Robust Estimation of Background Noise and Signal Detection in Climatic Time Series

Michael E. Mann & Jonathan M. Lees
Dept. of Geology & Geophysics
P.O. Box 208109
Yale University
New Haven, CT 06520-8109

February 7, 1996

Abstract

We present a new technique for isolating climate signals in time series with a characteristic "red" noise background which arises from temporal persistence. This background is estimated by a "robust" procedure that, unlike conventional techniques, is largely unbiased by the presence of signals immersed in the noise. Making use of multiple-taper spectral analysis methods, the technique further provides for a distinction between purely harmonic (periodic) signals, and broader-band ("quasiperiodic") signals. The effectiveness of our signal detection procedure is demonstrated with synthetic examples that simulate a variety of possible periodic and quasiperiodic signals immersed in red noise. We apply our methodology to historical climate and paleoclimate time series examples. Analysis of a ≈ 3 million year sediment core reveals significant periodic components at known astronomical forcing periodicities and a significant quasiperiodic 100 year peak. Analysis of a roughly 1500 year tree-ring reconstruction of Scandinavian summer temperatures suggests significant quasiperiodic signals on a near-century timescale, an interdecadal 16-18 year timescale, within the interannual El Nino/Southern Oscillation (ENSO) band, and on a quasibiennial timescale. Analysis of the 144 year record of Great Salt Lake monthly volume change reveals a significant broad band of significant interdecadal variability, ENSO-timescale peaks, an annual cycle and its harmonics. Focusing in detail on the historical esimated global-average surface temperature record, we find a highly significant secular trend relative to the estimated red noise background, and weakly significant quasiperiodic signals within the ENSO band. Decadal and quasibiennial signals are marginally significant in this series.

1 Introduction

As we seek to establish more confidently the extent to which anthropogenic influences may be affecting the climate, it becomes increasingly essential that we be able to distinguish “signals” in the climate record from the “noise” background in which they are immersed. Theoretical investigations suggest, for example, that signals on interdecadal and longer timescales could arise from the coupling of ocean-atmosphere processes and the dynamics of ocean gyres (e.g., Latif and Barnett, 1994) or the ocean’s thermohaline circulation (see e.g., Stocker and Mysak 1992, Delworth et al. 1994) . As such natural signals may partially mask an underlying anthropogenic trend (see e.g., Schlesinger and Ramankutty, 1994a; Santer et al, 1995) isolating them may allow for a more confident identification of anthropogenic climate change.

The level of confidence we attribute to a potential signal depends quite sensitively on our *a priori* assumptions regarding the nature of the background noise, as well as the appropriateness of our statistical model for the signal itself. Most paradigms associate climate signals with a direct response to periodic external forcing such as that by the yearly cycle in insolation, with unstable internal oscillations in the climate system (e.g., the El Nino/Southern Oscillation or “ENSO”) or, possibly, an interaction between such external and internal mechanisms. Signals of the first type can appropriately be modeled by a sinusoidal time-dependence, and are readily identified as narrow peaks in the power spectrum of the series with a coherent phase spectrum. In the latter two cases, however, such a simple time-domain model is not sufficient because the signal does not exhibit long-term memory of its phase. Nonetheless, such signals typically acquire a dominant timescale from the underlying dynamics (perhaps in resonance with external forcing) and can be detected as relatively narrow peaks in the power spectrum with greater amplitude than would be expected from chance noise fluctuations. The success of signal detection will thus hinge, in such cases, on the proper isolation of the underlying noise and an appropriate criterion for analyzing the power spectrum for signal contributions. In this paper, we describe a procedure that seeks to provide a more accurate means for signal detection and noise background estimation in climate studies.

In section 1 we will motivate and describe the "red noise" model for the underlying noise component of the spectrum. In section 2, we introduce a signal detection procedure that (a) seeks to optimize the fundamental signal-to-noise ratio problem through the use of multiple-taper spectral analysis, (b) invokes assumptions regarding signal (narrow-band but not necessarily periodic) and background noise ("red") that are faithful to our understanding of the physics governing the climate system, and that (c) enhances the detectability of signals by using a "robust" procedure to estimate the noise background. In section 3, we apply our analysis to synthetic examples to demonstrate how our procedure provides improved signal detection and noise estimation in the case where signal and noise are precisely specified and known. In section 4, we apply our analysis to a small number of historical and paleoclimatic time series, discussing implications for the existence of low frequency climate signals. Finally, in section 5, we apply our analysis to the instrumental record of estimated global-average surface temperature, focusing on the presence of low frequency signals, and providing a new perspective on the greenhouse signal detection problem.

2 The "Red Noise" model

"Red noise" is used to describe noise with relatively enhanced low-frequency fluctuations arising from the interaction of white noise forcing with the slow-response components of a system. There is theoretical justification for a red noise description of the noise background in climatic time series, as the thermal inertia of the oceans has been shown in models (e.g., Hasselmann, 1976; Wigley and Raper, 1990) to provide memory, effectively integrating atmospheric "weather" forcing. There is also considerable empirical evidence that the red noise model provides a reasonable description of the noise spectra for a variety of climatic and hydrological time series including long-term climate proxy records (Kutzbach and Bryson, 1974), historical sea and air surface temperature (Allen, 1992; Allen and Smith, 1994) and station precipitation data (Gilman et al., 1963).

The simplest statistical model for a discrete finite red noise series is the first-order autoregressive "AR(1)" process

$$r_n = \rho r_{n-1} + w_n \tag{1}$$

where $n = 1, \ldots N$ denotes the discrete time increment in units of the sampling interval Δt, $0 \leq \rho < 1$, the lag-one autocorrelation coefficient, describes the degree of serial correlation in the noise, and w_n is Gaussian white-noise sequence with variance σ^2. Note that the trivial limit $\rho = 0$ yields a white-noise process. For the AR(1) red noise process, autocorrelation decays exponentially as a function of time

$$\rho_n = \exp\left(-n\Delta t/\tau\right) \tag{2}$$

with a characteristic decay time of τ. Thus, if the lag-one autocorrelation coefficient $\rho \equiv \rho_1$ is known, the characteristic noise decay time scale can be determined as

$$\tau = -\frac{\Delta t}{\log \rho} \tag{3}$$

For periodicities much larger than τ, the spectrum behavies like a white spectrum.

The power spectrum of the AR(1) process is given by (e.g., Bartlett, 1966),

$$S(f) = S_0 \frac{1-\rho^2}{1 - 2\rho \cos \pi \frac{f}{f_N} + \rho^2} \tag{4}$$

for frequency f where S_0 is the average value of the power spectrum, related to the white-noise variance by,

$$S_0 = \sigma^2 / \left(1 - \rho^2\right) \tag{5}$$

(within some overall normalization of the power spectrum that depends on convention). $f_N = 2/\Delta t$, the Nyquist frequency, is the highest frequency that can be resolved for sampling rate Δt. More elaborate models might seek to describe the noise resulting from two or more such processes, each with different levels of autocorrelation (ie, a "compound" red-noise process—see Julian, 1961) or to determine the noise spectrum based on a more detailed model of the physics governing the climate system (e.g., Wigley and Raper, 1990–note that AR(1) red noise spectrum (4) is a limiting case of their more general noise model). An application of our methodology to these generalization of the AR(1) noise model adopted here would represent a worthwhile extension of the present study.

