Geosci. Model Dev. Discuss., 3, C562–C567, 2010
www.geosci-model-dev-discuss.net/3/C562/2010/

Geoscientific
Model Development
Discussions

Interactive comment on "Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0)" *by* G. A. Schmidt et al.

G. A. Schmidt et al.

gschmidt@giss.nasa.gov

Received and published: 23 November 2010

Response to Reviewers

We would like to thank the reviewers and editor for their comments.

All of the comments touch on the issue of whether we should recommend a 'best' (or highest priority) set of forcings for modeling groups to use. We have thought about this a lot, and have come to the conclusion that for both short term purposes (related to what will be ready for the IPCC report in 2013) and long term purposes (the length of the CMIP5 project), we are better off not favoring one set over the other. As stated by Jules Hargreaves, it is essential that the uncertainty in forcings (both in structure and amplitude) be properly characterised. Favoring one set of forcings over the other,

equally plausible, ones will almost certainly skew the initial results towards a uniformity of forcings that might mean that a key axis of uncertainty is not addressed in the first sets of papers. Over the long term this is less of a problem since groups will have had time to do more runs.

We deal with specific issues raised by the reviewers below.

Response to Referee 1

1, I understand the point of view of the authors that providing one single forcing set could give the false impression that this particular set is the most reasonable and that uncertainties are small. However, one of the goals of intercomparison exercises is precisely to compare model results is the same conditions. As a consequence, while it is instructive to propose a variety of forcing representing current uncertainties, selecting one set that should be run in all the groups would be very useful for the intercomparison itself. All the groups should then perform a simulation with this set at least and then with additional ones if possible.

We address this above, but we note that the real point of the intercomparison exercise is the comparison of the models with the observed data. If there are significant uncertainties in the forcings but everyone uses the same ones, it is impossible to attribute any model-data inconsistency to the models or the forcings. The simulations may be off because the models are incomplete, or they just used the wrong forcing. By varying the forcings we span a greater range of uncertainty and formal attribution studies will be able to do a much better job of assessing whether there are true model inconsistencies or not.

2/ Several simulations covering the last millennium were performed in the last decade. In particular, a summary is given in the 4th assessment report of the IPCC (Jansen et al 2007). That would be very instructive to precisely know how the proposed forcings differ from the ones in those previous simulations. This is very briefly discussed page 1562 but not mentioning figure 6.13 of Jansen et al. (2007) where the forcing used to *drive models are given. A figure for instance showing those solar and volcanic forcings compared to the ones proposed here by the authors, to see if there is an overlap of if they are clearly distinct, would be very instructive in the interpretation of the result and in the comparison between the various simulations.*

Indeed. We do mention figure 6.13 on page 1565 (line 25), but we accept that we could have expanded on this. The solar reconstructions we propose are all different to those used previously (with the exception of WLS), in that a) they extend longer back in time, b) they have solar cycle and spectral variation throughout, and c) they use more, and more varied, inputs from the cosmogenic isotope archives. With respect to the volcanic forcings, again, the ice core sulphate input data is significantly more extensive than was available pre-2006, and our inclusion of different assessments allows (for the first time) a test of the structural uncertainty in this reconstruction.

3, A Table describing all the solar and volcanic forcings would be very useful, in particular to follow easily all the acronyms used in the different figures (GRA, CEA, PMOD, WLS (noback), WLS (back), ...).

This information will be added to Table 1. PMOD is now defined in the text.

Specific points

1, Page 1551, Line 25. The "classical Medieval period" is defined as 1000-1200 CE. I think that would be necessary to explain that this definition is not very strict, that the timing of the "classical Medieval period", also referred to as the "Medieval Warm Period" or "Medieval Climate Anomaly", is different between the studies and to explain in one sentence why it has been "of particular interest in previous work".

The literature is extremely confused about the definition of the MWP/MCA as the reviewer is aware. We will add a couple of lines on that issue.

2, Page 1566, line 14. It is mentioned that "regional climate changes driven by the various forcings will be of most utility in assessing model skill". However, the contribution

of internal variability is also larger at regional scale than at global scale making more complex any attribution of a signal to a particular forcing. I think that at least mentioning this point would be useful for the readers.

