We thank the reviewer(s) for providing precise and valuable feedback on our manuscript. We will be happy to submit a revised 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.

Response to reviewer #3:

In this paper, Obase and co-authors compare transient deglacial simulations that have been run as part of the Palaeo Modelling Intercomparison Project (PMIP), and assess the climate evolution in Antarctica and the Southern Ocean. The paper covers the general comparison of important climate metrics like ocean overturning circulation, surface air and ocean temperatures, etc. and link changes in those metrics to each other. They use two simplified models, a multiple linear regression (MLR) model and a bipolar seesaw model, to assess the response of Southern Ocean SST to changes in Atlantic meridional overturning circulation (AMOC) and atmospheric CO2.

This is version 2 of the manuscript, which has been revised based on the comments of two other reviewers (not me). The introduction of the paper reads well and covers relevant literature. The methods have been clarified, but I suggest some additional edits. While the results and discussion of the paper are useful and generally presented in a reasonable way, these parts of the manuscript still require substantial polishing to make the material more easily accessible to the reader. At the moment, some concepts are not carefully defined, acronyms are introduced without definition and/or to describe essentially the same thing as another (e.g. WDC/EDC becomes WAIS/EAIS), figures and captions are missing some details, and some phrasings are not clear. This makes the manuscript a lot more difficult to follow than it needs to be. I therefore suggest major revisions of the paper before it is accepted for publication.

Thank you for carefully reading and giving us fruitful comments. We have

revised the manuscript not to make confusion. The manuscript changes are addressed in the specific lines below.

One useful outcome of the study is the comparison in deglacial Southern high latitude (SHL) climate response between the different FWF forcing strategies. As mentioned in the manuscript, PMIP models are, so far, incapable of reconciling FWF resulting from the reconstructed ice sheet retreat with the reconstructed AMOC behaviour in this time period (meltwater paradox, Ivanovic et al., 2018; Snoll et al. 2024). While this may have been discussed for the Northern Hemisphere (Snoll et al., 2024), and assessed in individual model papers, this paper allows us to compare the model behaviours that result from the different FWF strategies also for the SHL, and get an idea of the biases that can result from this lack in model skill. Therefore, I would have liked to see a clearer discussion of this. The first half of the final paragraph in Section 4-3 talks about this but does not pointedly link back to the two different FWF strategies. To me, not discussing this in a more concrete way seems like a missed opportunity.

Thank you for giving us fruitful comments. We have created one discussion subsection to have further discussion on FWF by referring to sensitivity studies from individual models. The change in the manuscript is addressed in the specific comment L477-494.

General Comments

[1] The definition of the reference period is not entirely clear. In some instances (e.g., Section 3-2-2), pre-industrial is mentioned but never defined in terms of what that means for these simulations. For the most part, 11ka seems to be the reference period but, in my experience, that would not be referred to as preindustrial but rather Holocene.

Generally, most figures are analyzed change since Last Glacial Maximum, 21 ka. The comparison with pre-industrial are mainly conducted to evaluate the Last Glacial Maximum, which is the initial condition of the deglaciation experiments. We clarify the reference periods in the manuscript accordingly.

[2] Bolling-Allerod (BA, 14.7 – 12.8 ka, or 13 ka depending on where in the manuscript you look) and Atlantic Cold Reversal (ACR, 15 – 13 ka) are used nearly interchangeably in the manuscript, which causes unnecessary confusion for the reader. While BA is used in the figures and earlier results sections, ACR starts being used from Section 3-6 and onwards through the Discussion, while still referring to figures with a shading gap that encompasses the BA (which is no longer mentioned). As the ACR is more relevant to this region and to the results of the paper, and even highlighted in the Abstract, it is not clear why it is not indicated earlier in the results section nor in the figures.

ACR is replaced with BA in several parts of the manuscript. As section 4-3 discusses the Antarctic cooling event, we keep the term ACR, but in some places we clarify the time period of ACR corresponds to BA:

…and an abrupt AMOC increase at the end of HS1 is necessary to simulate the high southern latitude cooling during BA (corresponding to the ACR)

[3] Figure captions require some clarification

We have addressed them in the specific lines below:

[4] Figure references are somewhat lacking in 3-2-1 and 3-2-2 (see specific comments). References to subpanels are very rare, even though they would be appropriate. In a paper that has so many figures, more substantial figure referencing could be very helpful to the reader.

We have added subpanels to clarify the reference of figure panels. We have addressed them in the specific lines below.

[5] In the response to previous reviewers, the authors include a figure R.1. with results from the bipolar seesaw model, which nicely demonstrate that all models would have simulated the ACR if their AMOC changes looked as those in iTRACE. This is a useful result but it is not clearly mentioned in the updated version of the manuscript. I suggest adding the figure to Supplementary figures, and discuss

the result in section 4-2.

