We thank the reviewers for providing precise and valuable feedback on our manuscript. The recommendations of the reviewer have been carefully incorporated into the revised manuscript and are described in the following. The reviewer's comments are indicated in black text, and our answers follow in blue text. The revised sentences of the manuscript are indicated in green text. We will be happy to submit a revised manuscript that reflects these changes.

Response to reviewer #1:

This manuscript analyses transient simulations of the last deglaciation Paleoclimate Modelling Intercomparison Project PMIP4 from six different climate models. The paper focuses on the evolution of Southern Ocean sea surface temperature (SST) and Antarctic air temperatures in response to increasing atmospheric greenhouse gas concentration (GHG) and millennial scale Atlantic Meridional Overturning Circulation variations. The set of simulations are of great interest to assess model dependent representation of a fundamental climate transition during a period of dramatic global warming.

The six analyzed experiments don't follow a strict unified experiment protocol and in particular differ in their freshwater forcing scenario. In consequence AMOC chronologies which are known to be highly model dependent are additionally reflecting differences in the forcing, which hampers a simple intercomparison. The authors deconstruct simulated climate change into AMOC and CO2 driven changes using a multilinear regression model and a thermal bipolar see-saw model. This approach might allow to compare the predominantly CO2 driven response independent of differences in freshwater forcing and AMOC response. The applied approach is original and interesting but I found the discussion to be too superficial in some places.

Technically, the paper is mostly well structured and the provided figures are of good quality, but I recommend to condense especially the descriptive parts to improve the readability and to also possibly sharpen some statements where the authors quite generally state the need for better system understanding and more modeling. In summary I recommend major revisions before publication.

Thank you for carefully reading and giving us fruitful comments. We have revised the article so that more specific statements are included. Our point-by-point reply follows.

General comments:

Abstract: The findings of the analyses need to be spelled out more explicitly. The sensitivity to freshwater forcing and AMOC is not really new or unexpected. Consider mentioning the sea-saw model and being specific about the results.

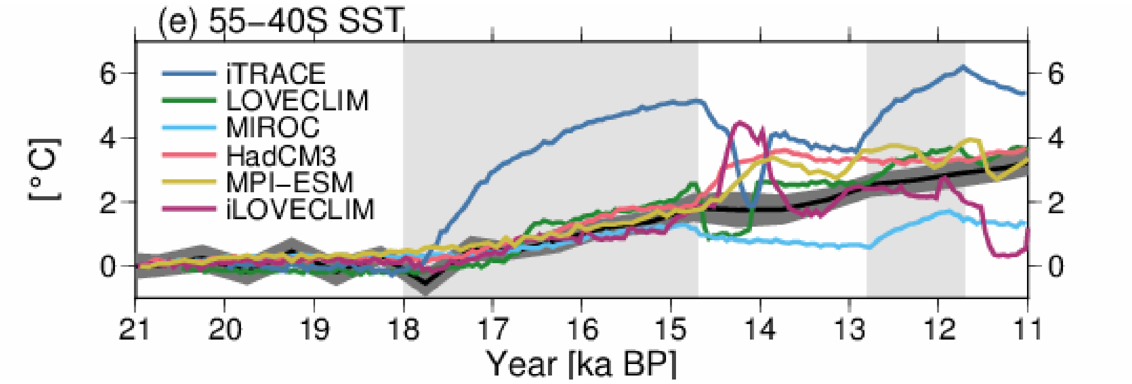
We have revised the abstract to clarify the use of the bipolar seesaw model. We also add a quantitative description of our results.

3-1 – 3-2-4: This is a bit tiresome to read and should be summarized in some way.

We have reduced the number of subsections by combining several subsections. By merging subsections [3-2][3-2-1][3-2-2] , we now begin with a description of SST and SAT changes during the onset of warming and the HS1 warming. The now merged subsections [3-2-3][3-2-4] were describe AMOC and bipolar climate changes. By this, we hope that the results are presented in a more structured way.

MLR analysis and see-saw model: it would be desirable to have a reference from reconstructed SSTs in Fig. 9 or to also apply these methods to SAT at the two core locations. It even might be interesting to use the see-saw model with the parameter combinations of table 5 but applying it to the same inputs for all models, e.g. the CO_2 and AMOC reconstructions from Figs 1b and 2bd.

We have decided to overlay the reconstructed SST data on Figures 3 and 9. We use the Southern Ocean SST stack dataset from Anderson et al. 2020:



Revised Figure 3 (only showing panels 3e): Color lines indicate simulated Southern Ocean SST. The black lines and grey shades indicates Southern Ocean SST stack data and its standard error from Anderson et al., (2020), respectively.

And comparison of SST with proxy data is added on section 4-2 (discussion of temperature change).

Second, we applied the see-saw model with inputs of CO2 and AMOC (Figure R1), The atmospheric CO2 variations are taken from Bereiter et al. (2015) as PMIP4 forcing, while the AMOC evolution is taken from the iTRACE experiment. This AMOC estimate provides a good model-data agreement over the last deglaciation (Liu et al. 2009; He et al. 2013; He et al., 2019). The results indicate that all models would have simulated ACR if the time-series of AMOC is the same as iTRACE.

