

# Review - Evaluation of regional climate features over Antarctica in the PMIP past1000 experiment and implications for 21st-century sea level rise

December 17, 2024

## 1 Overall quality

The topic discussed in the paper is exciting, innovative and relatively unknown, making a firm need for research like this to be published in CP. With the changes that I discuss below, I hope this work is also published in CP.

I applaud the authors for their choice of topic, the structure and writing of the paper, and the scrupulous analysis carried out. The majority of the introduction, data, method, discussion and the conclusion sections are clear, well written and easy to follow. The sheer amount of results presented is impressive. However, the fashion they are currently presented in deems to it hard to confidently understand the results section. Perhaps a more pedagogical way of conveying the results can be found, this is discussed below. The article is abundant with analysis, discussion and evaluation, meaning it is a very interesting read. This also means, however, it is lengthy. I think the article could benefit for being shortened slightly, with the most important conclusions honed in on. I feel the authors have tried to cover all discussion points - which yes is important with a model evaluation paper, but perhaps a middle ground can be reached, where the authors are slightly more ruthless, producing a more concise work which will still deliver this important subject matter.

## 2 Scientific questions

I appreciate that the authors are very aware of the limitations of working for example, with proxy data, or coarse resolution, and introduce this from the start, in the data and method sections. This, to a degree, reinforces the scientific quality of the study, that no conclusions are being based on assumptions, but rather critical analysis. However, perhaps by section 5, I began to wonder if we can pull any robust conclusions from the models/proxies? It seems, from the perspective of the authors, the majority of the data cannot be trusted. Models struggle to simulate realistic patterns of decadal climate variability, they overestimate current Antarctic precipitation buy sometimes 100%, simulate the wrong sign of snow accumulation, fail to reproduce warming in the DML and cannot be properly evaluated on their ability to capture SST trends. So considering all these limitations, why should this data be used? It is stated in line 53 "The LM is also a useful time period in which to evaluate model skill, as there is a

relative abundance of proxy data available" , so does abundance mean it can be used for the evaluation purposes the authors are aiming for?

I understand the vitality of highlighting the flaws in the proxies and models and that this is an evaluation of model performance, but I feel the authors need to include more references and more detail as to why they had the confidence to use these datasets in the first place to perform this highly important study. If the overarching conclusion of the study is that we can't yet use models to investigate paleo behaviour in this region, such as is hinted in line 332 "capturing regional trends of such small magnitudes that we observe in LM is still beyond current models' ability" then this needs to be reinforced and discussed further. This is a very important finding in terms of encouraging research development for the Southern hemisphere modeling/proxy community. As a member of the community I would like to know what is limiting these models, or what it is that we need to focus on to improve research in this area.

## **2.1 Abstract**

From the abstract, I gather the conclusion of the study is PMIP past1000 models reasonably capture SATs but not reconstructed regional snow accumulation nor Niño 3.4 index, and have some skill with SSTs but a slight cold bias. But I feel the real conclusion should be what does this mean for investigating Antarctic climate. What do these findings mean in respect to your study? What does this mean in regards to the opening line of the abstract - "Surface mass balance (SMB) of the Antarctic Ice Sheet (AIS) is an important contributor to global sea level change"? Relating the findings to this subject is what makes your study innovative and exciting.

## **2.2 Section 1**

From the introduction, it is not entirely clear that this is a model/data evaluation paper. The importance of the subject and the subject area is discussed well, the need for understanding paleo conditions to navigate the future unknowns is clear, but perhaps, from the tone of the introduction, I was expecting more an analysis of Antarctic paleo climate as opposed to model performance. I wish the core purpose - the evaluation - had been stated in the abstract and early on in the introduction, instead of waiting until the end of the introduction. In line 405, the authors state "The goal of this study is to provide a fair evaluation of the strengths and weaknesses of GCMs in simulating LM regional climate changes in Antarctica." This is a great and clear sentence, that in my opinion should have been used much earlier on.

## **2.3 ENSO (Section 3 and 4.4)**

I don't quite understand the purpose of exploring the ENSO Index. Where in the text does it explain how exactly it affects Antarctic climate or SMB? What is the purpose of including this analysis in regards to an Antarctic paleo climate study? The only line explaining this I found was 401, connecting ENSO with sea ice.

## **2.4 Section 4.1, 4.2, 4.3**

The results section starts with listing accumulation values but these lines (161-185) are hard to follow. Simply noting the values with no benchmark, does not help the reader understand how rates change. Considering Figure 2 shows the mean, trends and temporal variability values

of reconstructed and modeled time series for each ice core, plus table S1 shows details of the ice core records, is writing the various accumulation values necessary? Can this section be restructured, so the reader can easily follow which cores show which rates, why this is and how they compare? I understand that Figure 2 aims to do this, and, whilst it displays a lot of information is perhaps slightly overloaded and therefore adds to the complexity of the section. The use of notation to show which region the data sets respond to is not very pedagogic. If we are discussing regions, it is important to have some spatial context to this, or at least clearer marker areas on the figure. Figure 3 is on the other hand very easy to understand. Could these figures be interchanged so Figure 3 is introduced first? This perhaps would help the reader understand first what is a standard accumulation rate for that area, and then a break down of how the records vary within this. In the same manner, perhaps Figure 5 and Figure 4 should be switched around (and again 6 and 7), again so the reader has more clarity understand how temperatures vary per region and then per model. The values are again listed in pros form and it adds a level of confusion to the section. I want to highlight that Figure 10, the heatmap is a very clear way to deliver a summary of the vast quantities of information presented.

It benefits the study that the clear headlines from the accumulation rates, SAT and SST sections are delivered first, before delving into the point by point values. But perhaps this could be made even clearer. This would help guide the reader to the findings of the study more efficiently.

## 2.5 Section 5.1

Lines 314-320 talk about the interesting topic of how the models recreate the trends, but I wish for more discussion of literature on this section. The line 317 "they reproduce these regional trends for the wrong reason" teases a very interesting point, but then no literature backs it up. It feels almost clickbait like. What is the wrong reason? Or does it not exist and hence why there is no literature to fulfill the argument?

## 3 Technical

- Line 23 "Projections of 21st-century SMB span a large range and contain deep uncertainties (Li et al., 2023)". What does deep uncertainty mean?
- The authors tend to rely on non scientific, quite ambiguous phrases. For example; "not too different from the present, " is used often (line 51/77). Also " and by and large" (line 366). These phrases take away from the clarity of the method and results. How is it not too different? Is it similar enough to be used as an analogue? To what extent does by and large mean? I would suggest checking the paper and replacing these colloquial phrases to more scientific and suitable language.
- In methods, Figure 1 - would be very helpful with a legend to show which colors represent which regions, rather than just noted in the caption.
- I understand the authors do not want to have too many arbitrary tables, but it would be helpful to have a least of abbreviations of the Antarctic regions that the reader can easily refer back to. Currently, I had to keep going back to check lines 109-111, which felt cumbersome.