# BAYESIAN PERIODIC SIGNAL DETECTION APPLIED TO INTCAL98 DATA

V Palonen[1] • P Tikkanen

Accelerator Laboratory, P.O. Box 43, FIN-00014 University of Helsinki, Finland.

**ABSTRACT.** A Bayesian multiple-frequency model has been developed for spectral analysis of data with unknown correlated noise. A description of the model is given and the method is applied to decadal atmospheric INTCAL98 $\Delta^{14}$C data. The noise of the INTCAL98 data is found to be red, and there seems to be no support for continuous harmonic frequencies in the data.

## INTRODUCTION

The discrete Fourier transform (DFT) has been shown to be analogous to a Bayesian one-frequency model with several additional assumptions (Bretthorst 1988). The one-frequency model has been extended to take into account multiple sinusoids (Andrieu et al. 2001; Bretthorst 2003). Already these Bayesian models for spectral analysis have advantages compared to DFT. However, the analysis of proxy data with these models is hard because the models use parametric functions to handle the trend in the data. Their use is also doubtful because the models assume uncorrelated noise.

The power spectrum of uncorrelated noise is uniformly distributed for all frequencies. This is why it is also called white noise. Correlated noise has a power spectrum with more weight in the lower frequencies and is, therefore, called red noise. There is considerable evidence for red noise in the climatic time series (Gilman et al. 1962), as also a glance at power spectra of most proxies will reveal.

Our previous Bayesian model for spectral analysis assumed that the trend is random walk, i.e., we assumed red noise (Palonen and Tikkanen 2004). While it can be argued that this is a good assumption, we cannot be sure that the noise is red either. In this work, we develop a new Bayesian model where the noise color is allowed to vary. As an example of the new model, we analyze the decadal INTCAL98 atmospheric $\Delta^{14}$C data.

## MODEL

Let $M = \{m_1,\ldots, m_N\}$ be the set of $N$ measured data values with the corresponding uncertainties $S = \{s_1,\ldots, s_N\}$ recorded at different times $T = \{t_1,\ldots, t_N\}$. It is assumed that there is no uncertainty in $T$.

We will take the measurements $M$ to consist of a parametric function $g(t, \Theta)$ and noise:

$$m_i = g_i + \varepsilon_i + x_i \tag{1},$$

where $g_i = g(t_i, \Theta)$ and $\Theta$ is a vector containing all model parameters. The noise consists of measurement error $\varepsilon_i$, which is taken to be Gaussian white noise with given variances $s_i$, and of a first-order autoregressive noise $x_i$. The autoregressive noise represents all else in the data, i.e., the trend. It is given by the following equation:

$$x_i = \alpha x_{i-1} + \nu_i \tag{2},$$

[1]Corresponding author. Email: vesa.palonen@helsinki.fi.

979

where $\nu_i$ are normally distributed errors and $\alpha$ is the noise correlation coefficient, included as a parameter in the model. This enables analysis of data with noise of unknown color, from uncorrelated white noise ($\alpha = 0$) to heavily correlated red noise ($\alpha = 1$). We assume that only first-order correlations are present and that $\alpha$ is constant for the whole data set. We limit our model to first-order correlations instead of the whole correlation structure matrix in order to keep the model simple and computationally feasible.

Now, we would expect the variance of $\nu_i$ to scale linearly with the time step for red noise (random walk), and to remain constant for white noise. We will, therefore, include a linear change from a constant variance to a time-step-scaled variance according to the noise color parameter $\alpha$:

$$\nu_i \sim N\left(0, \left(\left|\frac{(t_i - t_{i-1})}{\langle t_j - t_{j-1} \rangle_j}\right| \times \alpha + (1-\alpha)\right)\Sigma^2\right) \tag{3},$$

where $\Sigma$ is the parameter representing the square root of the variance.

