Response to reviewers' comments on "Evolution of the Antarctic Ice Sheet from 2000–2300 and beyond: model sensitivity and uncertainty analysis using MPAS-Albany Land Ice" by Trevor R. Hillebrand, Matthew J. Hoffman, Holly K. Han, Mauro Perego, Alexander O. Hager, Andrew Nolan, Xylar Asay-Davis, Stephen F. Price, Jerry Watkins, and Max Carlson

Response to Reviewer #1

We thank the anonymous reviewer for the helpful comments on the manuscript. Our responses to their suggestions are in blue.

l.63: Consider adding references to the datasets mentioned

We will add these references.

l.135-142: A figure illustrating the modifications to the bedrock dataset would be very helpful. Perhaps this could be included as supplementary material.

We will add a figure to the Appendix that shows the changes to the bed topography.

l.144: This is the first mention of "sectors." Since multiple basin delineations exist, it would be helpful to specify which sectors you are referring to.

We will add a statement clarifying that we follow the sectors definitions used by ISMIP6.

l.168-169: Is the melting of grounded marine termini a significant process in Antarctica?

Good question! Drewry et al. (1982) noted that 38% of the length of the ice-sheet margin consists of grounded "ice cliffs". Where this ice is exposed to positive thermal forcing, some melt can be expected. However, there does not appear to be much observational work on this. Dryak and Enderlin (2020) found significant positive correlation between iceberg melt rates and glacier front ablation rates (i.e., calving plus melting) with Spearman's $\rho = 0.71$, p-value = 0.003, which could suggest significant melt undercutting of grounded glacier termini. Davison et al. (2024) point out that the available data for Antarctic Peninsula glaciers points to termini that are hundreds of meters thick, which makes them thick enough to be exposed to circumpolar deep water. Over our 2000–2015 historical simulation, we find an average melt flux of ~7.7×10^{13} kg yr^{-1} from grounded marine termini, compared with 1.2×10^{15} kg yr^{-1} average ice-shelf melt flux. So, in our historical simulations, melt at grounded marine termini is about 6% of the total ocean-forced melt, which is a small but non-negligible contribution. Furthermore, including this parameterization allows for ocean forcing of grounded marine termini through the course of the projection simulations, which is especially important due to the absence of prognostic calving in our model configuration. In future scenarios with very warm ocean water, neglecting this process could lead to a significant underestimation of ocean-forced melt, especially once ice shelves have largely melted or calved away. We will add an explanation to this effect where this parameterization is introduced.

l.169-171: Do I understand correctly that you do not use the deltaT values provided by Jourdain et al., 2020, but rather calculated your own?

That is correct. The deltaT values provided by Jourdain et al. (2020) did not give us a good fit to overall observed ice-shelf melt rates, likely because they used BedMap2 to define the ice and seafloor geometry and we used BedMachine v2. We used the Jourdain et al. (2020) values as a starting point for calculating our own deltaT values based on ice-shelf-average melt rates. We will add a statement to this effect.

l.184: This was also shown in Coulon et al. (2024), where significant retreat in the ASE is found under constant present-day forcing, using a different initialisation procedure than van den Akker et al. and while accounting for parametric and climate uncertainty.

We will add this reference to the text as further substantiation.

l.185-186: It would be useful to add a short explanation of what motivated the choice of forcings to be extended. I assume that the goal was to sample a wide range of forcings?

Yes, we wanted a control scenario, a low-emissions scenario (UKESM-SSP1-2.6-2300-extended), and a high emissions scenario with and without hydrofracture (CCSM4-RCP8.5-2300-extended and CCSM4-RCP8.5-2300-h-extended). We chose the CCSM4-RCP8.5 scenario rather than the HadGEM2-RCP8.5 scenario because in the HadGEM2 baseline simulations, West Antarctica has already largely deglaciated by 2300 (see Fig 4d), so we assumed that further changes beyond 2300 were likely to be relatively small. We will add a statement clarify these choices.

l.272-273: I believe fig.2 from Seroussi et al. (2024) only shows ice thickness and velocity RMSE, not the historical mass change trend.

