

Reponse to reviews of “PMIP4-CMIP6: the contribution of the Paleoclimate Modelling Intercomparison Project to CMIP6” by Kageyama et al.

[for clarity, we reproduce *the comments by the reviewers, editor and CMIP panel in blue/italic* and provide our answers in black]

We thank the reviewers for their comments which helped focussing and clarifying the manuscript. We have attempted to reply to all their comments. In particular, we have:

- reduced sections 1 and 2 and removed figures 3 to 5 which were published elsewhere
- expanded sections 3 and 4, with a more complete protocol and a more complete analyses plan, including an overview of the analyses plans for individual experiments.

We have also attempted to harmonise the text and to remove inconsistencies.

Reponse to Anonymous Referee #1

There is little doubt that PMIP made significant contributions in assessing the role of different forcings on past climate changes. The strength of a MIP comes of course from the clear definition of boundary conditions, forcings etc. Hence, it is absolutely useful to describe these as clearly as possible in the peer-reviewed literature. Having said this, I am struggling with the purpose of this ms. Large parts read like an (unconvincing) attempt to justify PMIP4. Instead of repeating what was done in PMIP 3, the ms. would be much more convincing if the authors could outline which insights into climate processes were gained that would have been impossible without PMIP (using a few examples). In my view, there is too much description of modeled changes and matches/mismatches with proxy data but too little information on real insights into climate processes - especially on processes that are of relevance outside the paleoclimate community.

We have shortened Sections 1 and 2 in agreement with this comment, removed Figures 3 to 5 which were describing previous results in terms of model-data comparisons and added examples of how PMIP can help and illustrate the processes at work in climate model responses to forcings. The available paleoclimatic reconstructions, which were also illustrated on these figures, are now the topic of a special section of the analyses plan (Section 4.1). We have attempted to complete the protocol wherever it lacked information, so that all necessary information to run the PMIP4 experiments for CMIP6 can be found in the present manuscript.

My understanding is that the ms. should serve as an entry point to a series of more specialized descriptions of the experimental setup for the individual time slices. While such a publication strategy seems very useful, it also means that repetition can and should be avoided. In my view this overarching ms. should have a clear focus on what is now sect. 3 as well as on common issues among the experiments.

This particular topic has been discussed between the CMIP Panel and the Editor. We have followed the CMIP Panel recommendation and added details on the protocols so that their description is complete. The complete justification of these protocols, though, would take too much space in the present manuscript and is developed in the companion papers on each period. These papers also detail the sensitivity experiments based on the PMIP4-CMIP6 experiment. They are being submitted

to GMD. The present manuscript will be held until acceptance of the companion paper and modified if necessary to guarantee consistence between all papers.

Here, I am surprised that some important aspects are only mentioned in passing:

- River routing: the recommendations are quite vague for the LIG and the Pliocene (recalling, for example, that the modern river-system draining into the Arctic developed to a large part after the LGM).

We have now devoted a section to river routing. This was previously in the section about implementing the ice sheets and it is true that there was no specification for periods other than the LGM. Our understanding, though, is that we are currently lacking data sets to take these changes in river routing into account for the LIG and mPWP experiments. We have therefore made the conservative choice of not requiring changes in river routing in those experiments. This could of course be the topic of additional sensitivity experiments, as we know this could be of importance for the oceanic circulation (e.g. Alkama et al, 2008, for the Last Glacial Maximum). This is now clearly stated in the manuscript (Section 3.3).

- Which plans exist to assess the results from LGM experiments using different ice sheet configurations?

It is the first time that we consider several possible ice sheets in our LGM experiments, and this is in acknowledgement that several approaches lead to significantly different reconstructions. It is therefore very important to assess these results and this is now better highlighted in our manuscript (Section 4.2). A first analysis will be to evaluate those simulations in comparison to paleo-climatic reconstructions and assess whether this comparison yields different results for the different ice sheet reconstructions. A second step will be to understand the differences between the climates simulated with the different reconstructions. This is why we encourage groups which can afford it to run several simulations with these different reconstructions. The “LGM ice sheets” working group has been set up to investigate these questions.

