Statistical analysis of two climate time series

By J o h n R e i d*,
P.O. Box 279
Cygnet
Tasmania 7112
Australia

31 January 2016

A B S T R A C T

Unlike the variance density spectral estimate or 'power spectrum', the sample periodogram is a rigorously defined statistic and, as such, can be used to distinguish between deterministic cycles and stochastic processes in a given time series. Under the null hypothesis, whereby the time series is assumed to be white, the periodogram ordinates are independent and identically distributed and, when normalized, have a χ^2 distribution with two degrees of freedom. This facilitates the testing for significance of both non-white trends and individual spectral peaks. Here we examine two climate time series: the EPICA Dome C Deuterium ice core time series from 490 kyr to the present and the GISS time series of annual global average surface temperature from 1880 to 2014. The EPICA time series showed significant peaks at 41.0 and 23.3 kyr indicating the presence of deterministic cycles with periods close to the periods of obliquity and precession of the Earth's orbit. On the other hand the GISS time series showed no evidence of decadal or multidecadal cycles and is well described by a purely stochastic ARIMA(0,1,1) model. Neither was there evidence of a deterministic trend in the GISS data, the apparent trend being attributable to spurious regression.

Keywords: periodogram, time series, chi-squared periodogram, random walk, spurious regression, spurious cycles, climate time series.

1. Introduction

The periodogram of a time series is defined as the squared modulus of the discrete Fourier transform of the time series. The periodogram is often regarded as a first-pass estimate of the variance density function or power spectrum of the process which gave rise to the time series. As such it has long been deprecated as a poor estimator because it does not converge in probability as the length of the time series increases.

In practice this does not matter. The variance density spectral estimate or 'power spectrum' is a property of an idealized stationary continuous function of infinite extent. In real life we only ever have a finite number of elements to deal with and, in any event, we can expect different underlying processes predominate at different time and frequency scales. The asymptotic properties of the periodogram are irrelevant and, at times, misleading. Furthermore the assumption of stationarity is unnecessarily stringent and precludes analysis of non-stationary phenomena such as the random walk.

On the other hand the periodogram has excellent statistical properties. Unlike the power spectrum, it is a rigorously defined sample statistic which can be used for hypothesis testing. When the process is assumed to be unselfcorrelated with zero mean each normalized ordinate of the periodogram has a χ^2 distribution with two degrees of freedom. Hence if we take as the null hypothesis, the assumption that the sample is the outcome of a white noise process, then it becomes possible to use the periodogram to test for the presence of deterministic cycles; an individual spectral peak whose χ^2 probability exceeds a specified value can be rejected as the outcome of a purely stochastic process and must be the outcome of a deterministic cycle. Likewise the null hypothesis can be rejected when stochastic trends such as band-limited noise or, say, a power law frequency relationship cause more periodogram values to lie outside the (flat) χ^2 confidence levels than would be expected by chance alone. This approach, known as the 'χ^2 periodogram', was first used by Eldridge (1947) for finding circadian rhythms in physiological data.

* Corresponding author.
e-mail: johnsinclairreid@gmail.com

2. Stochastic and deterministic processes

2.1. Definitions

For present purposes a stochastic process is a process in which the state at any time is defined as a probability conditional on previous states, e.g. as a function of previous states plus an additional random variable termed the 'innovation'. A deterministic process is a process in which the state at any given time can be expressed explicitly as a function of the time, e.g. as a trigonometric function or a polynomial. Defining a deterministic process implicitly, say, in terms of differential equations is problematic because chaotic states may occur which have many of the characteristics of a stochastic process. Numerical models of chaotic systems are, in effect, stochastic models because the finite computer word length leads to a quasi-random, round-off error which serves as a *de facto* innovation.

A time series may be the outcome of a combination of stochastic and deterministic processes. One purpose of any statistical analysis is to distinguish between the two components. The χ^2 periodogram is suited to this purpose.

2.2. Random walks

Hasselmann (1976) pointed out that measured intrinsic quantities such as temperature are often the outcome of the integration by natural processes of quasi-random, extrinsic quantities such as heat. As a consequence such measurements can be regarded as the outcome of a stochastic process and can be expected to exhibit a power law spectrum with negative exponent due to such integrating effects. The best known and simplest example of such a process is the 'random walk' obtained when white noise is integrated or summed. It has a power law spectrum with an index of -2.

Pelletier (2002) has shown that variance spectrum of atmospheric temperature exhibits a variety of power law relationships over a wide range of time scales from 2 years to 100,000 years, i.e. over a given time scale

$$S = Af^{\nu} \tag{1}$$

where S is the variance density at frequency, f, and A and ν are constants. Pelletier found that ν lay in the range -2.0 to -0.5 indicating random walks and random-walk-like behaviour.

2.3. Spurious Cycles

Cycles in time series in the time domain are manifested as narrow peaks in the spectrum in the frequency domain. It is well known the unfiltered, unwindowed periodogram is very 'noisy'; random peaks in the periodogram can easily be mistaken as evidence of deterministic cycles in the time domain. This is particularly true when the time series is a random walk. A random peak at the low frequency end of the periodogram of such a time series will be exaggerated by the rapid power-law decline of variance density with frequency. Furthermore such an exaggerated peak will be 'real' in the sense that corresponding quasi-cyclical behaviour may well be discernible to the naked eye in the time domain data. Such spurious peaks are characteristic of random walks and, as such, cannot be removed by either windowing the time series nor by using arcane spectral techniques such as the Maximum Entropy Method (e.g. Reid, 1971). The only way of dealing with them is by the accepted methods of frequentist statistics, i.e. by testing the null hypothesis that a given peak is the outcome of a purely random process. In order to do this a sample statistic with a known distribution must be used; the sample periodogram is such a statistic.

It might be argued that such peaks are not 'spurious' at all since they do indeed reveal information about the time series. This is to misunderstand the nature of statistical modelling which seeks a description of the data with predictive power. The presence of a truly deterministic cycle in the data allows us to make predictions about the system under investigation whereas a spurious cycle tells us no more about future behaviour than does a random walk.

2.4. Spurious correlation and regression

Random walk time series also typically exhibit a rising or falling trend which can mislead observers into assuming that a deterministic linear regression relationship holds between the dependent variable and the time when, in fact, no such relationship exists.