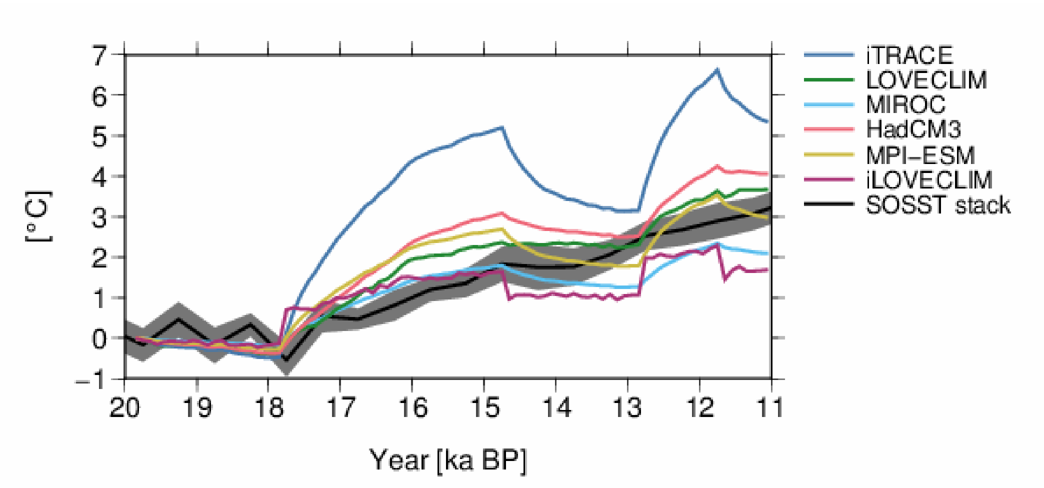


Figure R1: bipolar seesaw model results overlayed with SST stack (Anderson et al., 2020)

While northern hemisphere temperatures are also influenced by the retreat of ice sheets, it is a fundamental assumption in this study that Southern Ocean SST and Antarctic SAT are (primarily) driven by CO2 change and AMOC variations- this should be spelled out explicitly (maybe in the introduction), substantiated with references and discussed with respect to its limitations. Using CO2 and AMOC as input in the MLR analysis and see-saw model may provide a good fit but this does not exclude that for instance the retreating ice sheets also influence southern hemisphere temperatures. So I am reluctant to interpret the CO2 coefficients in table 4 and 5 as sensitivity to CO2.

We agree with your point: while Figure 8 indicates that CO2 is the primary forcing of the gradual warming, ice sheet changes and orbital forcing can also contribute to the warming. We have added sentences in the method section to clarify that CO2 and AMOC are the only inputs. We also refer to a climate modeling study evaluating the contributions of each forcing, which shows that ice sheets and orbital forcing have minor impact on Southern Ocean SST and Antarctic SAT during 19 to 15 ka (He et al. 2013).:

Methods: Because the CO2 forcing is shown to be the most important forcing contributing to Southern Ocean SST change (Figure 8), we use CO2 as a representation of a gradual forcing as the input of the MLR model. We note that other forcing, such as from ice sheets and orbital changes can contribute to the warming. On the other hand, sensitivity experiments evaluating the contribution of each forcing shows that they have a minor impact on Southern Ocean SST and Antarctic SAT changes during 19 to 15 ka (He et al. 2013).

We have also revised the end of section 3-5 (results of MLR and bipolar seesaw model) to clarify that the CO2 sensitivity of the MLR and the bipolar seesaw model may include other forcing:

We note that the values of the CO2 sensitivity from the MLR and bipolar seesaw model may include gradual forcing from other greenhouse gasses, the ice sheets, and orbital forcing. In addition, a sharp cooling associated with meltwater in the Southern hemisphere (~14.5 ka of iTRACE and LOVECLIM, ~11.5 ka of MPI-ESM and iLOVECLIM) is not considered in both MLR and bipolar models.

And we have also removed one sentence in the discussion because the sentence assumed that the CO_2 sensitivity is only response to the CO_2 forcing: "(deleted sentence) Estimates from the MLR and bipolar seesaw models indicate that both the increase in CO_2 during HS1 (~ 40 ppm) and the reduction in AMOC contributed to this warming."

Specific comments:

l. 29-31: maybe use "in phase/concurrent with" instead of "in response to", also this is
mixing observations with modeling results which makes the sentence ambiguous.

