Detection of the Climate Response to the Solar Cycle

MARK J. STEVENS AND GERALD R. NORTH

Climate System Research Program, Department of Meteorology, Texas A & M University, College Station, Texas

(Manuscript received 27 November 1995, in final form 5 April 1996)

ABSTRACT

Optimal space–time signal processing is used to infer the amplitude of the large-scale, near-surface temperature response to the "11 year" solar cycle. The estimation procedure involves the following steps. 1) By correlating 14 years of monthly total solar irradiance measurements made by the *Nimbus-7* satellite and monthly Wolf sunspot numbers, a monthly solar irradiance forcing function is constructed for the years 1894–1993. 2) Using this forcing function, a space–time waveform of the climate response for the same 100 years is generated from an energy balance climate model. 3) The space–time covariance statistics in the frequency band $(16.67\ \text{yr})^{-1}$–$(7.14\ \text{yr})^{-1}$ are calculated using control runs from two different coupled ocean–atmosphere global climate models. 4) Using the results from the last two steps, an optimal filter is constructed and applied to observed surface temperature data for the years 1894–1993. 5) An estimate of the ratio of the real climate response, contained in the observed data, and the model generated climate response from step 2 is given, as well as an estimate of its uncertainty. A number of consistency checks are presented, such as using data from different regions of the earth to calculate this ratio and using data lagged up to ±5 yr. Our findings allow us to reject the null hypothesis, that no response to the solar cycle is present in the data, at a confidence level of 97.4%.

1. Introduction

Global climate models are developed to simulate the climate system on timescales from months to decades and centuries. A major test of these models is their ability to simulate the seasonal cycle of surface temperature. Currently, beyond the timescale of 1 yr, there are no known natural examples of periodic forcing that could be used to test the long-term response of these models. The eruption of Mount Pinatubo in 1991 presented a natural experiment to test how well global climate models simulate the climate response to an impulse forcing. The models did fairly well at reproducing the effects of the additional aerosol forcing on the climate system (Hansen et al. 1992). This was encouraging to climate modelers, although not unexpected since the models are tuned to reproduce the seasonal cycle and could be expected to do well on timescales of several years. An important use of global climate models is to attempt to predict the response the climate will have to a doubling of the CO_2 concentration in the atmosphere. However, this response is on the timescale of decades to centuries, well beyond the timescale of any known test of the models. If a periodic forcing of the climate system on the decadal timescale could be found, it could bring added confidence in the ability of these models to predict the near-term climatic effects of the increase in atmospheric greenhouse gases. In this paper, we present an estimate of the response of the climate system to just such a forcing on the decadal timescale.

In 1843, the German amateur astronomer Heinrich Schwabe discovered an approximately 10-yr cycle in the number of dark spots visible on the sun's surface (Schwabe 1844). Since that time, many people have tried to correlate this cycle of sunspot numbers with various measurements of the earth's climate, especially with the surface temperature. For many years it was believed that the total radiative output of the sun decreased as the number of sunspots increased. This was reasoned to occur because the sunspots are cooler than their surroundings and hence would radiate less energy. However, it is now known that the radiative output of the sun actually increases with the sunspot number. This increase in total solar irradiance is due to the presence of areas of greater radiative output, known as faculae, which accompany the sunspots. The net output change is very small, but has been measured by sensors on several research satellites over the past 16 years. The amplitude (half the range) of this solar cycle of irradiance is approximately 0.6 W m^{-2} (Willson and Hudson 1991; Lee et al. 1995). However, the measurements only span one and one-half solar cycles, so the long-term variation of the solar irradiance is still unresolved.

The response of the climate system to such a small perturbation is most likely to also be very small. Esti-

Corresponding author address: Dr. Mark J. Stevens, Climate System Research Program, Department of Meteorology, Texas A&M University, College Station, TX 77843-3150.
E-mail: stevens@diffuse.tamu.edu

mates of the amplitude of the climate response to the solar cycle, as measured by the global annual mean surface temperature, range in size from 0.01°C to 0.08°C (Schneider and Mass 1975; Eddy et al. 1982; North et al. 1983; Hansen and Lacis 1990; Ardanuy et al. 1992; Crowley and Kim 1993). The estimated standard deviation of the global annual mean of the surface temperature is about 0.18°C. Thus, the solar cycle produces a very faint signal embedded in the natural variability of the climate (the noise). There have been many attempts made over the last 50 years to find this signal, usually by employing Fourier analysis of surface temperature time series data. These attempts have proven unsuccessful largely because of the small signal-to-noise ratio found in the data.

Signal processing techniques for detecting a signal amidst background noise have been developed by mathematicians and electrical engineers over the past 40 years (Selin 1965). Only recently have these techniques been applied to the problem of detecting the global warming signal due to increases in greenhouse gases (Wigley and Raper 1990; Wigley and Barnett 1990; Barnett 1991; Barnett et al. 1991). Even more recently, what are known as statistically optimal methods have been developed to detect such a signal (Hasselmann 1979, 1993; Bell 1982, 1986; Hegerl et al. 1994; Hegerl and North 1995; North et al. 1995; North and Kim 1995; Hasselmann et al. 1995). In this paper we use what is known as a space–time optimal filter to look for the signal of the climate response to the solar cycle in the historical record of the earth's surface temperature. In this case the signal is a spatial pattern that changes quasi-periodically over time. The filter is built using both spatial and temporal information about the expected signal and the background noise. The filter is optimal in the sense that, for the given signal, the maximum possible signal-to-noise ratio is found. A first attempt at this problem was described in North and Kim (1995).

2. Building the optimal filter

To construct an optimal filter, we must make the basic assumption that the data is a linear combination of the signal and the noise. This assumption is supported by a few model simulation studies involving the atmosphere alone (North et al. 1992; Marshall et al. 1994). On the other hand, the climate system is highly nonlinear, and it could be that even infinitesimal forcings could lead to responses that are not linear. While we cannot rule out such possibilities, we find no compelling reason not to ignore them at this time.

a. Filter theory

In this section, we sketch a brief derivation of the space–time optimal filter as presented in North et al. (1995) and North and Kim (1995). Suppose we are given several M component vectors of discrete time series, each of length N, $T_{m,n}$, where $m = 1, \cdots, M$, and $n = 1, \cdots, N$. These might be from, for example, M stations or regions on the earth, each with a record of N years. Let the time series be composed of two parts, a signal and noise:

$$T^{\text{data}}_{m,n} = T^{\text{signal}}_{m,n} + T^{\text{noise}}_{m,n}. \tag{2.1}$$

The first of these time series is deterministic and must be computed from theory or from a model. The second term above represents the natural variability of the climate system and is taken to be a random field defined over the earth with N time units. This too must be generated from a model since the observed data contains both the natural climate variability and the hypothesized signal. We must have an ensemble of records of $T^{\text{noise}}_{m,n}$ in order to compute the necessary statistics for use later. The ensemble can be considered as a collection of segments of a Gaussian stationary time series, but with correlations between the stations as well as serial correlations in time. A filter is sought that provides an estimator of the signal from the observed data:

$$\hat{T}^{\text{signal}}_{m,n} = \sum_{m'=1}^{M} \sum_{n'=1}^{N} \Gamma^{m,m'}_{n,n'} T^{\text{data}}_{m',n'}, \tag{2.2}$$

where $T^{\text{data}}_{m,n}$ is the observed data from the M stations and $\Gamma^{m,m'}_{n,n'}$ is the smoothing filter. The mean square error for the estimate is

$$\epsilon^2_{m,n} = \langle (\hat{T}^{\text{signal}}_{m,n} - T^{\text{signal}}_{m,n})^2 \rangle, \tag{2.3}$$

where $\langle \cdot \rangle$ denotes the ensemble average. If we minimize the mean square error with respect to the functional form of the filter and subject to the constraint that the estimate be unbiased ($\langle \hat{T}^{\text{signal}}_{m,n} \rangle = T^{\text{signal}}_{m,n}$), we obtain a linear matrix equation for the unknown optimal filter. The solution for the optimal filter may be expressed as

$$\Gamma^{m,m'}_{n,n'} = \frac{T^{\text{signal}}_{m,n}}{\gamma^2} \sum_j \sum_k \frac{(T^{\text{signal}}, \mathbf{e}_{j,k}) \mathbf{e}^{m',n'}_{j,k}}{\lambda_{j,k}} \tag{2.4}$$

with

$$\gamma^2 = \sum_j \sum_k \frac{(T^{\text{signal}}, \mathbf{e}_{j,k})^2}{\lambda_{j,k}}, \tag{2.5}$$

where the $\mathbf{e}^{m,n}_{j,k}$ are the space–time empirical orthogonal functions (EOFs), or principal components; $\lambda_{j,k}$ are eigenvalues of the noise covariance matrix; and where $(T^{\text{signal}}, \mathbf{e}_{j,k}) \equiv \Sigma_{m,n} T^{\text{signal}}_{m,n} \mathbf{e}^{m,n}_{j,k}$ is the projection of the signal onto the space–time EOFs. The index j is the eigenvector index, with $j = 1, \cdots, M$, and the index k is the frequency index, with $k = -N/2, \cdots, N/2$. The space–time EOFs are used to characterize the spatial and temporal structure of the noise. The eigenvalues $\lambda_{j,k}$ in Eq. (2.4) can be considered as optimal weights of the space–time EOFs. The quantity in Eq. (2.5) is the square of the theoretical signal-to-noise ratio. It is

deemed theoretical since all the terms involved in its calculation are derived from models. Figure 1a shows a schematic diagram of the steps involved in the construction of the optimal filter. Details will be covered in later sections.