In order to illustrate this effect, one million synthetic time series of length $N_{ts} = 100$ were generated by summing successive normally distributed random numbers using the Python programming language. The Pearson sample correlation coefficient, r, holding between each time series value and its index was found for each time series. The correlation coefficients were binned and normalized and the resulting frequency distribution, f, is shown in Figure 1a. It can be seen that f is symmetrical about zero and bimodal with peaks near -0.91 and 0.91 respectively.

The cumulative distribution, F, of the absolute values, $\mid r \mid$, was also calculated and some values are listed in Table 1. Further numerical experiments indicated that both f and F are largely independent of N_{ts} for $N_{ts} > 10$ or so. It can be seen from Table 1 that the probability of a 100-long random walk time series having a correlation coefficient of 0.5 or greater is 0.657, i.e. nearly two thirds of 100-long random walks correlate with the time index to better than 50 percent. Hence, not only is it possible for a random walk to appear to have a rising or falling trend, it is likely. This is so, even though there is no underlying connection or causality. Such correlations are spurious.

This phenomenon is well known in econometrics where it has been described described as 'spurious regression' by Granger and Newbold (1974). However it is evidently still not understood in the physical sciences where the uncritical fitting of supposed 'trends' by linear regression methods is common.

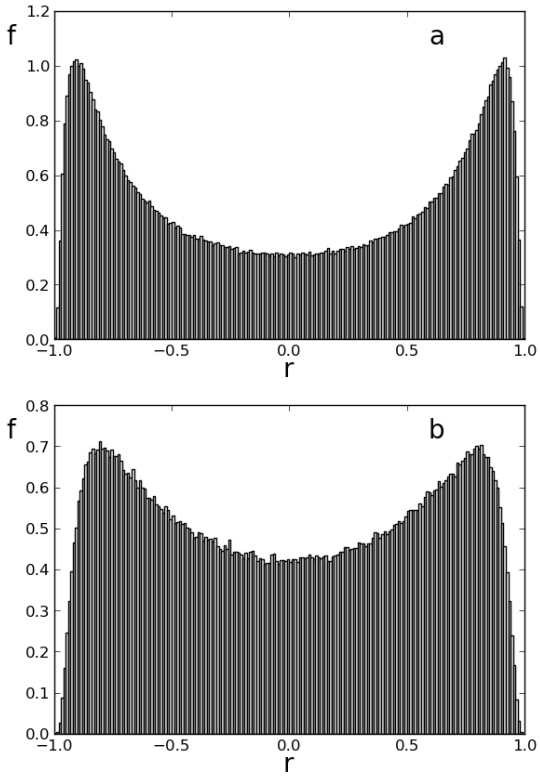

Fig. 1. The observed frequency distribution, f, of the correlation coefficient, r, of value vs. index of one million, synthetic, 100-long, sample sequences: (a) of a true random walk, (b) of an AR(1) process with coefficient, $a_1 = 0.99$

.

It is commonly assumed that spurious regression is an unfortunate consequence of the non-stationarity of the time series due, in turn, to the presence of a unit root in the characteristic equation of the model. The Dickey-Fuller test or the Phillips-Perron test may be used to test for the presence of such a unit root. It is presumed that the absence of a unit root implies that any observed correlation or regression must be real. However that is clearly not the case. Figure 1b shows the observed frequency distribution of the correlation coefficient for a first order autoregressive (AR(1)) process for which the single coefficient has the value $a_1 = 0.99$. The characteristic equation has a root which is close to, but not identically equal to, unity and the process is stationary. Nevertheless the frequency distribution has a similar bimodal shape to that of Figure 1a indicating that spuriously high correlation coefficients are likely in this case as well.

Clearly it is not the presence or absence of a unit root which leads to spurious regression, rather it is the presence of an integrating tendency and the resulting concentration of variance at low frequencies. Although not strictly a random walk, a stochastic process with roots close to unity exhibits similar regression-like behaviour to a true random walk. Indeed it may be impossible to discern the difference between them for a sample time series of finite length. We propose using the term 'wide-sense' random walk for such processes.

$\mid r \mid$	F	$1 - F$
0.50	0.343	0.657
0.60	0.435	0.565
0.70	0.544	0.456
0.80	0.683	0.317
0.90	0.865	0.135
0.91	0.885	0.115
0.92	0.905	0.095
0.93	0.926	0.074
0.94	0.945	0.055
0.95	0.963	0.037
0.96	0.978	0.022
0.97	0.990	0.010
0.98	0.998	0.002
0.99	1.000	0.000

Table 1. The observed cumulative frequency distribution, F, of the absolute value of the correlation coefficient, $\mid r \mid$, for 1,000,000 synthetic independent random walk time series each of length 100.

The periodogram is a useful tool for testing for the presence of deterministic cycles in the data but not for deterministic trends. A rising or falling trend in the time domain data corresponds to behaviour near zero in the frequency domain and, in turn, to divergence of the power law relationship (1) for negative index, ν. Furthermore in most practical circumstances, the ensemble mean is unknown and is assumed zero after subtraction of the sample mean. The value of the periodogram at zero frequency is, in effect, indeterminate and is left undefined at $k = 0$ below. All that the periodogram can reveal is the presence of a wide-sense random walk, power law relationship between variance and frequency. Such a power law relationship indicates that spurious regression is a possibility.

2.5. Time domain modelling

An apparent linear trend in a time series whose periodogram exhibits power-law behaviour can be tested using time domain models. At issue is whether the apparent trend is deterministic or the outcome of a purely stochastic process.

The testing and comparing of models themselves, rather than just the model parameters, has been a persistent topic in the statistical literature since the 1970s. Akaike (1973) proposed an information criterion for comparing models which compared the information accounted for by different models based on the number of parameters and the likelihood function. The AIC (Akaike Information Criterion) only permits an intercomparison of models and not the significance of any particular model. The

AIC finds application in the estimation of the optimum number of parameters of autoregressive (AR) models. Reid (1979) pointed out that this task can be done more effectively by testing subsets of regression coefficients for significance.

There are also a number of other statistical tests of time series models in which it is the residuals that are tested, viz.: Durbin-Watson, Box-Pierce, Ljung-Box and Bruesch-Godfrey. The residuals must be unselfcorrelated. In intuitive terms, if the residuals are self-correlated, then information remains behind which has not been properly accounted for by the model.

