Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-61-RC3, 2016

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "An evaluation of current capabilities of modelling low-frequency climate variability" *by* Heikki Järvinen et al.

Anonymous Referee #3

Received and published: 23 August 2016

This manuscript focuses on the capability of current climate models to simulate low-frequency climate variability, as determined through a randomised multi-channel singular spectrum analysis (RMSSA) of near-surface air temperature. On the basis of this analysis, the authors conclude that state-of-the-art climate models tend to exhibit variability that is too periodic, under-active at multidecadal timescales, and over-active at decadal timescales. On the positive side, I thought that the manuscript was clearly and smoothly written. However, I was left with many questions about the authors' choices. The title of the manuscript is very broad and ambitious, but the authors only analyse one variable with one method in only 12 climate models, so any conclusions that are drawn are much narrower in scope than the title would suggest. By focusing on statistically significant periodicities, the authors really do not directly address whether or not models have too much or too little low-frequency variability (particularly since all time

Printer-friendly version

Discussion paper

series are standardized prior to the analysis). All comparisons between the models and reanalysis are informal and subjective, and all formal significance testing is limited to red noise null hypotheses rather than model/reanalysis differences. Overall, I was hoping that this study would provide a more thorough and objective evaluation of model performance that goes beyond previous studies, or if that was not the intention, that the scope of this study would be more clearly articulated. I describe my concerns more thoroughly below.

Major Comments

1) Lines 41-48: The attribution of variance at different timescales by the authors is too simple and not entirely accurate. A substantial portion of variance at interannual to interdecadal timescales can be attributed to "climate noise" associated with processes with intrinsic timescales that are much shorter than interannual. That is the nature of red noise. For example, the North Atlantic Oscillation (NAO) is a teleconnection pattern with broad impacts and pronounced interannual and interdecadal variability, and yet much of that can be attributed to internal atmospheric variability (Wunsch 1999; Feldstein 2000). Therefore, it is not accurate to say that interannual variability is primarily attributed to ENSO or that decadal-to-multi-decadal variability is attributed to ocean dynamics. These comments may be true for periodic variability, but then the authors need to explain why they are focusing on oscillatory behavior and neglecting other dominant sources of interannual and multi-decadal variability.

Wunsch, C.: The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. Bulletin of the American Meteorological Society, 80, 245-255, 1999.

Feldstein, S. B.: The timescale, power spectra, and climate noise properties of teleconnection patterns. Journal of Climate. 13, 4430-4440, 2000.

2) Why do the authors choose the 12 models that they choose? Given that there are so many more simulations available, this choice seems arbitrary.

3) Line 166: Again, perhaps this relates to my misconception about what the authors are trying to address, but the decision to standardize the data sets has made it challenging for me to interpret the authors' results. The climate models may have very different temperature standard deviations, which would impact the temperature variability from interannual to multidecadal timescales (e.g., Thompson et al. 2015). However, by standardizing the data, the authors essentially are artificially adjusting the climate models and reanalyses to have common variance at every grid point. Therefore, the authors are erasing potentially important differences in variance between the models and reanalyses that would impact reanalysis/model differences at all timescales. The motivation for this decision and the consequences for interpretation should be discussed.

Thompson, D. W. J., Barnes, E. A., Deser, C., Foust, W. E., and Phillips, A. S.: Quantifying the role of internal climate variability in future climate trends. Journal of Climate, 28, 6443-6456, 2015.

4) Lines 230-234: The decision to evaluate model performance subjectively is unsatisfying. It is difficult to compare power spectra with short records, and visual inspection can be deceiving. Combined with my previous comment, I have difficulty interpreting the authors' results. There may be truth in the authors' conclusions in lines 252-254, but I would like more support.

5) Line 272: How are "false alarms" defined? Again, the authors did not determine if there are significant differences between the reanalyses and models, and so I do not see how the determination of false alarms was made.

6) Line 289: Why did the authors subjectively choose the Nino3.4 region to base the composites? Although it seems reasonable that the 3.5-yr mode would be related to ENSO, by basing the composites on a subjectively chosen region, the authors seem to be predisposing the analysis to highlight ENSO-like variability.

More generally, I am not sure why Section 3.4 is entirely focused on ENSO and its

teleconnections, given that these topics have been covered extensively in other studies and that the authors argue that five periodicities exist. It would seem less arbitrary to let the analysis direct the content and to focus on all identified periodicities.

Minor Comments

1) Lines 137-138: This relates to my first comment, but the authors are not really addressing whether the models "capture the observed temperature distribution."

2) Line 179: Is there any sensitivity to this choice of lag window?

3) Line 202: I do not understand why those components are called "trend components" if the data were detrended.

4) Lines 208-209: How is it determined that ENSO variability has a decadal component in 20CR?

5) Line 210: I would not consider the similarity of the 20CR and ERA-20C spectra to be striking, given that the reanalyses assimilate similar data.

6) Line 245: I am not convinced of five key periodicities. Physically, it seems that all identified periods may relate to one phenomenon (ENSO), and these five frequencies just happened to pass the significance threshold.

7) Discussion: Isn't it possible that the existence of too many significant periodicities in the climate models could be due to ENSO being too periodic in some models, which has been discussed previously?

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-61, 2016.