In order to model multiple sinusoids, the function $g$ is taken to be:

$$g(t, k, A_{1, \ldots, k}, \phi_{1, \ldots, k}, f_{1, \ldots, k}, m) = \sum_{j=1}^{k} A_j \sin(2\pi f_j t + \phi_j) + m \tag{4},$$

where $k$ is the number of sinusoids, and $A_j$, $\phi_j$, and $f_j$ are the amplitudes, phases, and frequencies of the sinusoids, and $m$ is the mean value of the data.

From Equations 1 and 2, we get:

$$m_i - g_i - \alpha(m_{i-1} - g_{i-1}) = \varepsilon_i - \alpha\varepsilon_{i-1} + \nu_i \equiv \beta_i \tag{5}.$$

The likelihood, namely, the probability of the data with known parameter values $\Theta$, is then:

$$p(\boldsymbol{M}|\Theta, \boldsymbol{S}, \boldsymbol{T}) = \prod_{i=2}^{N} \frac{1}{\sqrt{2\pi\sigma_{i,\,\text{tot}}^2}} \exp\left\{\frac{-\beta_i^2}{2\sigma_{i,\,\text{tot}}^2}\right\} \tag{6},$$

where

$$\sigma_{i,\,\text{tot}}^2 = s_i^2 + \alpha^2 s_{i-1}^2 + \left(\left|\frac{(t_i - t_{i-1})}{\langle t_j - t_{j-1} \rangle_j}\right| \times \alpha + (1-\alpha)\right)\Sigma^2 \tag{7}.$$

The joint posterior probability distribution for the parameters $\Theta$ can now be obtained from Bayes' theorem:

$$p(\Theta|\boldsymbol{M}, \boldsymbol{S}, \boldsymbol{T}) = \frac{p(\Theta)p(\boldsymbol{M}|\Theta, \boldsymbol{S}, \boldsymbol{T})}{p(\boldsymbol{M}|\boldsymbol{S}, \boldsymbol{T})} \tag{8},$$

where $p(\Theta)$ is the combined prior for the parameters and $p(\boldsymbol{M}|\boldsymbol{S}, \boldsymbol{T})$ is a normalization constant.

Table 1 shows the prior types and ranges used for each parameter. Normalized priors were used for all parameters. More information on the selection of priors in spectral analysis can be found in Bretthorst (1988).

Table 1  Priors for the parameters.

| Parameter | Symbol | Prior type | Prior range |
|---|---|---|---|
| Number of sinusoids | $k$ | uniform | [0, 10] |
| Noise color parameter | $\alpha$ | uniform | [−0.1, 1.1] |
| Standard deviation of RW step | $\Sigma$ | $\Sigma^{-1}$ | [0.01, 10] |
| Data mean | $m$ | uniform | [−100, 300] |
| Amplitude | $A_j$ | uniform | [0, 50] |
| Phase | $\phi_j$ | uniform | [0, 2π] |
| Frequency | $f_j$ | $f_j^{-1}$ | [1/5000, 1/20] |

As usual in Bayesian analyses, inference on the parameters $\Theta$ from Equation 8 cannot be made analytically and computational methods are needed.

**COMPUTATIONAL METHODS**

Inference on the posterior in Equation 8 is done with a reversible jump Markov chain Monte Carlo (MCMC) algorithm.

In MCMC, a chain of parameter space points distributed as the posterior distribution is constructed. This is done by using only the previous point $\Theta$ (hence the word Markov) in the chain to propose the next point $\Theta'$ at random (hence the words Monte Carlo) and then selecting or rejecting the proposed point according to some acceptance probability that is proportional to the posterior (Gilks et al. 1996). Reversible jump MCMC algorithms are capable of jumping between points in different parameter spaces, e.g., from parameter space of 1 sinusoid to parameter space of 2 sinusoids. In the present implementation, a jump to a new point is realized via one of the following 3 moves (Andrieu 2001); birth of a new frequency, death of an old one, or update of all at the moment existing parameters (normal MCMC move). The new point $\Theta' = (k', z')$ is then accepted with an acceptance probability (Green 1995; Waagepetersen and Sorensen 2001):

