# THE 11-YEAR SOLAR CYCLE IN THE LOWER STRATOSPHERE EXTRACTED BY THE EMPIRICAL MODE DECOMPOSITION METHOD

K.T. Coughlin[1] and K.K. Tung[1]

[1] *University of Washington, Box 352420, Seattle, WA 98195, USA*

*Submitted to Advances in Space Research, November 2002.*

**ABSTRACT**

We apply a novel method to extract the solar cycle signal from the atmospheric data in the lower stratosphere and troposphere. An alternative to traditional analysis is a nonlinear EMD (empirical mode decomposition) method. This method is adaptive and therefore highly efficient at identifying embedded structures, even those with small amplitudes. Using this analysis, the geopotential height in the Northern Hemisphere can be completely decomposed into five non-stationary temporal modes including an annual cycle, a QBO signal, an ENSO-like mode, a solar cycle signal and a trend. High correlations with the sunspot cycle unambiguously establish that the fourth mode is an 11-year solar cycle signal.

**Introduction**

Although there have been many reports of the 11-year solar cycle in atmospheric data, there is considerable debate on the spatial and temporal extent in which the atmosphere is influenced and on the validity of the statistical significance of these claims. A major problem is the shortness of the data record, which prevents a straightforward extraction of an 11-year signal in the energy spectrum of dynamical variables. In chemical records, like ozone, the solar cycle signal is much clearer (Hood et al. 1993, Haigh 1994, McCormack and Hood 1996, Hood 1997). Ozone heating involves wavelengths shorter than 200 nm. Unlike longer wavelengths, irradiance at these frequencies changes significantly between solar minima and solar maxima (WMO 1987). Changes of UV irradiance thus represents a potentially important influence on the stratospheric circulation. In tropospheric and lower stratospheric dynamical variables, however, the Fourier mode with an 11 year period has very little power.

This small amplitude solar cycle signal has been found in the quiescent regions of the data set. For the extratropical middle atmosphere, this means during summer time (away from the time periods of most variability), over the mid-latitude Pacific (away from the spatial areas of the most variability), and during the westerly phases of the QBO (away from the periods of the most dynamical disturbance). Originally, Labitzke (1987) discovered an association between the 30-mb winter mean polar temperature and the Sunspot numbers during the westerly phase of the equatorial QBO, with the phase defined by the equatorial 50-mb zonal wind. During the more disturbed easterly phase of the QBO though, she found that the correlations are inexplicably negative. Many other studies have shown similar results (Labitzke and van Loon 1988, van Loon and Labitzke 1994, Ruzmaikin and Feynman 2002). During these winter months, the solar cycle signal is weak compared to large atmospheric variations and the signal is more difficult to find (Labitzke 1987, Labitzke and van Loon 1988, Kodera 1991, Dunkerton and Baldwin 1992, van Loon and Labitzke 1994). However, during summer times, Labitzke and van Loon (1989, 1990) demonstrate that stratospheric data has large positive correlations with the solar cycle without grouping the data according to the phases of the QBO. And when considering annual means of the stratospheric data, positive correlations are found with the solar cycle over the eastern Pacific and the United States in the mid-latitudes (Labitzke and van Loon 1992, 1994, Labitzke 2001). This is the spatial area that contains the least amount of atmospheric variability.

During these quiescent regions of time and space, the stratosphere is positively correlated with the solar cycle. What remains is to observe the solar cycle signal in a more general and global sense, without resorting to any specific grouping of the data.

In the present work, we shall reexamine the association between the 11-year solar cycle and the lower stratosphere using unstratified data, i.e. no grouping according to the seasons, spatial regions or the phases of equatorial QBO. The passage of time and the recent developments in data analysis have given us two advantages: There is now one more solar cycle added to the time series since Labitzke and van Loon (1988). And the recent development of the Empirical Mode Decomposition (EMD) method (Huang et al., 1998) is suitable for isolating small amplitude embedded signals. With the EMD analysis we are able to perform a complete decomposition of geopotential height resulting in 5 modes and a trend. The fourth mode is a clear solar cycle signal, highly correlated with the sun spot time series. Statistical tests are also presented here which show that the mode itself is significant in the decomposition and that, because of the clarity of the decomposition and the addition of the most recent data, the correlation is significant.