This sentence was supposed to refer both to Fig. 2 in Seroussi et al. (2024) and to our own Fig 2, but we agree that it was worded poorly. We will rephrase to make it clear what comparisons are being made.

l.274-275: Maybe I am looking at the wrong figure, but from Figure 4 in Seroussi et al. (2024), I don't find that DOE_MALI contributions are closer to NCAR_CISM than to other models (for example VUW_PISM). Also, I am not sure what to take away from this information.

We will remove this statement.

l.294-296: I’m not sure I understand what you are trying to say here. Could you clarify?

Thanks for pointing out this confusing wording. What we mean to say here is that if one assumes that the Jourdain et al. (2020) parameterization is correct, then the existing calibration for gamma0 seems to provide a reasonably narrow constraint on future sea-level change. We will update the text for clarity.

Figures 16 & 18: These figures would benefit from error bars or whiskers, as the overlapping shaded regions are difficult to distinguish.

We agree that the regions can be difficult to distinguish, but whiskers would make for busy-looking plots. We will experiment with the shading opacity and adding darker curves to the borders of the shaded region to make the ranges more distinguishable.

l.376: I believe the reference should be to Fig. 16b, not 17b.

Thanks for catching that. We will change it.

l.379-381: Consider splitting this sentence into two for clarity.

We will change this to read: "The spread in ensemble members is small for the FRIS and Ross sectors. However, many ensemble members predict a close-to-modern-day grounding-line position in the ASE by 2200, while many others predict hundreds of kilometers of grounding-line retreat by this time."

l.406-407: Could this interaction term be explained by the fact that some ESMs (e.g., CESM) generate more surface melt, making them more susceptible to triggering hydrofracture?

Yes, that's a clear way of describing it. We will include this.

l.422-423: This could support my earlier comment regarding the e–h interaction term.

We will add, "…, likely reflecting the wide range of ESM-predicted surface melt available to drive hydrofracture."

l.438: I could not find the 1.5 m value in Stokes et al. (2025). Are you referring to the 134 cm reported in their Table 2?

The 1.5 m value refers to the curve labeled "+1.5°C median" (the multiple occurrences of the value 1.5 is just coincidental) in Figure 4 of Stokes et al. (2025). We can add a specific reference to their Figure 4 for clarity.

l.441: I agree with your point, but it may also be worth noting that the extended experiments rely on a single model configuration. The onset time of retreat is also likely to be strongly influenced by structural and parametric uncertainties.

Yes, but this also falls under the umbrella of "better model calibration". However, we can add these caveats.

l.455-456: This is an interesting result. Could you provide a tentative explanation for the different behaviors of ASE versus Ross and FRIS? Also, did I understand correctly that all simulations with different q values start from the same initial state, with the basal friction field rescaled? If so, please specify this explicitly to rule out influences from the initialisation.

We think this is likely due to the more complex basal topography in the ASE, but there is no straightforward way to know for sure. We will add this as one possible explanation.

Yes, we solved the inverse problem for the basal friction field using a value of q=1/3, and then recalculated it to produce the same basal traction at the initial time when using the different assumed values of q. We will reiterate this here.

l.517-578: This is an important point, but I think it would be worth noting that Willams et al. expect the effects of model resolution on the upper tail to be model dependent.

We will add a clarifying statement to this effect.

l.518: Is it the case for MALI as well? This is what I think I understand from section 2.2, although it is not specifically mentioned.

Yes, MALI currently lacks adaptive mesh refinement. We will clarify this.

l.594: One possible impact of surface meltwater on ice dynamics could be its influence on the temperature profile of ice shelves, potentially leading to warmer conditions than those shown in Fig. 13 using the temperature solver.

This is an interesting point. We will add a statement about this and cite Hubbard et al. (2016) as documentation of the impact.

Figure C3: There is something strange with the colorbar.
Oops, thanks for catching that! We will reduce the number of tick labels to prevent them overlapping.

Minor corrections:

l.51: ‘can further’ → ‘can be further’

This will be corrected.

l.566: ‘hypothesize’ → ‘hypothesis’

This will be corrected.

