Application of a Hierarchical Model to Paleoenvironmental Time Series with Latent Times

Aaron Springford* David J Thomson†

Abstract

Observation times in time series data are usually assumed to be known and evenly spaced. These assumptions enable many methods for statistical inference of the measured process, such as ARMA models and spectral estimation using the Fast Fourier Transform. Thankfully, the measurement process is usually controlled by the experimenter, limiting the effect of these assumptions by design. However, for paleoenvironmental studies based on core data, measurement times must be inferred from depth and dating information (using radioisotopes, for example). Inference of measurement times is required for depths with and without dating information. We have previously described a method to estimate chronologies – the relationship between depth and time – that provides posterior distributions of sampling times (Springford 2013). In this paper we extend the results of Springford (2013) to examine the effect of sampling time uncertainty on time series analysis estimates.

Key Words: Time series, Chronology model, Hierarchical, Bayesian, Computation, Spectrum estimation

1. Introduction

Time series data can be expressed as

$$x(t(1)), x(t(2)), \ldots, x(t(N))$$

where the observations are represented by $x()$, and the times at which the observations were collected are represented by $t()$ such that $t(i) \leq t(j)$ for $i < j$. It is usual that the $\{t(i), i = 1, 2, \ldots, N\}$ are selected when designing the data collection program or experiment, and are thus considered to be known quantities. It is also usual that the time increments $d_i = t(i) - t(i-1)$ are equal for all $i = 2, 3, \ldots, N$. In fact, the large majority of popular time series analysis methods assume that the observations are collected at known and evenly spaced times – two examples are auto-regressive/moving-average (ARMA) models and spectrum analysis using the fast fourier transform.

However, the observer doesn't always control the observation times. Moreover, there are cases in which the observation times themselves are not known and this uncertainty should be considered when performing time series analysis.

Perhaps the most common examples of time series in which the times are not known are paleoenvironmental records based on cores extracted from material that has accumulated over time. For example, paleolimnology studies often collect cores from lakebed sediments. These cores are then sectioned into samples as a function of depth, and the samples are analyzed for proxies of past environmetal conditions (Last et al. 2001). Although the samples are collected at known depths, variations in the accumulation rate of the core material, compression of the core, uncertainties in dating estimates, and so on, mean that the corresponding times are unknown. In order to perform time series analysis on these data, it is necessary to estimate the unkown times. Moreover, the uncertainty in the estimates of

*Queen's University at Kingston, Department of Mathematics and Statistics, Jeffery Hall, University Avenue, Kingston, Ontario, Canada, K7L 3N6. Email: aaron.springford@queensu.ca

†Queen's University at Kingston, Department of Mathematics and Statistics, Jeffery Hall, University Avenue, Kingston, Ontario, Canada, K7L 3N6. Email: djt@mast.queensu.ca

the unknown times should be incorporated into the overall time series analysis (Blaauw and Heegaard 2012). We have previously described a method to estimate chronologies – the relationship between depth and time – that provides posterior distributions of sampling times (Springford 2013). In this paper we extend the results of Springford (2013) to examine the effect of sampling time uncertainty on time series analysis estimates.

2. Motivating dataset and methods

Following Springford (2013), the motivating dataset is from a study of climate on the Tibetan Plateau, in particular the strength of summer and winter monsoons in the region (Yu et al. 2006). The data come from a peat core that was sectioned into 649 one-centimetre intervals, covering roughly 11,000 years. Proxy records from each section were obtained, as well as ^{14}C isotope ratios from a subset of eleven sections. As described in Springford (2013), these raw ^{14}C age estimates were converted into likelihoods using the calibration software *Calib6.11* (Stuiver et al. 1998). The time series data $x()$ are the proxy measurements *greyscale* and *humification*, thought to represent the strength of summer monsoons, and the proxy measurement *ash content*, thought to represent the strength of winter monsoons (Yu et al. 2006).

