AR: We thank both reviewers for their important comments to our manuscript. In our revised manuscript, we addressed the issues raised as indicated in our response. The reviewer comments are shown in black and our response in red.

In addition to the points raised by the reviewers, we have made a minor change to the abstract. Instead of referring to the "mid-Pliocene", we now refer to the "mid-Pliocene warm period (mPWP)" in the abstract to improve the consistency with the rest of our manuscript and with other PlioMIP2 studies.

Review Anonymous Referee #1:

Weiffenbach et al. assess the state of the Southern Ocean in 15 simulations of the mid-Pliocene performed as part of the Pliocene Model Intercomparison Project (PlioMIP2). 7 simulations with the full mid-Pliocene forcing (Eoi400) are also compared with experiments in which only the mid-Pliocene CO2 gas concentration was implemented (E400). This allows to separate the effects of the higher greenhouse gas concentration from the mid-Pliocene changes in orography. It is found that both the greenhouse gases and Pliocene orography contribute to the high southern latitudes climate change: i.e. warming, higher precipitation, stratification… The manuscript is well written, and the results are interesting. As such I recommend publication in Climate of the Past after the comments below are taken into account.

AR: We thank the reviewer for their positive evaluation and constructive comments and suggestions. Please find our point-specific responses below.

1. **Experimental design:** At the moment the manuscript is targeted for readers who are familiar with the PlioMIP2 experiments. There is no detail in the methods regarding the experimental design used in these experiments. Readers are simply referred to Haywood et al. 2016. Please add some information in the methods regarding the experimental design used in the PlioMIP2 experiments so that a reader of this manuscript does not have to read Haywood et al., to fully understand the study.

AR: The introduction (L39-L51) describes the mid-Pliocene boundary conditions and the experiments, but we agree that a brief description of the PlioMIP2 experimental design explanation could be added in the methods section for those unfamiliar with the ensemble. We have added a subsection "PlioMIP2 experimental design" to the methods section in which we give this overview and refer the reader to Haywood et al. (2016) for the full details.

2. **Discussion:** The discussion needs to be expanded and restructured.

While the discussion focuses on the relevance for future climate, there is a real paucity of references to other studies in the Discussion. There is no discussion of previous multi-model studies of the mid-Pliocene, previous multi-model studies focusing on high southern latitude climate, or previous studies discussing changes in SO stratification and the relation with the abyssal cell. For example, the paragraph starting L. 301 could be expanded to compare the results of the study with previous results.

This is consistent with \citet{de_lavergne_cessation_2014}, who show weakening of Southern Ocean deep convection due to surface freshening under the RCP8.5 scenario, as well as recent work by \citet{chen_reduced_2023} demonstrating that freshening of the Southern Ocean surface due to Antarctic meltwater induces weakening of deep convection.

AR: We agree that it would be good to have more multi-model studies to compare our results to, however, for the mid-Pliocene there are no previous multi-model studies focusing on the Southern Ocean or Southern high-latitude climate other than those we already refer to. We discuss CMIP5 and CMIP6 studies in L302-L304 as well as in section 2.4.3. As suggested by the other reviewer, we will add a reference to Purich et al. (2021) at L304 to consider the effect of (future) SO conditions on AABW formation in CMIP6. We have added the following sentence at L304: "CMIP6 models also show future Southern Ocean surface warming and sea ice decrease, as well as mixed layer shoaling that suggests slowdown of AABW formation (Purich et al. 2021)."

Regarding the relation between the SO and abyssal cell, we discuss this issue in section 2.4.3 in relation to biases in multi-model present day studies (CMIP5/6). We have expanded the paragraph starting at L301 further to include de Lavergne et al. (2014) who found reduced deep convection in the Southern Ocean due to surface freshening in CMIP5, as well Chen et al. (2023) who show that Antarctic meltwater can reduce deep convection and the volume of bottom water in CMIP6.

The model-data comparison is limited and only shown in the SI. I understand there is limited data availability for this time period, but have the authors performed a literature search on the topic? Even though the overlap is small, the authors could look at Grant et al., 2023 (https://egusphere.copernicus.org/preprints/2023/egusphere-2023-108/) for marine sediment cores in the SW Pacific.

