Coherence established between atmospheric carbon dioxide and global temperature

Cynthia Kuo, Craig Lindberg & David J. Thomson

Mathematical Sciences Research Center, AT&T Bell Labs, Murray Hill, New Jersey 07974, USA

The hypothesis that the increase in atmospheric carbon dioxide is related to observable changes in the climate is tested using modern methods of time-series analysis. The results confirm that average global temperature is increasing, and that temperature and atmospheric carbon dioxide are significantly correlated over the past thirty years. Changes in carbon dioxide content lag those in temperature by five months.

DURING the past century (see, for example, refs 1, 2), scientists have studied the possibility that the climate is influenced by changes in the atmospheric concentration of CO_2 caused by industrial and agricultural activities[3–5]. Recently, because of the potentially serious consequences of the greenhouse effect[6,7], the problem has been studied more intensively[8,9] with a view towards observing climatic effects attributable to the increase of atmospheric greenhouse gases. For example, the output of numerical models of the global climate has been compared with measurements[10,11], and the significance of the temperature increase has been tested using various straight-line segment and parametric models (ref. 12 and J. Seater, manuscript in preparation). But present numerical models of the atmosphere are crude, and comparisons between the time series representing the real data and predictions of the atmospheric models are difficult to interpret. Because the available data are short time series, conventional statistical methods are unreliable, and detection of the greenhouse effect remains controversial[13–18].

The longest modern series of precise CO_2 concentration measurements begins in March 1958 and consists of monthly values collected by Keeling at the summit of Mauna Loa in Hawaii[19]. Although these data are from a single station, they are typical of measurements made since 1974 at several sites[20]. Because we are interested in the time-series aspects of the problem, we use the longer Keeling series. The data have an upward trend that is readily visible (the upper curve in Fig. 1), an obvious annual component and irregular fluctuations. Five missing values were interpolated using a stochastic least-squares procedure on the residuals from a quadratic polynomial plus the first five annual harmonics. These interpolated values have an estimated error of 0.35 p.p.m. and are noted in Fig. 1.

Hansen and Lebedeff[21] created a time series of monthly global-average surface-air-temperature changes from January 1880 to December 1988. This series is a weighted average over stations, formed by subtracting the average January temperature during the reference period 1951–80 from all the January data, and repeating this for the other months, eliminating seasonal variations. Displayed as the lower curve in Fig. 1, the temperature series also increases with time but its fluctuations are relatively larger than those in the CO_2 record.

Here we apply multiple-window time-series methods (which are efficient for short series) to estimate the trends and power spectra of the Hansen–Lebedeff average global surface temperature series and Keeling CO_2 concentration measurements as well as the coherence between the two. This analysis shows that from 1880 to 1988, the average global temperature increased by 0.0055 ± 0.00096 °C yr^{-1}, and the probability that this slope is positive exceeds 99.99%. Furthermore, the monthly CO_2 concentration and global temperature series from 1958 to 1988 are coherent over much of the Nyquist frequency band from 0 to 6 cycle yr^{-1}; the probability that the level of coherence observed from 0 to 2 cycle yr^{-1} occurred by chance is $\sim 2 \times 10^{-6}$. Not only do both series have increasing trends that are highly significant, but there are linear relations between many of their oscillatory components. We interpret this as evidence that the changes in atmospheric CO_2 concentration are closely related to changes in global temperature.

Models

Many of the contradictions in the literature about the analysis of climate data can be traced to the use of inappropriate or oversimplified models (for review, see, for example, refs 22, 23). Statistical techniques that are vulnerable to difficulties include those based on parametric models (such as low-order autoregressive moving-average representations) and methods plagued by more insidious problems caused by implicit assumptions of time-series stationarity. The cavalier use of parametric models can lead to misspecification difficulties[24] because