Thank you for comment; we have put the figure in the Supplementary Figures, and add their descriptions in the manuscript:

Results: We also applied the bipolar seesaw model by using AMOC and CO2 coefficients from six different models (Table 5), but common inputs of CO2 (Bereiter et al., 2015) and AMOC from iTRACE. The results indicate that all models would have simulated cooling in the Southern Ocean SST during BA if there was an increase in the AMOC at the beginning of BA as iTRACE, LOVECLIM or MIROC (Fig. S2).

Discussion: This discussion is supported by results of bipolar seesaw model with common AMOC input (Fig. S2).

General comments:

L. 32: For clarity, consider adding “at this time” before the parenthesis.

Changed as suggested.

L. 49-51: Ocean processes should also be mentioned among the feedbacks, e.g., with a reference to Kohfeld and Ridgwell (2009).

We have clarified that atmospheric CO2 changes are related to ocean processes using the reference.

“including changes in atmospheric greenhouse gas (GHG) concentrations by ocean processes (Kohfeld and Ridgwell, 2009)…”

L. 58: Please define “early Holocene”.

In this article, we have defined Early Holocene as ~11 ka, which is an end of the analysis period.

“last deglaciation (transition from the LGM to the early Holocene, ~11 ka)”

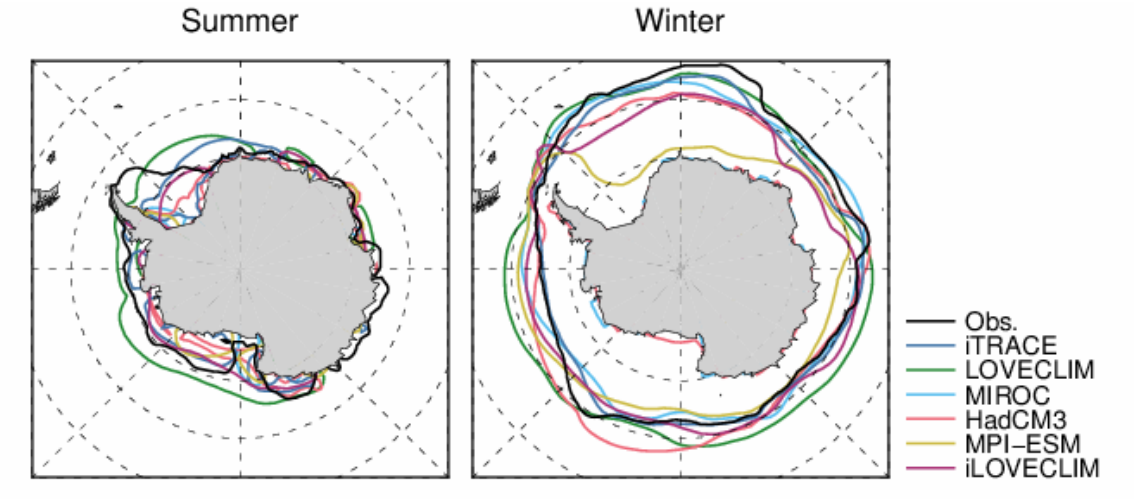
L. 150-151: “focus on the magnitude and rate of Antarctic […] (SST) and […] (SAT) changes” – This sentence would read better as “focus on the magnitude and rate of changes in Antarctic […] (SST) and […] (SAT)”

Changed as suggested.

L. 151: “difference between the AMOCs” – This should be more specific, e.g. “divergence between the AMOC’s strength and temporal evolution”.
Changed as suggested.

L. 162: “These simulations are initiated with glacial conditions” – It would be helpful if the authors would add here what time period the simulations cover, or at least what part of these simulations is analysed. This is not entirely clear from the manuscript text in its current form.
We have clarified the analysed time period in the method section.
We use the PMIP4 transient simulations of the last deglaciation performed with six atmosphere-ocean coupled climate models (Table 1), and analyse the time period from LGM to the Early Holocene 21–11 ka.

L. 167-168: “which shows models simulate reasonable sea ice extent” - I somewhat disagree with this statement. Two models severely underestimate the sea ice extent (SIE), and most others underestimate but to a lesser degree. I suggest changing to something along the lines of “which shows most models underestimate pre-industrial SIE, but mostly to an acceptable level, with more severe underestimates by LOVECLIM and, in particular, iLOVECLIM.”
We agreed with this, but after we have checked the figure script of Fig. S1, we have found some mistakes in data processing for three models (LOVECLIM, iLOVECLIM, MPI-ESM). The sea ice in LOVECLIM and iLOVECLIM were underestimated in the previous Fig. S1, now we have found that the underestimation of PI is reduced so we don’t need specific mention of these two models. Still, there’s some model biases so we have clarified that the model performance of sea ice are mostly acceptable level.
Text: “some models underestimate pre-industrial summer sea ice extent, but mostly to an acceptable level.”