We have addressed this and revised the sentences to clarify modeling results: "All models simulate a warming during Heinrich stadial 1 (HS1, 18 to 15 ka) concurrent with the CO_2 increase at a greater rate than the early warming,…"

l.36: model's -> models'

l.99: to avoid misunderstanding: austral spring

We have changed phrases as indicated.

l.103ff: include the conclusion of Menviel et al. (2011)

We have revised the sentences: "Menviel et al. (2011) further assessed that the major part of the ACR was caused by an AMOC increase at the end of HS1, but found that a better model-data agreement could be obtained with increased Southern Ocean stratification driven by meltwater input from the Antarctic ice sheet."

l.128-139: distinguish findings from models and reconstructions

We have restructured the sentences to clarify the evidence from reconstructions s and modeling studies: Proxy records (Sigman et al., 2010, Skinner et al., 2010, Martinez-Garcia et al., 2011) and modelling studies (Bouttes et al., 2012, Menviel et al., 2016, Menviel et al., 2018, Gottschalk et al., 2019) indicate that physical and biogeochemical changes in the Southern Ocean may have significantly contributed to ocean carbon uptake during the last glacial period and the atmospheric CO2 increase during HS1. …

l. 159: please consider to include the Paleomist ice sheet reconstruction (Gowan et al. 2021) in Fig. 1 for ice mass change.

We have chosen to focus Figure 1 on the GLAC-1D and ICE-6G_C ice sheet reconstructions since our manuscript presents results from the PMIP4 last deglaciation simulations, for which the protocol recommends specifically the GLAC-1D and ICE-6G_C ice sheets as the most highly resolved and globally complete ice sheet reconstructions available at the time (Ivanovic et al., 2016).

We analyzed the PaleoMIST ice sheet reconstructions datasets (Gowan et al. 2021) in the same format as Figure 1 of the article (Fig. R2). We find that the time interval of the PaleoMIST dataset (2,500 years) might lead to gradual demise of the northern hemisphere ice sheets of the last deglaciation than GLAC-1D or ICE-6G_C. One major difference can be found in rapid sea-level rise corresponding to Meltwater Pulse 1a. We prefer not to include PaleoMIST in Figure 1 in order to keep the manuscript focused on the PMIP4 results. Nevertheless, we have added one sentence in the discussion section, because we think it would be valuable to assess the uncertainties from other ice sheet reconstructions, as different ice sheets can have different AMOC variabilities: (Discussion second paragraph) It would be valuable to assess the uncertainties from ice sheet reconstructions, as new reconstruction has been published (e.g., Gowan et al., 2021), as different LGM ice sheets can induce different AMOC variabilities (Prange et al., 2023).

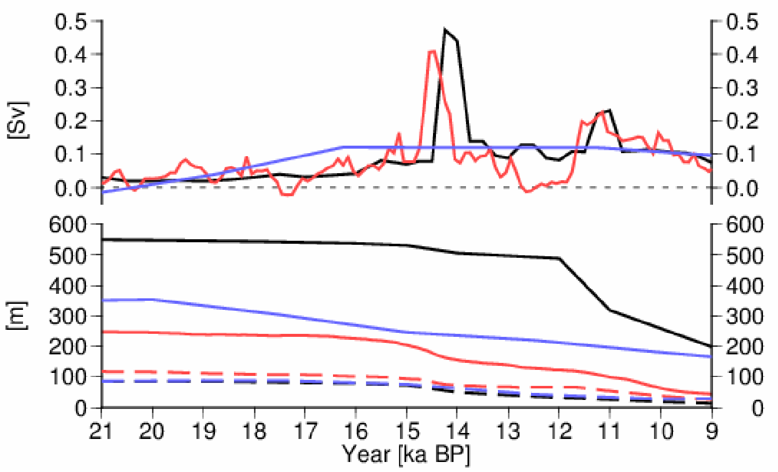


Figure R2: Freshwater based on ice sheet volume and surface elevation at EDC and WDC (same format as Figure 1 of the manuscript). Blue lines represent PaleoMIST (Gowan et al. 2021), while black and red lines represent Ice-6G_C and GLAC-1D, respectively. The bold and dashed lines represent surface elevation anomalies at WDC and EDC, respectively.

l. 191: what is respectively referring to? What is here the northern boundary of the arealy averaged Southern Ocean SST?

The statement refers to both zero-dimensional value of areal mean SST (55-40S), and 2-dimensional SST fields. We divided the sentences to clarify this. “The MLR analysis is applied to the 2-D fields of the Southern Ocean SST. The same analysis is applied to the SST averaged over 55-40S.”

l. 200: SST is here mean Southern Ocean SST?

Yes. We have clarified to mean Southern Ocean SST. “Southern Ocean SST“

2-2-2: What are arguments against using normalized AMOC?

Thanks a lot for your comment. We revised the sentences to clarify two assumptions in the bipolar seesaw models. The first assumption is that the bipolar seesaw models assume that the AMOC has only two modes, unlike the continuous values used in the MLR model. The second assumption is setting the threshold of AMOC dividing into weak and strong modes: "The term m(t) represents the modes of the AMOC from the climate model outputs. When using the simulated AMOC within the bipolar seesaw model, it is assumed that the AMOC has only two modes, unlike the continuous values in the model. And the AMOC state has to be divided into weak and strong modes; based on Figure 2, we assume that the AMOC is in a strong mode (m(t)=0) if the AMOC is greater than 14 Sv in all models."

l. 333-334: „which is able to explain about half of the total deglacial changes during HS1"- is this statement referring to a specific figure, to observed SST or SAT or simulated SST?