**Data**

NCEP Daily Global Analyses data is provided by the NOAA-CIRES Climate Diagnostics Center in Boulder, Colorado, USA, from their Web site at http://www.cdc.noaa.gov/ (Kalnay et al. 1996). It is used to create the geopotential height timeseries. The data used ranges from January 1958 to July 2002. The total geopotential height at 17 levels, from 1000 mb up to 10 mb, and at latitudes from the equator to the North pole are used throughout this analysis. The total geopotential height is averaged from 20N to 90N at each level. This averaged Northern Hemisphere time series is decomposed using the EMD method and the results are analyzed here. To investigate the latitudinal dependence, zonally averaged latitudinal strips of the total geopotential height are decomposed and analyzed as well. The latter results will be presented in a later paper.

The sunspot record is provided by the SIDC, RWC Belgium, World Data Center for the Sunspot Index, Royal Observatory of Belgium, at the web address: http://sidc.oma.be/ (Cugnon 1999). It consists of monthly data from January 1749 to September 2002. Only the data from 1958 to 2002 is utilized in this paper.

**Methodology**

In order to overcome the difficulty of searching for a small amplitude signal embedded in noisy atmospheric data, a nonlinear, non-stationary method called the empirical mode decomposition, or EMD, method (Huang et al. 1998) is used. This method is empirical because the local characteristic time scales of the data itself are used to decompose the timeseries. The number of modes and frequencies of each mode are inherently determined by these time scales. The structure of each mode is determined by the natural amplitude variations in the timeseries. Higher frequency oscillations are captured in the first mode and subsequent modes have lower average frequencies. These temporal modes are called intrinsic mode functions, or IMFs, and are not constrained to have constant frequencies.

Although a powerful method, EMD must be used cautiously. One difficulty encountered when using the EMD method is the influence of the end points. The EMD method relies on envelope calculations derived from a cubic spline interpolation between local extrema. The mean of the derived envelope is subtracted from the original time series and then the envelope of the residual is again found using a spline interpolation. The mean of that envelope is then subtracted from the residual and the process is repeated. Once the mean of the envelope is close enough to zero, the first IMF results. After the first IMF is found and subtracted from the original timeseries, the procedure is repeated to find the second IMF. These steps are iterated until there is nothing left in the residual. Because the structure of the subtracted means depends on the shape of the envelopes which are created using a cubic spline, the decomposition will be somewhat dependent on the spline calculation. However, splines are notoriously sensitive to end points. It is important to make sure that end effects do not propagate into the interior solution. Here this problem is dealt with by extending both the beginning and end of the data by the addition of typical waves,

$$\text{wave extension} = \text{A} \sin\left(\frac{2\pi t}{P} + \text{phase}\right) + \text{local mean}.$$

The typical amplitude, A, and period, P, are determined by the nearest local extrema.

$$\begin{aligned} \text{A}_{beginning} &= \|\text{max}(1) \text{ - } \text{min}(1)\| \\ \text{A}_{end} &= \|\text{max}(\text{N}) \text{ - } \text{min}(\text{N})\| \\ \text{P}_{beginning} &= 2\ \|\text{time}(\text{max}(1)) \text{ - } \text{time}(\text{min}(1))\| \\ \text{P}_{end} &= 2\ \|\text{time}(\text{max}(\text{N})) \text{ - } \text{time}(\text{min}(\text{N}))\| \end{aligned}$$

where max(1) and min(1) are the first two local extrema in the time series and max(N) and min(N) are the last two local extrema. This calculation takes place every iteration so that the additional waves are continually changing in amplitude and frequency. Because the additional waves have the same amplitude as the nearest oscillations, the addition of these waves causes the slope of the envelope to tend toward zero at the beginning and end of the timeseries. This technique eliminates large swings in the spline calculation that may otherwise form when the slope is artificially forced to zero. Three points of approximately the same amplitude are needed in order to flatten the spline. Any more than this can adversely affect the low frequency modes by artificially leveling the ends of any long term trend present in the data.