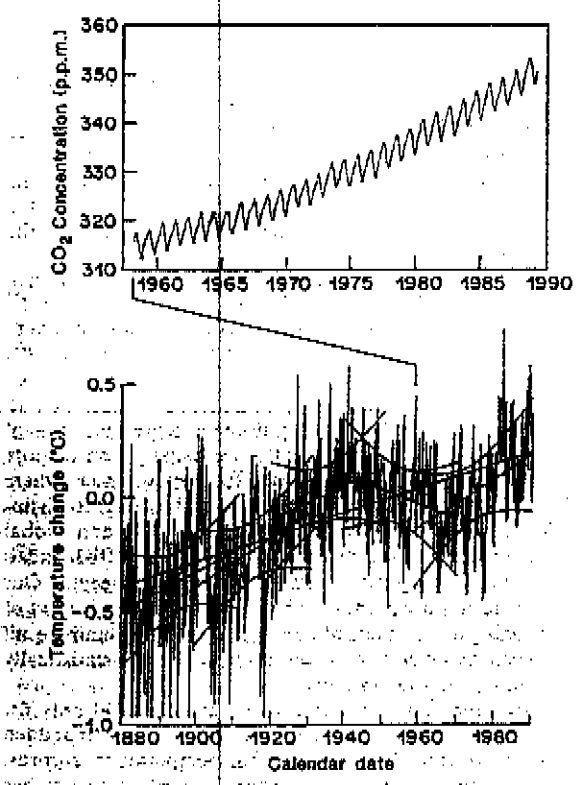

FIG. 1 The upper curve is the monthly Keeling CO_2 concentration data for March 1958 to December 1988. The five interpolated points are marked. The lower curve is the Hansen–Lebedeff average global temperature series for 1880 to 1988. The short line segments and hyperbolic arcs define trends and 95% confidence regions over 30-yr intervals, and the long straight line shows the general trend.

unanticipated effects (such as the modulation of the annual cycle discussed below) either are undetected, are attributed incorrectly or are simply lumped in as part of 'residual variance'. As shown below, the CO_2 concentration and global temperature series have complicated spectral density functions, so these series are not likely to be modelled by any simple form in either the time or frequency domains. The fundamental flaw of parametric models, including numerical atmospheric models and time-series models, is that they make no provision for the unexpected.

We represent each time series by the less restrictive decomposition into a parametric trend and a non-parametric residual. Because one cannot directly discriminate between the many possible functional forms for the trends from such short records, we represent the trend in each time series by the simplest non-constant function of time, a straight line of non-zero slope which can be thought of as the first two terms in the Taylor series expansion of the true trend. It may be argued that this semi-parametric representation is subject to the misspecification problems mentioned above, but, because we subject the residuals to greater scrutiny than we do the trends, features not included in the trends will still be present and studied in the residuals. Moreover, we find that including quadratic terms does not change the results of the following analysis significantly.

Therefore, we express each time series

$$\{x(0), x(1), x(2), \ldots, x(N-1)\}$$

as the sum of a constant, a linear trend and a stationary time series specified only by its spectral representation

$$x(t) = a_0 + a_1(t - t_{ref}) + e(t); \quad t = 0, 1, \ldots, N-1 \qquad (1)$$

where t_{ref} is a reference time, a_0 and a_1 are constants and the residual time series $e(t)$ has the spectral representation

$$e(t) = \int_{-1/2}^{1/2} e^{i2\pi ft}\, dX(f) \qquad (2)$$

for all t. $dX(f)$ is the differential of a generalized Fourier transform and is known as an 'orthogonal increment process'. (This is an extension of the representation used in ref. 25, where the CO_2 series was decomposed into a trend, an annual component and a residual.) The annual component is given by the first moment of $dX(f)$, and the power spectrum, or power spectral density, $S(f)$, of the residuals is, by definition, the second moment of $dX(f)$:

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

where E is the statistical expected-value operator. In a stationary series, values of $dX(f)$ at distinct frequencies are uncorrelated[26].

Estimation

Although the time- and frequency-domain representations of time series are formally equivalent[27], we usually find it more informative to analyse time series in the frequency domain where the effects of different physical processes can be easier to distinguish. With short time series, such as the CO_2 and global temperature records, it is difficult to resolve different frequencies and simultaneously obtain statistically significant results. Our approach is to use a variant of the multiple-window method of spectrum analysis[28,29] that, although it does not eliminate all the problems associated with short series, makes statistically efficient use of the available data.

Most frequency-domain methods are strictly valid only for the analysis of stationary data, not series with embedded trends such as the CO_2 and global temperature records. Using the multiple window procedure described below, we estimate the trends in each series, subtract off these terms, and estimate the spectrum of the residuals. Finally, we test the residuals for stationarity to see if our assumptions were violated.

Estimating the average a_0 and trend a_1 by ordinary least squares can produce misleading results if the residuals $\{e(t)\}_{t=0}^{N-1}$ have a non-white (non-flat) spectrum and are therefore correlated[30,31]. The residual spectrum need only be roughly flat in a small frequency interval around the origin for multiple-window regression, however, to produce valid estimates of a_0, a_1, the spectrum of the residuals and their errors.