This refers to total SST changes. We revised the sentences to clarify: "All models have a negative coefficient of AMOC (–0.3 to –2.4 ℃), indicating a Southern Ocean SST increase associated with an AMOC weakening. The negative coefficient of AMOC in all models suggest that an AMOC shutdown during HS1 has the potential to explain about half of the HS1 SO changes."

l. 386: delete or replace "although"

We deleted the phrase of "although" and changed the sentence for clarification: "The amplitude of the early warming in these models is comparable to a previous modelling study (Timmermann et al., 2009), while the other models show a slight cooling (iTRACE and LOVECLIM) or little change (iLOVECLIM)."

l.396: specify: smaller is here in relation to the other models?

Smaller than other models. We revised the sentence to clarify: "The smaller sea ice extent at the LGM, relative to other models, may lead to a high sensitivity to increased insolation during austral spring to summer, causing significant early warming associated with additional sea ice retreat (Timmermann et al., 2009; Roche et al., 2011)."

l.402-404: in the models, the early warming seems to be not related to AMOC, as there is no weakening of the AMOC and also the see-saw model does not explain the early warming.

We have revised the sentences to indicate early warming in the models were likely in response to insolation change. And if additional meltwater is added, weakening in the AMOC would contribute to more warming: "Although the local insolation changes are the likely cause of an early warming simulated in some of the models, the addition of freshwater could contribute to the AMOC weakening. For example, the consideration of an additional freshwater flux from the Fennoscandian ice sheet in the freshwater forcing, as suggested by Touccane et al. (2010), would weaken the AMOC and lead to a more pronounced warming in the southern high latitudes."

l. 420: typo: reache

corrected.

l. 425-440: This is hard to follow and seems a bit unstructured, as it goes from warming in general to local SAT to SST to ECS and global SAT.

We have revised the paragraph to start with SAT and SST changes and their comparison to proxy data, as we have added SST proxy (Anderson et al. 2020) in Figure 3: "iTRACE simulates the largest warming during HS1 among six models, with an increase of Antarctic SAT of 6–8℃ and Southern Ocean SST of 4–5℃ While the Antarctic SAT matches ice-core data, Southern Ocean SST is larger than the SST stack. Five models besides the iTRACE simulates a Southern Ocean SST change which compare well with the SST stack data, but these five models underestimates Antarctic SAT. This indicates different magnitudes of warming between Southern Ocean SST and Antarctic SAT is weakly simulated in models. While iTRACE exhibits the largest global mean SAT changes at the LGM (7.3 ℃, compared to the six-model mean of 5.3 ℃)., the ECS of iTRACE (3.6 ℃) is not the highest among the six models (Table 1); instead, MIROC4m has the highest ECS despite weaker deglacial warming.

l. 435: "SAT anomalies" maybe better "local SAT change"

we have changed as suggested.

l. 435: the fact that CO2 coefficients differ for models with similar ECS values could also be a result of different climate response to the (roughly at the same time) retreating ice sheets, please discuss.

We agree that the climate response due to ice sheet can be included CO_2 coefficient, which could contribute to the difference between ECS and CO_2 coefficient. This was partly discussed in the submitted manuscript (third paragraph of section 4-2), but we revised the paragraph to emphasize the ice sheet effects superimposed on the CO2 coefficient: "The MLR and thermal bipolar seesaw models in this study have several assumptions. Firstly, as the gradual forcing is represented only by the CO2 concentrations, they do not consider the effect from retreating ice sheets, meltwater flux in the Southern Ocean, or insolation changes explicitly, unlike climate models. Other forcings besides CO2 and AMOC, could be included in the CO2 or AMOC coefficients, for instance, other gradual forcings have fairly high positive correlations with the CO2 forcing. Antarctic ice sheet changes could impact Southern Ocean SST. Ice sheets in the Northern Hemisphere could also impact the Southern Ocean through deep-water formation in the North Atlantic. This may explain the CO2 coefficients from the MLR and bipolar seesaw model that is higher than expected from ECS value. On the other hand, the AMOC sensitivity of the LOVECLIM model is low compared to the 1.5 ℃ Southern Ocean SST increase found in the simulation of Heinrich stadial 4 (Margari et al. 2020, Fig. S2), and the CO2 coefficient is quite high.

l. 438: can you specify which forcings may be poorly constrained or which feedbacks might be misrepresented

We changed the sentences to clarify that the focus of the sentence is on the mechanism and extent of cooling in the LGM simulation and how this affects the deglacial warming: "Hence, understanding the mechanism and amplitude of cooling in the LGM simulations will contribute to a better understanding of multi-model differences in the warming of the last deglaciation."

l. 465-467: "gap between climate response and ice sheet reconstructions": awkward, maybe rephrase.

Rephrased by merging the previous sentences. (The revised sentences are indicated in the next specific comment)

l. 467ff: Maybe also discuss here literature regarding the sensitivity to the specific design of the freshwater scenarios on the northern hemisphere (location and depth of input, as freshwater or as icebergs).

