Bayesian Analysis (2008) **3**, Number 2, pp. 249–254

Comment on article by Blackwell and Buck

John Haslett* and Andrew Parnell†

1 Introductory remarks

It is a pleasure to discuss not just an interesting paper but in fact an excellent project: the successful infiltration of a body of physical science by modern Bayesian methods. The story of that project is a classical mix of cutting edge statistical methodology combined with years of dogged ground work, carefully building credibility in the right places. It is a most suitable story for a Case Studies meeting.

From the point of view of Statistics per se, this project is an instance of a wider theme. From a theoretical point of view the theme is of an underlying stochastic process - here a Gaussian random walk - given which the likelihoods of the observations decompose multiplicatively reflecting conditional independence. MCMC provides an algorithm for the inversion. From the point of view of environmental statistics, the wider theme is of 'data synthesis' - scraps of data scattered through space and time, being of variable quality, and brought together *for inference* by the underlying latent process, taken to be 'smooth' in some sense, and modelling assumptions of conditional independence.

We believe that projects such as this add to the credibility of the wider Statistics community, and this Case Studies paper should assist in developing both wider themes. We expand below on some examples, necessarily reflecting some of our work, some of which is collaborative with Professor Buck. The applied theme, which we address first, is of course much wider than environmental statistics, but my pointers in other directions will be weaker.

2 Scraps

In the present paper the scraps are data on ^{14}C from objects that contain information on calendar age. The underlying space-time process is the varying amount of ^{14}C in the past atmosphere. In this specific context the spatial dimension plays no role, as the atmosphere is supposed to achieve perfect mixing of ^{14}C. Studies in which we have been involved include eg Haslett et al. (2006) where the focus is the climate of the past atmosphere of Ireland for the past ∼13,000 years for example, or more generally of regions such as Europe. Now the scraps are proxies: eg pollen in cores extracted from lake sediment, where changes in the relative frequencies of different types of pollen reflect vegetation response to climate change.

The objective of palaeoclimate reconstruction is more accurately described as using modelling to reduce the uncertainties about past climate. A key step is to regard past

*Department of Statistics, Trinity College, Dublin, Ireland, mailto:John.Haslett@tcd.ie
†Department of Statistics, Trinity College, Dublin, Ireland, mailto:Andrew.Parnell@tcd.ie

© 2008 International Society for Bayesian Analysis DOI:10.1214/08-BA309A

climate as smoothly changing in time and in space; care is needed however, for it seems clear that there have been relatively sudden changes in climate. We discuss below a non-Gaussian random walk.

A further part of that uncertainty rests with the attribution of age to samples via a depth age functional relationship, where typically only a few samples are ^{14}C dated. Here the underlying process is the sedimentation rate at a point in space. Now we note that the depth age relationship must be monotone, and another non-Gaussian process is needed.

In the present paper note also that the data are of qualitatively different types, including tree rings and ^{14}C measured by very different types of technology. Some samples have (supposedly well defined) calendar ages; others the authors describe as 'floating sequences'. In these latter, the durations in calendar years (relative age) are (supposedly) known, even though it is not possible to state with authority the absolute ages. Indeed, one of the major achievements of this work is the inclusion within the inference process of such data. It is difficult to imagine how, without such an approach, the ^{14}C community might have used, and been persuaded to use, such sources of information. Indeed relative age is often as important as absolute age to the substantive science and is in some ways easier to quantify, given the right modelling; see Figure 2 in the present paper.

Climatologists, for example, are often more interested in past *rates* of climate change than in the absolute date of, for example, the rapid transition to the Holocene at the end of the last glacial period. For instance, the authors of the recent IPCC report on climate change observe that "During the last glacial period, abrupt regional warming (probably up to 16°C *within decades* (my emphasis) over Greenland) occurred repeatedly over the North Atlantic region" (Jansen et al. 2007).