Multiple windows

The multiple-window method uses the orthogonal sequences of N elements that optimally concentrate the spectral energy in a frequency band of width $2B$ centred on a frequency f, that is between the frequencies $f-B$ and $f+B$. These sequences are the lowest-order $[2BT]$ discrete prolate spheroidal sequences of ref. 32, where T is the duration of the observed series. These 'Slepian sequences' form a basis on which the data in the

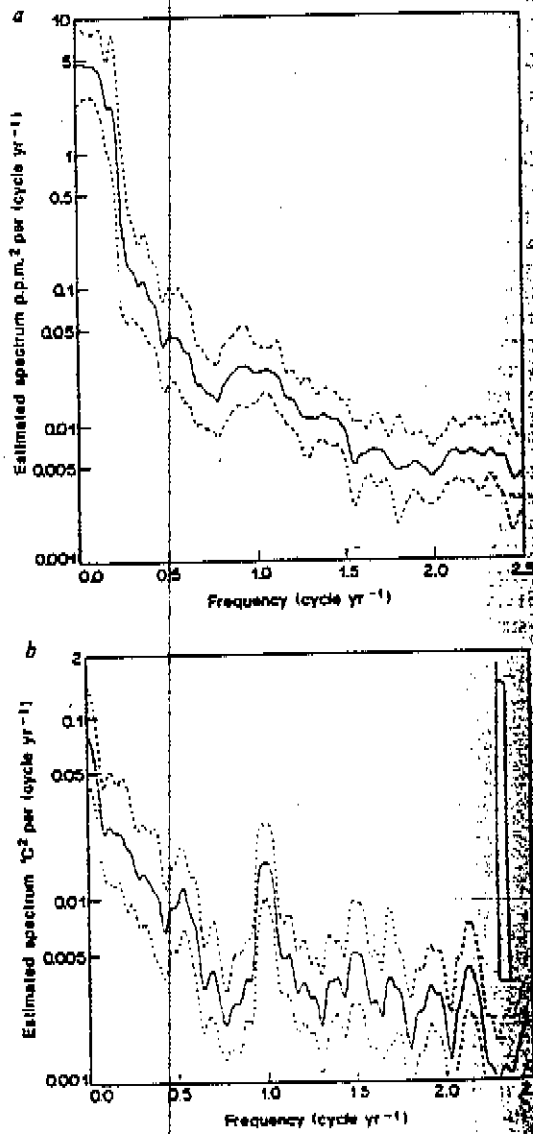

FIG. 2 a Multiple-window spectral estimate of trend-subtracted CO concentration series. The upper and lower dashed curves are 5 and confidence limits obtained by jackknifing on the windows. Harmonics annual cycle have been subtracted. b Multiple-window spectral estim detrended global temperature series for the interval 1880 to 198 dashed lines give the 5 and 95% confidence intervals as in a. The on the upper right shows the effective spectral window on the amplitude and frequency scale. Note that the power around 1 cycle not a periodic component. (Spectra done at higher resolution show feature is asymmetric with more power on the lower sideband tha upper, indicating combined amplitude and phase modulation of the cycle.)

frequency band can be expanded. The coefficients of this expansion depend on frequency and are obtained by taking the discrete Fourier transform of the product of the data with each Slepian sequence. The lowest-order coefficient is similar to a direct spectrum estimate using a conventional data window or taper. A multiple-window estimate of the spectrum is, however, an adaptively weighted average of the $2BT$ frequency-dependent coefficients.

Because of their energy-concentration properties, the Slepian sequences are the data windows that are most resistant to spectral leakage[28]. For example, the estimate of the temperature-series spectrum in Fig. 2*b* is produced using 11 Slepian sequences with $2B = 0.11$ cycle yr^{-1}; sidelobes of the effective spectral window shown in the insert are unobservable on the scale of Fig. 2*b*.

Choosing the parameter B for a multiple-window estimate of a spectrum involves a tradeoff between resolution and variance. The variance of this spectrum estimate is proportional to $1/(2BT)$. Thus, increasing B improves statistical reliability but decreases the resolution of the estimate.

Trend estimates

To estimate the coefficients a_0 and a_1 in equation (1) by conventional time-domain regression, one minimizes the sum of the squares of the residuals $\sum_{t=0}^{N-1} e^2(t)$ with respect to the coefficients. By Parseval's formula,

$$\sum_{t=0}^{N-1} e^2(t) = \int_{-1/2}^{1/2} |\tilde{e}(f)|^2 \, df \qquad (4)$$

where $\tilde{e}(f)$ is the discrete Fourier transform of the residuals $\{e(t)\}_{t=0}^{N-1}$, so minimization in either the time or frequency domain is equivalent. To avoid the energy associated with the harmonics of the annual cycle and other high-frequency processes, however, we minimize only the integral over the low frequencies

$$\int_{-W}^{W} |\tilde{e}(f)|^2 \, df; \quad W = BT/N < \tfrac{1}{2} \qquad (5)$$

and ignore the higher-frequency components of the residuals, resulting in estimates expressible in terms of the Slepian sequences.

