j} \sim N\{r(\theta_j) + \phi_{1j}\delta_{1j} + \phi_{2j}\delta_{2j}, \sigma^2\}$, with ϕ_{1j} and δ_{1j} defined as above, but with $\phi_{2j} \sim \text{binomial}(1, 0.001)$ and $\delta_{2j} \sim N(0, 100\sigma^2)$. The larger prior variance that is given to δ_{2j} has the desired effect of removing the influence of extreme outlying dates.

4.3. Previous attempts at chronology reconstruction

The approach of Blaauw *et al.* (2003) involves a concept which is known as wiggle matching (Van Geel and Mook, 1989). Given linear sedimentation rates over arbitrary periods of time, the calibration curve can be used to identify the radiocarbon ages of the given dates. A weighted least squares or maximum likelihood approach provides best fit linear segments. Blaauw and Christen (2005) extended this approach to define a parametric linear piecewise prior for the calendar ages. Given a set of k changepoints they used Markov chain Monte Carlo (MCMC) sampling to determine their ages and depths. The resulting calculations produce a large number of posterior chronologies. Predictions with uncertainties for other depths are obtained by repeated linear interpolation between changepoints. An important aspect of this model is that it does not explicitly parameterize the radiocarbon dates that are given for the core; the model does not contain our defined parameters θ_j (being the calendar ages of the radiocarbon dates). Rather, it is concerned only with creating posterior distributions for the depths and ages of changes in sedimentation rate. The value of k is chosen by considering several possible values and selecting that which maximizes an information criterion. Note that k is somewhat unsatisfactorily not treated as random.

An alternative approach was presented by Bronk Ramsey (2007), implemented in the radiocarbon dating program Oxcal (Bronk Ramsey, 1995, 2001). In this method, a Poisson process is defined with respect to the age axis. It may be thought of as modelling sedimentation as the accumulation in a given period T of a Poisson number of instantaneous increments, of size g. The mean sedimentation rate is thus λg per unit time, where g is a fixed physical constant characterizing the site, and inference involves λ only. There is a similarity here to the CPG model in that infinitely large α and β such that $\alpha/\beta = k$ leads to a scaled Poisson distribution (i.e. $T/k \sim$ Poisson). Both processes have a Poisson number of increments; ours are of random size, whereas those for Bronk Ramsey (2007) are of fixed and prespecified size, with unclear guidance for its determination.

5. Case-studies

We present two case-studies of the model for sites at Sluggan Moss (Northern Ireland) and Glendalough (Ireland). These two show a wide variety of features that are relevant to the theoretical implications of our model.

We adopt the CPG formulation as a prior on the calendar age gaps; now $\Omega_j = \theta(d_j) - \theta(d_{j-1})$, for each of the observed radiocarbon dates $x_j \pm \tau_j$ at depths $\{d_j, j = 1, \ldots, m\}$. In effect, our model requires that changes in the linear sedimentation rate can occur randomly throughout the core but are required to occur at least as often as every radiocarbon-dated depth.

The parameters of our model include the CPG parameters, (β, λ), as well as those for each of the calendar dates. All are updated by using single-component Metropolis–Hastings steps. At each step a set of calendar age simulations is passed to the CPG process through which β and λ are updated. We adopt vague inverse gamma distributions with shape and rate 0.01 as prior

distributions for β and λ. Since these approximately control the variability and mean rate of the process sections respectively, it is conceivable that more informative distributions may be used if these were available *a priori*.

The model is implemented in C++ and R (R Development Core Team, 2006) via the `Bchron` package (which is available from `http://lib.stat.cmu.edu/R/CRAN/`). Model performance is highly dependent on the number of radiocarbon dates, since each additional determination requires the inclusion of five extra parameters (that for the date themselves, two binary flag parameters and two outlier shift parameters). The MCMC runs require Metropolis–Hastings updates for each of the parameters. In practice, the two CPG process parameters add very little computational load to the model.