Revised Figure S1: Simulated summer and winter sea-ice edge for the pre-industrial (PI) experiments from six climate models used in this study (color lines). The black lines indicate summer and winter sea-ice edge from observational data (COBE-SST for 1981-2010, Hirahara et al. 2014).

L. 177-178: “elevation change since the LGM” – Make it clear that the elevation is lower in present-day climate by changing to “elevation reduction since the LGM”.

We have clarified the elevation reduction since LGM.

L. 181-182: [a] and [b] – Consider changing to capital letters to avoid confusion with panel letters in the figures cited in the same sentence. These groups could be indicated with similar line styles in the figures where all models’ time series are shown together. In the current version of the manuscripts, these groups are defined but never referred to with these attributes a and b again. I feel like you could either rewrite the sentence and skip the attributes altogether “classified into two groups; the first group with FWF adjusted […], and the second group with FWF consistent with […]” or, preferably, actively refer back to the two defined groups later in the text.

We have used “first group” and “second group” to distinguish FWF schemes.

And we have referred back to “FWF groups” in sections 3-5, 3-7 and 4-2.

L. 185-187: “In LOVECLIM and MIROC, the meltwater flux was uniformly applied […] while other models […] apply a spatially varying FWF (Table 2) – This particular difference in strategy is not clear from Table 2, and could be clarified in the table. The table mentions the freshwater schemes, but without explanations for what “TraCE-like” and “ICE-6G_C” mean, it is difficult to understand if this refers to the spatial distribution or something else.

We have added the areal information in FWF column of Table 2.

L. 202: “maximum and minimum values” a “temporal maximum and minimum values” (if this is what is used in the normalisation?)

We have clarified that temporal maximum and minimum values are used:

L. 204-206: “Every 100-years mean SST, AMOC and CO2 […], so each dataset has 90 timeslices.”– Does the selection of these 100-year intervals matter? Would the results differ if intervals were spaced/centered differently?

We have checked if we get the different results when 50-year interval is used. We have tested two models covering both FWF groups here as representatives (MIROC and HadCM3). The results of the thermal bipolar seesaw models are as follows: (The parenthesis value: 100-year interval)

MIROC: CO2 2.2K (2.4K), AMOC 0.8K (0.9K), timescale 500a (600a)

HadCM3: CO2 4.8K (4.8K), AMOC 1.3K (1.3K), timescale 600a (700a)

We got slightly different coefficients but within a range of rounding. We have revised the sentences to indicate that we take 100-year mean climatology to focus on long-term changes by reducing internal variabilities.

To focus on long-term climate change and reduce interannual variabilities, every 100-years, mean SST, AMOC,…

L. 209: At the end of Section 2-2-2, you mention the results table of the bipolar seesaw model (Table 5). You do not mention Table 4 (results of MLR) at all before then. As the table numbering should follow the text, you should mention it here,

at the end of Section 2-2-1, in a similar way as Table 5 is mentioned in the next section. That would lead to renumbering of Tables 3 and 4 (switch order) throughout the manuscript but would still allow for related tables to follow each other.

Table 4 is mentioned at the end of section 2-2-1, and renumbered Tables 3 and 4 accordingly.

L. 217: The term *m(t)* represents the AMOC modes. That would be even clearer from a quick glance at the equation if you rename it as *AMOCm(t)*.

Changed as suggested to clarify *AMOCm(t)* represents AMOC binary modes.

L. 223-224: "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." – I assume that AMOC < 14 Sv means m(t)=1, but it should be phrased more clearly that the modes are binary.

We have revised the sentences to clarification:

it is assumed that the AMOC modes are binary, unlike the continuous values

L. 236-237: "weaker LGM AMOC if GLAC-1D ice sheet was used" – I suggest "weaker LGM AMOC in sensitivity studies where GLAC-1D ice sheet was used"

Changed as suggested

L. 238: "compared to pre-industrial (PI)" – When reading the paper, it is quite difficult to follow what the different reference periods are. Sometimes, PI is used (and thus a different control simulation), and sometimes results are compared to 11 ka (thus the transient simulation). The difference is not made sufficiently clear, and I got lost several times reading the parts that make these comparisons.

Generally, most figures are analyzed change since Last Glacial Maximum, 21 ka. The comparison with pre-industrial are mainly conducted to evaluate the Last Glacial Maximum, which is the initial condition of the deglaciation experiments. We clarify the reference periods in the manuscript accordingly.

L. 248: “display an AMOC weakening due to” – It is not clear over what time interval the weakening is happening. Please clarify by changing to “display an AMOC weakening at <time> c.f. <reference time> due to”

We have clarified the time periods as suggested:

“…the other three simulations (HadCM3, MPI-ESM, iLOVECLIM) display an AMOC weakening at 14.7 ka compared to HS1 due to a significant increase in FWF…”

L. 250: “simulate an AMOC decline” – This wording suggests a continuous decline throughout the time period. Changing to “a weakened AMOC state c.f. LGM and the BA” would make the wording less open to interpretation.

