Author reply to Louise Sime

We thank Louise Sime for the very helpful comments and suggestions.

The original comments are displayed in red, our replies in black.

Manuscript summary:
This study introduces a novel comprehensive coupled atmosphere-ocean-vegetation-ice sheet-solid earth model, MPI-ESM/mPISM/VILMA, designed to simulate the last deglaciation with high realism. This model is unique because it includes interactive Earth system components such as ice sheets, icebergs, solid earth, and dynamic river directions, allowing for a comprehensive exploration of deglacial processes. The findings highlight that the model can reproduce abrupt millennial-scale climate events and that the timing of these events is influenced by initial conditions and model parameters rather than solely by external forcing. Additionally, the model reveals that changes in Arctic sea-ice export and freshwater dynamics significantly affect the Atlantic Meridional Overturning Circulation (AMOC) and North Atlantic climate.
I find this study really quite exciting. It provides a breakthrough in understanding how different components of the Earth's system interact over long timescales, particularly during the complex process of deglaciation. The ability of the new model to simulate abrupt climate events and reveal unexpected dynamics, including the influence of freshwater dynamics and ice sheet surges on ocean circulation, opens up new avenues for exploring past climate changes and understanding past climate tipping behaviours.
The analysis is mostly well-organised and constructed, and the paper is well-written. It effectively conveys the findings, highlights the novelty of the research, and provides a clear and concise overview of the study's objectives, methods, results, and implications. With the caveats below about it being on the long side, and whether the authors could extend parts of the analysis and split it into two, it is clearly suitable for publication in CP.

Major points:

This is not necessarily a criticism, but I find there is really a lot of material/work in this manuscript. It might be more digestible to most readers if it were split in two, to allow more focus on some aspects. Perhaps with the material on abrupt changes placed into a second manuscript?

We believe that it is an important aspect of this paper to demonstrate that the model's ability to reproduce a large variety of dynamic processes, on glacial-interglacial as well as centennial/millennial scales. Therefore we think it is important to represent them in one paper

The simulations are generally run from 26ka to 1850, allowing a climatological mean PI for each simulation to be specified as the last 1,000 years or so of the run. However, we never see what the PI simulations look like. The manuscript would benefit from an appendix to show how the PI states differ for each version of the model, and some indication of how far they deviate from the observed PI, and a comment on this to be added to 2.2 and 3.1.

Actually, this criticism applies only for SAT. For all other quantities (sea ice, ice sheets and AMOC at 30°N) we show absolute values of D2.1 and the ensemble (median, min and max) as well as observational products in figs. 3, 4 and 7. We will add a Figure to the Appendix that shows the deviations between modelled PI SAT and SAT from the ERA-20C reanalysis (see below). This

Figure also includes modelled and observation based sea-ice extent estimates. We will refer to the figure in the main text.