For the Glendalough sequence, a run of 1 million MCMC iterations (which provides practically uncorrelated samples and passes convergence checks such as those which are implemented in BOA; Smith (2005)) requires around 80 min on a 3.2 GHz Dell Precision desktop computer. Shorter runs of length 200000 provide useful indications about the chronology and can be run in under 10 min. In practice, we have found that chronologies with a large number of (non-outlying) determinations often converge faster because the monotonic constraints that are imposed on them allow for less variability in the parameter values.

5.1. Sluggan Moss

This core, taken from Sluggan Moss (Smith and Goddard, 1991) in County Antrim, has been extensively dated. The full data set contains 40 radiocarbon ages of between 985 and 12470 ^{14}C years BP which lie at depths ranging from 44.5 cm to 518 cm. In some cases, multiple radiocarbon ages lie at the same depth; we use only those dates which purport to be representative of the whole sample. There are 23 such radiocarbon determinations.

The pollen record for this core gives an indication of the palaeo-environment throughout the Holocene. In particular, Smith and Goddard (1991) drew attention to changes in the pollen assemblage at various depths across the core which correspond to large environmental shifts. The chronology model allows us to examine the timing of these shifts. The output of the CPG model is shown in Fig. 7.

Further to the CPG output, we can also examine the effect of variability in the chronology via the pollen diagrams for which they are created. Important events, such as the coincidental rise in Betula and fall in Juniperus (which are linked to the start of the Holocene), or the subsequent fall in Betula and rise in Corylus (which are linked to the occurrence of warmer summers), are now shown to be of uncertain age. Following the approach of Smith and Pilcher (1973), we can identify the start of the Holocene as occurring at depth 452 cm, and the onset of warmer summers at 424 cm. Thus, using our process we can identify a 95% HDR interval for the start of the Holocene as (10.6–11.0) $\times 10^3$ calibrated years BP, and the onset of warmer summers as (9.94–10.55) $\times 10^3$ calibrated years BP.

Many further possibilities for analysis are available given the joint nature of the chronologies that we construct, especially those concerned with the timing between events. In particular, Fig. 8 shows the posterior distribution of the time lapse between two depths (10 cm apart) both with and without the joint structure. There is a clear reduction in uncertainty. We expand on these analyses elsewhere.

5.2. Glendalough

The core at Glendalough has been used as part of the Bayesian climate reconstruction project by Haslett *et al.* (2006), though the reconstructions were presented in ^{14}C years. The core is

