Phil. Trans. R. Soc. Lond. A **330**, 601–616 (1990)
Printed in Great Britain

Time series analysis of Holocene climate data

By D. J. Thomson

AT&T Bell Laboratories, Murray Hill, New Jersey 07974, *U.S.A.*

Holocene climate records are imperfect proxies for processes containing complicated mixtures of periodic and random signals. I summarize time series analysis methods for such data with emphasis on the multiple-data-window technique. This method differs from conventional approaches to time series analysis in that a set of data tapers is applied to the data in the time domain before Fourier transforming. The tapers, or data windows, are discrete prolate spheroidal sequences characterized as being the most nearly band-limited functions possible among functions defined on a finite time domain. The multiple-window method is a small-sample theory and essentially an inverse method applied to the finite Fourier transform. For climate data it has the major advantage of providing a narrowband F-test for the presence and significance of periodic components and of being able to separate them from the non-deterministic part of the process. Confidence intervals for the estimated quantities are found by jack-knifing across windows.

Applied to ^{14}C records, this method confirms the presence of the 'Suess wiggles' and give an estimated period of 208.2 years. Analysis of the thickness variations of bristlecone pine growth rings shows a general absence of direct periodic components but a variation in the structure of the time series with a 2360-year period.

1. Introduction

To most scientists attempting to analyse their data, the subject of statistical time series analysis presents a bewildering array of methods and papers, each invariably purported to have or to show some optimal property. The purpose of this paper is to summarize some features of the multiple-window method as they apply to the analysis of Holocene climate data, to show some results of this analysis, and also to caution against some currently popular spectrum estimation methods whose use, I believe, can result in misleading conclusions. The analysis of Holocene climate data merits special attention as it is beset by a combination of problems that occur in other disciplines, but rarely in such concentration. Given current speculations on the 'greenhouse effect' and related global climate problems, it is mandatory to provide an accurate historical base against which changes may be measured. Specific problems affecting the time series part of data analysis, as opposed to the more difficult problems of interpretation and establishing causality, are the following.

1. Most of the available records are proxies and, as such, record a combination of effects. Changes in ^{14}C, for example, record a confounding of solar activity, geomagnetic changes, and various reservoir effects. Reliable analysis requires proper attribution.

2. Geophysical data almost invariably include a combination of periodic, or almost periodic, signals embedded in a non-deterministic, or noise-like, background. The background signal is usually 'red' that is, it has much more power at low frequencies than at high. The Wold decomposition theorem (Doob 1953), requires that the deterministic and random components be treated separately. Statistically, this must be done as well, but with the complication that both components be estimated simultaneously.

3. Sampling times in geophysical records are often unavoidably sporadic and, in addition to being irregularly sampled, the record may contain large gaps. It is usually assumed that cautious interpolations will allow the recovery of low-frequency terms, but the incommensurable periods of many known effects and large power variations at high frequencies exacerbate the inability to do anti-alias filtering.

4. Climate data contain statistical outliers both from the inevitable recording and measurement errors and also from natural causes. The former, as demonstrated in Kleiner *et al.* (1979), are bad enough; the latter may be seen even more insidious, for example, it has been suggested (see, for example, De Mendoça Dias 1962) that volcanic and solar activity are related. Thus long-term climate data may contain solar effects observed through indirect as well as more conventional mechanisms.

5. Radiometric timescales contain random inaccuracies whereas those obtained from dendrochronology may have jumps. Both have the effect of making a spectrum appear unnecessarily complicated.

6. There is much circumstantial evidence, and even more claims in the literature, for numerous low-level and closely spaced periodic signals in the climate and solar activity records (see, for example, the reviews by Pittock (1978), Shapiro (1979) and Sonett (1984)). Most of the suspected periods are unknown, not simply related, and the existence of closely spaced lines makes confirmation difficult.

7. The length of available data records is short in comparison with the frequency separation of the various suspected lines. Perversely, there is usually too much data to allow numerical calculation of 'exact' procedures.

8. Data for the complete Holocene, even if available, cannot be treated as a stationary process both because of the large residual transient from the last Pleistocene glaciation and to continuing Milankovitch processes. Further, in §7, I give evidence for a process with a 'period' of about 2400 years that alters the structure of the process.

In this paper we cannot address all the above in detail, and so concentrate on those where the multiple window approach is advantageous. The outline of the paper is as follows: the next section gives an abbreviated background on stationary processes, followed by the fundamental equation, with §4 outlining an inverse theory solution of it. Section 5 describes procedures for line detection and mixed spectra, §6 covers miscellaneous topics, with §§7 and 8 containing the examples. Finally, §9 discusses some of the problems with other spectrum estimation procedures in current use.

2. Background on stationary processes

The power spectral density, or power spectrum, or simply the spectrum of a random process is firmly established as the most useful description of most time series encountered in the physical sciences; its definition, and the accompanying spectral representation, despite their general description by Cramér in 1940, are unfortunately less familiar. For a discrete-time harmonizable process with unit time step (so the Nyquist frequency is $\frac{1}{2}$) the *Cramér* or *spectral* representation is written as a generalized Fourier transform,

$$x_t = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{e}^{\mathrm{i}2\pi\nu t}\,\mathrm{d}X(\nu), \tag{1}$$

and describes the process for all time. If the process is stationary, i.e. if its statistics are invariant

under changes of time origin, dX has zero mean, energy at different frequencies is uncorrelated, and the expected value of $|dX(f)|^2$ is, by definition, the spectrum $S(f)\,df$. Detailed properties are available in numerous books (see, for example, Priestley 1981; Brillinger 1975). For many problems in the physical sciences, however, a better, if even less well-known, approach is given by the extended, or Munk–Hasselman (Munk & Hasselman 1964), representation where the deterministic component, or first moment, is explicitly given by

$$\mathrm{E}\{dX(f)\} = \sum \mu_j \delta(f-f_j)\,df, \tag{2}$$

where δ is the Dirac delta-function, the f_j are the frequencies of periodic, or line, components, and μ_j their amplitudes. The continuous part of the spectrum is defined by

$$S(f)\,df = \mathrm{E}\{|dX(f) - \mathrm{E}\{dX(f)\}|^2\}, \tag{3}$$

E denoting the statistical expected value operator. That is, the non-deterministic part is given by the second central moment. Such processes are known as centred or conditionally stationary (as opposed to strict or covariance stationary) and are often noted as having mixed spectra. The distinction between first and second moment properties is exceptionally important.