We discuss freshwater inputs scenarios in the fourth paragraph of the discussion [4-2]: "On the other hand, simulations that are forced with a large NH meltwater pulse consistent with ice sheet reconstructions do not simulate an ACR (Ivanovic et al., 2016; 2018; Kapsch et al., 2022; Bouttes et al., 2023). This so-called meltwater paradox (Ivanovic et al., 2018; Snoll et al., 2024) suggests a need for a better assessment of freshwater scenarios, and the potential sensitivity of climate model biases to freshwater forcing. We also note that the routing location of meltwater input (Roche et al. 2010; He et al., 2020) and the consideration of icebergs and meltwater discharge into the ocean (Schloesser et al., 2019; Love et al., 2021) may induce quite different AMOC changes".

4-3: The models used in this study are of relative coarse resolution. Maybe discuss if resolution (or unresolved processes) may explain discrepancies between models and observations.

We added discussion in the paragraph of section 4-3 on the horizontal resolution and possible unresolved processes. We propose oceanic eddy as one representative. " The coarse resolution of the used atmospheric models (2.5 to 5.6 degrees in the horizontal) may impact the contrast of warming between EAIS and WAIS through an inherent smoothing of the surface topography of the Antarctic ice sheet and the associated impact on the atmospheric circulation (Figure 4). In addition, relatively coarse resolution of the ocean models (1 to 3 degrees), may impact the AMOC sensitivity to iceberg and freshwater flux in the North Atlantic (Condron and Winsor 2012), or parameterizations of mesoscale processes in the Southern Ocean and their response to the deglaciation.'

Conclusion: The conclusion reads like a summary. Maybe shorten.

Thank you for comments, we have revised the conclusion to shorten it and are now addressing the questions raised in the summary already in the introduction.

Table 4: maybe averaged Southern Ocean SST, also please list γ.

We have clarified that Table 4 (and Table 5) is from averaged Southern Ocean SST and list intercept γ.

Table 5: unit of AMOC coefficient: I understood from the text that the AMOC is not normalized but binary. Also it would be good to evaluate the goodness of fit between original and reconstructed see-saw SST and to compare it to a respective metric (correlation coefficient) in table 4.

Thank you, we corrected the unit of AMOC coefficient in Table 5, and added coefficient of determination as in MLR analysis.

	CO_2 coefficient [K/83 ppm]	AMOC coefficient [K]	Response timescale [year]	Coefficient of determination
iTRACE	6.0	–2.9	500	0.97
LOVECLIM	4.4	–0.6	300	0.94
MIROC	2.4	–0.9	600	0.97
HadCM3	4.8	–1.3	700	0.99
MPI-ESM	3.4	–1.4	500	0.95

iLOVECLIM	2.0	–0.8	100	0.54

Revised table 5: Results of bipolar seesaw model.

References:

Anderson, H. J., Pedro, J. B., Bostock, H. C., Chase, Z., and Noble, T. L.: Compiled Southern Ocean sea surface temperatures correlate with Antarctic Isotope Maxima, Quaternary Science Reviews, 255, 106821, https://doi.org/10.1016/j.quascirev.2021.106821, 2021.

Anderson, Harris J; Pedro, Joel B; Bostock, Helen C; Chase, Zanna; Noble, Taryn L (2020): Southern Ocean Sea Surface Temperature Anomaly Stacks [dataset]. PANGAEA, https://doi.org/10.1594/PANGAEA.912158

He, F., Shakun, J. D., Clark, P. U., Carlson, A. E., Liu, Z., Otto-Bliesner, B. L., and Kutzbach, J. E.: Northern Hemisphere forcing of Southern Hemisphere climate during the last deglaciation, Nature, 494, 81-85, https://doi.org/10.1038/nature11822, 2013.

Prange, M., Jonkers, L., Merkel, U., Schulz, M. and Bakker, P: A multicentennial mode of North Atlantic climate variability throughout the Last Glacial Maximum, Science, 9, 44, https://www.science.org/doi/10.1126/sciadv.adh1106, 2023.

Roche, D.M., Wiersma, A.P. & Renssen, H. A systematic study of the impact of freshwater pulses with respect to different geographical locations. Clim Dyn 34, 997–1013. https://doi.org/10.1007/s00382-009-0578-8, 2010.

He, C., Liu, Z., Zhu, J., Zhang, J., Gu, S., Otto-Bliesner, B. L., Brady, E., Zhu, C., Jin, Y. and Sun, J.: North Atlantic subsurface temperature response controlled by effective freshwater input in “Heinrich” events, Earth and Planetary Science Letters, 539, 116247, https://doi.org/10.1016/j.epsl.2020.116247, 2020.

Schloesser, F., Friedrich, T., Timmermann, A., DeConto, R. M., and Pollard, D.: Antarctic iceberg impacts on future Southern Hemisphere climate, Nat. Clim. Change, 9, 672–677, https://doi.org/10.1038/s41558-019-0546-1, 2019.