References

Davison, B. J., Hogg, A. E., Moffat, C., Meredith, M. P., & Wallis, B. J. (2024). Widespread increase in discharge from west Antarctic Peninsula glaciers since 2018. *The Cryosphere*, *18*(7), 3237-3251.

Drewry DJ, Jordan SR, Jankowski E. Measured Properties of the Antarctic Ice Sheet: Surface Configuration, Ice Thickness, Volume and Bedrock Characteristics. *Annals of Glaciology*. 1982;3:83-91. doi:10.3189/S0260305500002573

Dryak, M. C., & Enderlin, E. M. (2020). Analysis of Antarctic Peninsula glacier frontal ablation rates with respect to iceberg melt-inferred variability in ocean conditions. *Journal of Glaciology*, *66*(257), 457-470.

Hubbard, B., Luckman, A., Ashmore, D. W., Bevan, S., Kulessa, B., Kuipers Munneke, P., ... & Rutt, I. (2016). Massive subsurface ice formed by refreezing of ice-shelf melt ponds. *Nature communications*, *7*(1), 11897.

Response to Reviewer #2:

Hillebrand et al. present an analysis of the ISMIP6 AIS 2300 simulations (Seroussi et al., 2024) they performed with MALI as well as some additional sensitivity experiments. This is a useful study to help understanding the large spread of sea level projections for Antarctica.

The manuscript is well written and clear, however lacks some explanations and motivations. I currently do not understand why exactly those parameters, values, and modelling choices were tested, and not others. Furthermore, more explanation on the "non-linearity" of the sliding parameter sensitivity is required, and I cannot follow the argument that the melt parameter choice is less relevant. The argument about initialisation uncertainty in ISMIP6 AIS 2300 being the dominant driver for uncertainty is not clear to me.

I suggest minor comments concerning the framing of the results which should be addressed.

We thank the reviewer for these helpful comments. The reviewer's main points are well taken: we will motivate our choices of experiments more carefully; we will be more rigorous and quantitative in determining the sensitivity of the model to these different choices; and we will consolidate and clarify our reasoning behind our hypothesis that initialization uncertainty could account for a large fraction of the inter-model spread in ISMIP6-Antarctica-2300.

Comments:

Abstract

Line 8: "only moderate sensitivity". I disagree with this point in contrast to claiming higher sensitivity for other parameters later on. Looking at Figures 7a (melt parameterisation parameters) and Figure 9a and 11a, the melt parameter induces differences >50cm SLE in 2300 for each experiment, which is similar to changes induces by the energy balance solver (which is claimed to induce "very strong sensitivity" in the abstract) and the sliding law exponent (except for the q=1 case for the HadGEM forcing; claimed to induce strongly non-linear dependence). I do hence not see how the authors say that melting has "only moderate sensitivity" in contrast to the other parameters, as the overall influence on the sea level results is at the same order?

We acknowledge that the boundaries between "moderate", "strong", etc are very subjective. We will make our assessment of the sensitivity to the various parameter value and fidelity settings more quantitative and update the text to reflect this. However, the q=1 run should not simply be discounted when discussing model sensitivity, since plenty of models (at least two in Seroussi et al., 2024) still use linear viscous basal friction laws.

Line 9: "ranging from the 5th to the 95th percentile values" – unclear where these come from, think about adding "of the AntMean ISMIP6 calibration" or something similar. I suspect that the choice of your range of parameters has a strong influence on your sensitivity results, and if you had included other values, you would get a different answer (as you mention yourself in the discussion). The statement you can make is that using this parameter tuning, you find a less important contribution in the ANOVA results.

Yes, if we extended our range of parameter values we would certainly see stronger sensitivity to sub-shelf melt. However, there is not a good justification for extending these values further than the 5th–95th percentile values. Burgard et al. (2022) showed that the MeanAnt calibration target performs significantly better than the higher-melt PIGL calibration against reference NEMO simulations (see their Table 3), so we use the only viable calibration target with our chosen parameterization. We do not investigate the structural uncertainty inherent in using the Jourdain et al. (2020) quadratic parameterization versus other available parameterizations, but we have been clear about this limitation and refer to appropriate studies like Lambert & Burgard (2025) that rigorously investigate other choice of melt parameterizations. We mention in the Methods section that we use the MeanAnt calibration target, but we can add this to the abstract as well.