First moments correspond loosely to what scientists usually term harmonic analysis, that is the study of periodic phenomena. Typically, a process will contain a few such lines, each described by their amplitude, frequency, phase, and perhaps frequency drift rate. Consequently, such parameters may be estimated by using the well-established methods of maximum-likelihood. Moreover, super-resolution, i.e. discrimination of frequencies spaced closer than the Rayleigh resolution of $1/T$, T being the total observation time, is possible in principle (and indeed in practice to some extent) and the accuracy with which closely spaced lines may be estimated is a function of the signal:noise ratio, defined as the ratio of power in the first moment to power in the second moment as a function of frequency.

Second moments, on the other hand, are 'noise-like' and are the non-predictable or non-deterministic part of the process. An important property is that energy at different frequencies is uncorrelated, $\mathrm{Cov}\{dX(f), \overline{dX(g)}\} = 0$ for $f \neq g$, Cov the covariance operator. Here, in contrast to the line spectrum of the first moments, the spectrum is typically continuous and often smooth. Because the second-moment spectrum cannot vanish over an interval unless the process is deterministic, it is often referred to as the continuum spectrum. Also in contrast to the first moments, one is now trying to estimate a function of frequency, not just a handful of parameters. Consequently, one is *not* doing maximum-likelihood estimation but a quadratic counterpart of Backus–Gilbert inverse theory. Again, there is a distinction in achievable resolution. It is essentially impossible to resolve details of the continuum spectrum separated by less than $1/(2T)$, half the Rayleigh limit. Typically, useful resolution is between $2/T$ and $50/T$, much poorer than Rayleigh.

Confusing the two moment properties will result in absurdities like 'smoothing' line spectra and, equally unwarranted, applying super-resolution criteria to noise-like processes. Finally, one must remember that, classically, spectra are defined only for stationary processes. However, I make the usual assumption that the structure of the process varies slowly enough to assume local stationarity (see Priestley 1988; Daubechies 1988).

3. The fundamental equation

I assume that there are N observations from a stationary series, $x_0, x_1, \ldots, x_{N-1}$, at equally spaced sample times $0, 1, \ldots, N-1$. Because my goal is to estimate the statistics, in particular the first *two* moments of $\mathrm{d}X$, we take the Fourier transform of the available data,

$$\tilde{x}(f) = \sum_{t=0}^{N-1} x_t \, \mathrm{e}^{-\mathrm{i}2\pi ft}, \tag{4}$$

and combine it with the spectral representation (1): the result is the fundamental equation of spectrum estimation

$$\tilde{x}(f) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{K}_N(f-\nu) \, \mathrm{d}X(\nu), \tag{5}$$

where the kernel is given by

$$\boldsymbol{K}_N(f) = \sum_{t=0}^{N-1} \mathrm{e}^{-\mathrm{i}2\pi ft} = \exp\left[-\mathrm{i}2\pi f(N-1)/2\right] \frac{\sin N\pi f}{\sin \pi f}. \tag{6}$$

Note that $\tilde{x}(f)$ may be inverse transformed to recover the data and is thus a trivially sufficient statistic; the same is not true for $|\tilde{x}(f)|^2$. Equation (5) is traditionally treated as a convolution; the difference here is that we choose to treat it as a Fredholm integral equation of the first kind and, although it does not have a unique solution, attempt to find an approximate solution with 'reasonable' statistical properties.

4. Spectrum estimation as an inverse problem

In multiple-window spectrum estimation one chooses a bandwidth W and constructs a local least-squares eigensolution of the fundamental equation (5) on the interval $(f-W, f+W)$ for a set of frequencies f. The key to this solution is a set of special functions known as Slepian functions, or discrete prolate spheroidal wave functions, that are the time-limited functions whose Fourier transforms are most concentrated in frequency. Details are available in Slepian (1978), and references given therein. Slepian defines the discrete prolate spheroidal sequences $v_n^{(k)}(N, W)$ as an orthonormal solution to a matrix eigenvalue problem. Their Fourier transforms,

$$V_k(f) = \sum_{n=0}^{N-1} v_n^{(k)}(N, W) \, \mathrm{e}^{-\mathrm{i}2\pi fn}, \tag{7}$$

are solutions of the integral equation,

$$\lambda_k \, V_k(f) = \int_{-W}^{W} \boldsymbol{K}_N(f-\nu) \, V_k(\nu) \, \mathrm{d}\nu, \tag{8}$$

where the kernel is again given by (6). To agree with common definitions of the discrete Fourier transform, I use V in place of Slepian's U with

$$V_k(f) = (1/\epsilon_k) \exp\left[-\mathrm{i}\pi f(N-1)\right] U_k(-f),$$

where ϵ_k is 1 for k even, i for k odd, and the dependence of the λs, Us and Vs on N and W have been suppressed. Note carefully that the fraction of energy retained within the band $(-W, W)$ by the kth function is given by λ_k, and that there are about $K = [2NW]$ functions with eigenvalues near 1. These K functions each have most of their energy concentrated within this band, and thus are classified as 'good' windows from the point of view of spectrum estimation.

Although details of the solution are available in Thomson (1982), and, in easier to read form, in Park *et al.* (1987*b*), a summary is given here. Because the Slepian functions form a complete set, I assume that the observable portion of $\mathrm{d}X$ has the expansion,

$$\mathrm{d}X(f-\nu) = \sum_{k=0}^{\infty} x_k(f) \cdot \overline{V_k(\nu)}\, \mathrm{d}\nu, \tag{9}$$

for $|\nu| < W$. By the assumed stationarity, the part of $\mathrm{d}X$ in the inner or local domain, $(f-W, f+W)$ is uncorrelated with $\mathrm{d}X$ on the rest of the frequency domain, and it is helpful to think of the 'inner' energy as 'signal', the rest as 'noise'. Using the basic integral equation and orthogonality properties one obtains the expansion or eigencoefficients:

$$x_k(f) = \sum_{n=0}^{N-1} \mathrm{e}^{-\mathrm{i}2\pi f n} \cdot v_n^{(k)}(N, W) \cdot x_n. \tag{10}$$