Another important component to the analysis as implemented here is the inclusion of the annual cycle. Unlike more traditional methods, where a twelve month climatology is subtracted from the data and only the anomalies are studied, this method decomposes the total atmospheric signal. In fact, the linearized removal of climatology can actually degrade the nonlinear analysis. Here, the annual cycle is an important component and is retrieved as the first IMF in the decomposition. Since we are not interested in intraseasonal variations, a three-month running average can be applied to the time series to damp month-to-month variations. Near the surface, in the lower troposphere, where there is typically more variability more aggressive smoothing (5 and 7 point running averages) should be applied to damp higher frequency variations. If the high frequencies are retained, mode splitting becomes a problem. The higher month-to-month frequencies are either intermittent or appear intermittent. In either case, the resulting modes contain a mixture of frequencies and these split modes are much more difficult to interpret. The minimal amount of smoothing is therefore performed so that the first mode will contain the annual cycle.

**Results**

Figure 1 shows the complete decomposition of the 30-mb geopotential height spatially averaged over all longitudes and latitudes from 20N to 90N. Although the EMD modes are empirically determined, they remain locally orthogonal to one another. The time series is separated into 5 modes and a trend. The first IMF is the annual cycle. The second IMF has an average period of 28 months and is anti-correlated with the equatorial QBO. The third ENSO-like mode has an average period of 4 years and the fourth mode, with an average period of 11 years, is highly correlated with the solar cycle. The trend indicates cooling in the stratosphere over time. The fourth IMF has the same frequency as the solar cycle proxy and also contains amplitudes comparable to previous estimates of geopotential height solar cycle variations. In our analysis, the peak-to-peak amplitude variation of the solar cycle signal is about 50 meters for the geopotential height averaged over the Northern Hemisphere. In Labitzke's 2001 paper a 3-year running mean of geopotential height at 30 mb and at a selected location, 3N/150W, from 1958 to 1997 has peak-to-peak amplitude variations ranging from 40 to 90 meters. The significance of this IMF can be verified by comparing its average power to the average power of a pure noise decomposition. For a random time series, the EMD method produces modes which have an average power that decays as the average period increases.