- Spin up: how is an insignificant trend defined in the framework of PMIP?

The criteria for defining a spun-up experiments are now quantified in section 3.6, with the same criteria as for PMIP2: “We recommend that the spin-up should be run until the trend in global mean sea-surface temperature is $<0.05\text{K}$ per century and the Atlantic Meridional Overturning Circulation (AMOC) is stable; a parallel requirement for carbon-cycle models and/or models with dynamic vegetation is that the 100-year average global carbon uptake or release by the biosphere is $<0.01 \text{Pg C a}^{-1}$ ”

- I was surprised to read nothing about the calendar problem [Joussaume and Braconnot,1997], that featured high in earlier cycles of PMIP.

This is absolutely right and the calendar “problem” should still be dealt with, of course. We acknowledge that it would be very difficult for most modelling groups to adapt their online averaging procedures to account for the different definitions of the months for *midHolocene* and *lig127k*. This is why we require daily data for a few key variables so that the differences due to the changing calendars can be computed offline. This is now explained in section 4.5 and illustrated (Figure 5).

While I am generally very much in favor of the PMIP community to publish their experimental strategy, the current ms. needs considerable re-writing to become a useful contribution. I would suggest to drastically reduce the length and to focus it on new aspects:

- Sections 1 and 2 contain almost no new information and should be largely replaced by a concise summary of dynamical insights gained from earlier PMIP phases (see above)

We have considerably shortened these sections and added a few examples of “dynamical insights”.

- Figs. 2-5 were published elsewhere and should be removed

We have kept Fig. 2 which was created for this manuscript but removed all others as required.

- Focus should be on what is now sections 3 and 4.

These sections have been extended with a complete protocol and more detailed analyses plan.

Finally, someone out the large group of authors should read the ms. from beginning to end to ensure that the wording/style is consistent (incl. the ref. list; “et al” vs. et al.”; paleo vs. palaeo) and that geological ages are correct and consistent among the ms. (mPWP 3.2 Ma vs. 3.3-3 Ma; incorrect start of the Quaternary at 2.5 Ma).

Thank you for spotting these. We have done our best to have fully consistent text in this new version. The confusion about the dates for the mid Pliocene Warm Period might arise from the fact that while PRISM, the project which specifically deals with climate reconstructions for this period, has considered the full mid Pliocene Warm Period (from ~3.3 to 3 million years ago), the second phase of PlioMIP specifically focuses on a warm interglacial within the mid Pliocene Warm Period, and this interval, termed KM5c, has been dated to 3.205 Ma ago. We have added a sentence to clarify this, in section 1.3: “The mPWP experiment focuses on a specific interglacial period, dated at ~3.2 Ma before present, during the mid-Pliocene interval (3.3 to 3 Ma before present).”

Response to Steve Sherwood

This paper nicely lays out the plan for PMIP4, which is part of CMIP6. It describes the history and rationale for this project, then reviews the experiments which cover five periods in recent Earth history and expand on those of previous PMIPs. The protocols for these experiments are described, a few potential pitfalls are noted for modellers to beware of, and useful links to other MIPs are noted. I think the project sounds exciting and hope that plenty of modelling centres are able to participate. This article would benefit from some very minor revisions to improve clarity, but otherwise I am happy with it and don't see any major omissions.

We thank the reviewer for his enthusiasm. Despite this support, we have largely revised the manuscript according to the other reviewers' comments and the editor's comments.

The minor revisions to suggest are:

6, 18-29: Please mention (as I gather later though am still not 100% sure) that all of these experiments are "time-slice" experiments, i.e., the model is run to statistical equilibrium with time-invariant forcing specified based on the quoted point in time to give a representative mean state, except for the last millennium which is transient and includes time-varying forcings such as volcanic eruptions.

