Response to anonymous Referee #2.

We thank the referee for the time spent on our manuscript and for these careful and valuable comments.

General

This is an interesting paper that addresses an important problem: when can we expect runoff from Antarctic ice shelves, leading to mass loss and/or meltwater ponding which is commonly regarded as a precursor for ice shelf hydrofracture. The authors use multiple MAR simulations of future Antarctic climates, with which they calibrate a simple statistical model to emulate melt, runoff so as to approximate SMB and (instantaneous) firn saturation. Not considered are sublimation, rain and the time it takes to saturate the firn. When compared directly to MAR output, the emulator gives mixed results, yet provides valuable insights in the uncertainties that arise from the driving global models. The writing can be clarified in places (see below), the figures are generally clear.

Major comments

The abstract is at places confusing. It goes in one step from describing runoff emulation to AIS SLR contribution. Although it is stated that the mass loss numbers are based on SMB alone, the link that is made to hydrofracturing in the title and the presented numbers in terms of sea level rise could trick the reader into thinking that dynamical effects are also considered. It should be made clear from the outset is that this study treats runoff in two ways: as a source of mass loss and as a threshold to induce ponding and hydrofracturing (after which the dynamical effect on the ice sheet is not considered).

⇒ This is a good suggestion. We will improve the clarity of these aspects throughout the revised manuscript and in particular in the abstract.

Also in the abstract: "Based on a more limited and uncorrected ensemble, we find a considerable uncertainty in the contribution to sea level from 2000 to 2200". It is unclear whether this enhanced uncertainty derives from the fact that the ensemble is smaller and uncorrected, or from the method, or all of these. Please only include major results in the abstract, the details can be elaborated upon in the text.

⇒ In the description of our results, we will make it clearer that the larger uncertainty after 2100 is largely due to the diverging behavior of CMIP models: for some CMIP models, SMB over the grounded ice sheet remains over present-day values after 2100 while it strongly decreases in other models. For this reason, and given the limited sample size of 8 models, we do not provide an average after 2100 and only indicate the spread. The sentence quoted from the abstract will be rewritten as "*Based on a more limited and uncorrected ensemble, we find a considerable spread across CMIP models in the contribution to sea level from 2000 to 2200*".

l.59: The statistical model is based on MAR data, which is a fair choice given that this model has been used for multiple future runs. But for the statistical model to be useful, MAR must provide reliable runoff. Here it is stated that the atmosphere in MAR is coupled to a 30 layer snow model "...representing the first 20 m of snow/firn with refined resolution at the surface". But in many places the firn layer in Antarctica is considerably deeper than 20 m? What does this imply for runoff simulated by MAR? See also my comments on treatment of percolation and runoff evaluation below.

⇒ Only using MAR is certainly a caveat but this was the only RCM for which we had an ensemble of projections sufficiently large to be robustly emulated. This caveat was already acknowledged in the Discussion section. We nonetheless agree that simulating 100 m of firn would be better than 20 m, although the potential presence of ice lenses makes it difficult to anticipate the exact effect. The firn depth is expected to change the timing of firn saturation/air depletion, but not the threshold for firn saturation which is more related to the melting and snowfall rates. We have added the sentence in red to the relevant paragraph in the Discussion section:

"*Our approach has consisted of emulating MAR simulations. Other RCMs, possibly combined to elaborated firn models, have similar skills in representing typical Antarctic conditions (Mottram et al., 2021), but there is likely a considerable spread in their response to surface warming. For example, the depth and vertical resolution of firn models probably make important differences in the timing of runoff production. The 20-m firn layer simulated in MAR thus likely reaches liquid-water saturation earlier than models with a thicker firn layer. One of the next priorities will therefore be to emulate the diversity of RCM sensitivities, which would make the uncertainty ranges much more comprehensive.*"

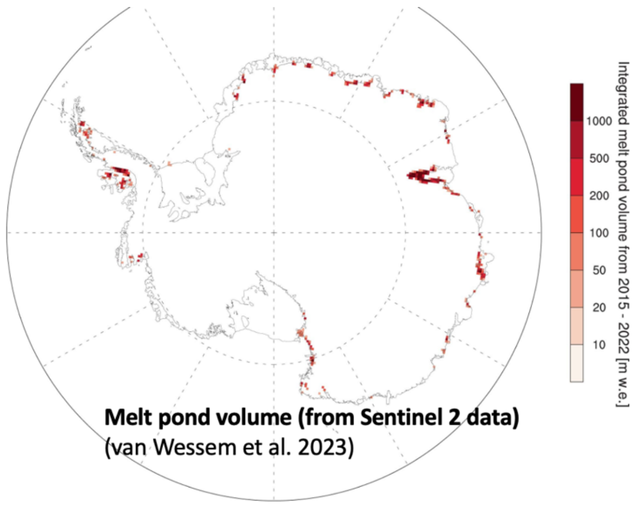
l.65: " If liquid water is not able to percolate further down, then it fills the entire porosity space of surface layers, and the excess is considered as runoff and removed from the snow/firn model (there is no representation of ponds or horizontal routing)". From this description it is not clear to me how water percolation interacts with ice lenses in the firn layer and how the process of filling up works before runoff starts.