The expansion coefficients are obtained by windowing the data with a Slepian sequence and taking the Fourier transform. This is identical to the better windowed spectrum estimates except, conventionally, only the first term in this expression has been used. Again I emphasize that the $K = [2NW]$ coefficients with eigenvalues $\lambda_k \approx 1$ are used to represent the information in the signal projected onto the local frequency domain for all subsequent inference. For example, a crude multiple-window spectrum estimate is

$$\bar{S}(f) = \frac{1}{K} \sum_{k=0}^{K-1} |x_k(f)|^2. \tag{11}$$

Temporarily restricting ourselves to the continuous part of the spectrum, the dependence on the bandwidth W of the estimate may be deduced from the existence of $K \approx [2NW]$ windows with eigenvalues near 1. If the spectrum is flat within the local domain the coefficients are uncorrelated (because the windows are orthogonal) and each contributes two degrees of freedom; thus estimate (11) has $2K$ degrees of freedom. If W is too small one has poor statistical stability, but if W is too large the estimate has poor frequency resolution. Typically W is chosen between $1/N$ and $20/N$ with a time-bandwidth product NW of 4 or 5 being a common starting point. (Thus $W = 4/N$ or $5/N$ with corresponding K usually taken conservatively as 6 or 8 giving estimates with 12 or 16 degrees of freedom.) The case $NW = 4$ is interesting for comparison: the bandwidth is the same as that obtained when a Parzen (1957) or Papoulis (1973) window is used as a data taper except that one obtains 10–12 degrees of freedom instead of 2. Obviously a fast Fourier transform (FFT) may be used for efficient computation of the expansion coefficients. Also, the Slepian sequences are eigenvectors of a simple symmetric tridiagonal matrix, equation (14) of Slepian (1978), and so are trivial to compute.

In practice, one uses a local least-squares solution for the integral equation instead of simply truncating the series after the first K coefficients. This gives a data adaptive weighting with superior protection against leakage and bias. An approximate solution is

$$\sum_{k=0}^{K-1} \frac{\lambda_k(\hat{S}(f) - \hat{S}_k(f))}{[\lambda_k \hat{S}(f) + (1-\lambda_k)\sigma^2]^2} = 0, \tag{12}$$

where $\hat{S}_k(f) = |x_k(f)|^2$ and σ^2 is the process variance. Details are given in §V of Thomson (1982). Note that this is the estimate usually used: form (11) is useful for an introduction, (12) is more accurate and reliable. Note also that estimate (12) is not a simple quadratic form but

the solution of an implicit equation involving a ratio of two such forms. One consequence is that the estimate of autocovariance, $\tilde{R}(\tau)$ obtained by taking the Fourier transform of the spectrum obtained by solving (12),

$$\tilde{R}(\tau) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{S}(f)\, \mathrm{e}^{\mathrm{i}2\pi f\tau}\, \mathrm{d}f,$$

can be non-zero for lags, τ, greater than N.

5. Harmonic F-test for periodic signals

As emphasized earlier, a major problem with most spectrum estimates is that they ignore differences between first and second moments. Here the multiple-window method is advantageous because it allows us to separate line and continuum components and permits an analysis-of-variance test for significance of possible line components. Given a periodic signal at frequency f_0 the expected value of the eigencoefficients is, by (2) and (7)

$$\mathrm{E}\{x_k(f)\} = \mu V_k(f-f_0). \tag{13}$$

The complex amplitude is estimated by minimizing the residual local least-squares error,

$$e^2(\mu,f) = \sum_{k=0}^{K-1} |x_k(f) - \mu(f)\, V_k(0)|^2, \tag{14}$$

with respect to μ, giving

$$\hat{\mu}(f) = \sum_{k=0}^{K-1} \overline{V_k(0)}\, x_k(f) \Big/ \sum_{k=0}^{K-1} |V_k(0)|^2, \tag{15}$$

which is simply a linear regression of the $x_k(f)$s on the $V_k(0)$s. The fit may be tested for significance by the usual regression F-test, the ratio of the energy explained by the assumption of a line component at the given frequency to the residual energy:

$$F(f) = \frac{1}{\nu} |\hat{\mu}(f)|^2 \sum_{k=0}^{K-1} |V_k(0)|^2 \Big/ \frac{1}{2K-\nu} \mathrm{e}^2(\hat{\mu},f), \tag{16}$$

where ν is two degrees of freedom when the line frequency is known, and is three if frequency is estimated. The location of the maximum value of F provides an estimate of line frequency that has resolution within 5–10 % of the Cramér–Rao bound. This test works well if lines are isolated, i.e. if there is only a single line in $(f-W,f+W)$. The total number of lines in the spectrum is not important as long as they occur singly. For lines spaced closer than W one uses a multiple-line test, a similar but algebraically more complicated regression of the eigencoefficients on a matrix of Slepian functions $V_k(f-f_j)$ with the simple F-test replaced by partial F-tests (see, for example, Draper & Smith 1981). Multiple-line tests should be used with caution, because the Cramér–Rao bounds for line parameter estimation degrade rapidly, when the line spacing becomes less than $2/N$ (Rife & Boorstyn 1976). It is important to remember that in typical time-series problems hundreds or thousands of uncorrelated estimates are being dealt with; consequently one will encounter numerous instances of the F-test giving what would normally be considered highly significant test values that, in actuality, will only be sampling fluctuations. A good rule-of-thumb is not to get excited by significance levels less than $1-1/N$.

One aspect of the F-test of particular significance for the analysis of climate data is its use in estimating the mean value of the series. It is uncommon to have measurements free from an

average offset, so the usual practice is to subtract the average, and possibly slope, from the data, and then treat it as though it were a sample from a process with known zero mean. The problem is that this approach leaves a hole in the spectrum around zero frequency that, combined with the very red spectra typical of geophysical processes, often leads to the appearance of a 'peak' at a period commensurable with the length of the series. A better approach is to include a term in (2) at zero frequency, use (15) at $f = 0$ to estimate $\hat{\mu}(0)$, and use the residuals $\hat{x}_k(f) = x_k(f) - \hat{\mu}(0)\, V_k(f)$ in place of $x_k(f)$ at frequencies less than W. This is particularly crucial when the periods of interest are nearly as long as the data; here multiple-line estimates are necessary, as is the inclusion of both positive and negative frequency terms. Further information on F-tests is in Thomson *et al.* (1986) and Park *et al.* (1987*a*).