Like the AIC some of these tests such as Durbin-Watson only generate a statistic for intercomparison of models rather than for hypothesis testing in the strict sense. They also vary in efficiency. What is needed is an efficient test which allows rejection of a null hypothesis at a given level of significance. The Ljung-Box whiteness test (Ljung and Box, 1978) satisfies these criteria.

3. The periodogram

3.1. Statistical properties

We start with a finite sequence of real numbers, $\{x_n, n = 0, ..., N-1\}$, which is assumed to comprise samples of a varying physical quantity separated by equal intervals of time or covering equal intervals of time, i.e. a 'time series'. The behaviour of the underlying continuous process, should there be one, at time scales shorter than the sampling period, Δt, is of no concern. If there is evidence of aliasing such as, for example a concentration of variance at the high-frequency end of the spectrum then it is the responsibility of the researcher to change the sampling regime to one with a smaller Δt, if that is possible. Likewise the length, N, and span $N\Delta t$ of the time series are assumed to be finite.

This approach differs from power spectrum developments which are based on continuous functions of infinite extent which are assumed statistically stationary and in which the asymptotic behaviour is of major importance. These concepts are most important in the design of electronic circuitry but are seldom if ever encountered in nature where differing physical mechanisms usually prevail at different temporal and spatial scales. In practice, we can deal with only a finite number of discrete quantities if we are to process them numerically and so their asymptotic behaviour is irrelevant.

We are concerned with $\{X_j, j = 0, ..., N-1\}$, the discrete Fourier transform (DFT) of the time series $\{x_n\}$ given by

$$X_j = \sum_{n=0}^{N-1} x_n . e^{-2\pi ijn/N} \quad , \quad j = 0, 1...N-1 \tag{2}$$

where $i = \sqrt{-1}$. In general, the X_j are complex numbers and, because the x_n are real

$$X_{-j} = X_{N-j} = X_j^* \tag{3}$$

where X_j^* is the complex conjugate of X_j. Because of the cyclical character of the DFT, X_j can be defined for arbitrary j since for any integer m

$$X_{j+mN} = X_j \tag{4}$$

The underlying concept is encapsulated by Parseval's theorem which, for the finite, discrete case under discussion here, can be expressed as:

$$\sum_{n=0}^{N-1} x_n^2 = \frac{1}{N} \sum_{j=0}^{N-1} \mid X_j \mid^2 . \tag{5}$$

We define a new sequence, the periodogram, $\{\hat{P}_k\}$, where

$$\hat{P}_k = 2 \left(\mid X_k \mid /N \right)^2 \quad , \quad k = 1, 2...N/2 - 1 \tag{6}$$

$$\hat{P}_{N/2} = \left(\mid X_{N/2} \mid /N \right)^2 .$$

Assuming N is even , then, after substituting from (3), (5) becomes

$$s^2 = \frac{1}{N} \sum_{n=0}^{N-1} x_n^2 = \sum_{k=1}^{N/2} \hat{P}_k . \tag{7}$$

since $\hat{P}_0 = 0$.

If $\{x_n\}$ has zero mean, then s^2 is the sample variance and (7) can be regarded as a conservation law which tells us how the variance is distributed with frequency, $\hat{P}_k$ being that component of the variance associated with the frequency $k\Delta f = k/N\Delta t$. Note that $\hat{P}_k$ has the same units as s^2; to make it a variance density, it must be divided by Δf.

The given time series, $\{x_n\}$, can be considered one realization of an ensemble of random sequences, $\{\xi_n\}$. Likewise $\{X_k\}$ is the corresponding realization of the random sequence $\{\Xi_k\}$, the DFT of $\{\xi_n\}$, and the computed periodogram, $\{P_k\}$ is the realization of the ensemble periodogram, $\{\Pi_k\}$.

The ensemble periodogram is given by the equivalent equation to (6), viz.:

$$\Pi_k = 2 \left(\mid \Xi_k \mid /N \right)^2 \quad , \quad k = 1, 2...N/2 - 1 \tag{8}$$

$$\Pi_{N/2} = \left(\mid \Xi_{N/2} \mid /N \right)^2 .$$

where

$$\Xi_j = \sum_{n=0}^{N-1} \xi_n . e^{-2\pi ijn/N} \quad , \quad j = 0, 1...N/2 \tag{9}$$

The mean ensemble periodogram, P, is then defined as

$$P_k = \mathrm{E}(\Pi_k) \quad , \quad k = 1, 2...N/2 \tag{10}$$

With this formalism it follows, trivially, that the sample periodogram, $\{\hat{P}_k\}$, is a consistent, unbiased estimator of its ensemble mean value.

3.2. Hypothesis testing

When using the periodogram for testing for cyclic behaviour we start with the null hypothesis that the given time series, $\{x_n\}$, is the outcome of a white noise process with zero mean, i.e. that it is unselfcorrelated. Should the time series be demonstrably non-white with a periodogram showing a power law trend with index, $\nu \approx -2$ then we create a new series $\{y_n = x_n - x_{n-1}\}$ which can be assumed white. Equation (10) becomes

$$P_k = \sigma^2/N \ , \ \ k = 1, 2...N/2 - 1 \tag{11}$$

where σ^2 is the variance of each ξ_n.

The random variable, Ξ_k defined by (9) can be written

$$\Xi_k = \Re_k - i\Im_k \tag{12}$$

where

$$\Re_k = \sum_{n=1}^{N-1} \xi_n \cos(2\pi kn/N) \tag{13}$$

and

$$\Im_k = \sum_{n=1}^{N-1} \xi_n \sin(2\pi kn/N) \tag{14}$$

For sufficiently large N, by the central limit theorem, both $\Re_k$ and $\Im_k$ must be Gaussian whatever the distribution of the ξ_n may be. They are also unselfcorrelated, uncorrelated with each other and have zero mean. By symmetry we can assume

$$\mathrm{Var}(\Re_k) = \mathrm{Var}(\Im_k) = \sigma'^2 \ , \text{ say} \tag{15}$$

so that $\Re_k/\sigma'^2$ and $\Im_k/\sigma'^2$ are both normal(0,1) and so

$$| \Xi_k |^2 /\sigma'^2 = \Re_k^2/\sigma'^2 + \Im_k^2/\sigma'^2 \tag{16}$$

has a χ^2 distribution with two degrees of freedom. It follows that

$$\mathrm{E}(| \Xi_k |^2)/\sigma'^2 = 2 \tag{17}$$

so that from (8) and (10)

$$P_k = \frac{2}{N^2}\,\mathrm{E}(\Xi^2) = \frac{4\sigma'^2}{N^2} \tag{18}$$

i.e. $\sigma = 2\sigma'$.