Other alternative models such as higher order autoregressive (AR(n), $n > 1$) models can provide a better fit to the overall spectrum, but often through parameterizing features that might actually be associated with signals. Such models can, for example, admit oscillatory solutions (see Allen 1992 for a more detailed discussion of this point) corresponding to narrow spectral peaks at non-zero frequency. Thus, the signal/noise distinction motivated earlier becomes blurred if a higher order AR noise model is invoked. "Long-range dependence" models for the underlying spectra of hydro-climatic time series exhibiting a $1/f$ scaling have also been argued (e.g. Mandelbrot and Wallis, 1969; Bloomfield, 1992). Others have directly argued against the AR(1) noise model based on conventional spectral analyses of certain climatic time series (e.g., Currie and Fairbridge, 1985; Newell et al, 1989). We will show below, however, that a *robustly* estimated AR(1) red noise background approximates quite well the the noise background spectrum in a variety of climate series. The AR(1) red noise is thus best justified because it has both theoretical and empirical motivation in the context of climate studies, and can easily be generalized to accommodate more detailed physics of the climate system.

3 Signal detection and "Robust" noise estimation

We assume an underlying time series model,

$$y_n = \sum_{l=1}^{L} {s_n}^{(l)} + r_n \tag{6}$$

where r is the red noise process described above, and $s^{(l)}$ denotes a finite, *a priori* unknown set of signals. Under the assumption that each signal $s^{(l)}$ has a distinct narrow-band signature in the frequency domain, spectral analysis affords a particularly useful means for detection. The efficiency of signal detection, however, depends on the resolution and statistical variance properties of the spectral estimate (see Thomson, 1982). Numerous spectral analysis studies of climatic time series (e.g., Gilman et al., 1963; Kutzbach and Bryson, 1974; Stocker and Mysak, 1992; Deser and Blackmon, 1993; Delworth et al. 1993) have taken into account the red noise background, but using a single-tapered periodogram. With only a single data taper, a spectral estimate with usefully low-enough statistical variance can only be obtained through a periodogram smoothing or related ensemble-averaging (e.g., Brillinger, 1981). Such a procedure bears a heavy cost in terms of the variance/resolution tradeoff and invokes fairly strict assumptions of stationarity. The multiple-taper or "multi-taper" method of spectral analysis (henceforth, "MTM"–Thomson, 1982; Percival and Walden, 1993) affords a modestly better tradeoff between spectral resolution and statistical variance than conventional single-taper methods (see e.g., Park et al, 1987)–an advantage which may be essential in the case where signal-to-noise ratios are low. Furthermore, MTM relaxes some of the strict stationarity assumptions, and the multiple degrees of freedom used to describe the spectrum provide for a local test for the presence of a sinusoidal signal against the assumption of a smoothly varying "locally-white" spectrum background. For these reasons, MTM is particularly well-suited for the spectral analysis of climatic time series (e.g. Thomson 1990ab, Kuo et al. 1990, Ghil and Vautard, 1991; Birchfield and Ghil 1993; Park and Maasch 1993; Mann and Park, 1993). However, to use MTM to test for the presence of both periodic and quasiperiodic signals against a red-noise background null hypothesis, several adaptations of the traditional MTM procedure are required.

We seek first to isolate any periodic signals corresponding to singular peaks in the power spectrum (within the confines of finite frequency resolution) associated with a coherent phase spectrum. This is accomplished by Thomson's (1982) reshaping procedure with some small modifications. The residual "continuous" spectrum component may still contain narrowband features associated with quasiperiodic signals. To identify these features, we separate out the estimated noise background from the residual spectrum.

This noise component must be estimated in such a way that the noise parameter estimation is not distorted by the potentially sizeable signals imbedded in the noise. For example, when a significant secular trend is present, the series is often "detrended" through subtraction of the assumed low-frequency contribution before noise parameters are fitted (e.g., Stocker and Mysak, 1992; Elsner and Tsonis, 1994). Similarly, an iterative procedure may be used to remove signal contributions from the series as they are detected (see e.g., Allen and Smith, 1994), eventually resulting in a residual series which is consistent with the underlying noise model. Because such a procedure is (a) not especially efficient in an exploratory data analysis context and (b) can lead to undesirable effects on the spectrum of the residual (see section 4), we prefer our own alternative "robust" noise background estimation procedure. This procedure makes use of a median smoothing of the spectrum to provide an estimate of the underlying noise background that is insensitive to the contributions of narrowband signals but does not alter the spectrum itself. The median-smoothed empirical estimate of the background will, however, rarely yield the true underlying noise spectrum, due to sampling fluctuations and some small residual bias that even the median smoothing cannot eliminate. This is seen below in our synthetic examples (section 3) where the noise is generated from a pure red noise process.

Assuming that the underlying noise was generated by an AR(1) red noise process, we can better isolate the true noise background by fitting an analytical red noise spectrum (4) to the median-smoothed background estimate. This involves fitting two free parameters, the average noise level S_0 (or equivalently, through (5), the average white-noise level σ) and the lag-one autocorrelation coefficient ρ, to the median smoothed background. This "robust" estimate of ρ will provide an improved estimate of the true persistence timescale τ of the noise through (3) since signal contributions present in the nominal lag-one correlation coefficient of the time series are largely eliminated. This value of τ may be compared with characteristic relaxation timescales of potential slow-response components of the climatic system for possible physical insight into the noise process.

Finally, we gauge the significance of periodic or quasiperiodic peaks in the spectrum relative to the estimated red noise background using elementary sampling theory (Tukey, 1950; Percival and Walden, 1993). Our entire procedure is illustrated schematically in Figure 1. Details of the spectrum estimation, reshaping, median-smoothing, red-noise background fit, and confidence level determination are provided in Appendix A.

4 Synthetic examples

We focus on two synthetic examples consisting of periodic and quasiperiodic signals superimposed on equal amplitude, moderately-red noise (Table 1). A comparison of the true, “raw” (ie, non-robustly estimated) and robust analyses of the noise background (Table 2) demonstrates the effectiveness of the robust red noise level determination.

The first synthetic time series (Figure 2a) describes a 120 year monthly sampled series (1440 months) with a periodic annual square-wave cycle, quasiperiodic interannual and decadal signals, a harmonic interdecadal 25 year oscillation, and a linear trend. We separately analyze the signal (Figure 2a-top) and signal-plus-noise (Figure 2a-bottom) time series. The “adaptive” (see Appendix A) MTM spectrum of the signal time series (Figure 2b) shows distinct peaks corresponding to each of the signals described above, including odd harmonics generated by the square wave. Leakage from the lower frequency signals produces a notably higher background floor to the spectrum at lower frequencies, but even at low frequency, the signals are at least two orders of magnitude above the floor–this leakage becomes negligible in the presence of added noise.

The signals are visually evident in the spectrum of the signal-plus-noise time series (Figure 2c), but the AR(1) background and confidence limits are artificially raised by the low frequency signals. This artifact leads to the clearly incorrect inferences that) there are no statistically-significant signals in the low frequency range (e.g., $f < 0.5$ cycle/yr) and that most of the high frequency ($f > 2$ cycle/yr) part of the spectrum is significant relative to red-noise at the $> 99\%$ level.

For the robust analysis, we choose a median smoothing window of width $\Delta f_{smooth} =$

1.2 cycle/yr, consistent with our guidelines for window width selection (Appendix A). We require a 99% significance level for phase-coherence and power in the reshaping procedure (see Figure 1; Appendix A). This analysis yields a far more accurate determination of the true red noise (Figure 2d; Table 2). The largest fractional discrepancy is found for τ (estimated at 0.11 yr rather than the true value of 0.21 yr) which is highly sensitive to the estimated value of ρ through its logarithmic dependence. The synthetic signals are each properly identified at the 99% confidence level. A periodic annual peak is detected along with quasiperiodic odd harmonics (in the latter case, the power level criterion is not met in the reshaping procedure). The interdecadal (25 year) sinusoid is detected as significant, but as a quasiperiodic rather than periodic signal. In this case, there too few repetitions (less than 5) of the cycle in the series for good enough statistical constraint in the phase-coherence test of the reshaping procedure. This failure underscores the difficulty in distinguishing *periodic* and quasiperiodic signals at very low frequencies. The relatively small number of spurious peaks above the 99% level are consistent with the expectations of chance occurance in a random noise sampling.