Indeed. There is greater internal variability at more regional scales, but our hope is that the greater use of ensembles (both initial conditions and with different forcings) will allow for the signal to be more easily discerned. With long simulations like these, there is more opportunity for superposed epoch analyses than there would be with simulations of the 20th C for instance. We will add a line or two to reflect this.

Response to Referee 2

General comments:

The experimental design in some respects follows PMIP traditions by identifying recommended and alternate boundary conditions. However, it also includes a very significant departure from previous phases of PMIP in that even within the recommended experimental design there are a number of options that can be chosen. I think the rationale of this is well made in the paper - it has always been part of PMIPs job description to examine the effects that uncertainty in boundary conditions have on climate model predictions for different intervals of time. That is very important when the time comes to compare model outputs to proxy data. In many respects it is also a pragmatic approach in that it recognises that it is impossible to provide different models with identical forcings even if the same boundary conditions are used due to structural differences within the models themselves (vegetation being a classic example).

Nevertheless from a practical point of view such an approach, whilst laudable, gives me some cause for concern. How many groups will really run the different permutations to test the forcing uncertainty properly across a large enough sample of structurally different models? As far as I know the primary objective of PMIP is to compare models runs for past climates and to do that part of PMIP it seems sensible to try as hard as one can to get the models to run with the same forcings at least in 1 experiment for each *time interval. So I agree with reviewer 1, I think it's useful to identify one permutation of the PMIP3 preferred boundary conditions for the last millennium that all groups sign up to and then provide the groups with the option of running further sensitivity tests if time and computing resources allow.*

See above, and response to Rev. 1. We stress that we are trying to shift the focus from model intercomparisons, to model-data comparisons in keeping with the idea that originally motivated PMIP. Note that we anticipate that multiple groups will run multiple forcing sets over the 5 year length of PMIP3.

I also agree that a table showing all forcings would be useful and perhaps highlighting the one permutation that groups are encouraged to perform before moving on to other simulations.

The extra descriptive information will be added to Table 1.

Response to Editor

*I am interested to see that both reviewers would like to see a base configuration to be run by all groups! Personally I can't see the point of this *if* the main uncertainty about the Last Millennium (LM) is the forcing. In that case, the experiment should be designed such that this is investigated; the result required is not the model sensitivity to identical forcing, but the sensitivity across the ensemble to different forcing. Remembering that the LM is a "heavy" computation, groups will not be able to do many runs. In addition there are many other PMIP and CMIP runs which compare the models under identical forcing. Surely scientists analyzing the LM results will have those runs also on hand to enable them to understand the basic model differences.*

We basically are in agreement here (see above). It is certainly true that the forcings are *a* major uncertainty. Selecting one specific set of forcings will be a hindrance to getting as wide as spread in the ensemble as we need for formal attribution studies, particularly early on. It is true that there will be many other same forcing experiments

within CMIP5 using the same models that will provide input on model-to-model biases.

Perhaps some further consideration of the experimental design is required. I thought of two approaches (probably there are many more) depending on how much expertise/ time/enthusiasm you consider there to be among the modelling groups. The first way is, if you believe that all the modelling groups have sufficient resources to consider carefully all the options and make good judgements, then leave each group to make its own best judgement as to which boundary conditions to use. The result will be an ensemble that samples the uncertainty of the experts, but you risk losing runs from groups with smaller resources. A second way would be for the authors of this paper as "LM experts" themselves, to derive prior distributions for the boundary conditions, and then randomly assign boundary condition packages to the different groups, in much the same way as climateprediction.net.

This is an interesting thought, but probably impractical. The motivations of different groups are not identical, and each group will likely have preferences for their 'first tries' that are informed by more than simply the need to add an ensemble member to the pot. Secondly, the number of variations is not so large that this really makes sense. Over 5 years, we anticipate that most interested groups will do a number of simulations that will cover most of the variations. Secondly. we are not in position to assign anything other than uninformative priors to these reconstructions.

Interactive comment on Geosci. Model Dev. Discuss., 3, 1549, 2010.