Response to reviewers: The effect of the present-day imbalance on schematic and climate forced simulations of the West Antarctic Ice Sheet collapse

We would like to start with a general expression of gratitude to the reviewers for their positive attitude, constructive feedback and helpful suggestions to improve the manuscript. Please find our replies in blue.

Referee #1 (anonymous)

The paper presents a series of model experiments on the evolution of Thwaites Glacier over the next centuries using an ice sheet model forced by several climate models. The paper demonstrates that the historical imbalance of the glacier matters a lot for its future stability and its potential for 'collapse'. The authors also put forward a limit in global temperature increase for which the glacier could 'collapse'.

The paper is rather lengthy and could benefit from some trimming, which would make the message clearer. Especially the experiments of steady state versus transient initialization are of interest, followed by the forcing experiments. The introduction on the different ways of initializing models could be shortened, as the importance for the paper is to make the distinction between steady state and including imbalance.

Overall, I find this an interesting study that with some polishing and a few clarifications (see below) I would recommend for publication.

We thank the referee for their positive review and constructive comments that will improve the manuscript, and will discuss below how we implemented these comments.

Line 14: model -> models

Adopted

Line 14: what models are meant here. I guess climate models and not ice sheet models. Please specify.

We meant ice sheet models, forced by climate models. We clarify this in the revised manuscript, by rephrasing this sentence to: "....".

Line 19: Collapse occurs 58 times in the text, but it is never defined what is exactly meant by collapse of the ice sheet. Later on 'onset of collapse' is used, which also requires a clarification.

By "collapse" we mean accelerated deglaciation leading to considerable grounded ice mass loss. We identify the onset of collapse as the first timestep at which the bedrock ridge 50 km upstream of the current grounding line (as shown in Figure S6) becomes entirely free of grounded ice. This is a convenient definition since the deglaciation accelerates once the ice is free of this pinning point. We will clearly define these terms in the revised manuscript as: collapse (i.e. accelerated deglaciation leading to considerable grounded ice mass loss)

The introduction is quite long giving a complete overview of different methods of initialization. It is quite interesting in itself, but is not necessarily guiding the reader towards the core of the paper, i.e., that starting a historical simulation from an observed imbalance results in different response of TG compared to starting from steady-state conditions. It is not so much the way an initialization is done, but what the imbalance is that counts for understanding the remaining of the manuscript.

We will shorten the introduction and focus on the forced collapse of the WAIS with and without the present-day imbalance.

Line 91: Our null hypothesis is that the GMSL rise from the present-day mass loss rates is independent of the GMSL rise caused by an increase in ocean thermal forcing, i.e. that the present-day mass loss rates do not influence future forced projections.

Quite confusing. I would suggest to remove the mention to GMSL. It is about mass loss either caused by a given imbalance due to a grounding line retreat some time ago, or due to the current applied ocean forcing. I don't see how GMSL rise can be caused from thermal forcing (except thermal expansion, but that is not what you are talking about I presume).

We will rewrite these lines, also in reply to referee #2, to: 'With these simulations we will investigate the importance for the future AIS evolution of the current mass imbalance,compared to that of future changes in ocean temperature and SMB. Our null hypothesis is that incorporating the present-day imbalance does not influence the forced deglaciation of the WAIS.'

Line 100: Is a spatial resolution of 4km sufficient to guarantee grounding line migration (see for instance discussion in Pattyn et al., 2013). Maybe briefly state what is done to facilitate grounding line migration at such spatial resolution.

Yes, we think 4km resolution is sufficient to capture grounding line dynamics, especially with the grounding line parameterization of Leguy et al (2021).

'We run CISM on a uniform 4 km grid, justified below. using the grounding line parameterization from Leguy et al. (2021) which scales the basal sliding and melt rate proportionally to its grounded and floating area fraction respectively. Doing so, Leguy et al. (2021) showed that this resolution is adequate to capture grounding line dynamics using CISM and idealized marine ice sheet experiments, Lipscomb et al. (2021) showed that this resolution in combination with the scaling reduces the model result's grid resolution dependency when modelling the AIS.'

Line 109: The regularized Coulomb friction law was already used in Joughin et al (2019), and is based on the work of Schoof and Gagliardini. (Joughin, I., Smith, B. E., and Schoof, C. G.: Regularized Coulomb Friction Laws for Ice Sheet Sliding: Application to Pine Island Glacier, Antarctica, Geophys. Res. Lett., 46, 4764–4771, https://doi.org/10.1029/2019gl082526, 2019.)

We will rewrite line 109 as: 'We use the regularized Coulomb sliding law suggested by Joughin et al. (2019) and confirmed with laboratory experiments by Zoet and Iverson (2020)'

Line 121: The reference that marine sediments are likely more prevalent in submarine basins may be a bit outdated. There are more recent studies that have investigated the probability of sediment versus hard bed of Antarctica. See for instance:

https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2021RG000767 and https://www.nature.com/articles/s41561-022-00992-5

It shows a larger diversity of possible outcomes for regions lying below sea level.

These are good suggestions and references, and challenge the first-order approximation of Aschwanden et al (2013). We will add after line 121: 'The parameterization and associated values used to represent marine sediments in Aschwanden et al. (2013), which rely solely on bedrock elevation, are challenged by the more recent findings of Li et al. (2022), who show that the likelihood of marine sediment does not directly correlate with bedrock height. Incorporating the likelihood map of Li et al. (2022) into CISM and analysing the influence of marine sediments on marine-based ice sheets is beyond the scope of this study, but represents a promising direction for future work.'

Line 127: What are these parameters (Ho, tau, r L)? How do they influence your optimization? It is not defined what these parameters are about.

Thank you for pointing this out. We revise the manuscript in the following manner to explain these parameters:: 'In 1.4a, H_0, τ, r and L are scaling constants, used to adjust the relative weights of the different terms. Increasing/decreasing H_0 makes the optimalization less/more sensitive to ice thickness errors; increasing/decreasing τ makes the changes in C_l per timestep smaller/larger, increasing/decreasing r draws C_l more/less to the relaxation target C_{lr}; and increasing/decreasing L results in a smoother/spikier 2D pattern of optimized C_l. Their values were tested and chosen to represent AIS thickness and surface velocities well, with minimal drift. Table S2 shows their values.'

Line 158: See my remark of Line 100: is this the way grounding line migration is dealt with? Interpolation of friction within partially floating cells AND subshelf melt as well? It has been shown in Seroussi and Morlighem (https://tc.copernicus.org/articles/12/3085/2018/) that it increases the sensitivity of grounding line retreat big time. Some discussion is needed.

Yes, friction and subshelf melt in grounding-line grid cells are scaled to their respective grounded and floating area fraction. This choice is based on the results of Leguy et al. (2021), who showed that: (i) when using this so-called partial melt parameterization (PMP), the resolution dependency of grounding line movement in CISM is the lowest, (ii) grounding line parameterization sensitivity is model dependent, so what works for ISSM doesn't necessarily work for CISM, (iii) using a comparable simulation setup as in Seroussi and Morlighem (2018), results using a resolution of 4km capture grounding line migration accurately– and are comparable to those with a resolution of 2km and coarser. We agree with the referee that this point is not made clearly in the manuscript, and thank the referee for suggesting this reference. We will add to line 170:

"Schematic tests by Seroussi and Morlighem (2018) showed that applying basal melt in proportion to the floating area fraction can lead to an overestimation of grounding line retreat rates; they therefore discouraged the use of PMPs. However, Leguy et al. (2021) conducted similar schematic tests with CISM and found that using a PMP reduced the CISM's sensitivity to grid resolution more than the No-Melt Parameterization (NMP) recommended by Seroussi and Morlighem. In more realistic AIS applications of CISM, Lipscomb et al. (2021) found that the PMP produced a moderate sensitivity of grounding line migration rates to grid resolution, lower than the sensitivities to basal melt rate and basal friction parameterizations. These results suggest that the optimal GLP is model-dependent and that for CISM, 4-km grid resolution using a PMP is sufficient for modeling continental-scale ice sheets on multi-century timescales."

Line 165: isn't this not too much different than keeping the calving front fixed, as you probably need quite high melt rates to have the front retreating through melting alone.

Yes, it is. We will add 'In practice, this means that the calving front will retreat only when the basal melt rates are increased greatly, and for present-day conditions, the calving front position is fixed at its observed location'.

Line 181: Is an initialization of 10 ka enough for the temperature field to reach an equilibrium?

