Reviewer#2:

General comments

In this study by Du et al., the authors propose a modelling intercomparison study of the simulated surface temperature, precipitation, moisture balance and winds at the LGM, focusing on the Australian region.

Such a study would fit well within the scope of Climate of the Past. The case study of Australia is interesting for several reasons such as the impact of sea level change on coastlines, the specific location (SH, in proximity to the Maritime Continent and Southern Ocean), and the overall disagreement of models and proxies especially concerning the latitudinal shift and variation of strength of the westerly winds, which seem both poorly represented in PMIP models (Chavaillaz et al., 2013) and poorly constrained by conflicting paleodata records (Kohfeld et al., 2013). The manuscript is also clear and an easy read.

However, as such (and as is often the case with intercomparison studies), the paper reads as very descriptive and superficial, so we are struggling to learn something new. Knowledge gaps, uncertainties in both model and data and processes/mechanisms are either barely mentioned or simply not highlighted enough so when we reach the end, the impression is underwhelming.

I am providing more concrete illustrations as to what could be improved and how with points 1-3. I also have suggestions related to methodology in points 4-8. So I would like to recommend a number of improvements before publication, in the hope of helping give this study more weight.

1. Please elaborate on the reasons why a case study of the past climatic changes over Australia is interesting. It would be great to mention climate processes that are key in this region. An example since dust transport is mentioned (L93): does this aridity have the potential to significantly enhance iron fertilisation in the ocean?

[Response]: Thank you so much for your comments and suggestions. We will add more discussion of why case study is interesting with reference to key climate processes in the region. While the issue of dust transport and iron fertilisation is beyond the scope of the present study, we will include discussion of the major changes in Southern Hemisphere climate and circulation which may have influenced Australia during the LGM, e.g., changes to the SST gradients in the tropical Pacific and Indian Ocean, and related changes in the Walker circulation and Hadley Cell, changes to the position or intensity of mid-latitude westerlies and changes in the strength of the Australian-Maritime Continent monsoon.

2. Consider adding one or two last sentences to the abstract and a paragraph in the discussion/conclusion to give the reader a broader perspective and hindsight on what we have learned and how significant are these new findings. A few questions to help brainstorm: So what? In the basis of the existing literature and these new findings, have we achieved a better understanding of the processes which influenced the past Australian regional climate? If not, what are we missing? Do we understand the model response to LGM forcings? What does it entail?

[Response]: Thank you for your suggestions. We will expand the Abstract and Discussion to include more discussion of the significance of the findings.

3. Please underline the knowledge gaps in the introduction. As such, the introduction is very descriptive (not impactful), with a structure (global changes / changes in different regions of Australia) that doesn't help guide the reader very logically towards understanding the knowledge gaps, their importance, the scientific question and the methodology used in this study. If the authors would like to keep this regional description structure, then it would be welcome to also point out the contrasts between these different regions, and also with the global climate, with a few short sentences to conclude this subsection.

[Response]: Thanks. The structure of the Introduction follows many previous proxy-based studies examining regional changes in Australian climate during this period (e.g. Reeves et al., 2013b; Fitzsimmons et al., 2013; Petherick et al., 2013; etc.). As Australia spans several distinct climate zones, we thought it would be logical to structure the Introduction around these distinct climate zones. We will add several sentences comparing these with each other and with global climate.

Still, I feel like the sometimes conflicting proxy records and the 'uncertainty about the drivers of the LGM climate changes' (L74-75), the 'ambiguous results' of models (L117) should be arguments brought to the reader's attention in a more convincing order to justify the need of this particular intercomparison study and its methods. For exemple, it is not clear what is the advantage of using an intercomparison method, nor for which reasons modellers simulate the LGM period (L102-107). It is also not clear why the authors chose to examine these three specific climate variables. I believe there are ways to reinforce the visibility of the scientific reasoning behind this approach.

[Response]: Thank you for the comments. We note that this approach is widely used in numerous papers using PMIP LGM simulations. We will more clearly explain the motivation for simulating the LGM and for using multiple climate models in Section 2. We will also explain why the particular climate variables were chosen.

4. Only PMIP4 outputs available on the ESGF were included. This is a bit of a shame because it limits comparison with the Kageyama et al. (2021) results, and the model ensemble size (and therefore the robustness of the results). CESM2 is also excluded for very good reasons (L143-144), but I believe the authors have found the source of the exagerated cooling in a cloud microphysics parameterization and run a corrected simulation. This is of course up to the authors, but they could consider contacting both the Kageyama et al. (2021) and Zhu et al. (2021) authors to request the model outputs. This would also enable the authors to compare individual model versions (CMIP5 vs CMIP6) and discuss potential improvements between the two generations. For now, only MPI-ESM is in the two ensembles.