AR: McClymont et al. 2020 is currently the only SST dataset of the KM5c time slice (3.205 Ma, 20 kyr time slice), which is the time slice that PlioMIP2 is centered around. This time slice was chosen to avoid any effects by orbital forcing (Haywood et al., 2016), which can impact seasonal and regional distribution of incoming solar radiation. The effect of orbital forcing is especially important in the high latitudes, and therefore we do not want to consider data that falls outside of this time slice. As the dataset from Grant et al. (2023) averages over several interglacials between ~3-3.3 Ma, we have not included it in our model-data comparison.

Finally, the order of the sub-sections in the Discussion should be thought through so that there is a better flow: i.e. The model-data comparison could come first, then discussion of results in the context of previous studies, relevance for future studies and caveats associated with SO biases.

AR: We agree with this suggestion and have changed the order of the discussion to the suggested order.

3. Minor points:

- L. 200: high-latitude North Atlantic?

  AR: This has been corrected.

- L. 219: Relationship between changes in abyssal cell strength and OHT is interesting and consistent with previous studies (e.g. Menviel et al., 2015, EPSL).

  AR: Thank you for pointing this out. We have added this comment and reference to the manuscript.

- L. 301-302 There is not enough information to understand the relevance of the sentence (i.e. which scenarios? When?).

  AR: We have added the missing information to this sentence and rephrased it to: "CMIP5 studies investigating RCP4.5 and RCP8.5 scenarios show a decline in Antarctic Bottom Water formation due to fresher and warmer Southern Ocean surface conditions, and consequentially a slowdown of deep ocean circulation by the end of the 21st century (Meijers et al., 2014; Ito et al., 2015)."

- L. 310-311: I don't understand this sentence

  AR: To clarify, we have rephrased L310-311 to: "Since these geographical changes are not relevant for the future, it is uncertain whether the strength of decline in abyssal cell circulation, or the lack of decline exhibited in some models, resembles a plausible future scenario."

- L. 319: "sea-ice cover"

  AR: This has been corrected.

- Figure 7: Is the cooling within the annual sea-ice zone due to the increased stratification?It does correspond to the area of large negative SSS anomalies. Is there reduced deep ocean convection in the Ross Sea sector in most models?

  AR: Figure 7 does not show cooling within the annual sea-ice zone as it shows the SST anomalies minus the global mean SST anomaly. This means that any areas in blue indicate areas that warm less than average, but do not necessarily cool. The below-average warming is most likely related to the persisting sea ice cover reducing air-sea fluxes and thus inhibiting warming of the surface.

  The area that warms less than average indeed corresponds to the area with large negative SSS anomalies, as well the area with a large stratification increase. It is not possible to make any strong claims about the causal relationship between the sea ice cover and the stratification because they influence each other. We think it is most likely that the increased stratification is caused by the large negative SSS anomalies, but it is not completely clear where these large negative SSS anomalies come from. This last point is further addressed in our response to the next comment.

- Figure 8; The fit between precipitation and SSS anomalies is not particularly obvious. I am not sure P-E is a better fit. Should you look into changes in sea-ice melt/freezing, Antarctic runoff and eventually changes in ocean currents to understand the changes in SSS?

AR: Indeed, the spatial fit is not always obvious because there are many other factors at play that could influence the SSS other than precipitation/P-E. It is likely that changes in sea-ice melt/freezing, Antarctic runoff and ocean currents also significantly influence the SSS and are possibly even more important than the surface freshwater flux in areas where the fit between the SSS and precipitation is not obvious. Unfortunately, the data for these variables is not available for most models. We think an in-depth analysis of the freshwater balance would be very interesting, but this would require a detailed investigation into one or a few models. Such an analysis would be beyond the scope of this study in our opinion.

- Figure S11: Historical and not pre-industrial simulations should be compared with present-day observations.

  AR: It is true that the present-day (1979-1999) sea ice observations used in Figure S11 would preferably be compared with historical rather than pre-industrial simulations. Unfortunately, these simulations are not part of the PlioMIP2 protocol and were thus not performed. As there are only sparse observations of the pre-industrial sea-ice extent, we use the present-day observations to show that the MMM sea-ice extent of the pre-industrial is spatially comparable to the historical observations, and that the model ensemble therefore likely shows a realistic sea-ice extent for the pre-industrial. To emphasize this point, we have changed L357-L361 to:

  "Supplementary Figure S11 shows that the MMM sea-ice cover and extent in the PI E^{280} simulations is spatially similar to historical observations (1979-1999 NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Meier, W.N. et al. (2021)), with a MMM E^{280} sea-ice area of 10.4 million km^2 and historical observed sea-ice area of 9.5 million km^2. While the historical observed sea-ice area should ideally be compared to historical simulations, Figure S11 gives some confidence that the simulated MMM E^{280} sea ice cover is realistic for the PI. However, the sea-ice area variation among models is 5.2-15.6 million km^2, which is substantially larger than the discrepancy between the MMM E^{280} and historical observations."