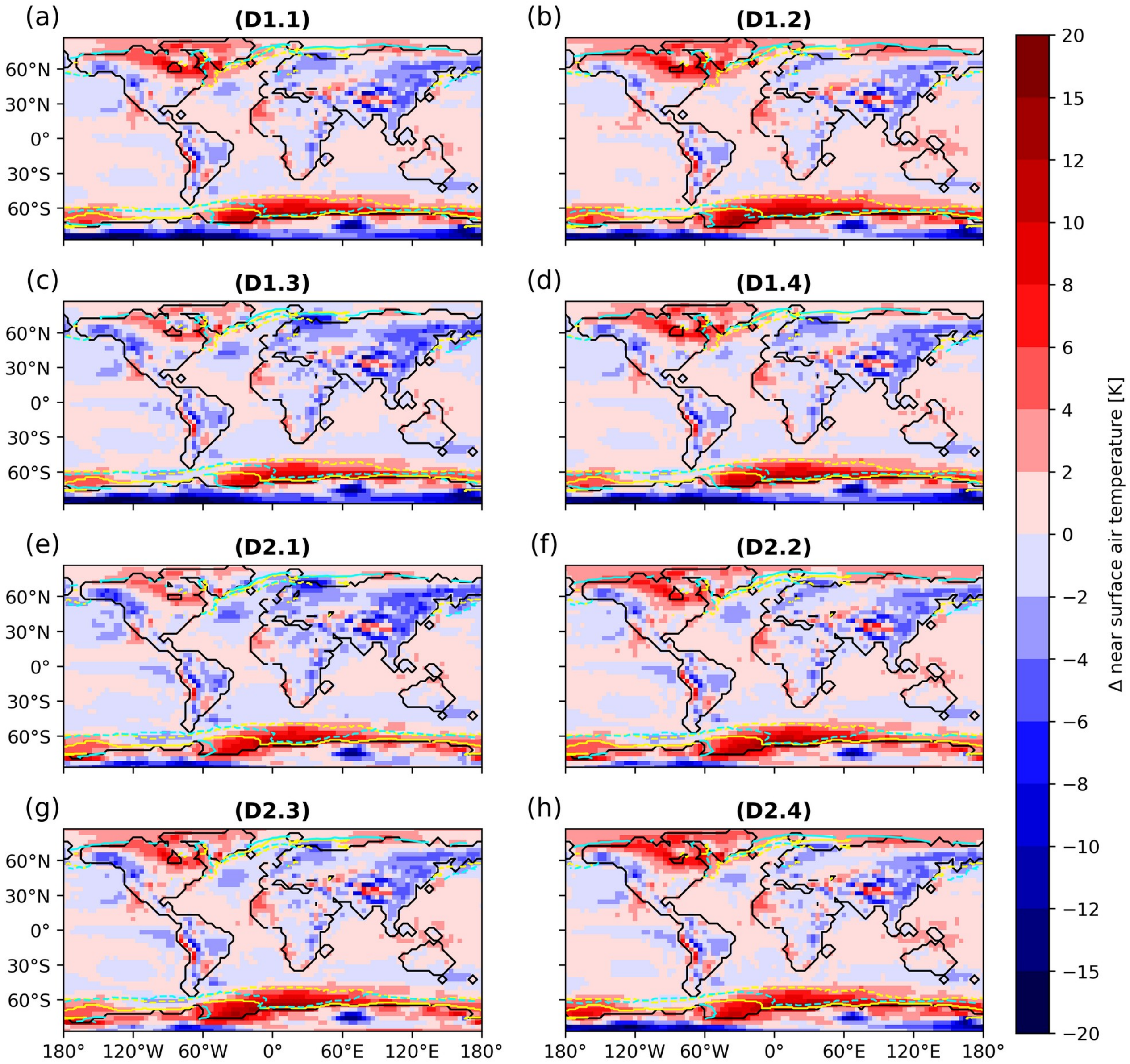

Fig.: Difference in annual mean near-surface air temperature between model and ERA-20C reanalysis (colours). Panels (a-h) show individual ensemble members. The isolines show sea-ice extent (> 0.15 sea-ice coverage in the long-term mean seasonal climatology) from the HadiSST dataset for summer (yellow solid) and winter (yellow dashed) and from the model for summer (cyan solid) and winter (cyan dashed). The land-sea mask is indicated by the solid black lines.

The description of the pre-26ka model spin-up seems to imply that the ocean is spun up for just 10 years, and then the first 1,000 years of each simulation are disregarded. This would give only 1,010 years to spin up the ocean. I think this may not be correct because the text also states that the simulations are initialized from pre-existing glacial simulations. However, this is not very clearly explained. Like the PI point above, spin-up is relatively important, so this should be carefully clarified in 2.1.

This is a misunderstanding. The asynchronous spin-up was run for 19,000 years for ice sheet and solid earth and for 1900 years for MPI-ESM (incl. ocean). We will add some text to make this more clear.

Some clarification/justification for modifications to ocean mixing/vertical diffusion would be useful, both in the context of expected glacial-interglacial ocean mixing changes (due to bathymetric and stratification effects on the dissipation of internal tides/waves). Alongside this, a strengthened discussion of how AMOC and mixing may be affecting the results from this model, given abrupt changes results are relatively strongly dependent on the strength of the AMOC & mixing i.e. discussion of would we expect other GCMs/ESMs to behave similarly (if they also had the same extra component coupled onto them, and were run for similar experiments), or are these unique to this model?

The background mixing is a rather poorly constrained parameter. Therefore we tested in ensemble 1 its effect on the simulation of LGM and the deglaciation. The effect on the AMOC was quite large, as discussed in the paper. The effect on the amplitude of the abrupt events was rather moderate. The size of the ice surges is determined by the ice sheet, only the strength of the reaction of the AMOC on these surges is affected by the ocean mixing and thus the AMOC. The most notable effect was the delay of the simulated deglaciation in case of reduced background mixing as a consequence of the colder polar climate and an early deglaciation in case of stronger mixing. There was, however, a strong negative effect of these changes on the ability of the model to reproduce the observed PI water mass age distribution in the North Pacific. Exp. D1.1 yielded the most realistic distribution in terms of PI radiocarbon (not shown and discussed in this paper, going to be a separate paper). Therefore we modified in ensemble 2 the background mixing only in the upper 1000 m to avoid the strong effects on age distribution. We will add a remark on this in the text.