We can also build a nonoptimized filter (the so-called "fingerprint" filter):

$$\Gamma_{n,n'}^{m,m'} = \frac{T_{m,n}^{\text{signal}}}{\gamma_{\text{nopt}}^2 \sigma^2} \sum_j \sum_k (T^{\text{signal}}, \mathbf{e}_{j,k}) \mathbf{e}_{j,k}^{m'n'}, \quad (2.6)$$

where the square of the nonoptimized signal-to-noise ratio is given by

$$\gamma_{\text{nopt}}^2 = \frac{1}{\sigma^2} \sum_j \sum_k (T^{\text{signal}}, \mathbf{e}_{j,k})^2 \quad (2.7)$$

and the total variance is given by

$$\sigma^2 = \sum_j \sum_k \lambda_{j,k}. \quad (2.8)$$

b. *The space–time EOFs*

Assuming that the time series of noise is stationary, we can, to a first approximation, factor the space–time EOFs into a product of temporal and spatial parts (Kim and North 1993). We write the time factor as a complex exponential. The space–time EOFs are then written as

$$\mathbf{e}_{j,k}^{m,n} = \mathbf{u}_{j,k}^m \frac{\exp(i2\pi f_k t_n)}{\sqrt{N}}, \quad (2.9)$$

where $f_k = k/N$ is the frequency and $t_n = n$ is the time. The spatial factors $\mathbf{u}_{j,k}^m$ are the eigenvectors of the noise covariance matrix. In all our calculations we will restrict the frequencies to the band from 0.06 to 0.14 yr^{-1}. The reason this is done will be explained in a later section. This frequency band, which includes periods from 16.67 yr to 7.14 yr, is hereafter referred to as the solar-cycle frequency band. In principle the eigenvalues and spatial eigenvectors should also depend on frequency, but we have assumed that over the solar cycle band they do not, so we drop the k index on the $\lambda_{j,k}$ and $\mathbf{u}_{j,k}^m$. We will describe how the noise covariance matrix is computed in a later section. In this notation the phase information is handled by the real and imaginary parts of the complex exponential.

The EOFs are orthonormal and complete, so that we can expand functions into series of them:

$$T_{m,n}^{\text{signal}} = \sum_j \sum_k \tilde{T}_{j,k}^{\text{signal}} u_j^m \frac{\exp(i2\pi f_k t_n)}{\sqrt{N}}. \quad (2.10)$$

Similarly, realizations of the noise can be expanded into such representations. The covariance of the noise coefficients have the special property

$$\text{cov}(\tilde{T}_{j,k}^{\text{noise}}, \tilde{T}_{j'k'}^{\text{noise}}) = \lambda_j \delta_{j,j'} \delta_{k,k'} \quad (2.11)$$

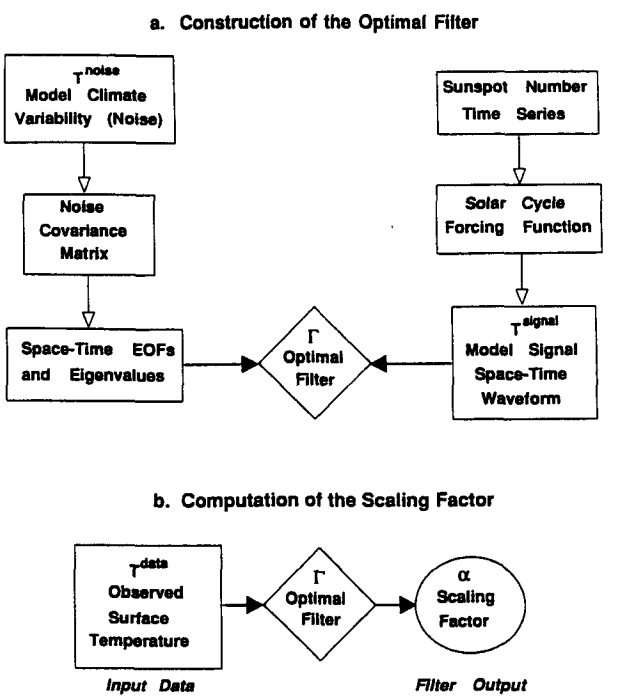

FIG. 1. Schematic diagrams showing (a) the construction of an optimal filter and (b) the calculation of the scaling factor.

based upon EOF theory (e.g., North 1984). The λ_j are the natural variances associated with each space–time EOF mode. We have assumed the dependence of the λ_j on frequency is negligible. The distribution of power in the λ_j is the generalization of the frequency spectrum in pure time series analysis.

c. *The scaling factor*

If we substitute Eq. (2.4) into Eq. (2.2) we can write the estimated signal as

$$\hat{T}_{m,n}^{\text{signal}} = \frac{T_{m,n}^{\text{signal}}}{\gamma^2} \sum_j \sum_k \frac{(T^{\text{signal}}, \mathbf{e}_{j,k})(T^{\text{data}}, \mathbf{e}_{j,k})}{\lambda_j}, \quad (2.12)$$

where $(T^{\text{data}}, \mathbf{e}_{j,k}) \equiv \Sigma_{m',n'} T_{m',n'}^{\text{data}} \mathbf{e}_{j,k}^{m',n'}$ is the projection of the data onto the space–time EOFs. We then define the scaling factor as

$$\alpha \equiv \frac{1}{\gamma^2} \sum_j \sum_k \frac{(T^{\text{signal}} \mathbf{e}_{j,k})(T^{\text{data}}, \mathbf{e}_{j,k})}{\lambda_j}. \quad (2.13)$$

The space–time waveform of the estimated signal can then be expressed as

$$\hat{T}^{\text{signal}} = \alpha T^{\text{signal}}. \quad (2.14)$$

Thus, the scaling factor α is a random variable that represents the ratio of the estimated signal in the observed data to the model-generated signal. A value of

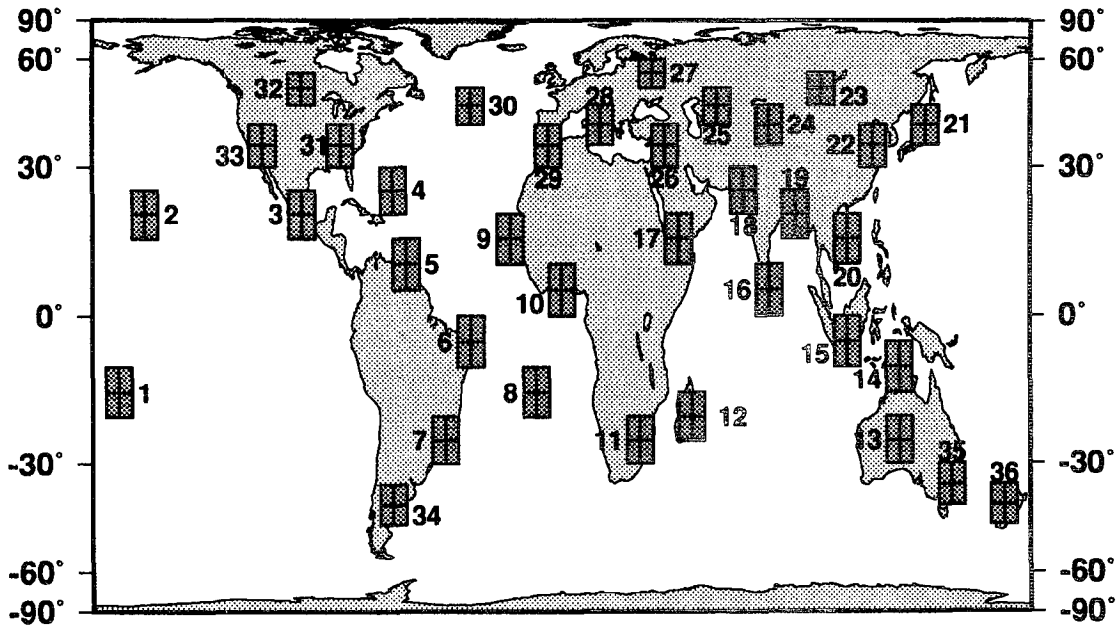

FIG. 2. Geographical locations of the 36 detection boxes.

unity would imply that the signal in the data has exactly the same amplitude as the model-generated space–time waveform. In Fig. 1b a schematic diagram shows the computation of the scaling factor.

As can be seen from Fig. 1a and Eq. (2.4), construction of the optimal filter requires both the model-generated space–time climate variability and the model-generated space–time signal. In addition, from Eq. (2.13) and Fig. 1b we see that computation of the scaling factor requires the observed space–time surface temperature data. In the next three sections we give details about each of these.