Changed using your suggestion:

“iTRACE, LOVECLIM, and MIROC simulate a weakened AMOC state compared to BA”

L. 252: “a gradual AMOC reduction” – Over what time period? Still YD? If so, I suggest changing to “a small but gradual AMOC reduction throughout the YD”

Changed using your suggestion:

“HadCM3 simulates a small but gradual AMOC reduction throughout the YD”

L. 257-260: Please improve figure referencing (specific panels)

We have revised the manuscript using figure panel references.

L. 262-263: For the EDC warming, refer to Fig. 4a. Noting here that MPI-ESM overestimates the EDC warming c.f. the ice core estimate.

We have referred to Fig. 4a and add note on MPI-ESM overestimates SAT at EDC core site (Fig. 4a).

This simulated EDC warming is comparable with EDC ice core estimates, with MPI-ESM overestimates warming at EDC site (Parrenin et al., 2013Fig. 4a)

L. 268-269: What temperature is referenced when talking about larger/largest

warming (+6-8 degC)? The SATs in Fig. 3? If so, refer to Fig 3 c-d, and clarify the text accordingly.

We have clarified that the sentences refers to SAT at WDC and EDC.

L. 276-278: Noting that iTRACE also has the largest sea ice area to start with, which may give it a larger potential for reduction compared to other models.

We have added this information on the largest sea ice area here:

Result:

L. 280-282: Please improve figure referencing (specific panel for AMOC) and add reference to Fig. 3 with panels for SAT and SST.

We have added figure references to clarify the reference to each panels (Figures 3 and 4).

L. 296: “The total deglacial (21-11 ka) warming” - I suggest changing to “The proxy-record based estimates of total deglacial (21-11 ka) warming”

We have revised the sentences

The proxy-record based estimates on total deglacial (21–11 ka) warming

L. 297-300: Please add figure references including specific panels.

We have added figure references to clarify the reference to each panels (Figures 3-5).

L. 299-300: “The Southern Ocean sea ice edge retreats poleward by 10 deg latitude in most models” – This might not be realistic, as most models underestimate the PI sea ice area.

This is right, we have revised the sentence to indicate models that exhibit the largest sea ice reduction is 10 deg latitude.

sea ice edge retreats poleward by 10° latitude in most models with the largest sea ice retreat (Fig. 5)

L. 300: “In this area” – Please, be more specific. I am not sure which area this refers to.

We have clarified the area refers to Southern Ocean near winter sea ice edge:

A SST increase of up to 6 ℃ is simulated in the Southern Ocean near winter sea ice edge in iTRACE…

L. 309-311: “This indicates that […] based on ice core reconstructions” – I am struggling to follow this sentence. Is it models that would require additional Holocene warming, or is it proxy record temperatures at 11 ka that are similar to pre-industrial? Also, the onset of the Holocene has not been clearly defined in the manuscript.

We have clarified that models would require additional Holocene warming in their simulations (11 to 0ka). As discussed in the discussion paragraph (last paragraph of section 4-1), one cause of model-proxy difference can be explained by ice sheet changes in the early Holocene to 0 ka in the experimental design (Fig. 1d). In addition, we have clarified the text to state the reference is 11ka.

…total temperature change between LGM and PI for five out of six models, indicating additional warming would come during 11 to 0 ka in the model simulations. This is different from ice core reconstruction, in that the temperature at 11 ka is comparable to the pre-industrial values (Parrenin et al., 2013; Buizert et al., 2021).

L. 319: “significant AMOC changes” - Throughout the manuscript, changes are several times referred to as significant (see e.g. in Section 3-6), though I do not get the impression that the actual statistical significance is tested. I suggest using a different wording, e.g. “substantial AMOC changes” in these cases, to make it clear that these types of statements do not refer to actual statistical significance.

As most parts do not indicate the statistical significance, we have revised the sentences:

L. 320: “an AMOC weakening” a Not every blue circle shows a weakening c.f. to

the previous point. You need to specify that you are referring to a weaker AMOC with respect to the long-term mean.

We have revised the sentence:

a weaker AMOC compared to long-term mean of respective models (blue circles)

L. 330-331: “has the potential to explain about half of the SST changes” – What is this estimate based on when the model range is so large? Is it just a mean of all coefficients? Also, I would recommend being more specific as to what SST changes you are referring to, e.g. rephrasing to “Southern Ocean SST anomaly”.

We have removed the sentence “The negative coefficient of AMOC in all models…” from here (results) because the previous sentence is rather speculation. The quantitative impact of AMOC changes during HS1 is addressed in the section of additional sensitivity experiments on HS1 freshwater (section 3-7), bipolar seesaw model with common AMOC inputs (Fig. S2), and discussed in the new discussion subsection (4-3)

L. 337: It does not quite make sense to me to have the bipolar seesaw model as a separate subsection when the entire section talks about both MLR and bipolar seesaw models and compare them. This section seems more like a continuation of 3-4.