6. Various aspects of multiple-window estimates

A classical problem in spectrum estimation is the choice of the bandwidth, W. Although several methods for making such a choice exist, they typically depend on an unverifiable smoothness assumption. Here again, multiple-window methods have the advantage of allowing a test for unresolved structure. To do this, take the vector of eigencoefficients

$$\boldsymbol{X}(f) = [x_0(f), x_1(f), \ldots, x_{K-1}(f)]^{\mathrm{T}}$$

and compute its covariance matrix by averaging over different realizations or sections $\hat{\boldsymbol{C}}(f) = \mathrm{ave}\{\boldsymbol{X}\boldsymbol{X}^{\dagger}\}$. The expected elements of $\boldsymbol{C}$ are $C_{jk}(f) = \mathrm{E}\{x_j(f)\,\overline{x_k(f)}\}$ or, up to terms in $1-\lambda_j$, approximately

$$C_{jk}(f) \approx \int_{-W}^{W} V_j(\nu)\,\overline{V_k(\nu)}\, S(f-\nu)\,\mathrm{d}\nu.$$

If S is constant in $(f-W, f+W)$ then, by the orthogonality of the Slepian functions, $C_{j,k} \approx S(f)\,\lambda_j \delta_{j,k}$ or $\boldsymbol{C} \approx S(f)\,\boldsymbol{I}$. Consequently one may use a complex analogue of the sphericity (Anderson 1984) or uniformity (Watson 1983) tests on $\hat{\boldsymbol{C}}(f)$ for unresolved structure. Note that assuming S is constant across the local band is the same as assuming the spectrum is locally white, or that the bandwidth is such that details are resolved: practically, S is rarely constant, but if the structural components in a spectrum estimate are much smaller than the random components, the estimate is probably reasonable. Use of quadratic inverse theory gives a better, but more complicated, test for unresolved structure that does not require subdivision or replication of the data. Details will be given in a forthcoming paper.

Cross-spectra and coherences may also be computed without the usual phase-averaging problems. Letting $x_k(f)$ and $y_k(f)$ be the eigencoefficients from series x and y respectively, the cross-spectrum estimate corresponding to (11) is

$$S_{xy}(f) = \frac{1}{K} \sum_{k=0}^{K-1} x_k(f)\,\overline{y_k(f)}$$

from which coherences may be computed as usual. If several series are available, for example solar and climate series, linear transfer functions are easily computed; the paper by Lanzerotti *et al.* (1986) contains an example. Generally speaking, multiple-window methods allow the use of all the inferential techniques of multivariate statistics without the loss of resolution implicit in data sectioning methods but, of course, with the unavoidable complications of having almost everything a complex function of frequency.

Tolerances and confidence limits may be found for multiple-window estimates without

gaussian assumptions by jack-knife or bootstrap methods (Efron 1982) by treating the eigencoefficients (10) as data. Details are given in a forthcoming paper by Thomson & Chave (1990).

7. An example: thickness variations of growth-rings in bristlecone pine

This data-set consists of 5405 measurements of the thickness of the annual growth-rings of bristlecone pines from Mount Campito (east-central California) starting at 3434 B.C., generously provided by Professor C. P. Sonett and attributed by him to the late Professor V. C. La Marche of the University of Arizona's tree-ring laboratory. The data is of high quality with few outliers.

The time dependence of this data, however, exhibits unusual non-stationary effects. Because of the numerous claims for periodic effects in the climate literature, the harmonic F-test (16) was used to look for these using the complete data record. F-tests at 8.86 and 19.61 years were statistically insignificant. Periods of 5.429, 5.071, 3.276, 2.500, and 2.143 years had F-tests with significances between the 99.9% and 99.99% levels but, unless confirmed by independent data, are best considered sampling fluctuations. (Remember that this set contains 5400 samples, so one should expect about five spurious peaks above the 99.9% level.) At periods longer than 5.5 years nothing of significance above 99.9% was found. The general level of the F-test at long periods is suspiciously low. A possible, although one hopes improbable, explanation is that this band consists of a continuum of unresolved lines. Next, the series was broken into overlapping subseries of various lengths and the tests repeated. Numerous lines with high apparent significances were found in individual blocks but did not persist between blocks. This could result if the time base had several step discontinuities from errors in the dendrochronology, but the otherwise high quality of this data makes this explanation unlikely. Another mechanism to explain this behaviour is that line components do exist, but as unresolvable multiplets. If true, the isolated line F-test would fail, and its behaviour in short blocks would be apparently erratic. In blocks where the singlets are in phase, a signal at the average frequency would be detected; but in blocks where they are out of phase, the singlets would largely cancel. To check this possibility several approaches were tried: first, a robust prewhitening filter of the type described by Kleiner *et al.* (1979) was used to reduce the dynamic range of the spectrum and the influence of the few outliers. Then a single prolate data window with time-bandwidth product of 1.5 was used to compute the spectrum, part of which is shown in figure 1. The portion shown, from 0.011 to 0.015 cycles per year (or periods between 66 and 91 years) includes the Gleisberg cycle (Sonett 1982). The solid curve shows a roughly symmetric array of peaks, suggestive of a modulation process. To check this, the analysis was repeated on the data listed in table 3 of Ferguson (1970). This data is a longer bristlecone pine series, beginning in 5142 B.C.: because of numerous anticorrelated errors in adjacent years, the higher frequencies are unrealiable; low frequencies are similarly unrealiable because of the filtering. The corresponding estimate (scaled so the largest peak in this band has magnitude 1) is shown as the dashed line. Remembering that the Rayleigh resolution here is 2×10^{-4} cycles per year, only one major peak, at 0.013 cycles per year (77-year period), is in good agreement with the Campito set, despite the presumably high correlation between the two data-sets.

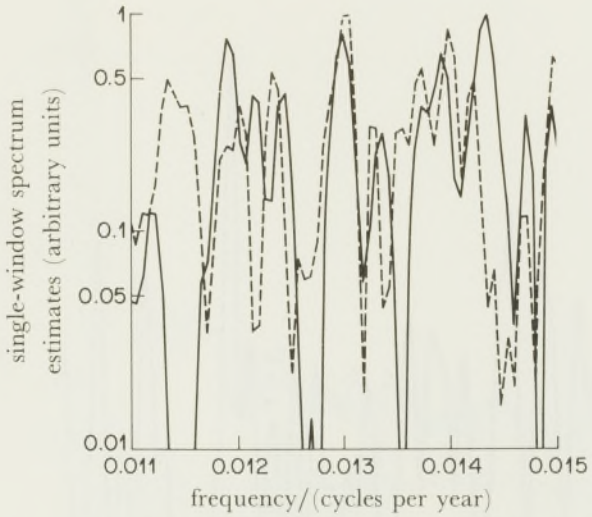

FIGURE 1. Single window (time-bandwidth product of 1.5) unsmoothed spectrum estimates of the Campito (solid line) and Ferguson (dashed line) bristlecone pine series. Note that the only peak with good agreement is the one at 0.013 cycles per year (77-year period). Both estimates have been scaled to have a maximum value of 1 in this band.