Love, R., Andres, H. J., Condron, A., and Tarasov, L.: Freshwater routing in eddy-permitting simulations of the last deglacial: the impact of realistic freshwater discharge, Clim. Past, 17, 2327–2341, https://doi.org/10.5194/cp-17-2327-2021, 2021.

Condron, A. and Winsor, P.: Meltwater routing and the Younger Dryas, P. Natl. Acad. Sci. USA, 109, 19928–19933, https://doi.org/10.1073/pnas.1207381109, 2012.

Response to reviewer #2

Review of manuscript "Multi-model assessment of the deglacial climatic evolution at southern high latitudes" by Obase et al.

In the manuscript, the authors used six transient simulations for the last deglaciation to assess similarities and differences of their Southern Ocean/Antarctic changes, e.g. temperature, responses to AMOC and CO2, etc. The main issue is that different models have very different forcing histories (especially the freshwater schemes), which undermines the foundation of a MIP work. Under this context, the two simple models seems to a key to improving comparability of multi-model outputs, which I do acknowledge.

Nevertheless, I did not see enough substantial improvements for our understanding of dynamics of Southern Ocean and roles of AMOC and CO2 in Southern Ocean. It seems that the experimental designs of HadCM3, MPIESM and iLOVECLIM in general follow the protocol, while the others are on their own. What about summarizing common features and mechanisms from the former and discussing potential uncertainties based on the latter? In addition, it would be important to first introduce detailed performances of each model on LGM and PI before discussing their results. There have been some for sea ice, which I think is really helpful. One potential way is to also provide absolute values with the same axis-scale for different models. Attentions might also be given to statements that are based on results from models with different forcing schemes. I do like the Discussion which discussed potential contributions of different factors and existing uncertainties, directing next steps. Honestly, I'm not very familiar with the expected outcomes of a MIP paper especially in this case. Dynamic-wise, I would encourage the authors provide additional sensitivity runs to substantiate key inferences, for instance roles of the early FWF forcing/sea ice on the early SO warming, etc., even if based on one model.

Thank you for careful reading and the helpful comments. We adopt three major changes in our manuscript following your comments.

The T1 protocol paper leaves the FWF as flexible "The Core experiment protocol is flexible on whether or not to include prescribed ice melt (i.e. freshwater fluxes) delivered from the ice sheets to the ocean and how to do it" (Ivanovic et al. 2016, section 2-5). A PMIP multi-model study of the last deglaciation found that freshwater forcing in the North Atlantic had a dominant impact on northern hemisphere climate during HS1 (Snoll et al. 2024). Nevertheless, a comparison of six climate models can give the opportunity to assess the magnitude of warming or cooling. In the final paragraph of the introduction, we refer to one new multi-model PMIP deglaciation article (Snoll et al. 2024) and clarify the objective of this study:

Introduction (end of paragraph): The first PMIP multi-model study of the last deglaciation focusing on the northern hemispheric climate during HS1 found that different freshwater approaches (melt-routed, trace-like, bespoke, Snoll et al. (2024)) have a dominant impact on the North Atlantic climate variability. While this finding could be drawn due to the flexibility of the PMIP deglaciation protocol (Ivanovic et al., 2018) in regards to the choice of the method on how to distribute the freshwater forcing, this flexibility makes it challenging to properly compare the simulations. Nevertheless, a comparison of six climate models gives the opportunity to assess the magnitude of a climate warming or cooling, which is here carried out with help of simple models and the analysis of additional sensitivity experiments.

Regarding the performances of each model during the LGM and PI, we added sentences to refer to previous LGM multi-model articles for evaluation of model performances: We use the PMIP4 transient simulations of the last deglaciation performed with six atmosphere-ocean coupled climate models (Table 1). These simulations are initialised with LGM conditions, and the LGM climate fields have been evaluated by previous studies, particularly for global temperature changes (Kageyama et al. 2021), sea-ice and SST changes in the Southern Ocean (Green et al. 2022), and SAT changes over the Antarctic ice sheet (Buizert et al. 2021). A part of the transient simulations utilized in this study have also been compared to proxy-reconstructions (Weitzel et al. 2024).

For the third point: we have decided to add one set of sensitivity experiments from MIROC and HadCM3 in the manuscript focusing on AMOC and Southern Ocean SST changes during HS1. In these new MIROC and HadCM3 simulations, a larger freshwater flux is added during HS1, as in the iTRACE and LOVECLIM simulations (Figure 2 of the manuscript). We have added a fourth paragraph in section 4-2 (discussion) to discuss these new results: "We additionally show two simulations run with the MIROC and HadCM3 models to assess the role of the potential impact on southern high latitude climate of a significant AMOC decrease during HS1.

In the MIROC simulations, the FWF during HS1 is increased to 0.1 Sv or 0.2 Sv between 18 and 15.5 ka (Figure 11a, red and orange lines) instead of 0.03 Sv in the original experiment (Figure 2, Obase and Abe-Ouchi 2019). This larger meltwater input further weakens the AMOC (Figure 11a) and leads to an larger increase by 1 ℃ in SST in the SO compared to the standard experiment. The 1 ℃ warming in response to AMOC reduction of ~5 Sv is significantly higher than results from the MLR and bipolar seesaw models.