Focusing on the low-frequency range of 0-0.5 cycle/yr (Figure 3) corresponding to periods ($1/f$) greater than 2 years, we see in more detail how the raw AR(1) determination and associated confidence levels overpredict the low-frequency noise level (Figure 3a). The true noise spectrum is essentially white on interannual timescales owing to the fact that $1/f >> \tau$, but the non-robust analysis prescribes a strongly red background as an artifact of the influence of the secular, interdecadal, and quasidecadal signals. None of these signals breach the 99% confidence level relative this poorly estimated noise background (the secular trend doesn't even breach the estimated mean red-noise level). The robust analysis, in contrast, accurately determines the noise background on interannual timescales, and correctly isolates the low-frequency signals (Figure 3b).

The second synthetic example is a 120 year length annually sampled series (Figure 4a) consisting of a similar set of signals (Table 1) superimposed on moderate *interannual* timescale red noise ($\rho = 0.2$, $\tau = 0.6$ year). This example is a model for the relatively short, coarsely sampled time series that are often encountered in historical climate data. Because the frequency width (0.5 cycle/yr) of the spectrum is smaller, and spectral resolution is poor, due to the short nature of the time series, there is less independent information available for noise background estimation. However, the robust noise level determination represents a clear improvement over a raw determination of the noise level and confidence limits (Table 2). The adaptive MTM spectrum of the signal time series shows the expected peaks centered at 4-5 yr, 9-11 year, and on secular (>50 year) time scales (Figure 4b). In the raw AR(1) determination of the signal-plus-noise series, the highly overestimated values of ρ and S_0 (Table 2) lead to poor signal detection. Only the quasidecadal signal is found to exceed the 99% confidence level, with the 4-5 year signal not even significant at the 90% level (Figure 4c). The robust determination provides fairly accurate estimates of the true red noise (Figure 4d, Table 2). Each of the three signals is found to be significant at $> 95\%$ confidence level with the expected one chance spurious peak (near 2.2-2.4 year period) exceeding the 90% confidence level.

5 Application to climate time series

We examine three diverse climatic time series including (1) an ≈ 3 million year (Ma) long Ocean Drilling Project (ODP) sediment core reconstruction with 1 kyr (1000 year) sampling, (2) a ≈ 1500 year long tree ring temperature reconstruction with annual sampling, and (3) the ≈ 150 year long record of monthly volume change record of the Great Salt Lake (GSL). The results of the red noise parameter estimation procedure are summarized in Table 3, and the detected signals are discussed below.

5.1 ODP site 607 core

The 2.7 Ma ODP site 607 δ^{18}O reconstruction (Figure 5a) is a proxy for low-frequency Pleistocene ice-related climate fluctuations (see e.g., Birchfield and Ghil 1993, Park and Maasch 1993, Lees and Park 1994). The adaptive MTM spectrum (Figure 5b) exhibits a secular peak, in addition to peaks near 100 kyr, 40 kyr, and 18-25 kyr. The non-robust analysis however finds the secular peak insignificant, and unrealistically determines much of the higher frequency continuum (f>0.04 cyc/kyr) as significant above the 95% confidence level. The values of ρ and S_0 (Table 3) are overestimated by the several narrowband, low frequency peaks which artificially elevate the estimated noise level.

The robust determination provides a far more reasonable fit to the noise background, with a lesser degree of autocorrelation and a lower average level (Table 3) although the noise background varies somewhat more abruptly than can be accommodated by the simple AR(1) red noise description–the compound red noise process described earlier may be more appropriate in this case where the dynamic range is quite large. Consequently we show results based both on fitting $S(f)$ and $\log S(f)$ in the red noise estimation procedure (Figure 5c) which differ somewhat in how they choose to fit an AR(1) dependence to the noise background (see Appendix A). However, significance inferences at the 99% level are independent of the two choices, and the estimated noise parameters are similar (Table 3). Several high-frequency peaks isolated as significant in the raw red-noise significance level determination, are dismissed as insignificant in the robust determination. Periodic signals near 19, 23 and 24 kyr, likely due to Milankovitch precessional forcing, and 40 kyr, corresponding to obliquity forcing are, detected at the 99% confidence level for significance. The 40 kyr peak has a strong quasiperiodic component which may be associated with problems in the way the series has been "tuned". The 100 kyr peak is determined as strictly quasiperiodic, consistent with the intermittent nature of the oscillation evident in the time series (Figure 5a); the 100 kyr variability displays greatest strength during the recent ≈800 kyr, but is absent before that point, demonstrating an apparent transition in the climate system consistent with non-linear internal variability (see e.g. Park and Maasch 1993). Such non-stationary signal behavior suggests the usefulness of an "evolutive" generalization (e.g. Birchfield and Ghil 1993; Yiou et al. 1991) of the present procedure in which the spectrum is evaluated in a moving window through the time series. The secular peak, determined as highly significant in the robust analysis, can be attributed to long-term tectonic forcing (e.g., Mitchell, 1976). The robust determination gives a timescale of τ ≈10-20 kyr (depending on which method of noise background fitting is used), consistent with estimated response timescales of stochastic low-frequency ocean circulation variability (e.g. Keigwin et al., 1994).

5.2 Fennoscandinavian tree ring reconstruction

A 1482 year long tree ring reconstruction of Scandinavian summer temperatures (Briffa and Jones 1992; Bradley and Jones 1993) is analyzed here. Significant interdecadal-to-century scale variability is visually evident in the series (Figure 6a). The strong, relatively narrowband peaks associated with signals on these time scales (Figure 6b) artificially elevates the raw estimated red noise background. The associated confidence levels imply that the secular trend is the only significant low-frequency peak. Evaluated robustly (Figure 6c) there is evidence for a significant secular trend and significant quasiperiodic variability on century timescales (50-70 year timescale peak), 30-36, 16-

18, 4.8-5.0, 3.6-3.8, 3-3.2, and 2.2 year time scales. The interannual signals correspond to well-known ENSO and quasibiennial timescales (see e.g., Ghil and Vautard, 1991; Mann and Park, 1994). The quasibiennial peak is close to the Nyquist frequency and should be interpreted cautiously. The interdecadal signal has also been isolated in instrumental temperature records (Ghil and Vautard, 1991; Mann and Park, 1994), long-term proxy (Mann et al, 1995*b*) and historical (Plaut et al, 1995) data. Century scale oscillations in the North Atlantic have been noted elsewhere (e.g., Stocker and Mysak, 1992; Mann et al, 1995*b*) as well as predicted theoretically (e.g., Stocker and Mysak, 1992; Delworth et al. 1994). The latter studies have associated this variability with instabilities in the thermohaline circulation. The non-robust red-noise analysis (Table 3) yields an estimate of the persistence time $\tau = 2.6$ years which is quite different from typical estimates for gridded instrumental surface temperature records ($\tau \approx 1$ year–see Allen and Smith, 1994). The robust analysis (Table 3–case "a") provides a far more consistent estimate $\tau = 0.7$ year. As summer seasons tend to erase the climatic memory associated with persistent cold-season anomalous circulation patterns (consistent with the "spring barrier" often discussed in the context of medium-range climatic forecasting), $\tau \approx 1$ year is a physically reasonable timescale for the noise persistence.

We chose this time series example to test the sensitivity of the procedure to the parameter the Δf_{smooth} and the threshold level required in the reshaping procedure. The values of ρ and S_0 obtained with a larger value $\Delta f_{smooth} = 0.11$ cycle/year (within the range of widths specified for our selection rules–0.05-0.13 cycle/yr in this case) yields a very similar estimation of the noise background (Table 3–case "b") and significant signals (Figure 6d). Lowering the reshaping threshold from the 99% to 95% confidence level, however, now leads us to infer periodic structure in the the interannual "ENSO" variations. This distinction demonstrates the ultimate subjectivity in deciding whether or not a given spectral peak is considered as representing periodic or quasiperiodic variability. Our red-noise significance criterion, however, isolates the same peaks as significant whether they are considered periodic or quasiperiodic.