Secondly, a narrow-band eigenvalue procedure similar to that described in Thomson (1986) was tried. This, as Pisarenko-style estimators do, produced spectacularly narrow peaks at similar frequencies to those shown. These frequencies were then used in a multiple-line F-test, but the significance of the fit was low. Finally, deleting each line in turn and computing partial F statistics again resulted in mediocre fits. None the less, this band probably should be modelled as a dominant multiplet split into two or three singlets, perhaps by precession, and the whole modulated. The reason for this is that if one writes (9) in the time domain, to give the complex demodulate,

$$x_f(t) = \sum_{k=0}^{K-1} x_k(f)\, v_t^{(k)}(N, W),$$

the magnitudes $|x_f(j/2W)|$ taken at the filter Nyquist rate, have an empirical probability distribution that is a poor match to the expected Rayleigh distribution. Development of such a complicated model, however, is beyond the scope of this paper and the limitations of this data.

Given the failure of the line test procedures to show significant levels even at known astronomical periods that are suspected of being present in climate data, one must attribute most of the variations in the spectrum to the non-deterministic or continuum part of the process. Computing a multiple-window spectrum from the complete record results in an estimate that varies rapidly in frequency, reminiscent of a collection of medium-Q modes. However, when confidence limits are set by a jack-knifing procedure, where individual windows are delected from the estimate, an excessive variance is obtained. As before, the data were subdivided, spectra computed on each block, then tested for stationarity by using the procedure described on p. 1994 of Thomson (1977) with the results shown in figure 2. For this test five sections each of 1800 years duration and offset by 900 years were used. The spectrum for each section was computed by using $K = 5$ windows with a time-bandwidth product $NW = 4$ and the adaptive weighting (12). At each frequency, the stationarity test is a Bartlett M statistic (see Pearson &

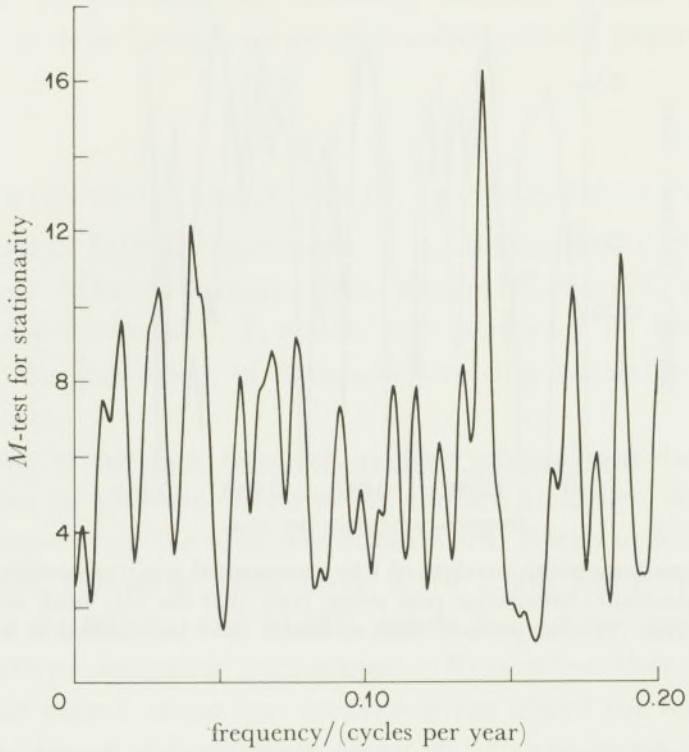

FIGURE 2. Stationarity test for the Campito data. The full data-set was broken into five overlapping segments each 1800 years long and offset from each other by 900 years, starting from an initial date of 3434 B.C. On each segment five tapers with a time-bandwidth of 4 were used, giving spectrum estimates each having 10 degrees of freedom. The expected value of Bartlett's M statistic for homogeneity of variance is 4.16 and the 90% point is at 8.1. The plotted data have been slightly smoothed.

Hartley 1969), so here would have a distribution which is approximately χ^2 with four degrees of freedom. As the expected value of M is 4.16 and the average, across frequency, is 5.2, the series is only mildly non-stationary.

Figure 3 shows a robust average, a trimmed logarithmic mean, of the five subset estimates. Conventional methods would assign factors of 0.8 and 1.2 for the lower and upper 5% points, but in view of the non-stationarity, factors of $\frac{1}{2}$ and 2 are more appropriate. Even so, there are many statistically significant features and, at a minimum, one can say that the probability of the spectrum being white is zero.

To explore the non-stationarity and associated excess variability further, spectra were computed on 16 sections each 900 years long and offset by 300 years by using 10 windows with a time-bandwidth product of 6 on each block. Again, to minimize effects of sampling noise, only frequencies below 0.2 cycles per year were kept. Logarithms of the different spectrum estimates were arranged in a rectangular matrix $\boldsymbol{G}$, the average over blocks subtracted at each frequency, and a singular value decomposition of the result computed. Taking $\delta\boldsymbol{G} = \boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}$ the left eigenvectors, the columns of $\boldsymbol{U}$, describe frequency dependence; the right eigenvectors, the columns of $\boldsymbol{V}$, the time dependence; and the singular values σ the scale of the effect. The first three eigenvectors are shown in figures 4 and 5. The frequency dependence (figure 4) is complicated: $\boldsymbol{U}_1$ shows a basic split in behaviour at 0.025 cycles per year; $\boldsymbol{U}_2$ also has this break but has related peaks near 0.044 cycles per year (22.7 years) and 0.011 cycles per year (9 years), possibly a modulation of low-frequency solar processes by the 8.86-year lunar cycle. $\boldsymbol{U}_3$ also has a 22-year peak and a related third harmonic at 0.14 cycles per year. The time

