GEOPHYSICAL RESEARCH LETTERS, VOL. 21, NO. 10, PAGES 883-886, MAY 15, 1994

Investigating the origins and significance of low-frequency modes of climate variability

Myles R. Allen

Department of Physics, University of Oxford *and* Rutherford Appleton Laboratory, U.K.

Leonard A. Smith

Mathematical Institute, University of Oxford, U.K.

Abstract. An analysis of the 130-year record of the Earth's global mean temperature reveals a significant warming trend and a residual consistent with an autocorrelated ("red") noise process whose predictability decays with a timescale of two years. Thus global temperatures, in isolation, do not indicate oscillations at 95% confidence against a red noise null hypothesis. Weak signals identified in the global series can, however, be traced to significant sea surface temperature oscillations in the equatorial Atlantic (period ~10 years) and the El Niño region of the Pacific (3–5 years). No robust evidence is found in this data for interdecadal oscillations. The 10-year Atlantic oscillation corresponds to a pattern of temperature anomalies which has been associated with interannual variations in West African rainfall and in U.S. hurricane landfall frequency.

Testing against red noise in SSA

When analysing complex systems, we must allow for noise (or processes indistinguishable from noise) being autocorrelated in time. The simplest discrete autocorrelated noise is the AR(1) process [*Davis & Vinter,* 1985]:

$$u_{t+1} = \gamma u_t + \alpha z_t \qquad (1)$$

where the first term on the RHS represents the memory of the process, and z_t is a white noise (independent, uniformly distributed) forcing. γ and α are real coefficients. The power spectrum of AR(1) noise contains no spectral peaks, although it is biased towards low frequencies (i.e. "red"). Predictability decays exponentially with a time-constant $\tau = -[\ln(\gamma)]^{-1}$. A variety of systems, both deterministic and stochastic, generate red power spectra, so processes like AR(1) noise must be considered in evaluating evidence for trends or oscillations in climate data.

Singular Spectrum Analysis, or SSA [*Broomhead & King,* 1986; *Fraedrich,* 1986; *Vautard & Ghil,* 1989], extracts information from a discrete scalar time series by finding the eigenvalues (λ_k^{data}) and eigenvectors ("EOFs", $\mathbf{e}_k$) of the series' lag-covariance matrix:

$$\lambda_k^{data} \equiv \mathbf{e}_k^T \mathbf{C}^{data} \mathbf{e}_k \qquad (2)$$

where C_{ij}^{data} is the estimated covariance of the series at lag $|j-i|$. Figure 1 shows λ_k^{data} for the 1861-1990 record of global, annual-mean, combined land and sea near-surface temperatures as compiled for the Intergovernmental Panel on Climate Change (IPCC) [*Houghton* et al., 1992]. Here M, the rank of $\mathbf{C}^{data}$, is 40 sampling intervals (years), but conclusions are not sensitive to varying M from 20 to 60 years [*Allen,* 1992].

The λ_k^{data} in figure 1 are arranged in the conventional "rank-order", i.e. by decreasing size. When SSA is used for signal detection, a symmetric/anti-symmetric pair of approximately sinusoidal EOFs, with similar, high-ranked eigenvalues, has been taken to indicate a physical oscillation. If the series contains red noise, however, then sinusoidal EOFs which are entirely due to noise may have high-ranked eigenvalues. Standard SSA applied to short realizations of AR(1) noise with the same lag-1 autocorrelation as the IPCC series (γ=0.8) indicates low-frequency oscillations in over 50% of cases, even when the EOF-selection criteria of [*Vautard* et al., 1992] are applied. Thus we cannot conclude that such oscillations exist on the basis of conventional SSA alone.