[Response]: Thanks for your comments. We decided to make use of only those models which are publicly available via the ESGF to ensure our results could be easily reproduced. We follow the standard approach in CMIP climate model studies which typically only use publicly archived simulations. Many other PMIP-based studies have also used this approach, e.g. a recent paper on LGM ITCZ changes from PMIP3/4 models used a similar set of models (Wang et al., 2023).

Regarding CESM2 models, we had some trouble linking the available CESM2 model simulations on ESGF with documentation and relevant publications. We will now include the CESM2-WACCM model as we now understand this model does not have an unrealistic climate sensitivity.

The arge-scale comparison between CMIP5 and CMIP6 models was provided in Kageyama et al. (2021) so we do not need to repeat this analysis. We do not assume that the two generations of a particular climate model (e.g. MPI-ESM) will produce a similar simulation of LGM climate, and we do not wish to focus on comparing generations of specific models – this is typically done by the relevant modelling groups.

The focus of the present study is to summarise the simulation of LGM climate of Australia from available PMIP3 and PMIP4 models. This information is of interest to the Australian palaeoclimate community and will provide a basis for comparison with numerous proxy-based reconstructions of standard climate variables such as surface temperature and precipitation, as outlined in the Introduction.

Reference:

Wang, T., Wang, N., & Jiang, D. (2023). Last glacial maximum ITCZ changes from PMIP3/4 simulations. Journal of Geophysical Research: Atmospheres, 128, e2022JD038103. https://doi. org/10.1029/2022JD038103

5. The authors mentions that all model outputs are regridded (L170), and it seems that all the following analysis use these regridded outputs. Of course, this is necessary to compute and plot the multimodel mean, but I am wondering whether it would be worth extending the use of the model native grids when plotting the individual maps and wind profiles. Is the latitudinal shift of westerly winds affected by the model resolution? Are there differences in the land-sea mask of each model which could impact the simulated temperature and precipitation patterns? Please consider plotting the PI or LGM land-sea mask (e.g. as a grey or dashed contour) on maps.

[Response]: Thanks for the suggestion. We have recalculated boundary lines between SH easterlies and westerlies on the native model grids (Table 4), however, we decided to remove it from the paper. This is because the results differ quite a lot on the different grids, and we think they are not robust enough to be included. The whole Section 3.2 will be then rewritten, including adding some new figures focused on the large-scale shift in winds.

Regarding the influences from the land-sea masks in different models, we will plot the LGM land masks for individual models in the Supplement document to show the potential impacts from the different land fractions configured in each model.

6. On maps, the authors should also represent the multimodel agreement significance as hashes (when >90%) and the proxy data as scatter points, whenever quantitative reconstructions can be found. I consider important that the reader is able to compare visually the performance of the models with the available proxy data, especially as the authors conclude that there is a general good agreement between model and data.

[Response]: Thanks. We will add stippling for model agreement. As noted for reviewer 1, we have provided already that the sign of the change from proxy records is uncertain, especially when taking into account the CO_2 effect on vegetation records, and therefore we refer to the literature but don't include any proxy records in our plots.

7. The authors mention using the first 100 years of simulation for the analysis. Why not the last 100 years? Please check that the simulations are in equilibrium, e.g. by computing the drifts.

[Response]: Thanks. Same as for Reviewer 1, most climate models used in this study only have 100-year length of simulation based on the number of years of data available on the ESGF (Earth System Grid Federation). According to Kageyama et al. (2017), the models have been spun up until equilibrium following the PMIP protocols (refer to Kageyama et al. (2017) for details of the spin-up). At least 100-year data from the equilibrium part of the simulation is required to store on the ESGF (Kageyama et al., 2017). Therefore, the data stored on the ESGF has already been in equilibrium and it does not matter anymore whether it is the first or the last 100 years. There will be no significant differences.

8. The paper would deserve more quantifications. An example is in L273: 'Some models show weakening and other model show strengthening but there are other instances (e.g. L372). It would be great to provide precise figures, e.g. phrasings like '5 models out of 12 show a weakening of at least 20%...'

[Response]: Thank you so much. We will change the wording to quantify model agreement and specify the magnitude of changes as suggested.

Specific comments

L2, L12 and L13: 'changes at the LGM', 'to cool by 2.6 at the LGM' and ,'decreased'. Unlike in the rest of the paper, it is sometimes unclear in the abstract that we are mentioning changes relative to the PI. This has to be indicated in some way (e.g. LGM-PI anomaly, with respect to PI…) or else the verbs are inconsistent with the time direction.