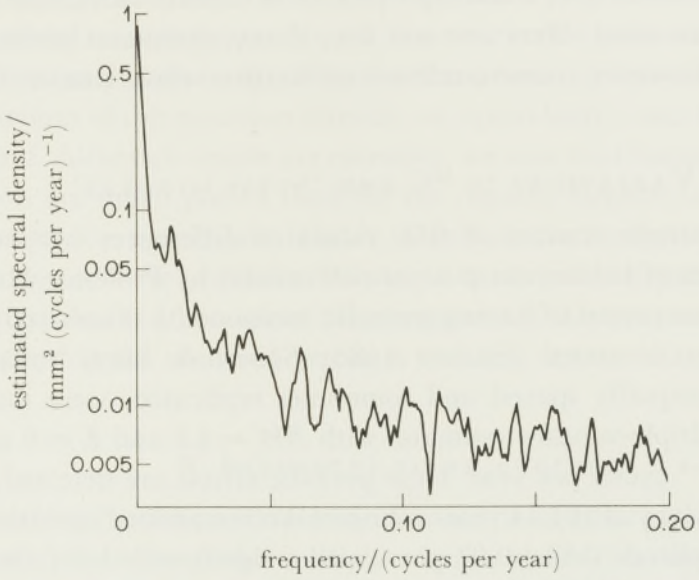

FIGURE 3. An estimate of the spectral density of the Campito data. The estimate is made by averaging the logarithms of the central three of the five sorted section estimates at each frequency. The data segmentation and windowing is the same as in figure 2. Because the sorting introduces discontinuities the estimate shown has been slightly smoothed.

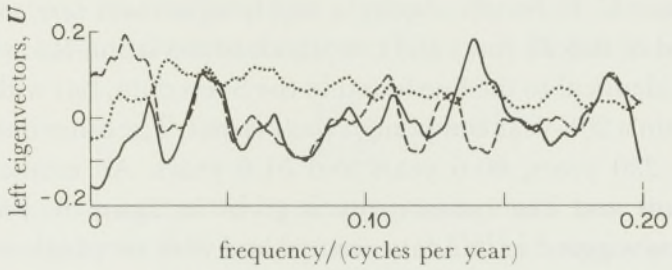

FIGURE 4. Frequency components, the left eigenvectors $\boldsymbol{U}$, produced by a singular value decomposition of a time-dependent spectrum estimate of the Campito data. The first eigenvector is shown dotted; the second, dashed; and the third, solid, with relative eigenvalues of 259.4, 95.3 and 71.9 respectively. The corresponding right eigenvectors are shown in figure 5.

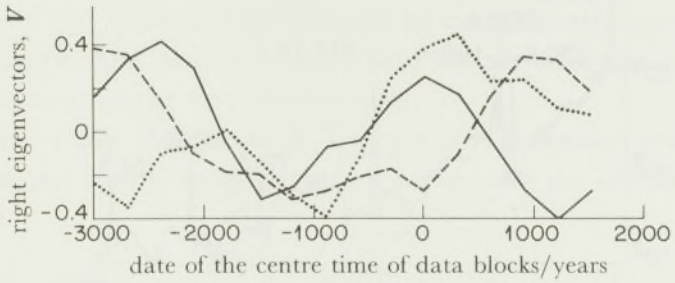

FIGURE 5. The time-dependent part, or right eigenvectors, of the SVD for the Campito data. The line styles are as in figure 4. Note the 2360-year period of the third eigenvector (solid line).

dependence of the second and third components is more interesting: $\boldsymbol{V}_2$ appears to vary with about a 4000-year period, whereas $\boldsymbol{V}_3$ has two cycles of a 2360-year period. Applying the harmonic F-test to $\boldsymbol{V}_3$ gives a significance level of 99.99%. Note carefully, however, that this effect is not a direct periodic component embedded in the signal but rather the structure of the process varies in a periodic manner.

Taken together, these results give a different picture of climate variability (as expressed by the tree-ring record) than usual. Here one sees few, if any, dominant periodic effects in the direct signal. There are, however, numerous low-level features whose statistical structure varies with time.

8. Variations in ^{14}C and 'Suess wiggles'

Our second data example consists of 644 values of differences between the ^{14}C and dendrochronological ages of bristlecone pine as determined by Professor Hans Suess. These records have long been suspected of having periodic components (Suess 1965; Sonett 1984), but the point has been contested (Stuiver 1980; Sonett & Suess 1984). The original measurements, being unequally spaced and sometimes replicated, were interpolated to an equally spaced grid. Multiple-window estimates with $NW = 4.5$ and $K = 6$ were used, giving a bandwidth of 5.4×10^{-4} cycles per year. Here periodic effects are detected. In particular a line with estimated frequency of 104.14 years (one jack-knife standard deviation limits of 104.0 to 104.28 years), an amplitude 0.84 ± 0.07 years, and a significance level above 99.99 % was found. Possibly related lines with periods of 208.2 years at 95 %, and 34.8 years at 99.6 % were also present.

The analysis was repeated, starting with 282 samples of differences between the ^{14}C and dendrochronology ages of samples from Irish oaks taken from G. W. Pearson's thesis, the data again courtesy of Professor C. P. Sonett. Again, a highly significant line, at 99.7 %, was found with an estimated period of 103.93 years and one standard deviation limits of 103.72 to 104.15 years, that is essentially identical to the line found in the Suess data, but with a lower amplitude of 0.48 ± 0.12 years. Again a 208-year component was present. The other lines common between the two data-sets were 230 years, 60.6 years and 51.6 years. An estimate of the spectrum showing both continuum and line components is given in figure 6. From this it may be concluded that the 'Suess wiggles' in ^{14}C data are real and with very high confidence, that they have a probable fundamental period of 208.15 years, a strong harmonic at 104.07 years, and probably additional frequencies.

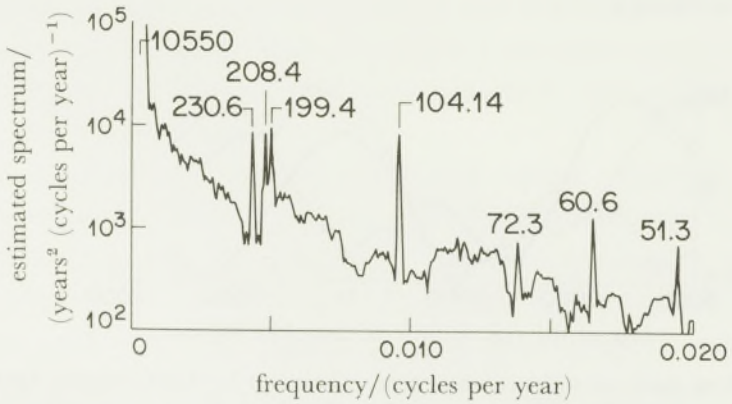

Figure 6. Reshaped multiple-window spectrum estimate of the Suess ^{14}C data. A time-bandwidth product of 4.5 with six windows was used with adaptive weighting resulting in an average stability of 11.6 degrees of freedom. At frequencies where the harmonic F-test exceeded the 95 % level the spectrum was 'reshaped' to show the line. The width of each line is proportional to its frequency uncertainty (from the F-test) and the height is scaled to conserve power. Significant peaks are labelled with their estimated period in years. The estimated low-frequency spectrum, clipped in the figure, was 2×10^6 but had unresolved components; the true value would be somewhat higher.