5.3 Great Salt Lake

The 144 year (1846-1989) historical record (see Sangoyami, 1993) of monthly volume change in the Great Salt Lake ("ΔV"–Figure 7a) is analyzed here. The raw red-noise determination (Figure 7b) isolates an annual cycle and its first two harmonics as significant, but does not indicate any significant low-frequency variability, other than a narrow 16-17 year peak. The estimated noise level is elevated by the annual cycle, its harmonics, and a peculiar broad band of elevated variance on interdecadal time scales. The robust analysis (Figure 7c), in contrast, suggests that a fair amount of the low frequency variability in the record is significant, including peaks in the interannual ENSO band and a broad 10-70 year band (Figure 7d). Specific modes of quasidecadal (10-12 year) and secular ($>$ 50 year timescale) variability have been shown to correlate well with large-scale modes of climatic variability (Mann et al, 1995; Lall and Mann, 1995). The significant broad-band of interdecadal variability is unusual for a hydroclimatic time series and may relate to the tendency of large scale forcing on decadal and interdecadal timescales to interact with the lakes own low-frequency internal non-linear dynamics (see Lall et al, 1995). The robust analysis indicates all of the harmonics of the annual cycle (ie, up to f_N) as being significant. It is noteworthy that the annual peak is found to contain a significant quasiperiodic (ie, non phase-coherent) component. This feature is consistent with observations that the annual cycle in western precipitation has exhibited a phase drift during the last century (Rajagopalan and Lall, 1995) with potential connections to anthropogenic climate change (Thomson, 1995).

6 Instrumental "global average" temperature record

We analyze here the 137 year long record of annually-averaged estimated global-average surface temperature anomalies of the Climatic Research Unit of East Anglia University (see e.g., Jones et al 1986). Visually the series exhibits a clear long-term warming trend (Figure 8a). The character and significance of this trend is a subject of considerable recent interest and controversy (Kuo et al. 1990; Ghil and Vautard, 1991; Bloomfield, 1992; Bloomfield and Nychka, 1992). Less visually obvious signals in this record on interannual, decadal, and interdecadal timescales have also been detected (Ghil and Vautard, 1991; Elsner and Tsonis, 1991; Allen and Smith, 1994). Mann and Park (1993;1994) favor *multivariate* analysis of global temperature gridpoint data because coherent, high-amplitude signals (e.g., ENSO) partially cancel in coarse spatial averaging due to spatial variation in the phase. Multivariate methods can also provide alternative "local" confidence tests that do not require detailed knowledge of the overally shape of the noise spectrum (Mann and Park, 1994).

Proper determination of the noise background is essential to detecting an anthropogenic signal (see e.g., Wigley and Raper 1990; Bloomfield and Nychka 1992). A non-robust analysis of the un-detrended series leads to the highly questionable inferences (Figure 8b) that the secular trend is *not* significant relative to red noise, and that nearly the entire interannual band (all periods of 7 years or shorter) of variability is significant well above the 99% confidence level. Furthermore, the value of $\rho = 0.93$ (Table 3) provides a noise persistence timescale estimate $\tau = 14$ years which is unphysically long compared to the typical 1 year timescale estimated for regional surface temperatures data (see section 4.2). In contrast, if the estimated low-frequency trend is first removed before noise parameters are estimated (Allen and Smith, 1994), a more reasonable estimate, albeit still slightly larger estimates of $\rho \approx 0.6$ and $\tau \approx 2$ years) are obtained–Allen and Smith (1994) argue for an iterative procedure to successively remove all detected signals from the series. Our robust procedure is insensitive to the contributions of the secular warming trend or any residual century-scale signals (see Schlesinger and Ramankutty, 1994a) and decadal and interannual timescale signals. This procedure yields noise parameters of $\rho = 0.385$ and $\tau \approx 1$ year (Table 3) which are highly consistent with that obtained for regional surface temperature records. The discrepancy between the raw estimates of τ for globally-averaged and regional surface temperature records arises from the subtle ways in which signal and noise add differently in the global-averaging process (see Mann and Park, 1994). The apparent enhanced noise persistence in the global-average series is thus an artifact which disappears in robust noise parameter estimation. In this sense, there appears to be no fundamental difference in the characteristics of typical regional and 'global' surface temperature noise.

The robust analysis argues for a highly significant secular trend (Kuo et al 1990; Wigley and Raper, 1990; Bloomfield and Nychka, 1992; Ghil and Vautard 1992; Allen and Smith, 1994; Mann and Park 1994) relative to red noise. Whether this signal arises as a response to greenhouse warming, from organized low-frequency climate variability, or some combination can not be addressed from an empirical analysis alone, but is potentially addressed by model prediction/observation intercomparison studies (e.g., IPCC 1990; Santer et al, 1995). Weakly significant signals on decadal and typical interannual ENSO timescales are also observed here, consistent with evidence from spatiotemporal analysis (Mann and Park 1994). A quasibiennial (≈ 2.1 year period) signal is isolated as significant only slightly below the 90% confidence level, but is close enough to the Nyquist frequency ($f_N = 0.5$ cycle/year) that aliasing effects may be problematic. As previously noted by Mann and Park (1993), no *periodic* low-frequency signals are detected in the global temperature record.

A recent study by Schlesinger and Ramankutty (1994a) has argued for the existence of a low-frequency ≈ 70 year timescale oscillation in the global-average temperature series superimposed on an anthropogenic warming trend. In their multivariate analysis of globally-distributed long-term instrumental temperature data, Mann and Park (1994) observe a similar signal centered in the North Atlantic. A recent multivariate analysis of proxy data (Mann et al, 1995*b*) suggests a robustness to such oscillations over several centuries. Whether or not signal can be detected in the global-average series, however, appears to depend quite strongly on assumptions regarding the nature of the secular warming signal and the underlying red noise background (Elsner and Tsonis, 1994; Schlesinger and Ramankutty, 1994b). We revisit this issue in our present analysis of this series. Although more sophisticated detrending is often invoked (e.g., Ghil and Vautard, 1990) and a model-predicted response to combined anthropogenic greenhouse/aerosol forcing was used by Schlesinger and Ramankutty (1994a) to define the secular trend, we choose to assume a simpler linear "anthropogenic" warming trend to illustrate some of the subtle issues involved. The spectrum of the detrended series (Figure 8d) is noticeably altered at timescales as short as the decadal ($f \approx 0.1$ cycle/year) by the detrending process which indicates one of the potential problems in detrending: the removal of a simple trend can lead to an undesirable bias of higher frequency variability. The "raw" red-noise analysis shows a peak near 70 year period, barely breaching the 90% confidence level. We no longer detect a significant quasidecadal peak, and the significance of the quasibiennial and ENSO peaks are altered. We are led to conclude that the simple process of trend removal followed by noise parameter estimation can lead to biased inferences. If instead, however, we apply the *robust* analysis to the detrended series, we are led to nearly identical conclusions to those in the robust analysis of the original series with the same peaks isolated at nearly identical significance levels, and very similar noise parameters estimated (Table 3–case b). Furthermore, the residual ≈ 70 year peak is found to be significant at the 99% level relative to the robustly-estimated red noise level. The robust analysis yields consistent results whether applied to the original or detrended series because it is insensitive to local features in the spectrum such as a narrow secular peak.