Review Chris Brierley:

Overall, I think this is good paper and should be published in climates of the past with little modification. The topic is appropriate for the journal and quite important in general. I have some suggestions that I think would make the manuscript appeal to a wider audience, which are mainly technical.

AR: We thank Chris Brierley for his helpful comments and positive words. Please find our point-specific responses below.

L35: you refer to a concentration pathway as being an emission scenario. This is confusing, as there is an esmrcp85 variant, but not a esmrcp45.

AR: Yes, we have corrected this by removing "emission" from L35.

L97: justify why cosmos is treated differently

AR: COSMOS is treated differently because the curvilinear grid of COSMOS is not suitable for calculating the transports and using the regular 1x1 grid is more straightforward and introduces less error. We have changed L97 to: "Ocean freshwater and heat transport have been calculated on native model grids, except for COSMOS where computation is done on a regular interpolated 1x1 degree grid as the native curvilinear grid is unsuitable for this calculation."

L112: What is the magnitude of the errors introduced by using monthly mean temperatures and salinities to compute the potential density and therefore stratification index?

AR: As not all models are able to provide potential density fields, we use the standard TEOS-10 (International Thermodynamic Equation of Seawater – 2010, IOC et al. (2010)) equation of state to calculate the 100-year mean potential density from the 100-year mean fields of temperature and salinity. This method has already been used for other multi-model studies such as Bourgeois et al. (2022) and Muilwijk et al. (2023). Due to averaging effects as well as not all models implementing the same approximation of the equation of state, there may be a small difference with the model-calculated potential density. The magnitude of this difference would vary between models and it is therefore difficult to provide an estimate of its magnitude without having all the model-calculated potential density fields available. We do not believe, however, that the magnitude of error is large enough to significantly impact our results. As the alternative is to not use models that don't have the potential density field available, we consider the possible error to be acceptable.

Table 1: The first 4 models in this table are all variants of CCSM4. You should at least acknowledge this unbalanced ensemble design, and preferably comment on whether this could skew your findings (I suspect not).

AR: There are indeed 5 CESM models included in the PlioMIP2 ensemble. With their respective ECSs of 4.1°C and 5.3°C, CESM1.2 and CESM2 both have a significantly higher ECS than the reported 3.2°C ECS of the CCSM4 models (Haywood et al., 2020), which suggests that the models do respond to forcing in considerably different ways. In addition, the results presented in several published PlioMIP2 studies show that the CESM models all present significantly different results in both the atmosphere and ocean, as can be seen in, for example, Haywood et al. (2020) and Zhang et al. (2021). This is likely related to individual modelling groups using different settings and parameters in their model. For instance, Table 1 in Zhang et al. (2021) lists the vertical mixing parametrization employed by the different modelling groups, which shows that all CCSM4 models use a different vertical mixing parametrization. Therefore, in our opinion, the results would not be skewed by the unbalanced ensemble.

We agree it is important to acknowledge the imbalanced in the ensemble and have added the following sentence to L89: "It is important to note that 5 out of the 15 models in the ensemble are CESM models, of which 3 are CCSM4 models, creating an imbalance in the ensemble. However, due to differences in model versions as well as model settings of individual groups, earlier PlioMIP2 studies show that the climate response varies significantly among the CESM members (e.g. Haywood et al., 2020, Zhang et al., 2021) and we therefore consider it unlikely that this imbalance skews our results."

Fig 1: At no point do you say in this figure caption what temporal averaging is being used. I presume this is annual mean, but showing a winter and/or summer sea ice edge is more conventional.

AR: The temporal averaging is the 100-year annual mean, we have altered the caption to include this information. While a winter and/or summer sea ice edge may be more conventional, we have chosen to only show an annual mean sea ice edge in our analysis because all other variables we present are annual mean variables.

