Response to reviewer #1

We believe the reviewer has a point in suggesting a more precise title of the paper. Thus, in the revision we have changed the title to:

“Long-range memory in internal and forced dynamics of millennium-long climate-model simulations.”

The main changes made in the revised manuscript are:

1. The data set used from the HadCM3 simulation was wrong (it only contained ocean data). We also wrote in one place that this simulation was a control run, which is not correct. We have presented analysis of the correct data, and since it is a forced run, we have also added a new figure with results from using the response model. Figure 1 representing the forcings has been modified and extended with the forcing of HadCM3. The results are much more in line with those for the other models studied, and enforces our conclusions.

2. We have addressed the comments by dr. Henriksson and the second reviewer by plotting the power spectra with monthly time resolution when available, and estimating the spectral index beta also from these spectra. Figures have been revised accordingly, and the captions have been extended and made more precise. Table 2 has also been modified.

3. We have included a discussion of the results of Henriksson et al. on the COSMOS model.

Response to reviewer #2

1. We are well aware about the analytic results of Lennartz and Bunde 2009 on the bias of the ACF estimators, and we have discussed the same thing in a paper in JGR from 2013 based on a Monte Carlo study. In the revision we mention very briefly why we don’t see the ACF as a good estimator an give reference to those two papers.

2. We believe the reviewer has a good point in also using the PSDs for beta-estimation, so we have done that in the revision. See our comment to dr. Henriksson.

3. We also agree that we should not use F for both DFA and WVA fluctuation function, so we have changed that. We cannot agree that the Haar wavelet serves our purpose better than the Mexican hat. The Mexican-hat wavelet is very standard in textbooks and programming packages, and can be intuitively interpreted as a local Fourier transform. A known weakness of the Haar wavelet is stronger spectral leakage (see e.g., textbook of Percival and Walden, 2008). Our main purpose to show the result for the Mexican-hat wavelet (and this was stated in the text) is to show how different methods can have different sensitivity to oscilliations. The Haar wavelet (there are different versions) will perform more similar to the DFA.

4. As will be apparent in the descriptions of the revisions given below, the results we obtained from HadCM3 were wrong. The true results are more similar to those obtained from the other models.

Main revisions:

We believe that reviewer #1 has a point in suggesting a more precise title of the paper. Thus, in the revision we have changed the title to:

"Long-range memory in internal and forced dynamics of millennium-long climate-model simulations."

The main other changes made in the revised manuscript are:

1. The data set used from the HadCM3 simulation was wrong (it only contained ocean data). We also wrote in one place that this simulation was a control run, which is not correct. We have presented analysis of the correct data, and since it is a forced run, we have also added a new figure with results from using the response model. Figure 1 representing the forcings has been modified and extended with the forcing of HadCM3. The results are much more in line with those for the other models studied, and enforces our conclusions.

2. We have addressed the comments by dr. Henriksson and the second reviewer by plotting the power spectra with monthly time resolution when available, and estimating the spectral index beta also from these spectra. Figures have been revised accordingly, and the captions have been extended and made more precise. Table 2 has also been modified.

3. We have included a discussion of the results of Henriksson et al. on the COSMOS model.

<u>Response to Reviewer #3, Svante Henriksson</u>

Thanks to Svante Henriksson for the comment and for making us aware of their recent paper, which we will of course cite in the revised version.

Henriksson et al. analyse COSMOS global temperature series and obtain lower spectral exponents than we do, in particular for the unforced simulation. The reasons for this discrepancy are the following:

The COSMOS simulation has a very strong ENSO-like signal, considerably stronger than in instrumental or reconstruction data. This leads to a distinct spectral "hump" in the global temperature on frequencies from 1 to 1/6 (1/yr), but with a long tail towards lower frequencies. The spectral shape of the ENSO signal is apparent from their Fig. 8a and their Fig. 8b gives an indication of a much wider power-law range with an index β closer to 1 in temperature series recorded remote from the ENSO region. The globally averaged temperature is a superposition of signals having these two spectral shapes, and the result is that the range of frequencies f>1/6 (1/yr) are influenced both by the ENSO hump and the underlying power law which dominates on higher lattitudes. This leads to an apparent lower slope (lower β) of the log-log spectrum in this range.

The authors make a point of splitting the spectrum of a non-scaling global signal into short pieces (ranges) and compute slopes for each of these ranges. We cannot see that this has physical meaning, since what appears as different ranges are results of combining (averaging over) signals with different scaling properties; in this case power-law scaling signals from temperate and polar zones, and quasi-oscillatory signals from the tropical oceans.

The goal of our analysis has been to uncover the underlying power-law scaling which prevails awy from the ENSO region, and is why we have focused on DFA. It has been shown that the DFA analysis it is relatively insensitive to oscillatory perturbations, which is quite apparent from our DFA fluctuation function for the COSMOS simulations. In comparison, the WVA is much more sensitive (see attached figure). However, the underlying scaling also becomes apparent from the power spectrum if we analyse monthly, deseasonalised, data. This is shown in panel (b) of the attached figure for the COSMOS control simulation. This figure will replace our old Figure 7.

In the revised paper we have estimated exponents from spectra for all data sets, and used monthly, deseasonalised data where such data are available.

This should resolve the particular discrepancy pointed out by Henriksson. There are many other interesting issues discussed in their paper that would be interesting to discuss in more depth, but we believe the right place for this is in a separate paper.