$$a(\Theta, \Theta') = \min\left(1, \frac{\text{likelihood}(\Theta')}{\text{likelihood}(\Theta)}\frac{\text{prior}(\Theta')}{\text{prior}(\Theta)}\frac{h(\Theta', \Theta)}{h(\Theta, \Theta')}\left|\frac{\partial(z', \boldsymbol{u}')}{\partial(z, \boldsymbol{u})}\right|\right) \tag{9},$$

where the function $h(\Theta, \Theta')$ is a proposal probability of a MCMC jump from $\Theta$ to $\Theta'$ with the random vector $\boldsymbol{u}$ used in construction of the new point, e.g., electing the amplitude, phase, frequency, and location of the new frequency. The random vector $\boldsymbol{u}'$ is used to construct $\Theta$ from $\Theta'$. Our algorithm differs from the one in Andrieu (2001) in that we also select the location of the new sinusoid among the old ones at random, thereby fulfilling the dimension matching requirement. Further details of the birth and death moves are discussed in the appendix.

The above form of the acceptance probability ensures the reversibility of the MCMC chain, and it has been shown that the chain will converge to the posterior, i.e., the points of the chain will be distributed according to the posterior, when the chain is long enough. When this is the case, calculating the expectation values for the parameters from the chain points, integrating over nuisance parameters, and getting the probability density as a histogram of the points is straightforward. An open problem is when the chain has converged. In this work, 7 separate chains starting from different points in the parameter space were simulated in order to strengthen the conclusion for convergence. The chains gave the same result. Trace plots do not show any problems either.

Our analysis on similar but synthetic data also ensured that the method finds frequencies when they are present. The method produces the synthetic frequencies well, but, as expected, smaller frequen-

cies in red noise begin to be less believable even if present in the synthetic data. We also checked the algorithm without data, i.e., with a constant likelihood. In this case, the posteriors for all parameters and the model order parameter $k$, in particular, were equal to the priors in Table 1.

## ANALYSIS OF INTCAL98 DATA

We will analyze the INTCAL98 past atmospheric $^{14}C$ concentration (Stuiver et al. 1998), shown in Figure 1. $^{14}C$ concentration is influenced by Earth's magnetic field, cosmic ray flux, solar activity, and various sinks and sources of carbon. Therefore, periodic oscillations in the $^{14}C$ concentration data are possible and, if present, are important in understanding the complex Sun-Earth system.
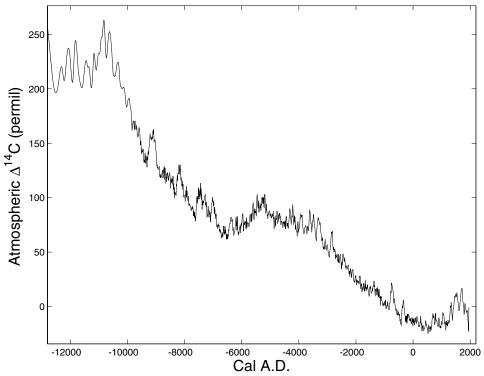


Figure 1 Atmospheric $^{14}C$ concentration. Vertical units represent a ‰ difference from the AD 1950 value. The data is taken from the INTCAL98 calibration measurements (Stuiver et al. 1998).

We made an MCMC simulation of $10^8$ points to obtain a posterior probability distribution of the model parameters for the INTCAL98 $\Delta^{14}C$ data from –9905 to 1955 cal AD. In Figure 2, a discrete probability distribution is shown for the number of sinusoids in the INTCAL98 $\Delta^{14}C$ data. The most probable number of sinusoids in the data is 0. The sinusoids, therefore, do not explain enough of the data to be believable with the present model and priors. This result contrasts with previous studies (Stuiver et al. 1991, 1993; Damon and Peristykh 2000) that used DFT or MEM with no or unspecified confidence levels and reported several frequencies in the INTCAL98 or INTCAL93 data.