Going further, how much data are the authors *not* using? It seems clear that the ^{14}C community has been rigorous in the selection of data upon which the calibration curve has been built, using only high quality data with well defined calendar ages. But the boundaries have been pushed, and the authors describe how, in 2004, it became possible to use 'aggregated' data spanning a small range of calendar years; the methodology of Section 5.3 outlines the approach. But we suspect that the community has rejected truly vast amounts of data carrying *some* information on calendar age. See, for example, the process of translating relatively poor ^{14}C information into calendar ages, with appropriate caveats, as conducted by the IntCal group and their collaborators. Could (possibly large amounts of) data with relatively poor calendar age precision play any constructive role?

To take a rather extreme case, joint work with Buck and others presents findings with possible implications for calendar age determinations, based on pollen count data at four locations in the British Isles. This studies the calendar dates, derived from ^{14}C data, of the sudden (from a climatological point of view) rise in Alder pollen, taken to signify the onset of the Holocene. Denoting these calendar dates as $\theta_i; i = 1, 2, 3, 4$ they offer a distribution of $D = max(\theta_i) - min(\theta_i)$, a measure of the simultaneity of this rather dramatic climatological event. Although not as temporally well defined an event as a volcanic eruption, the IPCC evidence from Greenland ice for example (Jansen et al. 2007) that the very large climate transition took but a very few decades. It is thus surprising that the width of a 95% HPD for D is perhaps 3,000 years, for the magnitude of the climate change indicates that it must have occurred almost simultaneously across the British Isles.

Indeed this surprise is a source of information itself, information that reflects, inter-alia, on the ^{14}C data from which it is derived. In saying this we recognize that calibration uncertainties constitute but a part of the picture, in which a greater role must be played by the uncertainties in the dynamics of plant ecology, which dynamics led the Alder to expand from its refuges, encouraged by the new climate. But several thousand years? More generally, are the authors aware of untapped sources of information? Can it be tapped with their current methodology? If so, would the wider research community be prepared to accept it?

What are the other untold stories of this project? Applied statistics, Bayesian or not, must concern itself with data quality. We know for example that the determination of ^{14}C relies on another stage in calibration, that concerning the instruments in the labs themselves. We are aware that this itself has been the subject of statistical scrutiny (Scott 2003) and that outliers are not infrequently encountered. But what checks have been conducted on the quality of the supposedly known calendar ages? More specifically, and in the broader context of this *type* of Bayesian statistical modelling, what are the routine tools - akin to tests for outliers and to studies of influence in regression - available to applied Bayesian modellers of the caliber of the present authors? And what do they think of these tools?

3 Processes

We turn now to the underlying technology, the use of suitably smooth stochastic processes to act as joint prior. Note that the interest here, and in many of the other applications mentioned, lies in the *joint* posterior. In this context, this means that the authors wish to sample entire calibration curves *en bloc*; here it may be sufficient to sample the curve at annual intervals. Effectively this amounts to what may be called 'stochastic interpolation', by which we mean (at least sometimes) generating functions such as $\mu(\theta)$ at a resolution in θ rather finer than the data per se can give guidance. In some contexts this is called 'down-scaling'. We are told that the radiocarbon community has not yet signed up to more than point-wise summaries; marginal samples for each year would suffice for this. As we see, these are insufficient to deal with what the authors describe as the covariances in Section 6.

In climate research we have referred to such joint samples as being entire climate histories. They may be represented as (possibly multivariate) random functions in time at a location in space. In principle we aspire to complete realisations, that is for *all* times. In many applications samples at discrete time points will suffice, but sampling at arbitrarily fine resolution is desirable. In fact we aspire to entire space-time climate histories for Europe, which may be represented as a movie; but that is for another day.

The key is that prior knowledge imposes a structure on such joint samples from the posterior, and realisations are conditioned by data, perhaps loosely, at a finite number of points on the curve.

