JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

# Reply to Comments by Gleisner, Thejll and Christiansen

K. T. Coughlin

Department of Meteorology, University of Reading, Reading, UK

K. K. Tung

Applied Mathematics Department, University of Washington, Seattle,

Washington, USA

---

K. T. Coughlin, Department of Meteorology, University of Reading, Reading, UK. (k.coughlin@rdg.ac.uk)

K. K. Tung, Applied Mathematics Department, University of Washington, Box 352420, Seattle, WA 98195, USA. (tung@amath.washington.edu)

**Abstract.** Gleisner and Thejll's (2003) methodology and choice of indices produce a large and structured residual. This accounts for the difference in both the solar cycle pattern and the area of statistical significance between their analysis and Haigh's (2003), mentioned by Coughlin and Tung (2004). Additional criticism by Gleisner, Thejll and Christiansen of the EMD method seems to result from a misunderstanding of how the method was implemented in Coughlin and Tung (2004).

## 1. Introduction

The comment by Gleisner et al. (2005), hereafter GTC, is directed at the following paragraph in the Introduction of our paper, *Coughlin and Tung* [2004b] , hereafter CT:

> *Multiple regression analysis is one way to separate the signals. Gleisner and Thejll (2003) used multiple regression analysis on annual mean temperatures to disentangle the atmospheric signals. Their multiple regression includes not only the solar cycle, but also El Nino-Southern Oscillation (ENSO), a trend and volcanic signals. Haigh (2003) includes these as well as other signals in their multiple regressions, notably the quasi-biennial oscillation of the equatorial winds, and she recovers a different structure of solar influence. Since the solar cycle response is dependent on the choice of other signals used in the multiple regression analysis, it is important to find an objective way to disentangle the atmospheric signals.*

In their Comment, our statement was taken to mean, "in Coughlin and Tung regression is rejected in favour of a method referred to as Empirical Mode Decomposition". That was not our intention at all. If a phenomenon, in this case, the 11-year solar cycle, can be reproduced with a variety of methods, it only adds credence to our claim that the signal is real.

CT's statement quoted above concerns only the different spatial structure obtained when different indices are included in the multiple regression analysis in different regression studies: one (Haigh (2003)) includes the QBO and the NAO in her regression and the other (Gleisner and Thejll (2003)) does not. As we will show in this note, it turns out that the difference is not only the subjective choice of indices but also the fact that the multiple regression model of Gleisner and Thejll explains very little of the atmospheric variability, although the latter is a consequence of the former.

## 2. Comparison between Gleisner and Thejll (2003) and Haigh (2003)

GTC claim that the difference in the tropospheric spatial structure between Gleisner and Thejll (2003) and Haigh (2003) was a result of our comparison of the *correlation* coefficient in the former with the *regression* coefficient in the latter. It is important to point out that the solar regression coefficient was not shown in Gleisner and Thejll (2003) nor in GTC. Gleisner and Thejll (2003) showed only a correlation coefficient in the left panel of their figure 1c and in the right panel they showed a "solar max-min difference in temperature". This quantity is *not* the regression coefficient, but is the difference between their "corrected zonal mean temperature", T*, at solar max and at solar min. This "corrected" temperature contains the large, and structured, residual, $\epsilon$, of their multiple regression analysis.

Following the description of their methodology, we have repeated their multiple regression analysis and show here in figure 1 the solar regression coefficient, $k2$ . Comparing this figure with figure 3(b) of Haigh (2003) shows the differences that we referred to in CT. For comparison with *Haigh* [2003] the regression coefficients here are calculated using an index of the 10.7 cm flux, normalized by the max-min flux difference of 120 solar flux units. Another difference in the two analyses is that *Gleisner and Thejll* [2003] used linear least squares to calculate the regression coefficients instead of an autoregressive model as in Haigh's analysis.