This was already mentioned in section 3.1 but we have added the information in section 1.3 too, following this comment.

8, 19: The word "observed" should be reserved for the instrumental period with real observations. Can we instead say "proxy-estimated" (or similar)?

We distinguish between primary observations, such as pollen counts, geomorphic features, or isotopic measurements, and the climate inferences or reconstructions made using these observations. We have therefore now distinguished between palaeoenvironmental observations and climate and/or vegetation reconstructions consistently throughout the text.

8, 29: By "mean values" do you mean global means? Time averages? Changes in global mean relative to modern?

This sentence has disappeared from the text following our decision to shorten Section 2.

10,40 By "trends" do you mean differences (relative to modern)?

We actually meant both differences relative to modern and trends, for transient simulations of the last interglacial. This has now been clarified.

11, 36: please fix error message

The error message has been fixed.

11,34: suggest "current" rather than "modern" (which can mean many things... from a palaeo perspective preindustrial could be viewed as modern)

ok, done.

17, 36-37: I don't think the polarity of the forcing is the real problem, but rather, the fact that some radiative forcing agents produce larger responses per unit global-average power input than do others, and/or provoke "rapid adjustments" to the forcing that are unrelated to global-mean warming or cooling. This is due to the spatial pattern of the forcing. Given the fact that a major goal here is to test model responses to forcings, and given that past forcings are different from the dominant ones we worry about for the future (greenhouse gases), this topic may deserve a bit more discussion.

It is true that the spatial patterns of the different forcings (and their responses) can be very different and they will be probably even more different when the dust forcing is included, because it introduces a radiative forcing which is quite regional. We have rephrased the sentence which could be misinterpreted due to the allusion to different forcings being responsible for cooling and warming. We have also introduced the single forcing experiments which are coordinated in PMIP4 and will help disentangling the response to individual forcings and their contribution to the total response.

Table 2: some of the cells in the table are blank, and I am not sure what this means. Suggest every box should say something (even if it is "see text")

All boxes of table 2 are now filled.

Figure 1. I found this figure confusing; maybe you are trying to cram too much into one panel. The caption refers to panel labels (a, b, ...) but there aren't any. It is hard to figure out what each curve is and which axis refers to what (especially when there is one on each side, or where it switches from right side to left going from one column to another). I would prefer the axes to be individually labelled with the quantity and units, or give the quantity in the title or above the curve and the units on the axis so we know we are looking at the right one. It is not at all obvious that the grey banded stripe is meant to represent the possible range of CO₂ in the left column; nothing wrong with showing it this way but please make clearer what everything is!

We apologize for this confusion. The labels have been added. Colored axes are used so that it is clear which axis a given line corresponds to.

Figure 2a. I presume there are black dots hiding underneath the red ones for mid- Holocene and Historical? To help avoid confusion perhaps the caption could mention this.

Yes, this is right and indicated in the caption.

Figure 2b. Please give units! Also, please explain in the caption what "OAV" and "OAC" refer to.

The unit (mm/year) has been added to the caption.

Figure 3. Please be consistent in labelling the panels (they are given (a,b) in the figure but you say lhs and rhs in the caption). Is this for land only, or land+ocean? What is STSI and ssTSI? Does 31 points mean 31 years? (give the time width of the smoothing window rather than the number of points)

Figure 3 has been removed following the other comments on this manuscript.

Figure 4. It doesn't appear to me that the proxy data are able to tell us anything about the seasonal cycle, given that the differences are small compared to the scatter --- so is it worth including the two

right panels in a review paper on PMIP? Especially since the figure is reproduced from another source so anyone who really wants to see the seasonal results can find them. I see later you are already requesting permission to use a portion of another IPCC figure, I'd suggest making a similar request here.

Figure 4 has been removed following the other comments on this manuscript.