The multiple-window approach has several advantages when the residuals are autocorrelated. First, the 'observations' in the multiple-window regression are closer to independent gaussian random variables than the original time-domain data $\{x(t)\}_{t=0}^{N-1}$, and therefore the multiple-window coefficient estimates are closer to maximum-likelihood estimates than are the estimates from ordinary least-squares analysis. Similarly, the number of degrees of freedom and error estimates are easily calculated using the Slepian sequences, and the multiple-window method is often more statistically efficient than ordinary least-squares analysis (C. Lindberg, manuscript in preparation). In conventional regression, the endpoints of the time series (for example, the abnormally high temperatures of the last decade) can have an inordinately large influence on the coefficient estimates[33]. This effect is largely eliminated in this approach. Finally, it is relatively easy to check the assumptions that have been made in representing the data by a particular model.

Trends

For the monthly Keeling CO_2 series (March 1958 to December 1988, $T = 29.8$ yr), estimates of the average and trend were obtained using $t_{ref} = 1975.0$, and a bandwidth of $2B = 0.39$ cycle yr^{-1}. This value of B confines the spectral energy associated with the trend components to frequencies f with $|f| \leq 0.195$ cycle yr^{-1}, avoiding energy associated with the annual cycle and its harmonics. This value of B also results in a spectrum that is locally white in the resolution band (at frequencies above 0.2 cycle yr^{-1}, the CO_2 residual spectrum drops rapidly, as shown below). The time–bandwidth product $2BT = 11.6$ gives eleven leakage-resistant windows. The multiple-window procedure results in the estimates $a_0 = 331.9 \pm 0.44$ p.p.m. and $a_1 = 1.191 \pm 0.053$ p.p.m. yr^{-1}.

For the monthly Hansen–Lebedeff global temperature series from January 1880 to December 1988, we obtained estimates of the average and trend using $t_{ref} = 1934.5$ and a resolution bandwidth of $2B = 0.128$ cycle yr^{-1}. This value of $2B$ was chosen to avoid the frequency band above 0.07 cycle yr^{-1} where the residual global temperature spectrum decreases rapidly, as shown below. Using 12 Slepian sequences, we obtain $a_0 = -0.106 \pm 0.030$ °C and $a_1 = 0.00554 \pm 0.00096$ °C yr^{-1}.

For comparison, estimates from ordinary least-squares analysis agree with the multiple-window estimates to three significant figures but, because of the lower values of the spectrum at higher frequencies, underestimate their standard deviations by a factor of five.

To assess the significance of this estimated global temperature slope, we note that Milankovitch theory predicts that at present the Earth should be cooling by -0.0004 °C yr^{-1} (refs 34, 35). Thus, solar variability aside, the null hypothesis is that the temperature trend should be slightly negative. To test this hypothesis we use a t statistic $t = [0.00554 - (-4 \times 10^{-4})]/0.000961 = 6.18$ which, as 12 windows were used and two parameters were estimated, is characterized by approximately 10 degrees of freedom. Therefore, given that the low-frequency spectrum is approximately white (see Fig. 2*b*), the slope is greater than the Milankovitch prediction with probability 99.995% (ref. 36). Using a narrower bandwidth leads to fewer Slepian sequences with resistance to spectral leakage, fewer degrees of freedom and an underestimate of the slope significance. A wider bandwidth results in an invalid t statistic, as the residual spectrum decreases rapidly at frequencies > 0.07 cycle yr^{-1} and so violates the 'locally white' assumption. Neither a jackknife variance estimate[37] (a non-parametric statistic sensitive to both non-gaussianity and non-stationarity, determined from the set of spectrum estimates computed with each of the windows deleted in turn), nor the stationarity test of ref. 38 show the residual series to be non-stationary. High-resolution quadratic inverse spectrum estimates provide some evidence for a ripple on the spectrum at frequencies < 0.07 cycle yr^{-1}, consistent with a 'recurrence time' in the temperature data of ~ 28 yr, but the amplitude of this ripple is not enough to change the significance of the slope.

