

## **The PMIP4 contribution to CMIP6 – Part 4: Scientific objectives and experimental design of the PMIP4-CMIP6 Last Glacial Maximum experiments and PMIP4 sensitivity experiments (GMDD manuscript gmd-2017-18)**

The reviewers' comments are reproduced in black and our replies are provided in blue.

### **Summary of the planned changes to the manuscript**

-introduce check points for simulation set up

- improved requirements for the documentation of the simulations:

- river routing

- tidal dissipation

- criteria for equilibrium

- spinup

- discussion on the LGM AMOC

### **Responses to reviews**

#### **Comment SC1: 'AMOC related issues', Andreas Schmittner (28 Feb 2017)**

Muglia and Schmittner (2016; doi:10.1002/2015GL064583) have shown recently that the PMIP3 models all simulate a stronger and deeper AMOC in the LGM experiments compared to piControl, which is inconsistent with reconstructions. They have also shown that changes in wind stress over the North Atlantic due to the additional LGM ice sheets causes stronger wind driven circulation, salt transport and AMOC. I suggest to cite this paper.

Although we did not intend this paper to provide an extensive review of the PMIP LGM results, we will include, in our discussion, a short discussion on the LGM AMOC, which will cite the work by Muglia and Schmittner, as well as other relevant papers, in particular Klockmann et al. (CP, 2016, doi:10.5194/cp-12-1829-2016, 2016), which also relates to the sensitivity experiments that are proposed in the manuscript.

Another effect that has been documented recently by Schmittner et al. (2016; doi:10.1002/2015GL063561) is changes in tidal energy dissipation, which also have the potential to increase the AMOC. I wonder why this effect is not part of the experimental design. Sensitivity experiments could be done with and without changes in tidal energy dissipation in those models that include a tidal mixing scheme. I think discussion of this issue is warranted even if it is decided not to include sensitivity experiments.

We are aware that some ocean models now have the capability to include the impacts of tidal mixing with diverse complexities and that this is a topic for current research even for present conditions (e.g. De Lavergne et al., JPO, 2016, DOI: 10.1175/JPO-D-14-0201.1, Griffiths and Peltier, J. Clim., 2009, DOI: 10.1175/2008JCLI2540.1, Mashayek et al., GRL, 2017, doi: 10.1002/2016GL072452). There were also earlier works (e.g. Arbic et al., Nature 2004, DOI: 10.1038/432460a) on the role of tidal mixing for Heinrich events. We will discuss this topic in the experimental design around the reference provided by the reviewer as well as these other references. As pointed out by the reviewer and after discussion with the modelling groups intending to perform the PMIP4 LGM experiments, we think that the diversity of representation of tidal mixing in the current models is too large to perform rigorous coordinated sensitivity experiments for this topic, but we will encourage the modelling groups which can perform such experiments to do so. This is also clearly a point to be precisely documented for the LGM run, which will therefore be added in the “documentation” section.

Page 9 lines 9, 11: perhaps include “albedo” with “extent and height” if this is intended

This precision will be added.

Section 4.3 Step 3: it would be good to document what was done with Bering Strait in both piControl and LGM experiments

Bering Strait should open in *piControl* and closed in the *lgm* experiment. The closure is actually set up at Step 1. We will make sure this point appears in the documentation of the simulations, along with changes in land sea mask and bathymetry.

## Review RC1 (27 March 2017)

### General comments

This paper describes the LGM and related experiments for the next phase of the PMIP, which are expected to aid the coming IPCC AR6 publication, and makes conscious effort toward the consistent comparisons with a series of the CMIP6 experiments. The manuscript is well written overall, and new aims to bring ESM components into the scope and to disentangle the multiple forcing-response relations are clear. Unfortunately, description of some figures is unsatisfactorily. I recommend to accept this manuscript after such minor revisions.

We thank the reviewer for these positive comments and will try and improve the figure quality.

### Major specific comments

1. I think that the description of Fig. 2 in the text is insufficient. The authors can at least address why the suggested reconstructions are different and what are the main differences before asking the community to use them. How about showing the differences between the three maps? The authors do not have to write in details as they are in the original references, but a summary would be useful.

