Reviewer #2

General

Compared to the original version, the paper is completely overhauled. Much of the unclarities have been resolved, and the readability of the paper has much improved, although the amount of information can still be overwhelming to the unprepared reader. Although one could still argue about the robustness of some of the results, I feel it is a valuable contribution, especially for the ice sheet modelling community in search of an easy way to implement a first-order estimate of snowfall, melt and firn saturation in their offline Antarctic ice sheet models.

⇒ We sincerely thank the reviewer for this second careful examination of our manuscript. We have revised our manuscript following most of the suggestions.

Remaining not so minor comments

(line numbers refer to document including tracked changes)

Paragraph starting line 29: This is a useful summary of the conditions that must be met before runoff occurs. As it is presented now, it can also be read that ANY of these conditions must be met rather that ALL. Can you make this a little more specific?

⇒ Yes, this sentence has been rewritten as:

*Runoff is a negative contribution to the surface mass balance. It is produced if surface melt and/or rain rates are high enough to **successively** (i) percolate and bring the temperature of underlying snow and firn layers to the freezing point, (ii) saturate the pore space in the snow and firn layers, which is sometimes referred to as firn air depletion (Pfeffer et al. 1991, Kuipers Munneke et al. 2014, Alley et al. 2018), and (iii) flow into the ocean.*

Have you considered using the simpler phrase "excess liquid water" rather than "liquid water in excess"?

⇒ Thank you for this suggestion, we have used "excess liquid water" in the revised manuscript.

l. 115: It remains unclear whether the MAR simulations are fully transient, and how/over which period the model snow/firn layer was initialized with regards to e.g. density and liquid water content. Please provide this information and how it could influence the results. Along the same lines: how should we see these results in the light of this recently published paper comparing RCMs over the Greenland ice sheet: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL111902? Consider including a remark or two about this in your discussion following line 520.

⇒ We have added the following paragraph to clarify this point:

For computational reasons, all simulation years were run in parallel with 20 years of spin-up to equilibrate the firn properties (e.g., 2051 is spun up from the transient 2031–2050 period). The initial state (e.g., 01-JAN-2031 for the simulation of 2051) is taken from the MAR–ACCESS-1.3 RCP8.5 simulation (Tab. 2), itself spun up from a previous version of MAR at 50 km resolution driven by NorESM1-M under RCP8.5 and spun up for 30 years from a present-day

MAR simulation. A spin up of 20 years is generally sufficient to remove any sensitivity to the initial state (Donat-Magnin et al., 2021, their Fig. 12).

About the suggestion to add a reference to Glaude et al. (2024), there was some specific tuning for the Greenlandic MAR configuration that may explain the strong future surface mass loss compared to the RACMO model. This tuning has not been used in our Antarctic configuration, so we do not think it is relevant to go too far in the reference to Glaude et al. (2024). We have nonetheless added a reference to this paper in the discussion section to support the extension to other RCMs.

Eq. 4: I find the use of RU as a negative number confusing. If runoff is positive, it represents a mass loss, so SMB = SF - RU is a more logical convention.

⇒ We think this is a matter of taste and we prefer keeping RU as a negative number so that removing the runoff contribution from the SMB can be written as SMB – RU (1st line of eq. 4).

l.494: I was surprised to read this so I looked it up but I think that Van Wessem and others (2023) do not claim the Ross ice shelf being susceptible to hydrofracturing. In their study, the near surface air temperature in some CMIP models and scenarios rather passes the warming threshold allowing initiation of conditions that in time can lead to meltwater ponding (note that saturating the firn layer can still take a lot of time especially if melt is weak). See also your own remark about this delay in l. 514.

⇒ We agree. Only half of the models in van Wessem et al. (2023) are prone to ponding over the Ross Ice Shelf under SSP5-8.5, which is not sufficient to reach our likely range (defined as 66% probability). We have therefore included van Wessem et al. in the previous citations:

*The giant Ross and Ronne-Filchner ice shelves are unlikely to experience hydrofracturing before the early 22nd century in SSP5-8.5, consistently with previous studies (Kuipers Munneke et al. 2014, **van Wessem et al. 2023**, Dunmire et al. 2024, Veldhuijsen et al. 2024).*

Remaining minor comments

(line numbers refer to document including tracked changes)

Some references ran from the page in my pdf.

⇒ This is an issue of the track-change version, the version with no tracked changes is ok. We didn't find an easy way to fix this using the *LaTeX/Overleaf* editor and the *latexdiff* command line.

l. 62: "Because of these difficulties, only a limited number of RCM-based projections are usually produced, which is generally insufficient to sample the CMIP model diversity regarding their representation of the recent period (Barthel et al., 2020) and their sensitivity to increasing anthropogenic emissions (e.g. Hausfather et al., 2022)." This sentence is formulated a little awkwardly; consider breaking up and/or reformulating.