To resolve this problem, we introduce Monte Carlo SSA. Suppose we suspect that a series consists of a nonlinear trend contaminated with AR(1) noise. We generate an ensemble of "surrogate" series [*Smith,* 1992; *Theiler* et al., 1992], each consisting of the observed >40-year variability (reconstructed from EOFs 1 & 2 following [*Ghil & Vautard,* 1991]) plus a segment of AR(1) noise, generated using (1). The coefficients α and γ are chosen to give the surrogates the same expected variance and lag-1 autocorrelation as the original data: the best-fit $\gamma = 0.6$, giving a decay-time τ of ~2 years. We calculate lag-covariance matrices, $\mathbf{C}^{surr}$, for each of the surrogates and project them onto the EOFs of $\mathbf{C}^{data}$. The power in the k^{th} EOF is defined as $\lambda_k^{surr} \equiv \mathbf{e}_k^T \mathbf{C}^{surr} \mathbf{e}_k$: see (2). Vertical bars in figure 1 show the 95% and 80% limits of the λ_k^{surr} distributions.

If the k^{th} data eigenvalue, λ_k^{data}, lies above the 95% limit of the corresponding λ_k^{surr} distribution, this would indicate that we can reject the null hypothesis "power associated with EOF-k is attributable to the observed

Paper number 94GL00978
0094-8534/94/94GL-00978$03.00

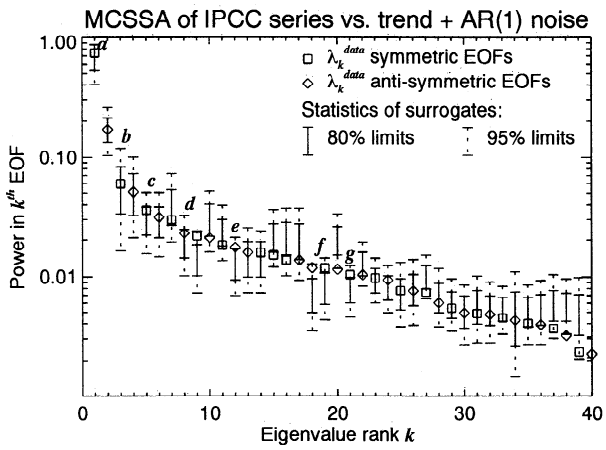

Figure 1. Squares and diamonds: eigenvalues, λ_k^{data}, of $\mathbf{C}^{data}$ for the IPCC global temperature series. Bars: distributions of λ_k^{surr} for a 1000-member ensemble of surrogates, each consisting of the observed trend contaminated with AR(1) noise. 95% of the λ_k^{surr} lie below (above) the upper (lower) limits of the dotted bars; 80% for the solid bars. No λ_k^{data} lie above the upper 95% limits, and only 2 lie just below the lower limits, so this data gives no evidence of oscillations at 95% confidence against a trend-plus-red-noise null-hypothesis. EOFs labelled c to f may be significant at the 80% level, but the "interdecadal" pair, b, is not.

trend plus this *specific* AR(1) process" at the 95% confidence level. Out of 40 EOFs, we would clearly expect some such excursions to occur by chance. Quantifying the probability of a given number of excursions is difficult because (i) the λ_k^{surr} within each individual surrogate realisation are not independent, and (ii) the noise process parameters have been fitted to the data (not a serious problem provided the number of parameters is $\ll M$). However, since *no* excursions above the 95% limits occur in figure 1, no oscillatory signals are indicated at 95% confidence.

Excursions above the 80% limits occur in EOFs corresponding to periods of 5 (pair d and e), 3.6 (pair f) and 2.9 years (g). We associate a frequency, f_k, with an EOF by maximizing the correlation between $\mathbf{e}_k$ and a sinusoid. The spectral resolution of SSA is $\Delta f_k \simeq \pm\frac{1}{2M}$. If these EOFs are included in the signal-reconstruction along with the trend, and γ and α are reset to give the surrogates the same variance and lag-1 autocorrelation, we also find a 10-year signal at this lower significance level (pair c). Some of these excursions should occur by chance, so our confidence in these possible signals is even lower than 80%. We draw attention to them not as conclusive evidence of oscillations, but to stress that eigenvalue rank-order can fail as a significance criterion [*Allen* et al., 1992]. EOF pair f in figure 1 (which corresponds to the 3–4-year component of El Niño) is more significant than the 5 EOFs which precede it. Physically, this means that this EOF pair explains an improbably high proportion of the variance in the data given the frequencies associated with it. Against a red noise null-hypothesis, we do not *expect* the same variance in all EOFs even if the data is pure noise. No interdecadal oscillations are indicated at the 80% level: the λ_k^{data} labelled b lie near the centres of their corresponding λ_k^{surr} distributions.