We have merged MLR and bipolar seesaw model as these covers the same topic with slightly different method.

L. 353-374: This section lacks some motivation for why these additional metrics are valuable, and for why certain choices have been made (notably the focus on the onset of deglaciation (L. 354-355), and the subsurface temperatures south of 60oS at depth of around 500 m.

We have revised the first sentence as “topic sentence” to focus motivation and summary of additional variable.

…Southern Ocean (zonal mean winds averaged over 65°S–40°S), to discuss potential impact on deglacial CO2 changes, which is prescribed in the

experiments

L. 359-361: Models show different AABW behaviour but, meanwhile, they all show increased SSTs. This might be worth noting.

We have added a sentence to indicate the relationship between Southern Ocean SST and AABW during HS1.

Thus, while there is agreement in Southern Ocean SST warming during HS1, the trends in AABW differ.

L. 361-365: These sentences address only the strength of the zonal winds. However, the peak strength latitude could also be of importance, e.g. for SST fields due to shifts in upwelling (add refs.)

We have analyzed latitudinal variations in the SHW when we were preparing the manuscript, but we could not find significant latitudinal shift in the Southern Ocean winds. Meantime, we think detailed analysis is needed because we have found some zonal heterogeneity in the responses of Southern Ocean winds in some cases (at least some periods in MIROC), so we would like to leave the further analysis of Southern Ocean winds a future topic. In the revised manuscript, we have referred to the related articles here:

Discussion: as well as regional changes in SHW strength and their relation to other climatic variables (Rojas et al., 2009; Sime et al., 2013)

L. 366: It would be nice to mention here, and possibly discuss further, that the behaviour of AABW formation, and other variables in this section, are decoupled from the AMOC behaviour. AABW formation behaviour is also not consistent within the two groups of FWF forcing strategies.

Yes, timeseries of AABW seems decoupled from AMOC and also FWF schemes. We have added one sentence:

In addition, the behavior of AABW seems different from AMOC and FWF schemes (Fig 10a-b).

L. 369-370: “During the ACR (15-13 ka)” – In the rest of the results section, the focus is on the BA (14.7 – 13 ka), which almost overlaps completely with the ACR. Thus, it seems unnecessary to use two names for such closely overlapping time intervals. I suggest either choosing one of the two, or making it clearer how the two are related and why you want to highlight the ACR.

As the time period of BA overlaps with ACR, we have changed to “BA” here.

L. 376-389: This section lacks a few references to specific figure panels (L. 380-381, Fig. 11d; L. 384-385, Fig. 11h).

We have added subpanels (a-h) in Figure 11 and added references to subpanel accordingly.

L. 376: Indicate here that MIROC and HadCM3 belong to different groups of the FWF forcing strategies.

Thanks, that is important point, we have revised the sentence:

We remind that the MIROC and HadCM3 comes from different freshwater schemes (Fig. 2).

L. 384: “significantly reduces the AMOC” – I suggest changing to “reduces the AMOC by ~15 Sv”. Also, the figure panel letter should be 11f, not b, I think.

We have revised the sentence, and added subpanels (a-h) in Figure 11.

L. 384-385: “induces an additional ~1 °C increase in Southern Ocean SST compared to the standard simulation.” – How does this compare to the MLR and bipolar seesaw model results?

As already stated in case of MIROC, We have added the sentence to discuss whether this result is predicted by MLR model:

The simulated HS1 warming in HadCM3 is consistent with both MLR and bipolar seesaw models (Tables 3 and 5).

L. 422: ~1 °C per 100m – I presume that this is a decrease with increasing altitude,

but this should be made clearer.

We have clarified the sign of changes:

SAT (~1 °C warming per 100 m altitude reduction)

L. 426: “contrast between EAIS and WAIS” – These acronyms have not been previously defined. Here, it also seems unnecessary to introduce further acronyms to refer to the same areas that have, so far, been known to the reader as WDC and EDC. I suggest either sticking to the same acronyms as before or writing out “contrast between the East and West Antarctic ice sheets”.

We have revised the sentences by removing EAIS and WAIS acronyms:

warming contrast between East Antarctica (EDC ice core site) and West Antarctica (WDC ice core site) through…

L. 427-428: “and the associated impact on the atmospheric circulation” – Here, I would like to see a reference that explains what these impacts may be.

This discussion is based on Buizert et al. (2021) referenced above, which compares the representation in surface topography in HadCM3 and MIROC. We have clarified the reference.

…through an inherent smoothing of the surface topography of the Antarctic ice sheet and the associated impact on the atmospheric circulation (Buizert et al., 2021).