The section on Abrupt Events is very sensibly laid out; I like the analysis. However, I find it slightly surprising that the focus is solely on simulated abrupt cold events, given there is considerable community interest in the possibility of abrupt warmings too, and some deglacial events are indeed abrupt warmings, rather than solely coolings. If the authors intend to retain the focus solely on abrupt coolings, then it would be helpful to have a sentence or two added to the Introduction and Section to better justify this focus. Otherwise, the focus could perhaps be broadened to also include abrupt warmings, which are also visible in the timeseries provided. See also the first main comment about splitting this manuscript in two.

To keep the paper at a reasonable length, we are focusing on events, where AMOC changes play a major role. By nature, these are mostly abrupt cold events and some abrupt warmings due to AMOC recovery. We will add a remark in the introduction.

Minor points:

The model name is a currently a bit of a mouthful, have the authors also considered giving the model a shorter name too MPI-ESM-extended or similar, perhaps?

The name is a bit long, but there are many model versions around e.g. the version used for the deglacial runs with prescribed ice sheets (Kapsch et al. 2022), a similar version but with interactive methane Kleinen et al. 2020), a deglacial version with carbon cycle and without interactive ice sheets (Extier et al. 2022), and a version with interactive icebergs (Erokhina and Mikolajewicz

2024) plus some older versions with interactive ice sheets (e.g. Ziemen et al. 2019). We believe it is important to make clear, which model version we are using to avoid confusion. The original name is used already in the ongoing cmorization process. Using different model names in the paper and the corresponding data publication is undesirable. Therefore we stick to the name chosen.

L18, missing refs, for very different rates of changes

We will move the citation of Fairbanks (1989) as reference also for the variations to the end of the sentence and add the citation of Lambeck et al (2014) who also estimated time evolution of sea-level rate changes (their figure 4D).

L19-22, greater and lesser 'volumes' rather than changes?

We replaced 'changes' with 'values'.

L34 abrupt 'AMOC' changes, or what events?

L34, 'the quantification of' rather than 'the exact changes'

We will rephrase the sentence in L34 as follows:

These data indicate the existence of abrupt climate events which shaped the sediment record, but even a qualitative estimate of changes in the characteristics of the AMOC remain poorly constrained (e.g. sign of AMOC strength variations....)

L54, and after, clarify what 'CMIP-style' means. Either CMIP models, or perhaps models that use CMIP-atmosphere (AMIP?) models, or perhaps that are run at common CMIP atmosphere-ocean resolutions? Either way, replace 'CMIP-style' descriptions with something more meaningful.

We will not use the term "CMIP-style" anymore, but use "comprehensive climate model" instead plus a more accurate description.

L75, split this into two sentences.

Will be done.

L113, clarify what is meant here by radial directions – by depth, but not by lat or long?

We will change this to: "VILMA is employed in its 1D configuration, which assumes that the viscosity structure of the Earth varies only with depth but is horizontally homogeneous."

Section 2.2 and 3.1, please see above first two main points.

As mentioned above, we will add a plot of the mismatch of PI SAT for all runs including sea ice extents. Ensemble values for sea ice extent and ice sheets are already shown in the existing figures 3 and 7.

Table 1 headings, spell out the headings better. Information about spin-up is not clear. Better to replace the last column 'Parent run' with a much clearer verbal description. Exp file names can be omitted.

We will replace the term 'parent-run' by 'spin-up' and add some more explanations to the table caption. As the individual parameters of the spin-up run are matching the experiments, it would repeat the information already given in column 2. Much of the confusion seems to come from the misunderstanding about the spin-up. We introduced the experiment names of the spin-up more clearly. Experiment names are important, as all members of ensemble 1 use the same spin-up simulation.

~L125, consider to add a short subsection or para in 2.1 which describes the water tracer/dye methodology, including precise conditions for tracing/dying water, and how tracers/dye is reset (presumably without resetting on exposure, the ocean would eventually saturate?).

We will move the introduction of the dye tracer to the model description and include a figure of the source region in the Appendix. We will emphasise the resetting of the dye tracer to 0 outside of the source region.

L193, maximum?

Will be changed.

L200-202, better to replace this with an Appendix that more carefully shows what the PI states look like themselves (not just the LGM-PI anomalies).

Fig. 3 already shows absolute values of sea ice extent. Both PI values of the model ensemble (Fig. 3a-d) and ‘observed’ sea ice extents for summer and winter are shown (Fig. 3e+f).

