Reply to Reviewer 1

We thank Reviewer 1 for their careful reading of the manuscript and their constructive comments. We have carefully addressed all comments point by point in the sections below. For each comment, we indicate the proposed changes to the manuscript text, with reference to the relevant line numbers of the original submission.

Major comments

On the use of a single GCM (CESM2) and its limitations

We agree that the use of a single ESM to force the RCMs limits the representation of the full range of plausible future climates. We propose to add the following sentence in the final paragraph of the Introduction to explicitly acknowledge this limitation:

"However, forcing the RCMs with a single ESM (CESM2) limits the representation of the range of plausible future climates and associated uncertainties (e.g. Kittel et al., 2021)."

Regarding the model naming, we acknowledge that CESM1 and CESM2 are distinct models. In the current manuscript, we referred to CESM2 as "CESM" for consistency with the naming convention adopted for the RCMs (MAR, RACMO, HIRHAM). However, to avoid any ambiguity, we will revise the manuscript to explicitly refer to "CESM2" throughout.

On the introduction of satellite datasets (AMSR2 and SSM/I) in the Results section

We acknowledge this comment. We note, however, that the comparison with satellite observations is presented following the funnel structure of the Results section: the analysis begins with a broad evaluation of present-day SMB against ERA5, before zooming in on the specific component that diverges most between models (melt), where observational data are then introduced for context. We also note that the description of these datasets and the justification for using two products are already provided in the paragraph on page 12. We believe this placement remains appropriate given this narrative structure, but we will ensure the datasets are at least mentioned in the Methods section for completeness.

On the size of map plots (e.g. Fig. 3)

We agree that the current 4×2 layout makes it difficult to identify small-scale features such as ice shelves. We will reorganize the figure into a 2×4 layout, which will result in larger individual panels and improve readability of fine-scale features.

On standalone paragraphs in the Results section (e.g. around Fig. 4)

We agree. The standalone paragraphs will be integrated into the surrounding text by removing unnecessary paragraph breaks, ensuring a better narrative flow throughout the Results section.

Minor comments

L31 – Citation needed for warmer ice shelves

We will move the following citations to the end of both sentences:: (Kittel et al., 2021; Boberg et al., 2022; Van Wessem et al., 2023; Jourdain et al., 2025)

L33 – Mention of GCMs alongside ESMs

We agree. We will revise the sentence to explicitly distinguish GCMs from ESMs and acknowledge that GCMs lack certain biogeochemical processes. The revised sentence will read:

"The climate models used to estimate the SMB of ice sheets fall into two main categories: ***global climate models (GCMs), including the more complex*** *Earth system models (ESMs), and regional climate models (RCMs), each having its own strengths and limitations. The main advantage of ESMs for simulating ice-sheet SMB lies in their physically consistent global framework, with interactive coupling between the atmosphere, ocean, sea ice, and land surface,* ***as well as the inclusion of additional biogeochemical processes compared to traditional GCMs****."*

L36 – Number of RCM limitations

We agree that the current phrasing incorrectly implies that RCMs face only two major limitations. We will rephrase the sentence to make clear that the two limitations mentioned are illustrative examples among several, rather than an exhaustive list.

L53 – Citation for model uncertainty (Hawkins and Sutton, 2009)

We thank the reviewer for this suggestion. We will add the citation to Hawkins and Sutton (2009) at L53, which provides a useful framework for distinguishing model uncertainty from scenario uncertainty and natural variability in climate projections.

L57 – Repetitive mention of high-emission scenario

We agree that this sentence is repetitive. We will revise it to remove the redundant mention of the high-emission scenario while retaining the key information.

" To address this gap, the present study performs an intercomparison of SMB simulated by these three RCMs, all driven by the same ESM, the Community Earth System Model (CESM), under a high-emission scenario (SSP5-8.5)."

L81 – 'Resolving' vs. 'capturing'

We agree. We will replace "resolving" with "capturing" to better reflect the fact that some of these processes are parameterized rather than explicitly resolved.

L84 – Define CMIP6

We will add the full name "Coupled Model Intercomparison Project phase 6" at first use at L84.

L98 – CESM2 naming inconsistency

This inconsistency will be resolved by adopting “CESM2” consistently throughout the manuscript, as noted in our response to the major comment on model naming above.

L144 – ‘Few widely spaced layers’

We thank the reviewer for flagging this ambiguity. We will clarify what “few widely spaced layers” refers to in the context of the HIRHAM description, and revise the sentence accordingly.