More importantly, with their choice of regression indices, Gleisner and Thejll's (2003) multiple regression model does not actually capture most of the observed temperature variance. The percentage of variance captured by their analysis is shown in figure 2. This percentage is the coefficient of multiple determination, $R^2$. From this we can see that their multiple regression model describes less than 50% of the variability in the NCEP/NCAR temperature except in a couple small areas. In the Northern Hemisphere troposphere, where the data is most accurate, the analysis captures the least amount of variability. North of 30°N, no more than 10% of the variability is explained by this multiple regression analysis, which means that the majority of the variability remains in the residual and has not been accounted for.

## 3. Statistical Tests

The residual from the multiple regression model of Gleisner and Thejll (2003) is not only large, but also contains structure, implying that it is not white noise. Gleisner and Thejll did not give enough detail on how they did their spectral Monte-Carlo statistics using this residual. However in our reproduction of their analysis we obtain very similar areas of 90 and 95% significance using standard statistical tests only when we assume the residual consists of white noise. While Gleisner and Thejll may not have made the same assumptions, the fact that almost their whole troposphere was claimed to be statistically significant at 95% is in contrast to the result of Haigh (2003), where only two strips in the mid-latitudes were found to be so significant. We perform an independent test of statistical significance using the correlation coefficient shown on the left panel of their figure 1c, and an estimate of the number of degrees of freedom.

Using this approach, we suggest that Gleisner and Thejll (2003) overestimated the statistical significance of their solar cycle response in the left panel of their figure 1c. In fact, based on our calculation their signal in the troposphere is not statistically significant. Gleisner and Thejll (2003) showed in their figure 1c the correlation coefficient r between the solar cycle index and the corrected zonal mean temperature T*. Regions where $r > 0.41$ are darkly shaded and claimed to be above 95% confidence level. The correction involves taking out regressed variability related to ENSO, volcanic aerosols and linear trend. The remainder, which includes the regressed solar cycle signal, is then correlated with the F 10.7 solar cycle index. This "correction" improves the magnitude of $r$, as can be seen by comparing this result with that of Labitzke et al. (2002), who did a direct correlation of the annual mean temperature with the solar cycle index. Labitzke et al. (2002) finds correlation coefficients that are approximately a factor of 2 smaller than those obtained by Gleisner and Thejll (2003). To see the effect of this procedure if this correction were carried to the extreme, by removing all other variability except the regressed solar cycle, the correlation would have yielded $r = 1$ everywhere. The "correction" removes some of the signals with shorter period variability and the degrees of freedom should be reduced to account for this. A decrease in the number of degrees of freedom will raise the threshold for 95% confidence level. In Coughlin and Tung (2004), we argued that the interval between independent data points for solar cycle variability is 5 years after taking into account that ENSO has a mean spectral peak near 4 years. This consideration would drastically reduce the number of degrees of freedom. With independent data points every 5 years, a student t-test threshold for the 95% confidence level should be closer to $r = 0.58$, which is the level we used for an analysis of 46 years. Using this criterion very little of the troposphere in Gleisner and Thejll (2003)'s temperature correlation would be statistically significant.

## 4. The EMD Method

There is no need to attack the EMD method that we used. We have already stated in our published papers some of the technical problems with the method, and have communicated these and more to Dr. Gleisner in our private correspondence, which was acknowledged. These include:

1. *Robustness:* The larger area (20-90N) averaging gives a more robust solar cycle signal than averaging over local narrow strips. We emphasized this point in the Introduction of our paper:

*As pointed out by Pittock [1978], some of the earlier pitfalls in the statistical analysis of this signal include: selecting a posteri the most favorable regional results for statistical analysis from a global map. We therefore pay particular attention to these issues. We will show that our signal is global and ubiquitous by demonstrating that various near global means are statistically significant and that individual latitude strips of 20 degrees at each* *pressure level share the same property (with the exception of South Polar latitudes, where the data quality may be less than desired).*

We are glad that GTC concur: “we find that the decomposition of the northern hemisphere average (20-90N) of the 700 hPa geopotential heights presented in Coughlin and Tung (2005 sic) is robust to changes in q.” q is the number of iterations performed in the EMD method.