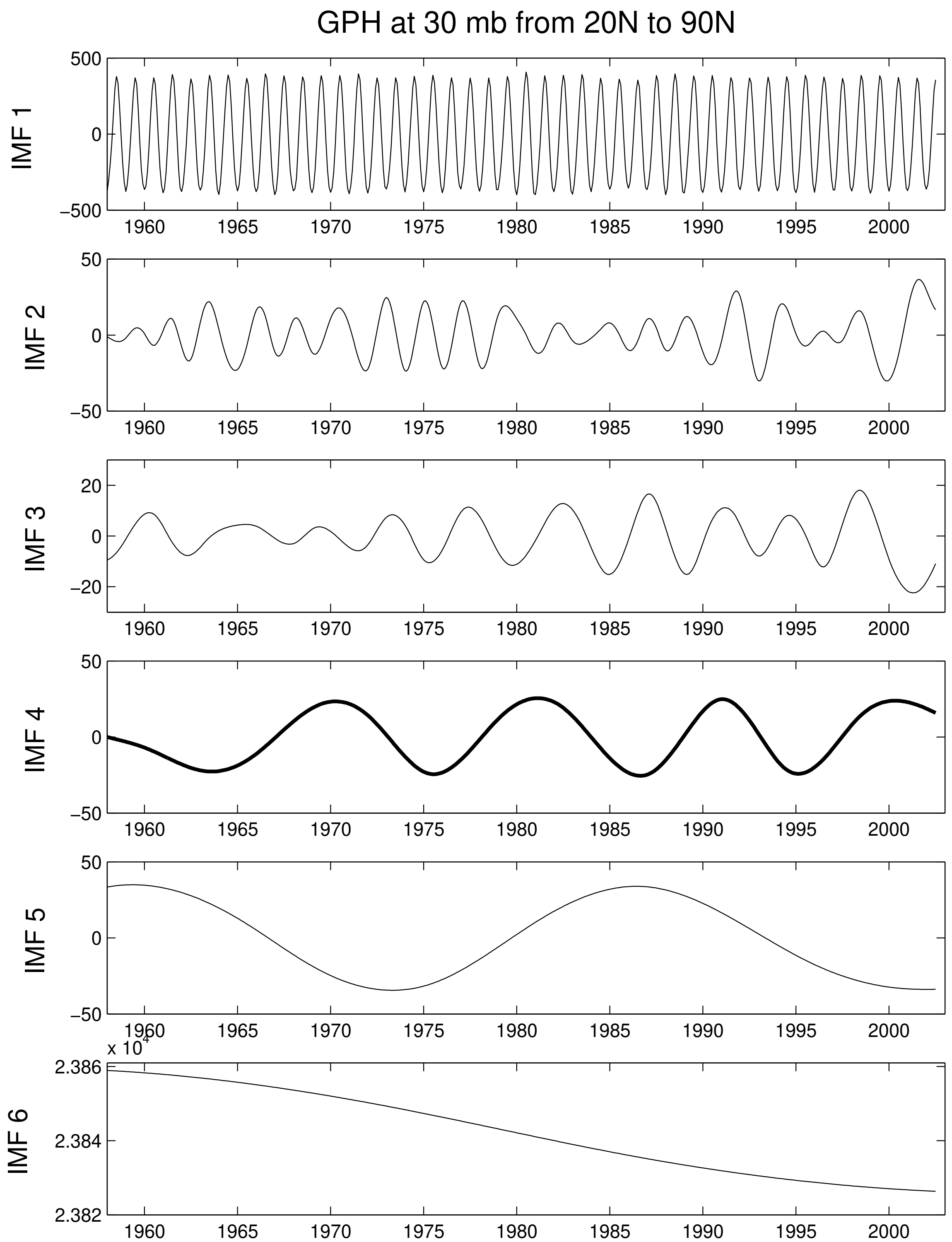


Fig. 1. The total geopotential height at 30 mb from 20N to 90N is decomposed into 5 modes and a trend. The first mode is the annual cycle. The second mode has an average frequency of 27 months and is anti-correlated with the equatorial QBO. The third ENSO-like mode has an average period of 4 years and the fourth mode is highly correlated with the 11-year solar cycle. The trend indicates cooling in the stratosphere over time.