2.1 Bayesian hierarchical chronology model

Details regarding the development of the chronology model can be found in Springford (2013). The chronology data consist of the section depths $\{d_i : i = 0, 1, \ldots, N\}$, assumed known, and the ages $\{a_j : j = 1, 2, \ldots, M\}$ with corresponding ageing errors $\{\sigma_j : j = 1, 2, \ldots, M\}$, from which the likelihoods $\{f(a_j, \sigma_j) : j = 1, 2, \ldots, M\}$ for the calibrated ages $\left\{a_j^c : j = 1, 2, \ldots, M\right\}$ were obtained. From the section depths, the depth increments are the first differences $\{v_k : k = 1, 2, \ldots, N-1\}$, $v_k = d_k - d_{k-1}$, each with a corresponding (latent) time increment u_k. We assume that $t_0 = 0$. To relate the ageing data with the latent times, let ϕ_j be the index k of t corresponding to the age a_j^c. Modelling the time increments and forcing the condition $u_k \geq 0$ for all k restricts the time-depth relationship to be monotonically increasing.

We assume that the standardized time increments follow the prior

$$w_k \sim \mathrm{N}(\mu, 1/\tau^2) \text{ for all } k \tag{1}$$

where μ is the mean and τ^2 is the information (reciprocal of variance) of the distribution. The likelihood is the probability distribution for the calibrated ^{14}C ages a_j^c, which are assumed to follow

$$a_j^c \sim f(a_j, \sigma_j) \tag{2}$$

where $a_j^c = t_{\phi_j} = \sum_{l=1}^{\phi_j} u_l = \sum_{l=1}^{\phi_j} w_l v_l$. These likelihoods were obtained using the calibration software *Calib6.11* (Stuiver et al. 1998).

Hyperpriors for the Normal prior (equation 1) are on μ and τ^2. Given τ^2, the conditional distribution for μ is

$$P(\mu | w_1, \ldots, w_{N-1}, \tau^2) \propto \exp\left(-\frac{(N-1)\tau^2}{2}\left(\mu - \frac{1}{N-1}\sum_{k=1}^{N-1} w_k\right)^2\right).$$

The conjugate hyperprior is

$$\pi(\mu | \mu_1, \tau_1^2) = \frac{\tau_1}{\sqrt{2\pi}} \exp\left(-\frac{\tau_1^2}{2}(\mu - \mu_1)^2\right) \tag{3}$$

with hyperparameters μ_1 and τ_1^2. The posterior for μ has hyperparameters

$$\begin{aligned} \mu_1^{'} &= \frac{\mu_1\tau_1^2 + \tau^2\sum_{k=1}^{N-1} w_k}{\tau_1^2 + (N-1)\tau^2} \\ \tau_1^{2'} &= \tau_1^2 + (N-1)\tau^2 \end{aligned} \tag{4}$$

Given μ, the conditional distribution for τ^2 is

$$P(\tau^2|w_1,\ldots,w_{N-1},\mu) \propto \tau^{N-1}\exp\left(-\frac{\tau^2\sum_{k=1}^{N-1}(w_k-\mu)^2}{2}\right).$$

The conjugate hyperprior is a Gamma distribution, however this hyperprior choice is not recommended in general because it is difficult to make weakly informative with respect to the posterior (Gelman 2006). We chose instead to use the uniform distribution on τ^{-1}

$$\pi(\tau^2|\sigma_L,\sigma_U) = \begin{cases} \frac{1}{\sigma_U-\sigma_L} & \sigma_L \le \tau^{-1} \le \sigma_U \\ 0 & \text{otherwise} \end{cases}$$

with hyperparameters σ_L and σ_U, $0 \le \sigma_L < \sigma_U$. For this hyperprior, the posterior is

$$\pi(\tau^2|\sigma_L,\sigma_U,\mu,w_1,\ldots,w_k) \propto \begin{cases} \tau^{N-1}\exp\left(-\frac{\tau^2\sum_{k=1}^{N-1}(w_k-\mu)^2}{2}\right) & \sigma_L \le \tau^{-1} \le \sigma_U \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

which was sampled numerically.