If we test against AR(1) noise without the trend, we find improbably high power at k=1 (a in figure 1): the >40-year variability in the IPCC series cannot be explained as a consequence of this AR(1) process.

The significance tests used by [*Ghil & Vautard,* 1991] only considered white noise contamination. The test introduced here can be generalised to distinguish between a data series and any well-defined process. If we *assume* that the noise is white, by setting γ=0 in (1), we find evidence for a trend and oscillations with approximate periods of 27, 14, 10, 6, 5 and 3.6 years, all apparently significant at 95%, and together accounting for almost 85% of the variance. Introducing one extra parameter in the noise (γ) explains away all these apparent oscillations. Thus assuming, without *a priori* grounds, that noise is white may lead us to underestimate the amplitude of the stochastic component of a series and thus overestimate system predictability.

Origins of possible signals identified

Signals which cannot be distinguished from AR(1) noise in the global mean temperatures *alone* may nonetheless have a genuine physical origin. To investigate possible oceanic origins, we used "regression-maps" [*Lim & Wallace,* 1991] to identify sea surface temperature (SST) patterns associated with the different modes of the global temperatures, as follows. SSA allows the components of a time-series which contribute to variability on different time-scales to be reconstructed independently. Figure 3a of [*Ghil & Vautard,* 1991] shows "reconstructed components" (RCs) of global temperatures corresponding to the >40-year (trend), 3–5-year, 10-year and 27-year modes. We shall refer to these as x_{1t} to x_{4t} respectively.

We applied the following regression model to local time series of annual-mean SSTs on a global 5° × 5° grid [*Bottomley* et al., 1990]:

$$y_t = b_0 + \sum_{m=1}^{4} b_m x_{mt} + u_t \qquad (3)$$

where y_t is the local SST and u_t is a red noise process defined as in (1) above. The 4 RCs of the global series, $x_{mt:\ m=1,4}$, were common to all points: the "reference time series" of the regression map. At each grid point, we estimated b_0 and the coefficients $b_{m:\ m=1,4}$ using the method of generalized linear regression (GLR) with missing data [*Kmenta,* 1981; *Allen,* 1992]. Because residuals are autocorrelated, using standard linear regression – replacing u_t with z_t in (3) – would result in serious underestimation of the errors, $\sigma(b_m)$ [*Davis & Vinter,* 1985]. In many regions, coverage is poor prior to *c.* 1950. Rather than requiring any "gap-filling", we make use only of the data available: an advantage over analyses based on spatial EOFs [*Folland* et al., 1991].

Optimizing γ, the lag-1 autocorrelation of u_t, for each grid-point individually gave a global average of 0.3, implying a decay-time $\tau<1$ year. Thus introducing land and ice data and/or taking a global average temperature appears to increase the level of autocorrelation (as suggested by M. Ghil). Individual estimates of γ are not very robust, so we imposed a global value of 0.3: the following results are not sensitive to varying γ, nor to omitting the (doubtful) 27-year mode.

Figures 2.1 to 2.4 show maps of regression coefficients, b_m. If $2<b_m<4$ in a particular square (medium-dark grey), then SSTs, y_t, at that point have varied in phase with, and with 2–4 times the amplitude of, the m^{th} RC of the global temperatures, x_{mt}. Points where $|b_m|>2\sigma(b_m)$ – locally significant at $\gtrsim$95% – are enclosed by squares. A significant relationship between y_t and x_{mt} confined to a particular region with $b_m > 1$ would indicate that the m^{th} RC of the global temperatures originates in that region.