We start an initialization with the Robin solution to the temperature profile. During this 10 ka, the AIS geometry does not change considerably (with the exception of some outlet glaciers, mainly in the peninsula). The average temperature profile of grounded, floating and all modelled ice grid cells is shown below, which we will add to the supplementary material:

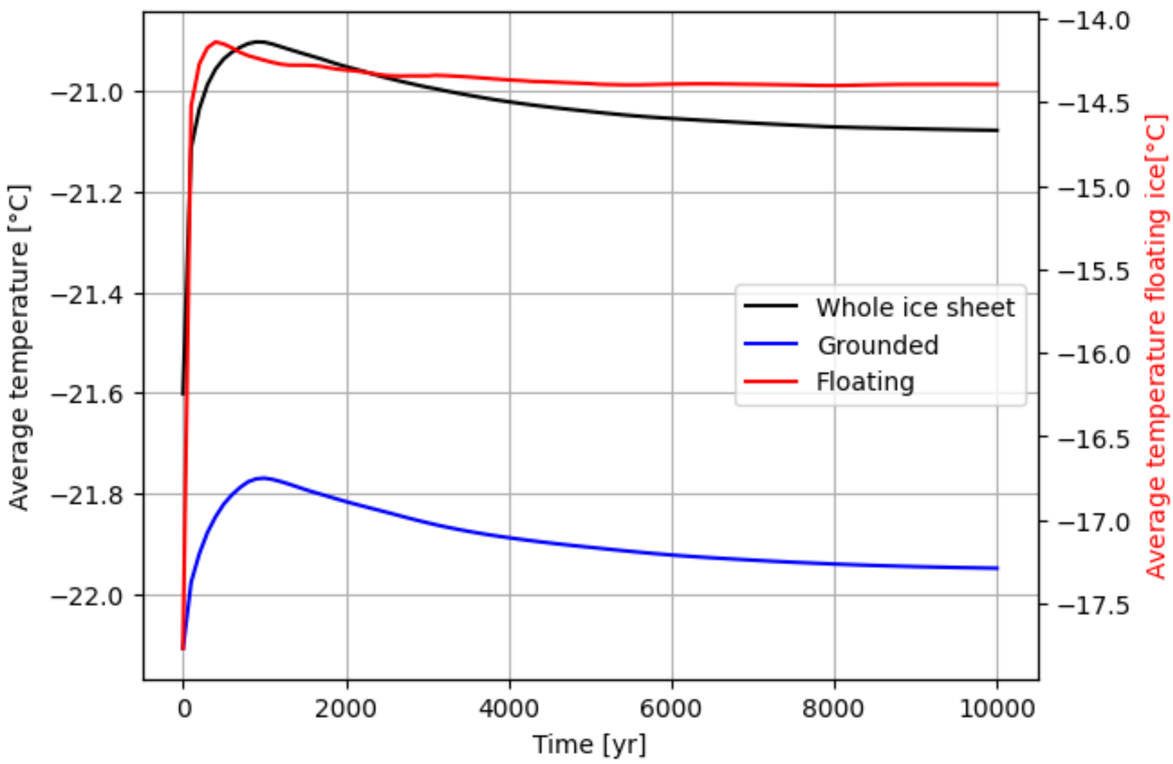


Figure S3. Grid point average temperature during the transient initialization (similar to the equilibrium initialization, not shown) for floating ice (red line, right y-axis), grounded ice (blue line, left y-axis) and all grid points containing ice (black line, right y-axis)

We will also add to line 189: Figure S3 shows the evolution of the average temperature during the initialization: it flattens out at 10 kyr, indicating that the ice sheet has reached thermal equilibrium.

Table 1: Overall, I found the figures in the supplementary material more of interest than the first few figures shown in the manuscript. Therefore, some information of figures S4 and S5 could be transferred to the main manuscript and replace table 1. One way of representing this is as in Martin et al, (2011) Figure 15 (https://tc.copernicus.org/articles/5/727/2011/tc-5-727-2011.pdf), so that different regions of the ice sheet/ice shelf system are represented.

This is a good suggestion. We will add the following figures to the manuscript, and add the values shown in Table 1 to the caption of the figures:

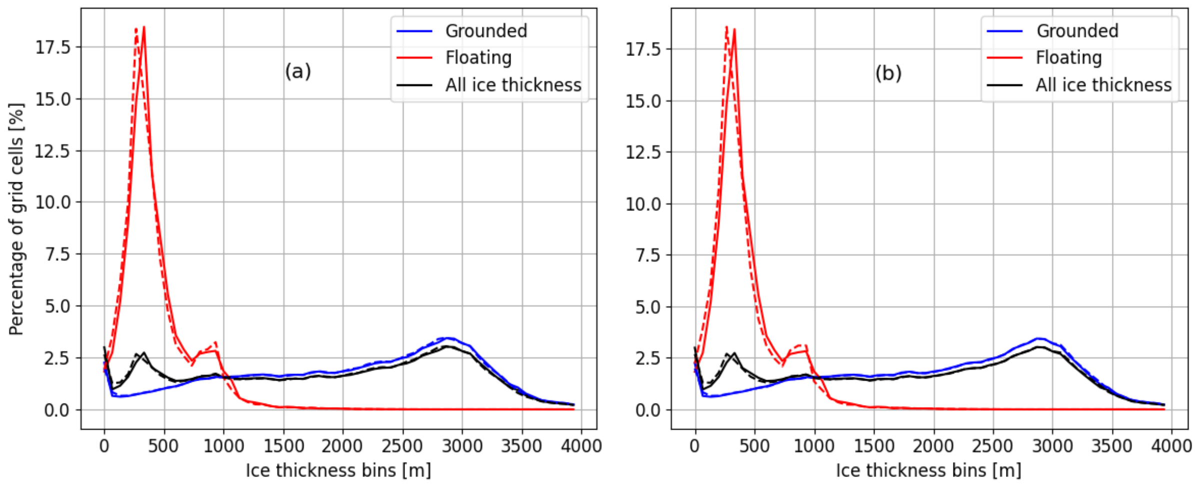


Figure 3. Binned ice thickness (m) for the observed (solid) and modelled ice (dashed lines). The present-day condition of the transient initialization is shown on the left, the equilibrium simulation on the right. For the transient initialization, the root mean square errors (RMSEs) for floating ice, grounded ice and in total are respectively 44, 31 and 35 meters. For the equilibrium initialization they are respectively 50, 23 and 30 m.

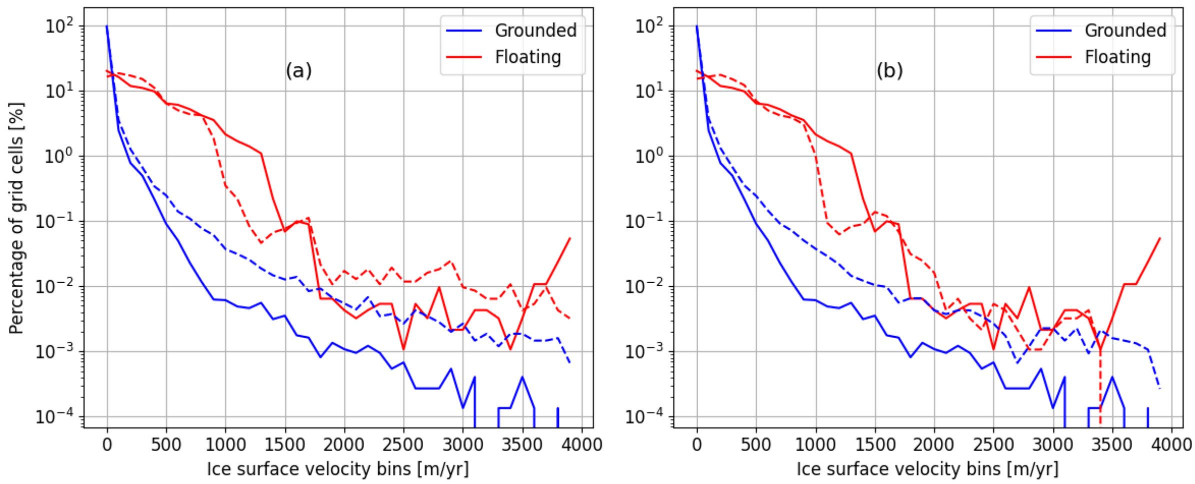


Figure 4. Binned ice surface velocities (m yr^{-1}) for the observed (solid) and modelled ice (dashed lines). The present-day condition of the transient initialization is shown on the left, the equilibrium simulation on the right. For the transient initialization, the Root Mean Square Errors (RMSEs) for floating ice, grounded ice and in total are respectively 201, 112 and 143 m yr^{-1}. For the equilibrium initialization they are respectively 202, 98 and 130 m yr^{-1}.

We will rewrite Section 3.1 accordingly.

Figure 3: Is not showing integrated mass loss, but mass contribution to SL in terms of GSLR and % of VAF. Mass loss also comprises that mass that is lying underneath floatation level.

That is correct. We will replace ‘integrated mass loss’ with ‘Loss of ice volume above flotation’ and check the whole manuscript where we now use "mass" if IVAF is a more appropriate quantity.

Line 307: I don't think that delta T can be considered an inverted parameter, as there is not inversion method used. Maybe use 'optimized'.

Yes you are correct. We will replace the words “inverted” and “inversion” by “optimized” and “optimization” respectively throughout the manuscript (we used similar wording after equation 1.4 (a,b) for example).

Figure 4 and Line 349: instead of pointing the readers to a supplementary figure S6 just to find out where a little line is drawn, it would be more informative to mention in the caption where this line is situated as a function of present-day GL position (i.e., XX km inland from the current GL position). This line is also defined as bedrock ridge and important as onset of collapse. What is meant by onset of collapse (see also remark on collapse in general)?