A hybrid Gibbs sampler was used to generate 1000 samples from the posterior distribution of model parameters as in Springford (2013). Time series analysis was then performed using each of the 1000 chronologies.

2.2 Time series analysis

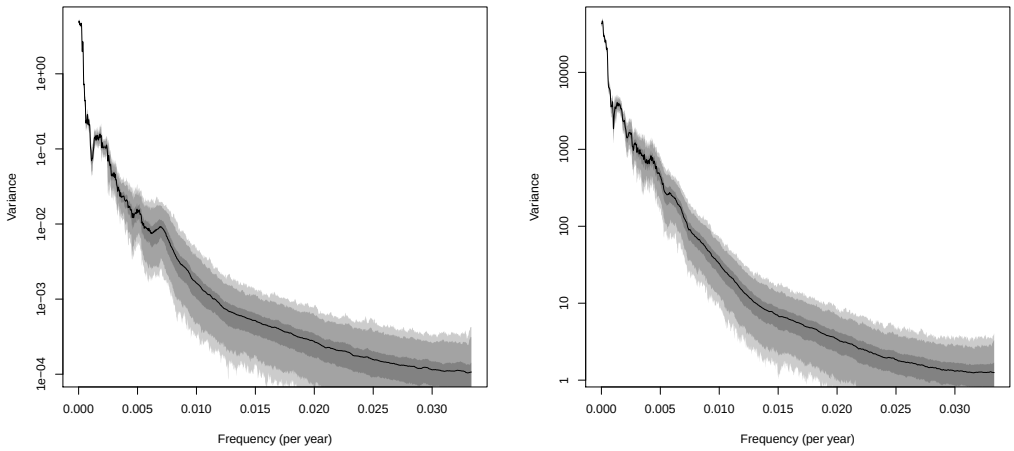
Because climate processes are well-known to exhibit periodic behaviour, we chose to investigate the effect of timing uncertainty on spectrum estimates of the proxies. We computed a multitaper spectrum estimate (Thomson 1982) for each of three proxies using the ensemble of 1000 equally probable time series. In order to make use of the fast Fourier transform when obtaining spectrum estimates, we resampled the time series using linear interpolation to obtain equally-spaced times for each of the 1000 series. This simple method is expected to perform nearly as well as more sophisticated methods, although it is expected to be biased slightly high when estimating variance at low frequencies and increase the amount of noise when estimating variance at high frequencies (Lepage 2009). The ensemble of 1000 spectrum estimates for each proxy represent the uncertainty introduced by the unknown sampling times.

3. Results

We show the results of our investigation into the effects of sampling time uncertainty on time series analysis. We investigated each of the three sampled proxies, but omit graphical results for *greyscale* which were characteristically similar to the results for *humification* and *ash content*.

The ensemble of 1000 spectrum estimates for the *humification* and *ash content* proxies are displayed in Figure 1. For each ensemble, we computed intervals containing 99, 95,

Figure 1: Ensemble of spectrum estimates showing the effects of uncertainty in sampling times. The left panel shows results for the proxy *humification* and the right panel shows results for the proxy *ash content*. The shaded areas, from lightest to darkest, contain 99, 95, and 50 percent of the spectrum estimates, respectively. The black line is the ensemble mean spectrum estimate.



and 50 percent of the spectrum estimates at each frequency, as well as the average spectrum value at each frequency. The results show that low-frequency signals are retained in spite of uncertainties in sample timing, but that the high-frequency portion of the spectrum is relatively featureless at frequencies above 0.01 cycles per year.

Multitaper spectrum estimation provides a way to characterize the periodicity of a signal at a given frequency using an F-statistic (Thomson 1982). The larger the F-statistic, the more a signal appears to be periodic. It is natural to consider both the variance of a signal at a given frequency as well as the periodicity of a signal at a given frequency when examining time series of paleoclimate because many processes in climate are periodic.