Regression coefficients on the trend (figure 2.1) indicate a fairly uniform "warming" ($b_1 \sim 1$) over the Atlantic and Indian Oceans, with enhanced warming in the north-west Pacific. Negative coefficients (light grey squares) occur off southern Greenland and in the data-sparse Arctic and tropical Pacific, but few of these are significant, so they may simply be due to inadequate coverage in these regions.

The clearest feature related to the oscillatory components of global temperatures is the El Niño signal associated with the subdecadal RC, x_{2t} (figure 2.2). The global nature of El Niño is clearly visible.

The key significant feature in the pattern associated with the 10-year RC, x_{3t} (figure 2.3), is a patch of positive coefficients in the northern equatorial Atlantic, with negative coefficients south of the equator. Such a pattern of SST anomalies has been associated with variations in Sahel rainfall [*Folland* et al., 1991] and U.S. hurricane landfall frequency: figure 9 of [*Gray,* 1990]. A second patch of significant coefficients is observed in the south-east Pacific, but coverage is very poor in that region. Assessing the overall significance of such features is difficult, since we expect ~5% of the grid-point regressions to appear significant by chance. The 10-year pattern in the equatorial Atlantic is, however, fairly robust: it is largely unchanged if (following [*Elsner & Tsonis,* 1991]) we repeat the analysis using only the years 1861–1970, 1881–1990 and 1910–1990.

Applying Monte Carlo SSA to the area-mean SST in the north equatorial Atlantic (figure 3), we found (in addition to >40-year variability and an ENSO signal) a 10-year period oscillation with a peak to peak amplitude of up to >0.4°C. The evidence for this oscillation is weak: it is only distinguishable from AR(1) noise at 95% confidence in the full 130-year series, not in either the first or the last 110-year segments. A reconstruction of this signal is, however, in phase with the 10-year component of the global series; both go through a quiescent period between 1910 and 1955 during which they both undergo a ~180° change of phase. The correspondence of phase and amplitude suggests this is the origin of the 10-year mode in the global temperatures, although in the early part of the series the appearance of this mode in the global record could be due to inadequate coverage of regions outside the Atlantic. No significant 10-year oscillation could be identified in an area-mean series from the south equatorial Atlantic (confirming