Finally, it has been argued that the negative trend in the temperature record between 1940 and 1970 invalidates the conclusion that the temperature is increasing over the long term. To test this, we repeated the trend calculations for overlapping 30-year subsections of the Hansen–Lebedeff series. The estimated trends for each interval are shown as the short straight lines through the temperature record in Fig. 1, and the 95%-confidence region of each is bounded by the hyperbolic arcs[39]. The line associated with the linear trend of the entire temperature record remains in the corridor collectively outlined by the subsection error bounds.

Spectrum estimates

We estimate the spectrum of the residual series $\{e(t)\}_{t=0}^{N-1}$ by subtracting the periodic components (detected by a statistical F-test[28]) from the detrended data, and by making an adaptively weighted multiple-window estimate of the power spectral density $S(f)$ of the residuals[28]. No frequency components above 2.5 cycle yr^{-1} are shown to avoid artefacts from the unequal lengths of the months.

Figure 2*a* is the low-frequency part of the spectrum estimate of the CO_2 residuals using $2B = 0.39$ cycle yr^{-1} and 11 windows. The spectrum is not white, so estimates using ordinary least-squares analysis of the CO_2 trend error bounds are invalid. The variance of the estimate has been calculated by jackknifing over windows[37,40,41], and the resulting 5% and 95% confidence limits are shown. These error bounds are consistent with the

χ^2-distributed estimate (with 22 degrees of freedom) expected from stationary gaussian data. Because of the shortness of this series, it is unlikely that details in the low-frequency end of the spectrum have been resolved; the plotted spectrum is a compromise between frequency resolution and statistical significance.

Figure 2*b* shows an estimated spectrum of the Hansen–Lebedeff global average temperature residual series. To allow resolution of more details in the spectrum, a bandwidth of $2B = 0.11$ cycle yr^{-1} was used, resulting in 11 Slepian sequences

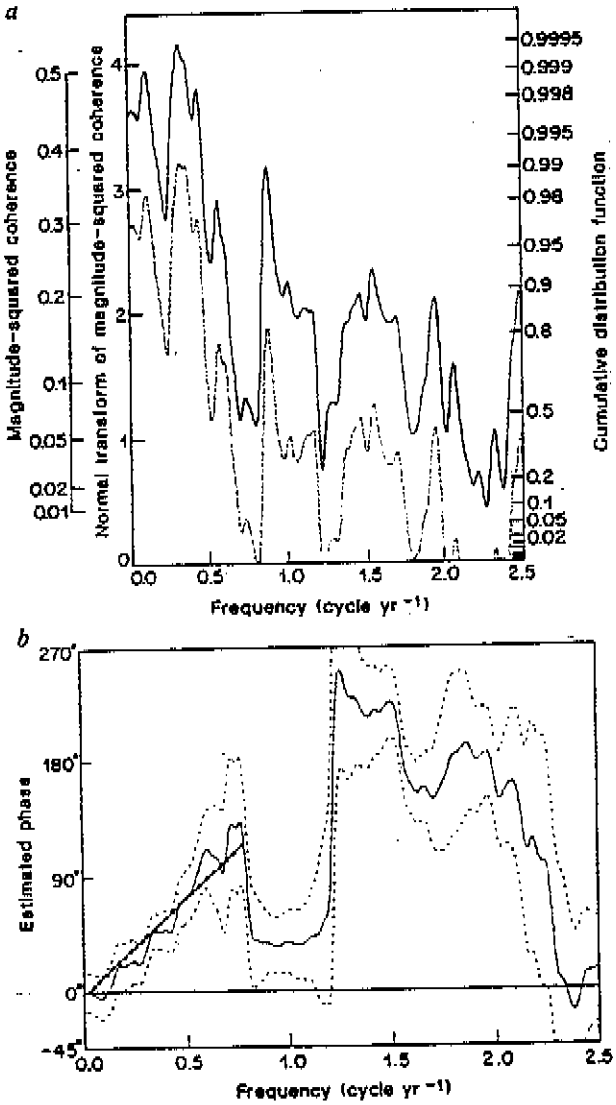

FIG. 3 *a*, Magnitude-squared coherence (MSC) between the Keeling CO_2 series and the global temperature series during 1958–1988. The vertical axis on the far left is the value of the MSC, and the axis on the left of the graph is the value of the coherence magnitude transformed by $\tanh^{-1}$. On this scale the coherence estimates should be roughly gaussian with unit standard deviation[48]. The vertical axis on the right gives the cumulative distribution function for incoherent series. The upper (solid) curve represents the transformed coherence values ($\tanh^{-1}(|C(f)|)$), and the lower (dashed) curve is offset by one standard deviation as determined by jackknifing over windows. *b*, Phase of coherence between the Keeling CO_2 series and the global temperature series during 1958–1988 corresponding to the MSC in *a*. The two dashed lines show the ±1 s.d. limits obtained by jackknifing. The low-frequency trend in the phase corresponds to a delay of −5 months, whereas the 'hole' near 1 cycle yr^{-1} reflects the predominance of other effects near this frequency. The poorer limits on the phase at higher frequencies is a consequence of the lower coherences there. These jackknife estimates agree with gaussian theory.