Section 2.1 – Introduction of snow and firn schemes

We agree that the model descriptions jump directly into snow and firn schemes without explicitly stating their role in computing melt, percolation, retention, refreezing, and runoff. We will add a sentence at the beginning of Section 2.1 to clarify this, and will also add a sentence in the Introduction to contextualize the statement about the rudimentary representation of the snowpack.

Section 2.2 heading

We agree that “analysis” is not appropriate here. We will remove it from the section heading.

Table 1 – Terminology and title

We will revise the column headers to “Surface albedo” and “Maximum water capacity of firn”. We will also update the table title to: *“Main characteristics and parameterisations of each model for the representation of the SMB.”*

Section 2.3 – Comment on the 10 km common grid

We agree. We will add a sentence noting that the 10 km common grid is finer than the native grids of several models, and in particular that it represents a substantial downscaling step relative to CESM2, which operates at resolutions of 100 km.

Table 2 – Integration domain

The table heading already specifies “integrated over the entire Antarctic ice sheet”.

L234 – Missing citation

We will add the citation Jakobs et al. (2021) at this location.

Figure 3 caption – Use of ‘anomaly’

We agree that “anomaly” is not the appropriate term here, as the figure shows differences between period means rather than deviations from a baseline. We will revise the caption to read:

“Future mean SMB (2080–2099) relative to present mean SMB (1995–2014)…”

Or we can talk about “change” rather than “anomaly” for this case.

We will also check and correct all other uses of “anomaly” in the text where the same issue may arise.

L275 - L276 – Identification of ice shelves on maps

We agree that many ice shelves are difficult to identify in the current maps. We will add a supplementary map labelling the main ice shelves and geographic locations referenced in the text, and will refer to it where relevant.

L295–296 – Standalone paragraph

We agree. The standalone paragraph will be merged with the surrounding text by removing the unnecessary paragraph break.

Figure 4 caption – Vague reference to ‘parts of the ice sheet’

We thank the reviewer for this comment. The caption already specifies the two parts referred to: "grounded (left) and floating (right) parts of the ice sheet." We will simply add "Antarctic" before "ice sheet" for completeness.

L332 – ‘Observed differences’

We agree that the term "observed" is neither appropriate nor necessary here. We will simply remove it.

Summary – Mention of Bayesian treatment

We agree that the mention of a “Bayesian” treatment in the abstract is not followed up in the manuscript. We will introduce a brief mention of this approach in the Summary section to ensure consistency with the abstract.

To better clarify the answer to the first comment, CESM2 was selected because it provides a physically consistent forcing for Antarctic applications and has been shown to realistically represent several key aspects of the Antarctic climate system, including the present-day surface mass balance (e.g. Dunmire et al., 2022). Furthermore, CESM2 belongs to the subset of CMIP6 models identified as providing plausible large-scale climate conditions over the Antarctic Ice Sheet, making it particularly suitable for high-latitude applications (Agosta et al., 2022; Mottram et al., 2025). In addition, CESM2 includes one of the most sophisticated snow schemes among CMIP-class ESMs (van Kampenhout et al., 2017).

We also now note that CESM2 exhibits one of the highest equilibrium climate sensitivities within CMIP6 (Meehl et al., 2020; Zelinka et al., 2020), resulting in a comparatively strong warming response under future high-emission scenarios. This strong climate forcing enhances surface melt and amplifies changes in Antarctic surface mass balance compared to lower-sensitivity CMIP6 models, particularly over ice shelves where increased meltwater production can partly offset or even exceed the increase in snowfall (Kittel et al., 2021; Jourdain et al., 2025). Regarding Antarctic sea ice, CESM2 provides a realistic representation of present-day Antarctic sea-ice characteristics and associated processes compared to previous model

generations (DuVivier et al., 2020; Singh et al., 2021). Future Antarctic sea-ice projections remain associated with substantial uncertainty across CMIP6 models, and CESM2 represents one possible realization of this range. Nevertheless, the combination of a realistic representation of Antarctic climate processes and a relatively strong climate response makes CESM2 particularly suitable for our objective of assessing the sensitivity of different RCMs under a common large-scale forcing.

While CESM2 does not represent the full range of uncertainty associated with CMIP6 projections, it provides a physically plausible high-end climate response that allows us to strengthen the climate signal and investigate the sensitivity of the different regional climate models to a common forcing.

Reply to reviewer 2

We thank Reviewer 2 for their thorough and constructive review of our manuscript. We appreciate the positive assessment of the manuscript's clarity and the relevance of the results, particularly in the context of ongoing ISMIP7 efforts. We have carefully addressed all comments and questions raised by the reviewer, point by point, in the sections below. For each comment, we indicate the proposed changes to the manuscript text, with reference to the relevant line numbers of the original submission.