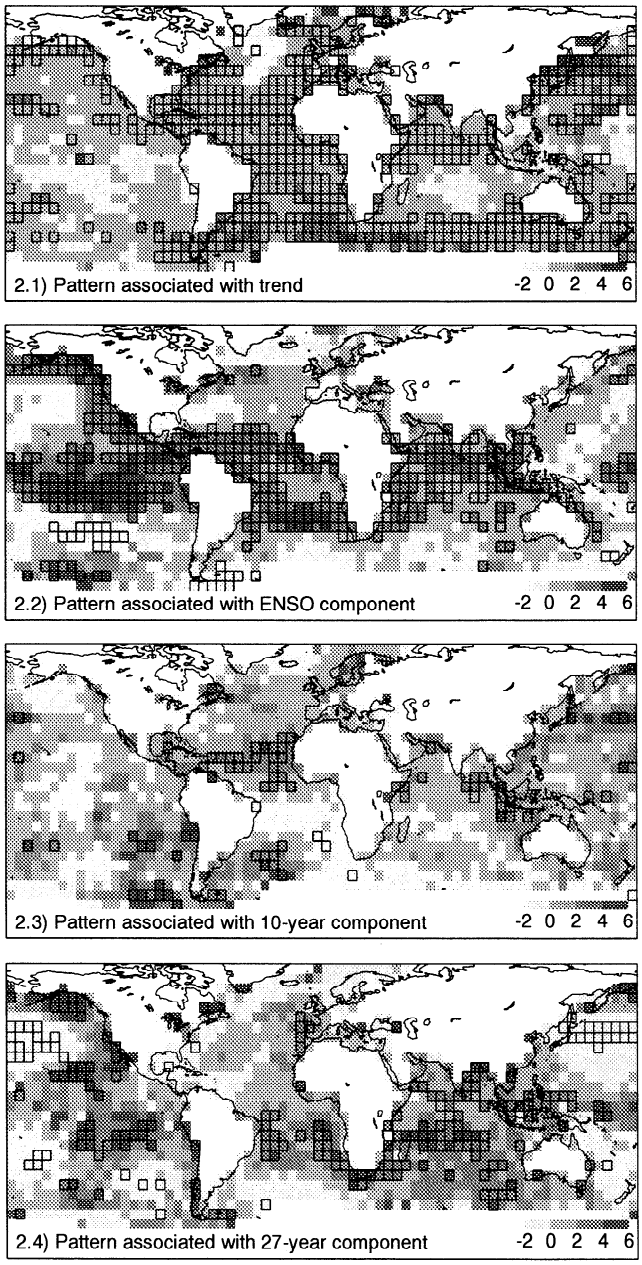

Figure 2. Patterns associated with (2.1) trend, (2.2) subdecadal, (2.3) 10-year and (2.4) 27-year RCs of global temperatures. Grey-scales indicate the coefficients b_m in eqn. (3) between local SSTs, y_t, and RCs x_{1t} to x_{4t} of global temperatures. Boxes indicate points where $|b_m|>2\sigma(b_m)$. Each y_t is the average of 4 seasonal anomalies, calculated from monthly anomalies about a 1951-80 mean annual cycle. Bucket-corrected data used from the MOHSST5 data-set. Years in which northern-winter or northern-summer season contain no data treated as missing. Points with ≤20 years of data shown white.

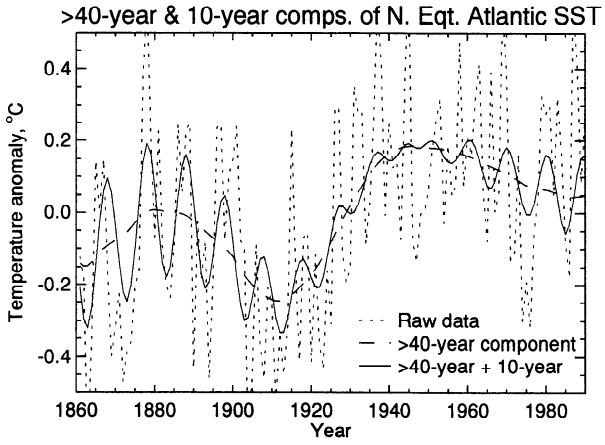

Figure 3. SST anomalies in the northern equatorial Atlantic (area-mean 70W–10E, 0–20N: ~80% spatial coverage since 1880). 10-year oscillatory component (solid line) is significant against AR(1) noise at 95%. Note correspondence of phase and amplitude with 10-year signal in figure 3a of [*Ghil & Vautard,* 1991].

[*Houghton & Tourre,* 1992] that this is not a simple north-south dipole), nor from the SE Pacific.

Figure 2.4 shows no such clear "origin" of the 27-year RC. Significant coefficients occur in the Indian Ocean ($b_4>0$) and in the northwest Pacific ($b_4<0$). Time series of area-means from these regions do not, however, indicate significant interdecadal oscillations. Significant positive coefficients also occur in the eastern Pacific, but the similarity of the patterns in figures 2.2 and 2.4 in this region suggests this is a consequence of data-sparsity and amplitude-modulation of the El Niño signal. This conclusion is supported by repeating the analysis on the years 1861-1970: the 27-year pattern in the Pacific then completely disappears.