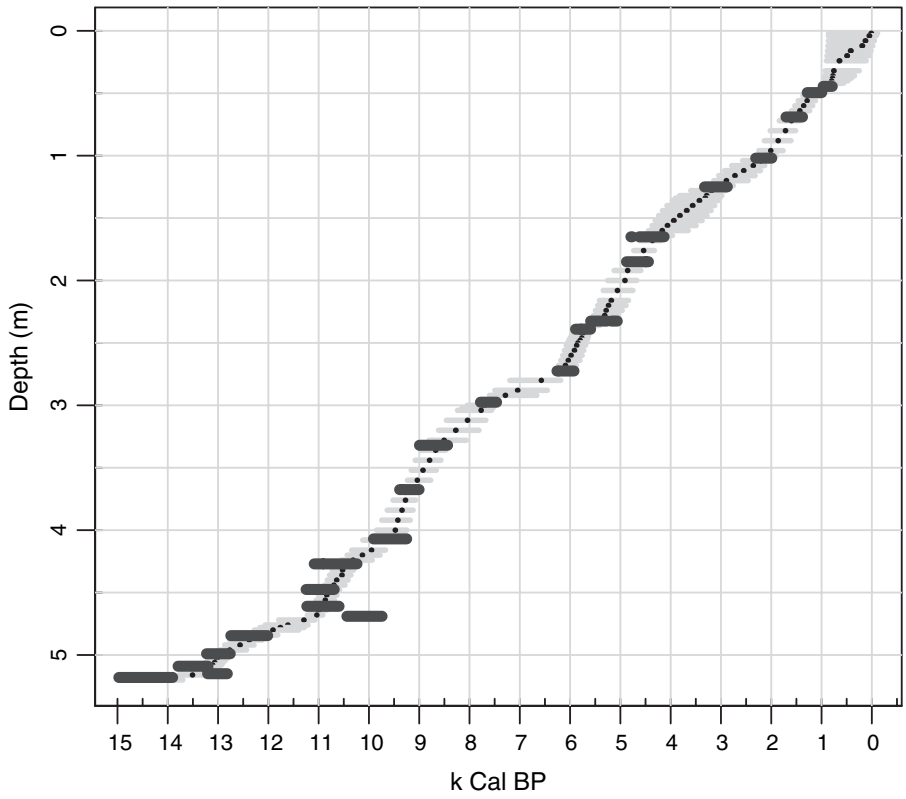

Fig. 7. CPG output for Sluggan Moss: ▬, 95% HDR regions for the uncertainty in the chronology model at each depth; •, modal estimate of the chronology; ▬, 95% HDR interval for the unconstrained calibrated ages that would be obtained if no prior model had been applied

interesting from a chronological viewpoint as it contains just five radiocarbon dates, with an additional restriction that the top of the core is the date of extraction.

A full run of the CPG model provides the chronology that is shown in Fig. 9(a). The small sample size that is present for the Glendalough core also induces a high degree of 'bowing' where little temporal information exists. We contrast this with Fig. 9(b), which shows the comparison between the CPG method and the best fit Blaauw and Christen method, known as BPeat. Clearly, the BPeat model ignores the extra uncertainty that is present when interpolating between dates.

As above, following the methods of Smith and Pilcher (1973), we can identify the start of the Holocene from the pollen record, occurring at an estimated 12 m. A 95% HDR interval for this depth is $(10.84–10.88)\times10^3$ and $(10.90–11.28)\times10^3$ calibrated years BP. Thus there appears to be some evidence of contemporaneity between the start of the Holocene at Sluggan Moss and Glendalough.

6. Discussion

We have introduced a new and simple monotone stochastic process $y(x)$ with piecewise linear sample paths. It is simply defined on a three-parameter CPG process, is Markov with respect to the y-axis and, under simple restrictions on a gamma shape parameter, has sample paths that are mean square continuous with respect to the x-axis.

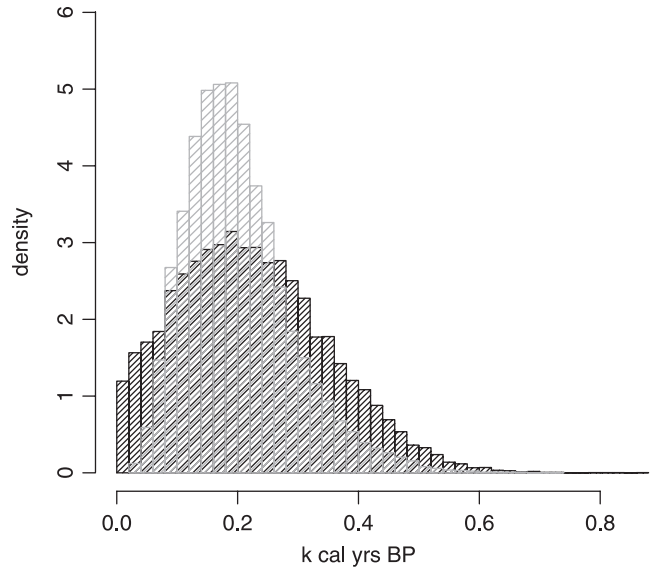

Fig. 8. Two overlaid histograms showing the posterior distribution of the time lapse between two depths utilizing joint structure (▨) and ignoring joint structure (▨)

We have focused on a Bayesian analysis of its simplest form, a two-parameter independent increments process, which corresponds to large values of the Poisson parameter. Given a sample of points on $y(x)$ we have shown that a close connection to the Tweedie distribution makes it simple to compute the likelihood. In particular, we can marginalize over the number and the lengths of the linear segments. It is therefore easy, given such data, to sample from the posterior distribution of the parameters. As it is particularly easy to sample from the conditional distribution of the process given data, it is thus straightforward to generate realizations of the sample path that are consistent with the data and the model. We have further shown by simulation that the model is robust, in the sense that such conditional realizations, given data, are consistent with the generating (monotone) stochastic process, even when this departs in several ways from the conditions of the independent increments CPG model.