L. 428-431: “In addition, the relatively coarse resolution […] may impact the AMOC sensitivity […] or parameterizations of mesoscale processes […] – I think you need to phrase more clearly what you mean when saying that the resolution impacts the parameterizations. On the one hand, there is the effect of the use of an eddy parameterization, but I am not sure if that is what you are referring to, or if you also mean other effects. Adding some helpful references would also be advisable.

The sentence refers to mesoscale oceanic eddies (Menviel et al. 2023), which is frequently parameterized in coarse-resolution ocean model. We have clarified

that in the text.

L. 433-434: “additional warming occurs after the onset of the Holocene” – As the onset of the Holocene has not been clearly defined, I struggle to find this warming in the figure.

We have changed “onset of the Holocene” to “11 ka” for clarification.

L. 436-438. “The higher surface elevation […] may contribute to the simulated Holocene warming” – This confuses me. Above, you say that lower elevation leads to warming. Please, clarify this!

We have revised the sentences:

The decrease in surface elevation of the Antarctic ice sheet after 11 ka (Fig. 1e) may contribute to the Holocene warming after 11 ka.

L. 442-443: “HS1 […] AMOC changes” – It could be useful to reiterate that these sensitivities are tested in the model.

We agree the point, we have added the information in this sentence by referring Fig. 11:

sensitivity of each model to CO2, and to AMOC changes comes from FWF schemes as shown in additional sensitivity experiments using MIROC and HadCM3 (Fig. 11).

L. 449: “weakly simulated” – I suggest changing to “not fully represented” .

We have revised the sentence:

different magnitudes of warming between Southern Ocean SST and Antarctic SAT are not fully represented in models

L 449-450: “the largest global mean changes at the LGM (7.3 °C, compared to the six-model mean of 5.3 °C)” – This phrasing makes it sound like the models are warmer at the LGM. I suggest “the largest cooling of global mean SAT at the LGM c.f. PI”.

We have revised the sentences as suggested:

L. 452: “LGM global mean SAT changes” - Please be specific in what those changes are in reference to (PI?)

The reference is PI here; we have clarified it.

Global mean SAT at the LGM compared to PI (7.3 ℃

L. 463-466: “However, the simulation length […] climate response at high southern latitudes” – The sentence is long and does not read well. I suggest splitting into three sentences, the first ending after “less than 420 years”. The second sentence can be rephrased as “In this study, we estimate the timescale for the bipolar seesaw to be 500-700 years (Table 5). Thus, longer simulations may be needed to evaluate the extent of the climate response at high southern latitudes.”

We have revised the sentences by dividing:

However, the simulation length in their study is less than 420 years. In this study, we estimated the timescale for the bipolar seesaw to be i~500–700 years. Thus, longer simulations are needed to evaluate the extent of the climate response at high southern latitudes.

L. 469-471: “Other forcings […] have positive correlations with the CO2 forcing.” – The message of this sentence is not quite clear to me. Do you mean that other gradual forcings that have positive correlations with the CO2 forcing could be included in the coefficients? I feel like there may be a word or two missing to make the meaning clear.

Yes, other gradual forcing, particularly NH summer insolation (positive correlations with the CO2 forcing) could be included in the coefficients. We have clarified NH summer insolation:

Other forcings could be included in the CO2 or AMOC coefficients in this analysis. F instance, insolation change in Northern summer have positive correlations with the CO2 because both exhibit gradual increase (Fig. 1a-b).

L. 471-472: “Antarctic and Northern Hemisphere ice sheet changes could impact Southern Ocean SST through deep-water formation” – How this happens should be briefly explained, adding appropriate references.

We have added one reference (Abe-Ouchi et al., 2015) to show that the different LGM ice sheet (PMIP3 compared to PMIP2) can induce further Southern Ocean SST cooling. Meantime, we have “deleted deep-water formation” because it is not mentioned in the above reference.

L. 477-494: I suggest making it clearer how this part of the discussion relates to the FWF strategies.

Following these lines and major comments above, we have created one subsection 4-3 to clarify the discussion of FWF. While most sentences come from the subsection 4-2 [rate of temperature changes] of the previous version of manuscript, we have added several sentences:

Subsection 4-3: Freshwater forcing and temperature changes in southern high latitudes

…

FWF in the NH and associated reduction in the AMOC lead to Antarctic warming during HS1 in the sensitivity experiments with individual forcing (He et al., 2013) and sensitivity experiments with FWF (Figure 11). A comparison of the six models reveals that capturing phase changes in the AMOC is necessary to simulate warming or cooling trends in the southern high latitudes, supported by results of bipolar seesaw models with a common AMOC (Fig. S2). While MIROC simulates a rapid increase in the AMOC at BA transition with a continuous freshwater flux (Fig. 2), a greater FWF by a factor of 1.5 weakens the AMOC throughout the last deglaciation (Obase et al., 2021). Thus, the simulated temperature changes in southern high latitudes are sensitive to FWF in the NH, and the amount of meltwater is still different between models and ice sheet reconstructions. The so-called meltwater paradox (Ivanovic et al., 2018; Snoll et al., 2024) is present for a major period of the last deglaciation including BA transition and HS1,

suggesting a need for a better assessment of freshwater scenarios, and a need to assess sensitivity of climate models to freshwater forcing, and to reduce potential climate model biases. …

L. 506-511: Note that an “accurate” position of the SHW in PI simulations c.f. observations is difficult to achieve when using an atmospheric model of insufficient resolution (Guemas and Codron, 2011; Hourdin et al., 2012). If the position in the control state climate is biased, this affects the potential range for a latitudinal shift in the wind position.