This is a good suggestion. We added to the caption: ‘Triangles indicate the timestep when the bedrock ridge approximately 50 km inland of the present-day TG grounding line ungrounds, depicted by the line AB in Fig S6’.

Line 400: I wouldn't call these simulations outliers. They are valid solutions for that given forcing. Just that these forcings are relatively low in melt and high in accumulation and therefore result in less mass loss than other forcings. This is not the definition of an outlier.

We agree. As the other reviewer pointed out, it is relevant that NorESM was run run until 2100, while the other four ESMs run till 2300.We will add the following text:

The simulations forced with NorESM output (RCP126 and RCP585) show much less mass loss than the others, as both NorESM simulations have less warming and more enhanced accumulation than the other ESM runs. The NorESM output assumes a constant climate after 2100, while output from the other ESMs assumes ongoing warming through 2300.

Referee #2 (Helene Seroussi)

Summary:

In this manuscript, the authors perform numerical simulations of the Antarctic ice sheet using the CISM model and test the impact of the initialization method (including or not the present-day thickness rate change during the model initialization phase) on the future simulations, based on both idealized forcings and projections using climate model conditions. They find that the inclusion of thickness change rate impacts the results mostly in the ASE as it is already experiencing large changes. They also conclude that the response to additional forcing is not linear, with the future changes compared to present-day conditions very different for the two initialization methods.

The paper is usually well-written and clear, and the figures appropriate. The conclusions are well supported by the results provided. However, there are number of aspects that can be improved to make the paper clearer, including the ESM and simulations used, the ability of the two methods to capture current thickness changes, the metrics to consider that a basin collapsed, and limit repetitions between the results and discussions. Below are some major and minor suggestions to improve the manuscript.

We thank the referee for their extensive review and constructive comments.

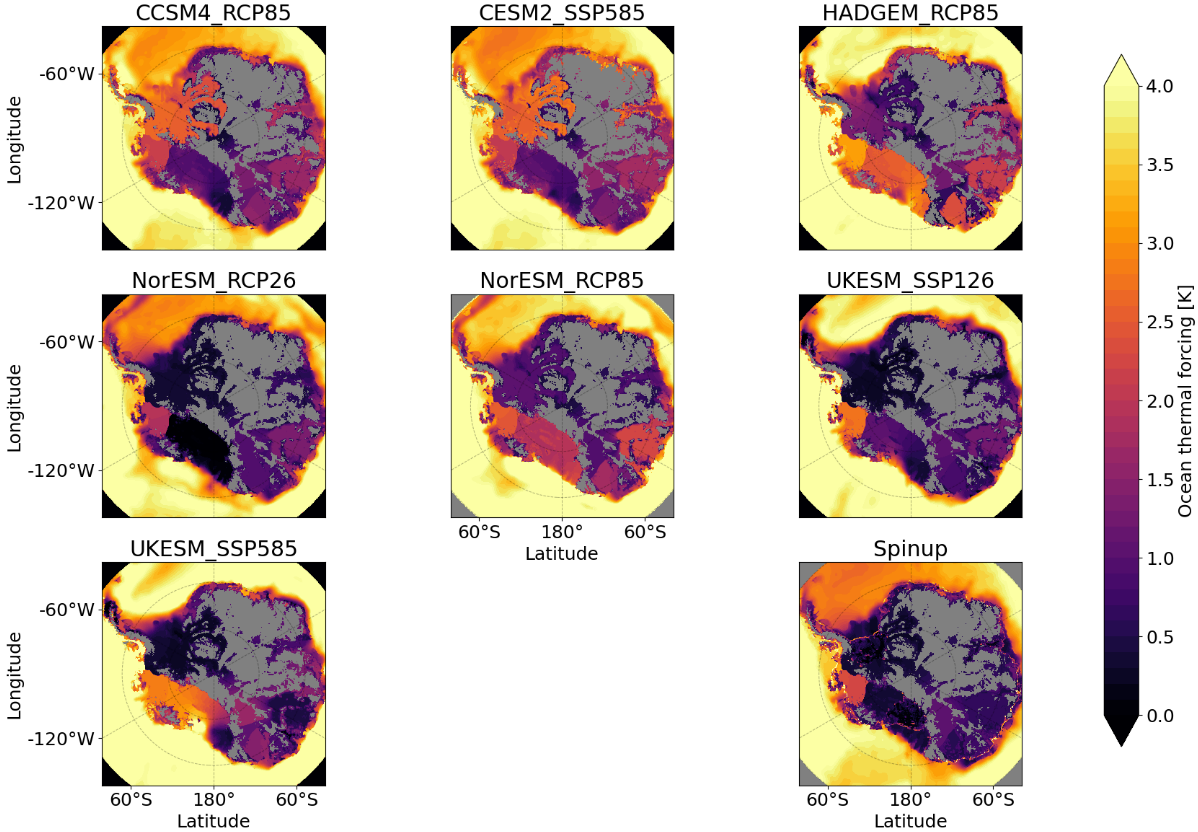
Major comments:

- The paper mentions 7 ESMs in several instances, but only 5 ESMs are actually used, from which 7 simulations are used. This needs to be clarified throughout the text.

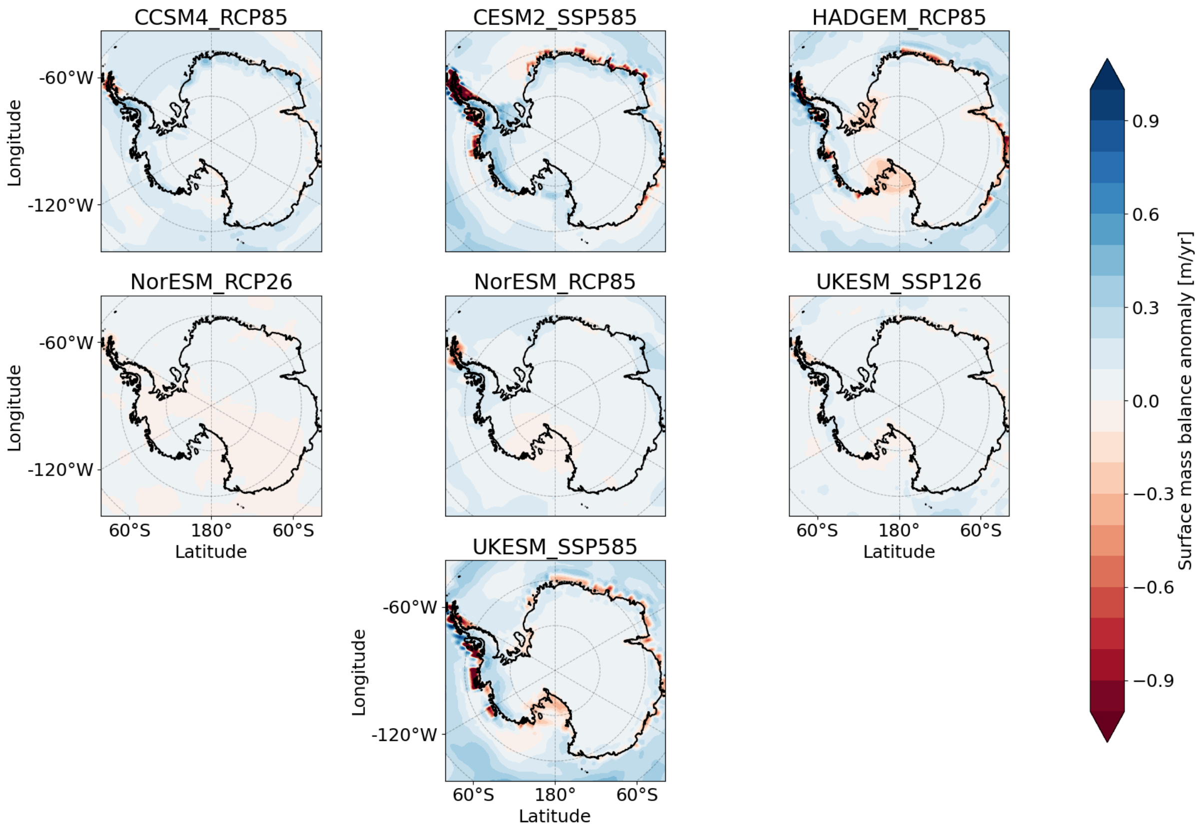
We will clarify this throughout the manuscript.

- Unlike the other ESMs, the NorESM simulations are only run until 2100 (and not 2300), so the patterns compared in figures 1 and 2 are shown for different times for the different simulations. Since high-emission scenarios are based on continued increases in emission after 2100, the temperature and overall conditions are very different at these two times. I am not asking to remove or redo these simulations, but this should be made very clear as soon as the simulations are described and reminded each time the patterns are compared and maybe additional figures are needed (for example for 2100).

We agree with this suggestion. We will add the following figure to the supplementary material, showing the thermal forcing of all ESM outputs for the year 2100:



We will also add the following figure to the supplementary materials, showing the surface mass balance anomalies of 2100:



We will add the following discussion early in the climate forcing section::

'Four of the ESMs (CESM4, HadGEN2-ES, CESM2 and UKESM) were run forward from the historical period until 2300, with high emissions continuing long after 2100 in the high-emission scenarios (RCP85 and SSP5-8.5). For these models, we fix the thermal forcing and SMB anomalies at the 2300 values during the last 200 years of the simulation. NorESM, however, was run only until 2100. After 2100, the NorESM forcing in our simulations is held fixed at late 21st century values. Thus, the NorESM forcing is not directly comparable to the other ESMs after 2100.'