Line 12: not clear what you mean with "vice versa". Do you mean that FRIS and Ross have less mass loss with a less plastic sliding behaviour? In that case, maybe simply remove it.

Yes, that's what we meant. We will remove the confusing "vice versa" phrasing.

Introduction

General: The introduction is missing a discussion of previous results on sensitivity analysis of ISMIP6 simulations. It might be worth checking all ISMIP publications. For example, I could find this study TC - ISMIP6-based Antarctic projections to 2100: simulations with the BISICLES ice sheet model which seems to do something very similar.

There are only a few such studies available, but we will add an overview of those studies to the Introduction.

Moreover, I cannot find a motivation as to why exactly the parameters/schemes tested were tested, at the moment their choice appears arbitrary. Make sure to motivate your choices!

We will add this motivation to the Introduction.

Adding these two points might make the introduction longer, you could consider removing information that is not directly relevant for your study (e.g., on MICI, calving, flow law exponents, etc).

After adding the suggested material to the introduction, we will be careful to prune away text that is less directly relevant.

Line 26: “Onset of such runaway retreat leads to……” – there is a subtly here. To my knowledge DeConto & Pollard, De Conto et al. do not check for the retreat being MISI driven in their results, hence it is not clear if the MISI-driven retreat causes the multiple metres of sea level rise in their simulations. MICI is then added as an afterthought, as if it was to enhance the already strong results. I suggest you reformulate.

These studies do in fact attribute their simulated ice-sheet retreat to MISI several times throughout each paper. From DeConto & Pollard (2016), for instance:

“Pliocene retreat is triggered by meltwater-induced hydrofracturing of ice shelves, which relieves backstress and initiates both MISI and MICI retreat into the deepest sectors of WAIS and EAIS marine basins.”

“Without atmospheric warming, the magnitude of RCP8.5 ocean warming in CCSM4 is insufficient to cause the major retreat of the WAIS or East Antarctic basins; and even with >3 °C additional warming in the Amundsen and Bellingshausen seas it takes several thousand years for WAIS to retreat via ocean-driven MISI dynamics alone.”

“Despite these limitations, our new model physics are shown to be capable of simulating two very different ancient sea-level events: the LIG, driven primarily by ocean warming and MISI dynamics, and the warmer Pliocene, in which surface meltwater and MICI dynamics are also important.”

Similarly, from DeConto et al. (2021):

“Similar to other models without ice-cliff calving, enhanced precipitation in East Antarctica partially compensates for MISI-driven retreat in West Antarctica”

“Setting these parameter values to zero (Extended Data Figs. 1, 6) effectively eliminates hydrofracturing and ice-cliff calving, limiting rates of ice loss to processes associated with standard calving, surface mass balance, sub-ice melt and MISI, as in most other continental-scale ice-sheet models.”

However, as you point out, there does not appear to be any rigorous check for whether the rapid retreat is truly due to the positive feedback that defines the marine ice sheet instability, although it seems quite likely that MISI does drive that retreat. We can rephrase to make it clear that this is their interpretation, however, and integrate the statement about MICI more completely with the rest of the statement.

Line 58 “regardless of their skill in reproducing historical ice-sheet behaviour”. I suggest reformulating, as it sounds like this was tested/required by ISMIP6, however, reproducing historical behaviour was not an aim of ISMIP6, and the trend analysed by Aschwanden et al. was never meant to be a representation of historical behaviour.

It is true that this was not a requirement of ISMIP6, but the fact that there is an experiment labeled “historical” suggests that matching historical observations is at least desirable. And our use of “regardless” makes it clear that historical behavior was not used for calibration or weighting of individual ISMIP6 contributions. As Aschwanden et al. (2021) point out, it is hard to trust projections using models that perform poorly over the observational period. However, we can clarify that ISMIP6 did not require good agreement to historical changes.

