General comments

This manuscript by Du et al. has been much improved since first submitted, thanks to the diligent commitment of the authors to address the reviewers' comments. The paper reads easily and the parts which were rewritten have clarified the overall structure of the manuscript. I think that this modelling intercomparison study at the LGM is well suited for publication in Climate of the Past, providing an interesting focus on the Southern Hemisphere hydroclimate thanks to the case study on Australia.

I do have a few additionnal comments to hopefully guide further improvement. Some I have still classified as major.

Major comments

1. Knowledge gap: (a) the knowledge gap outlined in the abstract (L7 : « The climate changes… remain uncertain. ») is extremely laconic. I would argue that the abstract's objective is also to convince the reader to continue reading, and thus to explicitly present him or her with a problem worth solving. So I would recommend elaborating a bit on the knowledge gap in the abstract as well. [If the authors are limited by abstract length requirements, I think that the relationship between CMIP5 and 6 models and PMIP phase 3 and 4 models is comparatively much less important (and could be explained in Section 1.2 only).]
(b) The introduction sentence of the knowledge gap on L170-171" The LGM is commonly recognised as a time of global cooling and lower sea levels, best estimates placing this at ca. 21 ka. However, ..." is extremely confusing until the temporal discrepancy is pointed out later in L.174. The knowledge gap is also underdeveloped, to my opinion. Could the authors clarify and elaborate on the knowledge gap, possibly describing the different temporality of the SH regarding the start of the deglaciation, the bipolar seesaw mechanism, etc...?

2. Ending note (L20-22 and L1284-1289): I am not a fan of the 'further analysis is required' statements as it doesn't spell out clear directions of where the research should move forward to make progress on the still unresolved knowledge gaps of the paper. It is a bit of a shame to end the abstract and conclusion on an underwhelming note. Could you maybe provide clearer recommendations for modellers and for experimentalists, based on what we have learnt in this study? To better identify model biases – and better resolve mechanisms, what analysis are we lacking? Which sensitivity tests could be made? As for data, what do we need from data to better constrain models in the Australian regions?

3. L141-145 and Figure 1 / L191 : (a) Placement: The 'Australia case study' is brought to the reader's attention too soon, before it is even justified, and before starting describing the LGM overall climate again. In addition, Fig. 1 shows the different regions that are starting becoming relevant from Section 1.1 onwards. I would advise moving this figure to later in the text.
(b) Additional proxy information : while useful as it is, Fig. 1 could also be enriched with, e.g., the coring locations of all the proxies described in Section 1.1.
(c) Justifying the regionalisation : finally, I would like to point out that the connection between the separation into 3 regions and the atmospheric circulation mechanisms explaining the existence of this specific regionalisation is not expicitly made (neither in the legend of Fig 1, in L191, or around L175-177), until the much too late Section 2.3.
Hence, I would suggest reorganizing things (Fig. 1 / Section 1.1 / Section 2.3) so that the relevancy of this regionalisation becomes apparent to the reader in a logical manner.

4. The land-sea masks and their potential impacts. (a) It is a bit of the shame that Fig. 1 doesn't show the difference between PI and LGM land-sea masks with different contours.

(b) The authors have indicated that they will show the LGM land-sea masks for individual models in SI. I would suggest it may also be relevant to show those with additional contours in Fig. 3 and 9 notably, for I have been constantly wondering about the impact of different coastlines on the simulated variables and their potential disagreements. When the 'maritime' continent is becoming less maritime, leading to less evaporation, how does that affect the precipitation patterns over the whole region? My point is that different models, of different resolution, and using different ice-sheet reconstructions for the LGM, are not likely to have implemented the exact same coastlines e.g. around the Sahul shelf, leading to possible model disagreement in this region (i.e. stippling where the coastlines differ). I would like for the reader to have a chance to examine this potential effect (and if some shows up, for the authors to also discuss this).

Specific comments

- L12-13 « with a multimodel mean 2.9°C decrease in annual average surface air temperature over land at the LGM compared to the pre-industrial » is so packed with information that it is a bit difficult to read.

