Dear Dr. Sutter,

We have addressed the suggestions from your editorial review and respond to each one separately below. Thank you for the care you have taken in handling our manuscript. We note that we have also corrected Figure B1, which was previously inadvertently missing changes to the bed at Lake Vostok, and added a sentence to the text at L175 (in tracked-changes version) describing this change. Please also note that we had to compress the pdfs to fit within the 50 MB limit for upload, so some figures may be slightly lower quality than in previous iterations. This will be remedied when we upload the high-resolution figures for production.

Sincerely,

Trevor Hillebrand

L10: is this change in SLE symmetric with respect to variations in the melt parameter? I.e. linear across the range?

The impact of using the 5th and 95th percentile values of the melt parameter gamma0 is roughly symmetric in terms of sea-level contribution with respect to the baseline simulations using the median value of gamma0 (see figure 7). Testing whether the response is linear within that range would require more simulations, so we make no claims about the behavior between these endmembers and the baseline. We have made no changes here.

L178: I do not follow this reasoning. Do you mean to say that the Bedmachine boundary conditions are more conducive to MISI than Bedmap2? Is an important pinning point absent in Bedmachine? Please elaborate.

The BedMachine ice surface corresponds to 2015, while BedMap2 corresponds to the early 2000s. Thus, there could be a stronger MISI signal already present in the BedMachine geometry, which could lead to strong grounding-line retreat even if our model and forcing were perfect. Given that there are large uncertainties in both the model and the forcing, these likely exacerbate any existing pre-disposition for retreat. We have rephrased to: "As MISI could already be underway in the ASE \citep[e.g.,][]{joughin2014marine}, the c. 2015 ice geometry from BedMachine v2 may already contain a signal of MISI-style retreat in the grounding-line position and/or thickness field that is not present in the Bedmap2 data, and uncertainty in both the model and the forcing could exacerbate this issue."

L202: I would suggest shifting this sentence to l195 so that it is directly connected to the discussion of melting of grounded marine termini.

We have moved this sentence.

L212 Is there a reason for the specific cut-of date 2775 instead of e.g. 3000 or 2800?

We ran the control simulation until it crossed the 1m sea-level contribution mark. While 2800 or 3000 are maybe more pleasing numbers, there is really nothing special about those time horizons, so we used the physically meaningful sea-level value as a stopping criterion. It's a little awkward to include this justification here, since it mixes results and methods, but we have added "We used a threshold of 1 m sea-level contribution threshold to determine the length of the extended control simulation" to section 2.3.2.

L294 do you explain "prescribed ice shelf collapse" somewhere?

We have changed this to:

"Including ice-shelf collapse via hydrofracture using masks provided with ISMIP6 forcing..."

L295 you repeatedly use the word "significant". Please replace by e.g. "substantial" if you don't refer to a statistical analysis.

We have used "substantial", "strong", and "major" to replace many instances of "significant." We have simply removed other instances that were not necessary.

L321: what do you mean by "rapid" retreat.

Fig 6c shows less than 100km of grounding-line retreat at Thwaites between 2000 and 2300, and 200–300km or so between 2300 and 2400. We added a reference to this panel and added "(hundreds of kilometers per century)" to the text.

L385: do you explain the difference between temperature and enthalpy-based formulations somewhere? If not please do so briefly.

In S2.1 we have added "Temperature- and enthalpy-based solvers are available for thermal evolution. The enthalpy-based solver differs from the temperature-based solver by additionally accounting for liquid water content of ice at the pressure-melting point."

L407-408: you briefly mention potential effects of vertical resolution on ice flow linked to the temperature profile in ice shelves. What about grounded ice? There could be considerable influence on ice flow exerted by temperature profile depending on the vertical resolution and e.g. the delineation of "temperate" and "cold" ice affecting the rheology. Given the fact that you employ very few vertical layers this merits a statement.

We have moved the text regarding the impact of vertical resolution to the Discussion (Section 4.3) now that it is longer that just one sentence. We have added: "While we use only five vertical layers, borehole measurements of ice temperature in fast-flowing regions of Antarctica generally reveal smooth, low-gradient temperature profiles with depth (Engelhardt, 2004), which can be well approximated with a piecewise linear fit. Some temperature profiles in slow-flowing and divide-flow regions exhibit high curvature in the upper ice column (Talalay et al., 2020) that may be more difficult for our coarser near-surface layers to resolve, but such slow-flowing regions have little impact on sea-level projections."