Gaussian processes have long been used for processes which may be regarded as evolving in time, for example Cox (1955). Nevertheless the area is still very active. One line arises in the context of diffusions associated with stochastic differential equations. In principle, studies of climate, past and future fall into this category, and indeed climate models are now being used to guide palaeoclimate reconstructions. But such models are still deterministic, and the use of deterministic models to help define stochastic process priors is less clear. In this particular case the Gaussian random walk in discrete time, as used here with constant variance for the increments, is amongst the simplest, and in the present context it seems to be more than adequate as a prior.

Another line of research is that being pursued by Havard Rue in Trondheim (Rue and Martino 2005). It presents the possibility of completely avoiding MCMC when working with Gaussian Markov Random Fields (GMRF) a trivial example of which is indeed the discrete random walk. The basic idea may be stated in terms that are not far from the present study. Given a GMRF for $\mu(\theta)$, with known variance parameters τ, and $[Y(\theta)|\mu(\theta)] = Poisson(\mu(\theta))$, then $[\mu(\theta)|Y]$ while not Gaussian, may be approximately Gaussian. When τ is not known, integration with respect to $[\tau|Y]$ may be achieved numerically for low dimensional τ. It may therefore be that the establishment of the calibration curve and its uncertainties may be achieved in a Bayesian setting without MCMC.

The converse, the task of subsequently using it to determine $[\theta|^{14}\mathrm{C}\ \text{data}]$ may not be so simple, as this distribution is manifestly non-Gaussian being possibly bimodal. But it may provide short-cuts. In the wider context of its use in projects in time and/or space (or more generally in a multidimensional space) it may be worth mentioning some limitations of Gaussian process priors. Two processes mentioned above - climate smoothness and the depth age relationship in cores - are not well catered for. Convenient alternatives do exist.

Rather dramatic changes in climate have almost certainly happened in the past and the use as a prior of the Normal distribution, with constant variance, most certainly over-smoothes climate reconstruction (Haslett et al. 2006). Several avenues are open. The simplest are long tailed Levy (independent increments) processes. The Normal Inverse Gaussian family (Barndorff-Nielsen 1998) is particularly attractive, for it has a well defined density and provides a smooth path from the Normal to the Cauchy. It may be thought of as a scale mixing of independent Normal random variables, whose variance is drawn from the Inverse Gaussian family. Other mixings (for example mixing via draws from a χ^2 distribution as in the t distribution as in Haslett et al. 2006) are not infinitely divisible; this inhibits working with data that are observed only irregularly in time.

The key to depth-age modelling is that the relationship is monotonic - deeper means older; further this is one fact that we can rely on in a world where so many other sources of information are subject to so much uncertainty. One approach: the sub-sampling of monotone sequences from the joint posterior of the random function of depth vs age, based on a Gaussian prior, is extremely inefficient; especially so when the data (here ^{14}C information constraining the curves) are well-spaced in the natural units of the study, here being years. Simple monotone stochastic processes are rare; processes with Gamma increments must be pure-jump processes in continuous time and are thus also inappropriate. We have recently proposed a very convenient model for such a process (Haslett and Parnell 2008)

One way to think of that way of modelling depth $Y(t)$ as a random function of time is as an integral over time of a sedimentation rate $S(t)$. If we define $S(t)$ as a positive piecewise-constant process and equip random rates and durations with suitable Exponential and Gamma distributions, the resultant process $Y(t)$ can be thought of as piecewise linear, including the possibility of almost flat patches. A Poisson process determines the breakpoints, so it is Markov. Crucially it is possible to marginalize with respect to this Poisson process, yielding a very simply described process. Indeed it seems to be sufficient to deal with an even simpler independent increments process.

