

Interactive comment on “A model comparison study for the Antarctic region: present and past” by M. N. A. Maris et al.

Anonymous Referee #1

Received and published: 28 November 2011

Overall quality:

The submitted paper by Maris et al. suggests the good idea of evaluating the coupled climate models of the PMIP2 database with respect to their performance in the Antarctic region. This evaluation of the single models is novel and indeed might be of value for both the ice-sheet and climate modeling community. As with other model inter-comparison projects, it would be interesting to see if the skill of a single model exceeds the multi-model mean with respect to a certain quantity and region (e.g., precipitation over the ice sheet). However, the evaluation procedure presented here lacks a clear strategy as the paper simply shows - rather randomly - comparisons of single models with the reference data. The overall aim of identifying the best performing model is not clearly supported by the applied analysis. In addition, the paper lacks a careful prepa-

C1900

ration as the text is sometimes hard to follow and the figures are of poor quality. Since correcting for the here mentioned points implies a complete revision of the manuscript, we conditionally reject the paper but encourage the authors to submit a restructured paper at a later stage.

Individual issues:

1. Does the paper address relevant scientific questions within the scope of CP?

Yes.

2. Does the paper present novel concepts, ideas, tools, or data?

Yes, the evaluation of the single PMIP2 models regarding performance in Antarctica is a novelty.

3. Are substantial conclusions reached?

Although the authors propose two models in their conclusions which are likely among the best ones, the way how they come up with this conclusion is completely missing or at least not shown to a sufficient extent. This is one of the major criticisms.

4. Are the scientific methods and assumptions valid and clearly outlined?

The methods are valid but overall rather poorly described (e.g., the correlation coefficients - which has limited value anyway in the context it is applied here). However, we suggest to use a more sophisticated and better quantifiable method for both the evaluation of the skill of models as well as the comparison of the different models. As a prerequisite, one has to define more precisely the requirements that a simulated temperature or precipitation field has to fulfill in order to be a useful input for the ice-sheet model (e.g., absolute values, relative changes over time, seasonality, ...). While some of these analyses have been performed in the paper, the results are not discussed in context of the ice-sheet model requirements. In the end, a ranking of the models would

C1901

provide a more differentiated picture of how the conclusions are reached. A conceptually well-structured and carefully carried out model inter-comparison has been done by, e.g., Stoner et al. (2009, J. of Climate).

5. Are the results sufficient to support the interpretations and conclusions?

No or at least this is not sufficiently shown with respect to some conclusions. Or in other words: often, we miss the reason for a conclusion or, more generally, the reason for mentioning something in the conclusions. In our view this is largely due to the unstructured approach taken. If a bias is found in one or two models (out of 14), it is not comprehensibly explained why this is important for the search of the best models - especially, if the authors do not chose to rule out certain models on basis of such biases. Page 3592, line 14ff: inverse temperature bias patterns are found for two models in present-day vs. MH. The implications of these findings are, however, not elaborated in much detail. On page 3589, first paragraph, this is called a "compensational behaviour" that "might lead to more realistic temperature at 6ka". What does this mean with respect to the model's performance? Does this make the model useless for ice-sheet modelling? Does this happen only in these two models (in section 4 it is two models, in the conclusions it is "some models")? At least, Figs. 1 and 2 have to be discussed in the context of this, as a small average bias over a region apparently does not guarantee a good spatial behaviour. In this respect, the correlation coefficients do not provide much information by themselves. Page 3592, line 9: the (apparently general) bias in Amundsen and Bellinghausen Sea coast precipitation is mentioned but not put in context of ice-sheet modelling. Is this a crucial feature that needs to be simulated correctly? Should this bias be weighted more than other biases in other regions?

6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?

The PMIP2-data is public and transparent and therefore sufficient. The ice core data, on the other hand, are poorly described, especially in light of the fact that the authors

C1902

conducted their own calculations on the data. As shown by the comment of [van Ommen] these calculations are not necessarily correct. The present-day RACMO output (temperature and precipitation) is poorly referenced - the Lenaerts et al. (2010) paper to our knowledge has not been published yet. As one can therefore not verify this reference it cannot be used in the paper. In addition, the authors have yet to show why the output of RACMO is a more appropriate reference for the present-day Antarctic climate than conventional reanalysis products. This could be done by referring also to Reijmer et al. (2005), who specifically addresses this question. It would be nice to shortly address the issue that RACMO output is interpolated back onto the much coarser resolution of the GCMs (i.e., what is gained then from the high resolution of RACMO?).