We agree that the biases in PI SHW is one probable cause, we have added in the manuscript:

Biases in the SHW in PI simulations can be a cause of little changes in SHW, however, additional studies should look in more details into potential changes in the SHW in these simulations,

L. 511-514: The message of the two sentences “In addition, no model […] and decreasing surface salinity” is unclear. Is the second sentence talking about the models or the real world?

The reference Marson et al., (2016) is model study. We have revised the sentences to clarify the discussion:

Instead, the long-term weakening in AABW by warming the Southern Ocean, enhancing sea ice melt are consistent with previous analysis of deglaciation experiments (Marson et al., 2016).

L. 516: (Marzocchi et al., 2020; Stein et al., 2020) – Please also add a reference to Ferrari et al., 2014: https://doi.org/10.1073/pnas.1323922111

We have referred to Ferrari et al., (2014) here.

L. 534-535: “smaller LGM sea ice extent” – It is not clear what you are referring to. The LGM sea ice extent is smaller compared to what?

“smaller” should be multi-model difference. We have revised the sentence:

The multi-model difference could be related to the difference in LGM sea ice extent,

L. 538: “a major warming occurs” – Please specify where the warming occurs, e.g. a major warming occurs in the high southern latitudes (HSL). Note that this acronym can also be used in several other places in the Conclusion.
We have clarified that a major warming occurs in southern high latitudes.

L. 540: “a larger warming” – Specify where the warming takes place by saying “a larger HSL warming”
We have clarified that a major warming occurs in southern high latitudes.

L. 544-546: “The simulations do not exhibit significant changes in […] meridional circulation in the Southern Ocean” – Where in the paper was this shown? AABW formation changes substantially.
L. 546: “which could contribute to enhanced CO2 outgassing from the Southern Ocean” – Does that outgassing really happen? If not, rephrase to “which could have contributed” and make it clear whether this is observed or not.
We revised the manuscript to indicate that there were some changes in AABW and winds, but appear to be weaker to account for atmospheric CO2 increase indicated from ice core reconstructions:
The simulations exhibit some changes in winds over the Southern Ocean or meridional circulation in the Southern Ocean, which could have contributed to enhanced CO2 outgassing from the Southern Ocean

Table 1, L. 573-574: It is not clear to me what the difference is between “Name” and “Climate model name”. References are missing for MPI-ESM and iLOVECLIM. If these simulations are not previously published, then please indicate this.
“Name” is a name of the simulation (iTRACE) or shorter name of model (e.g.

MIROC, HadCM3). The references are the description paper of each model, but we have added the reference for MPI-ESM and iLOVECLIM in Table 1.

L. 574: "ECS" – The acronym has not yet been used in the text when the table is introduced, so please also write it out here as "Equilibrium Climate Sensitivity (ECS)".
We have changed accordingly.

Table 2, L 576-577: In the text, PMIP 4 is referred to as "Bereiter et al. (2015)". Given that Kohler et al. (2017) is given for the other choice in the table, I would prefer if Bereiter et al. were also listed instead of PMIP4.
We would like to keep "PMIP4" in Table 2 because PMIP4 GHG is composed of three different references corresponding three GHG components (CO2, CH4, N2O), comes from (Bereiter et al. 2015; Loulergue et al., 2008; Schilt et al., 2010), which is described in Ivanovic et al. (2016). We have clarified it in the caption:
PMIP4 in the column GHGs indicates they use three GHG reconstructions (CO2, CH4, N2O) in PMIP4 protocol paper (Ivanovic et al., 2016)

Figure 1, L. 583: The caption is not entirely clear. Panel a) Does both insolation lines come from the same reference?, c) There is only one line per ice reconstruction in this panel, but it needs to be clarified that the same colours apply in the next panel, d-e) There is no panel e.
We have clarified that both insolation come from the same reference, and clarified black and red lines in panel (d), and deleted panel (e)

Figure 2: It would be helpful to indicate the groups of FWF in the figure.
We have indicated the FWF grouping in the figure caption, following the method section of the manuscript.
The top three model panels correspond to the first group with FWF adjusted to reproduce large-scale AMOC variability, and the bottom three model panels correspond to the second group with FWF consistent with the reconstructed ice

volume changes.