3. The observed data

For the historical record of surface temperature, we used the combined land and ocean dataset compiled by Jones and Briffa (1992) and Parker et al. (1994). This dataset, which was obtained from the National Center for Atmospheric Research (NCAR), contains monthly surface temperature anomalies from 1854 to 1993 on a 5° × 5° latitude/longitude grid. To reduce the noise in the data, we averaged four adjacent 5° × 5° boxes to create a 10° × 10° detection box. Since we are searching for a signal that has a period of approximately 11 yr, we require that each of the 5° × 5° boxes used contain 100 years of monthly data. To ensure adequate spatial sampling the boxes should be well distributed over the earth's surface, while at the same time the spatial correlation of climate variability between the boxes should be minimized. Within these constraints of data availability and desired spatial distribution, we selected 36 detection boxes. Figure 2 shows their locations.

In order to maximize the signal-to-noise ratio, it is necessary to decide which months to use from the observed data. Figure 3 shows the monthly area-weighted variances for the detection boxes in the Northern Hemisphere extratropics and Tropics (30°N to 30°S). It is clear that using all 12 months of the year will result in very small variances in the Tropics. It is also clear that the winter months in the extratropics have much larger variances than the summer months. Accordingly, we adopt the rule that only the 6 months with the smallest variances will be used in the extratropics, and all 12

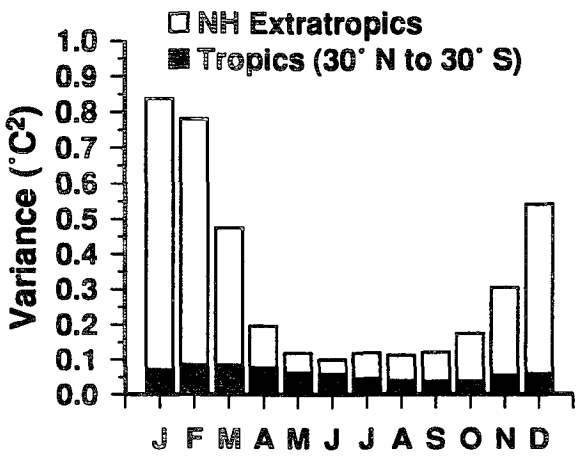

FIG. 3. Area-weighted monthly mean variances (°C²) of the observed surface temperature anomalies for the 13 detection boxes in the Northern Hemisphere extratropics and the 20 boxes in the Tropics. Monthly means for the period 1894–1993.

months in the Tropics. Using this rule, a 100-yr time series of the observed data was constructed for each of the 36 detection boxes. Having adopted this rule for the observed data, it was also applied to the construction of the model signal and model noise time series for each of the 36 detection boxes.

4. The model noise

Since the signal we are seeking is presumably embedded in the observational data, we cannot use the data to compute the natural variability of the climate. The only alternative is to use long control runs from coupled ocean–atmosphere climate models. We assume these models adequately represent the space–time structure of the natural variability, including the spatial correlations of the surface temperature (Kim et al. 1996b).

The control runs that we use in this study are from the Geophysical Fluid Dynamics Laboratory (GFDL) in Princeton, New Jersey (Stouffer et al. 1994), and from the Max-Planck-Institut für Meteorologie (MPI) in Hamburg, Germany (Cubasch et al. 1992; von Storch 1994). The GFDL dataset was obtained from the National Climate Data Center (NCDC). Both the GFDL and MPI models are three-dimensional global climate models with multilayer dynamical atmosphere and ocean components and with surface and sea bottom topography and include a sea ice model. The MPI model, known as ECHAM/LSG, does not generate any El Niño–Southern Oscillation effect (Hegerl et al. 1994), while the GFDL model produces a weak ENSO effect. Both datasets contain 1000 years of monthly mean data, which were interpolated onto 5° × 5° latitude/longitude grids. For each month the time series at each grid point was converted into a time series of surface temperature anomalies, as in the Jones dataset.

The following procedure was used to calculate the covariance matrix of the model noise. For each of the 36 boxes the monthly time series within that box was averaged using the rule described in the last section. The resulting thirty-six 1000-yr time series were then split into shorter time series of 100-yr length (same length as the observed data). Thus, the GFDL and MPI datasets each yielded ten 100-yr time series for each of the 36 boxes. Each set of 36 of these 100-yr time series is considered one realization, so we have ten realizations from each dataset. We assume that these time series are stationary. For each of these realizations the covariance between the 36 boxes was calculated at each of the nine frequencies in the solar-cycle band. Assuming that the variance has no frequency dependence, the covariances were averaged over the frequencies in the band. Finally, the ensemble mean of the covariances was calculated by averaging over the realizations. The end result of this procedure is a 36 × 36 real symmetric covariance matrix of the model noise. The eigenvectors and eigenvalues of this matrix were calculated by diagonalizing the matrix. Each of the 36 eigenvectors has 36 components and a corresponding eigenvalue. These 36 spatial eigenvectors are used in the construction of the space–time EOFs, as in Eq. (2.9). Following this procedure, a complete set of space–time EOFs and eigenvalues was produced from both the GFDL and MPI datasets.

5. The model signal

a. The solar-cycle forcing function

Measurements taken by several satellites over the past 16 years indicate that the total solar irradiance is correlated with the sunspot number. Figure 4 shows the monthly Wolf sunspot number and the monthly *Nimbus-7* total solar irradiance for the years 1979–92. Both datasets were obtained from NCDC. By lagging the *Nimbus-7* data by one month, the correlation is found to be 0.725. The record of monthly mean sunspot number since 1749 is available, and hence a plausible record of the monthly total solar irradiance can be constructed by using the calibration from the satellite measurements. Several other methods of calculating such a forcing function have been developed by other researchers (Foukal and Lean 1990; Hoyt and Schatten 1993). Figure 5 shows, for the period 1894–1993, the solar irradiance forcing function we constructed in just such a manner. The quasi-periodic 11-year solar cycle is apparent, as well as an increase in the mean irradiance over the last century. By restricting our calculations to the solar-cycle band we exclude any contribution from the low frequencies associated with the long-term trend of the mean. Notice that the maximum amplitude is approximately 1 W m^{-2}, while the mean amplitude over the 100 years is closer to approximately 0.5 W m^{-2}.

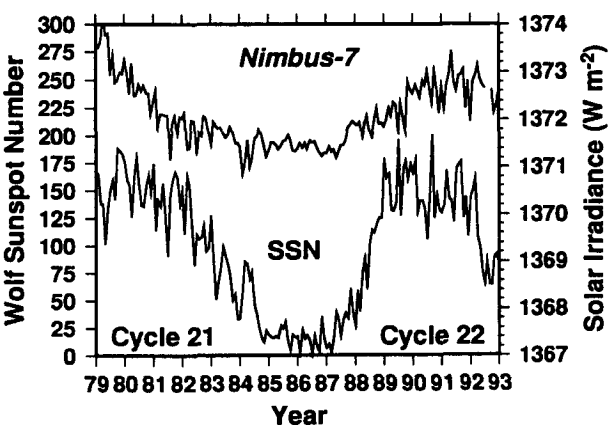

FIG. 4. Monthly mean Wolf sunspot number (SSN) and total solar irradiance from *Nimbus-7* for the period January 1979–December 1992.

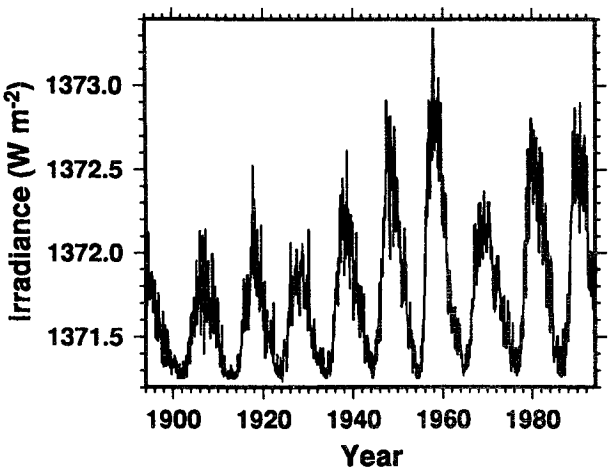

FIG. 5. Monthly solar-cycle irradiance forcing function for the period 1894–1993.

Figure 6 shows the periodogram of the annual means of the monthly time series in Fig. 5. This type of plot shows the power for each frequency in the solar-cycle band decomposed into the squares of the real and imaginary parts of the discrete Fourier transform of the forcing function $s(n)$:

$$Power(f_k) = \frac{1}{N}\left[\sum_{n=1}^{N} s(n)\cos\left(\frac{2\pi f_k t_n}{N}\right)\right]^2 + \frac{1}{N}\left[\sum_{n=1}^{N} s(n)\sin\left(\frac{2\pi f_k t_n}{N}\right)\right]^2.$$

In Fig. 6, (real part)^2 refers to the first term and (imaginary part)^2 refers to the second term in the above equation. Note that almost all the power appears in the square of the real part at the two frequencies 0.09 yr^{-1} (period of 11.11 yr) and 0.10 yr^{-1} (period of 10 yr). Decomposing the periodogram in this way allows us to see how the power is distributed between the two parts at each frequency.