3.3. Detecting a deterministic cycle

The periodogram of a cosine function of arbitrary (i.e. non-integral) frequency and arbitrary phase is given in Appendix A. Consider a time series, $\{x_j\}$, defined as

$$x_j = \epsilon_j + A\cos(2\pi j f_0/N + \phi) \ , \ \ j = 0, 1...N-1 \tag{19}$$

where $\{\epsilon_j\}$ is the outcome of a white noise process with zero mean.

Assuming $\{x_j\}$ is the realization of a random sequence $\{\xi_j\}$ with Fourier transform $\{\Xi_j\}$, then, recalling that $\mathrm{E}(\Xi_k) = 0$, from (19) and (8), the ensemble periodogram is given by

$$\Pi_k = W_k + Q_k \tag{20}$$

where W_k is the periodogram of the white noise component and Q_k is the periodogram of the cosine function given by (A18).

When testing for a spectral peak, we are, in effect, testing the null hypothesis that the amplitude, A, in (19) is zero. Since W_k/σ^2 has a χ^2 distribution, the null hypothesis can be rejected when $Q_k/\sigma^2 < 1 - F(p)$ where $F(p)$ is the cumulative χ^2 distribution with 2 degrees of freedom at significance level, p.

3.4. Detecting a random walk

The periodogram of a random walk is given in Appendix B. At low frequencies, when k is small, the mean ensemble periodogram of a random walk is given by

$$P_k = Ak^\nu \tag{21}$$

so that

$$\log(P_k) = log(A) + \nu\log(k) \tag{22}$$

and $\nu = -2$. i.e. there is a power law relationship which is manifested as a linear relationship on logarithmic scales; the linear relationship, should it exist, is then immediately obvious. The value of the slope coefficient and its standard error can be found by the conventional methods of linear regression. Owing to the skewed nature of the $\log(\chi^2)$ distribution the value of the constant term, $\log(A)$, is biased but the slope is unaffected.

3.5. Discussion

The null hypothesis, that the given time series is the outcome of a white noise process with zero mean, can be rejected when either one or more peaks, Q_k, in (A18) exceed the confidence level calculate from a χ^2 distribution with two degrees of freedom or when the slope, ν, in (22) is significantly less than zero. There may be other situations in which the null hypothesis can be rejected, e.g. the presence of band-limited noise or the presence of a power law trend with $\nu \neq -2$, but these lie beyond the scope of this article. The periodogram is a powerful tool for dealing with the behaviour of a time series in the frequncy domain. There is no requirement that the data be windowed or filtered in any way. In fact such windowing methods preclude the use of the periodogram as a well-defined statistic.

4. A paleoclimate time series

Milankovic (1941) was first to propose that the cycling of climate apparent in the geological record and known as 'the Ice Ages' could be due to secular variations in the earth's orbit around the sun. These variations are thought to give rise, in turn, to variations in solar insolation at latitudes north of 60 N and its effect on the growth and collapse of the Northern Hemisphere ice sheet . The pioneering work of Hays *et al* (1976) identified

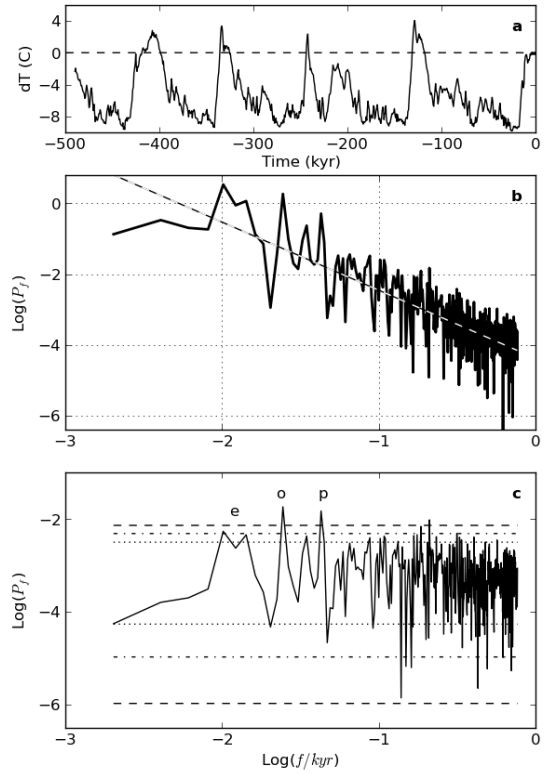

Fig. 2. (a) Time series of proxy global temperature derived from EPICA Dome C Deuterium measurements (Jouzel *et al* 2007). (b) Periodogram showing f^{-2} trend. (c) Periodogram of first differences showing frequencies of eccentricity, obliquity and precession peaks at 'e','o' and 'p' respectively. Upper and lower confidence limits at the 95 percent, 99 percent and 99.9 percent levels are shown by the dotted lines, dot-dash lines and dashed lines respectively.

clear peaks in the ocean sediment record and related them to the eccentricity, obliquity and precession of the orbit. The 'eccentricity peak' near 100 kyr dominated their spectra. However, more recently, a number of workers have questioned the astronomical origin of the 100 kyr peak, because variations in orbital eccentricity can have little effect on insolation, (e.g. Muller and MacDonald (1997) , Ridgwell *et al* (1999) , Lisiecki(2010).