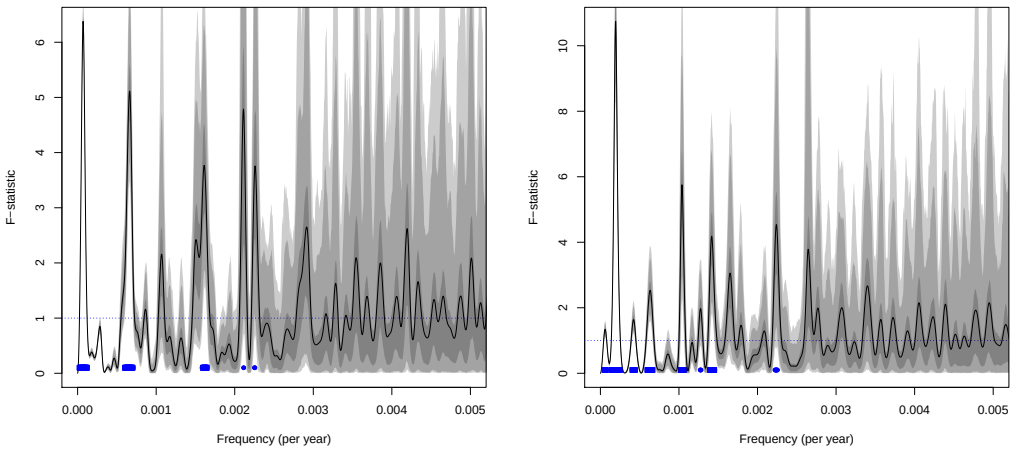
For each ensemble, we computed intervals containing 99, 95, and 50 percent of the F-statistic for periodicity at each frequency, as well as the average F-statistic value at each frequency. The results show that uncertainty in the value of the F-statistic due to sampling time uncertainty increases as the frequency increases (Figure 2).

The F-test results show some common frequencies being identified as significant between the three proxies. In particular, all three proxies were found to have significant variance on a period of 440-446 years. In addition, Humification and Greyscale showed matching significant variance at periods of 1508 and 14456 years. The rest of the significant periodicities seem to be proxy-specific: Ash Content has noticeably significant periodicities at 964 and 5120 years; Humification has noticeably significant periodicities at 474, and 620 years; Greyscale has a noticeably significant periodicity at 655 years. These results support the assertion of Yu et al. (2006) that the climatic dynamics of the Tibetan Plateau as reflected in the monsoon proxies are not simple.

4. Discussion

The focus of this analysis was on demonstrating the effects of sampling time uncertainty on a common time series analysis. One thousand posterior samples were obtained from the chronology model, and a spectrum estimate was obtained for each. We have demonstrated in this example that sampling time uncertainty can have a noticeable effect on time series

Figure 2: Ensemble of F-statistic estimates showing the uncertainty introduced by uncertainty in sampling times. The left panel shows results for the proxy *humification* and the right panel shows results for the proxy *ash content*. The shaded areas, from lightest to darkest, contain 99, 95, and 50 percent of the spectrum estimates, respectively. The black line is the ensemble mean F-statistic estimate. Also plotted for reference is the $F = 1$ line (dashed). Square points denote those frequencies with corresponding 99 percent F-statistic intervals that do not contain 1. Round points denote those frequencies with corresponding 95 percent F-statistic intervals that do not contain 1.



analysis and should not be ignored. Moreover, the effect is frequency dependent in the case of spectrum estimation, especially when considering the F-test for periodicity of Thomson (1982).

The ensemble of spectrum estimates can be viewed in one of two ways: (1) The estimate can be viewed as a parametric bootstrap procedure that is conditional on the posterior for the chronology model; or (2) The estimate can be viewed as a posterior distribution for the spectrum if we begin with an independent and flat prior on the spectrum for each of the three proxies. Here, we have avoided the issue of defining a proper prior for the three spectra associated with the three proxies, treating the analysis more as a parametric bootstrap procedure. Devising an appropriate prior is an obvious next step in the analysis. For instance, a prior that models the coherence between proxies (which is evident in this example) might be appropriate. Alternatively, a prior that treats certain known frequencies of interest separately may help to focus our inference on just those frequencies.