This forcing function can be used to generate the space–time waveform of the response from a suitable climate model. If a coupled ocean–atmosphere climate model were used to construct such a signal, many long realizations would have to be generated to compute the ensemble mean. This would entail much time and expense. It is for this reason that we have developed an energy balance model to generate our signal waveform.

b. *The energy balance model (EBM)*

The EBM we employ is a time-stepping, two-dimensional model with realistic geography and a slab ocean mixed layer. While there is no thermodynamic sea ice model, the perennial sea ice is assigned an appropriate heat capacity. The resolution of the model (2.8125° × 2.8125°) is similar to an atmospheric GCM with T42 truncation. The model does a good job of simulating the annual cycle, in both amplitude and phase. The EBM solves the elliptic partial differential equation, known as the energy balance equation, at each time step on a regular latitude/longitude grid on a sphere. The full multigrid method with *W*-cycling and red/black line relaxation in both directions is used to find the solution (Briggs 1987; Press et al. 1992). There are 65 × 128 grid points on the sphere, including the poles, and 48 time steps per year. The diffusion coefficient is proportional to the gradient of the surface temperature (Stone 1973) and is recalculated at each time step. Previous researchers had developed an EBM, with a constant diffusion coefficient, using the multigrid method on an irregular grid, but with only *V*-cycling and point relaxation (Bowman and Huang 1991; Huang and Bowman 1992). The monthly two-dimensional albedo field is derived from five years of monthly albedo data from the Earth Radiation Budget Experiment, obtained from the National Aeronautics and Space Administration (NASA). There is no ice-albedo feedback in the model. The radiation damping coefficient over the mixed layer is 2.2 W m^{-2}. The mixed layer depth varies from 60 m at the equator to about 86 m at the Arctic and Antarctic Circles. Together these parameters yield a thermal relaxation time that varies from about 3.6 yr at the equator to about 5.1 yr at the Arctic and Antarctic Circles.

c. *The model signal space–time waveform*

Using the solar-cycle forcing function to drive our EBM, we generate a monthly space–time waveform of the solar-cycle response for the 100 years, 1894–1993. From Figs. 4 and 5 we can see the solar cycle was near its peak during 1980. Figure 7 shows the model-gen-

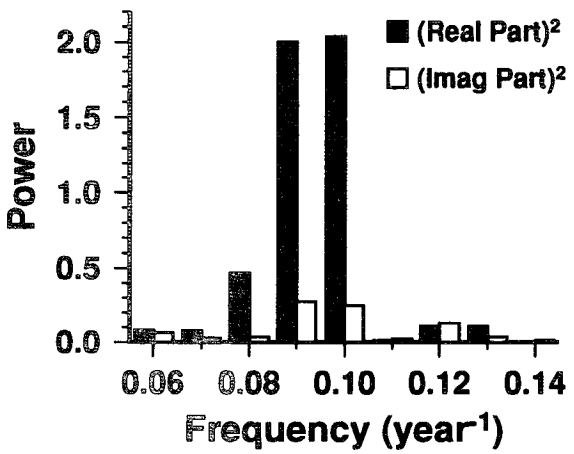

FIG. 6. Power spectrum of the solar irradiance forcing function (see Fig. 5) over the solar-cycle frequency band. The power at each frequency has been decomposed into the squares of the real and imaginary parts of the discrete Fourier transform.

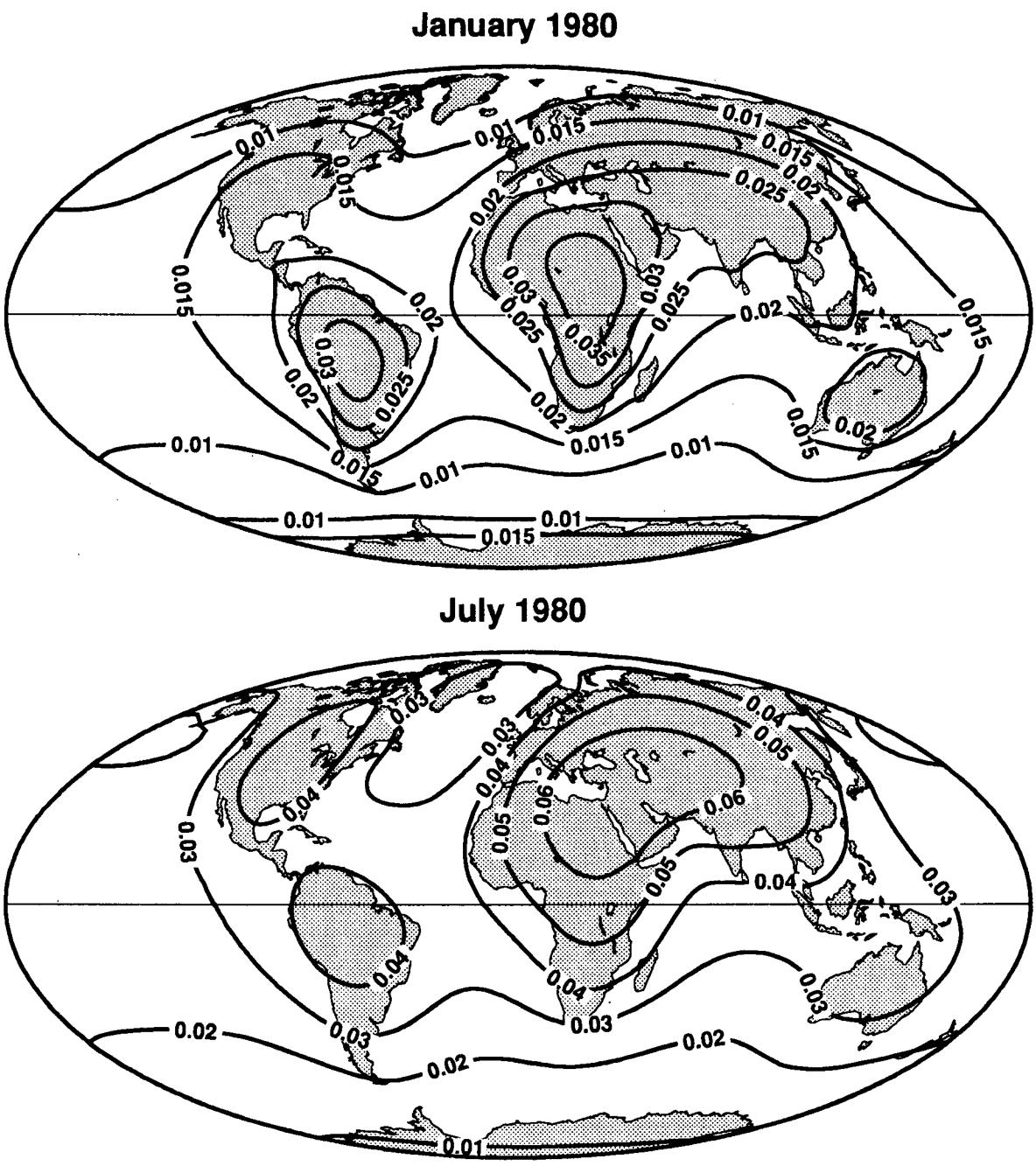

FIG. 7. The solar-cycle surface temperature response (°C) generated by the EBM for (top) January and (bottom) July 1980.

erated response (the signal) to the solar forcing function for January and July 1980. Notice that the response is smallest over the wintertime polar region. The power spectrum of the model signal from detection box 4 in the western Atlantic Ocean and from detection box 24 in central Asia is shown in Fig. 8. Comparing Fig. 8 with Fig. 6, we can see how the power in the forcing has been redistributed differently in the response of box 4 (ocean) and box 24 (land). This probably is due to the differing heat capacities of land and ocean. This type of geographically based phase shift is part of the discrimination capability of the optimal filter. For an artificial forcing function of constant amplitude (1.0 W m^{-2}) and period (10 yr), the geographical distribution of the amplitude of the response (°C) and the phase lag of the response behind the forcing (months)

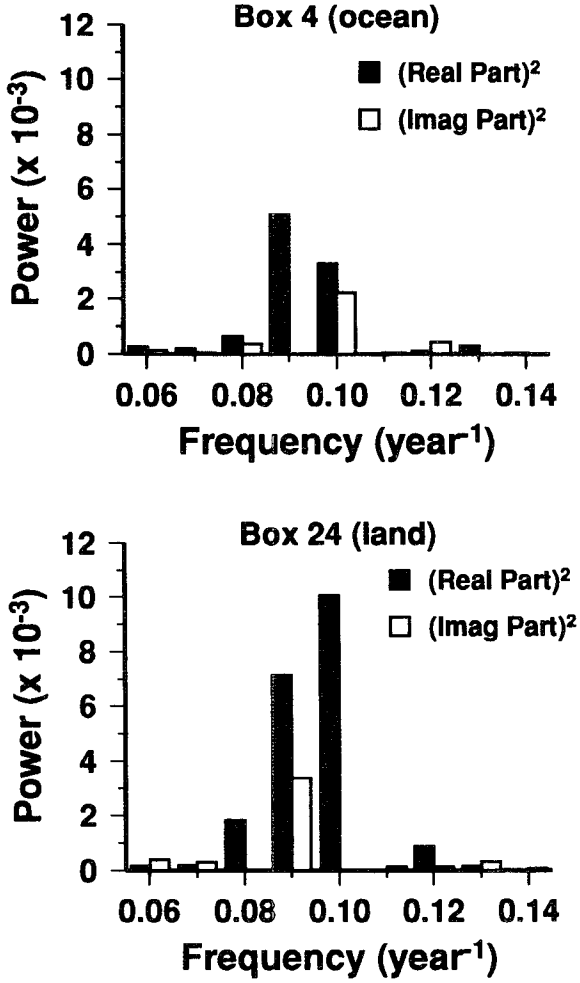

FIG. 8. Power spectrum of EBM solar-cycle response over the solar-cycle frequency band, as in Fig. 6, for (top) detection box 4 (ocean) and (bottom) box 24 (land). See Fig. 2 for locations of the detection boxes.

is shown in Fig. 9. Note that both fields vary considerably over the earth.