The time series from the EPICA Dome C Ice Core 800 kyr Deuterium Data and Temperature Estimates (Jouzel *et al*, 2007), was downloaded from the World Data Center for Paleoclimatology website. A time series of equally spaced values was generated by averaging all the values in each 653.3 year interval from 490 kyrs to the present epoch (750 samples). It is shown in Figure 2a. Its periodogram spectrum is shown in Figure 2b and resembles the spectra of Hayes *et al*; the classical three spectral

[!h]

Table 2. Orbital Peaks

	e	o	p
calc	.01 kyr^{-1}	.0244 kyr^{-1}	.043 kyr^{-1}
obs	-	$.024 \pm .001$ kyr^{-1}	$.043 \pm .001$ kyr^{-1}

Table 3. Ice Age Terminations

Termination	Time (kyr BP)	Reference
V	430	Augustin *et al*
IV	340	Cheng *et al*
III	246	Cheng *et al*
II	131	Cheng *et al*
I	12	Cheng *et al*

peaks attributed to orbital cycles and listed as 'calc.' in Table 2 can be clearly seen.

The spectrum has a negative slope for index $k > 5$ and a regression line was fitted in the range $5 \leqslant k \leqslant 150$ by the conventional linear regression method and the power law index ν found to be -2.06 ± 0.35, implying that, in this frequency range, the time series can be considered a random walk.

In order to test for the statistical significance of the candidate spectral peaks listed in Table 2, a new series of first differences, $\{y_n = x_n - x_{n-1}\}$ was derived and its periodogram plotted in Figure 2c. The spectrum is now flat with sporadic peaks. Our aim is to test these peak for significance. Upper and lower confidence limits at 95.0, 99.0 and 99.9 percent levels are shown.

4.1. Discussion

In Figure 2c the 'o' and 'p' peaks exceed the upper 99.9 percent confidence limit but the 'e' peak is barely significant at the 99.0 percent level. Since there are 375 ordinates we can expect 3 or 4 ordinates to exceed the 99.0 percent level by chance alone. Furthermore the 'e' peak is a broad peak and involving three neighbouring ordinates and so cannot represent a single deterministic cycle. Thus there is strong evidence of deterministic cycles at the frequencies of the 'o' and 'p' peaks, but it is likely that the 'e' peak is no more than a random walk excursion.

There are, in fact, no cycles with periods near 100 kyr evident in the time series itself. The last five Terminations have an average period close to 100 kyr but closer examination (Table 3) shows that successive differences between terminations are 90, 94, 115, and 119 kyr respectively. They are either about 90 kyr apart or about 120 kyr apart. Huybers and Wunsch (2005) have shown that the ice sheets terminate every second or third obliquity cycle.

Peaks in the rate of temperature change in Figure 2a associ-

ated with the 4 most recent terminations occurred at times of rapidly increasing Northern Hemisphere insolation as noted by Cheng *et al* (2009). However not all times of rapidly changing NH insolation gave rise to terminations. This implies that a fully deterministic causality was not the case. Rather, only the *probability* of a termination increased at times of rapidly incrasing NH insolation. The behaviour of global temperature at these time scales can perhaps best be described as a deterministically modulated stochastic process.

A feature of interest in Figure 1 is the almost constant variance density of the low frequency end of the log spectrum. Allowing for the inherent noisiness, for indices, $k < 5$ or so the spectrum could easily be white, i.e. the random walk $\nu = -2$ trend does not apply at low frequencies. Index $k = 5$ corresponds to an integrating time constant of $\tau \approx 100$kyr. Indeed visual examination of the time series in the upper panel of Figure 2a does indeed suggest a system recovering a stable equilibrium after being driven from stability by an impulse which precipitates terminations. It looks like an integrated Poisson process with an integration time constant, $\tau \approx 40 - 100$ kyr.

5. A contemporary climate time series

The GISS global average temperature anomaly time series, $\{x_i\}$, was downloaded from the GISS/NASA Web site (GISTEMP team, 2015, Hansen *et al*, 2010).

It is shown plotted in the time domain in Figure 3a. A linear trend, derived using the regression model discussed below, is represented by the dashed line. Figure 3b shows the $\{P_k\}$ of the time series of Figure 3a. The logarithmic axes allow power-law trends to be easily recognized as straight lines. It can be seen that there are two such trends: for $k \leqslant 20$ there is a power law trend, $P = Ak^{\nu}$, where the index, ν, is close to -2 and at higher values of k, P_k is almost constant. This suggests that the observed time series is the outcome of a random walk process with added white noise, i.e. of a stochastic, 'ARIMA' process.

Figure 3c shows the periodogram of the time series of first differences. It shows a rising trend consistent with added white noise but, more importantly, there are no significant peaks for $f < 10^{-0.6} yr^{-1}$, i.e. at periods greater than about 4 years. There are no decadal or multidecadal deterministic cycles.

5.1. Time domain model

In order to test the apparent linear trend in temperature we must resort to time domain modelling. The frequency domain analysis of Figure 3 has served a useful purpose in implying that the time series can be fitted with a very simple ARIMA process. The alternative hypothesis is that the time series can be fitted with a linear regression model.

Under the linear regression model the $\{x_i\}$ are assumed to be realizations of a sequence of random variables, $\{\chi_i\}$, deterministically related to the time, but with the addition of stochastic

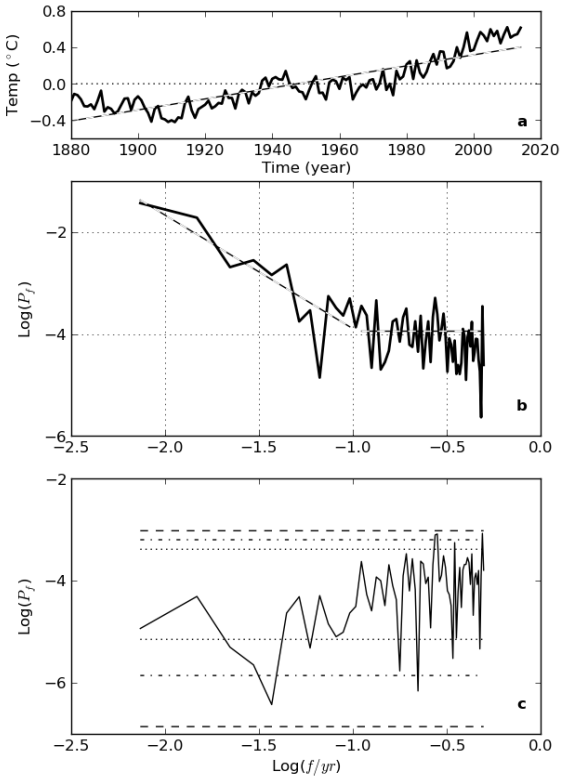

Fig. 3. (a) Time series of the GISS global average temperature anomaly, 1880-2014 (GISTEMP team, 2015, Hansen *et al*, 2010). The trend line has a slope of .605 °C/century. (b) The periodogram, $\{\hat{P}_k\}$, of this time series plotted using logarithmic scales. The dashed lines shows a power law component with a slope of -2 and a horizontal unselfcorrelated component. (c) The periodogram of the time series of first differences.