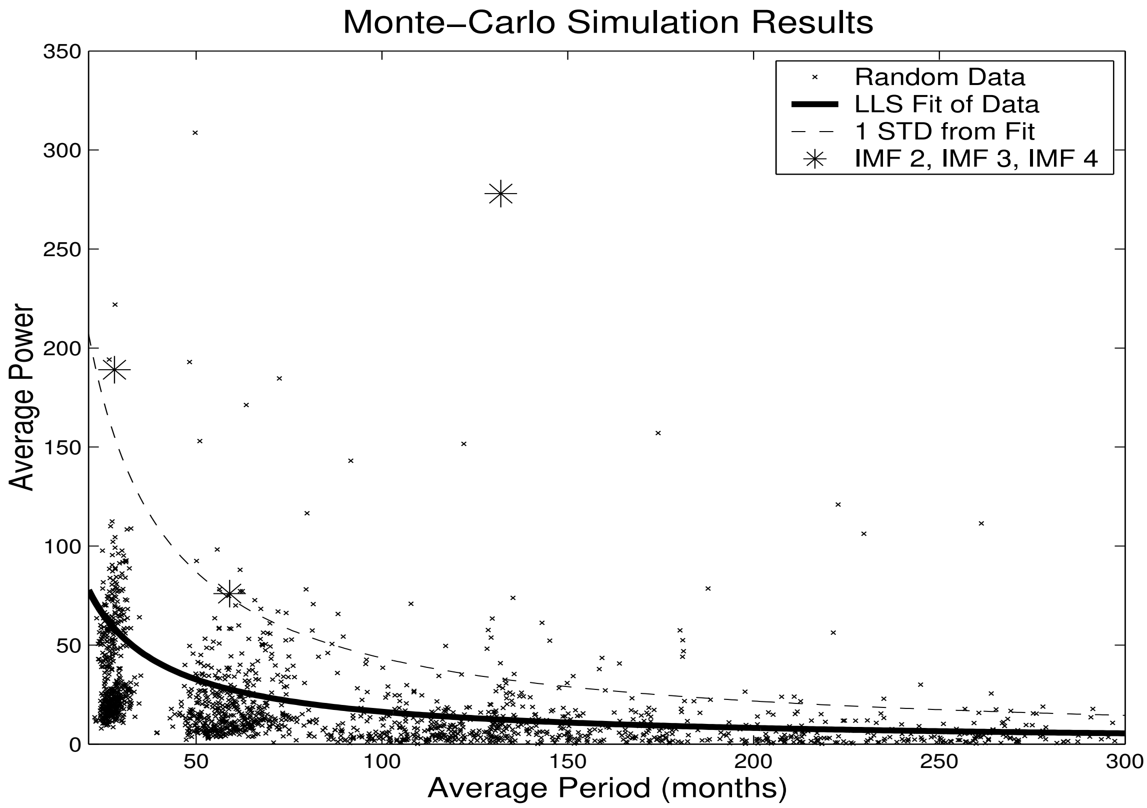


Fig. 2. 500 randomly generated time series are decomposed using the EMD method. The average power versus the average period of each of the subsequent modes are marked with an 'x'. The solid line represents the least squares fit of these points and the dashed line is one standard deviation from it. The stars represent the average power versus average period of IMF 2, IMF 3 and IMF 4 in the decomposition of the 30-mb total geopotential height.

In figure 2 the average power of modes 2, 3 and 4 in the 30-mb geopotential height are compared to modes 2, 3 and 4 of 500 randomly generated time series. The random time series are created by adding the 30-mb climatology to appropriately scaled noise. See Appendix A for a description of the noise normalization. For randomly generated time series, the EMD method exhibits a period doubling phenomenon so that each mode tends to have an average period of about twice that of the previous mode. This creates a clustering of the modes about certain periods in the Monte-Carlo plot (figure 2). It can also be seen that the average power of the random modes varies inversely to the average period (power $\sim 1/$period) so that the modes quickly decrease in power as the average period increases. The solid line represents the linear least squares fit of the random decompositions. The dashed line is one standard deviation above the mean. The three stars represent the average power of IMF 2, IMF 3 and IMF 4 of the 30-mb geopotential height at average periods of 28, 59 and 132 months respectively. These modes all fall above the confidence interval and therefore are significant. In particular, the fourth mode lies so far above the noise level that not one of the 4th modes in the Monte-Carlo simulation has comparable power. This confirms that the fourth IMF is a real signal, different from random noise.