P9: I am very surprised that you are able to discuss AMOC in PlioMIP2 without citing Zhang et al's 2021 paper called "Mid-Pliocene Atlantic Meridional Overturning Circulation simulated in PlioMIP2". This work predates the Weiffenbach et al (2023) paper that is cited on the topic by 2 years! More disturbingly, 13 of the 16 authors of the present manuscript were also authors of the 2021 work. So it seems implausible that you (collectively) are unaware of it. Most importantly, if you (collectively) now have reservations about your (collectively) earlier work then you should be airing them publicly so that the whole community is aware of them.

AR: We did not cite Zhang et al. (2021) due to an unfortunate but not intended mistake. We apologize for any confusion this may have caused. It is indeed important work that should be cited when referring to the stronger AMOC in PlioMIP2. We have referred to Zhang et al. (2021) in addition to Weiffenbach et al. (2023) in L310 and changed L180-L181 to: "The MMM AMOC strength increases by 3.1 Sv (+16%) in the mPWP simulations with respect to the PI, which agrees with Zhang et al. (2021) who show a consistently stronger mid-Pliocene AMOC in PlioMIP2."

Fig captions. Please do not start figure captions with uncommon acronyms. When I read any paper, I will first read the abstract and then look at the figures: only afterwards working my way through the main text, if I've found it interesting. I'm sure that I'm not the only person who approaches papers in a similar fashion. I recommend that you try to make your figure captions somewhat standalone. The current reliance on only PlioMIP2 terminology is pervasive. I was also confused by SMM, as even expanding it to "special model mean" does not help explain what the figure is.

AR: Thank you for this suggestion. We have reviewed all figure captions and clarified any acronyms in the captions. We think using the PlioMIP2 terminology for the experiments and the SMM in the captions is unavoidable. However, to avoid confusion, we now refer to the Methods section when defining the acronym SMM. We have also added a sentence at the end of each caption that clarifies the experiment names E^{280}, E^{400} and Eoi^{400}.

Fig 10: Consider adding the non-special models (boring models?) to this figure. They would not need to identifiable.

AR: We have added the non-special models to this figure as light grey circles to provide a more complete figure and clarified which models are included in the least-squares fits.

We do want to note that the non-special models are not boring or different from the special models in any climate relevant sense. The only difference from the special models is that they don't have the output data for the E^{400} experiment, which has a mid-Pliocene CO_2 concentration with a pre-industrial orography and other boundary conditions.

L290: Be wary of using "CMIP projections" when I think you mean the scenarioMIP projections. The ice sheet model intercomparison project is part of CMIP.

AR: We have changed "CMIP projections" into "CMIP ESM projections" to avoid any confusion. We have also changed "CMIP simulations" into "CMIP ESM simulations" in the next line for the same reason.

L294: I had forgotten what ~24m relates to, please remind the reader by providing more context.

AR: Thank you for this comment, as it also points out an error in our argument. The ~24 m sea level equivalent decrease in ice volume refers to the sea level increase in the mPWP after reconstructing both the Greenland and Antarctic ice sheet volume. We have corrected this error by changing ~24 m into ~21 m, which is the reconstructed sea level equivalent decrease in Antarctic ice volume in the mPWP (Dowsett et al., 2010), and clarified L294 by changing it into:

"It should be taken into account, however, that Antarctic Ice Sheet projections do not show a ~21 m sea level equivalent decrease in ice volume, which has been reconstructed for the mPWP Antarctic ice sheet (Dowsett et al., 2010)."

L302: You mention CMIP5 studies here. Has no-one published anything on the topic relating to CMIP6?

AR: Yes, we agree that information on CMIP6 is lacking here. Many CMIP6 Southern Ocean focus on model biases in historical runs, which we discuss in section 2.4.3. We were able to find only one CMIP6 study by Purich et al. (2020) that explicitly discusses future Southern Ocean surface conditions and AABW formation. We have added this reference in the following sentence after L302-304: "CMIP6 models also show future Southern Ocean surface warming and sea ice decrease, as well as mixed layer shoaling that suggests slowdown of AABW formation (Purich et al. 2021)."