6. The square of the signal-to-noise ratio

We have now described all the elements needed to build the optimal filter and to calculate the scaling factor. Before we proceed we examine in more detail the square of the signal-to-noise ratio, γ^2. Equation (2.5) allows us to calculate the contribution made to γ^2 by each eigenmode and frequency of the space–time EOFs. From Fig. 8, we expect the two frequencies, 0.09 yr^{-1} and 0.10 yr^{-1}, to contribute the most to γ^2. Figure 10 shows the contributions to γ^2 for the GFDL- and MPI-derived space–time EOFs. As can be seen, the contributions are spread over many of the higher eigenmodes and do not come from only the first few eigenmodes. In these higher eigenmodes the associated eigenvalues are small and the projection of the signal onto the space–time EOF relatively large, giving a large signal-to-noise ratio. For eigenmode numbers greater than about 21 the contributions to γ^2 may be spurious. Excluding these contributions to γ^2 reduces the signal-to-noise ratio by about 10% and, hence, increases the filter pass-thru error, but at the same time results in a reduction in the filter sampling error (see below). The net result is about the same as including all the eigenmodes.

7. Errors in the optimal filter

Each of the components used in the construction of the optimal filter can contribute to the uncertainty in the final estimation of the scaling factor. We now examine the sources of these errors.

a. Filter pass-thru error

Let us assume that we have a "perfect" optimal filter available. It might seem that this filter would not be a source of error in the calculation of the scaling factor. However, our observed data, which is used as input to the filter, is composed of a signal and natural climate variability. This "perfect" filter will let the signal pass through, but at the same time, will also pass through any noise that by chance has the same space–time structure as the signal. Thus, it is impossible to build a filter that does not pass through some component of the noise. It can be shown that this error variance, which we call the filter pass-thru error, has a theoretical value that is equal to the inverse of the square of the signal-to-noise ratio, $1/\gamma^2$, where γ^2 is given by Eq. (2.5). For the GFDL-derived filter $1/\gamma^2 = 0.21$, and for the MPI filter $1/\gamma^2 = 0.05$.

b. Filter sampling error

We have seen that all optimal filters will have pass-thru error, but how does this error vary from one realization of a filter to another? The filter sampling error σ^2_{samp} gives us an estimate of how much the scaling factor might vary if we used filters made from the same signal waveform but from different realizations of noise. Assume we have 10 000 years of a coupled-model control run. By dividing this into ten 10^3-yr segments we could create ten optimal filters. Each of these ten filters is built from a different 10^3-yr sample of noise and the same signal. For each filter we calculate α. The variance of these ten values of α gives us an estimate of the filter sampling error. To estimate the filter sampling error variance we used our EBM, with noise forcing, to generate a 10^4-yr control run. From this, ten optimal filters were constructed and the filter sampling error calculated. The results are shown in Fig. 11.

c. Filter bias error

The most obvious way in which bias error will enter into the construction of the optimal filter is through the

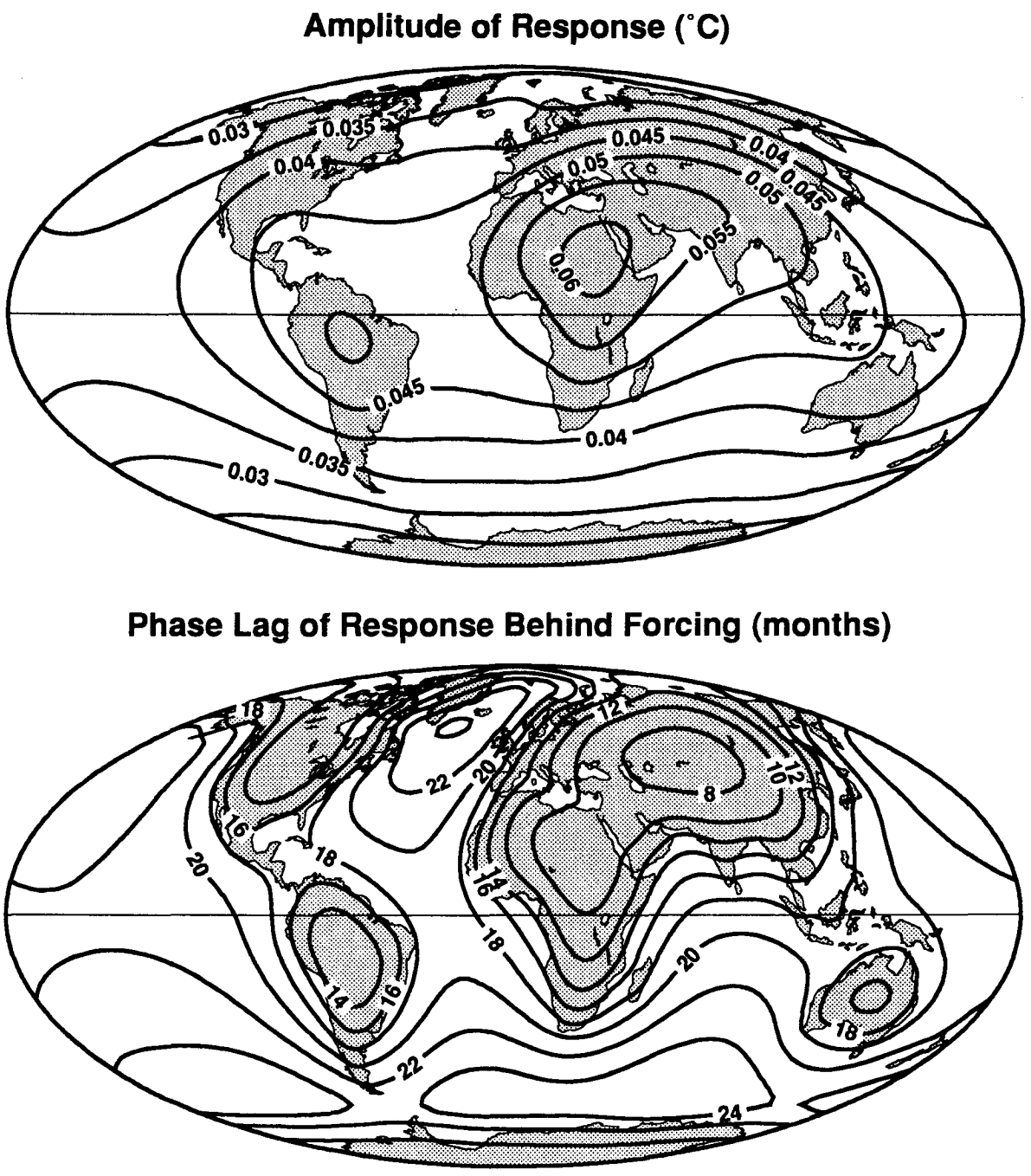

FIG. 9. (top) Contour lines show amplitude of EBM response (°C) to a forcing function of constant amplitude (1.0 W m^{-2}) and phase (0.10 yr^{-1}) for one complete cycle. (bottom) Contour lines show phase lag of EBM response (months) to the same forcing function.

model-generated climate variability. It is known that the space–time structure of the model noise is different for the GFDL and MPI models (Kim et al. 1996b). The amplitude of the noise in the MPI model is about half of that in the GFDL model. There are also differences in the spatial structure of the noise. The MPI model produces no ENSO cycle, while the GFDL model has only a weak ENSO cycle. While most of the power of ENSO events is in the higher frequencies outside of the solar-cycle frequency band, there is proba-