- One of the initialization methods includes the dH/dt rate as a constraint, so an additional metric to measure how the two methods are able to match dH/dt at the beginning of future simulations would be useful. This would also balance the other metrics that are all about present-day conditions at a given time and not about rate of changes.

This is a very good suggestion. RMSE does not show a clear difference because many grid cells in the interior have a low mass change rate which is close to the error in the surface mass balance, and dominate the statistics, while it is the regions with large observed thinning rates where a match is important. In the spirit of our reply to referee #1, we made new graphs showing the binned difference between modelled and observed ice thickness and ice surface velocity. We will include the binned difference between the modelled and observed mass change rates as shown in the following figure:

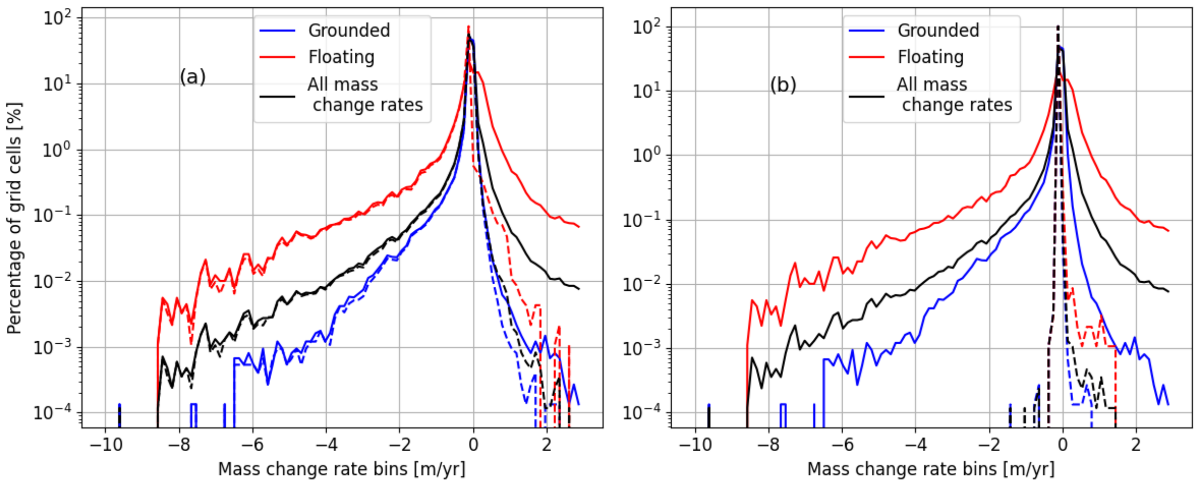


Figure 5. Binned mass change rates for the observed ice (solid lines) and the modelled ice (dashed lines). The transient initialization is shown on the left, and the equilibrium initialization on the right.

We will add the following text at ln 286:

'Figure 5 shows the binned mass change rates from the observations and both model initializations. The transient initialization reproduces the observed thinning rates almost perfectly. For positive rates, the modeled values begin to diverge from the observations in grounded regions where observed thickening exceeds 0.5 m yr^{-1}, which occurs in less than 0.01 % of the modeled grid cells. This discrepancy arises because the model can no longer reduce ice flux by increasing friction; C_c in Eq. 1.4 has reached its upper limit of 1. The surface mass balance and ice inflow into the cell become insufficient to reproduce the observed mass change. These cases of underestimated ice sheet thickening occur mainly in the interior of the EAIS, a region with little to no dynamic connection to theWAIS.

For floating grid cells, the model cannot reproduce observed thickening rates because ice accretion is not formally represented in the ISMIP6 basal melt parameterization. This restriction arises from the inability to specify where the positive mass change terms Eq. 1.9 should be applied, on top or on the bottom of a floating ice column.

The equilibrium simulation presents a different pattern. As expected, it fails to capture the observed present-day mass change rates completely, as it was initialized to be in equilibrium. Consequently, it exhibits minimal model drift, only 0.001 % of the modeled grid cells show a drift greater than 0.1 m yr^{-1}.'

In the collapse of the Ronne and Filchner-Ronne basins, the threshold for collapse initiation is 10% IVAF, but the collapse is considered for 100 mm GMSL. I don't understand why a similar absolute number is used for both basins? There is no reason why a similar number should be used for both basins since they don't hold the same amount of ice. So using also a percentage of the ice contained or IVAF for both basins would be more relevant and would lead to a better comparison of the timescale these two basins take to "collapse".

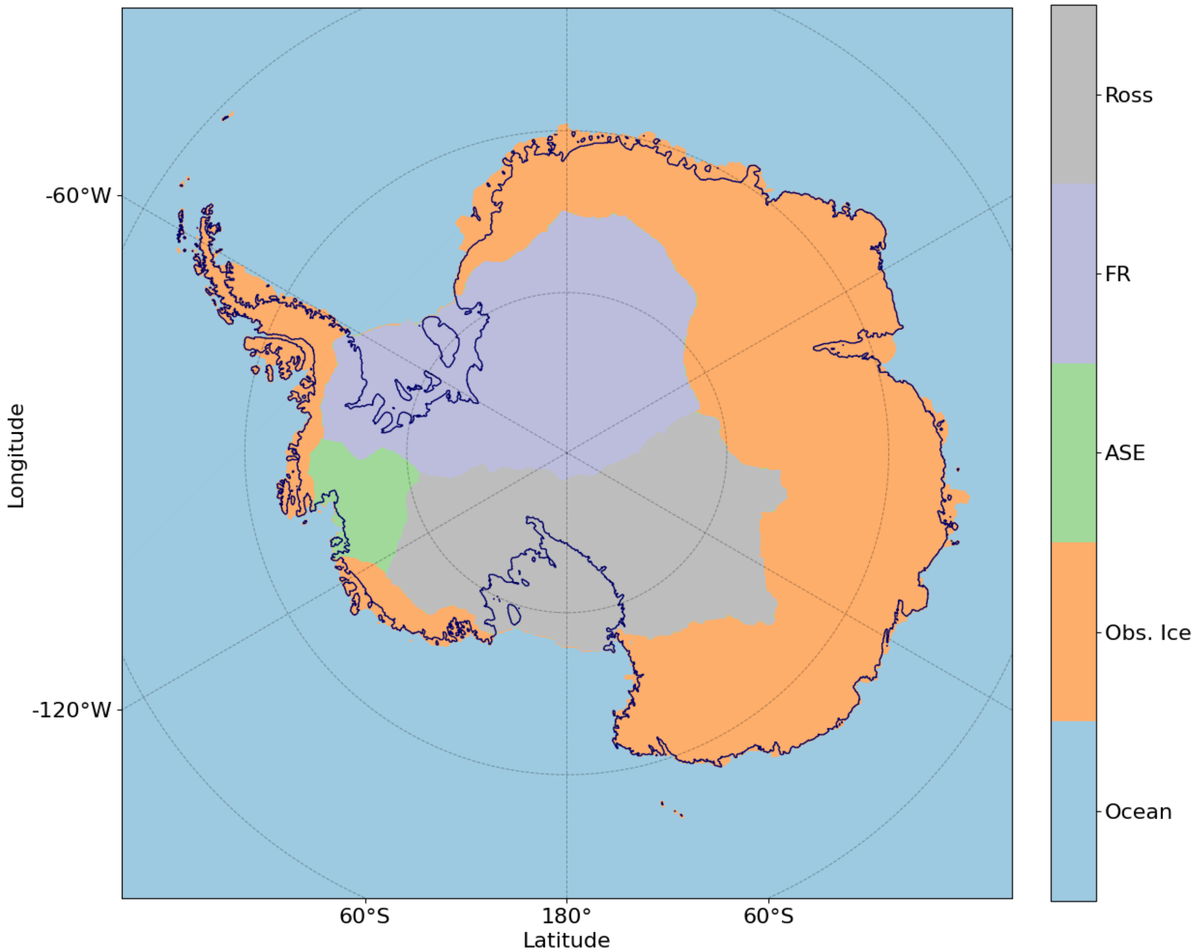
We agree that it would be better to replace the 100 mm GMSL with a percentage of IVAF. We choose a 5% target of IVAF loss, which is 58.7 cm SLR from Filchner-Ronne and 24.4 from

Ross instead of the fixed 100 mm GMSL limit before. This lengthens the delay, but the same general pattern emerges. Table 4 now looks like:

ESM forcing	Transient initialization delay (yr)		Equilibrium initialization delay (yr)	
	FRIS	Ross	FRIS	Ross
CCSM4_RCP85	145	280	130	305
CESM2_SSP585	130	195	120	210
HADGEM_RCP85	120	200	110	210
NorESM_RCP26	x	x	x	x
NorESM_RCP85	320	x	x	x
UKESM_SSP126	x	x	x	x
UKESM_SSP585	100	120	80	120

- S3 shows the extents used for the different basins. Ronne-Filchner does not include all of the Filchner ice shelf of the glaciers feeding in there. So it would be good to update this basin to include all of Filchner ice shelf and all the glaciers feeding it.