Line 152 “As MISI is likely already under way” the reference is quite old, and more recently, it has been more specified that “MISI being under way” needs to be tested by more than just finding continued retreat in the Amundsen Sea. You do not need MISI here, simply that the ASE is changing is sufficient.

We will reframe this sentence to reflect more recent work and hedge the statement about MISI being “likely already under way”, changing it to “may already be under way”.

Line 169: How much do your deltaT values differ from the original ones? Note that this might impact how sensitive your melt reacts to changes in the ocean forcing, see for example Lambert & Burgard 2025.

This is true, although the sensitivity to changes in deltaT is likely less than the sensitivity due to gamma0. We will add a table to the Appendix documenting difference between our deltaT values and the ones provided by ISMIP6.

Methods

Section 2.3.3 & 2.3.4. The choice of the parameter ranges for the sub-shelf melt and the sliding parameters appears a bit arbitrary. If they are not sampled equally broadly, how can you claim that one factor has a higher influence than the other?

The sub-shelf melt parameter samples from the 5th to the 95th percentile values reported by Jourdain et al. (2020) for Antarctica. Our sliding parameter spans linear viscous to nearly plastic behavior. These both represent what we take to be the widest defensible range of values for these parameters for the given parameterization and calibration target. We will clarify this in the text.

Results

Section 3.3 & 3.4 How do the control runs look with changed melt or sliding parameters? Is the drift again small enough to be ignored in your plots, or how much influence does this have on your results (especially with the "non-linearities" in the sliding exponent experiments)?

In the pre-print, we did not examine the impact of these parameters on the control simulation. For each sensitivity test, the relevant comparison is to the baseline simulation (i.e., q=1/5, median gamma0 value) that uses the same forcing, rather than the control simulation with the same parameter values. However, following your feedback, we have now run control simulations for each of the sensitivity experiments and will include these in the manuscript. We prefer to compare the simulations without subtracting the control run, in keeping with the approach used for ISMIP6-Antarctica-2300. In this way, we can demonstrate what the impact of a different parameter choice would have been had it been included in the ISMIP6 ensemble. But we will report the drift in the respective control runs and discuss these magnitudes relative to the forced response. For the most part, the differences are small, but we note that subtracting the control drift would have the effect of isolating the forced response from the unforced (control) response, rather than necessarily being the "correct" way to demonstrate sensitivity.

Line 289 "indicate only moderate sensitivity" as stated above, I disagree with this as your difference is still 50-60cm between the low and high end, which is more than a doubling of the contribution for the CCSM4 forcing.

We will be careful to report the quantitative values rather than relying on more qualitative descriptions of "low", "moderate", "extreme", etc.

Line 295 "This indicates… the calibrated range of gamma for this single paraeterization using the AntMean calibration is fairly well constrained." This appears grammatically odd, reformulate.

We will split this into two sentences to make for easier reading.

Section 3.4: How much of your non-linear behaviour comes from the following sources: (1) drift in the control state if you simply switch the sliding exponent and re-scale basal slipperiness, (2) the fact that the Amundsen Sea basin is simply "empty" at some point, (3) how much your choice of the sliding exponent determines if Pine Island also disappears (under control conditions)?

1) We have now run the control simulations for these different basal sliding exponent values. Most of the non-linear behavior still arises from the forced simulations, rather than from the unforced response.

2) The large slope of the purple mass-change curves in Fig 9b show that the Amundsen Sector does not run out of ice under CCSM4-RCP8.5 forcing. The slope begins to level off after 2250 in Fig 9c, when the sector begins to run out of ice under HadGEM2-RCP8.5 forcing.

3) Neither Pine Island Glacier nor Thwaites Glacier disappears under control forcing.

Line 306: How much of the mass gain, and “non-linear behaviour” for the simulations comes from the drift in the control runs when you simply change parameters? Especially the q=1 case with net mass gain might hint at some model drift generated by switching parameter values. Your figure 2 shows mass gains in that case until 2015, what happens if you extend this case? You might want to consider analysing results relative to the drift.