We have subsequently embedded this within a problem of radiocarbon dating by adopting the CPG distribution as a prior for the true chronology (which is equivalent to the true sedimentation history). We have illustrated this with two contrasting examples. In this context, we show that the method has several desirable attributes. Elsewhere we examine in greater detail the construction of such chronologies via this and other methods. Further applications will arise in related work on palaeoclimate reconstruction where the possession of joint posteriors for entire chronologies is important.

It is clear that several extensions to the CPG methodology are possible. The most interesting of these is the dropping of the requirement that the data are at renewal points of the CPG, which is related to our requirement that λ is large, for now—given a sample of data—it is neither simple to compute the likelihood nor to sample realizations conditional on the data. Progress here would be welcome, for there then become clear connections to 'broken stick' piecewise linear regression and the ability to marginalize with respect to the number of breaks (or knots) would be an attractive way to bypass Dirichlet process or reversible jump methodologies. This is closely related to another extension, the relaxation of the assumption that α is known, for, unless λ is large, the insensitivity to α cannot be guaranteed.

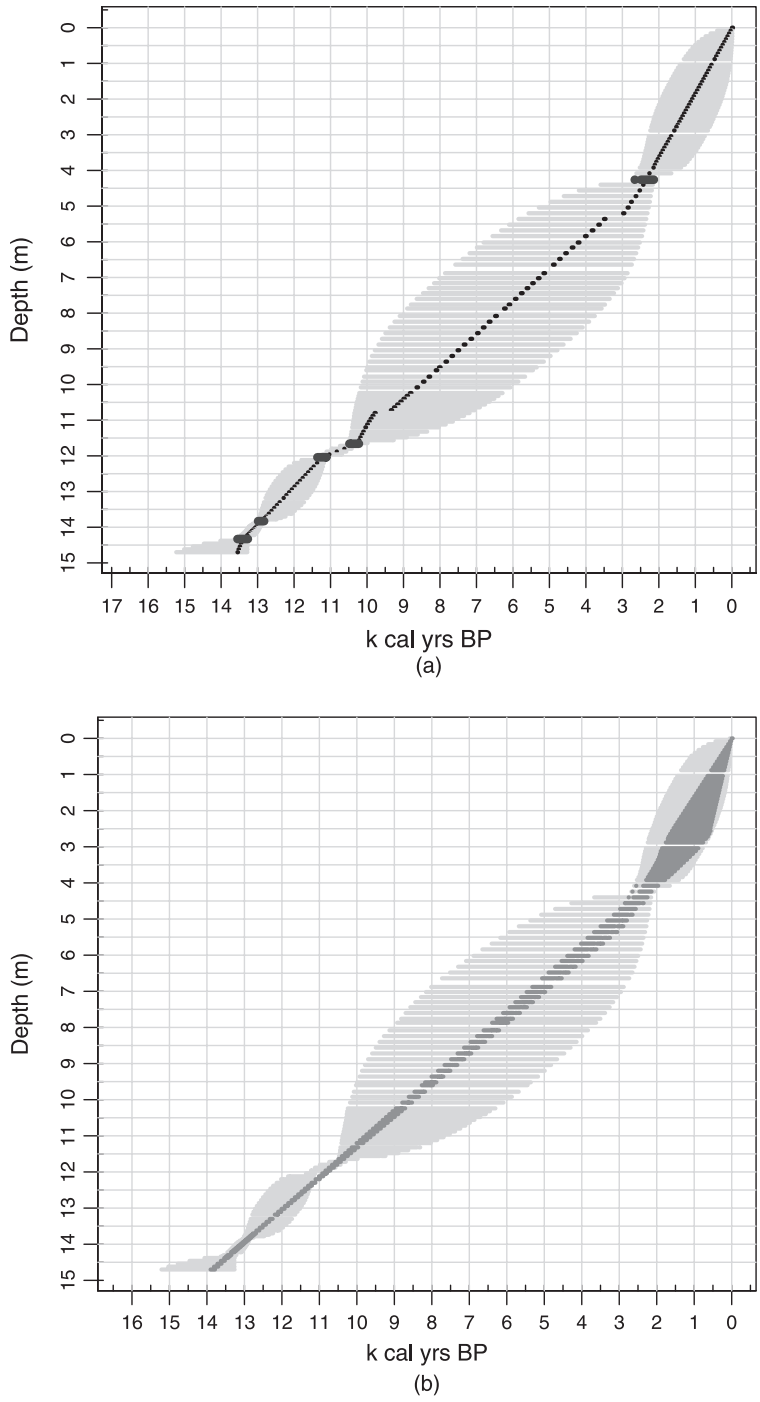

Fig. 9. (a) CPG output for Glendalough (▬, 95% HDR; •, mode; ▬, unrestricted calibrated dates; note that the radiocarbon determinations at depths 426 cm and 1166 cm both exhibit multimodal densities—thus the biggest mode (•) may shift discontinuously, giving the erroneous impression of discontinuous paths) and (b) comparison between the CPG method and the best fit Blaauw and Christen (2005) method with one changepoint (▬, 95% HDR CPG chronology; ▬, 95% HDR BPeat model)

Acknowledgements

This publication has emanated from research that was conducted with the financial support of Science Foundation Ireland.

References

Aitken, M. (1994) *Science-based Dating in Archaeology*. London: Longman.
Blaauw, M. and Christen, J. A. (2005) Radiocarbon peat chronologies and environmental change. *Appl. Statist.*, **54**, 805–816.
Blaauw, M., Heuvelink, G., Mauquoy, D., van der Plicht, J. and van Geel, B. (2003) A numerical approach to ^{14}C wiggle-match dating of organic deposits: best fits and confidence intervals. *Quat. Sci. Rev.*, **22**, 1485–1500.
Bowman, S. (1990) *Radiocarbon Dating*. London: British Museum.
Brezger, A. and Lang, S. (2006) Generalized structured additive regression based on Bayesian P-splines. *Computnl Statist. Data Anal.*, **50**, 967–991.
Brezger, A. and Steiner, W. (2005) Monotonic spline regression to estimate promotional price effects: a comparison to benchmark parametric models. In *Operations Research Proceedings* (eds H. Haasis, H. Kopfer and J. Schönberger), pp. 607–612. Berlin: Springer.
Bronk Ramsey, C. (1995) Radiocarbon calibration and analysis of stratigraphy: the OxCal program. *Radiocarbon*, **37**, 425–430.
Bronk Ramsey, C. (2001) Development of the radiocarbon calibration program OxCal. *Radiocarbon*, **43**, 355–363.
Bronk Ramsey, C. (2007) Deposition models for chronological records. *Quat. Sci. Rev.*, **27**, 42–60.