($2BT = 108 \times 0.11 = 11.88$). As noted in ref. 21, this spectrum exhibits substantial power at periods near integer multiples of the annual cycle. Because the *F*-test shows that this power is not a simple periodic signal in the data (recall that the Hansen–Lebedeff referencing procedure has already subtracted the annual cycle) we examine it further below.

Relations between the series

That both the CO_2 and global temperature data have positive slopes does not prove that the two series are related. As the discussion on spurious correlation in ref. 42 makes clear, the presence of the trends in each series will cause simple time-domain correlations between the two series to be high. If it is found, however, that the fluctuations of the two detrended series are coherent over a band of frequencies, then it is more likely that the two series are related.

The frequency-domain analogue of correlation, coherence[43–45], has been applied to meteorological data for many years[46]. Conventional estimates of coherence between single sections of two records are the smoothed (by a moving average) complex product of the discrete Fourier transforms of the two series, and so can be badly biased if the phase changes over the averaging band. Section-averaging methods are inappropriate for short series; not only does dividing these series into subsections result in poor frequency resolution, but the correlation between the subsections produces unreliable coherence estimates. The multiple-window approach described in refs 28 and 41 provides a less biased estimate of coherence $C(f)$ which is suitable for short series, allowing the extraction of more information from the same data.

Figure 3*a* shows the multiple-window magnitude-squared coherence between the Keeling CO_2 and global temperature residuals from 1958 to 1988 (produced using the same bandwidth and number of windows as the spectrum estimate of the CO_2 residuals). It is remarkable that the two series have a coherence above the 90% confidence level at frequencies <0.6 cycle yr^{-1}, with coherence exceeding 98% over much of this low-frequency band. Because multiple-window coherence estimates spaced $2B$ apart are essentially independent, the probability of observing such levels across a wide band by chance from independent series is very low, $\sim 2 \times 10^{-6}$. Thus, not only are the trend components (at frequencies ~ 0 cycle yr^{-1}) of both time series increasing, but the residuals of the two series are also coherent with high confidence in the low-frequency band.

Figure 3*b* is a plot of the phase of the multiple-window coherence between the two residual (trend-subtracted) series. The phase of the coherence at 0 cycle yr^{-1} is zero and, because both trends are positive, is independent of whether trends are included or not. Between 0 and 0.8 cycle yr^{-1} the phase is roughly linear, corresponding to the CO_2 series lagging the temperature series by −5 months (calculated by taking the slope of this linear section of the phase[44], $-120°/(360° \times 0.8$ cycle yr$^{-1})$) in agreement with arguments that natural positive feedback mechanisms in the carbon cycle causes carbon dioxide to lag temperature in some frequency bands[19] (R. Marston, personal communication). Current knowledge of these complicated interactions involving solar forcing, the Southern Oscillation and exchange of CO_2 with the oceans on various timescales, is summarized in section 6.6 of ref. 19. Also, in agreement with Keeling's hypothesis that ocean processes are dominant, we find that the coherence between the CO_2 and the Southern Hemisphere average temperature records is slightly higher than that for the Northern Hemisphere and that the observed delay of −5 months between the global temperature and CO_2 is also seen in the Southern Hemisphere phase. As the Northern Hemisphere average temperature leads CO_2 by −3 months, part of the −5-month delay must be due to the transport time from the Southern Hemisphere to Mauna Loa.

The hole in the phase curve near 1 cycle yr^{-1} occurs because the temperature spectrum there is dominated by a different