This latter analysis shows that the global-average series contains a significant residual peak centered near 70 year period relative to the assumption of red noise superimposed on a linear warming trend. Trend removal based on incorrect or overly simple *a priori* assumptions, however, could clearly lead to incorrect inferences. For example, if the true trend is an exponential increase, the residual of linear-detrending will yield an apparently significant low frequency oscillation purely as an artifact of detrending. In this case, both the linear trend and apparent low-frequency oscillation might appear to represent distinct statistically significant signals, when only one true low-frequency signal was actually present. The multivariate methods discussed above can resolve distinct similar timescale signals through their differing spatial as well as temporal characteristics, underscoring another advantage of multivariate methods when spatially-distributed data are available.

One argument against the proposition that the observed warming of the last century is due to enhanced greenhouse conditions is that certain *ad hoc* null hypotheses for the noise-like component of the system can explain away the observed warming trend (e.g., Bloomfield, 1992). Our analysis would argue that a parsimonious model for the behavior of global temperature series can be provided which is consistent with our understanding of the physics governing the climate system. The global-average temperature series is consistent with a a few weakly significant quasi-oscillatory components superimposed on a secular warming trend, and immersed in red noise with a roughly 1 year persistence timescale. The quasiperiodic nature of the oscillatory components is consistent with internal organized dynamics of the climate system, while the long-term warming trend is within the bounds of climate model predictions of the response to combined radiative and aerosol anthropogenic forcing of the climate.

7 Conclusions

We have developed a new technique for signal detection in climate time series that is faithful to the physics governing the climatic system. Our procedure uses "multitaper" spectral estimation in concert with a robust noise level determination to allow for optimal detection of periodic and quasiperiodic signals buried in climatic red noise. The effectiveness of this procedure is verified through an analysis of synthetic examples chosen to simulate a variety climate signals immersed in moderate red-noise backgrounds. The common detection of similar signals (e.g., a 3-5 year ENSO timescale signal) in disparate climate time series, and the consistent estimate of the ≈ 1 year noise persistence time in regional and global surface temperature, strengthens our confidence in the underlying red noise model as it is "robustly" estimated in our procedure. A highly significant secular warming trend, and moderately significant quasiperiodic decadal and interannual signals in the global temperature record are detect relative to a properly estimated red noise background by our analysis. A residual "oscillation" centered near 50-70 year period is found to be highly significant if a simple (linear) warming trend due to anthropogenic forcing is invoked. However, this feature could alternatively be viewed as a non-uniformity in the warming trend itself. Neither possibility can be ruled out by the analysis of this single series.

Appendix A

A1. Spectrum estimation and reshaping

We use Thomson's (1982) procedure for separating the continuous and broadband features of a spectrum which is described elsewhere (e.g., Percival and Walden, 1993) in more detail. The raw multitaper power spectrum $S(f)$ of a time series $\{y_n\}_{n=1}^{N}$ (from which one typically removes the mean) is constructed from of a group of independent $k = 1, .., K$ spectral estimates,

$$Y_k(f) = \sum_{n=1}^{N} w_n^{(k)} y_n e^{i2\pi f n \Delta t}$$

as

$$S(f) = \frac{\sum_{k=1}^{K} \lambda_k |Y_k(f)|^2}{\sum_{k=1}^{K} \lambda_k}$$

within some overall normalization. Δt is the sampling interval (e.g., monthly or annual), and the data taper $\{w_n^{(k)}\}_{n=1}^{N}$ is the kth member in an orthogonal sequence of K Slepian tapers determined by a variational condition to minimize spectral leakage outside of a band of $\pm p f_R$ (ie, of bandwidth $2pf_R$) about a given frequency f where $f_R = 1/(N\Delta t)$ is the Rayleigh frequency. λ_k is a measure of the fractional leakage associated with the kth data taper. Only the first $2p - 1$ tapers are usefully resistant to spectral leakage, so one forced to use $K \leq 2p - 1$ tapers in spectral estimation. The choice of K thus represents a compromise between spectral resolution and variance. In this study, we use either $p = 2$ or $p = 3$, with $K = 2p - 1$ in each case. Shorter time series require the lower values to achieve the desired frequency resolutions.

The contributions from each of the K eigenspectra can be variably weighted to further minimize broadband leakage through an *adaptive* spectral estimate

$$S(f) = \frac{\sum_{k=1}^{K} b_k^2(f)\lambda_k |Y_k(f)|^2}{\sum_{k=1}^{K} b_k^2(f)\lambda_k}$$

where $b_k(f)$ is a weighting function that further guards against broadband leakage for a non-white ("colored") but locally-white process. The adaptive spectrum estimate has an effective degrees of freedom ν that generally departs only slightly from the nominal value $2K$ of the non-adaptively weighted spectrum. Spectral power associated with the low-frequency trend in a data series is confined to a "secular band" about zero frequency, $0 < f < pf_R$ (for $p = 2$, and a 100 year data series, this corresponds to variability with timescale greater than 50 years). As broadband leakage is typically limited to far less than 1% in the adaptive MTM spectrum estimate, long-term trends generally should not bias higher frequency variability as long as dynamic ranges are not extremely large (ie, many decades in power). Such dynamic ranges are not found in any climatic time series we've analyzed.

A test for *periodic* signals (ie, for phase-coherent variability) can be performed based on the ratio of the variance explained by a sinusoid at a given frequency to that explained by the adaptively weighted spectral estimate,

$$F\,(f) = \frac{(K-1)\,|\mu\,(f)\,|^2 \sum_{k=1}^{K} {W_k^0}^2}{\sum_{k=1}^{K} |Y_k\,(f) - \mu\,(f)\,W_k^0|^2}$$

where W_k is the Fourier transform of the data taper w^k, and ${W_k}^0 = W_k\,(f=0)$

$$\mu\,(f) = \sum_{k=1}^{K} {W_k}^0 Y_k\,(f) \,/ \sum_{k=1}^{K} {W_k^0}^2 \tag{7}$$

For locally white noise, $F(f)$ is distributed with 2 and $2K - 2$ degrees of freedom, allowing a test for periodic components relative to the null hypothesis of locally-white noise. The influence of any periodic components can be removed by a reshaping procedure (Thomson, 1982; Lees, 1995), providing an estimate of the continuous background spectrum. Simple subroutines to perform MTM spectral analysis (Lees and Park, 1994) are available over the internet.

We have modified the original reshaping procedure. To detect a putative *periodic* component (and perform reshaping) at frequency f, we require that the F-test confidence level be sufficiently high at f (e.g., 95% or 99% confidence level), *and* that the overall spectral power at f be significant (see A3) relative to an interrum estimate of the background spectrum obtained by a median-smoothing (see A2) of the reshaped spectrum. Thus, our procedure for periodic signal identification is somewhat more conservative than that of Thomson (1982). Our approach assumes that a true climate signal must have power that is significant relative to a smooth estimate of the overall noise background, whether or not the phase information suggests the presence of a harmonic signal. The true noise background will subsequently be estimated by the red noise fit to the median-smoothed spectrum. Significance of periodic signals, as described above, and quasiperiodic signals (narrowband peaks in spectral power associated with an incoherent phase spectrum) will ultimately be measured with respect to this noise background (see A3).

A2. Median smoothing and red-noise background fit

We perform a robust (as earlier defined) estimate of the spectrum background by using a median-smoothing estimator of the reshaped spectrum: at each frequency point, the spectrum estimate is replaced with the median value in a moving window centered at that point. Near the edges of the spectrum (ie, $f = 0$ or $f = f_N$), the window is simply truncated to include fewer points. While the statistical constraint near these boundaries is thus not very good, the smooth red-noise spectrum that is fit to the median-smoothed estimate is insensitive to the precise boundary conditions of the smoothing. Methods with perhaps more desirable statistical properties are available (e.g., Thomson, 1977) but we favor our approach for its simplicity and ease of implementation. Our procedure is found to provide correct statistical results when applied to synthetic red noise processes (A3).