We will include the Filchner basins and update Figure S3 as follows:



We will also use these basins to calculate the new Table 4 in reply to the previous comment.

- The discussion and result sections are sometimes repetitive, with the discussion spending a lot of time repeating some of the results. Restructuring them would help to remove some of the repetitions. I think also the discussion could do a better job at comparing with previous studies looking at both the entire Antarctic ice sheets and specifically for Thwaites and Pine, whose collapse is discussed in details here.

We will rewrite the discussion section considerably in response to the minor comments.

Minor comments:

l.13: "projections": specify projections of what and over what period

We will replace 'projections' with 'Ice sheet model projections' and specify a time scale of the next few centuries.

l.14: "most models project some or severe" -> "while most models project some moderate to severe"

We will rewrite as "most models project moderate to severe mass loss"

l.16: "uniform ocean": the main text describes the perturbation to be only in the ASE sector

We will add '.. in the Amundsen Sea Embayment'.

l.18: "seven ESMs": there are only 5 ESMs from which outputs for 7 simulations are used. There are a lot of confusions about these ESMs, so make sure it is clear what is used. Also ESM has not been defined.

We will change this to 'seven forcing datasets produced by five Earth System Models (ESMs)'.

l.19: TG was not defined

We will replace 'TG' with 'Thwaites Glacier (TG)'.

l.20: GMSL was not defined so far. Also is that GMSL from Thwaites or also other basins?

We will rewrite as 'Global Mean Sea Level (GMSL) rise from the ASE'.

l.24: "coming 5 centuries": you want to be more nuanced, it does not happen for that long or at least not in all the cases.

We will remove this sentence as the next two sentences better explain the point.

l.29-31: this is the main conclusion, so start with this sentence before going into the details of the results.

We will move this sentence up to L.23.

l.35: "ice sheet modelled sea level" -> "modeled ice sheet contribution to sea level"

We will make this change.

l.37: "increases" -> "increases rapidly"

We will make this change.

l.43: "choice of the model" -> "choice of the ice sheet model"

We will make this change.

l.52: you also want to remind before this sentence that there are many other differences and then say a main difference is the initialization.

We will write: 'One of the most important model differences is the method used to simulate the present-day state of the AIS.

l.54: comma after AIS

No longer applicable, since we have rewritten this passage..

l.57: Larour et al. 2012 does not use the mass change rate: they use static data assimilation with the stress balance equation only and there is no additional term in the cost function related to the mass change. You want to check that it is really the case for the other citations.

We will remove Larour et al 2012, we will check the other citations when preparing the revised manuscript.

l.58-59: this sentence should go before the previous one, so first the generic method and then the specific about the mass change.

We will switch these sentences.

l.64: add a reference for the inconsistencies, for example Seroussi et al. 2011

We will add the following citation:
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011GL048659, which is Morlighem et al (2011)

l.66: "The other": the other what? Other models? '

We will rewrite as: 'Another ice sheet initialization method …',

l.74: remove "at grid point base"

We will do this.

l.77: remove "will"

We will remove 'will'.

l.80: I don't understand, you just said it is investigate the role of the present-day mass change a few sentences ago, so why are you saying here the present-day mass change rates was not included? You also want to explain what you mean by "forced" (in general I found this term a bit confusing throughout the text"

We meant to say that the forcing did not include future climate change. We will change this sentence to 'However, that study only showed simulations with a sustained present-day climate; it did not investigate how the present-day imbalance affects simulations that include future climate change (hereafter labelled as 'forcing')."););."We found not other ambiguities

l.81: again what does forced mean?

It means with additional atmospheric and oceanic forcing added compared to present-day, which we will clarify with the previous reply

l.81: "using firstly" -> "first using"

We will use this suggestion!

l.87: I don't understand why you say 21st century climate change since most climate model used simulate changes until 2300.

We will remove '21st century'.

l.90: rather than talking about a null hypothesis, I would rephrase to say that you are testing or investigating the impact of the initial conditions and the importance of capturing the present-day imbalance. You can even say it is common in intercomparison to subtract a control run with constant climate conditions to better compare model results, like what is done in ISMIP6

We will rewrite these lines, also in reply to referee #1, to:

With these simulations we will investigate the importance of the current mass imbalance for future AIS evolution, compared to the importance of projected changes in ocean temperature and SMB." We will no longer refer to a null hypothesis.

l.91: comma missing after i.e.

We will rewrite these lines as a reply to the previous comment and therefore skip the 'i.e.'.

l.94: "forced simulations" -> "future projections" or just "simulations" (or realistic simulations).

We have removed this sentence.

l.103: don't use higher-order here: this is typically used for the Blatter-Pattyn approximation and could be confusing for people used to that

We will remove 'higher order'.

Eq.1.3: I don't think you need the \beta u_b part here: this wrongfully gives the impression that the basal stress varies linearly with basal velocity, which is not the case as shown in the last part of the equation

We will remove it and also replace beta in Eq 1.1 and 1.2 with \tau_b,x and \tau_b,y , and rewrite Ln 106 – Ln 109 accordingly.

l.111-114: it is not clear what is done for the effective pressure N, can you add the equation if can be described easily with an equation or rephrase this sentence to make it more clear how N is calculated.

We will add the equations:

$$N = \rho_i g H \left(1 - \frac{H_f}{H}\right)^p \qquad (1.x)$$

where the flotation thickness H_f is given by

$$H_f = \left(0, -\frac{\rho_w}{\rho_i} b\right), \qquad (1.x)$$

with ρ_i the density of glacial ice, ρ_w the density of ocean water, g the gravitational acceleration and b the height of the bedrock

l.123: this linear description of the coefficient was derived for Greenland if it's from Aschwanden et al., 2013. Did you check how results from Antarctica compare to that and if a similar relationship can be used? Part of the reason I am asking is that we also find some good and simple relationship to describe the friction in Greenland but could never find a similar relationship for the friction for Antarctica.

Referee #1 pointed us towards the paper by Li et al (2022), who plotted the likelihood of marine sediment f for Antarctica and found that is is not well correlate with the bedrock height. We will add the following discussion to Ln 123:

The parameterization and associated values used to represent marine sediments in Aschwanden et al. (2013), which rely solely on bedrock elevation, are challenged by the more recent findings of Li et al. (2022), who showed that the likelihood of marine sediment occurrence does not directly correlate with bedrock height. Incorporating the likelihood map of Li et al. (2022) into CISM and analysing the influence of marine sediment on marine-based ice sheets is beyond the scope of this study, but represents a promising direction for future work.

l.130: mention what this observational dataset is (climatology combining all observations over the past couple of decades or something like that)

We will add: "The thermal forcing is derived from a climatology based on Southern Ocean observations of the past few decades (Jourdain et al., 2020)."

l.141: "matches": what is the target of this match? Extent? Or thickness? Or something else?

We will add 'thickness'..

Eq.1.7: what does this quantity represent?

When f_{float} is positive, it is equal to the distance between the ice shelf base and the ocean floor and hence, the ice is floating. When f_{float} is negative, it can be interpreted as the ice thickness change needed for the ice column to become floating. We will add this to Ln 156.

l.161: references for the calving laws: I would go back and cite the papers that originally introduced the different calving laws.

We will do this and add Choi et al. (2018) and Levermann et al. (2012).

l.167: the "ice shelf thinning" is included in your calving, it is actually the only thing included since calving happens when the ice gets thin enough. The impact of large stresses or strain rates on the other hand are not included.

We agree and will change this to 'The ice shelf front can retreat only through ice shelf thinning'.

l.172: "initialization" -> "initial state"

We will do this

Eq.1.8: there is no need to introduce a variable F for the flux divergence, since it is simply H \bar{u}. It is more common and will make things easier to simple use that instead of introducing a new variable.

Good suggestion, we will replace F with H \bar{u}.

l.184: I think that on top of looking at the long-term drift, it would be useful to also look at the instantaneous thickness change at the beginning of the projections (using present-day conditions for the forcing) to assess the impact of including the mass change term and seeing how the two runs compare to observations. It would be a very useful and complementary metric to the thickness and velocity misfit and you already have the data used for dH/dt. So I would compare the mass change and add the results in Table 1.

We will add a figure comparing the instantaneous dH/dt in the two runs to the observations, as suggested also by the first reviewer. .

l.184: what does "without forcing mean"? you always need to have forcings at the ice/atmosphere and ice/ocean interfaces? So do you mean that the forcing is constant or unchanged compared to the initialization phase?

We will rewrite as 'without additional forcing relative to present-day conditions'

l.187: For which run is the 0.04% change? I would put number for both cases

Both simulations; we will add this: …results in a WAIS change in IVAF over 1000 years of 0.5% for the Transient initialization and approximately 0% for the equilibrium initialization. For the whole AIS, the model drift in the transient initialization is 0.05% of the IVAF and -0.05% for the equilibrium initialization

l.191: "Continuation simulations" is not very clear. Maybe experiments or future projections would be a better description of this section.

