Dear Dr. Sutter,

It is our pleasure to submit the revised version of our manuscript, "Evolution of the Antarctic Ice Sheet from 2000--2300 and beyond: model sensitivity and uncertainty analysis using MPAS-Albany Land Ice", for consideration for publication in The Cryosphere. We have incorporated the comments from two anonymous reviewers that have improved the manuscript. While most individual comments were minor, the most significant changes we have made were (1) to rely less on subjective language when assessing model sensitivity; (2) to conduct additional sensitivity tests using control forcing; and (3) to add several figures to the Appendix detailing our new control simulations, our changes to the BedMachine v2 bed topography, and our calibrated δT values for ocean thermal forcing. We have also added motivating text to the Introduction for why these experiments were conducted and clarified specific questions about motivation for different parameter choices in the methods section. Our responses to the reviewers' comments can be found below, with our line-number references referring to the version with tracked changes.

Sincerely,

Trevor Hillebrand, Matthew Hoffman, Holly Han, Mauro Perego, Alexander Hager, Andrew Nolan, Xylar Asay-Davis, Stephen 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 have added references (Fretwell et al., 2013; Morlighem et al., 2020; Pritchard et al., 2025) and moved this text to Section 4.6 based on the other reviewer's suggestions for tightening up the Introduction.

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 have added Figure B1 to the Appendix that shows the changes to the bed topography and referenced it from the text.

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 have added a statement clarifying that we follow the sectors definitions used by ISMIP6, with some modifications: L 169 "…which we define following Jourdain et al. (2020) and Rignot et al. (2019), but with a single sector for each the Ross and Filchner-Ronne ice shelves as shown in Fig. 1c"

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. Note that while we adopt the parameterization and parameters values specified for ISMIP6-Greenland without modification, we assume subglacial discharge is zero, despite it being non-negligible in reality, so our parameterized melt rates may be underestimated.

We have added an explanation to this effect where this parameterization is introduced:

L195: "…the melt undercutting parameterization of Rignot et al. (2016) for melting of grounded marine termini, which comprise roughly 38% of the coastline of Antarctica (Drewry et al., 1982)."

L201: "In our historical simulation, melting at grounded marine termini accounts for about 6% of the total submarine melt flux."

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 of differences in ice-shelf geometry: they used BedMap2 to define the ice and seafloor geometry and we used BedMachine v2 in most of the domain, as well as perform a short relaxation. We used the Jourdain et al. (2020) values for gamma0 but then recalibrate our own deltaT values using our initial condition geometry to match observed ice-shelf-average melt rates. We have added a statement to this effect and provided a new figure to the Appendix (Figure B2):

L198: "Our calculated δT values are shown in Figure B2, along with the reported values from (Jourdain et al., 2020) for comparison."

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 have added this reference to the text in L 215 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 have rewritten much of Section 2.3.2 to clarify these choices:

"As none of the simulations submitted to ISMIP6 had reached a steady state by 2300, we selected a subset of the baseline experiments (low emissions, high emissions with and without hydrofracture) to extend to 2500, and we extended the control simulation to the year 2775. We extended the control simulation out further to determine whether the present-day forcing is sufficient to cause significant ice-sheet retreat in our configuration, which has been found by other studies using different ice sheet models and initialization and calibration procedures (Coulon et al., 2024; van den Akker et al, 2025). We chose CCSM4-RCP8.5-2300 as the high-emissions scenario because of the large ice-sheet mass and area remaining at 2300 our baseline simulations, which would allow for significant change beyond 2300 (Fig. 3, 4). We chose UKESM-SSP1-2.6-2300 because it is the only available low-emissions scenario run out to 2300. We extended the forcing to 2500 by randomly sampling forcing from the years 2280–2300, similar to the approach used by ISMIP6-Antarctica-2300 to extend the 21st century forcing out to 2300 in the "repeat" experiments."

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.

We rephrased to make it clear what comparisons are being made: L 311 "While our model fit to observations is among the best of the multi-model ISMIP6-Antarctica-2300 ensemble when quantified by ice thickness and velocity root mean square error (Figure 2 in Seroussi et al., 2024) and have a modest bias in historical mass-change trend (Fig. 2), our SLC predictions fall roughly in the middle of the predicted range at 2300 (Figure 4 in Seroussi et al., 2024)."