In the HadCM3 simulations, the freshwater flux from the "A" simulation of

Trace-21ka is used (Figure 11b blue lines). A larger freshwater flux during H1 significantly reduces the AMOC, and induces additional ~1 ℃ warming in Southern Ocean SST than the standard simulation with minimal freshwater forcing. The results from the MIROC and HadCM3 sensitivity experiments show that the simulated warming during HS1 can be twice as strong with an AMOC shutdown compared to the standard simulation of each model.

As in the LOVECLIM Heinrich stadial 4 simulation (Figure S2; Margari et al. 2020) the warming in the southern high latitude in response to AMOC strength is not necessarily linear, while MLR models assume linear temperature responses to the AMOC.

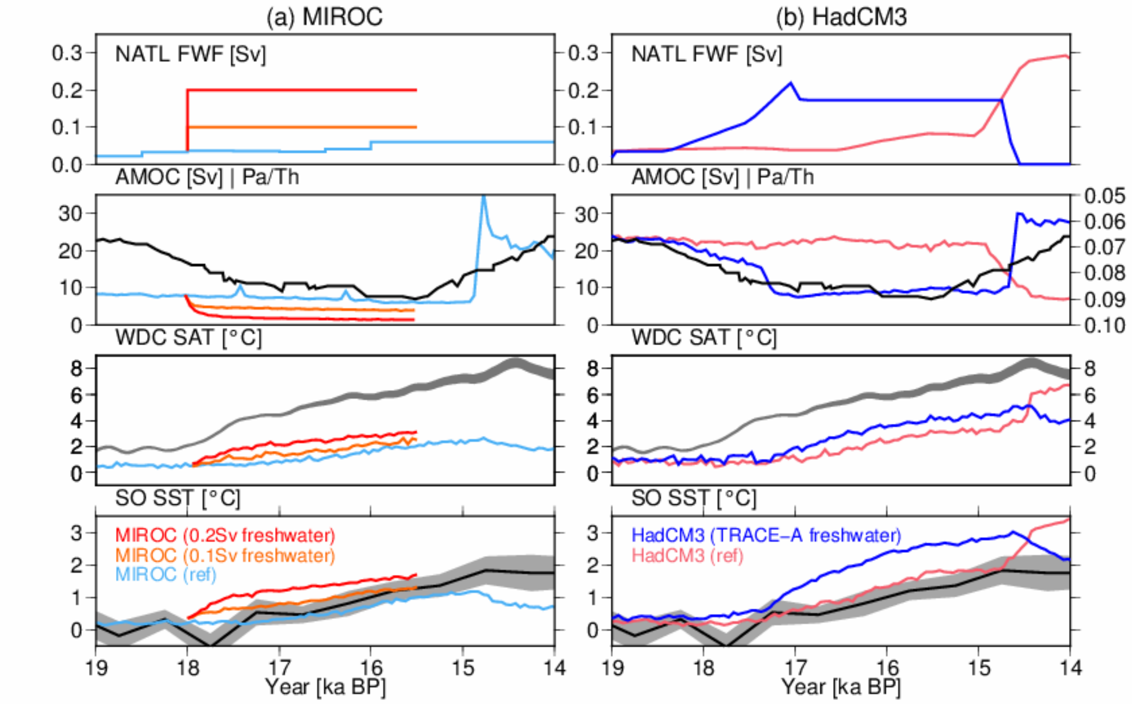


Fig. 11 (new figure): Results from transient deglaciation experiments performed with (a) MIROC and (b) HadCM3. The black lines in each panel represent proxy data, the same as Figure 3. In two MIROC sensitivity experiments, a larger amount of freshwater flux (0.1 or 0.2 Sv) is added into the North Atlantic (50–70°N) during 18-15.5 ka compared to the standard MIROC experiment (light blue lines). In the TRACE-A HadCM3 sensitivity experiment (blue lines), a larger freshwater flux is added in the North Atlantic following the Trace-21ka simulation (Liu et al., 2009), while the pink lines in (b) represent the HadCM3 simulation used in Snoll et al. (2022).

Minor comments:

Line 25-26: In the T1 protocol paper by Ivanovic et al 2016, it is suggested that the FWF scheme in the core simulation should at least follow the associated ice sheet history, which obviously is not the case for all the models here. Since these transient runs were initially conducted with different research focuses and none of them are really focusing on Southern Ocean (e.g. identical SO forcing, or referring to SO proxies, etc.), this avoids me from considering it as a typical MIP study for SO.

We have added one sentence in the abstract to clarify that the freshwater forcing is different between models. The comment is addressed in the introduction and method (general comment #1)

Note is that the protocol of the last deglaciation sets the choice of freshwater forcing as flexible, thus the freshwater forcing is different in each model.

Line 38-42: I'm not convinced since the models are forced differently and hence spread of results is expected.