This significant mode is a clear solar cycle signal. The correlation with the sunspot data is shown in figure 3. A student t-test (Hogg 1993) is used to calculate the significance of these correlations,

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}}$$

where $t$ is the student t value, $r$ is the correlation and $\nu = N-2$ is the degrees of freedom. Special care

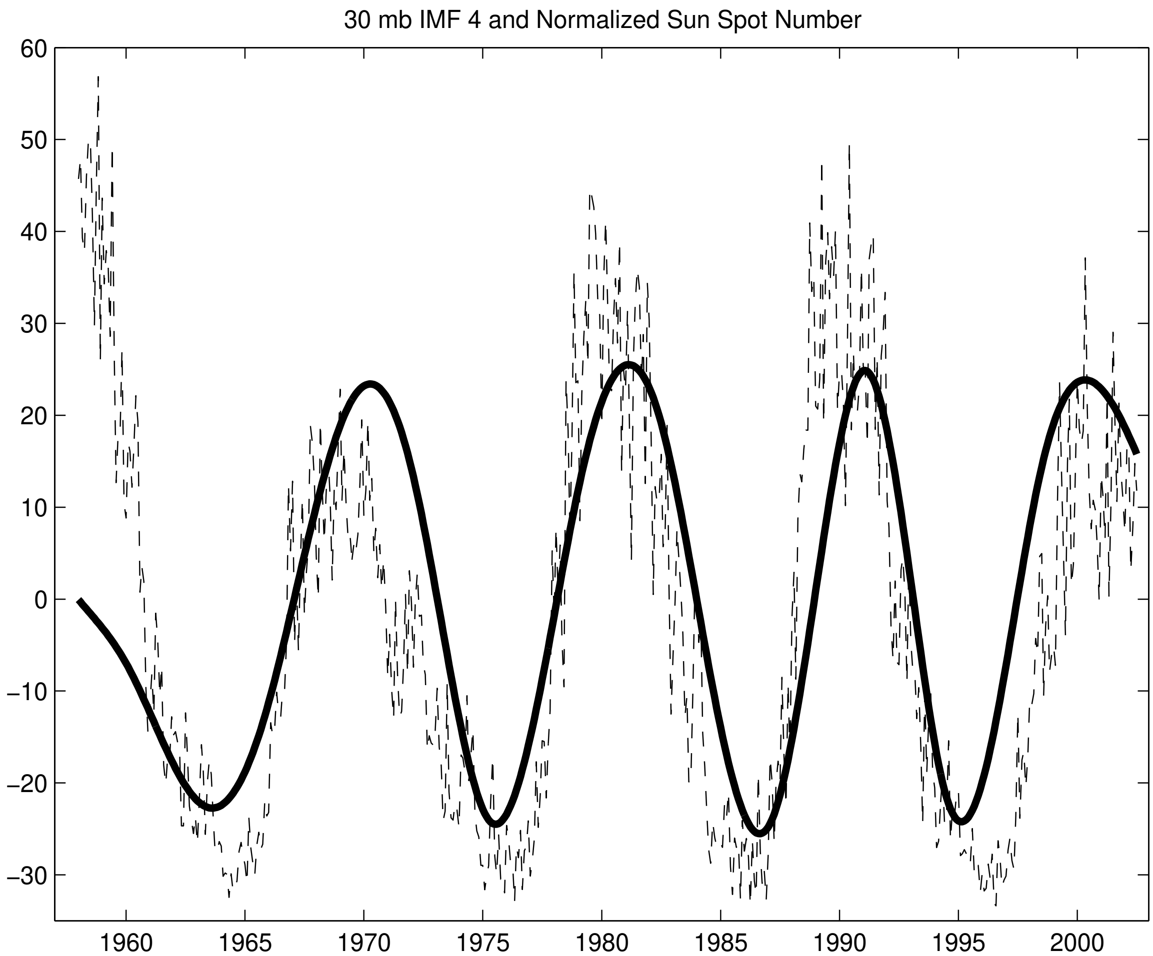


Fig. 3. The dashed line is the normalized sunspot number. The thick solid line is the 5-year mean of this time series and the thin solid line is IMF 4 of the 30 mb geopotential height. The two time series are significantly correlated at .70. The significance calculation is made assuming that the mode has only 7 dof.