We have now run the control simulations for each sensitivity experiment and will include these results in the manuscript. As we noted above, we prefer to follow the approach used by ISMIP6-Antarctica-2300 and not subtract the control drift from each run, but we will report the amount of drift due to each control simulation relative to the forced simulation. Comparison to these simulations does not change our interpretation.

Line 309: see comment before.

Addressed above.

Line 325: “Under …, all three values <=1/3 yield almost the same SLC” is that not because around 1m, the Amundsen Sea basin is empty?

The rate of mass loss is still high at 2300 in those simulations (although beginning to decrease), so the basin is not completely empty. At the initial condition, this basin contains 1.2 m sea-level equivalent ice. However, it could be true that ice on a back-sloping bed has begun to run out.

Section 3.5

Line 340 “strong effect” how do you see this is stronger than for the melt sensitivity in the magnitude and patterns?

As with previous similar comments, we will update this statement with a more quantitative assessment of the melt and energy balance sensitivities to make our interpretation clear.

Line 349: What do you think the effect of your coarse vertical resolution of only 5 layers is on this result?

We do not expect this finding to be particularly dependent on vertical resolution. This is an average vertical resolution of 100 m for a 500m-thick ice shelf, but the true resolution will be finer because the vertical layers are concentrated towards the ice base. We will add a sentence acknowledging this limitation.

Section 3.6

What happens in the control runs for the two different solvers?

We have now run the control simulation for the MOLHO solver, and it is very similar to the control using the Blatter-Pattyn solver.

Section 3.7

General: Are results analysed relative to the model drift?

No. Following Seroussi et al. (2024), we ignore the model drift in the control run because it is small compared to the forced response in the high emissions scenarios. The drift in our control run is ~40 mm by 2300, compared with ~750 mm by 2300 for the CCSM4-RCP8.5 forcing and ~2800 mm for the HadGEM2-RCP8.5 forcing. We will be sure to clarify this.

Figure 20: the two blues are hardly discernible, please update one.

The two shades of blue used here are very different and are distinguishable from each other using all filters (including greyscale) provided by the Coblis Color Blindness Simulator at https://www.color-blindness.com/coblis-color-blindness-simulator/.

Line 403: "Despite the wide range of melt parameter sampled", arguably, using the AntMean does not necessarily sample a wide range.

We sample a wide range within the constraints of the calibration (5th to 95th percentile). And as we noted above, the MeanAnt calibration is much more realistic than the only other available calibration target (PIGL) using this parameterization. We note in many other places in the manuscript that the calibration target or choice of the form of the parameterization could have a large impact, so we do not feel that we need to change this wording, but we can add a parenthetical specifying the constraints.

Comparing with Seroussi et al., 2024, Figure 15, your hydrofracture and climate terms reach similar values of around 60cm variance by 2300. However, the "ice model" term in Seroussi is much higher in variance than your "sliding exponent" and "melt parameter" terms, with overall variance of the Seroussi ensemble being larger (1.6m instead of 1m in your ensemble). It appears that the variance you find from hydrofracture and climate forcing is in line with the variance in the full ISMIP6 2300 ensemble. However, you sample less variance due to parameter choices than in Seroussi (where everything goes into the "ice model" term), which explains why overall the relative contributions of hydrofracture and climate are much higher in your case.

We agree, and will point this out.

Line 400 "Likely reflecting the importance of each of these processes on their own". Not sure about this, as their individual importance should be in the single contributions? Please explain.

What we mean here is that even a relatively weak interaction between two dominant terms will result in a larger share of the variance than a relatively strong interaction between two terms whose individual contributions are negligible.

Discussion

Line 431: Feldmann & Levermann 2015 do not make a statement about modern day melt rates causing retreat, Favier et al. 2014 is on Pine Island not Thwaites.