General comments

On the 2% water retention value in RACMO

The version of RACMO used in this study, as well as in Van Wessem et al. (2018), employs the original internal firn model described by Ettema et al. (2010), in which the irreducible liquid water retention capacity is prescribed as a fixed value of 2%. This parameterization differs from the more recent IMAU-FDM configuration described in Machguth et al., where the water holding capacity depends on firn porosity. We have therefore replaced the incorrect citation with Ettema et al. (2010) and clarified this point in the revised manuscript.

On the apparent discrepancy between RACMO and MAR runoff despite similar melt values

This is indeed a complex issue resulting from the interplay of several factors, and we agree that the discussion in L482–500 would benefit from being made more explicit. We identify at least four contributing mechanisms:

1. **Firn layer thickness and water retention capacity.** The thinner firn column in MAR (20 m) compared to RACMO (30 m) could limit refreezing capacity once the firn is saturated, potentially contributing to higher runoff in MAR. However, MAR's higher water retention threshold (~5%) largely compensates for the shallower column. A simple calculation illustrates this:

Density assumption	RACMO (30 m, 2%)	MAR (20 m, 5%)
Uniform 400 kg m^{-3}	339 L	565 L
Uniform 500 kg m^{-3}	273 L	456 L
Layered profile (0–2 m @ 300, 2–15 m @ 500, >15 m @ 720 kg m^{-3})	211 L	419 L

Across all tested density profiles, MAR's total water retention capacity is roughly twice that of RACMO. Firn layer thickness alone is therefore unlikely to be the primary driver of the runoff difference in most cases, and we prefer not to highlight it as a leading explanation without stronger supporting evidence.

1. **Rapid FAC depletion in MAR.** MAR produces substantially more liquid water (through both melt and rain, given its warmer climate), which rapidly saturates the available pore space and consumes firn air content (FAC), regardless of the higher retention threshold. Additional factors — such as differences in densification parameterizations, fresh snow density, and firn regeneration capacity — may further accelerate FAC reduction.
2. **Simulation continuity and firn thermal state.** RACMO is run as a single continuous long simulation; in a warming climate, its firn column is therefore further from equilibrium and more "eager" to refreeze meltwater. MAR simulations are typically composed of shorter segments, starting closer to equilibrium with reduced refreezing capacity. Sensitivity tests with MAR run as a single long simulation yield results very similar to the segmented approach, however, because melt events rapidly homogenize the firn state — suggesting this factor plays a secondary role.
3. **Treatment of runoff at ice lenses.** RACMO does not parameterize runoff when percolating meltwater encounters an ice lens — runoff only initiates once the entire firn column is saturated. MAR prescribes that a fraction of percolating meltwater runs off upon encountering an ice lens. This structural difference in runoff parameterization likely contributes significantly to the higher runoff in MAR for equivalent melt rates.

Given the number of competing and interacting factors, we are cautious about attributing the runoff discrepancy to any single cause. We will revise the discussion in L482–500 to present these mechanisms more clearly and concisely.

On the use of annual vs. summer (DJF) temperatures in Figures 6–8 (and Fig. S3)

We thank the reviewer for this suggestion. We agree that confining the analysis to summer (DJF) near-surface temperatures would be more physically meaningful for the processes discussed (melt, runoff, melt-albedo feedback). However, this is a limitation of the available model output: the projections were produced on an annual calendar basis through 2100, and monthly or seasonal components are not available to us for this analysis. We therefore cannot disaggregate the results by season, and the analysis is necessarily restricted to annual means and cumulative quantities. We will explicitly acknowledge this limitation in the revised manuscript.

On the introduction of satellite products and evaluation metrics in the Methods section

We agree with this comment. AMSR2, SSM/I, and the evaluation metrics (CMS, MMS, MMD, MOA) will be introduced and described in the Methods section rather than first appearing in the Results section.

Minor comments

Abstract

We agree that the last sentence of the abstract lacks sufficient context. We will revise it to better reflect the manuscript's conclusions. Regarding a more extensive rewrite, we will first verify the abstract length limit imposed by the journal and extend it accordingly if the limit allows.

L17

We will restructure the sentence as recommended: *"from -40 ± 9 Gt yr^{-1} over the 1979–1990 period to -252 ± 26 Gt yr^{-1} over the 2009–2017 period."*

Introduction – Mottram et al. 2021, Machguth et al. 2026, Medley et al. 2022

We will include Mottram et al. (2021) and Machguth et al. (2026) in the introduction as an additional RCM intercomparison study, noting that it focuses on the historical period. Both references will be contextualized with respect to the present study.

L22–24

We will add a reference to Medley et al. (2022) at this location to note that more sophisticated firn models also predict that most meltwater refreezes in the snowpack over Antarctica.