must be taken in estimating the degrees of freedom because the method used to generate the fourth mode is new and non-stationary, so typical techniques may not be applicable. Here we use the following argument: In the EMD method, each IMF is successively subtracted from the original time series, leaving behind only variations with timescales longer than those in the subtracted modes. It is reasonable then, that the data points in the residual time series will be dependent on timescales equivalent to the periods of the last subtracted mode. In the decomposition of the 30-mb geopotential height, the third IMF has an average period of 4.4 years. This implies that the fourth IMF is composed of data which is dependent on timescales of about 4.4 years. Conservatively then, we estimate that the fourth IMF has independent time intervals of 5 years. Using this estimate, nine independent measurements of the time series can be made for the 4th IMF and therefore $\nu = 7$ degrees of freedom for the correlation calculation. The sun spot data is then smoothed using a 5-year running mean and correlations are calculated by comparing the data every 60 months. These correlations average to .70. For comparison, the correlation coefficient, calculated using every month in the unsmoothed sunspot numbers, is .72. In fact, the correlation is very robust. And using the student t-test the .70 correlation is statistically (at the 95% confidence level) different from the null hypothesis of a zero correlation.

This statistical relationship between the fourth IMF and the sunspot numbers can also be verified using a regression with auto-regressive (AR) errors. For this analysis, we assume that the IMF contains the solar cycle plus an autoregressive error,

$$\text{IMF 4 (t)} = \text{b (SI(t) - } \mu) + \text{e(t)}.$$

SI is a solar cycle index defined here to be the sunspot number divided by 1000. The parameters, b and $\mu$, and their variances can be found by fitting the error, e(t), to an AR(p) process for some order, p (Newton 1988). Although the parameter values do not change very much, we find that the best fit for the errors is an autoregressive process of order 5. This gives the parameter values of $\mu = .082 \pm .013$ and $b = 42 \pm 7$ as the best fit for this model. Since $b = 42$ is about 6 times its standard error, we can effectively rule out the hypothesis $b = 0$. This implies that the solar cycle is directly related to the fourth IMF. The regression is highly significant and the sunspot cycle statistically explains a majority of the fourth mode.

**Conclusion**

A clear solar cycle signal is observed in the 30-mb geopotential height using the nonlinear, non-stationary EMD method. The total geopotential height at 30 mb is averaged over all longitudes and from 20N to 90N. No specific grouping of the data is used in this analysis. The entire timeseries is completely decomposed into 5 modes and a trend. Using a Monte-Carlo simulation, the power in each mode is compared to the power in 500 decompositions of random noise. The fourth mode is found to have an average power far above the noise level and therefore is a significant signal. The correlation between this signal and the solar cycle proxy is .70 which is also significant given our estimation of the degrees of freedom in the mode. Using a regression with AR errors, the significance of the correlation is verified. The result is both a statistically and visually convincing solar cycle signal in the total 30-mb geopotential height.

**Appendix A**

To distinguish the signals from noise in the EMD method, we examine the energy of all the IMFs and compare the energy in each mode to the energy distribution of white noise. Typically, we may expect the highest frequency mode to contain only noise and the power in this mode can then be used to calibrate the noise distribution. However, in this case the first mode contains the annual cycle which is obviously not pure noise. To estimate the power of the noise present in this time series, we subtract the climatology from the first mode and assume that the remainder is noise. This is still a very conservative estimate since the true annual cycle signal probably does vary from year to year and contains a real signal in addition to the climatology. However, using this criteria, we can normalize the random noise and perform a Monte-Carlo simulation. Five hundred random time series are added to the 30-mb climatology before the EMD method is used to decompose the noise into modes. The power of the original modes can then be compared to the power in the modes created from the noise.

$$\text{random time series} = A_{noise}\ \text{R} + \text{climatology}$$
$$A_{noise} = \text{standard deviation (IMF 1 - climatology)}$$

where R is a uniformly distributed random time series with amplitudes between 1 and -1. The climatology is the geopotential height averaged over each individual month of the total 30-mb geopotential height.

EMD modes are generated using these time series and then the statistics of this simulation are calculated to create the energy distribution of the noise. The power of each IMF is then compared to this distribution to determine its significance.