We will use 'Future projections'

l.197: there are not seven ESMs used in this study, only 5 ESMs from which outputs from 7 simulations are used, this should be made very clear right away. So this paragraph needs to be clarified. Maybe it would be easier to have a short section of idealized experiments and another one on projection experiments in which you talk about both the projections and the ESMs (so merging with section 2.4).

We will change this to 'seven forcing datasets produced by five different ESMs'. We will also combine sections 2.3 and 2.4, as suggested. We will add 2.3 as a paragraph on the idealized experiments as follows:

2.3 Idealized schematic warming scenarios

To assess how sensitive the simulations starting from the transient initialization are to imposed ocean warming, and to compare this with the sensitivity of simulations starting from the equilibrium initialization, we conduct a set of idealized experiments in which ocean temperatures in the ASE region are systematically increased after initialization. Each experiment is then run for 1000 years to allow for a possible WAIS collapse. Ocean temperatures are incremented from 0 K to 2 K in 0.2 K steps, yielding ten simulations for each initialization

l.200: explain how the repeating is done.

We will rewrite as 'After 2100, the NorESM forcing from the late 21st century (2080–2100) is repeated for the rest of the run.'

l.200: "one where the forcing" mention which one right away (maybe that would be easier after the ESMs are mentioned hence the suggestion to slightly reorganize).

We have rewritten this section, explaining right away the difference between NorESM and the other runs.

l.201: "last datapoint": what is the impact of using just 1 datapoint as opposed to an average over 10 or 20 years? If that year has anomalously high or low SMB (or ocean thermal forcing) it might bias the results. Without necessarily rerunning all the experiments with an average value, you can add some discussion (here or in the discussion section) about how 2300 compares to over years around that time.

We agree, and we will add this to the discussion as follows:

In this study, we employed seven forcing datasets derived from five ESMs . Although forcing data were available only up to the year 2300, our simulations extended to 2500. To enable CISM to run through 2500, we held the final available forcing fields (from 2300) constant over the years 2300–2500. This approach could introduce biases if conditions in 2300 differ substantially from those in the surrounding decades. To assess this, we compared the 20-year mean (2280–2300) thermal forcing and surface mass balance (SMB) anomalies, shown in Figs S9 and S10, with the values used for the year 2300 (Figs 1 and 2). For ocean thermal forcing, the year-2300 fields closely match the 20-year mean across all datasets. SMB anomalies show minor differences between the year 2300 and the average of 2280-2300, most notably in the NorESM_RCP26 dataset over the Weddell Sea, but overall the spatial patterns and magnitudes remain highly similar. These comparisons suggest that substituting the 20-year mean for the single year 2300 would not affect our results considerably.

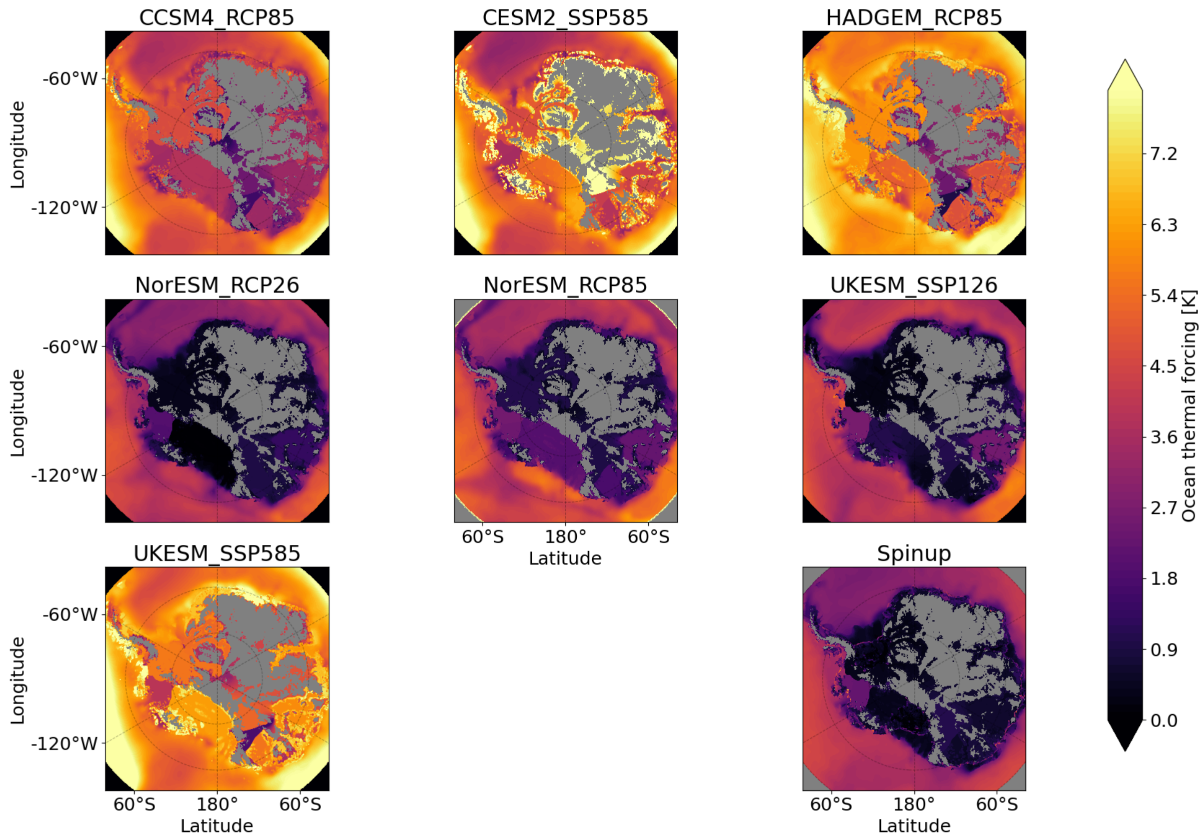


Figure S9. Thermal forcing ($TF_{base} + \delta T$ in Eq. (1.5)) from the five ESMs. Thermal forcing averages for the years 2080–2300 (except the present-day) are shown for a depth of -500 metres. Cells with bedrock above sea level are shown in grey.

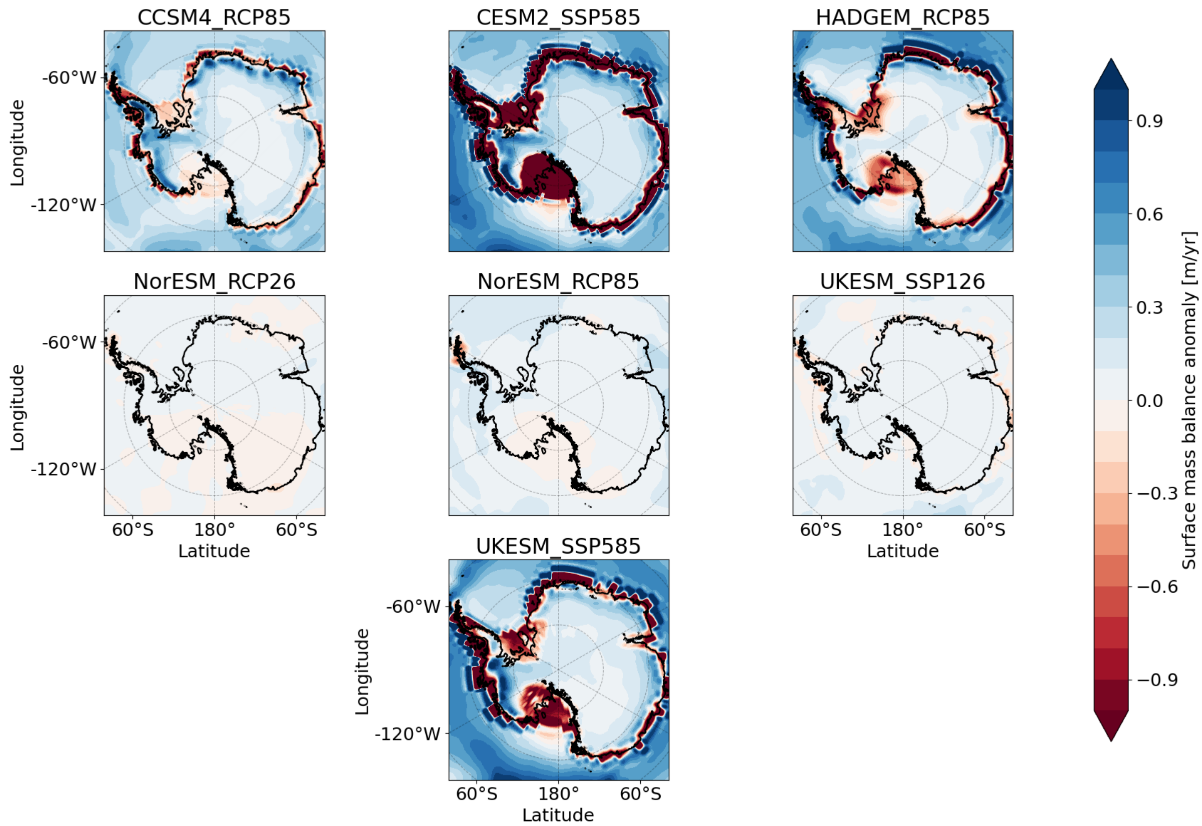


Figure S10. Surface mass balance (SMB) anomalies simulated by the five ESMs. The average of the years 2280-2300 is shown. The observed grounding line position is shown in black.

l.212: how about the other two climate models? What are their ECS and how does it compare with rest of the ESMs?