⇒ This has been rewritten as:

*Because of these difficulties, only a limited number of RCM-based projections are usually produced, which is generally insufficient to sample the CMIP model diversity**. This may affect the** representation of the recent period (Barthel et al., 2020) and **the** sensitivity to increasing anthropogenic emissions (e.g. Hausfather et al., 2022) **in the small RCM ensemble**.*

l. 70: later -> latter

⇒ This has been corrected.

l. 91: consider removing "with each other".

⇒ This has been modified as suggested.

Caption Table 1: SSP-5.85 -> SSP-5-8.5

⇒ This has been corrected.

Caption Table 1: for which at least a MAR -> for which at least one MAR

⇒ This has been corrected.

Tables 1 & 2: please try to diminish the overlapping information in these tables.

⇒ Two models in Tab. 2 are from the CMIP5 era and are not represented in Tab. 1. We have therefore kept the ECS and ensemble members in Tab. 2 even if the information was already provided in Tab. 1 for the CMIP6 models.

l. 148: "here we do not attempt to know" please reformulate.

⇒ This has been rewritten as “we do not attempt to determine”.

l. 154: Is this really proof of a positive mass balance?

⇒ This paragraph has been rewritten following a suggestion of Reviewer #3 (see response to their 3rd comment).

l. 156: contribution > contributions

⇒ This has been corrected.

Caption Fig. 1: non -> not (3x)

⇒ This has been corrected.

l. 325: surface air -> near-surface air

⇒ This has been modified as suggested (also in 5 other sentences).

l. 334: There appears to be a parenthesis missing

⇒ We did not find any missing parenthesis around L. 334.

l. 353: as high as ~1000 m high -> up to elevations of 1000 m asl

⇒ This has been modified as suggested.

l. 345: into -> onto

⇒ This has been corrected (L. 354).

l. 405: over -> above

⇒ This has been corrected.

l. 406: a SMB -> an SMB

⇒ This has been corrected (also in two other sentences).

l. 487: monitored -> considered

⇒ This has been corrected.

Reviewer #3

General comments

This study uses a mixed physical/statistical method to dramatically increase the number of model ensemble members available to project future Antarctic SMB changes, and consequently sea level rise contributions and ice shelf hydrofracturing risk. I find it to be both novel and interesting, and feel it will be a valuable contribution to the scientific literature. The manuscript has been revised significantly in light of two reviewers' comments, and the result is a much clearer and more instructive paper. The schematics were particularly helpful to me to understand the method. The authors have addressed many of the earlier concerns of the reviewers, however I feel the paper would still benefit from some minor revisions, as detailed below.

⇒ We welcome this new evaluation of our work, and we thank Reviewer #3 for these very useful suggestions.

Line numbers are in reference to the tracked changes document

L49-52 Because you say here that the timing of ice shelf collapses influences sea level rise, it could be helpful to explicitly mention somewhere that you do not include the contribution of these collapses to sea level rise, because you don't have a dynamical ice sheet in your simulations.

⇒ We have added the following sentence in the introduction paragraph of section 3:

Importantly, our estimates of sea level projections only contain the part related to SMB variations, and not the contribution from the ice sheet dynamics which is driven by ocean-induced melting and hydrofracturing.

Tab 1 caption "2300" ◊ "2200"

⇒ This has been corrected.

L152-155 Can you offer any quantification of how much snowfall >> rainfall in the present or melt >> rain in the future? For instance, what percentage of MAR's precip falls as rain in the historical simulation, and how much does this increase by the end of the century?

L155-157 Relatedly, can you estimate the uncertainty associated with the assumption that drifting snow and sublimation are negligible? Maybe you can put a number on the mass change associated with sublimation (either from Agosta et al. or your own MAR simulation)?

⇒ We have rewritten the corresponding paragraph in section 2.3.1 as:

> *In Eq. 3, the effect of rainfall, sublimation and drifting-snow erosion are assumed to be negligible. Sublimation remains below 10% of snowfall even in a warmer climate (Kittel et al., 2021, their Tabs. S2-S3), and drifting-snow erosion is at least an order of magnitude smaller than sublimation (Gadde and van de Berg, 2024). As shown in Appendix B, rainfall represents less than 15% of the total precipitation on the grounded ice sheet until 2200 and on the ice shelves until 2100. The impact of neglecting rainfall in our method is discussed in Appendix B.*
>
> We have added a new appendix (now Appendix B) that shows the snowfall, rainfall and melting time series for both the grounded ice sheet and the ice shelves in the IPSL–SSP5-8.5 simulation until 2200 (now Fig. B1). In this Appendix, we provide arguments supporting that rainfall can be neglected in our approach.

Fig 4 I agree with R1 that this figure and its description is slightly confusing (there's a lot of information in there!) but your edits have improved the clarity of the caption. I wonder if you could increase the clarity of the text by making explicit reference to each of the properties of the figure that lead you to certain conclusions, e.g. "First of all, the minimal differences between the black dots on each radial and the coloured dot corresponding to the same GCM show that the biases are small for MAR simulations derived from themselves (Fig. 4). This shows that our methodology and its implementation are robust."