component, a random variable, ξ_i. The value of the variable, χ_i, at time t_i is given by

$$\chi_i = c_1 t_i + c_0 + \xi_i \tag{23}$$

where t_i is the time, c_1 and c_0 are constants estimated from the data and the $\{\xi_i\}$ are independent and identically distributed (i.i.d.) random variables with zero mean. The constants of the linear regression model, (23), were estimated as $\hat{c}_1 = 0.00605$ and $\hat{c}_0 = -11.78$. The correlation coefficient, r, was 0.87. The regression residuals, $\{\hat{\xi}_i\}$, were found by subtracting the regression line values from the original data and are shown in Figure 4a.

The ARIMA model used was a (0,1,1) model, viz.:

$$\chi_i = \chi_{i-1} + \epsilon_i + b_1 \epsilon_{i-1} \tag{24}$$

where $\hat{b}_1 = -0.611$ was estimated from the data and the random variables $\{\epsilon_i\}$ were assumed to be i.i.d. with zero mean.

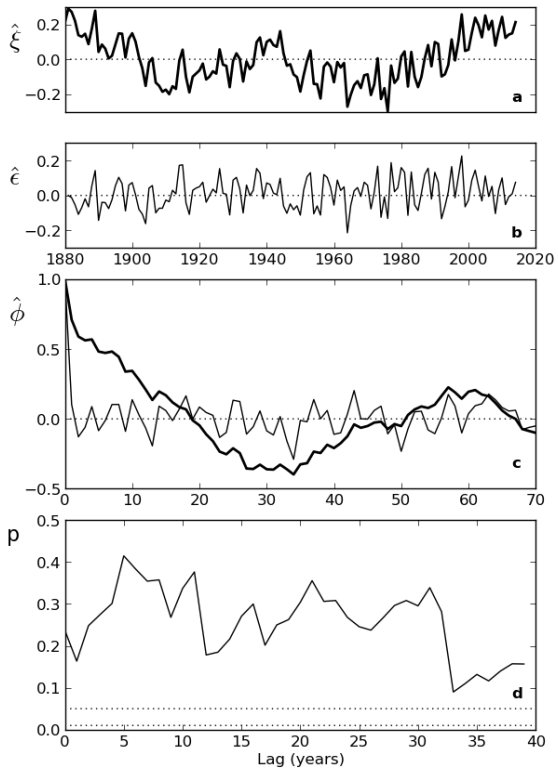

Fig. 4. (a) Time series of the regression-model residuals, $\hat{\xi}$. (b) Time series of the ARIMA-model residuals, $\hat{\epsilon}$. (c) Sample autocorrelation functions of $\hat{\xi}$ (thick line) and $\hat{\epsilon}$ (thin line). (d) Ljung-Box probability value, p, for the ARIMA-model residuals, $\hat{\epsilon}$. The regression-model residual probabilities are too small to be distinguished from the horizontal axis.

The residuals from the ARIMA model, $\{\hat{\epsilon}_i\}$, were found by rearranging (24), setting $\hat{\epsilon}_1 = 0$ and substituting sample values x_i for random variables X_i, viz.:

$$\hat{\epsilon}_i = x_i - x_{i-1} - \hat{b}_1\hat{\epsilon}_{i-1} \tag{25}$$

The resulting time series of ARIMA-model residuals, $\{\hat{\epsilon}_i\}$, were found recursively in this way and are shown in Figure 4b.

The sample autocorrelation functions, $\hat{\phi}$, of the two sequences of residuals are shown in Figure 4c. The autocorrelation function of the regression-model residuals, $\{\hat{\xi}_i\}$, is shown by the thick line. They are clearly self-correlated. The autocorrelation function of the ARIMA-model residuals, $\{\hat{\epsilon}_i\}$, is shown as the thin line in Figure 4c. These may well be unselfcorrelated since the values of $\hat{\phi}$ at lags greater than zero are comparatively small.

The Ljung-Box test is a strong test for self-correlation. It returns a set of probabilities that a sequence of residuals is unselfcorrelated. When this test was applied to the sequence of regression-model residuals, $\{\hat{\xi}_i\}$, it returned probabilities in the range 2.3×10^{-16} to 1.3×10^{-91}. Thus the null hypothesis that the linear regression residuals are unselfcorrelated can be rejected at a very high level of significance and the linear regression model itself must be rejected. The Ljung-Box test was applied to the sequence of ARIMA-model residuals, $\{\hat{\epsilon}_i\}$, derived using (25) and probabilities, p, were returned to a maximum lag of 40. These all lay in the range .089 to .415 and are plotted in Figure 4d. None of the probabilities lay below .05 so there was no reason to reject the null hypothesis that this sequence is unselfcorrelated.

5.2. Discussion

It follows that the ARIMA-model of equation (24) accounts for the observed time series, but the regression-model of equation (23) does not, because its residuals are highly self-correlated. Such self-correlation is commonly accounted for as a consequence of 'natural oscillations' such as the 'Pacific Decadal Oscillation' or the 'Atlantic Multidecadal Oscillation'. The fact that the stochastic ARIMA model satisfies the Ljung-Box test implies that, not only was there no linear trend or drift present in the data, but neither were there any deterministic cycles or oscillations. Apart from an added white noise component, the GISS global average temperature time series is a random walk.

In order for the ARIMA model of equation (24) be accepted, it was a necessary condition that the residuals be unselfcorrelated as proved to be the case. In a sense, it is also a sufficient condition because once the statistical character of the observations has been completely accounted for, then, by Occam's Razor, there is no point in making further assumptions.