We will add the exact values from Table 1 from Meehl et al 2020.

Fig.1: rather than spin-up, I would use present-day or observations or climatology.

We will change this to ‘present-day’.

Fig.1 caption: not all models have values for 2280-2300 since NorESM simulations stop in 2100. Make sure this is very clear throughout the text, since it explains to a large extent the patterns of SMB and ocean thermal forcing (not as much warming in high-emission scenarios by 2100 compared to 2300).

We will add this to the text, as well as additional figures with the 2100 values, as mentioned in a reply to one of the major comments.

l.219: “as” -> “similar to”

No longer applicable, since we have rewritten this passage..

l.226: Why do you show the average of the last 20 years (2280-2300) on the figure but then use only the last year (2300) to continue the projections? That seems a bit inconsistent.

In Figure 1, we will replace the averages with the 2300 values used to force the ice sheet.

l.235: NorESM simulations stop in 2100 that will explain a lot of the different for this model, especially for the high-emission scenario. Same for l.239: the only high-emission scenario that is different is NorESM, to a large extent because it is not comparing for the same year so the comparison is biased.

We have added text explaining that the NorESM runs stop in 2100; please see the major comment above.

Fig.2: here again, NorESM RCP585 stands out because it is warming by 2100 and not 2300 and the warming continues after 2100 in this scenario.

We will remove the NorESM panels from this figure, while adding figures of 2100 anomalies to the supplementary materials.

Fig.2 captions: what do you mean by "modelled" SMB here? Is that the SMB values used during the initialization?

We meant the SMB obtained from RACMO, which is a modelled product. We will clarify this in the caption.

Fig.251: You should also add a metric to compare the dH/dt trend at the very beginning of the future projections

This is a very good suggestion, also highlighted by the other referee. We will add a figure as addresses an earlier major comment.

l.256: Goelzer et al. 2020 is a paper on Greenland, so I am not sure how much we can compare velocity and thickness errors for the Greenland and Antarctic ice sheet simulations?

We agree and will remove this reference.

l.259: comparing the ice fluxes at the grounding line is a great idea. Could you add that to table 1?

We thank the referee and will add those as follows to Ln 261:

Using the same flux gates and calculation as Van Den Akker et al. (2025b), the observed ice fluxes in this study are at PIG and TG are 31.1 km2 yr-1 and 26.1 km2 yr-1, respectively. In the equilibrium initialization, those fluxes are 17.0 km2 yr-1 and 17.4 km2 yr-1 , hence much lower than observed. Using the present-day mass change rates and the initialization as described in Section 2 results in PIG and TG fluxes of 30.4 km2 yr-1 and 24.5 km2 yr-1. , which are in good agreement with observations.

l.270: you still have the same target data for both initialization, so I am confused why the error would be so much larger in one case if the same target is included in both?

Because the model cannot decrease the ice flux enough on the EAIS, to compensate for the low SMB and small positive mass change rates. Our inversion is bounded in the sense that the friction coefficient and ocean temperature correction are capped. In some areas, this causes the model to under/overestimate the mass fluxes with maxed or minimized out friction or temperature perturbation. We will add text explaining this model limitation, as an accompanying discussion for the new figures with the binned thickness, velocity and mass change rates differences.

Table 1: add a metric for the dH/dt change at the beginning of the projections, and one for the grounding line mass flux.

We will replace table 1 with three new figures, in a joint reply to referee 1, and we will elaborate on the grounding line fluxes in reply to the comment on L259.

Fig.3 caption: you have a and b on the panels, so I would use that in the caption instead of left and right

We will replace left and right in the caption with a and b.

l.304: "broader" –> "entire"

We will use this suggestion

l.305: "earlier": what are you referring to? Earlier than what?

The greater the warming, the earlier the collapse. We will clarify this point.

l.306: "increases linearly": it is not clear on Fig.3 that you are discussing here that there is a linear increase. It looks like this is more clear on Fig.5, so I would reorganize this part.

We will move this statement to Figure 5, where we show the linearity between added ocean warming and the maximum GMSL rise rate.

l.306: "more pronounced": what do you mean by that? In one case, it is steady-state (with no temperature change) and in the other on it changes because of the initialization method, so it sounds obvious that there will be a large difference in the case of the steady-state. But maybe I am missing something.

We mean the difference between warming and the control situation: the control simulation from the equilibrium initialization does not feature a WAIS collapse, all other simulations do. They both react differently to a difference (increase) in temperature. We will change 'more pronounced' to 'larger' in the text.

l.314: remove period after "Fig. S6)"

We will do that.

Fig.4: add a legend with the blue and red lines as well as the blue and red stars

We will do that.

Fig.4 caption: use a and b instead of left and right since you already have then on the figures

Thanks for the suggestion, we will do that.

l.326: I still think the null hypothesis is confusing, so I would reformulate to say something like: investigate the impact of different initial conditions

We have changed this to “To investigate the impact of different initial conditions”, removing the reference to a null hypothesis.

l.328: “simulations. These results are shown in Figure 4.” -> “simulations (Figure 4)”

We will make this change.

l.331: remove “already”

We will remove ‘already’.

l.333-334: rephrase this sentence: The response to warming conditions is impacted by the modeled initial conditions …

We have simplified the text:: ‘The simulations starting from the transient initialization have a much different response than those starting from the equilibrium initialization.’

l.337: remove “than in equilibrium”

We will do this.

Fig.5 caption: Overall title would be more representation with something like: Impact of ocean warming on collapse and maximum GMSL …

We will rewrite this: ‘Impact of ocean warming on the collapse onset and maximum GMSL rate’

Fig.5 caption: “added” -> “additional”

We will do this.

l.352: “an ice dynamical limit” -> “a limit on the onset of collapse”

We will make this change.

l.352: “2 degrees”: if the collapse would happen earlier with more than 2 degrees of warming, then that’s not the definition of a limit: it would the case if you had an asymptotic value regardless of how warm the ocean gets

We agree. We did test higher warming scenarios up to perturbations of 10 K and we could not find a real asymptote. We will remove this from the manuscript.

l.354: “rate of sea level rise keeps on intensifying”: is that also from the same ASE region or is there also mass loss from other sections or glaciers outside of Pine Island and Thwaites?

This is only for the the ASE region, which is the only place where we add ocean warming. This means that other glaciers in the same basin, like Smith, Pope and Kohler, can also be subject to a faster collapse, which eventually will 'attack TG from the side'. We will clarify this in the text.

l.362: "Until 2150": how do you find 2150 as the transition from the initialization to the forcing dominated uncertainty?

Because the differences between transient and equilibrium initializations are relatively the largest before 2150, after that, the forcing brings out large GMSL for both initializations. We will clarify this in the text as follows:

'The additional sea-level contribution can come not only from PIG and TG, but also from the neighbouring Pope, Smith and Kohler glaciers, which are in the same basin as TG and therefore receive the same ocean warming.'

l.363-364: Rephrase this sentence, it is very confusing.

We will rewrite this sentences as: After 2150, all simulations project eventual WAIS collapse, in most cases before 2500.

l.365-366: remove "blue" and "red" as there colors do not reflect the types of simulations.

We will do so.

l.366: "collapse": what do you use to consider that the ASE collapsed?

"Collapse" refers to accelerated deglaciation leading to considerable grounded ice mass loss. We identify the onset of collapse as the first timestep at which the bedrock ridge 50 km upstream of the current grounding line (as shown in Figure S6) becomes entirely free of grounded ice. This is a convenient definition since the deglaciation accelerates once the ice is free of this pinning point.

We consider the ASE collapse to be complete when the ASE basin is ~80% empty and the ice loss rates flatten. This is typically when there is ice left in only a few places on bedrock above sea level, and the marine basins are empty.

We will clearly define the terminology in the revised manuscript.

l.371: "The forcing then rules the response" rephrase or remove that sentence

We will remove this sentence.

l.379: the UKESM simulations actually has the largest/fastest AIS mass loss

Thanks for this sharp eye, we will add that UKESM has the largest and fastest AIS mass loss, and that HADGEM shows the largest/fastest retreat in the ASE region.

l.383: remove "in our simulations"

We will remove this

l.386: “profoundly”: quantify what you mean in terms of percentage difference or something like that.

In 2500 for the whole AIS, the difference between transient and equilibrium starting runs is about 1-2 % IVAF change. In a regional ASE simulation, or for the whole AIS before 2500, the difference can be up to 10-20% IVAF, a tenfold difference. We will add this to the text.

l.401: “only one”: there are several lines that are above the extrapolated mass change in the beginning of the run and until 2250

We will remove this paragraph, since the NorESM differences will have been explained earlier.

l.402: “absence of any future forcing”: is that future forcing or continuing with the same constant present-day conditions? Also it’s not possible to not have a forcing, there are always prescribed conditions, even if they are zero SMB and melt.