To conclude this example, note that the interpolation problem was largely ignored; this is not because it is trivial or unimportant, quite the reverse. To proceed rigorously, one develops an integral equation analogous to (6) using the time points (or intervals) observed, then either uses the eigenfunctions of this equation directly, or equivalently, uses them to interpolate to an equally spaced grid. Although details are excessive, we note that the eigenvalues of the random sampling equation are much poorer than for the regular Slepian functions, so that random aliases of high-frequency signals are difficult to exclude on analytic grounds. Bronez (1988) gives an introduction to these problems. Because a variety of interpolation schemes were tried here with similar results, and because of the agreement of the two sets with different sampling times, we conclude that the observed period is real and not an alias.

9. Estimates to be avoided

Having dwelt at length on the multiple-window methods, what of the myraid other spectrum estimation methods? Recall that the multiple-window method occurs as a natural solution of the inverse problem; the better of the single-window estimates, i.e. a Parzen or Papoulis window applied directly to the data, not to the correlations, give the first term in the series. (See Harris (1978) for properties of these, and many other, windows.) They are consequently less efficient and are unable to separate moments. Also, being *ad hoc*, the theory for such windows is much less satisfactory, not to mention algebraically more complicated, than multiple windows. The oldest spectrum estimate (from 1894) still in common use, the unwindowed periodogram,

$$P(f) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x_n \, \mathrm{e}^{\mathrm{i}2\pi nf}\right|^2, \tag{17}$$

is obsolete, but, because it is computed by simply passing one's data through a discrete Fourier transform (DFT) (usually as an FFT) and squaring, it has had an unfortunate resurgence. It is often disguised, for example: the usual frequency spacing of $\Delta f = 1/N$ gives exactly the Fourier series representation of the periodically extended observations; a least-squares fit to the data by a trigonometric series is again the same thing; and there are innumerable similar variants. As an estimate of the spectrum it is both inconsistent and so badly biased that it is extremely misleading and should not be used. This bias can be exceptionally severe; in Thomson (1977) an example using engineering data is shown where the periodogram is in error by more than a factor of 10^{10} over most of the frequency range.

Indirect or *Blackman–Tukey* (Blackman & Tukey 1958) estimates were developed before the invention of the FFT and traditionally began by computing the sample autocovariances,

$$\hat{R}(\tau) = \frac{1}{N}\sum_{n=0}^{N-1-|\tau|} x_n \, x_{n+|\tau|},$$

as a function of the lag, τ, for lags up to perhaps 40 % of the series length. The spectrum is then estimated by multiplying $\hat{R}(\tau)$ by one of a multitude of lag windows and Fourier transforming. Because the sample autocorrelations may be obtained by Fourier transforming the periodigram,

$$\hat{R}(\tau) = \int_{-\frac{1}{2}}^{\frac{1}{2}} P(f) \, \mathrm{e}^{\mathrm{i}2\pi ft} \, \mathrm{d}f \tag{18}$$

(and are now usually computed with a second FFT and zero padding), the convolution theorem shows that indirect estimates are simply linear smoothers applied to the periodigram. Thus apart from potentially serious round-off problems, indirect estimates give one a stable and incorrect estimate instead of the unstable and incorrect estimates obtained with the periodigram; again they are obsolete and should not be used.

Autoregressive or 'maximum-entropy' estimates are more insidious; misappropriating authority from statistical mechanics and Shannon theory, such approaches choose the estimate maximizing the Wiener entropy,

$$H\{\hat{S}\} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \hat{S}(f)\, \mathrm{d}f,$$

subject to the constraints that it reproduces the first few autocorrelations. Because the autocorrelations are given by (18), the 'maximum-entropy' estimate can be no better than the periodogram. Thus in addition to fitting a biased and inconsistent estimate, one imposes what may be totally inappropriate model whose forced fitting can introduce spurious features in the estimated spectrum. The Burg procedure (see Chen & Stegun 1974) avoids use of the autocorrelations but imposes both strong stationarity and time-reversal symmetry constraints. Bretthorst (1987) fits an all-line model by using ordinary least squares and a geophysically unrealistic white noise assumption. For dangers of model misspecification, see Tukey (1982); also Kay & Marple (1982) for some examples.

One positive aspect of Wiener entropy that should not be overlooked is that $\exp\{H\}$ is the variance of the best linear predictor of a stationary process; note that it is a logarithmic mean of the spectrum. Thus a small change in total power can drastically alter predictability so spectra should amost invariably be plotted with a logarithmic ordinate.

Finally, the method of 'linear spectral analysis' recently used in the geological literature, is badly biased, erratic and untrustworthy except when one knows *a priori* that there is a single periodic signal with high signal:noise present (see Stigler & Wagner 1987).

To conclude this section, note that it can be argued that careful prewhitening will reduce the bias problem (and although strongly advocated by Blackman & Tukey, the advice is apparently usually ignored); this is usually the case, but there are examples where it is not possible. Moreover, the same argument can be applied, and with greater success, to multiple-window methods. Even ignoring the overwhelming advantages of the multiple-window methods with mixed spectra, careful analysis of bias and variance invariably show multiple-window methods to have the advantage.

10. Conclusions

The major point of this paper is that, for the analysis of Holocene climate data, the multiple-data-window method of spectrum estimation is better than the various *ad hoc* methods of time-series analysis. With such methods I have shown that climate variations, as recorded through growth rings of bristlecone pines, have few, if any, direct periodic components but do show evidence for changes in structure with a period about 2360 years. I have also shown that the 'Suess wiggles' in ^{14}C variations have a component with a period of 208.2 years, plus harmonics, with very high certainty.

References

Anderson, T. W. 1984 *An introduction to multivariate statistical analysis*, 2nd edn. New York: Wiley.

Blackman, R. B. & Tukey, J. W. 1958 *The measurement of power spectra*. New York: Dover.