In addition, we will add a new figure to the appendix showing PI SAT biases and sea ice extent for each ensemble member.

Figure 3, obscures more than it shows. It might be more instructive to see the LGM-PI anoms subtracted from the equivalent Anna et al and Osman et al anoms.

We do not share this point of view. For entirely observation-based estimates showing model obs differences might be a good approach. However, the Annan et al. and Osman et al. estimates are strongly influenced by their a priory choice of the models from the PMIP ensemble. So the plots would greatly differ depending on our choice of reference. Therefore we prefer to show the original figures.

Figure 5, I really like these dye/tracer sections.

Thank you

Figure 5 and Figure 6, and thereabouts in text, please ensure there is sufficient information in the text to reassure the reader that these results are not due to spin-up issues. See also main points, and minor point aboves about better clarification on spin-up procedures and ocean run durations.

Prior to the LGM section time slice, the AOGCM was run either 3000 (D1.2 and D1.3) and 4900 years (all other runs) without parameter changes starting from a glacial state. For the LGM climatologies we show the average over model years 3001 to 8000 since the begin of the synchronous simulations. So drift should not be a real problem. PI had about 25,000 synchronous years (+ 1900 asynchronous). In a transient simulation with transient forcing, no full equilibrium can be achieved, see e.g. fig 10b. We do not see any obvious drift in this figure prior to the LGM.

Page 6, and generally dye/tracer results. Please clarify how dye/tracer is removed. Is it reset to 0 on exposure to the atmosphere? Or something else?

We will mention that the dye tracer is set to 0 in the surface ocean outside the source region which guarantees a potential saturation with time (see also reply to L125).

L295, it is interesting that these coupled model simulations do not help much with this Antarctic ice sheet extent problem, possibly it is worth highlighting that this results supports the idea that the problem is in the representation of ice sheet physics in PISM rather than in the forcing/coupling?

We do not believe that this supports the hypothesis that the representation of ice-sheet physics in PISM is insufficient. This is corroborated by the fact that previous studies using PISM for the Antarctic ice sheet over similarly long time scales did manage to reproduce the advance and retreat pattern of the Antarctic ice sheet (e.g. Albrecht et al. 2020, Albrecht et al. 2024). Rather, we think that the parameter space in which our model setup can successfully simulate the advance and retreat pattern of the Antarctic ice sheet could be smaller. We will add a sentence reflecting this to the revised manuscript.

L310-311, rewrite this sentence – it is very hard to understand.

We will reformulate this sentence to: "This variability is not evident in the proxy based products due to the lower temporal resolution of the reconstructions, which is a consequence of the methodological design and the quality of the underlying proxy data."

L375, missing punctuation/sentence issue.

Will be changed.

Table 2, this table would be easier to digest if the rows were shaded to reflect whether the simulated even occurs earlier, later, or (within the uncertainties) at commensurate with the proxy evidence for a similar event.

Thanks for this suggestion. As the timing of the opening of the straits is not consistently earlier or later in individual simulations, a shading reduces the readability of the table. Hence, we decided that we would like to keep it as it is.

L529, salinity twice

Will be changed.

L531, 'varies significantly' rather than the significantly varies

Will be changed.

L581, see first comment on model name

We will keep the model name, see our reply on that above.

L602, remove 'also'?

Will be removed.

L607-610, there are some odd clause orders in here. Check ordering for English, and improve the sentences.

We will rework the mentioned sentences and improve their readability.

L632, either clarify what is meant by unexpected, or possibly rewrite this sentence to focus on the successfully model simulation of hereto uncaptured processes? (river rerouting, arctic freshwater sign changes, strait flow impacts, and other ice-sheet change related climate-land surface-ocean related processes.).

Good suggestion. We will do this.

References not already cited in the paper:

Albrecht, T., Bagge, M., and Klemann, V. (2024). Feedback mechanisms controlling Antarctic glacial-cycle dynamics simulated with a coupled ice sheet–solid Earth model, *The Cryosphere*, 18, 4233–4255. doi:/10.5194/tc-18-4233-2024

Extier, T., Six, K., Liu, B., Paulsen, H. & Ilyina, T. (2022). Local oceanic CO2 outgassing triggered by terrestrial carbon fluxes during deglacial flooding. *Climate of the Past*, *18*, 273-292. doi:10.5194/cp-18-273-2022

Kleinen, T., Mikolajewicz, U. & Brovkin, V. (2020). Terrestrial methane emissions from the Last Glacial Maximum to the preindustrial period. *Climate of the Past*, *16*, 575-595. doi:10.5194/cp-16-575-2020