Note that the assumption that the given time-series is a pure random-walk is not necessary. It only needs to be random-walk-like and to have an integrating tendency in order that spurious regression occur. For example an ARMA(1,1) model works equally well, wherein equation (24) becomes

$$X_i = a_1X_{i-1} + \epsilon_i + b_1\epsilon_{i-1} \tag{26}$$

and a_1 is close to unity. In practice a_1 could be as small as 0.985 without affecting the whiteness of the residuals. It follows that testing whether the time series is a 'true' random walk using, say, the Dickey-Fuller test (Dickey and Fuller, 1979) is irrelevant; what matters is that residuals of the stochastic processes (24) or (26) are unselfcorrelated and those of the regression model (23) are not.

Similar reasoning applies to deterministic cycles. A deterministic cycle such as that described by (19) corresponds to a pole on the unit circle in the z-transform of the sequence. For a sequence of finite length it is impossible to determine whether such a pole lies $exactly$ on the unit circle. The relevant peak may instead be the outcome of an AR(2) process with a pole very close to the unit circle, i.e. physically, to a resonance with

a very narrow bandwidth or 'high Q'. Such a stochastic resonance has the same effect as a truly deterministic cycle in that its influence will be felt well beyond the time interval under examination. It will have predictive power.

6. Conclusions

Researchers may be unaware of the random walk character of much of their data and the implications of spurious regression and spurious cycles that it entails.

Statistical inference is a means by which the principles and mechanisms underlying the natural world may be understood; it is the scientific method in numerical form. Theories which do not fit observations, models which do not fit data, are rejected. The testing of spectral peaks for significance and the testing of residuals for self-correlation are both part of this process. The periodogram avoids the trap of mistaking random excursions for real phenomena.

References

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle, in Petrov, B.N.; Cski, F., 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadmiai Kiad, p. 267-281.

Cheng, H., Edwards,R. L., Broecker, W. S., Denton,G. H., Kong, X., Wang, Y., Zhang, R. and X. Wang, 2009. Ice Age Terminations. *Science* **326**, 248-252

Dickey, D. A. and Fuller, W. A. 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* **74** 427431.

GISTEMP Team. 2015. GISS Surface Temperature Analysis (GISTEMP)., *NASA Goddard Institute for Space Studies.* Dataset downloaded 2015-01-21 at `http://data.giss.nasa.gov/gistemp/`.

Granger, C.W.J. and Newbold, P. 1974. Spurious regressions in econometrics. *J. Econometrics* **2**, 111-120.

Hansen, J., R. Ruedy, M. Sato, and K. 2010. Global surface temperature change. *Rev. Geophys.*, **48**, RG4004.

Hays, D.J., Imbrie, J., Shackleton, N.J., 1976. Variation in the Earths Orbit: Pacemaker of the Ice Ages. *Science*, **194**, 4270, 1121-1132.

Hasselmann, K. 1976. Stochastic climate models. Part I. Theory. *Tellus* **XXVIII**, 6, 473-485.

Huybers, P., Wunsch, C., 2005. Obliquity pacing of the late Pleistocene glacial terminations. *Nature* **434**, 491-494.

Jouzel, J. *et al*, 2007. EPICA Dome C Ice Core 800KYr Deuterium Data and Temperature Estimates. IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series #2007-091. NOAA/NCDC Paleoclimatology Program, Boulder CO, USA. EPICA Dome C Ice Core 800KYr Deuterium Data and Temperature Estimates `http://doi.pangaea.de/10.1594/PANGAEA.683655?format=html#download`

Lisiecki, L.E., 2010. Links between eccentricity forcing and the 100,000-year glacial cycle., *Nature Geoscience* **3**, 349 352.

Ljung, G. M. and Box, G. E. P. 1978. On a Measure of a Lack of Fit in Time Series Models., *Biometrika* **65**, 297-303 .

Milankovic, M., 1941. *Kanon der Erdbestrahlung und seine Anwendung auf das Eiszeitenproblem.*, Koniglich Serbische Akademie, Belgrade.

Muller, R.A., MacDonald, G.J., 1997. Spectrum of 100-kyr glacial cycle: Orbital inclination, not eccentricity. *Proc. National Acad. Sciences USA* **94**, 8329-8334.

Pelletier, J.D. 2002. Natural variability of atmospheric temperatures and geomagnetic intensity over a wide range of time scales., *Proc. National Acad. Sciences USA* **99 suppl. 1**, 2546-2554.

Reid, J.S. 1979. Confidence Limits and Maximum Entropy Spectra.,*J.Geophys.Res.* **84**, 5289-5301.

Ridgwell, A.J., Watson, A.J., Raymo, M.E., 1999. Is the spectral signature of the 100 kyr glacial cycle consistent with a Milankovitch origin? *Paleoceanography*, **14**, 437-440.

APPENDIX A: Periodogram of a cosine function

In the absence of asymptotic considerations, the periodogram of a finite time series cannot exhibit a delta function. Nevertheless the periodogram of a sinusoid will exhibit a narrow peak. Of particular interest is the shape that this peak takes when the sinusoid has a frequency which is not an integral multiple of the frequency resolution, i.e. when it does not occur at one of the DFT frequencies.

Consider a time series $\{c_j, j = 0, ..., N-1\}$, defined explicitly by a cosine function with arbitrary frequency, f_0, and phase, ϕ:

$$c_j = A\cos(2\pi j f_0/N + \phi) = x_j + y_j \tag{A1}$$

where

$$x_j = Ae^{2\pi i j f_0/N + \phi i)}/2 \tag{A2}$$

and

$$y_j = Ae^{-2\pi i j f_0/N + \phi i)}/2 \tag{A3}$$

Substituting x_j into (2) gives

$$X_k = \frac{A}{2N}\sum_{j=0}^{N-1} e^{-2\pi i j(k-f_0)/N)+\phi i} \tag{A4}$$

After multiplying both sides by $e^{-2\pi i(k-f_0)/N}$

$$X_k e^{-2\pi i(k-f_0)/N} = \frac{A}{2N}\sum_{j=0}^{N-1} e^{-2\pi i(j+1)(k-f_0)/N+\phi i} \tag{A5}$$