L. 591-592: Use these two lines about the shading in all other figures that use this shading (Fig. 3, 6, 10). Each figure caption should be standalone.
We have added the information on gray shading in other figures.

Figure 3: Check if this figure is readable for people with colour vision impairment (e.g. using https://www.color-blindness.com/coblis-color-blindness-simulator/). If not, try and also distinguish lines that do not stand out enough by using different line types. Add description about shading (see L. 591-592) to caption.
We have followed the color schemes using the following: https://cran.r-project.org/web/packages/khroma/vignettes/tol.html
And we have added the information on gray shading (consistent with Figure 2)

L. 595: e) "Southern Ocean SST" – This should be more specific, i.e. "zonal mean SST in the latitude band 40-55 °S"
We have clarified the definition of SO SST in the caption.
(e) Southern Ocean SST (zonal mean SST in the latitude band 55–40°S),

Figure 5: Colour bar title and unit are missing. This is nicely done in Figure 4, but missing here. I am also somewhat confused as to why SATs are not shown here as they are in Fig. 4.
We have added color bar title in Figure 5. SAT for 21-18ka (Fig. 4a) is important for the onset of Antarctic warming, but SAT for 21-11ka is less discussed in the manuscript.

Figure 6, L. 613: In general, the approach used in this figure is informative, but why did you choose the EDC of Parrenin et al. (2013) for the grey line in b? Noting that Buizert et al. (2013) provide both WDC and EDC and could thus also be used. Please, motivate this choice or use both references. Add description about shading (see L. 591-592) to caption.

We have added Buizert data (for EDC) to make consistent with Figure 3, and added vertical gray shades same as Figure 3.

Figure 7: I would choose to use other colours to indicate the AMOC strength, because blue is instinctively thought of as cold and red as warm, but here, the dots fall on the other end of the temperature spectrum.

We understand the point the color information may overlap temperature change indicated by vertical axis. However, we would like to keep the red-blue colour scheme because AMOC anomaly is often illustrated by red (increased AMOC streamfunction -> strengthening) to blue (reduced AMOC streamfunction -> weakening).

Figure 8: Colour bar title and unit are missing.

L. 622: Please specify what you mean by LGM sea ice edges. In the text, you mention winter sea ice edge, and it looks like this is what you are showing, but this should also be clear in the caption.

We have added colour bar title and units including figure captions. We have revised the caption to clarify winter sea ice edges are displayed in Figure 8.

Figure 10 caption: Add description about shading (see L. 591-592)

We have added the information on gray shading in other figures.

Supplementary Materials: Supplementary figure numbering is not consistent with the numbering and order in the main text

This can be a result of we have updated the Supplementary Materials during the first revision of the manuscript in September 2024. We have carefully checked the manuscript and made consistent with Figures in Supplementary Materials.

Minor comments

L. 147:”feedback” a feedbacks

L. 183: (Fig. 2a, black lines) a (Fig. 2a, upper panel, red and black lines,

respectively)

L. 226: “bipolar seesaw models” a bipolar seesaw model

L. 259: Switch the order of LOVECLIM and iTRACE to follow the same order as in the figure.

L. 264: “is not simulated” a is not fully simulated

L. 265: “also” after MPI-ESM is not needed

L. 280, L. 290 and L. 292: Switch order of MIROC and LOVECLIM in text to follow order in figure.

L. 446: “simulates” to simulate (talking about the five models)

L. 473: “from ECS value” to from the ECS value

L. 476 “separations” to separation

L. 511 “increase in AABW” to increase in AABW formation

L. 628: “850hPa winds” to 850 hPa eastward winds

L. 577, Table 1: Typo in MPI-ESM and iLOVECLIM Ice sheets “Ice-6G_c” to Ice-6g_C

We have reflected these changes.

L. 394, L. 408: Fig. S2 to Fig. S3

L. 453: Fig. S3 to Fig. S4

L. 475 Fig. S4 to Fig. S2

We have changed Supplemental Figure numbers consistently.

[references]

Menviel, L. C., Spence, P., Kiss, A. E., Chamberlain, M. A., Hayashida, H., England, M. H., and Waugh, D.: Enhanced Southern Ocean CO2 outgassing as a result of stronger and poleward shifted southern hemispheric westerlies, Biogeosciences, 20, 4413–4431, https://doi.org/10.5194/bg-20-4413-2023, 2023.

Abe-Ouchi, A., Saito, F., Kageyama, M., Braconnot, P., Harrison, S. P., Lambeck, K., Otto-Bliesner, B. L., Peltier, W. R., Tarasov, L., Peterschmitt, J.-Y., and

Takahashi, K.: Ice-sheet configura-tion in the CMIP5/PMIP3 Last Glacial Maximum experiments, Geosci. Model Dev., 8, 3621-3637, doi:10.5194/gmd-8-3621-2015, 2015.