Examining the spectra for the three proxies, it seems that any structure at high frequencies is obliterated. This could be due to the uncertainties in the times, or due to the linear interpolation before calculating the spectrum, or both. It makes intuitive sense that structure at high frequencies would be lost if times are jittered - the standard error in time increments is on the order of nine years (Springford 2013), meaning that time increments are jittered over approximately forty years or more. When attempting to estimate the spectrum at frequencies shorter than 80-100 years, the effect of the jitter is apparently to smear out any details (Figure 1).

We view the example analysis presented as a first step in the development of a coherent framework for analysis of time series in which the times are latent. Future developments aim to round out this framework and improve estimates of latent times and time series models:

1. Determination of suitable prior(s) for the spectrum. Including a group-level prior on the spectrum might help with multiple comparisons issues that arise when examining the significance of many frequencies simultaneously. Instead of focusing on rejection of the null hypothesis (usually, a spectrum that is locally white) for each of the many frequencies present in the spectrum, the focus is on detection of frequencies that carry an appreciable amount of power compared to the larger group of frequencies (Gelman et al. 2012).

2. Examination of the use of time-series models (ARMA, state-space models). These time-series models provide a complementary approach to spectrum estimation, especially when modelling the underlying processes that gave rise to the data.

3. Obtaining combined time series model or spectrum estimates for multiple proxies from the same core using correlation or coherence. It could also be advantageous to allow core accumulation to be correlated with proxy measurements, since in many cases it would make sense that both accumulation and proxy values are driven by the same underlying process. In our example, the strength of monsoons might affect both the growth rate of peat and the measured proxies. A key challenge will be computational efficiency; depending on the prior structure used for the spectrum of each proxy, generating posterior samples will require repeated calculation of the spectrum.

4. Obtaining combined time series model or spectrum estimates for multiple proxies from different locations. Often it is the case that paleoclimate is examined regionally, using multiple records collected as different locations. The main challenge here will be determining a suitable model relating the proxies across sites. The default option might be similar to the case of multiple proxies from the same core - that is, provide no particular structure when estimating covariance or coherence between proxies. However, an alternative option is to model the strength of relationship between proxies parametrically, for example as a function of the distance between cores. This stronger class of assumption might lead to improved estimates, or easier interpretability of results, or both.

5. Acknowledgement

This work was funded in part by a scholarship from the National Science and Engineering Research Council of Canada.

References

Blaauw, M. and Heegaard, E. (2012), *Estimation of Age-Depth Relationships*, Springer, chap. 12, Developments in Paleoenvironmental Research.

Gelman, A. (2006), "Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)," *Bayesian analysis*, 1, 515–534.

Gelman, A., Hill, J., and Yajima, M. (2012), "Why we (usually) don't have to worry about multiple comparisons," *Journal of Research on Educational Effectiveness*, 5, 189–211.

Last, W. M., Smol, J. P., and Birks, H. J. B. (2001), *Tracking environmental change using lake sediments*, vol. 1-004, Dordrecht: Kluwer Academic Publishers.

Lepage, K. (2009), "Some advances in the multitaper method of spectrum estimation," Ph.D. thesis, Queen's University at Kingston.

Springford, A. (2013), "A Bayesian Hierarchical Chronology Model for Time Series Analysis of Paleoenvironmental Data," in *JSM Proceedings, Statistics and the Environment Section*, American Statistical Association, Alexandria, VA, pp. 3358–3369.

Stuiver, M., Reimer, P., and Braziunas, T. (1998), "High-precision radiocarbon age calibration for terrestrial and marine samples," *Radiocarbon*, 40, 1127–1151.

Thomson, D. (1982), "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, 70, 1055–1096.

Yu, X., Zhou, W., Franzen, L., Xian, F., Cheng, P., and Jull, A. (2006), "High-resolution peat records for Holocene monsoon history in the eastern Tibetan Plateau," *Science in China Series D: Earth Sciences*, 49, 615–621.