We will add a short description of Fig. 2 and a short explanation of the differences between the ice sheet reconstructions. These are described in detail in the provided reference (Ivanovic et al., 2016) for the new (ICE-6G\_C and GLAC-1D) reconstructions, so the main point here will be to summarize how the PMIP3 ice sheet was built and why it is different from the latest reconstructions.

2. I think that the description of Fig. 5 in the text is insufficient. Could you provide what is the main message that the readers should take from Figs. 5c-5e?

The reviewer is right, we will add explanations on these panels.

3. I think that the description of Fig. 6 is insufficient. The authors can address at least briefly why they are different and what are the main differences. I also do not understand the meaning of labels (PI, PI&LGM, LGM) in Figs. 6a and 6b.

This information will be added in section 4.6 and the caption will be completed: PI stands for regions which are sources for dust only for pre-industrial conditions, LGM for regions which are sources for dust only at LGM, and PI&LGM are regions of dust sources for both PI and LGM conditions.

4. Given many choices in forcing, would it be difficult to compare model results even though the use of different forcing may cover the range of forcing uncertainty. I think it is useful to discuss the philosophy of why the authors choose the strategy of free choice in forcing, rather than defining one set of 'standard' forcing for all models to use before conducting optional experiments.

It is true that considering several choices for the forcings is new to the PMIP LGM set up but not new to PMIP, since the PMIP3 last millennium simulations (Schmidt et al., GMD, 2011, doi:10.5194/gmd-4-33-2011 and 2012, doi:10.5194/gmd-5-185-2012) already considered spanning different forcing files to account for the uncertainties in our knowledge of the forcing. We explain this new philosophy for the LGM in the introduction, on page 4, lines 22-25. We will add more about this topic in the outlook which the other reviewer 3 required to expand.

## Minor specific comments

1. It would be easier for readers, including non-modellers and modellers who have never done the LGM experiment, to overview the LGM experimental design before going into the details if two tables are provided which separately summarise the forcing and the sensitivity experiments (e.g. LGM\_PI\_ghg, LGM\_PI\_ice, etc.)

We will add summary tables in the revised manuscript. These will be added to Section 3, which are actually giving this overview before the detailed description of the implementation of the boundary conditions. It should indeed be useful to add these summary tables, to give a concise view of the experiments as well as to provide the modeling groups a quick way to check their set up and to give them the main outline for documenting their simulations.

2. Based on the PMIP3 experience, it would be very useful to define the 'equilibrium' or to define some indices which show the closeness to the true equilibrium. Some descriptions were already given (e.g., SST trend, but not entire ocean trend), but is it possible to make all indices a bit more objective (e.g., the trend of variable X for the last 100 years of the spinup)? I do not think it is necessary that all models are very close to the equilibrium, but the information on how far from the equilibrium is useful to safely interpret the result.

The reviewer is right that the description of the spin-up is crucial and we will provide more detail in the revised manuscript, in response to his comment as well as other comments to the manuscript.

3. The authors mention climate sensitivity and close collaboration with the RFMIP. They also suggest a LGM AGCM simulation. Would it be useful to request each model to provide effective radiative forcing as in the RFMIP (if AGCM simulation is to be done already as suggested by the manuscript)?

The AGCM run mentioned in the paper is intended as a step for initializing the coupled model. Since we allow modelling groups to start from a previous glacial simulation to gain time on spin up, the AGCM simulation is not the one exactly needed for a thorough RFMIP-style calculation. This would require an AGCM run with *piControl* SST and sea-ice, the rest of the set up being the *lgm* one. Using such a simulation poses questions, though, because of changes in land sea mask (cf. Braconnot et al., NCC, 2012, DOI: 10.1038/NCLIMATE1456) and because the sea-ice and SST also play an important role in setting up the LGM climate. Another point of view would be to evaluate the difference in radiative forcing between *lgm* and *piControl* conditions taking the LGM as a reference and consider the *piControl* as a perturbation. Although this should be symmetric compared to the RFMIP protocol, in theory, it should give comparable results. In practice, due to the differences in land sea mask and SST/sea-ice basic states, it is unlikely to yield exactly the same results. As a result, we think that the topic is not really mature to coordinate experiments within PMIP4-CMIP6 but should rather be investigated first within a working group, which would ideally also relate to the Past2Future PMIP working group (<https://wiki.lsc.ipsl.fr/pmip3/doku.php/pmip3:wg:p2f:index>), which works on climate sensitivity. Several groups mentioned their interest in the topic. This would strongly benefit to interactions between RFMIP and PMIP.

