

Interactive comment on “Temperature response to external forcing in simulations and reconstructions of the last millennium” by L. Fernández-Donado et al.

Anonymous Referee #2

Received and published: 24 October 2012

Review of Fernandez-Donado et al.

Summary: This manuscript presents a characterization and comparison between large-scale temperature reconstructions and pre-PMIP3 last millennium (LM) simulations. In addition to a very thorough and detailed review of the reconstructions and the simulations, the authors perform several analyses to test the agreement between the model responses, the forcing estimates and the temperature reconstructions and to estimate climate sensitivity.

General Remarks: This is a well-done paper that compiles a tremendous amount of disparate information, particularly on the pre-PMIP3 LM simulations, all of which will

C2038

be very useful to the community. The ensemble of LM simulations has grown very rapidly in the last few years and a compendium of the now-available simulations has not been previously completed. In many cases, some of the model and simulation information (e.g. forcing choices) is very tedious to track down for the LM runs, and it is nice to see the simulation descriptions now all in one place. I should also add that it makes sense to provide this pre-PMIP3 compilation as a benchmark for the now larger collection of PMIP3 LM runs, given that the pre-PMIP3 collection used a wider range of forcings and model configurations (e.g. resolutions). All of this is to say that the collection of information in this manuscript is timely, thorough, and very useful.

If there is a criticism of the manuscript, it would be that the latter analyses of the reconstructions and simulations are a bit underdeveloped. The authors propose and characterize several straightforward means of comparing the reconstructions and simulations, as well as estimating climate sensitivity in the models and the reconstructions. This is all done rather simplistically using large-scale temperature indices and linear assumptions about the climatic response to forcing. The authors are nevertheless transparent about their assumptions. Moreover, it should be said that regardless of the level of sophistication, this is the first study to do this with so many reconstructions and model simulations and there are already some interesting results to ponder. I therefore do not think it would be fair to criticize the authors for not achieving an ideal concept of what might be done, given the novelty of their first step and the large challenge of assimilating so many reconstructions and simulations into their analyses.

With all of the above as summary, I suggest that the paper be published with just a few minor revisions listed below.

Minor Suggestions

Consider the following title change: Large-scale temperature responses to external forcings in simulations and reconstructions of the last millennium

Pg. 4007, Ln. 6: The Jones et al. (Holocene, 2009) review article is appropriate here.

C2039

Pg. 4007, Ln. 14: The Tingley et al. (QSR, 2012) methodological review article is appropriate here.

Pg. 4007, Ln. 28: Why isn't Ammann et al. (PNAS, 2007) included here?

Pg. 4010, Ln. 9: Forward modeling of what? If general, Evans et al. (JGR, 2006), Thompson et al. (GRL, 2011) and Ohlwein and Wahl (QSR, 2012) should be included as additional examples of forward-modeling proxy studies.

Pg. 4017, Ln. 7: It would be useful to expand why the knowledge of the forcing is considered to be a "low Level of Scientific Understanding" and what exactly the criteria for this classification were.

Pg. 4023, Ln. 3-4: "...is not to be expected in the model response." At the 11-yr cycle?

Pg. 4023, Ln. 22: I think the authors mean the Ammann and Wahl (2007) paper here. Also, while there were very slight improvements in that paper over the implementation of the Mann et al. (1999) paper, the method itself has been abandoned by Mann and coauthors for an alternative (RegEM). It therefore is a little strange to call Ammann and Wahl (2007) the replacement choice, when it is arguably the Mann et al. (Science, 2009) paper that represents an update from previous work, both methodologically and in terms of the expanded network that was used.

Pg. 4024, Ln. 2: What is meant by annual borehole data? I do not believe these data are ever presented as such. Table 4 also presents Huang (2004) as a replacement for Pollack and Smerdon (2004). That is not entirely accurate (AR4 could have used the Huang paper instead of Pollack and Smerdon given that they were both published in 2004). Both papers are based on the same borehole data, but the Huang paper tries to blend them with the Mann et al. (1998) reconstruction. There are differences of opinion on whether or not it is a good idea to do that. One should therefore not be presented as a replacement of the other.

Pg. 4024, Ln. 3: It is strange to say that many of the reconstructions that follow have

C2040

"low-variance on inter annual timescales." Many of these reconstructions explicitly have been filtered or only target decadal or lower signals.

Pg. 4027, Ln. 14: It strikes me that there may be another reason for excluding the earlier part of the simulations: initial conditions. This brings up a larger point: how consistent were the initial conditions across the simulations or how consistently were they chosen? This is worth mentioning and potentially something that could be included in one of the simulation summary tables.

Pg. 4029, Ln. 7: "Since the quality of the...forcing factors...can be considered stable through time..." I think this is a dubious statement even for estimates besides the volcanos. The authors might want to provide more justification for this statement.

Pg. 4035: It would be useful to show the Mann et al. (2009) proxy pattern. Even better, the authors might consider calculating pattern correlations between the proxy pattern and the model patterns. This would provide a quantitative measure of the agreement (or disagreement).

Pg. 4035, Ln. 19: It would be really nice to know what is going on with the CNRM model. I do not expect a full diagnosis, but it is such an outlier that the authors might want to address it. Do they have any thoughts on why it shows such different behavior in the SH?

Pg. 4036: It would be appropriate to point out a few things regarding the spatial pattern of the Mann et al. (2009) reconstruction. While it is a valuable estimate, I am a little nervous to point to it as the pattern estimate gold standard. There are of course many uncertainties in the reconstructions and the estimated spatial patterns are subject to the largest of them. A few recent studies in the literature have begun to articulate this view (Smerdon et al., GRL, 2011; Li and Smerdon, Environmetrics, 2012; Annan and Hargreaves, Clim. Past, 2012; Werner et al., J. Clim., 2012). The authors should at least point out these potential spatial uncertainties as a caveat regarding the pattern estimates in the reconstructions.

C2041

Pg. 4038, Ln. 3-6: I do not understand how the authors have accommodated the decadal resolved reconstructions in this moving window scheme. Do these correlations mean anything if the true DOF are approximately 3 in the moving window?

Figs. 3-4: I am not a big fan of including the Huang et al. (2008) curve in this analysis. That particular borehole inversion targeted the last 20,000 years. It did this by merging three data types: the instrumental record, the high-quality temperature profiles used for the 500-yr inversions, and much noisier heat flux data for the period prior to about 500 years ago. This data merge has its advantages and disadvantages, but it may particularly impact the character of the smaller events like the MWP/LIA relative to the much larger LGM signal. The temporal resolution of the inversion is also much lower (and progressively less back in time), making the timing and amplitude of the MWP/LIA events less constrained than the other reconstructions. For these reasons, it is an apples-to-oranges comparison in this context.

Language Note: The manuscript would generally benefit from some language improvements. I am not going to list all of the examples, but the authors would do well to read over the manuscript carefully for language choices. One formulation that occurs repeatedly is exemplified by the following: "can contribute to extend the knowledge." This should be replaced by: "can contribute to extending the knowledge." Similarly, "that allows to compare" should be changed to "that allows comparisons of."

Interactive comment on Clim. Past Discuss., 8, 4003, 2012.