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 have removed 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 have updated the text for clarity, taking the other reviewer's comment into account as well:

L 336: "While the proper form of the sub-shelf melt parameterization might be highly uncertain, the calibrated range of $\gamma 0$ for this single parameterization using the MeanAnt calibration results in a 0.6–0.7 m range in sea-level equivalent mass change. This is modest relative to the multi-meter model-to-model differences for a given forcing scenario in ISMIP6-Antarctica-2300 (Seroussi et al., 2024)."

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

We have darkened the outlines of the shaded regions in Figures 16 and 18, which makes them much easier to distinguish.

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

We have changed this reference.

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

We have changed this to read: L434 "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 have changed this text to: L461 "The interaction between the ESM and hydrofracture forcings is the largest of the two-way terms, likely because some ESMs generate more surface melt than others, which makes them more likely to drive hydrofracture."

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

We have added: L 479 "…, 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 have added a specific reference to their Figure 4 in L 496 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 have added these caveats in L499: "…although the timing of retreat could still be sensitive to remaining structural and parameter uncertainties."

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 have added this as one possible explanation: L514 "potentially due to the more complex bed topography in this region."

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 have reiterated this here: L 520 "We note that this behavior cannot be attributed to initial condition uncertainty because all of our sensitivity tests begin from the same initial state with the basal friction coefficient rescaled based on the assumed value of q (see Section 2.3.3)."

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 have added clarifying statement to this effect: L581 "…although they acknowledge some model dependence of the required resolution."

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: L 581 "(including MALI)" 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.

We have added a statement about this and cited Hubbard et al. (2016) as documentation of the impact: L677 "…for instance through modulation of ice temperature (Hubbard et al., 2016)."

Figure C3: There is something strange with the colorbar.
We have fixed the overlapping colorbar tick labels.

Minor corrections:

l.51: 'can further' → 'can be further'

Corrected (now L 59)

l.566: 'hypothesize' → 'hypothesis'

This sentence was removed based on the other reviewer's comments.

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 now motivate our choices of experiments more carefully; we are more quantitative in determining the sensitivity of the model to these different choices; and we have consolidated and toned down 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 have updated the text of the abstract to be more quantitative. 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 have added 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 have removed the “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 have added an overview of those studies to the Introduction in L69–76.

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 have added this motivation to the Introduction in L89–98.

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).

We have moved the relevant paragraph to the beginning of Section 4.6.

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.

We have changed the reference to Seroussi et al. (2024), since this is more relevant to our study, and removed the mention of MICI (L33–34)

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 have clarified that ISMIP6 did not require good agreement to historical changes: L66 "…although historical accuracy was not a prerequisite for participating."

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 have changed this from "is likely already underway" to "could already be underway" (L178).

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 have added Figure B2 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 have clarified this in the text:

L232 "…which we take to be a comprehensive range of values for this parameterization."

L237: "We take this to represent a comprehensive range of this parameter, since q>1 is unphysical and q=1/10 is among the smallest values used in the literature and should behave similarly to smaller values (Gillet-Chaulet et al., 2016} and to a regularized Coulomb law for a given effective pressure (Joughin et al., 2019)."

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 have included 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 have reported the drift in the respective control runs where relevant and discussed 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.

New figure in Appendix: C1

Changes to text:

Re-write of first paragraph of Section 2.3.3 to include mention of control runs (L220–227)

L336: “In the control simulations, the SLC differs from our baseline by <30 mm (Figure C1).”

L356, L363, and L411 now reference Fig C1

L369: “ In the simulations with control forcing, the ordering of sensitivity differs from the forced experiments, with q=1 yielding ~350 mm SLC from the ASE by 2300, following by q=1/3 (~245 mm), q=1/10 (~150 mm), and q=1/5 (~115mm).”

L387: “The control simulation with fixed temperature loses mass equivalent to 180 mm SLC by 2300, compared with a net gain of 50 mm in our baseline control run (Fig C1).”

L393: “In the control simulations, both the Ross and FRIS are insensitive to the thermal solver, but the mass in the ASE increases by ~150% when using a fixed temperature field (Fig. C1).”

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 have rewritten this to be more quantitative: L329: "Our simulations with 5th and 95th percentile values of the melt sensitivity parameter, gamma_0, result in ~±40% and ~±13% differences from our baseline simulations for the CCSM4-RCP8.5-2300 and HadGEM2-RCP8.5-2300 scenarios, respectively (Figure 7)."

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.

Rewritten as "While the proper form of the sub-shelf melt parameterization might be highly uncertain, the calibrated range of γ_0 for this single parameterization using the MeanAnt calibration results in a 0.6--0.7 m range of sea-level equivalent mass change. This is modest relative to the multi-meter model-to-model differences for a given forcing scenario in ISMIP6-Antarctica-2300."