⇒ We have followed this suggestion and this paragraph now starts as:

> *The method is evaluated in Fig. 4. First of all, the minimal differences between the black dots on each radial and the coloured dot corresponding to the same climate model show that the biases are small for MAR simulations derived from themselves. This shows that our methodology and its implementation are robust. We nonetheless note larger differences between the black dots on each radial and the coloured dots corresponding to the other climate models in Fig. 4c,d, indicating significant biases in melt rates when a MAR simulation is derived from another one.*

L260-262 It doesn't follow so well that you talk about the biggest limitation of this part of the method (that runoff is underestimated at the end of the 21st century in the ssp3-4.5 emulated scenario) and then immediately say that you conclude the method is adequate. Another statement to qualify this would help the reader understand your conclusion, particularly as it makes the end of century emulated SMB a different sign to the original MAR simulation. e.g. can you say something about the relative magnitude of the underestimate in runoff, or say something about the performance of the emulated simulation over the course of the entire period?

⇒ We agree. This paragraph has been rewritten as:

> *The SSP1-2.6 and SSP2-4.5 emulated fields are quite accurate over the grounded ice sheet during the 21st century: the biases indicated in Fig. 3 for the emulated values are relatively low compared to the mean values simulated by MAR. The bias in emulated runoff becomes larger at the end of the 21st century in the SSP2-4.5 scenario (Fig. 3e-f). This bias has little impact on the grounded ice sheet SMB (Fig. 3a) but cancels the small negative SMB anomaly simulated*

over the ice shelves near 2100 (Fig. 3b). These biases are small and limited to the end of the 21st century, so we conclude that our method is suitable for the emulation of multiple SSP scenarios based on an existing MAR simulation in a warmer scenario (here SSP5-8.5). The spatial patterns are also well represented by this method (Appendix D, Fig. D2).

L297-298 I think this statement could be clearer. You could explicitly state that changing the r value would not be consistent with the results of the previous sections, nor reflect physical processes.

⇒ This has been modified as suggested.

L303 Given that averaging across CMIP-model emulations is required to achieve a useful result (which makes sense given the uncertainty associated with clouds and other parameterisations in GCMs), can you offer guidance on the way these results and/or the method should be interpreted by others wishing to replicate your approach?

⇒ We have added the following statement in the Discussion section:

For any method, the results presented in section 2.3.5 stress the importance of using a training dataset made of RCM simulations driven by multiple CMIP models for obtaining robust results.

Figs 1, 6 These schematics are really helpful for aiding understanding!

⇒ Thank you!

L377 should this be "17th-83rd"?

⇒ This has been corrected.

L378-380 I think some subtle changes here would make this clearer: "These percentiles account for the uncertainty of the CMIP models weighted to account for the likelihood of their ECS, and for the uncertainty of the threshold on liquid water production in excess when we investigate the potential for ice shelf hydrofracturing."

⇒ This has been modified as suggested.

Fig 8 caption missing reference ("section??")

⇒ This has been corrected.

L484 Earlier you use the spelling Dröning Maud Land – now it's Dronning Maud Land. Please use one for consistency.

⇒ This has been corrected to "Dronning".

Discussion – a brief discussion of MAR's limitations (e.g. lack of ponding/routing, no dynamic ice sheet etc.), as well as how well the method may be applied to other RCMs would be interesting to include.

⇒ We have added the following sentences to the Discussion:

The absence of physical representation of ponding and horizontal routing of liquid water nonetheless remains a major caveat of all current RCMs. Another important limitation of RCMs is the use of a constant ice sheet elevation, although the melt–elevation feedback seems to become important only after 2200 in the SSP5-8.5 scenario (Coulon et al., 2023).

L518 “Other RCMs, possibly combined to elaborated firn models” – do you mean “combined with elaborated firn models”?

⇒ Yes, this has been corrected.

L554 “built” ◊ “build”

⇒ This has been corrected.

L558 “models” ◊ “models’” or “model”

⇒ This has been corrected.

Figs C1-C3 I’m impressed by the similarities between the directly simulated and emulated results. The only major difference I can really pick out is a different sign between MAR/emulation over the Ross ice shelf, which is evident in all three figures. Can you comment on this?

⇒ We have added the following comment to Appendix D:

For a reason that remains elusive, the largest mismatch is found on the Ross ice shelf, where the emulation produces a negative SMB around 2100 for the three applications, while it remains mostly positive in the original MAR simulation.

L619-621 Missing figure references

⇒ This has been corrected.

L620 Think this should be “In East Antarctica, a runoff anomaly first prevails”

⇒ This has been corrected to “In East Antarctica, runoff anomalies first prevail”

Fig D1 caption missing section reference

⇒ This has been corrected.