[Response]: Thanks. We will change the wording here to avoid confusion.

L13-14: Why are the changes in temperature and precipitation indicated over two different defined regions?

[Response]: Thanks for asking. The calculation for MMM precipitation changes was over a smaller domain aiming to exclude the land areas between 0 to 10°S to allow comparison with previous studies focused on northern Australian rainfall.

L18-19: I find the sentence explaining the potential reasons for model-data disagreement to be rather vague and could be reformulated.

[Response]: We will rewrite it to improve clarity.

L23: The time window proposed for the LGM is unusual and only justified later in the text.

[Response]: Thanks. We will change it to "22-18 thousand years ago (ka)".

L31: Why such a gap between the Annan et al. (2022) and the Tierney et al. (2020a) estimates?

[Response]: This disagreement is due to the "choice of prior", i.e. the particular climate model used in the Tierney et al. (2020) study, as discussed by Annan et al. (2022). We will briefly note this.

Fig 1 and 2: These figures do not bring a lot of information to the table. The authors could consider combining them, combining Fig 1 with e.g. Fig 3 (the land-sea mask could be indicated on another map with a grey contour), or enriching them with more information (e.g. the Sunda and Sahul shelves mentioned in L34 could be annoted on the map to show the reader their location). As for Fig 2, consider using a different style (than the red contour, e.g. hashes) for the southwest box.

[Response]: Thanks for the suggestion. We decided to combine Fig 1 and Fig 2 with the LGM land mask for CCSM4 model as background shading. The southwest box will be plotted using a different style.

L85: The authors could elaborate on the reasons why they used the word ‘possibly‘.

[Response]: According to Kohfeld et al. (2013), "If a single cause related to the southern westerlies is sought for all the evidence presented, then an equatorward displacement or strengthening of the winds would be consistent with the largest proportion of the observations." This was summarised with the word "possibly", but we will rewrite this as "is consistent with" to better summarise the findings of the Kohfeld et al. (2013) study.

L110: ‘not fundamentally different‘ could be true for the variables examined in the study and not others. Please check if this is the case for all variables (including ocean circulation).

[Response]: The global study of Kageyama et al. (2021) has already addressed the comparison between PMIP3/CMIP5 and PMIP4/CMIP6 models. While individual models will differ for any variable considered, the sample spread of PMIP3 and PMIP4 models overlaps for all the variables of interest in this study, therefore we think that it is reasonable to treat them as a single ensemble. This same approach is used in numerous studies, e.g. Brown et al. (2020); Wang et al. (2023). The alternative option would be to examine PMIP3 and PMIP4 LGM simulations separately, which we do not think would provide greater insight and which would result in two very small ensembles.

Table 2 does not feel very necessary. The authors could consider moving it to SI to save some space.

[Response]: Table 2 summarises information from Kageyama et al. (2017) for PMIP4 models and from the PMIP3 website for PMIP3 models and therefore we feel that it is worthwhile to include.

L168: Why use both ts and tas?

[Response]: As suggested by reviewer 1, we will change all of our ts analysis in Section 3.1 to surface air temperature (tas) to allow a clearer comparison between our study and previous work.

Fig 4 / Table 3: What is the GMST simulated by these models? Are the models which are cold on a global scale also the ones simulating cold temperatures over Australia?

[Response]: This is a good question. We will add the GMST values to new Figure 3 (reproduced from Table 3) to allow comparison between global temperatures and Australian temperatures and comment on this relationship in the text.

Table 3: Tables are not great to visualize data (also true for Table 4 and 5). The authors could consider a different type of plot to show the reader the large intermodel differences in the seasonal amplitude (not commented in the text?) in a single glance.

[Response]: Thanks for the suggestion. We will change Table 3 to a scatter plot (new Figure 3) which is similar to Figure 1 (b) from Kageyama et al. (2021) to compare the seasonal and global changes in individual models. Table 4 has been removed. We will keep 5 because that it is easier to compare each column in the table to identify differences between precipitation changes and P-E changes and to see values averaged over land only.

L250/L269: Please consider using transitions between subsections (here as in other instances). It is the opportunity to remind the reader of your scientific reasoning (e.g. how these variables are linked).

[Response]: Thanks. We will add transitions as suggested.

L272: I am wondering whether Fig S1 which shows very large model disagreements should not be part of the main text. Also, please explain why you chose to plot the JJA season specifically.

[Response]: The JJA season is when the westerlies bring rainfall to southern Australia. We therefore wanted to see whether there was a shift in the westerlies in this season in order to help understand changes to Australian climate.

We will add Fig S1 to the main text in Section 3.2, and the whole section will be rewritten.