Buck, C. E., Cavanagh, W. G. and Litton, C. D. (1996) *Bayesian Approach to Interpreting Archaeological Data*. Chichester: Wiley.
Christen, J. A. (1994) Summarizing a set of radiocarbon determinations: a robust approach. *Appl. Statist.*, **43**, 489–503.
Davis, M. H. A. (1984) Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models (with discussion). *J. R. Statist. Soc.* B, **46**, 353–388.
Dellaportas, P., Friel, N. and Roberts, G. (2006) Bayesian model selection for partially observed diffusion models. *Biometrika*, **93**, 809–825.
Dunn, P. and Smyth, G. (2005) Series evaluation of Tweedie exponential dispersion model densities. *Statist. Comput.*, **15**, 267–280.
Dunson, D. (2005) Bayesian semiparametric isotonic regression for count data. *J. Am. Statist. Soc.*, **100**, 618–627.
Haslett, J., Whiley, M., Bhattacharya, S., Salter-Townshend, M., Wilson, S. P., Allen, J. R. M., Huntley, B. and Mitchell, F. J. G. (2006) Bayesian palaeoclimate reconstruction (with discussion). *J. R. Statist. Soc.* A, **169**, 395–438.
Heinricher, A. and Stockbridge, R. (1993) Optimal control and replacement with state-dependent failure rate: dynamic programming. *Ann. Appl. Probab.*, **3**, 364–379.
Holmes, C. and Heard, N. (2003) Generalized monotonic regression using random change points. *Statist. Med.*, **22**, 623–638.
Hunter, J. J. (1974) Renewal theory in two dimensions. *Adv. Appl. Probab.*, **6**, 220–221.
Jørgensen, B. (1987) Exponential dispersion models. *J. R. Statist. Soc.* B, **49**, 127–162.
Kaas, R. (2001) Compound Poisson distribution and GLMs—Tweedie's distribution. In *Handelingen van het Contactforum*, pp. 3–43. Brussels: Brussels University Press.
Kong, M. and Eubank, R. (2006) Monotone smoothing with application to dose-response curves. *Communs Statist. Simuln Computn*, **35**, 991–1004.
Lam, Y. and Zhang, Y. (2003) A geometric-process maintenance model for a deteriorating system under a random environment. *IEEE Trans. Reliab.*, **52**, 83–89.
Lee, M., Whitmore, G., Laden, F., Hart, J. and Garshick, E. (2004) Assessing lung cancer risk in railroad workers using a first hitting time regression model. *Environmetrics*, **15**, 501–512.
Neelon, B. and Dunson, D. (2004) Bayesian isotonic regression and trend analysis. *Biometrics*, **60**, 398–406.
Ramsay, J. O. (1998) Estimating smooth monotone functions. *J. R. Statist. Soc.* B, **60**, 365–375.
R Development Core Team (2006) *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
Reed, W. and McKelvey, K. (2002) Power-law behaviour and parametric models for the size-distribution of forest fires. *Ecol. Modllng*, **150**, 239–254.
Reimer, P. J., Baillie, M. G., Bard, E., Bayliss, A., Beck, W. W., Bertrand, C. J., Blackwell, P. G., Buck, C. E., Burr, G. S., Cutler, K. B., Damon, P. E., Edwards, L. L., Fairbanks, R. G., Friedrich, M., Guilderson, T. P., Hogg, A. G., Hughen, K. A., Kromer, B., Mccormac, G., Manning, S., Ramsey, C. B., Reimer, R. W., Remmele, S., Southon, J. R., Stuiver, M., Talamo, S., Taylor, F., van der Plicht, J. and Weyhenmeyer, C. E. (2005) Intcal04 terrestrial radiocarbon age calibration, 0–26 cal kyr bp. *Radiocarbon*, **46**, 1029–1058.
Scott, E. (2003) The fourth international radiocarbon inter-comparison. *Radiocarbon*, **45**, 135–408.