**ACKNOWLEDGMENTS**

We would like to thank Don Percival at the Applied Physics Laboratory at the University of Washington for providing valuable help and insight into the statistical approach of regression using autoregressive errors. This research is supported by the National Science Foundation, Climate Dynamics, through grant ATM 9813770.

**REFERENCES**

Cugnon, P., Solar Influences Data Analysis Center, http://sidc.oma.be/, 1999.

Dunkerton, T. J. and M. P. Baldwin, Modes of Interannual Variability in the Stratosphere, *Geophys. Res. Lett.*, **19**, 49-52, 1992.

Haigh, J. D., The Role of Stratospheric Ozone in Modulating the Solar Radiative Forcing of Climate, *Nature*, **370**, 544-546, 1994.

Hogg, R. V. and E. A. Tanis, Probability and Statistical Inference, Macmillon Publishing Company, pp. 543, New York, NY, 1993.

Hood, L. L. et al., Quasi-decadal Variability of the Stratosphere - Influence of long-term Solar Ultraviolet Variations, *J. Atmos. Sci.*, **50**, 3941-3958, 1993.

Hood, L. L., The solar cycle variation of total ozone: Dynamical forcing in the lower stratosphere, *J. Geophys. Res. - Atmos.*, **102**, 1355-1370, 1997.

Huang, N. E. et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. Lond. A*, **454**, 903-995, 1998.

Kalnay, E., et al., The NCEP/NCAR Reanalysis 40-year Project, *Bull. Amer. Meteor. Soc.*, **77**, 437-471, 1996.

Kodera, K., The Solar and Equatorial QBO Influences on the Stratospheric Circulation during the Early Northern-Hemisphere Winter, *Geophys. Res. Lett.*, **18**, 1023-1026, 1991.

Kodera, K., Quasi-decadal modulation of the influence of the equatorial quasi-biennial oscillation on the north polar stratospheric temperatures, *J. Geophys. Res.*, **98**, 7245-7250, 1993.

Labitzke, K., Sunspots, the QBO, and the stratospheric temperature in the north polar region, *Geophys. Res. Lett.*, **14**, 535-537, 1987.

Labitzke, K. and H. van Loon, Associations between the 11-year solar cycle, the QBO and the atmosphere: I. The troposphere and stratosphere in the northern winter, *J. Atmos. Terr. Sci.*, **50**, 197-206, 1988.

Labitzke K. and H. van Loon, The 11-year solar cycle in the stratosphere in the northern summer, *Annales Geophysicae*, **7**, 595-598, 1989.

Labitzke, K. and H. van Loon, The 11-year solar cycle in the stratosphere in the northern summer, *Annal. Geophys.*. **7**, 1990.

Labitzke, K., The global signal of the 11-year sunspot cycle in the stratosphere: Differences between solar maxima and minima, *Meteor. Z.*, **10**, 83-90, 2001.

McCormack, J. P. and L. L. Hood, Apparent Solar Cycle Variations of Upper Stratospheric Ozone and Temperature: Latitude and Seasonal Dependences, *J. Geophys. Res. - Atmos.*, **101**, 20933-20944, 1996.

Newton, H. J., TIMESLAB: A Timeseries Analysis Laboratory, Wadsworth and Brooks/Cole Publishing Company, pp. 241-243, Belmont, CA, 1988.

van Loon, H. and K. Labitzke, Association between the 11-Year Solar Cycle and the Atmosphere. Part IV: The Stratosphere, Not Grouped by the Phase of the QBO, *J. Climate*, **3**, 827-837, 1990.

WMO, Atmospheric Ozone: Assessment of Our Understanding of the Processes Controlling its Present Distribution and Change, World Meteorological Organization, NASA, Office of Mission to Planet Earth, Two Independence Square, 300 E Street SW, Washington DC, 1987.

E-mail address for K.T. Coughlin: katie@amath.washington.edu