This sentence is followed by: “However, results that depend on EMD applied to local variables, such as in figure 3 of CT, are fragile and error prone.” We entirely agree that a decomposition at a single latitude, like the 700 hPa GPH at 40 degrees N which GTC show in their figure 2, would be “fragile and error prone”. CT repeatedly cautioned against using EMD decomposition on local variables. GTC incorrectly assumed that our figure 3 used single latitude decompositions, when in fact 20-degree latitude strips were used. These decompositions are largely robust, except in the regions we mentioned.

2. *Volcanoes:* There is, in the research community, a perennial concern of volcanoes contaminating the solar cycle signal, and this issue is raised again by GTC here. This is not a new concern and we were fully aware of it. With Fourier decomposition, which is global in time, two volcano eruptions 9 years apart will project onto the solar cycle frequency band. This problem is much diminished, but not eliminated, with the EMD method, which is local in time. Furthermore, we argued that whatever contamination remains in the EMD decomposition will not affect our conclusion about the existence of

the solar cycle signal, which warms the troposphere, since the effect of the volcanoes are generally to cool the lower atmosphere.

GTC mention, without substantiation, EMD's "inability to properly account for the effects of volcanic emissions". It is not clear what kind of "volcanic" experiment was done, and what amplitude and duration of emissions were used in their "wedge". To shed some light on this problem, we include here an example using realistic years and durations of volcanic emissions based on the aerosol index used by Haigh (2003). See figure 3. A synthetic time series composed of an annual oscillation, the NCEP/NCAR QBO index (*Kalnay and Coauthors* [1996]), normalized to $\frac{1}{2}$ based on the maximum tropospheric amplitude from Haigh's (2003) regression, and an 11-year sine wave, normalized to 1 based on Haigh's maximum tropospheric values, are decomposed using EMD. This decomposition is shown in blue. A second decomposition is produced by applying EMD to the above signal plus the volcanic aerosol index (*Sato et al.* [1993]) for the northern hemisphere, normalized to -0.5, also from Haigh's (2003) maximum tropospheric amplitude. This decomposition is shown in red. The difference in the decompositions can be seen in the figure. The peak amplitudes during "solar max" and the peak amplitudes during "solar min" are scarcely changed in the 4th mode. The volcanic emission in 1991 serves to slightly shift the time of "solar max" response. Other volcanic emissions do not seem to have any discernible effect on the 4th mode. The majority of the volcanic signal is actually absorbed in the first and second modes. This is because the volcano aerosols typically last one to two years in the atmosphere and so, in the local EMD analysis, they predominantly affect the annual and the QBO modes.

3. *Mode mixing:* The EMD method simply separates the original signal into 5 modes and a trend. No data are filtered out other than by the original smoothing of the time series. Another way of looking at this is that when we add up the resulting modes and the trend, we get back the original data that went into the analysis. When mode mixing occurs, a high amplitude noise or a glitch of high frequency may locally (at the time location of the glitch) push the QBO oscillation from the second EMD mode into the third mode, and the ENSO signal, originally in the third mode, into the 4th mode. So the solar cycle oscillation, which should occupy the entire 4th mode, now contains some ENSO oscillation during the years of this glitch. Often, when mode mixing occurs, an extra mode, or sometimes more than one, will be formed to accommodate the lowest frequency signal which has been pushed out of its original mode, as happens in GTC's example. There is still nothing wrong with this decomposition, because all the signals are still there. It just complicates the physical identification of, say, the 4th mode, as the solar cycle. When this happens, the mode mixing can usually be picked up visually. The remedy is to smooth the original data before it is fed back to the EMD apparatus. Since the smoothing that needs to be done never exceeds 5 months in the extratropical time series shown in CT, it is not likely that by itself, it will create a false 11-year signal. Therefore, while we concur that mode mixing is a technical problem with the EMD method—we in fact pointed this out to the authors—we do not believe that it affected the results we showed in CT.