A probability density for the noise color parameter $\alpha$ is shown in Figure 3. The noise is highly correlated red noise with $\alpha = 1.0006 \pm 0.0008$. Previous studies with DFT or MEM do not take this into account and, therefore, they probably have incorrect confidence levels, if any. The number of frequencies and the noise color roughly confirm our previous analysis with a random walk model
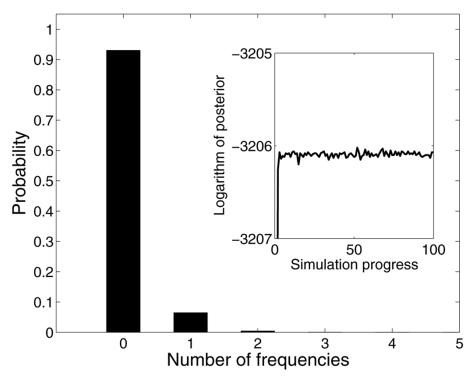
Figure 2  A probability distribution for the number of continuous harmonic frequencies in the INTCAL98 data. Inset shows the logarithm of prior times likelihood.
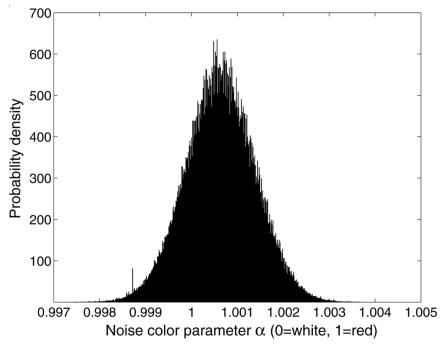


Figure 3  A probability distribution for the noise color parameter $\alpha$ for the INTCAL98 data

(Palonen and Tikkanen 2004) which is equivalent to $\alpha = 1$, although the selection of the prior range for the amplitude has an effect upon the model probabilities.

Probability distribution for period in the case $k = 1$ is shown in Figure 4. Because the distribution is smooth and the peaks are in accord with previous analyses, we conclude that there is sufficient mixing in the present MCMC chain.
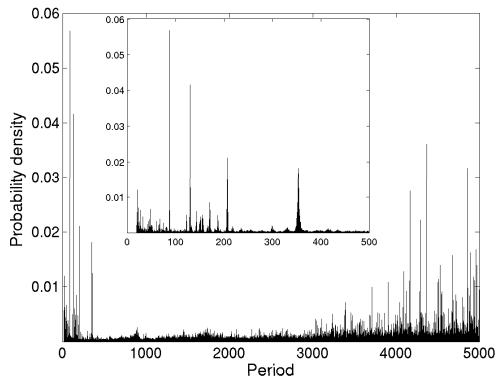


Figure 4  A probabilty distribution for the period assuming there is one frequency in the data ($k = 1$). The inset shows the lower-period part in more detail.

## CONCLUSIONS

The noise model in this work uses only first-order correlations. It is the simplest model capable of modeling white, pink, and red noise. The noise model is a step in the right direction—the data should determine the type of noise to be used. The present model also incorporates a simple and effective way to handle the trend. Artificial differentiation between the trend and noise is not needed. A problem of the present method is the computational implementation. For large parameter spaces, there are problems with convergence of the MCMC chain.

Not every peak is a real signal. The present model indicates that there are no believable continuous harmonic frequencies in the INTCAL98 $\Delta^{14}C$ data. However, there are indications that some frequencies in climatic proxies are not continuous (Ogurtsov et al. 2001) and, in these cases, continuous frequency models like the present model, DFT, or MEM are not the best option. Wavelets provide a better option but, like DFT and MEM, they also have a problem with confidence levels and noise color. We are currently working on a Bayesian model with noncontinuous sinusoids.