Yes, this is true. We have removed this text.

l.404: “run towards 2100” -> “run only until 2100 and conditions from 2100 are repeated after that”. Also this needs to be made clear much earlier, as soon as you describe the EMS simulations

We have removed this text.

Fig.6 caption: remove “on the right”. Also “unforced” is not clear, and you always have a forcing, even if it’s zero.

We will remove ‘on the right’, and we will replace ‘unforced’ throughout the manuscript with ‘constant present-day forcing’

l.420: “high-warming” -> “high-emission”

We will change this.

l.424: “respectively disappear and completely unground entirely by 2500” –> “completely ungrounded by 2500”

We will change this.

l.426: “does” -> “do”

We will change this.

l.427: “simulations” -> “in simulations”

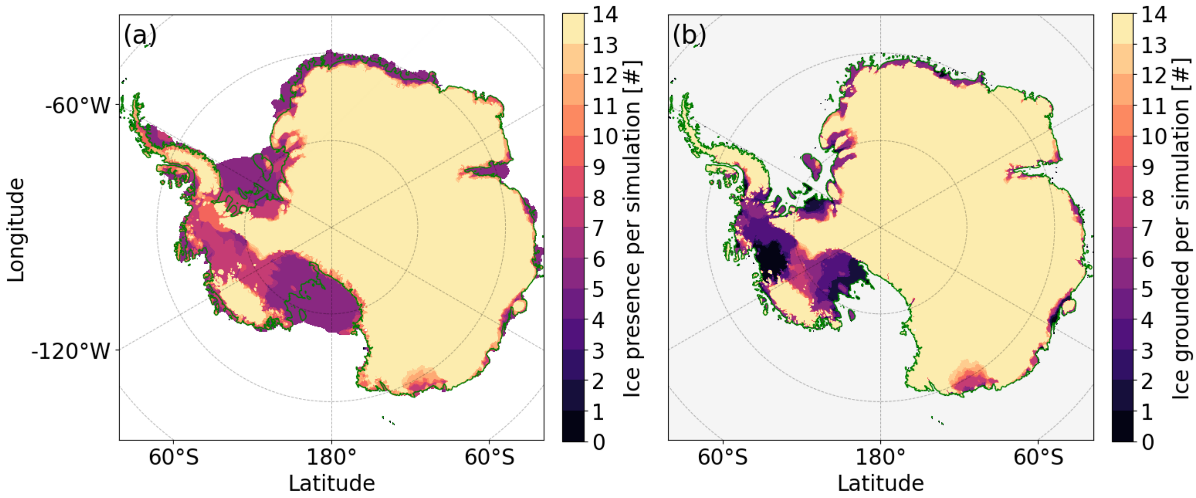
We will change this.

l.428: “deglaciate” -> “also deglaciate”

We will change this.

Fig.7: remove the parts that are already floating at the beginning of the experiments on panel b, so it is more clear what areas do unground

We will do this:



l.441: "considerably": quantify what you mean

We mean that the mass change rates are close to zero there, so adding them does not matter. We will remove 'considerably'

l.444: If you want to use the abbreviation FR, I suggest you introduce earlier in the paper and consistently use it throughout the text

We will do this.

Fig.8: I am not sure why different colors are used for the three basins because it does not add much. Instead I would continue using the same colors for the different ESM simulations for consistency with the other figures

This is a good suggestion, we will do this.

Fig.8: Why do you show the thermal forcing for PIG, thwaites and Larsen when the other panels show Filchner-Ronne, Ross and ASE? Try to keep it consistent for the three metrics reported

We do not show the thermal forcing of PIG, Thwaites and Larsen. The top row is Ross, the middle Filchner-Ronne and the bottom row the ASE. To clarify, we will change the caption to:

Ice volume above floatation (IVAF), floating ice mass and average thermal forcing for Ross (upper row), Filchner Ronne (middle row) and ASE (bottom row) per basin until 2300. The left column shows the IVAF as a percentage of what was present at the initialization; the middle column shows the total mass of floating ice per region; the right column the average thermal forcing at 510 m below sea level at the calving front.

Fig.8 caption: "per basin": mention which basins you are looking at instead.

Yes, we will add: Filchner-Ronne, Ross and ASE.

Fig.8 caption: remove the last sentence since the figure is only until 2300

We will do this.

l.468: "100 mm of GMSL": why use an arbitrary number of GMSL and not a fraction of the basin contribution to GMSL as is done for the initiation? I don't know the difference between the Ross and Filchner-Ronne basins, but there is no expectation that they hold the same amount of ice. That could also make changes in the results reported in table 4 and give insights into the timing to deglaciate these two basins, and how they compare. Finally, why not include also ASE in table 4?

We agree with the reviewer that it would be better to replace the 100 mm slr to a percentage of IVAF. We choose a 5% target of IVAF loss, which is 58.7 cm SLR from FR and 24.4 from Ross instead of the 10 fixed limit before. This makes the delay longer, but the same general pattern emerges. Table 4 now reads:

ESM forcing	**Transient initialization delay (yr)**		**Equilibrium initialization delay (yr)**	
	FRIS	**Ross**	**FRIS**	**Ross**
CCSM4_RCP85	145	280	130	305
CESM2_SSP585	130	195	120	210
HADGEM_RCP85	120	200	110	210
NorESM_RCP26	x	x	x	x
NorESM_RCP85	320	x	x	x
UKESM_SSP126	x	x	x	x
UKESM_SSP585	100	120	80	120

We excluded the ASE from this analysis, since buttressing in this region currently has little impact on the ice dynamics.

l.476: remove "typically"

We will do this

Fig.9 caption: why "extra"? I am not sure I understand what you mean here, what are you comparing it to?

We will remove 'extra'. This is the mass loss from the ESM-forced simulation with the present-day forcing continuation subtracted.

l.520: "present" -> "show"

We will do this

l.522: "considerably": add some numbers to quantify

We quantify this as

‘In 2500 for the whole AIS, the difference between transient and equilibrium starting runs is about 1-2 % IVAF change. In regional ASE simulations and before 2500, the difference can be up to 10-20% IVAF, a tenfold difference.’

l.530: “does not occur before 2100”: you want to compare your numbers with previous studies such as Robel et al. 2019, Joughin et al. 2014, etc.

We will do this and add it to the discussion

l.532: “extreme warming”: you see already a very large difference, and 2 degrees is quite a lot, so I would rephrase this sentence

Since we cannot rule out that there is a limit, we will remove this sentence.

l.533: Why does the peak rate continues to increase? Where is the ice coming from and what is the mechanism that leads to this continued increase?

This is likely the partial melt parameterization feedback, which means that a cell containing the grounding line receives basal melt according to the percentage of the cell that is floating. This holds for small to modest increases in basal melt rates but not for extreme increases.

l.536-539: rephrase this positively to highlight the new finding (also will has three “l”s in the manuscript)

We will rewrite this as ‘Hence, we we find that including the present-day mass loss rates increases the modelled GMSL rise from the AIS, mainly in the short term (before 2100) when mass loss is dominated by the ASE.’

l.551 and following: this is all repeating information that are already mentioned in the results section, so try to go beyond the results reported and compare with other studies for example

We will move the differentiation of the thermal forcing to the result section, and use the space to compare our study with other studies here. We will move up and rewrite the paragraphs on global mean sea level rise and add comparisons with different studies. We will remove repetitions on the importance of the ESM forcing versus the importance of the mass change rates.

l.570: missing parenthesis

We will add them

l.576: remove “which “

We will do this

l.588: missing comma after i.e.

We will add this

l.590: “our ESM” -> “the ESM”

We will do this

l.591: "decreasing": for what period is that? It looks like it first increases and then decreases?

Over the whole AIS, so from present-day until 2300. We will add this.

l.602: "use" -> "improve this aspect by using"

We will add "better represent thermal forcing by using…"

l.607: "from 2100": as mentioned earlier, this should be made clear from the beginning

We will make this clear earlier.

l.615: "calving ceased entirely": I don't understand, it can continue retreating if the ice is thin enough to allow ice front retreat

We meant that when ice front retreat is caused by basal melting and not by frontal ablation of ice shelves, the calving flux is effectively zero. The ice front retreats, not due to calving but due to basal melting. We will clarify this in the revised manuscript in the description of the calving algorithm.

l.617: "modeled front positions" -> "ice thickness". This seems contradictory to the calving scheme described based on ice thickness only

We will remove "rather than relying solely on modeled front positions", since this limitation has already been mentioned.

l.622: "by" -> "with" and remove "up"

We will do this.

l.628: "whether or not include the present-day" –> "to include or not the present-day"

We will rephrase this.

l.630: "with" -> "over"

We will do this.

l.633: "depedent" -> "dependent"

We will do this.

l.636: "shelfs"???

We will add 'ice-shelves'. Apologies for this typo.

Fig.S3: why is Filchner ice shelf and the glaciers feeding it not entirely in the Filchner-Ronnee domain?

We will update this figure as explained to an earlier comment, and redo the basin calculations accordingly.