We will add a few sentences about the interactions with RFMIP and the link with the "climate sensitivity" question in Section 5, which Reviewer RC2 asked to expand.

## Review RC2 (12 April 2017)

This is a clearly written and scientifically well motivated paper describing the numerical experiments designed to mimic the effect of the last glacial maximum (LGM) on the climate system as a contribution to CMIP4. Much of the paper, appropriately so, reads like a technical guide to performing the simulations. This seemed to have been well done, but a meaningful evaluation would require an attempted implementation of this guide, to actually perform the experiments. This is not something I was in a position to do, nor do I have experience in directly configuring a model to perform numerical experiments with this degree of modification to the model; hence my ability to substantially review this, the most important part of the paper, is limited. However this perspective motivates one of my comments below.

1. The scientific backdrop, in the form of the four questions that place the LGM experiments in the context of CMIP6 was developed well, but more specificity and follow through would have strengthened it. For instance, on the climate sensitivity question wouldn't it be helpful to draw out more specifically the important role of tropical sea-surface temperatures, both with respect to some controversial aspects of the reconstructions (cf. Annan and Hargreaves, 2015 Quat. Sci. Rev.) and their perceived potential to rule out particularly high values of ECS (eg. Stevens et al., 2016, Earths Future). Scientifically the manuscript would be stronger, and the subsequent analysis would be easier, if time was spent articulating a few specific hypotheses as to how the PMIP simulations might contribute to better bounding climate sensitivity, or informing estimates on the bounds of forcing. Coming back to the questions at the end of the manuscript would also unify the presentation.

[This comment is well taken and we will add precisions on the works which motivated the four topics of analyses overviewed in section 2 as well as developing section 5 \(Analyses and outlook\).](#)

2. Very much related to the above, the manuscript needs a more thoughtfully prepared and substantive conclusions. At the moment it leads with platitudes such as "The LGM ... provides a demanding test of model reliability..." or 'will create an unprecedented data set'. The first point is false, the second says nothing. The LGM might very well test a group's ability to create a model that is adaptable to our understanding of two different climate states, and be no measure of its fidelity to predict the state of the system where the answer is not known ahead of time. Concentrating more on the specifics of the questions raised in section 2, and the hypotheses that can be drawn from them would strengthen the manuscript.

[We will elaborate more on the questions outlined in section 2 in section 5, along the lines suggested by the reviewer for the "climate sensitivity" question \(+ the link with RFMIP, as outlined in the response to Reviewer RC1\). We will come back to the questions raised in section 2 and expand on the benefits of interacting with other CMIP6 MIPs.](#)

3. Given the complexity of the experiments, some checkpoint mechanism should be included. My suggestion is that after configuring the land-seas masks groups should be asked to publish in their documentation, but also simply check, their clear sky upward shortwave, either from the runs themselves or from a control period of a year or so. Here I imagine that the LMD group provides netcdf output of annually averaged clearsky reflected shortwave radiation, and numbers for its value over ice-free ocean, and over land/land ice sheets. The numbers, and plots should be given in the manuscript in the form of a figure and a table, and the full 2D fields should be provided as netCDF files for groups to compare to. The clearsky reflected shortwave radiation is not everything,

but it is a good first indicator of the properties of the ice-sheets and (over the ocean) of the strength of the dust forcing. Having these in the same ballpark is an important check on the plausibility of the experiments, and avoids the problem of groups unwittingly using a very different and less plausible forcing. I also suggest checkpointing the bathymetry (page 11).