Subtracting (A5) from (A4) and transposing gives

$$X_k = \frac{Ae^{\phi i}\left[1 - e^{2\pi i(k-f_0)}\right]}{2N\left[1 - e^{2\pi i(k-f_0)/N}\right]} \tag{A6}$$

$$= \frac{Ae^{\pi i(k-f_0)+\phi i}\left[e^{-\pi i(k-f_0)} - e^{\pi i(k-f_0)}\right]}{2Ne^{\pi i(k-f_0)/N}\left[e^{-\pi i(k-f_0)/N} - e^{\pi i(k-f_0)/N}\right]} \tag{A7}$$

i.e.

$$X_k = \frac{Ae^{\pi i(k-f_0)+\phi i}\sin\{-\pi(k-f_0)\}}{2Ne^{\pi i(k-f_0)/N}\sin\{-\pi(k-f_0)/N\}} \tag{A8}$$

Likewise, $\{Y_k\}$, the DFT of $\{y_j\}$ is given by

$$Y_k = \frac{Ae^{\pi i(k+f_0)+\phi i}\sin\{-\pi(k+f_0)\}}{2Ne^{\pi i(k+f_0)/N}\sin\{-\pi(k+f_0)/N\}} \quad \text{(A9)}$$

where, from (A1) the squared modulus of the DFT of $\{c_j\}$ is given by

$$|\ C_k\ |^2 = X_kX_k^* + X_kY_k^* + X_k^*Y_k + X_k^*Y_k^* \quad \text{(A10)}$$

From (A8) and (A9)

$$X_kX_k^* = \frac{A^2\sin^2\{\pi(k-f_0)\}}{4N^2\sin^2\{\pi(k-f_0)/N\}} \quad \text{(A11)}$$

$$X_kY_k^* = \frac{A^2e^{2\pi if_0+2\phi i}\sin\{\pi(k+f_0)\}\sin\{\pi(k-f_0)\}}{4N^2e^{2\pi if_0/N}\sin\{\pi(k+f_0)/N\}\sin\{\pi(k-f_0)/N\}} \quad \text{(A12)}$$

$$X_k^*Y_k = \frac{A^2e^{-2\pi if_0-2\phi i}\sin\{\pi(k+f_0)\}\sin\{\pi(k-f_0)\}}{4N^2e^{-2\pi if_0/N}\sin\{\pi(k+f_0)/N\}\sin\{\pi(k-f_0)/N\}} \quad \text{(A13)}$$

and

$$Y_kY_k^* = \frac{A^2\sin^2\{\pi(k+f_0)\}}{4N^2\sin^2\{\pi(k+f_0)/N\}} \quad \text{(A14)}$$

Adding the cross-modulation terms gives an error term E_k,

$$E_k = X_kY_k^* + X_k^*Y_k \quad \text{(A15)}$$

Thus

$$E_k = \frac{A^2\cos\{2\pi(1-1/N)f_0+2\phi\}\sin\{\pi(k+f_0)\}\sin\{\pi(k-f_0)\}}{2N^2\sin\{\pi(k+f_0)/N\}\sin\{\pi(k-f_0)/N\}} \quad \text{(A16)}$$

Substituting (3) in (A9) gives

$$Y_kY_k^* = X_kX_k^* \quad \text{(A17)}$$

and the periodogram of a cosine function of arbitrary frequency and phase becomes:

$$Q_k = \frac{A^2\sin^2\{\pi(k-f_0)\}}{2N^2\sin^2\{\pi(k-f_0)/N\}} + E_k \quad \text{(A18)}$$

Note that the error term, E_k, is dependent on the phase, ϕ. Its absolute value $|E_k|$ has a maximum value of $0.025A^2$ when $f_0 \approx 1$ and $k = 1$ and is bounded above by $|\ 1.5A^2/2N\sin(2\pi k/N)\ |$.

APPENDIX B: Periodogram of a random walk

Let the sequence of random variables, $\{\zeta_n, n = 1, ..., N\}$, be the outcome of a random walk process, i.e.

$$\zeta_n = \zeta_{n-1} + \xi_n \quad \text{(B1)}$$

where the sequence $\{\xi_n, n = 0, ..., N\}$ is the outcome of a zero mean white noise process. Then

$$\xi_n = \zeta_n - \zeta_{n'} \quad \text{(B2)}$$

where $n' = n - 1$ and taking Fourier transforms of both sides gives

$$\Xi_j = Z_j - Z'_j \quad \text{(B3)}$$

where $\{Z_j\}$ and $\{Z'_j\}$ are the periodograms of $\{\zeta_n\}$ and $\{\zeta_{n'}\}$ respectively, i.e.

$$Z_j = \sum_{n=0}^{N-1}\zeta_n e^{-2\pi ijn/N} \quad \text{(B4)}$$

and

$$Z'_j = \sum_{n'=0}^{N-1}\zeta_{n'}e^{-2\pi ijn'/N} = \sum_{n=1}^{N}\zeta_{n-1}e^{-2\pi ij(n-1)/N} \quad \text{(B5)}$$

However

$$Z_j = \zeta_0 + e^{-2\pi ij/N}\sum_{n=1}^{N}\zeta_n e^{-2\pi ij(n-1)/N} - \zeta_N \quad \text{(B6)}$$

$$= e^{-2\pi ij/N}Z'_j + \delta \quad \text{(B7)}$$

where $\delta = \zeta_N - \zeta_0$. Substituting into (B3) gives

$$Z_j = \frac{\Xi_j - \delta e^{2\pi ij/N}}{1 - e^{2\pi ij/N}} \quad \text{(B8)}$$

We now substitute Z_k for Ξ_k in (8) to give the ensemble periodogram of $\{\zeta_n\}$, viz.:

$$\Pi_k = 2\Big[\frac{Z_kZ_k^*}{N}\Big]^2 \ , \quad k = 1...N/2 - 1 \quad \text{(B9)}$$

Given that $\mathrm{E}(\delta) = 0$ and substituting (B9) and (11) into (10), then the mean ensemble periodogram of $\{\zeta_n\}$ is given by

$$P_k = \frac{\sigma^2 + N\,\mathrm{E}(\delta^2)}{N\sin^2(\pi k/N)} \approx Ak^{-2} \quad \text{(B10)}$$

where σ^2 is the variance of the white noise process, $\{\xi_n\}$, and A is a constant. The approximation is very good when $k << N$, i.e. at the low frequency end of the periodogram. The $\mathrm{E}(\delta^2)$ term in the denominator accounts for the time series, $\{\zeta_n\}$, being non-stationary.