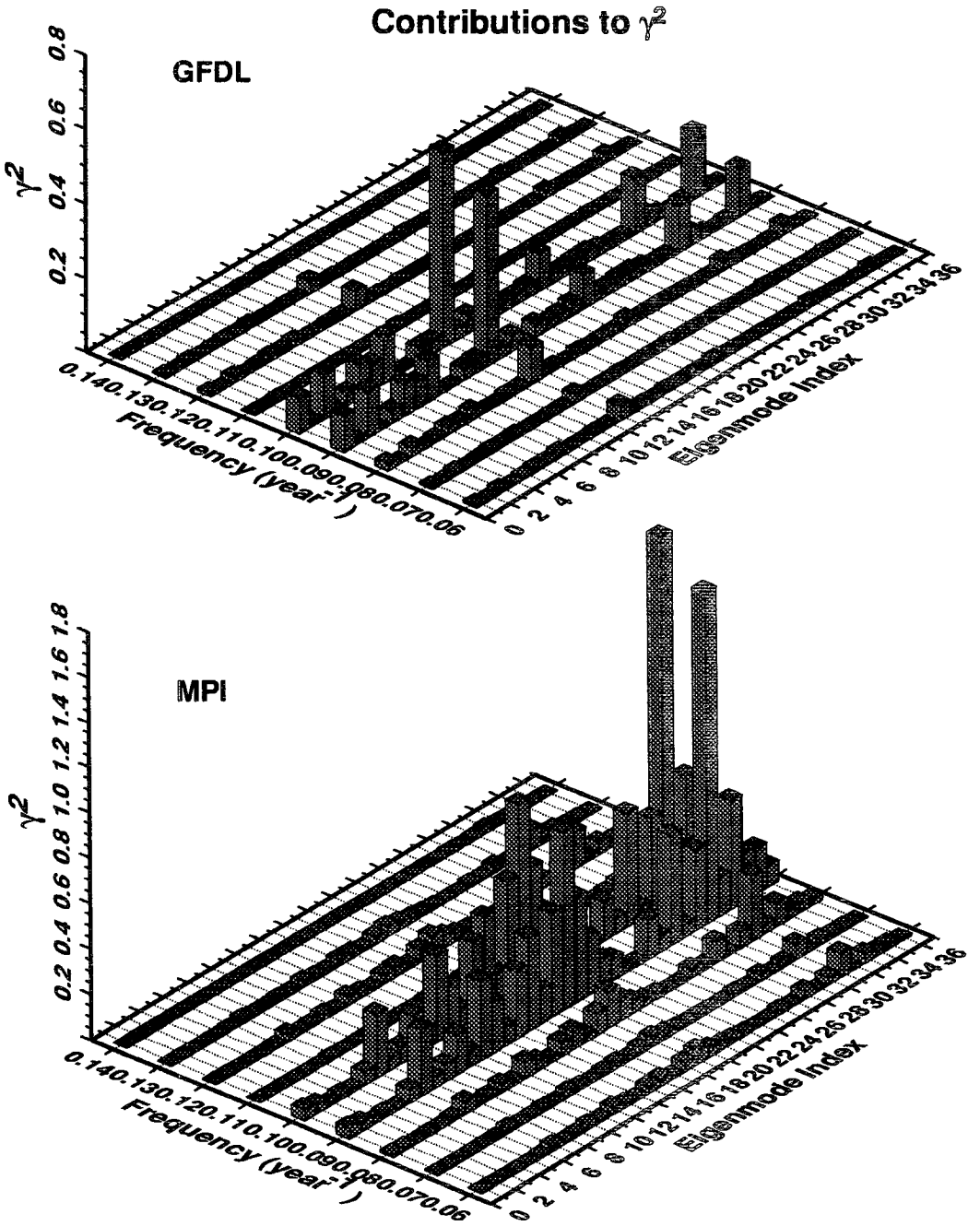

FIG. 10. Contributions made to the square of the signal-to-noise ratio γ^2 at each frequency in the solar-cycle band and for each eigenmode. (top) Contributions using the GFDL-derived space–time EOFs in Eq. (2.5) and (bottom) MPI-derived space–time EOFs.

bly some power in the solar-cycle band (Mann and Park 1994). This will result in an underestimate of the noise at those frequencies. These differences will directly affect the space–time EOFs and eigenvalues, and through them affect the calculation of the signal-to-noise ratio and the scaling factor. However, we have no way of satisfactorily estimating this type of model bias error.

The second way in which bias error will enter into the construction of the optimal filter is through the generation of the signal waveform. The filter bias error σ^2_{bias} gives us an estimate of how much the scaling fac-

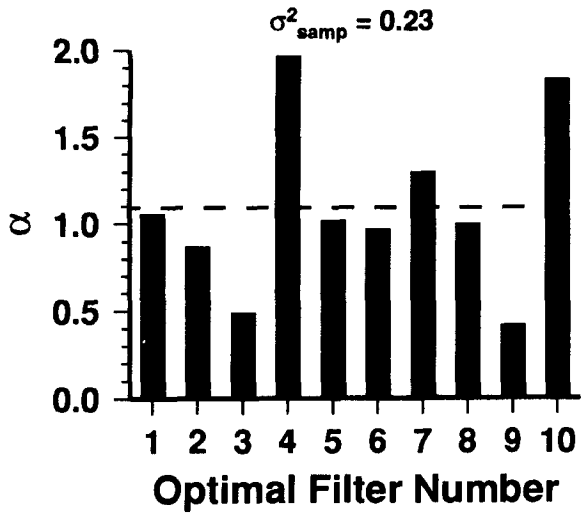

FIG. 11. Estimation of the filter sampling error by using the observed data as input to the 10 EBM-derived optimal filters. The dashed line is the mean of the resulting 10 values of the scaling factor α.

tor might vary if we used filters built from the same realization of noise, but with different signal waveforms. To estimate the filter bias error, we used the EBM to generate different signal waveforms by varying the depth of the mixed layer over a reasonable range of values. We then built a set of optimal filters using these different signals and the same noise realization (GFDL or MPI). For each filter we calculated α, as for the filter sampling error. The variance of these values of α is an estimate of the filter bias error. The results for the GFDL and MPI cases are shown in Fig. 12.

d. Total error estimate

We can now form an estimate of the total error in the scaling factor due to these errors in the construction of the optimal filter. Assuming these errors are independent of each other, we have

$$\sigma^2_{total} \approx 1/\gamma^2 + \sigma^2_{samp} + \sigma^2_{bias}. \quad (7.1)$$

For the GFDL-derived filter $\sigma^2_{total} \approx 0.21 + 0.23 + 0.009 \approx 0.45$ or $\sigma_{total} \approx 0.67$ and for the MPI-derived filter $\sigma^2_{total} \approx 0.05 + 0.23 + 0.002 \approx 0.28$ or $\sigma_{total} \approx 0.53$.

8. Scaling factor results

We now present the results of the calculation of the scaling factor and its uncertainty. For the GFDL-derived filter $\alpha = 1.30 \pm 0.67$, where the uncertainty is $\pm\sigma_{total}$. For the MPI-derived filter $\alpha = 0.93 \pm 0.53$. To evaluate the statistical significance of these results we make the assumption that α is a random variable with a Gaussian distribution. We want to test the null hypothesis that there is no solar-cycle signal in the observed data. If the null hypothesis were true, we would expect a mean value for α of zero. However, we assume a signal is present in the observed data and that it has the same phase as the model-generated signal, which implies a scaling factor greater than zero. This allows us to use a one-tailed test of the null hypothesis. In the GFDL case, the value for α of 1.30 is $1.94\sigma_{total}$ from zero. This implies a risk in rejecting the null hypothesis of 2.6%, or a confidence level of 97.4%. In the MPI case, the value for α of 0.93 is $1.75\sigma_{total}$ from zero. This implies a risk of rejecting the null hypothesis of 4.0%, or a confidence level of 96.0%.

These are significant results, but we must remember there is unestimated bias error in the calculation of the noise covariance matrix and the signal. The estimate of the filter pass-thru error is theoretical and, hence, only strictly valid when there are a large number of estimates of α. In summary, these results imply a highly significant value for the scaling factor α, but the estimated error is probably underestimated.

a. Results from different regions

The results presented so far have used all 36 of the detection boxes in all the computations. However, we can repeat the same procedures using subsets of the 36 boxes from different regions of the earth. For each subset of boxes, a complete set of space–time EOFs and eigenvalues can be calculated and, hence, an optimal filter constructed. The results for the global case and six regions are presented in Tables 1 and 2, as well as in Fig. 13. In the estimation of the total error, we have assumed that the filter sampling and bias errors for each

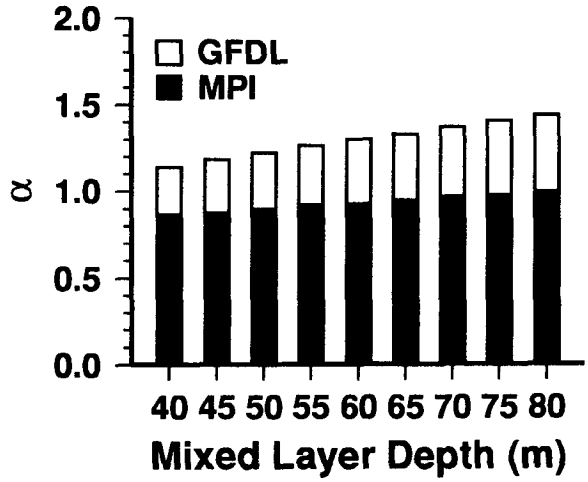

FIG. 12. Estimation of the filter bias error. For each mixed layer depth a new signal waveform was generated and a new optimal filter built. The observed data was used as input to the different filters and the scaling factor calculated. Results shown for both GFDL and MPI based optimal filters.

TABLE 1. Calculated statistics based on the optimal filter built from the GFDL coupled-model noise and the EBM signal. The column labeled γ_{nopt} is the nonoptimized signal-to-noise ratio computed by using Eq. (2.7). The column labeled σ_{total} is the total error estimate for the scaling factor α and is calculated using Eq. (7.1). The column labeled "c.l." is the confidence level of rejecting the null hypothesis that no response to the solar cycle is present in the observed data.