It is indeed a good idea to provide check points before the groups engage on running their full simulations. Before examining the clear-sky short wave radiation, which could depend quite heavily on the model's parameters, we strongly encourage the groups to check their land sea mask, land ice mask and orography, not only in the boundary condition files but also in the output files from the first year of simulation. Radiative fluxes such as upward surface clear-sky radiation can be compared to the previous PMIP2 or PMIP3 results, rather than providing a benchmark from the LMD model only, since the flux can be model dependent. Rather, we prefer to suggest the modelling groups to compare their first results to previous PMIP2 (Braconnot et al., CP, 2007, doi:10.5194/cp-3-279-2007) or PMIP3 runs (Braconnot and Kageyama, 2015, 10.1098/rsta.2014.0424), which show the shortwave radiative forcing estimated via the simple method of Taylor et al., 20067, DOI: 10.1175/JCLI4143.1). Numbers and maps can be found in the references cited above, which we will clearly refer to in the section 4.10 (documentation of the simulations) which will be expanded.

Another checkpoint is that the altitude of the ice sheet is actually implemented in the model. This can be checked by examining the atmospheric circulation over and around ice sheet areas. This point will also be added to the manuscript.

4. I suspect experience will show that the description of the procedures is insufficient to perform the experiments. Here I think a procedure for collecting such experiences to improve the technical documentation for setting up and running the experiments would be beneficial. I could imagine that this be handled in the form of a forum, but ideally this would be in some way connected to the manuscript, and/or lead to a revised procedural description after these experiences have been collected. I encourage the authors and editor to work together on a way to incorporate this feedback in the PMIP documentation.

We strongly encourage the modelling groups to document their simulation and provided a guideline for this in section 4.10. This section will be completed with the checkpoints as recommended by the reviewer in point 3. This documentation could be published as a GMD paper in the PMIP4 special issue, which we have asked with this goal in mind. Precisions on the protocol could also be published, if necessary, as a follow-up of the present manuscript, as was done in PMIP3 for the last millennium (Schmidt et al, GMD, 2011 and 2012).

5. The manuscript/protocol should be more demanding of the documentation (§10). For modelling groups to be considered within PMIP the PMIP community should set some standards of documentation. Point 2 above is an example of such a possible standard, but there are certainly more. Indeed having some more synthetic measures included as tables in each groups documentation will be very helpful for subsequent meta studies. A clear minimum standard of documentation should be demanded for the participating groups to be considered for entry into the PMIP community.

There will be several ways to document the simulations, among which publishing a description in GMD, along the guidelines of section 4.10, which will be completed. We are also in contact with the

group working on the CMIP6 documentation so that the required documentation on the DECK reference experiments matches the documentation required for the LGM one.

6. The figure quality is poor. Figure 1 does not even present the units of what is being plotted and the caption is not complete. In other figures the projection is changed for no apparent reason and in some cases (Fig 2) apparently sub optimal. For instance why not polar projections for Fig 2. Fig 5 needs improvement as it is difficult to use quantitatively, and some homogenization of color scales is required. Earth is distorted in its aspect ratio for no apparent reason in Fig 6.

Finally having the terminator at 190 deg (or centering the maps at 10 E better separates Asia from North America.

We will try and improve the figure quality. For figure 2, we will not choose polar projections because this figure also shows differences in altitude outside glaciated regions (of the order of 80 m) due to global sea-level drop and this is important too.

Additional, somewhat more minor points include

- P4.L11: 'carbone' should be 'carbon'

This will be corrected.

- P10: In the discussion of ice sheets, I find it hard to imagine that it would be possible to ascertain the effect of the differences being discussed. Is there experience to suggest otherwise?

The differences between the ice sheet reconstructions are quite large (several hundred metres for the North American ice sheet, as illustrated on Fig. 2). Beghin et al (Clim Dyn,2016, DOI 10.1007/s00382-015-2720-0) showed that differences in the implementation of the PMIP3 ice sheets for the different models likely led to differences in the jet stream position over the North Atlantic and these were much smaller than the differences between the ice sheets considered in the PMIP4 protocol. We therefore do expect differences due to the ice sheet reconstructions and this is why we considered these options in the protocol. We discuss this in section 2 but could expand the discussion if required. This is a topic which could be added to the expanded "Analyses and outlook" final section.