We are glad that GTC report better results when they smooth the data with 3 month means in the lower panels of their figure 2, after observing mode mixing with the unsmoothed data: “In CT it is mentioned that in some cases smoothing of the time series can be used to avoid a splitting of the annual mode. The lower panels of figure 2 show the results of the EMD applied to the same time series as before, but now smoothed with a three-month running average as suggested in CT. Now the annual cycle is retrieved as the first mode for both values of q.” Their criticism of this result is that it is still somewhat dependent on the stopping criterion on q. This is discussed next.

4. *Stopping criteria:* GTC state incorrectly that “In CT, as well as in Coughlin and Tung (2005), the stopping criterion was that the requirement (a) was fulfilled for three consecutive iterations.” Requirement (a), in GTC, was that an oscillation have a “negative minima and positive maxima”. This was not our criterion. Our stopping criterion is instead that once the number of extrema becomes within ±1 of the number of zero crossings, this number remains the same for 3 iterations. The EMD method is an iterative method. The number of iterations, q, determines how close the local mean is to zero and how close the resulting modes resemble a pure Fourier mode. With each iteration more of the local mean is subtracted from the mode causing the local mean of the mode to tend towards zero as well as causing the modes to become more symmetrical. One is in effect obtaining different mathematical representations of the time series as q varies. In the limit of large q the representation approaches, for practical purposes, a Fourier series representation, which is less compact than an EMD decomposition. This explains why additional EMD modes are created in GTC’s example as q increases. It also explains why it does not make sense to use convergence with respect to q as the stopping criterion, as

GTC seem to be using. GTC are correct in pointing out: “It has previously been noted by Riling et al. (2003) that with few iterations the local mean is not zero and with too many iterations physical modes may spread over adjacent modes (mode mixing).” Our aim was to obtain the “physical modes” without significant mode mixing, while achieving an exact zero local mean was not necessary in our analysis and not our primary objective.

## 5. Conclusion

In their comment of our paper, Gleisner et al. (2005) suggests that despite their not having a sufficient number of regression indices in their multiple regression model, the regressed solar cycle is very similar to that of Haigh (2003), who additionally included QBO, NAO and an AR noise model in her regression model. We demonstrate here that this is not true. The confusion was caused by the authors’ use of ”corrected” temperature differences instead of the regressed coefficient in making the comparison. Additionally we show here that the model of Gleisner and Thejll (2003) explains no more than 10% of the observed variance north of 30 N, one region of focus in their original paper, casting doubt on whether the solar cycle result so obtained is statistically significant. While Gleisner et al. is correct in pointing out that no method is completely objective and that in the EMD procedure some choices are need in its implementation, their result and ours are in agreement over the fact that if correctly implemented, the EMD decomposition is relatively robust. We run a test decomposition of a time series with and without volcanoes and show here that the criticism of Gleisner et al. regarding the effect of volcanoes on the 4th IMF of the EMD decomposition is unsubstantiated.

**Acknowledgments.** We would like to thank Simon Crooks and Teresa A'mar for their contribution in useful discussions regarding statistical testing and regression analysis. The research was supported by National Science Foundation, under grant ATM-3 32364 and the UK Natural Environment Research Council (NERC).

## References

Coughlin, K., and K. Tung (2004), 11-Year Solar Cycle in the Stratosphere Extracted by the Empirical Mode Decomposition Method, *Advances in Space Research*, *34*(2), 323–329.

Coughlin, K., and K. Tung (2004b), Eleven-year solar cycle signal throughout the lower atmosphere, *J. Geophys. Res.*, *109*, doi:10.1029/2004JD004873.

Covington, A.E. (1969), Solar Radio Emissions at 10.7cm, *J. Royal Astron. Soc. Canada*, *63*(125).

Gleisner, H., and P. Thejll (2003), Patterns of tropospheric response to solar variability, *Geophys. Res. Lett.*

Haigh, J. (2003a), The effects of solar variability on the earth's climate, *PHILOS T ROY SOC A*, *361*(1802), 95–111.