Useful guidelines for selecting the width of the smoothing window Δf_{smooth} are,

a) $\Delta f_{smooth} \lesssim f_N/4$

b) $\Delta f_{smooth} \gtrsim \min(pf_R, 0.05)$ cycle/year

These rules reflect two assumptions:

a) to resolve the overall shape of the spectrum, the full bandwidth of the spectrum (ie, the Nyquist frequency, f_N), which depends on the sampling rate of the series, should be several (at least 4) multiples of the window width.

b) the minimum smoothing width should be greater than both of two characteristic bandwidths: the spectral resolution, which depends on the time-frequency bandwidth product p chosen for spectral estimation as well as the length of the time series (through f_R), and the maximum intrinsic bandwidth of typically observed climate signals (e.g., the $\approx$4-6 year interannual ENSO signal, which is about 0.05 cycle/year).

The latter rule (b) represents the subjective element in window selection, and is based on *a priori* assumptions regarding the nature of signals in the record. For a time series of length 100 years, annual time resolution, and time-frequency product $p = 3$ for example, $\Delta f_{smooth} \approx 0.05 - 0.13$ cycle/yr would be appropriate, while for a time series of length 50 years, monthly resolution, and time-frequency bandwidth $p = 5$, $0.1 < \Delta f_{smooth} < 1.5$ cycle/yr should be employed. Our experience shows that the results of the analysis are relatively insensitive to the exact choice within such specified ranges.

Once the background has been estimated through the median-smoothing procedure, we use Brent's method (Press et al., 1992) to perform an efficient search for the best fit (in the least-squares sense) values of ρ and S_0 in (2) to the median-smoothed spectrum. A fit to $S(f)$ will be more weighted towards the low-frequency, high amplitude portion of the spectrum than a fit to $\log S(f)$. The latter procedure may provide a preferable alternative when the dynamic range of the spectrum is large. Any significance estimates that are strongly dependent on whether the red noise background is fit to $S(f)$ or $\log S(f)$ should not be interpreted with great confidence.

A3. Confidence level determination

While sophisticated non-parametric methods have been developed (Thomson and Chave, 1991), we have found the elementary sampling theory of the power spectrum (Tukey, 1950; Percival and Walden, 1993) to provide satisfactory results. The latter sampling theory, used by Gilman et al (1963) in their seminal investigation of red-noise confidence testing in climate spectra, assumes that the spectrum are χ^2 distributed with ν degrees of freedom where ν is the number of degrees of freedom in the spectral estimate. As described above, $\nu \approx 2K$ degrees of freedom is an approximation for the adaptive multitaper spectrum estimate. The ratio of power associated with a peak in the spectrum to the local power level of the background is assumed to be distributed as $\frac{\chi^2}{\nu}$, and can be compared to the tabulated χ^2 probability distribution to determine peak significances. Although the elementary sampling theory is strictly valid only for stationary processes (e.g., processes exhibiting a white noise background), it provides a suitable description for the weak-to-moderate red-noise found in climate timeseries, which can be treated as locally white. In such cases, the K spectral estimates are only weakly correlated. We tested the validity of our confidence testing procedure by performing Monte Carlo simulations to produce pure AR(1) processes. We verified that the number of peaks exceeding given confidence levels equaled (within small sampling fluctuations) the expected rates of chance incidence (ie, roughly 1% of the spectrum was found to exceed the 99% level, 5% of the spectrum found to exceed the 95% confidence level, etc.).

Acknowledgments

Helpful comments regarding the manuscript from Barry Saltzman, Jeffrey Park, David Thomson, Michael Dettinger, Myles Allen, and an anonymous reviewer are gratefully acknowledged. M. Mann has been supported by NSF grant ATM-9222591 (Climate Dynamics Program). J. Lees was supported under NSF NEHRP grant EAR-9011441 and the donors of the Petroleum Research Fund, PRF 26595-G2, administered by the American Chemical Society.

References

Allen, M.R. and Smith, L.A.: 1994, 'Investigating the origins and significance of low-frequency modes of climate variability', *Geophys. Res. Lett.*, **21**, 883–886.

Allen, M.R.: 1992, *Interactions between the atmosphere and oceans on time-scales of weeks to years*, Ph.D. thesis, University of Oxford. 202 pages.

Bartlett, M.S.: 1966, *An introduction to Stochastic Processes*, Cambridge, Cambridge University Press, 362 pp.

Birchfield, G.E., and Ghil, M.: 1993, 'Climate Evolution in the Pliocene-Pleistocene as Seen in Deep Sea δ^{18} Records and in Simulations: Internal Variability Versus Orbital Forcing', *J. Geophys. Res.*, **98**, 10385–10399.

Bloomfield, P.: 1992, 'Trends in Global Temperature', *Climatic Change*, **21**, 1–16.

Bloomfield, P. and Nychka, D.: 1992, 'Climate Spectra and Detecting Climatic Change', *Climatic Change*, **21**, 275–287.

Brillinger, D. R.: 1981, *Time Series, Data Analysis and Theory*, Holden-Day.

Currie, R.G. and Fairbridge, R.W.: 1985, 'Periodic 18.6-year and cyclic 11-year induced drought and flood in northeastern China and some global implications', *Quat. Sci. Rev.*, **4**, 109–134.

Delworth, T., Manabe, S., and Stouffer, R.J.: 1993, 'Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model', *J. Clim.*, **6**, 1993-2011.

Dickey, J.O., Marcus, S.L., and Hide, R.: 1992, 'Global propagation of interannual fluctuations in atmospheric angular momentum', *Nature*, **357**, 484–488.

Ghil, M., and Vautard, R.: 1991, 'Interdecadal oscillations and the warming trend in global temperature time series', *Nature*, **350**, 324–327.

Gilman, D.L., Fuglister, F.J. and Mitchell, J.M. Jr.: 1963, 'On the power spectrum of "red noise"', *J. Atmos. Sci.*, **20**, 182–184.

Hasselmann, K.: 1976, 'Stochastic climate models. Part I: Theory', *Tellus*, **28**, 473–485.

IPCC: 1990, *Climate Change, the IPCC Scientific Assessment,* J.T., Houghton, G.J. Jenkins, and J.J Ephraums, Eds., Cambridge University Press.

Jones, P. D., Wigley, T. M., and Wright, P.B.: 1986, 'Global temperature variation between 1861 and 1984', *Nature*, **322**, 430–434.

Julian, P.R.: 1961, *A study of the statistical predictability of stream-runoff in the upper Colorado River Basin*, Boulder, Bureau of Economics Research, University of Colorado, 98 pp.

Keigwin, L.D., Curry, W.B., Lehman, S.J., and Johnsen, S.: 1994, 'The role of the deep ocean in North Atlantic climate change between 70 and 130 kyr ago', *Nature*, **371**, 323–326.

Keppenne, C.L., and Ghil, M.: 1992, 'Adaptive filtering and prediction of the Southern Oscillation index', *J. Geophys. Res.*, **97**, 20449–20454.

Kuo, C., Lindberg, C. and Thomson, D.J.: 1990, 'Coherence established between atmospheric carbon dioxide and global temperature', *Nature*, **343**, 709–713.

Kutzbach, J.E. and Bryson, R.A.: 1974, 'Variance spectrum of Holocene climatic fluctuations in the North Atlantic Sector', *J. Atmos. Sci*, **31**, 1958–1963.

Lall, U., Mann, M.E., The Great Salt Lake: A Barometer of Low-Frequency Climatic Variability, *Water Resources Research*, **31**, 2503–2515, 1995.

Latif, M. and T.P Barnett, 1994: 'Causes of decadal climate variability over the North Pacific and North America', *Science*, **266**, 634–637.