L525 -> as this could explain a ...

Fixed

L527 -> Here, we have ...

Removed "here".

L532: I find it difficult to follow this argument: if the grounding line retreats 100s of km the effective pressure parameterization would still be applicable close to the new grounding line?

Yes, but the friction coefficient we calculated via optimization would have used an effective pressure that was very from reality. Therefore, the friction coefficient would contain large errors to compensate for the error in effective pressure, meaning that when the grounding line approaches these regions the basal traction will be very inaccurate even if the effective pressure comes within a reasonable range. We have changed this to: "Therefore, in simulations that predict tens to hundreds of kilometers of grounding-line retreat, effective pressure-dependent sliding laws may be strongly biased due to inaccurate effective pressure used to solve for the friction coefficient during model initialization."

L556: your discussion of the stress balance treatment is very interesting and it would be nice to hear a take on the potential causes of the differences. I am aware that this is difficult to disentangle it would be great to read your interpretation if only in a short paragraph.

This is indeed very difficult to disentangle because we cannot distinguish between the impacts of initial condition uncertainty and model uncertainty in the forward simulations. We have added Figure A1, showing the slightly faster flow speeds predicted by the MOLHO model for a given ice geometry (the initial condition, since this is the only time the geometries are exactly equivalent between MOLHO and BP runs), and we have added "The MOLHO solver predicts slightly higher flow speeds at the initial condition (Fig. A1)" in Section 3.6. In Section 4.4 we have added a few sentences explaining that the differing amounts of vertical shear in the MOLHO (and especially in the SSA) formulations likely lead to more sliding to achieve stress balance and thus lead to higher velocities.

L582: You mention here that MALI lacks an adaptive mesh refinement. So I assume the mesh is defined at the beginning of the simulation based on, as you mention in the methods, velocities and distance to grl. If the grounding line retreats sufficiently, as it does in many of your high emission runs, new grl positions would find themselves on coarse resolution sections of the model domain. It would be good to mention this already in the methods with a reference to the caveat section.

In S 2.2, we add: “As MALI lacks an adaptive mesh refinement capability, this cell spacing is held fixed in time. In simulations predicting major grounding-line retreat in West Antarctica, the grounding line will eventually retreat into lower-resolution regions, leading to less accurate solutions later in the simulations. We discuss this limitation further in Section 4.5.

L615: in the case of bed-products the only real choice I currently see is between Bedmap or Bedmachine but not between different iterations as the most recent usually provides the most accurate boundary conditions for present day initializations.

This is true, but it takes a huge time investment to configure and calibrate a model using new geometry data sets. Therefore, a group that has a well-calibrated configuration using an old geometry is not very likely to update with each new version of BedMachine. We have also shown very different behavior in our model when using a geometry product representing the early 2000s versus the mid-2010s, so there could be reasons to use e.g. Bedmap2 instead of Bedmap3 ice surface elevations if a specific time-period is desired. We think that our statement is general enough to stand without revision.

L659: I see a note on the absence of bed response in the simulations presented here on page 35 and 37 but not previously (apologies if I overlooked this). It should be noted prominently already in the methods section as this provides important context for the results. Absence of bed response usually leads to a stronger tendency for onset of MISI.

In Section 2.1 we have added: “We recently incorporated coupling to a sea-level model \citep{han2025improving}, but these changes were not included here and thus the bed topography is held fixed in time for all simulations in this study.”

L660 – 665 noting these discrepancies already in the introduction would provide additional motivation.

In the Introduction we have added: “Even when using the same model codebase, the choices made by different modeling groups contributing to ISMIP6 lead to fundamentally different ice-sheet responses to forcing.”

L664: While it is the same model, different initializations, boundary conditions (such as ghf) and parametrizations can easily lead to widely different results. I think this aspect is not surprising.

It might not be surprising to experienced ice-sheet modelers, but it lends support to the idea that individual modeling groups should be publishing in-depth analysis of their contributions to model intercomparison studies. We have not actually used the word “surprising”, opting for “striking” instead. But the main point of the sentence is that Thwaites is supposed to be very vulnerable to ocean forcing, so it is important to call out the fact that models can give *fundamentally* different results there even when applying strong forcing and using fairly common assumptions. We have made no changes.

L681: while I appreciate the comical relief I’d suggest a more scientific formulation, e.g. computationally prohibitive/unfeasible/etc.