L321: You are missing a closing bracket

AR: This has been corrected

References:

Bourgeois, T., Goris, N., Schwinger, J., & Tjiputra, J. F. (2022). Stratification constrains future heat and carbon uptake in the Southern Ocean between 30°S and 55°S. *Nature Communications*, *13*(1), 340. https://doi.org/10.1038/s41467-022-27979-5

Chen, J.-J., Swart, N. C., Beadling, R., Cheng, X., Hattermann, T., Jüling, A., Li, Q., Marshall, J., Martin, T., Muilwijk, M., Pauling, A. G., Purich, A., Smith, I. J., & Thomas, M. (2023). Reduced Deep Convection and Bottom Water Formation Due To Antarctic Meltwater in a Multi-Model Ensemble. *Geophysical Research Letters*, *50*(24), e2023GL106492. https://doi.org/10.1029/2023GL106492

Dowsett, H., Robinson, M., Haywood, A. M., Salzmann, U., Hill, D., Sohl, L. E., Chandler, M., Williams, M., Foley, K., & Stoll, D. K. (2010). The PRISM3D paleoenvironmental reconstruction. *Stratigraphy*, *7*(2–3), 17.

Grant, G. R., Williams, J. H. T., Naeher, S., Seki, O., McClymont, E. L., Patterson, M. O., Haywood, A. M., Behrens, E., Yamamoto, M., & Johnson, K. (2023). Amplified surface warming in the south-west Pacific during the mid-Pliocene (3.3–3.0 Ma) and future implications. *Climate of the Past*, *19*(7), 1359–1381. https://doi.org/10.5194/cp-19-1359-2023

Haywood, A., Dowsett, H., Dolan, A., Rowley, D., Abe-Ouchi, A., Otto-Bliesner, B., Chandler, M., Hunter, S., Lunt, D., Pound, M., & Salzmann, U. (2016). The Pliocene Model Intercomparison Project (PlioMIP) Phase 2: Scientific objectives and experimental design. *Climate of the Past*, *12*(3), 663–675. https://doi.org/10.5194/cp-12-663-2016

Haywood, A. M., Tindall, J. C., Dowsett, H. J., Dolan, A. M., Foley, K. M., Hunter, S. J., Hill, D. J., Chan, W.-L., Abe-Ouchi, A., Stepanek, C., Lohmann, G., Chandan, D., Peltier, W. R., Tan, N., Contoux, C., Ramstein, G., Li, X., Zhang, Z., Guo, C., … Lunt, D. J. (2020). The Pliocene Model Intercomparison Project Phase 2: Large-scale climate features and climate sensitivity. *Climate of the Past*, *16*(6), 2095–2123. https://doi.org/10.5194/cp-16-2095-2020

de Lavergne, C., Palter, J. B., Galbraith, E. D., Bernardello, R., & Marinov, I. (2014). Cessation of deep convection in the open Southern Ocean under anthropogenic climate change. *Nature Climate Change*, *4*(4), Article 4. https://doi.org/10.1038/nclimate2132

McClymont, E. L., Ford, H. L., Ho, S. L., Tindall, J. C., Haywood, A. M., Alonso-Garcia, M., Bailey, I., Berke, M. A., Littler, K., Patterson, M. O., Petrick, B., Peterse, F., Ravelo, A. C., Risebrobakken, B., De Schepper, S., Swann, G. E. A., Thirumalai, K., Tierney, J. E., van der Weijst, C., … Zhang, Z. (2020). Lessons from a high-CO2 world: An ocean view from ~ 3 million years ago. *Climate of the Past*, *16*(4), 1599–1615. https://doi.org/10.5194/cp-16-1599-2020

Muilwijk, M., Nummelin, A., Heuzé, C., Polyakov, I. V., Zanowski, H., & Smedsrud, L. H. (2023). Divergence in Climate Model Projections of Future Arctic Atlantification. *Journal of Climate*, *36*(6), 1727–1748. https://doi.org/10.1175/JCLI-D-22-0349.1

Purich, A., & England, M. H. (2021). Historical and Future Projected Warming of Antarctic Shelf Bottom Water in CMIP6 Models. *Geophysical Research Letters*, *48*(10), e2021GL092752. https://doi.org/10.1029/2021GL092752

Zhang, Z., Li, X., Guo, C., Otterå, O. H., Nisancioglu, K. H., Tan, N., Contoux, C., Ramstein, G., Feng, R., Otto-Bliesner, B. L., Brady, E., Chandan, D., Peltier, W. R., Baatsen, M. L. J., von der Heydt, A. S., Weiffenbach, J. E., Stepanek, C., Lohmann, G., Zhang, Q., … Abe-Ouchi, A. (2021). Mid-Pliocene Atlantic Meridional Overturning Circulation simulated in PlioMIP2. *Climate of the Past*, *17*(1), 529–543. https://doi.org/10.5194/cp-17-529-2021