- P10: The first time the variables 'sftxx' are introduced it might be helpful to explain their naming convention so as to better fix them in the reader's head.

This precision will be added.

- P10.L5: Wouldn't it be helpful to give some numbers to compare to. For instance downward insolation in JJA in NH and SH separately.

As discussed above in the "check point" question, we will point to references giving these numbers, based on the PMIP2 and PMIP3 results.

- P12.L19: 'illustrating the impacts ...' should be 'illustrates the impacts ...'

This will be corrected.

- P13.L15: The apparent un-availability of the river routing makes it appear that the experimental description is not complete. These files need to be included and their main differences from the present noted before the manuscript is published.

After rediscussing this specific point with the modelling groups, we agreed to give major guidelines on which rivers are affected by the ice sheets and changes in land sea mask. Precise river routing files could be provided later on but do not appear as crucial in the set up for the participating groups. We will therefore update this part of the protocol.

- P16: Please explain the rfip naming convention, i.e., realizaiton, forcing, initialization, physics?

The meaning of rfip will be added.



## Comment SC2: 'Deep-ocean circulation and equilibration issues', Alice Marzocchi and Malte Jansen (25 Apr 2017)

It has been found that the AMOC does not adjust monotonically to glacial forcing (e.g. Stouffer and Manabe, 2003; doi:10.1007/s00382-002-0302-4) and that the integration time in many PMIP simulations is likely insufficient to reach deep-ocean equilibration (e.g. Zhang et al., 2013; doi:10.5194/cp-9-2319-2013). Moreover the criteria for equilibration requested by the PMIP protocols are likely to be insufficient (Zhang et al., 2013). While this issue is addressed in the manuscript (page 15 at lines 25-28), the specific requirements remain somewhat unclear and should be rephrased. Specifically, we believe that a quantitative criterion for the drifts in the AMOC (see Figure 1 as an example), deep-ocean temperature and salinity should be provided. While it is expected that not all simulations in the archive will be integrated to full equilibrium, the level of deep-ocean equilibration needs to be clarified in order to avoid erroneous interpretations of modeling output that is dependent on transient effects. This is already discussed to some degree on page 15 (lines 29-31), but it should be clarified what exactly is the information requested as part of the documentation. Among other things, we believe that time-series of abyssal temperature, salinity and AMOC need to be provided. Finally, while we understand that the complete time-dependent data for all simulations cannot be stored on ESGF, the modeling groups should be encouraged to provide time-dependent data upon request.

A number of studies have pointed towards the important role of temperature and sea ice changes around Antarctica in controlling deep-ocean circulation and stratification (e.g. Shin et al., 2003 doi:10.1007/s00382-002-0260-x; Ferrari et al., 2014 doi: 10.1073/pnas.1323922111; Jansen and Nadeau, 2016 doi:10.1175/JPO-D-16-0084.1; Klockmann et al., 2016 doi:10.5194/cp-12-1829-2016; Jansen, 2017 doi:10.1073/pnas.1610438113). The PMIP models show a widely varying sea-ice extent in the LGM, with most models likely underestimating the abundance of sea ice (e.g. Roche et al., 2012 doi.org/10.1016/j.quascirev.2012.09.020; Goosse et al., 2013 doi.org/10.1016/j.quascirev.2013.03.011). It may therefore be worth to point out that the representation of climatic changes around Antarctica – in particular sea ice expansion and export – should receive careful attention when setting up and analyzing the PMIP4 LGM simulations.

The lack of deep-ocean equilibration, and the unrealistic representation of sea ice formation and export, may explain many of the apparent inconsistencies between different PMIP models and with the LGM geological record.

We thank the authors of this comment, as this, together with the comments from the other reviewers, will help refining the required documentation, in particular in terms of spin-up, which will not be archived but should be described in the documentation. We will include the suggested variables (deep temperature and salinity, sea-ice) in the list of variables with which the spin-up should be documented.