Conclusions

In our analysis of the global temperature record, the only component of variability which we were able to distinguish, at 95% confidence, from a simple form of red noise was the warming trend. Generalized regression analysis between the reconstructed trend and gridded SSTs indicated it is spatially ubiquitous, which is consistent with (although it does not, of course, establish) a global change in external forcing. We also identified subdecadal variability associated with El Niño and a possible 10-year period oscillation which appears to originate in the northern equatorial Atlantic. No evidence was found for oscillations on 20–30 year time scales, although this may simply indicate that this series of 130 years is just too short for such oscillations to be distinguished from red noise.

Acknowledgments. We would like to thank D. Anderson, D. Broomhead, A. Darbyshire, T. Davidson, J. Elsner, C. Folland, M. Ghil, J. Huke, M. Jackson, G. King, B. Legras, M. Muldoon, D. Parker, P. Read, S. Ruth, R. Smith, T. Tsonis, R. Vautard and the reviewers. MRA was supported by a CASE award from the UK NERC and Met. Office, LAS by a Senior Research Fellowship from Pembroke College, Oxford.

References

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.

Allen, M. R., Read, P. L., & Smith, L. A. 1992. Temperature timeseries? *Nature*, **355**, 686.

Bottomley, M., *et al.* 1990. *Global Ocean Surface Temperature Atlas.* London: H.M.S.O. Updated data provided by M. Jackson and D. Parker, UKMO.

Broomhead, D. S., & King, G. 1986. Extracting qualitative dynamics from experimental data. *Physica D*, **20**, 217–236.

Davis, M. H. A., & Vinter, B. B. 1985. *Stochastic modelling and control.* Chapman and Hall.

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

Folland, C., Owen, J., Ward, M., & Colman, A. 1991. Prediction of seasonal rainfall in the Sahel using empirical and dynamical methods. *J. Climatology*, **10**, 21–56.

Fraedrich, K. 1986. Estimating the dimension of weather and climate attractors. *J. Atmos. Sci.*, **43**, 419–432.

Ghil, M., & Vautard, R. 1991. Interdecadal Oscillations and the Warming Trend in Global Temperature Time Series. *Nature*, **350**, 324–327.

Gray, W. M. 1990. Strong Association Between West African Rainfall and U.S. Landfall of Intense Hurricanes. *Science*, **249**, 1251–1256.

Houghton, J. T., *et al.* (eds). 1992. *Climate Change 1992, Supplement to the IPCC Scientific Assessment.* Cambridge Univ. Press.

Houghton, R. W., & Tourre, Y. M. 1992. Characteristics of low-frequency sea surface temperature fluctuations in the tropical Atlantic. *J. Climate*, **5**, 765–771.

Kmenta, J. 1981. On the problem of missing measurements in the estimation of economic relationships. *Pages 233–257 of:* Charatsis, E. G. (ed), *Proc. Econometric Society European meeting 1979.* Amsterdam: North Holland.

Lim, G. H., & Wallace, J. M. 1991. Structure and evolution of baroclinic waves as inferred from regression analysis. *J. Atmos. Sci.*, **48**, 1718–1732.

Smith, L.A. 1992. Identification and Prediction of Low-Dimensional Dynamics. *Physica D*, **58**, 50–76.

Theiler, J., Eubank, S., Longtin, A., Galdrikan, B., & Farmer, J. 1992. Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, **58**, 77–94.

Vautard, R., & Ghil, M. 1989. Singular Spectrum Analysis in nonlinear dynamics with applications to paleoclimatic time series. *Physica D*, **35**, 395–424.

Vautard, R., Yiou, P., & Ghil, M. 1992. Singular Spectrum Analysis: A toolkit for short, noisy and chaotic series. *Physica D*, **58**, 95–126.

M. R. Allen, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK (atmmra@vax.ox.ac.uk)

L. A. Smith, Mathematical Institute, 24-29 St. Giles, Oxford OX1 3LB, UK (lenny@maths.oxford.ac.uk)

(received June 29, 1993; revised December 1, 1993; accepted February 23, 1994.)