The reviewer is correct that we are missing some nuance about the experimental design in Feldmann & Levermann (2015). However, Feldmann & Levermann (2015) do in fact demonstrate retreat using modern day melt rates. Here is a relevant quote from their abstract: "In our simulations, at 5-km horizontal resolution, the region disequilibrates after 60 y of currently observed melt rates. Thereafter, the marine ice-sheet instability fully unfolds and is not halted by topographic features." And another quote from their Methods: "Our equilibrium simulation is forced by time-averaged FESOM data for the period 1970–1999, representing 20th-century conditions. Compared with observations, sub-ice-shelf melting from FESOM is reasonably captured for Filchner–Ronne and Ross ice shelves, but significantly underestimated beneath the ice shelves in the Amundsen Sea (figure 10 in ref. 21). For this region, average FESOM melt rates are around 3 m/y, whereas present-day average melt rates are observed to be one order of magnitude larger (15, 29). We thus scale the field of melt rates below the Amundsen ice shelves toward values that are consistent with current observations to represent the observed enhanced melting below these ice shelves, perturbing the equilibrium ice sheet locally. The perturbation is applied for lengths of 20, 40, 50, 60, 80, 100, 150, and 200 y (Fig. 3). After the end of the perturbation pulse, the original (but weaker than observed) 20th-century melt rates are applied, and simulations are integrated for several thousand years until a stable grounding-line position is reached." So, present-day melt rates do drive retreat in their simulations, but that is relative to an equilibrium state that may or may not represent the late 20th century. However, this is far too nuanced to include in this very general statement about present-day melt rates potentially driving long-term retreat, so we opt to keep the wording as it currently stands.

Thanks for catching the misrepresentation of the Favier et al. study. We will rephrase this sentence to be more general to the Amundsen Sea, rather than just to Thwaites Glacier.

Line 433: Feldmann does not show that reducing melt rates would halt retreat. They show that after kicking the Amundsen Sea with a strong melt water pulse for a short while, reversing to current melt rates is not sufficient to stop long term retreat.

Feldmann & Levermann (2015) show continued retreat for modern-day melt-rate forcing of longer than 60 years. In addition to the text quoted above, here is a quote from their Results section (underlining is our own): "For simulations with a perturbation duration of at least 60 y, an abrupt increase in ice loss then marks the onset of the self-sustained ice loss (Fig. 3*B*). In these destabilized simulations, the major part of West Antarctica's marine ice disintegrates". In their Fig 3, the blue curves represent simulations with elevated melt-rate forcings of <60 yr, which do not lead to destabilization of the ice sheet. We believe the text is correct as we currently have it.

Line 444: "Our .. remarkable insensitivity" see previous comments.

We will address this by being more quantitative, as in previous comments.

Section 4.3: How big do you think is the influence of vertical resolution on your finding?

As noted above, we do not expect the impact of vertical resolution to be large, but we will note that this has not been tested.

Line 525: Do you mean with grounding line parameterization a scaling of basal friction around the grounding line with the grounded fraction or do you mean a schoof-type assumption on grounding line flux?

We mean the former, and we will clarify the text.

Line 547: “We take this as further support…” I think this is the first time this hypothesis is raised in the paper. Moreover, I do not understand the jump from discussing model fidelity to the initial conditions.

Thanks for pointing this out. We will move the text regarding initial condition uncertainty to its own subsection of the discussion and lay out our hypothesis more systematically.

Line 565: “historical behaviour” I would call this trends in the initial conditions or so, as there was not experiment with historical forcing to test for their behaviour.

There was an experiment with historical forcing, called “Historical” in the ISMIP6-Antarctica-2300 protocol. All forced runs branch from the end of the historical run, so its behavior could well be critical to determining the forced response. The reviewer is correct that there was no specific analysis of the trend of the Historical simulation, but we do not consider that relevant to this sentence.

Line 566: “This further supports”.. again, the hypothesis has not been stated anywhere clearly and argued for. The variance in your ensemble is overall smaller than in the original ISMIP6 AIS 2300 simulations, and as you say, the reasons could be multiple, ranging from the initial condition uncertainty to you not sampling parametric uncertainty as widely. I do not see how you can exclude all other factors based on your experiments?

We do not exclude all other factors, nor is our hypothesis based solely on our experiments, but rather on a combination of our experiments, the ISMIP6 ensemble, and other published work cited in this section that rigorously quantifies the impact different choices across multiple models. We have been clear that this is simply a hypothesis to be tested or considered in future model intercomparisons, and which we cannot test with our current experimental design. But we agree that we have spread the text out across too many different portions of the manuscript. We hope that consolidating the text regarding initial condition uncertainty will help clarify this.