- L13-14 « while models show consistent patterns of regional cooling » confused me at first as it felt like a repetition of L11-12. Can’t both of these informations be presented all at once ?

- L16-18 « […] vary greatly between modes […] shows little change […] are also uncertain, with wide model disagreement » : I was confused by the ‘also’ L18 since a sentence describing surface moisture balance changes was placed between the two sentences pointing out the model disagreement.

- L296: While the newly added summary paragraph works well, I would add here a transition sentence to the next section to justify the use of models to complete the picture formed with the proxy data, something along the lines of : "In this context, climate models could thus provide precious insights into the mechanisms responsible for this observed climate."

- Table 1: While the use of the first 100 years of model outputs has been justified in the reviews and in the text, I would suggest avoiding the terms "run length" and "length of simulation" (e.g. in Table 1 or in Section 2.2 title) to refer to the length of the model outputs available on the ESGF. The authors could use something like "output length after spin-up" or some other equivalent so as not to confuse the reader. I would even argue that actually, the third column of Table 1 is irrelevant, for (1) the authors are using only the first 100 years anyway, and (2) only the spin-up duration can give the reader an idea of how well equilibrated the LGM simulations are. So the authors could consider replacing this column with the spin-up duration numbers, or removing it altogether.

- Fig. 2: I admit to finding the he first occurence of the 70% stippling peculiar, for their is no stippling appearing in Fig. 2. The authors could either choose a higher standard (e.g. good agreement on the amplitude of the change?), or simply warn the reader with e.g. "A stippling indicating areas where less than 70% of ensemble members agree on the sign of the anomaly has been chosen, consistently with the following figures. As a result of the high agreement between models in terms of the sign of the temperature anomaly, no stippling is shown here."

- Fig. 5: Please consider further commenting on Fig. 5 in the text. What do we see in terms of model disagreement? Do we see an increased seasonality at the LGM wrt. PI? The authors could also consider quantifying the model spread or commenting on the obvious two outliers (one PMIP3 and one PMIP4) in the global mean temperature plot.

- L696: Please consider adding a transition sentence to connect the previous section with the one starting now.

- Fig. 6: I am wondering whether showing wind **changes** as vectors is the best data vizualisation choice. The reader may assume that arrows stand for the winds themselves.

- L793-795: Please also mention that some models do not show any significant change. Also, I would start with the change of westerly strength (L796) before mentioning the latitudinal shifts, as the latter is not the first metric that usually comes to mind.

- L829 "as noted in Section 3.2": It was not clearly spelled out in Section 3.2. I would suggest reformulating the transition, so that the writing can flow more in-between sections.

- L952: This mention to Table 3 may confuse the reader as to whether he should read Table 3 or Figure 12 first. I suggest it is not necessary here.

- L1063 "While a decomposition of thermodynamic and dynamic drying components of precipitation change is not included in this study, it is evident that the thermodynamic drying response is dominant": (a) Why not? It would be interesting to see such an analysis.
(b) On what basis can you say it is dominant (without doing the decomposition)? I may have missed the 'evident' fact which helps conclude this. Please justify, or nuance (for this is a strong statement, comparatively to e.g. L1059 "it is likely that...").

- L1230: (a) It feels like the winds subsection lacks a clear conclusion, along the lines of "This also reveals that much progress remains to be made with respect to...".
(b) Please also consider refering to this recent paper by Gray et al. (2023): https://doi.org/10.1029/2023PA004666, to elaborate on your discussion.

- L1279-1280 "again suggesting that caution is required" is redundant with a previous statement, so it does not bring anything to the table. I suggest removing it.

Technical comments

- L139-140: Why not use 6.1 ± 0.4 °C, to provide uniformity with the previous figures?

- L305 "a slight increase" of what?

- In L315 and a couple of other times, the poleward/equatorward shifts are refered to as southward /northward. I think it is best to stick to poleward/equatorward even if the study is focusing only on the SH.

- L316 equatorward shifts (reversed word order)

- L318 "more recent" repetition

- L358 "the" -> their

- L1142 "has evaluated" -> evaluates

- L1158 "the models may not have resolved" -> the models do not resolve