The method is implemented in the R package `Bchron` and is available for download at `http://lib.stat.cmu.edu/R/CRAN/`. The software allows the user to create chronologies based on different types of dated sediments (including but not exclusive to ^{14}C dates), and produces posterior piece-wise linear functions as described above. Extra uncertainty is accounted for in terms of the measurement of the depth of the sediment and, more relevantly to this paper, by treating and correcting outlying radiocarbon determinations. Key to the treatment of outliers is the ability to use certain parts of the determination distributions according to their chronological information content.

4 Implementation

We are intrigued by the authors' remarks that their algorithm may not be the most efficient. We consider only the simplest case, with independent errors. Here ^{14}C determinations are denoted by a vector X whose elements can be modeled as $X_i = \mu(\theta_i) + u_i$ where the zero-mean errors are independent and Gaussian. Thus $X = \mu(\Theta) + U$, where U has known diagonal variance V_U. The θ terms are unknown but we observe $T = \Theta + W$, where W is also zero-mean with known variance matrix V_W. We seek $[\mu(\Theta), \theta | X, T]$ and study it by marginalizing drawing samples from this distribution. The challenge arises from the fact that the second moment structure of $\mu(\Theta)$ depends on the unknown Θ. But the full conditionals are

$$[\mu(\Theta)|X, T, \Theta] = [\mu(\Theta)|X, \Theta]$$

and

$$[\Theta|X, T, \mu(\Theta)] = [\Theta|T, \mu(\Theta)].$$

It seems that it should be possible to iterate between these, sampling sequentially for Θ and $\mu(\Theta)$, in Gibbs fashion, with block updating.

Indeed, it seems that this extension to the general case may be possible. Now Θ and $\mu(\Theta)$ are longer than X (to allow for the blocking, as in Section 5.2) and $X = A\mu(\Theta) + U$

for known A. Indeed as 'floating sections' yield ^{14}C data of the form of contrasts $X(\theta_1) - X(\theta_2)$ and calendar age data of the form $\theta_1 - \theta_2$, such an algorithm may also be available. The parameter r is of course unknown, but they already use Gibbs here and this should be possible within this formulation too. It is quite possible that we have missed some subtlety here.

5 Conclusion

In conclusion, the authors have show-cased a successful and important application of modern Bayesian modelling. They have enabled others to proceed to studies where ^{14}C dating uncertainties are just part of the battle. But more importantly, we hope their work will encourage others - particularly but not exclusively environmental statisticians - to rise the new challenges being presented by the availability of scraps of data in space and time.

References

Barndorff-Nielsen, O. (1998). *Processes of normal inverse Gaussian type*. Springer-Verlag. 252

Cox, D. R. (1955). "Some statistical methods connected with series of events." *Journal of the Royal Statistical Society - Series B*, 17: 129–164. 252

Haslett, J. and Parnell, A. (2008). "A simple monotone process with application to radiocarbon dated depth chronologies." *Journal of the Royal Statistical Society - Series C*, To appear. 253

Haslett, J., Whiley, M., Bhattacharya, S., Mitchell, F., Allen, J., Huntley, B., Wilson, S., and Salter-Townshend, M. (2006). "Bayesian palaeoclimate reconstruction." *Journal of the Royal Statistical Society - Series A*, 169(3): 395–438. 249, 252

Jansen, E., Overpeck, J., Briffa, K. R., Duplessy, J.-C., Joos, F., Masson-Delmotte, V., Olago, D., Otto-Bliesner, B., Peltier, W. R., Rahmstorf, S., Ramesh, R., Raynaud, D., Rind, D., Solomina, O., Villalba, R., and Zhang, D. (2007). "Palaeoclimate." In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L. (eds.), *Climate Change 2007: The Physical Science Basis. Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, 433–497. Cambridge University Press. 250, 251

Rue, H. and Martino, S. (2005). "Approximate Inference for Hierarchical Gaussian Markov Random Fields Models." Technical report, University of Science and Technology, Trondheim. 252

Scott, E. (2003). "The fourth international radiocarbon inter-comparison." *Radiocarbon*, 45(2): 135–408. 251