Lall, U., Sangoyomi, T. and Abarbanel, H.D., 1996: 'Nonlinear dynamics of the Great Salt Lake: nonparametric short term forecasting', *Wat. Res. Res.*, in press.

Lees, J.M.: 1995, 'Reshaping spectrum estimates by removing periodic noise: Application to seismic spectral ratios', *Geophys. Res. Lett.*, **22**, 513–516.

Lees, J.M. and Park, J.: 1995, 'Multi-taper spectral analysis: A stand-alone C-subroutine', *Computers & Geology*, **21**, 199-236.

Mandelbrot, B.B. and Wallis, J.R.: 1969, 'Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long Run Statistical Dependence', *Water Res. Res.*, **5**, 967–988.

Mann, M.E. and Park, J.: 1993, 'Spatial correlations of interdecadal variation in global surface temperatures', *Geophys. Res. Lett.*, **20**, 1055–1058.

Mann, M.E. and Park, J.: 1994, 'Global scale modes of surface temperature variability on interannual to century time scales', *J. Geophys. Res.*, **99**, 25819–25833.

Mann, M.E., Lall, U., Saltzman, B.: 1995, 'Decadal-to-century scale climate variability: Insights into the Rise and Fall of the Great Salt Lake', *Geophysical Research Letters*, **22**, 937–940.

Mann, M.E., Park, J., Bradley, R.S.: 1995, 'Global Interdecadal and Century-Scale Climate Oscillations During the Past Five Centuries', *Nature*, **378**, 266–270.

Mitchell, J.M.: 1976, 'An overview of climatic variability and its causal mechanisms', *Quaternary Res.*, **6**, 481–493.

Newell, N.E., Newell, R.E., Hsiung, J. and Zhongxiang, W.: 1989, 'Global marine temperature variation and the solar magnetic cycle', *Geophys. Res. Lett.*, *16*, 311–314.

Park, J., Lindberg, C.R., and F. L. Vernon III: 1987, 'Multitaper spectral analysis of high-frequency seismograms', *J. Geophys. Res.*, **92**, 12675–12684.

Park, J., and Maasch, K.A.: 1993, 'Plio-Pleistocene time evolution of the 100-kyr cycle in marine paleoclimate records', *J. Geophys. Res.*,**98**, 447–461.

Percival, D.B. and Walden, A.T.: 1993, *Spectral Analysis for Physical Applications*, Cambridge University Press, Cambridge, UK, 583 pp.

Plaut, G., Ghil, M. and Vautard, R.: 1995, 'Interannual and Interdecadal Variability in 335 Years of Central England Temperatures', *Science*, **268**, 710–713.

Press, W. H., Flannery, B. P., Teukolsky,S. A., and Vetterling, W. T.: 1992, *Numerical Recipes*, Cambridge University Press, Cambridge, UK, 994 pp.

Rajagopalan, B. and Lall, U.: 1995, 'Seasonality of precipitation along a meridian in the western United States', *Geophys. Res. Lett.*, **22**, 1081–1084.

Sangoyomi, T.: 1993, 'Climatic Variability and Dynamics of Great Salt Lake Hydrology', Ph.D dissertation, Utah State University, Logan, Utah, 247pp.

Santer, B.D., Mikolajewicz, U., Bruggemann, W., Cubash, U., Hasselmann, K., Hock, H., Maier-Reimer, E. and Wigley, T.M.L.: 1995, 'Ocean variability and its influence on the detectability of greenhouse warming signals', *J. Geophys. Res.*, **100**, 10693–10725.

Schlesinger, M.E. and Ramankutty, N.: 1994, 'An oscillation in the global climate system of period 65–70 years', *Nature*, **367**, 723–726.

Schlesinger, M.E. and Ramankutty, N.: 1994, 'Reply to "Low-frequency oscillation"', *Nature*, **372**, 508–509.

Elsner, J.B. and Tsonis, A.A.: 1994, 'Low-frequency oscillation', *Nature*, **372**, 507–508.

Stocker, T.F. and Mysak, L.A.: 1992, 'Climate fluctuations on the century time scale: A review of high-resolution proxy data and possible mechanisms', *Climatic Change*, **20**, 227–250.

Thomson, D.J.: 1977, 'Spectrum estimation techniques for characterization and development of the WT4 waveguide-I and II', *Bell. Syst. Tech. J.*, **56**, 1769–2005.

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

Thomson, D.J.: 1990, 'Time series analysis of Holocene climate data', *Phil. Trans. R. Soc. Lond. A*, **330**, 601–616.

Thomson, D.J.: 1990, 'Quadratic-inverse spectrum estimates: applications to palaeoclimatology', *Phil. Trans. R. Soc. Long. A*, **332**, 539–597.

Thomson, D.J., Chave, A.D.: 1991, 'Jackknifed error estimates for spectra, coherences, and transfer functions' in *Advances in Spectrum Analysis and Array Processing*, (ed. S. Haykin), Prentice Hall.

Thomson, D.J., 1995, 'The seasons, global temperature, and precession', *Science, 268*, 59–68.

Tukey, J.W.: 1950, 'The sampling theory of power spectrum estimates', *Symposium on applications of autocorrelation analysis to physical problems*, Woods Hole, Mass, 13-14 June 1949, Office of Naval Research, Washington, D.C., 47–61.

Trenberth, K.E., and Shin, W.T.K.: 1984, 'Quasibiennial fluctuations in sea level pressures over the Northern Hemisphere', *Mon. Wea. Rev.*, **112**, 761–777.

Elsner, J.B. and Tsonis, A.A.: 1991, 'Do bidecadal oscillations exist in the global temperature record?', *Nature*, **353**, 551–553.

Vautard, R., Yiou, P., and Ghil, M.: 1989, 'Singular-spectrum analysis: A toolkit for short, noisy chaotic signals', *Physica D* **35**, 395–424.

Vernon, F. L., J. Fletcher, L. Carroll, A. Chave and Sembrera, E.: 1981, 'Coherence of seismic body waves from local events as measured by a small-aperture array', *J. Geophys. Res.*, **96**, 11981–11996.

Wigley, T.L, and Raper, R.: 1990, 'Natural variability of the climate system and detection of the greenhouse effect', *Nature*, **344**, 324–327.

Yiou, P., Genthon, C., Ghil, M., Jouzel, J., LeTreut, H., Barnola, J.M., Lorius, C., Korotkeitch, Y.N.: 1991, 'High-frequency paleovariability in climate and in CO_2 levels from Vostok icecore records', *J. Geophys. Res.*, **96**, 20365–20378.

Figure Captions

Fig. 1

Schematic illustration of the procedure for signal identification and confidence level estimation. The reshaping step is accomplished by using a 95% or 99% confidence cut-off for both the F-test and the spectral power amplitude. Non-linear fits are calculated using Brent's method to minimize the misfit of the AR(1) model and the robust estimate of the smoothed spectrum.

Fig. 2

Synthetic example representing 120 years of monthly sampled data with a red-noise background generated from $\rho = 0.4$, and the signals described in the text.

a. Time series of signal (top), noise (middle), and signal-plus-noise (bottom)

b. Multitaper spectrum of signal component of time series using $K = 3, p = 2$, with peaks labeled by periodicity in years.

c. Multitaper spectrum of signal-plus-noise showing red-noise fit to spectrum (solid curve) and 90,95,99% confidence limits (dashed curves) based on raw AR(1) fit to spectrum. As in later plots, the non-reshaped (dashed) spectrum is plotted on top of the reshaped (solid) spectrum so that "periodic" signals are indicated by narrow dashed peaks, while more "quasiperiodic" signals are indicated by solid peaks. For a signal that has both significant periodic and quasiperiodic components, the area between the reshaped and unreshaped spectra is a measure of the relative fraction of variance that appears to represent periodic variability. The median-smoothed background (thick dot-dashed) spectrum is also shown. Higher-frequency peaks are labeled by periodicity in years, while lower-frequency region (surrounded by box) is expanded in the following figure. In this and all similar plots, unless otherwise noted, a 95% confidence level for the F-test and spectral power is used in the reshaping procedure.

d. Multitaper spectrum of signal-plus-noise, as in (c) but based on robust estimate of AR(1) background. The true red-noise background (heavy solid) is also shown for comparison.