Smith, B. J. (2005) Bayesian Output Analysis program (BOA), version 1.1.5. University of Iowa, Iowa City. (Available from `http://www.public-health.uiowa.edu/boa`.)
Smith, A. G. and Goddard, I. C. (1991) A 12500 year record of vegetational history at Sluggan Bog, Co. Antrim, N. Ireland (incorporating a pollen zone scheme for the non-specialist). *New Phyt.*, **118**, 167–187.
Smith, A. and Pilcher, J. (1973) Radiocarbon dates and vegetational history of the British Isles. *New Phyt.*, **72**, 903–914.
Stein, M. (1999) *Interpolation of Spatial Data*. New York: Springer.
Stuiver, M. and Pollach, H. (1977) Discussions of reporting ^{14}C data. *Radiocarbon*, **19**, 355–363.
Tweedie, M. (1984) An index which distinguishes between some important exponential families. In *Statistics: Applications and New Directions: Proc. Indian Statistical Institute Golden Jubilee Int. Conf.*, pp. 579–604. Calcutta: Indian Statistical Institute.
Van Geel, B. and Mook, W. (1989) High-resolution ^{14}C dating of organic deposits using natural atmospheric ^{14}C variations. *Radiocarbon*, **31**, 151–155.
Wang, L. and Dunson, D. (2007) Bayesian isotonic density regression. *Technical Report 11*. Duke University, Durham.