7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

The authors give a short statement about other PMIP2 model inter-comparisons. It would be helpful to have an introduction including references discussing the issues when using temperature and precipitation output of climate models as input for an ice-sheet model. There have certainly been some attempts to use climate model data to force an ice-sheet model. These publications most likely also include a verification of the input data, providing an opportunity for the authors here to refer to (e.g., Murphy et al., 2002, JGR).

8. Does the title clearly reflect the contents of the paper?

Yes.

9. Does the abstract provide a concise and complete summary?

Yes. However, considering our critique of the paper's concept, the abstract has to be rewritten with a stronger focus on how the best models were chosen.

10. Is the overall presentation well structured and clear?

C1903

The coarse structure with the subsections focusing on present-day, MH, and LGM is reasonable.

11. Is the language fluent and precise?

There are several imprecise statements and some which seem to be out of context, which makes the manuscript hard to read. A few examples: Third sentence of section 4: "This might be a timing problem, as the Antarctic climate optimum ended just before 6 ka (Ciais et al., 1992)". The relation to the rest of the paragraph is not apparent. Page 3592, line 23: "measured with 100% certainty" is a somewhat imprecise formulation as it is quite common for measurements to have some error and therefore not be absolutely certain. It would be better to give uncertainty ranges in the corresponding Tables 2-5. The same goes for timing issues with ice cores (same sentence). Page 3591, line 6ff: the "speckled results" of CNRM are supposed to be only a precipitation problem as temperature does not show such a pattern. Unfortunately, the reader might think that the authors suggest temperature to be alright in CNRM. However, one would never expect "speckled results" for temperature as this is generally a much smoother field anyway (both in reality and models). It is therefore not surprising that temperature and precipitation anomalies are not alike patternwise. It would be more interesting to learn about the actual reasons for the "speckled results" in precipitation. Could it be the changed land mask from LGM to present? If this were the case why would it occur in the present-day vs. RACMO comparison as well?

12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Yes. Minor detail: GCM originates from General Circulation Model rather than Global Climate Model.

13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

C1904

The choice of the figures should be better justified - in particular Figures 3-7 seem to be chosen randomly and do not really support the study's aim to find the best performing model. For example HadCM3, presumably one of the best models, is never shown and basically not discussed at all. In contrast, the bad models get described in some detail. Instead, one could rule out those models and focus on the good ones. Therefore, the important conclusion of which models perform best needs to be extended substantially. When a good model is found, the authors may go back and describe this model in more details so that the reader learns why and to what extent this is a good model. The short statement on Line 12+13 on page 3592 ("Considering all results for the present-day, the four models that perform best are ECHAM53, HadCM3, MIROC 3.2.2 and UBRIS") is not comprehensible for the reader. For the comparison with the ice cores it would be interesting to have some measure of uncertainty for both the model and the ice core temperature/precipitation. Also, one should comment on the fact that only two of fourteen models get the sign of change of EDC and Fuji MH-present day temperature difference correct (Tab. 2). Page 3592, line 14ff: it would be nice to have an extended discussion on what the implications of this conclusion are. To what extent is a model "useless" when it has a certain bias in present-day climate? Tab. 1: If for ECHAM53 only three years of output are available, it probably should not be included in the analysis as it does not represent a climatology. Figures featuring maps have a poor resolution. We recommend using eps or any other vector-based format to get rid of the resolution problem.

14. Are the number and quality of references appropriate?

The Pollard and DeConto (2009) reference is not really accurate (they are looking at glacial states over the last 5 mio. years but not at LGM specifically). The Lenaerts et al. (2010) paper has never or not yet been published and cannot be used as reference for the validation of the RACMO model. The conclusions section needs to be better supported by references or some conclusions drawn need to be shown (in more detail) in the results section. For example, page 3591, line 26ff: it is written, that ice shelves

C1905

are not simulated but are made up by sea ice. To our knowledge, most models have a land mask covering the ice shelf area, which is simply defined as a landmass covered by snow or ice. So the attribution of the temperature bias over the shelves to the lack of ice shelves in models is not convincing or needs to be shown. Page 3592, line 4ff: the model's energy balance is given as a possible explanation for the meridionally inverse temperature bias without providing support for this theory (either through a reference or within the manuscript). Page 3590, line 22: "Including other estimates, [...]". Which ones? Page 3591, line 1: the missing change of cyclonic systems is attributed to model resolution. Is there evidence from the study or the literature that support this claim?

15. Is the amount and quality of supplementary material appropriate? -

Interactive comment on Clim. Past Discuss., 7, 3583, 2011.