Kalnay, E., and Coauthors (1996), The ncep/ncar reanalysis 40-year project, *Bull. Amer. Meteor. Soc.*, *77*, 437 – 471, data provided by the NOAA-CIRES Climate Diagnostics Center in Boulder, Colorado, USA, from their web site at http://www.cdc.noaa.gov/.

Labitzke et al., K. (2002), The global signal of the 11-year solar cycle in the stratosphere: observations and models, *J. Atmos. Sol-Terr. Phys.*, *64*, 203–210.

Sato, M., J. Hansen, M. McCormick, and J. Pollack (1993), Stratospheric aerosol optical depths, *J. Geophys. Res.*, *98*, 22,987–22,994.

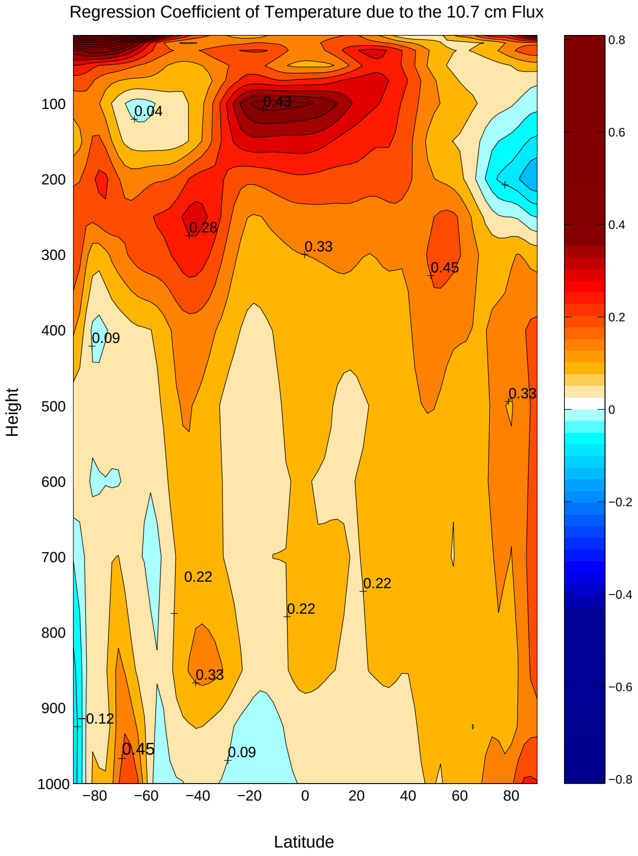

**Figure 1.** The Solar Regression Coefficient ($^\circ C$) from a multiple regression with 3 indices and a trend. Using a 3 month lagged NINA3 (from NOAA Climate Prediction Center (CPC) at http://www.cdc.noaa.gov/ClimateIndices/List/#Nina3) as the ENSO index, volcanic aerosols (*Sato et al.* [1993]), the 10.7cm flux (provided as a service by the National Research Council of Canada) and a linear trend, the regression coefficients of the NCEP/NCAR monthly temperature reanalysis (*Kalnay and Coauthors* [1996]) are calculated using MATLAB's regression code. For comparison with Haigh (2003), the solar index has been normalized by the max-min difference, 120 solar flux units. A solar flux unit (s.f.u.) is $10^{-22} W m^{-2} Hz^{-1}$.