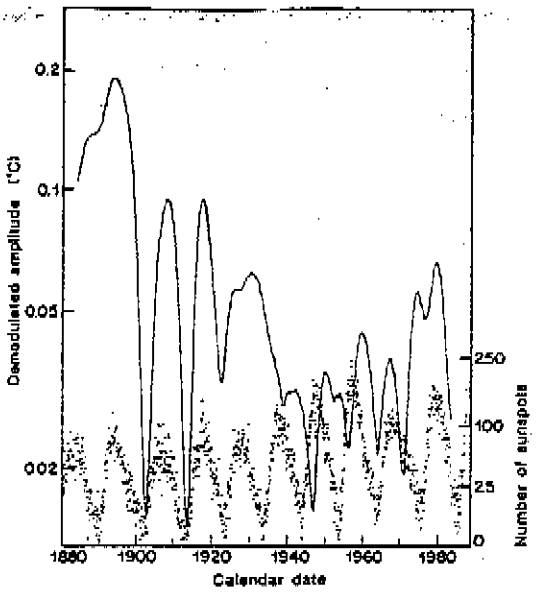

FIG. 4 The solid curve shows the amplitude of the complex demodulate of the Northern Hemisphere temperature record between 0.9 and 1.1 cycle yr^{-1}. A plot of the sunspot numbers is shown below. Detailed examination of the coherence between the demodulate and the solar cycle shows that these two have about a 1/(30 yr) difference in frequency (corresponding to the stronger lower sideband mentioned in Fig. 2*b*).

physical mechanism, and because of their different phase characteristics, we analyse the Northern and Southern Hemisphere temperature records separately in this frequency band. To do this, a multiple-window bandpass-filtered temperature series is formed by expanding the residual temperature series on a set of Slepian sequences that have most of their spectral energy confined between 0.9 and 1.1 cycle yr^{-1}. The resulting 'complex demodulate' for the Hansen–Lebedeff Northern Hemisphere temperatures is shown by the solid line in Fig. 4. The sunspot record (solid dots in Fig. 4) is clearly similar to the demodulated temperature data, suggesting that the annual signal is modulated by fluctuations in solar output. Detailed analysis shows that coherence between the sunspot and filtered temperature series is significant, and similar to the non-stationary structure found in ref. 47. The phase of the sunspot-filtered temperature coherence has a linear drift, corresponding to a frequency difference between these two oscillations of $1/\Delta$, where $\Delta = 27$–30 yr, possibly related to the low-frequency ripple in the temperature spectrum mentioned above. The results for the Southern Hemisphere are similar.

Discussion

The coherence results presented here provide significant evidence that the average global temperature and CO_2 concentration from 1958 to 1988 are linearly related at many frequencies. But caution must be exercised in interpreting this result as suggesting that the variations in atmospheric CO_2 are causing the changes in global temperature, even though there are plausible physical mechanisms linking the two series. Apparent correlations that are used to postulate causality can sometimes be misleading, as in the case involving timing of earthquakes[48]. In addition, one should be particularly cautious in interpreting the coherence when the series analysed are as short as these; climatic and solar variations often are of longer duration than these records.

From atmospheric chemistry, global temperature depends nonlinearly on CO_2 concentration (T. Graedel, personal communication). The procedure used here implicitly uses a linear dependence. Bispectral estimates provide evidence of quadratic terms, although the shortness of these series makes this difficult to quantify. Also, except for the solar modulation of the temperature series near 1 cycle yr^{-1}, we have ignored the cyclostationary properties of these two series (that is, the statistics of these series vary periodically).

A more complete analysis would include estimates of the coherences between the various global average temperature time series, records of atmospheric CO_2 concentration, human CO_2 production, sunspots, volcanic activity and the Southern Oscillation Index, which are all high in various frequency bands and have complicated phase interrelations. For example, the coherence between the detrended Keeling CO_2 series and the Southern Oscillation Index is high near 0.4 and 2.4 cycle yr^{-1}. If we calculate the coherence between the CO_2 and global temperature series from which terms describing their linear dependences on the Southern Oscillation Index and sunspot record have been subtracted (called a partial coherence), the low-frequency magnitude-squared coherence increases to almost 0.7 whereas it decreases near 0.3 cycle yr^{-1}. Several of these series also exhibit an oscillation at an apparent period of ~15 years. Further analysis of their multivariate relations will be described elsewhere in more detail. □

Received 30 October 1989; accepted 8 January 1990.