Figure 7. Please spell out "preindustrial" rather than PI since PI is not one of the study time periods and you haven't used this acronym much.

This information has been placed in the caption as it was not fitting in the title of the plot.

Response to Anonymous Referee #3

This manuscript details the contribution of the paleoclimate modeling community to the new Coupled Modelling Intercomparison Project. For this phase the PMIP community is planning to expand its contribution significantly, including 5 different periods or experiments. In order that as many modelling groups as possible participate in this effort a clear modelling setup. This manuscript gives an overview of the rationale and broadly outlines the experimental setup. Hence the importance of this paper is clear

General Comments

I find the manuscript a bit long without going into the specific details of each experiment. This is ok as it is mentioned that there will be special papers for each one of them. I would suggest shortening section 1, and section 2 could benefit for a clearer discussion on current modelling gaps.

We have indeed considerably shortened Sections 1 and 2 and removed figures 3 to 5. Section 2 was originally written to summarize previous work on each period, including aspects which models could not account for. We have chosen to keep this section short and have not added information but have also attempted to clarify what is new in the PMIP4-CMIP6 experiments compared to the previous phases, e.g. including the dust forcing, which is new for the *midHolocene* and *Igm* experiments, including the analysis of the sensitivity of the *Igm* results to the imposed ice sheets, using an improved volcanic forcing for the *past1000* experiment. In our opinion, this addresses “modelling gaps” in a positive manner.

I have also personal comment. Having worked with the past1000 simulations, it would be very useful if all the modelling groups planning to run this experiment, would run it up to present. (850-2005 CE instead of stopping in 1850).

This is an important point which has been added to section 2.2 and explained in the specific manuscript on the past1000 experiment (Jungclaus et al, 2016).

Figures: I am not familiar with the journal's policy in this respect, but it seems odd to me to use so many previously published figures. Also the quality and style of the figures is too diverse. For example, please use just one projections for all maps.

We have removed Figures 3 to 5. Figure 6 (now Figure 3) has been redrawn with projections similar to the other maps displayed in the paper.

There is a lot of inconsistent naming of the experiments throughout the text. Please use either “midHolocene” or “MH”. “LM” or “past1000” and so on.

We agree this could be confusing. We use different names for the periods and the experiments. This is now clearly explained in Section 1.3. We have been careful in updating the manuscript accordingly.

Specific comments

Page3, Line 13: Missing comma after However. Or are you using “however = to whatever extent” ?

A comma has been added.

Page3, Lines 14-15: used twice “well outside”

This repetition was made on purpose so that the two facts could be linked.

Page 5, lin34: include years for “the last interglacial”.

done.

Page 6, line 1: : :of participating MODELLING groups.

done

Page 6, line5: in this section maybe include the CMIP6 questions?

We did include the CMIP6 questions in the first paragraph, but do not feel it is necessary to lengthen the title of the section since we also discuss the link with the CMIP6 key questions in Section 4 (plan of analyses).

Page6, line 7: is “How does the Earth System respond to forcing” question 1? How many questions are there? Please introduce a bit better.

This is the exact formulation of the CMIP6 first key question, this is why we give it between quotes. The reference to Eyring et al (2016) has been added to indicate that this formulation does not depend on us. Examples of analyses of PMIP4 results to answer these questions are further given in the analysis plan.

Page 6, line35: define DECK

done, with a reference to Eyring et al, 2016.

Page 7, line 20: change “interesting” for “valuable”

done

Page 7, lines 41-42: please be consistent between title and text about MH and LGM.

As explained above, we use “MH” and “LGM” for the periods, and *midHolocene* and *lgm* for the experiments themselves. We hope that it is clearer now.

Page 10, line 24: question: is there a separate paper with specifics?

Yes, this is indicated in Section 3.

Page10, line 34: needs a reference.

A reference to Capron et al (2014) has been added

Page 11, line 36: there is a typo.