Fig. 3

Spectra of first synthetic example, focusing on on interannual range (0.0 to 0.5 cycle/yr frequency range)

a. Same as 2c, but restricted to interannual frequency range

b. Same as 2d, but restricted to interannual frequency range

Fig. 4

Synthetic example representing 120 years of annual sampled data with a red-noise background generated from $\rho = 0.2$, and the signals described in the text.

a. Time series of signal (top), noise (middle), and signal-plus-noise (bottom)

b. MTM spectrum of signal component of time series.

c. MTM spectrum of signal+noise using showing red-noise fit to spectrum (solid) and 90,95,99% confidence limits (dashed) based on raw AR(1) fit to spectrum. The median-smoothed spectrum is also shown (thick dot-dashed)

d. Multitaper spectrum of signal+noise, as in (c) but based on a robust estimate of AR(1) background. The true red noise background (heavy solid) is also shown for comparison.

Fig. 5

ODP site 607 core

a. Time series

b. MTM spectrum of the δ^{18}O series, showing red-noise fit to spectrum (solid) and 90,95,99% confidence limits (dashed) based on raw AR(1) fit to spectrum. The median-smoothed spectrum is also shown (thick dot-dashed)

c. MTM spectrum as above, but based on a robust estimate of the red-noise background. Thick curves are based on the log-fitting of the red-noise spectrum.

Fig. 6

Fennoscandinavian summer temperature tree-ring reconstruction

a. Raw annual time series (dashed) along with a 20 year moving average (thick solid) to visually emphasize lower frequency variability.

b. MTM spectrum of the above series showing red-noise fit to spectrum (solid) and 90,95,99% confidence limits (dashed) based on raw AR(1) fit to spectrum. The median-smoothed spectrum (thick dot-dashed) is also shown. A 99% confidence level cutoff is required in the reshaping procedure in this case.

c. MTM spectrum as above, but using robust estimate of the red-noise background.

d. Same as (c) but with the larger median smoothing and a 95% level cutoff. Lower cutoff leads to determination of some periodic signals, labeled by periodicity in years.

Fig. 7

Great Salt Lake month-to-month volume difference time series

a. Raw monthly time series in units of m^3/month (dashed) along with a 10-year moving average (thick solid) to visually emphasize lower frequency variability.

b. MTM spectrum of the above series showing red-noise fit to spectrum (solid) and 90,95,99% confidence limits (dashed) based on raw AR(1) fit to spectrum. The median-smoothed spectrum (thick dot-dashed) is also shown. The low-frequency region (surrounded by box) is expanded in part (d).

c. MTM spectrum as above, but using robust estimate of the red-noise background.

d. Same as (c) but focusing on interannual and longer time scales ($\tau > 2$-year period).

Fig. 8

Global-average surface temperature anomaly record

a. Raw annual time series from 1854-1990.

b. MTM spectrum of the above series using showing red-noise fit to spectrum (solid) and 90,95,99% confidence limits (dashed) based on raw AR(1) fit to spectrum. The median-smoothed spectrum (thick dot-dashed) is also shown.

c. MTM spectrum as above, but using robust estimate of the red-noise background. Peaks significant at near or above the 90% confidence level are labeled with periodicity in years.

d. MTM spectrum but based on time series with linear trend subtracted off (heavy dotted spectrum) along with the original (solid) spectrum, and raw (thin solid) and robust (thick solid) red noise fit. The 90,95,99% confidence limits based on raw (thin dashed) and robust (thick dashed) red noise fit to spectrum are shown. The median-smoothed version of the original spectrum (thick dot-dashed) is also shown.

SERIES	N	Δt	SIGNAL	T (years)	f (cycles/yr)	REL. AMP.
SYNTHETIC #1	1440	1 month	linear trend	∞	0.0	1.5
			periodic sinusoid	25	0.04	1.0
			frequency modulated oscillation	9-11	0.09-0.11	1.0
			frequency modulated oscillation	4.3-4.8	0.21-0.23	1.0
			periodic square wave	1.0	1.0	1.0
			AR(1) noise ($\rho = 0.4$)			1.0
SYNTHETIC #2	120	1 year	linear trend	∞	0.0	1.5
			frequency modulated oscillation	9-11	0.09-0.11	1.0
			frequency modulated oscillation	4.3-4.8	0.21-0.23	1.0
			AR(1) noise ($\rho = 0.2$)			1.0

Table 1: Description of the two synthetic examples, including length and sampling rate of each synthetic series, and details of each component in the time series including qualitative description, period (or period range) of oscillation, frequency (or frequency range), and relative amplitude

SERIES	p	Δf_{SMOOTH}	ρ_{TRUE}	S_{0TRUE}	τ_{TRUE}	ρ_{RAW}	S_{ORAW}	τ_{RAW}	ρ_{RBST}	S_{0RBST}	τ_{RBST}
SYNTHETIC #1	2	1.2 cyc/yr	0.4	$1.78 \cdot 10^3$	0.21 yr	0.91	$3.79{\cdot}10^3$	0.88 yr	0.46	$1.7{\cdot}10^3$	0.11 yr
SYNTHETIC #2	2	0.07 cyc/yr	0.2	$1.2{\cdot}\, 10^2$	0.62 yr	0.39	$2.3{\cdot}10^2$	1.1 yr	0.27	$1.6{\cdot}10^2$	0.76 yr

Table 2: Results of noise-background analysis of synthetic time series, giving time-frequency bandwidth p (where number of tapers used is $K = 2p - 1$), frequency-width of smoothing window, true values of ρ, S_0, and τ, raw (ie, non-robust) estimates of ρ, S_0, and τ and robust estimates of ρ, S_0, and τ

SERIES	Δt	T	N	p	Δf_{SMOOTH}	ρ_{RAW}	$S_{O\,RAW}$	τ_{RAW}	ρ_{RBST}	$S_{0\,RBST}$	τ_{RBST}
ODP 607 δ^{18}O	3.5 kyr	2.8 Ma	788	3	0.007 cyc/kyr	0.927	$1.0 \cdot 10^2$	46 kyr	0.689	$5.39 \cdot 10^1$	9.4 kyr
" (log fit)									0.781	$6.15 \cdot 10^1$	14.1 kyr
SCAN TR. RNG. (a)	1 yr	1482 yr	1482	3	0.06 cyc/yr	0.685	$6.0 \cdot 10^2$	2.6 yr	0.23	$4.4 \cdot 10^2$	0.68 yr
SCAN TR. RNG. (b)					0.11 cyc/yr				0.26	$4.3 \cdot 10^2$	0.75 yr
GSL ΔV	1 mo	144 yr	1731	2	0.05 cyc/yr	0.475	$8.4 \cdot 10^{13}$	1.3 mo	0.41	$5.5 \cdot 10^{13}$	1.1 mo
GLOBAL T_{AVE}	1 yr	137 yr	137	2	0.06 cyc/yr	0.932	$5.0 \cdot 10^0$	14 yr	0.38	$2.0 \cdot 10^0$	1.0 yr
" (detrended)									0.36	$1.9 \cdot 10^0$	1.0 yr

Table 3: Results of noise-background analysis of time series analyzed in this study, giving the sampling rate, length of series (in years), # of samples in series, time-frequency bandwidth p (number of tapers is $K = 2p - 1$ in each case), frequency-width of smoothing window, raw (ie, non-robust) estimates of ρ, S_0, and τ and robust estimates of ρ, S_0, and τ