Bretthorst, G. L. 1987 Bayesian spectrum analysis and parameter estimation. Ph.D. thesis, Department of Physics, Washington University, St Louis, Missouri, U.S.A.

Brillinger, D. R. 1975 *Time series, data analysis and theory*. New York: Holt, Rinehart and Winston.

Bronez, T. P. 1988 Spectral estimation of irregularly sampled multidimensional proceses by generalized prolate spheroidal sequences. *IEEE Trans. Acoust. Speech, signal Processing* **36**, 1862–1873.

Chen, W. Y. & Stegun, G. R. 1974 Experiments with maximum entropy power spectra of sinusoids. *J. geophys. Res.* **79**, 3019–3022.

Daubechies, I. 1988 Time–frequency localization operators: a geometric phase space approach. *IEEE Trans. Inf. Theory* **IT-34**, 605–612.

De Mendoça Dias, A. A. 1962 The volcano of Capelinhos (Azores), the solar activity, and the earth-tide. *Bull. volcan.* **24**, 211–221.

Doob, J. L. 1953 *Stochastic processes*. New York: Wiley.

Draper, N. R. & Smith, H. 1981 *Applied regression analysis*. New York: Wiley.

Efron, B. 1982 *The jackknife, the bootstrap, and other resampling plans*. Philadelphia: SIAM.

Ferguson, C. W. 1970 Dendrochronology of bristlecone pine, *Pinus aristata*. Establishment of a 7484-year chronology in the White Mountains of eastern-central California, U.S.A. In *Radiocarbon Variations and Absolute Chronology, Proc. 12th Nobel Symp.* (ed. I. U. Olsson), pp. 237–259. New York: Wiley; Stockholm: Almqvist & Wiksell.

Harris, F. J. 1978 On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE* **66**, 51–83.

Kay, S. M. & Marple, S. L. Jr 1982 Spectrum analysis – a modern perspective. *Proc. IEEE* **69**, 1380–1419.

Kleiner, B., Martin, R. D. & Thomson, D. J. 1979 Robust estimation of power spectra (with discussion). *Jl R. statist. Soc.* B **41**, 313–351.

Lanzerotti, L. J., Thomson, D. J., Meloni, A., Medford, L. V. & Maclennan, C. G. 1986 Electromagnetic study of the Atlantic Continental Margin using a section of a transatlantic cable. *J. geophys. Res.* B **91**, 7417–7427.

Munk, W. & Hasselmann, K. 1964 Super-resolution of tides. In *Studies on oceanography* (ed. K. Yoshida), pp. 339–344. Tokyo. (Reprinted 1965 University of Washington Press.)

Papoulis, A. 1973 Minimum bias windows for high-resolution spectral estimates. *IEEE Trans. Inf. Theory* **IT-19**, 9–12.

Park, J., Lindberg, C. R. & Thomson, D. J. 1987*a* Multiple-taper spectral analysis of terrestrial free oscillations: part I. *Geophys. Jl R. astr. Soc.* **91**, 755–794.

Park, J., Lindberg, C. R. & Vernon, F. L. III 1987*b* Multitaper spectral analysis of high-frequency seismograms. *J. geophys. Res.* **92**, 12675–12684.

Parzen, E. 1957 On consistent estimates of the spectrum of a stationary time series. *Ann. math. Statist.* **28**, 329–348.

Pearson, E. S. & Hartley, H. O. 1969 *Biometrika tables for statisticians*, 3rd edn. Cambridge University Press.

Pittock, A. B. 1978 A critical look at long-term Sun–weather relationships. *Rev. Geophys. Space Phys.* **16**, 400–420.

Priestley, M. B. 1981 *Spectral analysis and time series*. London: Academic Press.

Priestley, M. B. 1988 *Non-linear and non-stationary time series analysis*. London: Academic Press.

Rife, D. C. & Boorstyn, R. R. 1976 Multiple tone parameter estimation from discrete-time observations. *Bell System Tech. J.* **55**, 1389–1410.

Shapiro, R. 1979 An examination of certain proposed Sun–weather connections. *J. atmos. Sci.* **36**, 1105–1116.

Slepian, D. 1978 Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V: the discrete case. *Bell System Tech. J.* **57**, 1371–1430.

Sonett, C. P. 1982 Sunspot time series: spectrum from square law modulation of the Hale cycle. *Geophys. Res. Lett.* **9**, 1313–1316.

Sonett, C. P. 1984 Very long solar periods and the radiocarbon record. *Rev. Geophys. Space Sci.* **22**, 239–254.

Sonett, C. P. & Suess, H. E. 1984 Correlation of bristlecone pine ring widths with atmospheric ^{14}C variations: a climate–Sun relation. *Nature, Lond.* **307**, 141–143.

Stuiver, M. 1980 Solar variability and climatic change during the current millenium. *Nature, Lond.* **286**, 868–871.

Stigler, S. M. & Wagner, M. J. 1987 A substantial bias in nonparametric tests for periodicity in geophysical data. *Science, Wash.* **238**, 940–944.

Suess, H. E. 1965 Secular variations of the cosmic ray produced carbon-14 in the atmosphere. *J. Geophys. Res.* **70**, 5937–5952.

Thomson, D. J. 1977 Spectrum estimation techniques for characterization and development of WT4 waveguide. *Bell System Tech. J.* **56**: part I, 1769–1815; part II, 1983–2005.

Thomson, D. J. 1982 Spectrum estimation and harmonic analysis. *Proc. IEEE* **70**, 1055–1096.

Thomson, D. J. 1986 Combined multi-window eigenvalue spectrum estimates. In *Proc. Third IEEE ASSP workshop on spectrum estimation and modelling*, pp. 103–105.

Thomson, D. J. & Chave, A. D. 1990 Jackknifed error estimates for spectra coherences, and transfer functions. In *Advances in spectrum estimation and array processing* (ed. S. Haykin), ch. 2. Englewood Cliffs, New Jersey: Prentice-Hall. (In the press.)

Thomson, D. J., Lanzerotti, L. J., Medford, L. V., Maclennan, C. G., Meloni, A. & Gregori, G. P. 1986 Study of tidal periodicities using a transatlantic telecommunications cable. *Geophys. Res. Lett.* **13**, 525–528.

Tukey, J. W. 1982 Styles of spectrum analysis. In *The collected works of John W. Tukey*, vol. II (ed. D. R. Brillinger), pp. 1143–1153. California: Wadsworth.

Watson, G. S. 1983 *Statistics on spheres*. New York: Wiley–Interscience.