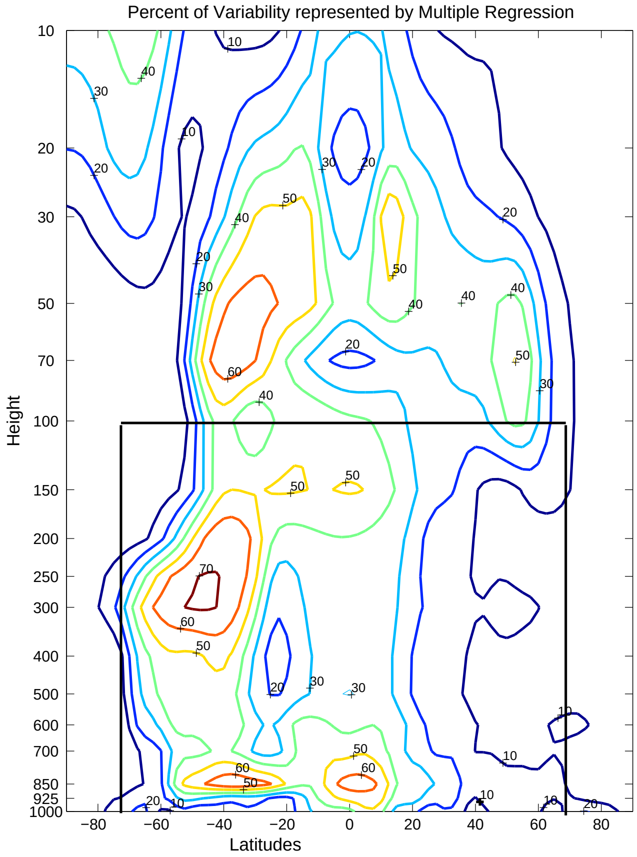

**Figure 2.** The percent of the temperature variability that is described by the multiple regression performed by Gleisner and Thejll (2003). The area marked by the bold black square denotes the area shown in their paper.

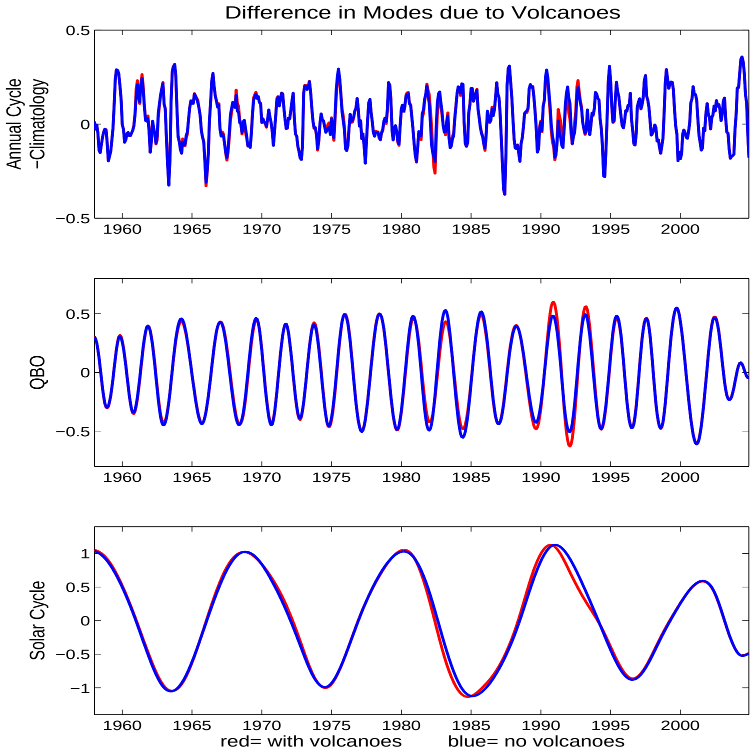

**Figure 3. Using EMD, the difference between a signal which contains volcanic aerosols (red) and a signal which does not (blue).** Both synthetic signals also contain an annual cycle, a QBO, and the changes in solar flux. The QBO is calculated at the CDC (30 hPa zonal wind at the equator) and can be found on their website at www.cdc.noaa.gov/Correlation/qbo.data. The volcanic stratospheric aerosol time series (*Sato et al.* [1993]) is used. The annual cycle is composed of a sine wave of amplitude 100. The top figure shows the differences in the annual cycle after the climatology has been subtracted. The middle figure shows the difference in the second mode and the bottom figure shows the differences in the third mode. In this example the volcanoes are given the maximum tropospheric amplitude seen in Haigh's 2003 regression.