The faulty reference has been removed. The following sentence has been reformulated.

Page 13, line 27: delete “web sites”. Its repeated.

The sentence has been fixed (and has no occurrence of “web sites” now).

Page 17: line 22: delete “the”

done

Figure 1: needs more information of the different panels. Very hard to understand at this point

We have modified the figure – cf. response to the second reviewer.

Response to the CMIP Panel

Dear PMIP authors,

The CMIP Panel is undertaking a review of the CMIP6 GMD special issue papers to ensure a level of consistency in answering the key questions that were outlined in our request to submit a paper to all co-chairs of CMIP6-Endorsed MIPs. These questions are outlined in the overview paper (Eyring et al, GMD, 2016) and the relevant section is summarised below:

'Each of the 21 CMIP6-Endorsed MIPs is described in a separate invited contribution to this Special Issue. These contributions will detail the goal of the MIP and the major scientific gaps the MIP is addressing, and will specify what is new compared to CMIP5 and previous CMIP phases. The contributions will include a description of the experimental design and scientific justification of each of the experiments for Tier 1 (and possibly beyond), and will link the experiments and analysis to the DECK and CMIP6 historical simulations. They will additionally include an analysis plan to fully justify the resources used to produce the various requested variables, and if the analysis plan is to compare model results to observations, the contribution will highlight possible model diagnostics and performance metrics specifying whether the comparison entails any particular requirement for the simulations or outputs (e.g. the use of observational simulators). In addition, possible observations and reanalysis products for model evaluation are discussed and the MIPs are encouraged to help facilitate their use by contributing them to the obs4MIPs/ana4MIPs archives at the ESGF (see Section 3.3). In some MIPs additional forcings beyond those used in the DECK and CMIP6 historical simulations are required, and these are described in the respective contribution as well.'

We very much welcome the PMIP contribution and the detailing of the experimental design, analysis plan and diagnostic output that you currently cover in sections 3 and 4. We also welcome the strong links that PMIP has clearly forged with other CMIP6 MIPs and look forward to the joint analysis that you describe.

We thank the CMIP Panel for these positive comments

Additionally, we would like to see some more detail on some of the issues raised above, notably;

a. More discussion on the specific goals of PMIP4 in CMIP6 and what science gaps it is attempting to fill. You describe the 3 CMIP6 science questions and PMIP links to them in Section 1.3 and the links to the WCRP GCs in section 4.3, but it would be good to see some discussion on what PMIP4 is hoping to achieve that is new since PMIP3.

We have largely re-written section 4 with this goal in mind.

b. The description of the experimental design for each experiment is comprehensive and very useful. There are however, a worrying large number of papers 'to be submitted'.

Is it clear that once this paper is published modelling groups will be able to rely on it to provide a comprehensive start point for setting up their experiments?

This is right, and we have clarified the situation in the revised manuscript. All necessary information to set up the experiments is now given in this manuscript. The accompanying manuscripts (now submitted except for the LGM paper which has now been sent to co-authors) are giving additional

justifications and technical explanations about this set-up and the design of complementary (sensitivity) experiments.

c. A lot of focus in the analysis plan is given to the multi-time period analysis, but not all modelling centres will be contributing to all (or indeed in some cases more than 1) of the entry card/tier 1 experiments. Could more be said about analysis of the specific experiments and what new we will learn from these experiments since PMIP3-CMIP5 (higher resolution, better data, more ES components)?

We rely on Table 1 to give a summary of the analyses which will be carried specific experiments and explain it in Section 4. We have chosen not to lengthen the manuscript with extended analyses plans for each periods because these are given in the companion papers, in which they are better articulated with the associated sensitivity experiments. On the other hand, we have given more detail on the reconstructions available to evaluate the models for the five PMIP4-CMIP6 periods, hence giving more details on the possible benchmarking analyses, in line with the fact that CMIP6 required data to be associated with each MIP.

d. You make the point that the comparison of these time periods to palaeodata is one of the key drivers but say very little about the observational data sources or whether these products will be made available to the community to facilitate comparison. In section 4.4 you describe the new metrics and forward modelling you request the models output. It would be good to document how these will be evaluated.