Region	Boxes	γ_{nopt}	γ	$1/\gamma^2$	σ_{total}	α	c.l.
Global	36	1.41	2.18	0.21	0.67	1.30	97.4
Tropics	20	1.78	2.02	0.25	0.70	1.11	94.3
Extratropics	16	0.82	1.11	0.81	1.02	1.14	86.9
N. Hemisphere	24	1.16	1.87	0.29	0.73	0.89	88.9
S. Hemisphere	12	1.43	1.78	0.32	0.75	1.59	98.3
W. Hemisphere	16	1.21	1.81	0.31	0.74	1.16	94.2
E. Hemisphere	20	1.20	1.83	0.30	0.73	1.34	96.6

subset of boxes are the same as for all 36 boxes. Notice that the error estimates are smallest for the global results. Also, notice that the error estimates for the MPI results are smaller than for the GFDL results. This is due to the smaller amplitude of the climate noise in the MPI model. In Table 3 we show the results from an optimal filter made from the 10^4 years of EBM noise used to estimate the filter sampling error. The consistent results from the global and regional calculations give us added confidence in the power and validity of the optimal filtering method.

b. Lagged data test

As a test of our method we can ask the following question: what result would we expect if we lagged the observed data and used it as input to the optimal filter? For a lag of zero we would expect α to be at its maximum value. As the lag is increased, α should decrease until the lag is equal to one-half of the cycle length. At this point, the signal in the input data would be exactly out of phase with the model-generated signal used in the filter. This behavior should be true for both positive and negative lags. Since we have only 100 years of observed data, we can perform this test by using the middle 90 years (1899–1988) of the data and lagging it from −5 to +5 yr. The results are shown in Fig. 14 for both the GFDL- and MPI-derived filters.

c. Estimated global response to the solar cycle

Using Eq. (2.14), we can estimate the global annual mean response of the climate system to the solar-cycle forcing. We simply multiply our estimated value of the scaling factor α by the global annual mean temperature response T^{signal} produced by the EBM. This is shown in Fig. 15 for estimates of α based upon the GFDL, MPI, and EBM derived optimal filters. The response is shown as the temperature departure from the 100-yr global annual mean of the model.

9. Discussion

There are a number of assumptions that are made about the ingredients needed to build the optimal filter. In addition, to estimate the signal in the observed data, there are assumptions made about the data. These assumptions are listed below.

a. Assumptions about the *forcing function* and *signal waveform*

1) The forcing function used to produce the EBM signal is an adequate representation of the solar variability over the last 100 years.

2) The space–time structure of the solar cycle signal produced by the EBM is a good representation of the true solar-cycle response in the observed data.

b. Assumptions about the *climate variability*

1) The space–time structure of the real climate variability is well represented by the GFDL and MPI coupled climate models.

2) The length of the control runs does not seriously bias the results.

TABLE 2. Same as Table 1 except for use of MPI coupled-model noise.

Region	Boxes	γ_{nopt}	γ	$1/\gamma^2$	σ_{total}	α	c.l.
Global	36	2.58	4.63	0.05	0.53	0.93	96.0
Tropics	20	2.49	4.01	0.06	0.54	1.09	97.8
Extratropics	16	1.40	1.94	0.27	0.71	0.96	91.2
N. Hemisphere	24	2.20	4.07	0.06	0.54	0.69	90.0
S. Hemisphere	12	1.69	2.84	0.12	0.59	1.52	99.5
W. Hemisphere	16	1.76	3.28	0.09	0.57	0.41	76.4
E. Hemisphere	20	2.15	3.94	0.06	0.54	1.20	98.7

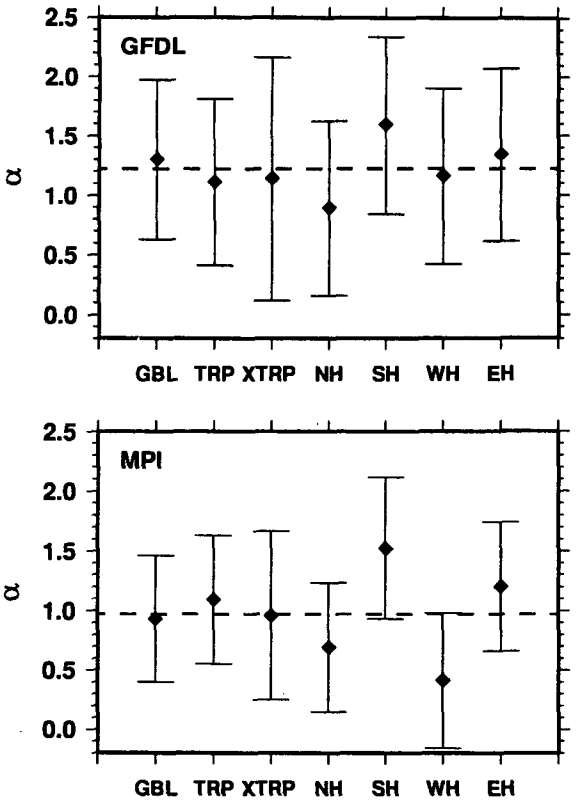

FIG. 13. Calculation of the scaling factor for the global and six regional subsets of detection boxes. (top) Results for the GFDL-based optimal filter and (bottom) for the MPI-based optimal filter. The dashed lines show the means of the seven values of the scaling factor calculated.

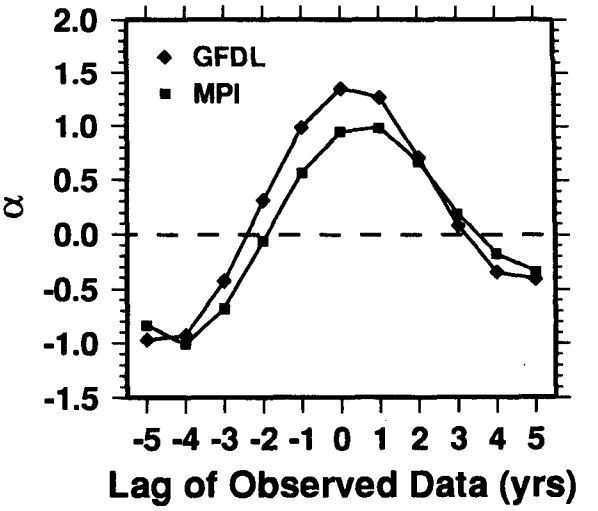

FIG. 14. The results of the lagged data test. Ninety years of observed data was lagged and used as input to the GFDL- and MPI-based optimal filters.

c. Assumptions about the *historical data record*

1) The 100-year historical record is sufficiently long.
2) The lack of coverage over large areas of the earth does not bias the results.

Some of these assumptions are untestable and some can be improved upon in the future. However, at the current stage of our research these assumptions are unavoidable. To increase the confidence in their validity, we propose the following:

1) the use of coupled ocean–atmosphere climate models to calculate the space–time waveform of the response to the solar cycle;
2) the generation of long control runs from other coupled ocean–atmosphere climate models;
3) examining the possibility of including some long historical surface temperature records of different lengths. This would allow us to utilize the long (>100 yr) instrumental records from several sites around the world.

Two possible ways to improve the performance of the method are

1) the use of cyclostationary EOFs (Huang and North 1996; Kim et al. 1996a);
2) calculation of the optimal months to use in the covariance matrix to minimize its eigenvalues and, therefore, maximize the signal-to-noise ratio.

The results we have presented here demonstrate the potential power of the method of optimal space–time filtering for detection problems of this type. We are very confident that the method has succeeded in de-

TABLE 3. Same as Table 1 except for use of EBM noise.

Region	Boxes	γ_{nopt}	γ	$1/\gamma^2$	σ_{total}	α	c.l.
Global	36	1.68	2.08	0.23	0.68	1.13	95.2
Tropics	20	1.70	1.99	0.25	0.69	0.92	90.8
Extratropics	16	1.26	1.55	0.42	0.81	0.77	82.9
N. Hemisphere	24	1.47	1.90	0.28	0.72	0.93	90.2
S. Hemisphere	12	1.41	1.52	0.43	0.81	1.45	96.3
W. Hemisphere	16	1.20	1.55	0.42	0.81	1.29	94.4
E. Hemisphere	20	1.33	1.56	0.41	0.80	0.77	83.2

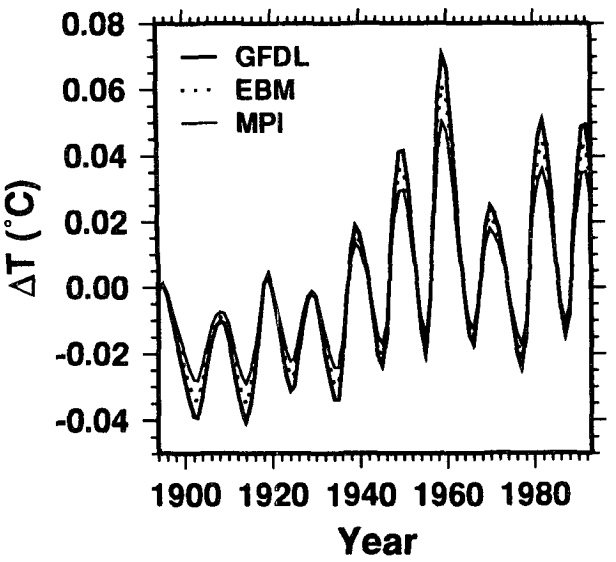

FIG. 15. Estimated global annual response to the solar-cycle forcing. The scaling factor was calculated using the GFDL-, MPI-, and EBM-based optimal filters and the observed surface temperature data.

tecting the solar-cycle response in the observed data. These results obviously will be subject to future revision with improvements in formulating the components comprising the optimal filter.