## APPENDIX

The total proposal distribution $h(\Theta, \Theta')$ is a product of the probability of a move from $k$ to $k'$ and the proposal distribution of the new point, $q(\Theta, \Theta')$. Because we use a uniform prior for the model parameter $k$, the move probabilities take a constant value $c$, except where new values of $k$ would be out of prior range, in which case the probability is 0. We use independent proposal distributions for the amplitude, phase, and frequency of the new sinusoid. The proposal distributions are uniform and truncated as the prior of the corresponding parameter. Therefore, the acceptance probability in Equation 9 for a birth move is simply:

$$a((k, z), (k+1, z')) = \min\left(1, \frac{p(\Theta'|\boldsymbol{M}, \boldsymbol{S}, \boldsymbol{T})c(k+1)}{p(\Theta|\boldsymbol{M}, \boldsymbol{S}, \boldsymbol{T})c(k+1)\frac{1}{R_A R_\phi R_f}}\right)$$

$$= \min\left(1, \frac{p(\Theta'|\boldsymbol{M}, \boldsymbol{S}, \boldsymbol{T})}{p(\Theta|\boldsymbol{M}, \boldsymbol{S}, \boldsymbol{T})\frac{1}{R_A R_\phi R_f}}\right),$$

where the $(k + 1)$ terms come from the selection of the frequency to be added or removed, and $R_A$, $R_\phi$ and $R_f$ are the prior ranges, e.g., $R_A = A_{max} - A_{min}$.

The probability for a remove-move is correspondingly:

$$a((k+1, z'), (k, z)) = \min\left(1, \frac{p(\Theta|\boldsymbol{M}, \boldsymbol{S}, \boldsymbol{T})\frac{1}{R_A R_\phi R_f}}{p(\Theta'|\boldsymbol{M}, \boldsymbol{S}, \boldsymbol{T})}\right).$$

## REFERENCES

Andrieu C, Djuric PM, Doucet A. 2001. Model selection by MCMC computation. *Signal Processing* 81:19–37.

Bretthorst GL. 1988. Bayesian spectrum analysis and parameter estimation. *Lecture Notes in Statistics, Vol. 48.* Berlin: Springer Verlag. 209 p.

Bretthorst GL. 2003. Frequency estimation, multiple stationary nonsinusoidal resonances with trend. *AIP Conference Proceedings* 659:3–22.

Damon PE, Peristykh AL. 2000. Radiocarbon calibration and application to geophysics, solar physics, and astrophysics. *Radiocarbon* 42(1):137–50.

Gilks WR, Richardson S, Spiegelhalter DJ, editors. 1995. *Markov Chain Monte Carlo in Practice.* Boca Raton: Chapman & Hall/CRC Press. 512 p.

Gilman DL, Guglister FJ, Mitchell JJM. 1962. On the power spectrum of red noise. *Journal of the Atmospheric Sciences* 20:182–4.

Green PJ. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82:711–32.

Ogurtsov MG, Kocharov GE, Lindholm M, Eronen M, Nagovitsyn YA. 2001. Solar activity and regional climate. *Radiocarbon* 43(2A):439–47.

Palonen V, Tikkanen P. Forthcoming. Spectral analysis of the INTCAL98 calibration curve: a Bayesian view. AMS-9 conference proceedings. *Nuclear Instruments and Methods in Physics Research Nuclear Physics B*.

Stuiver M, Braziunas TF, Becker B, Kromer B. 1991. Climatic, solar, oceanic and geomagnetic influences on late-glacial and holocene atmospheric $^{14}C/^{12}C$ change. *Quarternary Research* 35:1–24.

Stuiver M, Braziunas TF. 1993. Sun, ocean, climate and atmospheric $^{14}CO_2$: an evaluation of causal and spectral relationships. *The Holocene* 3:289–305.

Stuiver M, Reimer PJ, Bard E, Beck JW, Burr GS, Hughen KA, Kromer B, McCormac FG, van der Plicht J, Spurk M. 1998. INTCAL98 radiocarbon age calibration, 24,000–0 cal BP. *Radiocarbon* 40(3):1041–83.

Waagepetersen R, Sorensen D. 2001. A tutorial on Reversible Jump MCMC with a view toward applications in QTL-mapping. *International Statistical Review* 69:49–61.