Line 571 and following: This is an interesting observation.

Thanks!

Conclusions

Line 673: See comment on Favier and Feldmann before. An interesting study here is also Bett et al., 2024 (TC).

We will correct the statement with respect to the Favier et al. paper. See our previous response about Feldmann & Levermann (2015). We will add a reference to the Bett et al. (2024) paper.

Line 675: See comments before on claims of melt sensitivity being less important.

As above, we will rely on quantitative rather than qualitative description here.

Line 692: How generalisable are your results? Do other ISMIP6 AIS 2300 spin-off studies report similar findings? Otherwise, your findings might be model-specific?

We can only speculate about how generalizable our results might be, and there very few ISMIP6 AIS 2300 spin-off studies to date. We hope that this paper will encourage other groups to publish similar sensitivity-type analyses that could further clarify the reasons for the inter-model spread documented by Seroussi et al. (2024). Model-specific findings are important to document, as they can account for inter-model spread in multi-model ensembles.

Line 698: Again, I do not follow your argument about initialisation uncertainty.

Addressed above.

Appendix

Figures C1-C4: I do not think they were referred to in the main manuscript. Explain please why you include them. Figure C3 something went wrong with the colour bar labels.

We refer to these figures in the captions of Figures 8, 10, 12, and 15, but we can add references to these in the text where the West Antarctic figures are referenced. Figures 8, 10, 12, and 15 in the main manuscript are zoomed in to West Antarctica by request of the editor, but we wanted to retain the full continental results for completeness, so we added these to the Appendix.

We will fix the issue with overlapping text on the colorbar labels.

References:

Aschwanden, A., Bartholomaus, T. C., Brinkerhoff, D. J., & Truffer, M. (2021). Brief communication: A roadmap towards credible projections of ice sheet contribution to sea level. *The Cryosphere*, *15*(12), 5705-5715.

Burgard, C., Jourdain, N. C., Reese, R., Jenkins, A., & Mathiot, P. (2022). An assessment of basal melt parameterisations for Antarctic ice shelves. *The Cryosphere*, *2022*

Bett, D. T., Bradley, A. T., Williams, C. R., Holland, P. R., Arthern, R. J., & Goldberg, D. N. (2024). Coupled ice–ocean interactions during future retreat of West Antarctic ice streams in the Amundsen Sea sector. *The Cryosphere*, *18*(6), 2653-2675.

DeConto, R. M., & Pollard, D. (2016). Contribution of Antarctica to past and future sea-level rise. *Nature*, *531*(7596), 591-597.

DeConto, R. M., Pollard, D., Alley, R. B., Velicogna, I., Gasson, E., Gomez, N., ... & Dutton, A. (2021). The Paris Climate Agreement and future sea-level rise from Antarctica. *Nature*, *593*(7857), 83-89.

Feldmann, J., & Levermann, A. (2015). Collapse of the West Antarctic Ice Sheet after local destabilization of the Amundsen Basin. *Proceedings of the national academy of sciences*, *112*(46), 14191-14196.

Jourdain, N. C., Asay-Davis, X., Hattermann, T., Straneo, F., Seroussi, H., Little, C. M., & Nowicki, S. (2020). A protocol for calculating basal melt rates in the ISMIP6 Antarctic ice sheet projections. *The Cryosphere*, *14*(9), 3111-3134.

Lambert, E., & Burgard, C. (2025). Brief communication: Sensitivity of Antarctic ice shelf melting to ocean warming across basal melt models. *The Cryosphere*, *19*(7), 2495-2505.

Seroussi, H., Pelle, T., Lipscomb, W. H., Abe-Ouchi, A., Albrecht, T., Alvarez-Solas, J., ... & Zwinger, T. (2024). Evolution of the Antarctic Ice Sheet over the next three centuries from an ISMIP6 model ensemble. *Earth's Future*, *12*(9), e2024EF004561.