Acknowledgments. We dedicate this work to the memory of Tzvi Gal-Chen. This work was supported by a grant from the National Institute for Global Environmental Change (Department of Energy) through its South Central Regional Office at Tulane University. We are grateful to its director, Professor Robert Watts, for his support and encouragement. G. R. North has also benefitted from a DOE CHAMMP grant. The Department of Energy does not necessarily endorse the findings of this study.

We have benefitted from many helpful discussions with our colleagues K. Bowman, T. Crowley, K.-Y. Kim, W. Pennell, and S. Shen. We thank T. Bell, T. Wigley, and F. Zwiers for helpful criticisms of an earlier version of this work. We thank H. von Storch, G. Hegerl, and M. Dorn for making the MPI control run dataset available to us. Some of the ideas for this work were generated during a 1994 visit by G. R. North to Prof. R. Gurney's group at the University of Reading.

REFERENCES

Ardanuy, P. E., H. L. Kyle, and D. Hoyt, 1992: Global relationships among the earth's radiation budget, cloudiness, volcanic aerosols, and surface temperature. *J. Climate,* **5,** 1120–1139.

Barnett, T. P., 1991: An attempt to detect the greenhouse-gas signal in a transient GCM simulation. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. E. Schlesinger, Ed., Elsevier, 559–568.

——, M. E. Schlesinger, and X. Jiang, 1991: On greenhouse-gas signal detection strategies. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. E. Schlesinger, Ed., Elsevier, 537–558.

Bell, T. L., 1982: Optimal weighting of data to detect climatic change: Application to the carbon dioxide problem. *J. Geophys. Res.,* **87,** 11 161–11 170.

——, 1986: Theory of optimal weighting of data to detect climatic change. *J. Atmos. Sci.,* **43,** 1694–1710.

Bowman, K. P., and J. Huang, 1991: A multigrid solver for the Helmholtz equation on a semiregular grid on the sphere. *Mon. Wea. Rev.,* **119,** 769–775.

Briggs, W. L., 1987: *A Multigrid Tutorial.* Society for Industrial and Applied Mathematics, 90 pp.

Crowley, T. J., and K.-Y. Kim, 1993: Towards development of a strategy for determining the origin of decadal–centennial scale climate variability. *Quat. Sci. Rev.,* **12,** 375–385.

Cubasch, U., K. Hasselmann, H. Höck, E. Maier-Reimer, U. Mikolajewicz, B. D. Santer, and R. Sausen, 1992: Time-dependent greenhouse warming computations with a global coupled ocean–atmosphere model. *Climate Dyn.,* **8,** 55–69.

Eddy, J. A., R. L. Gilliland, and D. V. Hoyt, 1982: Changes in the solar constant and climatic effects. *Nature,* **300,** 689–693.

Foukal, P., and J. Lean, 1990: An empirical model of total solar irradiance variation between 1874 and 1988. *Science,* **247,** 556–558.

Hansen, J. E., and A. A. Lacis, 1990: Sun and dust versus greenhouse gases: An assessment of their relative roles in global climate change. *Nature,* **346,** 713–719.

——, ——, R. Ruedy, and M. Sata, 1992: Potential climate impact of Mount Pinatubo eruption. *Geophys. Res. Lett.,* **19,** 215–218.

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology over the Tropical Oceans,* D. B. Shaw, Ed., Royal Meteorological Society, 251–259.

——, 1993: Optimal fingerprints for the detection of time-dependent climate change. *J. Climate,* **6,** 1957–1971.

——, and Coauthors, 1995: Detection of anthropogenic climate change using a fingerprint method. Rep. 168, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 20 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstrasse 55, D-20146 Hamburg, Germany.]

Hegerl, G. C., and G. R. North, 1995: Statistically optimal approaches to detecting anthropogenic climate change. Rep. 167, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 18 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstrasse 55, D-20146 Hamburg, Germany.]

——, H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones, 1994: Detecting anthropogenic climate change with an optimal fingerprint method. Rep. 142, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 59 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstrasse 55, D-20146 Hamburg, Germany.]

Hoyt, D. V., and K. H. Schatten, 1993: A discussion of plausible solar irradiance variations, 1700–1992. *J. Geophys. Res.,* **98,** 18 895–18 906.

Huang, J., and K. P. Bowman, 1992: The small ice cap instability in seasonal energy balance models. *Climate Dyn.,* **7,** 205–215.

——, and G. R. North, 1996: Cyclic spectral analysis of fluctuations in a GCM simulation. *J. Atmos. Sci.,* **53,** 370–379.

Jones, P. D., and K. R. Briffa, 1992: Global surface air temperature variations during the twentieth century: Part 1. Spatial, temporal and seasonal details. *The Holocene,* **2,** 165–179.

Kim, K.-Y., and G. R. North, 1993: EOF analysis of surface temperature field in a stochastic climate model. *J. Climate,* **6,** 1681–1690.

——, ——, and J.-P. Huang, 1996a: EOFs of one-dimensional cyclostationary time series: Computations, examples, and stochastic modeling. *J. Atmos. Sci.,* **53,** 1007–1017.

——, ——, and G. C. Hegerl, 1996b: Comparisons of the second-moment statistics of climate models. *J. Climate,* **9,** 2201–2221.

Lee, R. B., III, M. A. Gibson, R. S. Wilson, and S. Thomas, 1995: Long-term total solar irradiance variability during sunspot cycle 22. *J. Geophys. Res.*, **100,** 1667–1675.

Mann, M. E., and J. Park, 1994: Global-scale modes of surface temperature variability on interannual to century timescales. *J. Geophys. Res.*, **99,** 25 819–25 833.

Marshall, S., R. J. Oglesby, J. W. Larson, and B. Saltzman, 1994: A comparison of GCM sensitivity to changes in CO_2 and solar luminosity. *Geophys. Res. Lett.*, **21,** 2487–2490.

North, G. R., 1984: Empirical orthogonal functions and normal modes. *J. Atmos. Sci.*, **41,** 879–887.

——, and K.-Y. Kim, 1995: Detection of forced climate signals. Part II: Simulation results. *J. Climate,* **8,** 409–417.

——, J. G. Mengel, and D. A. Short, 1983: Climatic response to a time varying solar constant. *Weather and Climate Responses to Solar Variations,* B. M. McCormac, Ed., Colorado Associated University Press, 243–255.

——, K.-J. J. Yip, L.-Y. Leung, and R. M. Chervin, 1992: Forced and free variations of the surface temperature field in a general circulation model. *J. Climate,* **5,** 227–239.

——, K.-Y. Kim, S. S. P. Shen, and J. W. Hardin, 1995: Detection of forced climate signals. Part I: Filter theory. *J. Climate,* **8,** 401–408.

Parker, D. E., P. D. Jones, C. K. Folland, and A. Bevan, 1994: Interdecadal changes of surface temperature since the late nineteenth century. *J. Geophys. Res.*, **99,** 14 363–14 399.

Press, W. H., S. A. Teukolsky, W. T. Vettering, and B. P. Flannery, 1992: *Numerical Recipes in FORTRAN. The Art of Scientific Computing.* 2d ed. Cambridge University Press, 963 pp.

Schneider, S. H., and C. Mass, 1975: Volcanic dust, sunspots, and temperature trends. *Science,* **190,** 741–746.

Schwabe, H., 1844: Solar observations during 1843. *Astron. Nachr.,* **21,** 233.

Selin, I., 1965: *Detection Theory.* Princeton University Press, 128 pp.

Stone, P. H., 1973: The effect of large-scale eddies on climatic change. *J. Atmos. Sci.,* **30,** 521–529.

Stouffer, R. J., S. Manabe, and K. Y. Vinnikov, 1994: Model assessment of the role of natural variability in recent global warming. *Nature,* **367,** 634–636.

von Storch, J.-S., 1994: Interdecadal variability in a global coupled model. *Tellus,* **46A,** 419–432.

Wigley, T. M. L., and T. P. Barnett, 1990: Detection of the greenhouse effect in the observations. *Climate Change. The IPCC Scientific Assessment,* J. T. Houghton, G. J. Jenkins, and J. J. Ephraums, Eds., Cambridge University Press, 239–255.

——, and S. C. B. Raper, 1990: Natural variability of the climate system and detection of the greenhouse effect. *Nature,* **344,** 324–327.

Willson, R. C., and H. S. Hudson, 1991: The sun's luminosity over a complete solar cycle. *Nature,* **351,** 42–44.