L304-305: 'In JJA the SH westerlies shift equatorward'. Would it be worth it to also investigate the seasonal shift of the westerlies in Sect. 3.2?

[Response]: We explore the shift in the position and intensity of the westerlies in Section 3.2 with a focus on the SH winter (JJA) season. We will add a new figure which shows the MMM change in 850 hPa winds and convergence for each season.

L397-398: Does this correspondance hold for all models?

[Response]: There is higher model agreement over northern Australian precipitation change for MAM than for other seasons – this will be indicated with stippling in the relevant figures. The drying occurs over the region of cooler land, particularly the exposed Sahul shelf in a large majority of models. It is not possible to include seasonal plots for all models for all variables or the paper and Supplement would be unreasonably long.

L412-413: Does this relationship hold if you use the change in strength of westerly winds over the same region?

[Response]: Section 3.3.2.1 (winds) will be rewritten.

L463-465: I will make a subjective comment here. While this is a valid reason to criticize the proxy records (well-explained in introduction), I feel like modellers should maybe not be too critical of proxy uncertainties when such large intermodel differences are observed. The primary reason why we are observing this model-data disagreement might be that, well, models are wrong. I will also point out here that the discussion and especially the conclusion seem lenient with models. I would expect the large intermodel difference observed to reflect a poorly-represented process.

[Response]: Thanks for this important comment. As the reviewer points out, given the model disagreement, it is indeed reasonable to argue that at least some models are wrong. We are interested in finding areas of model-proxy agreement, and areas where models show robust inter-model agreement. However, we will also expand the discussion of possible model biases and uncertainties.

L467-470: Could you discuss the potential reasons why you can find a displacement of the boundary line but no consistent latitudinal shift in westerly winds? Do you have any idea?

[Response]: Section 3.2 and Section 3.3.2.1 (winds) will be rewritten. As noted earlier, the table of boundary lines will be removed, instead, a new figure showing the MMM LGM

seasonal 850 hPa wind change and Fig S1 showing individual model 850 hPa wind change in JJA will be added into Section 3.2.

Fig. 11: Consider different marker styles or colors for individual models or model generation (CMIP5/6).

[Response]: Section 3.3.2.1 (winds) will be rewritten and Fig 11 will be removed.

Fig. 6: It would make sense to put MPI-ESM-P (CMIP5) and MPI-ESM-LR (CMIP6) in the same column on Fig 6 so that it is easier to compare the two generations visually.

[Response]: As discussed earlier, we do not assume that the two generations of a particular climate model will produce a similar simulation of LGM climate, and we do not wish to focus on comparing generations of specific models. Therefore, it is not necessary to put them in the same column on Figure 6.

Technical corrections

L16-17: 'many regions' is unclear

[Response]: This sentence will be rewritten to provide more detail of vegetation changes based on proxies and biome models, with citation of Prentice et al. (2011) as the appropriate source: "There was a large reduction in area covered by boreal and temperate forests in northern mid- to high latitudes, expanded lowland tundra in Eurasia, expansion of savanna and grasslands at the margin of Amazon tropical forests and replacement of some areas of tropical forest in Africa, China and Southeast Asia with savanna, woodland and grassland (Prentice et al. 2011)."

Reference:

Prentice, I. C., Harrison, S. P., & Bartlein, P. J. (2011). Global vegetation and terrestrial carbon cycle changes after the last ice age. New Phytologist, 189(4), 988-998. https://doi.org/10.1111/j.1469-8137.2010.03620.x.

L26: 'glaciers' instead of 'glaciation'

[Response]: Thanks. This will be corrected.

L43: 'reflect' or 'are associated with'?

[Response]: We will change the wording as suggested.

L44: Consider replacing 'related to combinations' with something like 'caused by a combination'

[Response]: We will change the wording as suggested.

L45: 190 ppm in Table 2

[Response]: Thanks for pointing it out. The 180 ppm in L45 are summarised from proxy records, whereas Table 2 shows the IPCC configuration values for the LGM experiment.

L110: 'drier conditions' would work but not 'drier changes'

[Response]: This will be modified.

L121: 'differences' or 'gaps' would fit better than 'variations'

[Response]: Thanks, this will be modified.

L124: 'more recent' instead of 'newer'

[Response]: We will change the wording as suggested.

L144-145: ‘Furthermore, the PMIP4 protocol highlighted…‘

[Response]: (Line 155) We will change the wording as suggested.

L170: 10 m

[Response]: (Line 172) We will change the wording as suggested.

L328: What is the Top-End region?

[Response]: We will change this to “central northern Australia” as the region “Top End” is only known in Australia.