We have reluctantly changed this to “prohibitively computationally expensive”.

L726: add a similar sentence in the methods so the reader understands the scope/limitations of the setup right away.

We have changed the last sentence of the Introduction to read, “While our investigation excludes many possible sources of uncertainty in our model (e.g., subglacial hydrology, iceberg calving, evolving fabric, and glacial isostatic adjustment), we focus on processes that we expect to have a substantial impact on projections while also being feasible to explore in our model framework.”

L761 I wonder whether the change in SLC is really symmetric as you vary the melt sensitivity parameter?

It really is roughly symmetric for the simulations we have run. But as we noted above, we can’t know if intermediate values would behave linearly without doing more (expensive) simulations. We have made no changes.

L767: I suggestion explicitly stating (e.g. in brackets) what the parameters represent (e.g. q=1 -> linear viscous).

We have added "(linear viscous)" "(hard-bedded)" and "(effectively plastic)" for q=1, q=1/3 and q=1/10, respectively.

Figures:

Figure 1: your sectors of the Ross and Filchner Ronne ice shelves are quite large as they encompass a considerable chunk of EAIS as well. This means that you mix the regional response of marine ice sheet/shelf systems with inland EAIS sectors which can lead to compounding effects in your regional analysis. You could use more refined basins which only include the Filchner-Ronne and Ross shelf sector respectively without including the EAIS. I am aware that this would involve redoing parts of the analysis. Thus if you prefer to keep those basins, please mention the caveats of the basin selection explicitly.

Our analysis for these sectors is done during run-time, and our entire workflow (including the initial parameter tuning all the way through the ANOVA) is based on these sector definitions, so we will not change them at this point. We have added a clarification in S2.3.1 that: "We chose to define basins that incorporate the entire catchments of the Filchner-Ronne and Ross ice shelves to avoid sharp transitions in our tuned deltaT values along an arbitrary boundary within the ice shelves. We note that these two sectors thus include portions of both East and West Antarctica, and an analysis that used smaller basins could provide different insights."

Fig 6, 17 & 19 : I suggest focusing on WAIS as your discussion also almost entirely covers this region. The continental scale illustration makes it difficult to identify the differences in grounding line position. You could add the continental scale in the supplementary material.

We prefer not to change these figures, for several reasons. As you noted above, our basin delineations include large portions of East Antarctica. We also include the Amery catchment in our time series and ANOVA plots (Figs 16, 18, 21), so it would be awkward to leave that sector out of the maps. For Figure 6, we include the whole continent for comparison with Figure 4, as there are more changes in East Antarctica in these longer simulations. We have made no changes here.

Figure 13: while you only have 5 vertical layers the figure seems to indicate variations within a single layer. I assume this is a plotting artifact. Also, it could be beneficial to show temperature differences between the temperature solver and the run with fixed temperature. This would illustrate differences more clearly. In this specific case I would replace the absolute temperature on panel b) with the difference between a temperature solver and assuming fixed temp.

We assume that you are referring to vertical variations, which should not be allowed within a layer, as opposed to horizontal variations that are expected and allowed for our 3D temperature field. The appearance of vertical variations within a layer are a trick of the eye and not actually present in the plot. We agree that there do appear to be vertical variations of temperature within a layer, but we have used the Apple Digital Color Meter tool to examine the columns within a layer and find no vertical variation in RGB values. The effect is probably due to the curved interface between the layers and the vertical exaggeration. We have made no change to the figure on this point, as we do not have control over this effect.

Regarding plotting the differences between the temperature fields, this is an interesting thought but would be rather confusing in practice. The ice geometries are not the same, so plotting the temperature difference would leave areas of undefined (or physically meaningless) temperature difference. Instead, we have added transects of the depth-averaged temperature in the lower panels. This highlights an important difference that we had not noticed before: the temperature difference extends significantly upstream of the grounding line. We attribute this difference to increased advection of cold ice from upstream in the thermomechanically coupled case, and we have added text to this effect at the end of Section 3.5

Figure 15: To a reader unfamiliar with these kinds of plots it might be confusing that the thickness change is only shown where there is ice cover in both the projection and the respective baseline run.

These plots are simply the thickness difference between the two simulations at 2300. Note that this is not a difference of the RCP8.5 projection run versus the control run, but the difference between two RCP8.5 projection runs using different stress balance solvers. If ice is present in one run and not the other, then the result is a non-zero value. We have made no changes.