1. Tyndall, J. *Phil. Mag.* **22**, 160–194; **22**, 273–285 (1863).
2. Arrhenius, S. *Phil. Mag.* **41**, 237–276 (1896).
3. Houghton, R. A. in *The Changing Carbon Cycle: A Global Analysis* (eds Trabalka, J. R. & Reichle, D. E.) 175–193 (Springer, New York, 1986).
4. Schlesinger, W. H. in *The Changing Carbon Cycle: A Global Analysis* (eds Trabalka, J. R. & Reichle, D. E.) 194–220 (Springer, New York, 1986).
5. Rotty, R. M. & Marland, G. *The Changing Carbon Cycle: A Global Analysis* (eds Trabalka, J. R. & Reichle, D. E.) 474–480 (Springer, New York, 1986).
6. Seidel, S. & Keyes, D. *Can We Delay a Greenhouse Warming?* Rep. No. EPA 23010:4001 (US Environmental Protection Agency, Washington, 1983).
7. Budyko, M. I. *Climatic Changes* Ch. 7 (Waverly, Baltimore, 1977).
8. World Meteorological Organization *Rep. World Conf. Changing Atmosphere* (World Meteorological Organization, Geneva, 1989).
9. National Research Council *Changing Climate* (Natn Acad. Press, Washington, DC, 1983).
10. Barnett, T. P. & Schlesinger, M. E. *J. geophys. Res.* **92**, 14772–14780 (1987).
11. Preisendorfer, R. W. & Barnett, T. P. *J. atmos. Sci.* **40**, 1884–1896 (1983).
12. Solow, A. R. *J. Clim. appl. Met.* **26**, 1401–1405 (1987).
13. Maddox, J. *Nature* **334**, 9 (1988).
14. Kerr, R. A. *Science* **244**, 1041–1043 (1989).
15. Broecker, W. S. *Science* **245**, 451 (1989).
16. Schlesinger, M. E. *Science* **245**, 451 (1989).
17. Risby, J. *Science* **245**, 451–452 (1989).
18. Solow, A. R. & Broadus, J. M. *Clim. Change* (in the press).
19. Keeling, C. D., Bacastow, R. B., Carter, A. F., Piper, S. C. & Whorf, T. P. *Am. geophys. Un. Geophys. Monogr.* **55**, 165–236 (1989).
20. Gammon, R. H., Komhyr, W. D. & Peterson, J. T. in *The Changing Carbon Cycle: A Global Analysis* (eds Trabalka, J. R. & Reichle, D. E.) 1–15 (Springer, New York, 1986).
21. Hansen, J. & Lebedeff, S. *J. geophys. Res.* **92**, 13345–13372 (1987).
22. Shapiro, R. *J. atmos. Sci.* **30**, 1105–1116 (1979).
23. Pittock, A. B. *Rev. Geophys. Space Phys.* **16**, 400–420 (1978).
24. Tukey, J. W. *Scripps Inst. Oceanogr. Ref. Ser.* **84–6**, 100–103 (1984).
25. Cleveland, W. S., Freeny, A. E. & Graedel, T. E. *J. geophys. Res.* **88**, 10934–10946 (1983).
26. Priestley, M. B. *Spectral Analysis and Time Series* (Academic, New York, 1981).
27. Brillinger, D. R. *Time Series: Data Analysis and Theory* (Holt, Rinehart & Winston, New York, 1975).
28. Thomson, D. J. *Proc. IEEE* **70**, 1055–1096 (1982).
29. Park, J., Lindberg, C. R. & Vernon, F. L. III *J. geophys. Res.* **92**, 12675–12684 (1987).
30. Grenander, U. *Ann. math. Statist.* **25**, 252–272 (1954).
31. Brillinger, D. R. *Biometrika* **76**, 23–30 (1989).
32. Slepian, D. *Bell System Tech. J.* **57**, 1371–1429 (1978).
33. Belsley, D. A., Kuh, E. & Welsch, R. E. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity* (Wiley, New York, 1980).
34. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
35. Imbrie, J. & Imbrie, J. Z. *Science* **207**, 943–953 (1980).
36. Pearson, E. S. & Hartley, H. O. *Biometrika Tables for Statisticians* 3rd edn Vol. 1. (Cambridge University Press, London, 1970).
37. Thomson, D. J. & Chave, A. D. *Advances in Spectrum Estimation* (ed. Haykin, S.) Ch. 2 (Prentice-Hall, Englewood Cliffs, New Jersey, in the press).
38. Thomson, D. J. *Bell System Tech. J.* **56**, 1769–1815; **56**, 1983–2005 (1977).
39. Daniel, C. & Wood, F. S. *Fitting Equations to Data* (Wiley, New York, 1980).
40. Lindberg, C. R. & Park, J. *Geophys. J. R. astr. Soc.* **91**, 795–836 (1987).
41. Vernon, F. L. III. thesis, Univ. of California (San Diego) (1989).
42. Yule, G. U. & Kendall, M. G. *An Introduction to the Theory of Statistics* 14th edn (Hafner, New York, 1965).
43. Tukey, J. W. in *Advanced Seminar on Spectral Analysis of Time Series* (ed. Harris, B.) 25–46 (Wiley, New York, 1967).