We have dedicated a new section to the available paleodata, pointing to available data sets (Section 4.1). This is by no means exhaustive but demonstrates the existence of reconstructions for each PMIP4 period. This is a very active area of research and there will be new syntheses, which we have indicated too.

We hope you agree that some level of consistency across the MIP papers in this special issue is valuable and that the above suggestions can be accommodated in your paper.

Other comments:

e. There is a lack of consistency in the naming of the experiments e.g. notably the use of LM and past1000 somewhat interchangeably. Please clarify

We use different names for the periods (LM: last millennium, MH: mid Holocene, etc) and the associated experiments (past1000, midHolocene, etc). This is now clarified as early as possible in the manuscript, in section 1.3.

f. In section 3.2, the implementation of the ice-sheets needs to be a bit clearer. For example do all points 2-5 refer to both midplioceneEoi400 and lgm?

This has been clarified:

Steps 2 and 3 are compulsory for both experiments,

With many thanks for your ongoing efforts in the CMIP6 process.

The CMIP Panel

Response to the Editor's comments

[21 July 2016]

I think most things have been covered by the reviewers and the CMIP panel's comment, so please enjoy responding to their comments. In common with one of the reviewers, I'm not enchanted by the history lesson. I understand that you want to draw people into considering paleoclimate, but I think the resulting length of the paper is more likely to turn people away.

We have shortened the history lesson and re-focussed the manuscript on the protocol and analysis plan.

However, the real problem for publication of this paper in GMD is the possibly incomplete protocols.

From Section 3: "The modified forcings and boundary conditions for each PMIP4-CMIP6 palaeoclimate simulation are summarised in Table 2. The complete details of the experimental protocols are given in a series of companion papers: Otto-Bliesner et al for the midHolocene and lig127ka experiments, Kageyama et al for the lgm, Jungclaus et al for the past1000 and Haywood et al (2016) for the midPliocene-eoi400 experiment. These papers also explain how the boundary conditions for each period have been built and constitute key references for the experimental protocol for each of the PMIP4-CMIP6 simulations."

The problem, as highlighted by the CMIP panel, is that most of these papers are not published. I can't control what is going to be in those other papers - all I can do is make sure that this paper meets the peer review requirements. Therefore, for all experiments please include in this paper, "the complete details of the experimental protocols" for the PMIP4-CMIP6 experiments. You can leave the details on "how the boundary conditions for each period have been built" to the still to be submitted papers. You can also leave all alternate experiments that are within PMIP4 but outside CMIP6 to those other papers. Please make sure that Table 2 is edited so that it does not appear that essential details required for setting up the experiments are included in these unpublished papers. Basically, a modeller should be able to set up the PMIP4-CMIP6 runs using the information contained in this paper.

The alternative is that we put this paper hold until the other papers have passed through peer review. This could be workable as I understand that you intend to submit the other papers also to GMD. Even in that case, I would like to see the complete protocols for PMIP4-CMIP6 detailed here (citing the other papers as required), as trying to extract the CMIP6 protocols from the much more elaborate single-interval papers is likely to be a trying process.

The accompanying papers have been or will soon be submitted. We agree that the present manuscript is held until the companion papers are accepted, and to update the protocol accordingly, if necessary, to guarantee consistency between the manuscripts.

[6 October 2016]

I'm also editing the DeepMIP paper, and they happen to have a rather prominently positioned section about data for comparison with the model output, which caused me to pay attention to it, and then criticise it as inadequate! I think that more or less the same criticism can be levelled at this PMIP4 manuscript. Section 4.2 touches on ways in which the models may be compared to data, but this is

less important than outlining the datasets that are available and including specifics about any new datasets that are planned to emerge within the timeframe of the project. Please add this information for all the PMIP4 periods in the revised manuscript. It doesn't have to be a huge amount of material - the point is a practical one - to lead users of PMIP model output in the right direction so they can discover the relevant datasets. I expect you would include citations in the revised section 4.2, but you might also wish to include some additional information (DOIs/weblinks etc) in the Data Availability section.