Figure 1 - The data shown are from the LGM simulation with CCSM4 described by Brady et al. (2013; doi:10.1175/JCLI-D-11-00416.1). The additional data for part of the spin-up and for a continuation of the PMIP3 simulation are available from the Earth System Grid at NCAR archive. Years of integration are calculated from a branching point of the PMIP3 spin-up. The final 100 years of the PMIP3 simulation are stored in the ESGF archive. The trend for the 100 years of simulation available from the PMIP3 archive clearly highlights that AMOC drifts cannot be determined from such short

timescales, due to the high internal variability. The manuscript should therefore specify that longer time series need to be analysed to test the equilibration criterion of a “stable Atlantic Meridional Overturning Circulation” (page 15, line 25).

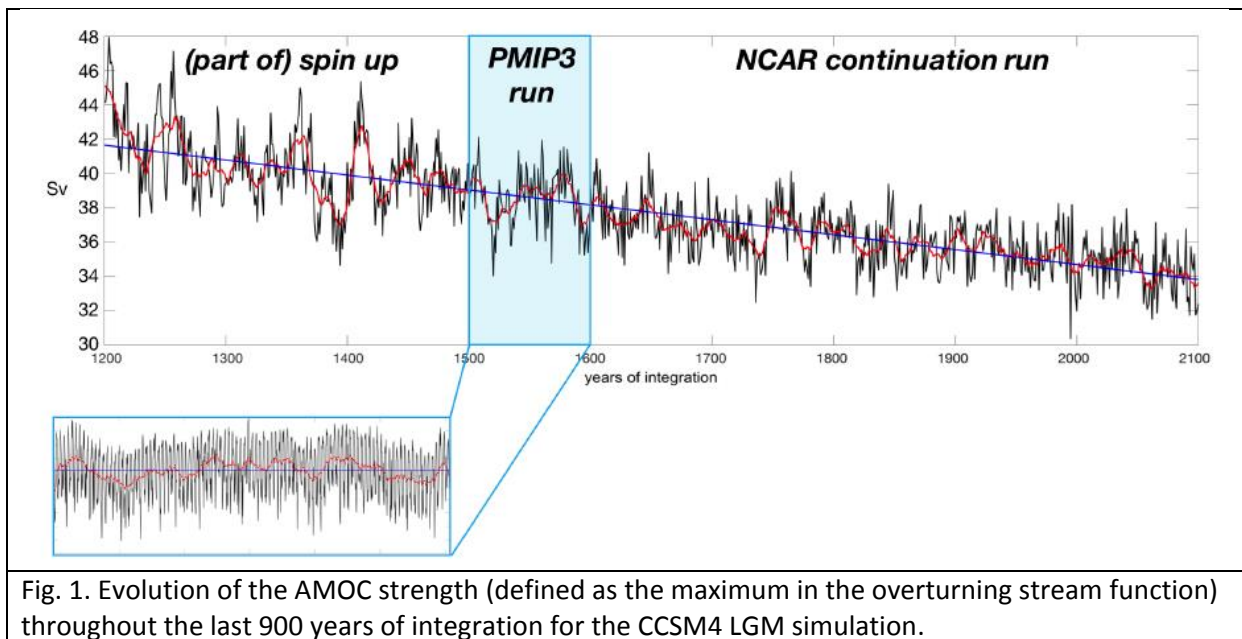


Fig. 1. Evolution of the AMOC strength (defined as the maximum in the overturning stream function) throughout the last 900 years of integration for the CCSM4 LGM simulation.

### Editor Comment (James Annan)

I hope that you will submit a revised version taking account of the reviews and comments received.

The manuscript doesn't say very much about outputs and their storage. Are there any plans for this (yet)? It would be useful to provide a list of outputs that you would like to see - perhaps tiered into those which are considered a minimum requirement for participation, and additional outputs which would be useful.

We indeed did not say much about the required output. We will be more specific for the spin-up phase (in line with proposed changes in the “documentation” section) as well as provide the output list as was done by the other PMIP4 manuscripts. The results will be stored on ESGF as for the PMIP4 and PMIP4-CMIP6 experiments.