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. We have shown these simulations in Fig C1 and integrated them into the text as described above.

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. We have added "as the sector begins to run out of ice grounded below sea level" in L369.

3) Neither Pine Island Glacier nor Thwaites Glacier disappears under control forcing, but we have added "—including Pine Island Glacier —" in L378 to point out that a choice of q=1/3 has a large impact on PIG grounding-line position.

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 now 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 have the amount of drift due to each control simulation in Figure C1 and as mentioned in the text above. 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. Addressed as above by adding "as the sector begins to run out of ice grounded below sea level" in L369.

Section 3.5

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

We have rewritten the first paragraph of S3.5 (L385–400) to be more quantitative, and we have pointed out that the ASE grounding line retreats up to 200 km further by 2300 than in the baseline simulation, which is an effect not seen in the melt sensitivity tests.

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 have added: "We note that we have not examined the impact of vertical resolution on this effect; however, we do not expect a strong resolution dependence, as our non-uniform vertical layer thickness results in higher resolution near the ice base (Fig. 13 a,b)."

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. This is shown in the new Figure C1.

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 have clarified this in Section 2.3.3:

"However, as ISMIP6-Antarctica-2300 did not subtract control simulation drift from forced projections, the results we present here are also not drift-corrected."

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

We have made no change. 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 added "within the constraints of the MeanAnt calibration" in L459 to reiterate this.

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 have pointed this out in Section 4.6, L644: "The variance due to ESM and hydrofracture in our ensemble each reach 0.6 m by 2300 (Fig. 20), in close agreement with Seroussi et al. (2024); however, the ~0.4 m variance (combining q and γ_0 contributions in quadrature) in our ensemble due to parameter uncertainty is much less than the ~1.15 m variance due to ice-sheet model uncertainty in the ISMIP6-Antarctica-2300 ensemble."

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.

We have removed this text and substituted the explanation suggest by Reviewer 1: L461 "…because some ESMs generate more surface melt than others, which makes them more likely to drive hydrofracture."

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 have removed "of Thwaites Glacier" in L487.

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 and have made no changes.

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

We have made this more quantitative and backed of from "remarkable insensitivity" to "moderate sensitivity": L502: Our sub-shelf melt experiments (Figs. 7–8}) show moderate sensitivity to the chosen value of the melt parameter γ_0, with a larger relative impact (+40%) under than weaker CCSM4-RCP8.5-2300 forcing than under the stronger HadGEM2-RCP8.5-2300 forcing (±13%). While these differences are not negligible, the spread due to parameter uncertainty is quite small compared with the spread due to forcing uncertainty, in keeping with the low contribution of the melt parameter to the ensemble variance (Figs. 20–21).

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

We addressed this in Section 3.5 and have made no changes here.

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 have clarified the text: L588: "Previous work with MALI has shown that a resolution of less than 1 km is required for a converged solution on a marine ice sheet-type domain when using the sub-element parameterization of grounding-line position used here (Hoffman et al., 2018)."

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.

We have removed this sentence.

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 and have made no changes.

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 and perhaps stated too strongly. We have removed this hypothesis from the Abstract and Conclusions. Here, we simply change the grammar to be: L649 "Thus, it could be that a significant fraction of the large ice-model uncertainty reported by Seroussi et al. (2024) originates from initial condition uncertainty".

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 corrected the statement with respect to the Favier et al. paper, using "ASE" instead of "Thwaites Glacier". See our previous response about Feldmann & Levermann (2015); no changes made regarding this. We added a reference to the Bett et al. (2024) paper.

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

We have made this point more quantitative: L758 "Varying the ice-shelf melt sensitivity parameter γ_0 within the 5th--95th percentile range found by Jourdain et al. (2020) for the MeanAnt parameterization results in a ±40\% and ±13\% difference in SLC from our baseline simulations for the CCSM4 and HadGEM2 RCP8.5 scenarios, respectively. This difference is greatly exceeded by the factor of 3.5 difference between the CCSM4 and HadGEM2 RCP8.5 baseline simulations. Our 72-member ensemble corroborates this finding, with uncertainty in \gamma_0 never accounting for more than 10\% of the ensemble variance."

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 are 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, as we state in the item in L792.

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

Addressed above. We removed this from the Conclusions, as it made this hypothesis feel too strongly stated.

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 have now 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 have fixed the issue with overlapping text on the colorbar labels in Figure C3, which is now D3.

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.