We mention that the models are forced differently in the final sentences. We revised the sentences as follows. "Finally, all simulations exhibit minimal changes in Southern Hemisphere westerlies and Southern Ocean meridional circulation during the deglaciation. Improved understanding of the processes impacting southern high latitude atmospheric and oceanic circulation changes, and their impact on deglacial atmospheric CO2 increase are needed, as well as reducing forcing uncertainties particularly for the meltwater."

Line 184: what's the physical meaning of γ?

The term γ is intercept of the MLR model. We clarify it in the paragraph. "CO2 is the forcing used in each simulation, and γ is intercept."

Line 200-202: α and β here are different from those in eq. (1). It would be clearer with subscripts.

Thank you for your suggestion. We use α_1 and β_1 for MLR models and α_2 and β_2 for bipolar seesaw models, respectively.

Line 210-212: why is the initial SST(0) equal to 0? Why should it be the LGM values? Please provide more details about the way to generate the best fit parameters.

We have changed the term SST(0) as $\Delta SST(0)$ to clarify that the SST in the bipolar seesaw model is defined as the temperature change since the LGM. We have changed the equation (2) and (3) accordingly.

Second, we revised the sentences to provide more details about the method to get the best fit parameters. "At first, we conduct systematic sensitivity experiments to calculate minimum root mean square error between actual ΔSST and bipolar seesaw models. We conduct 9610 sensitivity experiments for each model within the parameter ranges shown in Table 3. The combination of parameters that gives the minimum root mean square error, along with coefficient of determination between the climate models' SST changes and bipolar seesaw models are displayed in Table 5"

References:

Snoll, B., Ivanovic, R., Gregoire, L., Sherriff-Tadano, S., Menviel, L., Obase, T., Abe-Ouchi, A., Bouttes, N., He, C., He, F., Kapsch, M., Mikolajewicz, U., Muglia, J., and Valdes, P.: A multi-model assessment of the early last deglaciation (PMIP4 LDv1): a meltwater perspective, Clim. Past, 20, 789–815, https://doi.org/10.5194/cp-20-789-2024, 2024.

Kageyama, M., Harrison, S. P., Kapsch, M.-L., Lofverstrom, M., Lora, J. M., Mikolajewicz, U., Sherriff-Tadano, S., Vadsaria, T., Abe-Ouchi, A., Bouttes, N., Chandan, D., Gregoire, L. J., Ivanovic, R. F., Izumi, K., LeGrande, A. N., Lhardy, F., Lohmann, G., Morozova, P. A., Ohgaito, R., Paul, A., Peltier, W. R., Poulsen,

C. J., Quiquet, A., Roche, D. M., Shi, X., Tierney, J. E., Valdes, P. J., Volodin, E., and Zhu, J.: The PMIP4 Last Glacial Maximum experiments: preliminary results and comparison with the PMIP3 simulations, Clim. Past, 17, 1065–1089, https://doi.org/10.5194/cp-17-1065-2021, 2021.

Green, R. A., Menviel, L., Meissner, K. J., Crosta, X., Chandan, D., Lohmann, G., Peltier, W. R., Shi, X., and Zhu, J.: Evaluating seasonal sea-ice cover over the Southern Ocean at the Last Glacial Maximum, Clim. Past, 18, 845–862, https://doi.org/10.5194/cp-18-845-2022, 2022.

Buizert, C., Fudge, T. J., Roberts, W. H., Steig, E. J., Sherriff-Tadano, S., Ritz, C., Lefebvre, E., Edwards, J., Kawamura, K., Oyabu, I., and Motoyama, H., Kahle, E. C., Jones, T. R., Abe-ouchi, A., Obase, T., Martin, C., Corr, H., Severinghaus, J. P., Beaudette, R. Epifanio, J. A., Brook, E. J., Martin, K., Aoki, S., Nakazawa, T., Sowers, T. A., Alley, R. B., Ahn, J., Sigl, M., Severi, M., Dunbar, N. W., Svensson, A., Fegyveresi, J. M., He, C., Liu, Z., Zhu, J., Otto-bliesner, B. L., Lipenkov, V. Y., Kageyama, M., and Schwander, J.: Antarctic surface temperature and elevation during the Last Glacial Maximum, Science, 372, 1097–1101, https://doi.org/10.1126/science.abd2897, 2021.

Weitzel, N., Andres, H., Baudouin, J.-P., Kapsch, M.-L., Mikolajewicz, U., Jonkers, L., Bothe, O., Ziegler, E., Kleinen, T., Paul, A., and Rehfeld, K.: Towards spatio-temporal comparison of simulated and reconstructed sea surface temperatures for the last deglaciation, Clim. Past, 20, 865–890, https://doi.org/10.5194/cp-20-865-2024, 2024.

Kobayashi, H., Oka, A., Obase, T., and Abe-Ouchi, A.: Assessing transient changes in the ocean carbon cycle during the last deglaciation through carbon isotope modeling, Clim. Past, 20, 769–787, https://doi.org/10.5194/cp-20-769-2024, 2024.