L26–29

We agree that this sentence is too long and convoluted. We will break it into two shorter sentences for clarity.

L46–48

We agree that this sentence requires clarification. By this statement, we mean that RCMs, while refining the spatial representation of climate forcing, are not independent of the driving GCM and may inherit or amplify its biases, introducing an additional source of uncertainty in the context of future projections. We propose to rephrase the sentence as follows:

"However, RCMs need to be forced by an ESM at their boundaries, which may introduce additional uncertainty due to potential internal biases or differences in climate sensitivity, particularly in the context of future climate projections."

L75–78

We agree and will restructure the sentence as suggested: *"The omission of wind-driven erosion and deposition is not expected to compromise our intercomparison since runoff rapidly becomes…"*

L191

We refer to "future anomalies" as the difference between the mean of the 2080–2099 period and that of the reference period 1995–2014, and to "time series anomalies" as the deviation of each annual value from that same reference period mean. We will clarify this terminology in the manuscript.

L229

Current sentence : *« This refreezing partially suppresses melt, yet significant runoff still occurs,... »*.

"melt" is correct in this context. The sentence refers to the previous statement that HIRHAM simulates melt rates approximately five times greater than those of the other two RCMs, accompanied by refreezing about four times higher than in MAR and RACMO. The phrase "this refreezing partially suppresses melt" therefore means that the higher refreezing rate partially compensates for the higher melt rate, yet significant runoff still occurs. No change is needed.

QuikSCAT comparison

L256

We agree that this sentence is confusing and will rephrase it for clarity. The sentence discusses how changing the melt detection threshold from 1 to 3 mm day^{-1} affects the Mean Melt Duration (MMD) differently across models. For HIRHAM, increasing the threshold has only a limited impact on MMD, suggesting that most melt days in HIRHAM exceed 3 mm day^{-1} regardless of the threshold applied. In contrast, for MAR and RACMO, raising the threshold reduces the number of melt days counted, as fewer days exceed the more restrictive threshold, which has a more significant impact on their respective MMD values. We will revise the sentence to make this distinction explicit.

L322

We will add "(near-surface air temperature)" after "climate response" for clarity.

L337

This is indeed a strange mistake. We will correct to W m^{-2} K^{-1} both in the text at L337 and in the caption and axes of Figure 6.

L343–L345

We will specify the starting temperature for HIRHAM over ice shelves at L343, and correct the RACMO starting temperature to ~254 K at L345.

L346

We agree that the sentence is imprecise. What we mean is that HIRHAM stands out from the other models throughout the entire simulation period, from start to finish, in terms of the magnitude of its response. We will revise the sentence to clarify this, and explicitly acknowledge that panels c, e, and f are exceptions, where HIRHAM values are close to those of CESM2.

Figure 8 – MOA vs. temperature and legend

We will add a supplementary figure showing MOA as a function of near-surface temperature. We will also update the figure legend to explicitly indicate that open dots represent accumulation and closed dots represent melt.

L516

We will correct the citation to "The Firn Symposium team" as recommended.

Figure S5

We apologize for this omission. Figure S5 was inadvertently left out of the submitted document and will be included in the revised submission.

Technical corrections

We will implement all suggested technical corrections:

1. **L1:** Add "Global Climate Model" before CESM2 and "Regional Climate Models" before MAR, RACMO, and HIRHAM, and use CESM2 consistently throughout the manuscript.
2. **L84:** Move the Danabasoglu reference to parenthesis.
3. **L145:** Add missing period after "SMB model".
4. **L235:** Remove the duplicated sentence.
5. **L263:** Remove "can" from "We can also".
6. **L407:** Correct "F. 8" to "Fig. 8".

Minor clarifications throughout the text

We will implement all suggested clarifications: adding "SMB" or "change in SMB" at L7; "mass" at L16; "air" at L27; specifying the referent of "This figure" at L266; adding "SMB" at L279, L285, L286, and L306; adding "over ice shelves" at L318.

Jan De Rydt – topical editor

Minor corrections:

Line 12: start with “THE Antarctic Ice Sheet”…

Line 22: remove duplicate “snow”

Line 34: its -> their

Line 84: (Danabasoglu et al., 2020) between brackets

Line 145: full stop missing after SMB model

Line 179: clarify if the common ice mask is applied in each of the RCMs, or as a post processing step when analysing the data

Line 190: does “-future anomalies-“ need the hyphens, or did you mean this to be in italic font?

Line 191: same as above for “-time series anomalies-“

Line 200: remove “result”

Line 212: comma instead of a full stop before “Confirming”

Line 419: “Ross ICE shelves”