We have added section 4.1 and one table to give an overview on the available data sets for each period.

[27 October 2016]

In response to my previous comment which highlighted the problem of citing un-submitted works, the authors elected to put this manuscript on hold until those manuscripts were submitted to GMD. Those manuscripts are now almost complete, and the authors have enabled me to see two out of the three of them for approval before submitting to GMD.

*I find there is a real problem, which is that these two draft manuscripts both contain substantial discussion of the background and rationale behind the *CMIP6* PMIP4 protocols. These discussions are too entwined with the non-CMIP6 experiments within PMIP4 to be removed from those draft papers, and including the content within this overview manuscript under consideration would be entirely impractical.*

In terms of the peer review at GMD I find it unacceptable that we should in this overview manuscript be approving the protocols for the CMIP6 runs without first reading the rationale behind them!

All this suggests that the complete protocols should be peer reviewed within those other papers. Since those papers describe the rationale behind the CMIP6 protocol they must also be included in the CMIP6 special issue.

As for the present paper under consideration, a purely descriptive paper about PMIP does not fit into the peer review criteria and I have great difficulty accepting it as a standalone paper. However, if it were presented as the introductory part of a multi-part paper, I think it will be a useful contribution to the whole. This has led me to come up with a solution that solves all these problems, and I think it also produces a nice final product. I suggest that the papers be submitted as 4 parts of a multi-part paper. The titles should be made consistent with each other and include the Part number in the title. They do not have to be exactly this, but something like this would work...

PMIP4-CMIP6, the contribution of the Paleoclimate Modelling Intercomparison Project to CMIP6, Part 1: Introduction and(/or?) Overview

PMIP4-CMIP6, the contribution of the Paleoclimate Modelling Intercomparison Project to CMIP6, Part 2: Eemian and midHolocene

PMIP4-CMIP6, the contribution of the Paleoclimate Modelling Intercomparison Project to CMIP6, Part 3: The Last Millennium

PMIP4-CMIP6, the contribution of the Paleoclimate Modelling Intercomparison Project to CMIP6, Part 3: The Last Glacial Maximum

From the point of view of this paper, the authors are welcome to submit a revised version, but if outline protocols for the experiments are included then the final publication will need to wait until all the other three papers are accepted.

I appreciate the difficulties the authors have had with trying to write a single coherent paper as a contribution to the GMD CMIP6 special issue, but I think it was an intractable problem, because the resulting paper could not pass the GMD peer review criteria of including both the rationale and the protocols. In that context, I'd like to contrast this MIP with one of the others I edited. CFMIP have a huge number of experiments fully described within their paper, including tier 1 and tier 2 experiments. I think it is an elegant paper. However, CFMIP experiments are all highly idealised, and thus there is no great debate to be had in defining the protocols and they can be very simply described. PMIP, on the other hand is trying to model real and disparate intervals in the earth's climate history. This is an order of magnitude more complex, involving several different communities of scientists, and it so it is not appropriate to squeeze it all into a single paper. The multi-part paper will enable these communities to each take the responsibility for the defence of their own experiment protocols.

As stated before, we have accepted to publish this manuscript only after the other three, on each PMIP experiments, are accepted. Nonetheless, we have prepared this revision and entitled it "Part 1" to make explicit the changes we have already implemented following the comments from the reviewers and the editor and the fact that this manuscript is linked to three other manuscripts + the manuscript by Haywood et al, 2016, already published in CP. For the moment, we have kept the complete protocols in this manuscript as recommended by the CMIP Panel.