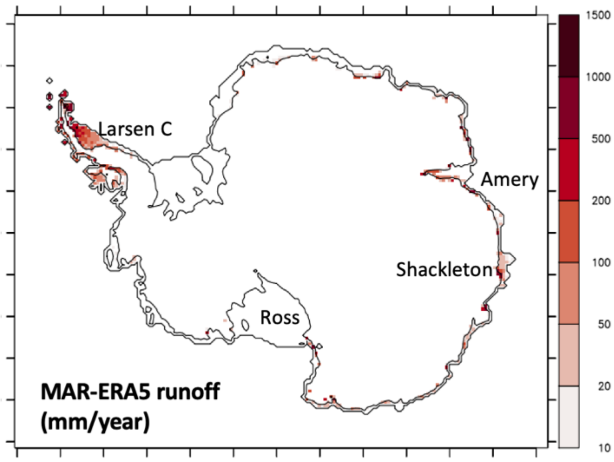
⇒ This sentence will be clarified so that the entire paragraph reads:

"*In the presence of surface melting or rainfall, liquid water percolates downward into the next firn layers with a water retention of 5% of the porosity in each successive layer. The firn layers are fully permeable until they reach a close-off density of 830 kg* m^{-3}. *To account for possible cracks in ice lenses and moulins, the part of available water that is transmitted downward to the next layer decreases as a linear function of firn density, from 100% transmitted at the close-off density to zero at 900 kg* m^{-3} *and beyond. If liquid water is not able to percolate further down, it remains where it is. When the entire porosity space in the surface grid cell is filled with liquid water or if the surface grid cell is denser than 900 kg* m^{-3}, *any additional surface melt is considered as runoff and removed from the snow/firn model. There is no representation of ponds or horizontal routing.*"

l.67: "The surface mass balance and melting conditions produced by MAR have been evaluated in comparison to observational." Has there also been an attempt to evaluate the modelled meltwater ponding/runoff? I seem to remember that Van Wessem et al. (2023) compared their predicted ponding potential to satellite observations of melt ponds.

⇒ Van Wessem et al. (2023) indeed presented a comparison to an observational estimate of melt pond volume (derived from Sentinel 2). This comparison is qualitative as neither MAR nor RACMO simulate ponding (they remove the excess of liquid water and do not simulate horizontal transport of liquid water). We nonetheless show here the runoff produced by MAR forced by ERA5 over 2015-2022 in comparison the melt pond volume over the same period:





We find that the areas of high runoff in MAR generally correspond to the areas where high melt pond volumes are estimated from Sentinel 2 data, even if the area of high runoff over Larsen C is larger than the area of large melt pond volume in the satellite product. We will add a sentence on this overall qualitative agreement although we may not include the figure.

We also note that already included this statement at the end of section 3.3, which may be considered as a kind of evaluation:
"*In our projections, conditions favorable to hydrofracturing become likely by ~2050 for several East Antarctic ice shelves (Fig. 10). This is the case of Nivl, Roi Baudouin, Amery and Shackleton ice shelves on which widespread melt ponds or aquifers are already observed in present-day conditions (Kingslake et al., 2017; Bell et al., 2018; Stokes et al., 2019)*".

Figure 1 shows results for the integrated grounded ice/ice shelves only. How do the spatial patterns look?

⇒ This is a good suggestion. The revised manuscript will include the maps for some of the cases integrated in Figs. 1, 2, 4 (in Appendix).

Minor comments

l. 12: "Based on a runoff criteria.." Please note that 'criteria' is plural, so either use "Based on runoff criteria.." or "Based on a runoff criterion.."

⇒ Thank you, this will be corrected.

l. 12: "we identify the emergence of surface conditions prone to hydrofracturing" Suggest: "we identify the timing of surface conditions that make ice shelves prone to hydrofracturing" or something similar.

⇒ This will be corrected as suggested.

l. 13: " A majority of ice shelves could remain safe" Suggest to reformulate: "Our results suggest that the majority..."

⇒ This will be corrected as suggested.

l.16: precipitation -> snowfall

⇒ Precipitation consists of snowfall + rainfall, and rainfall is a positive term in the surface mass balance that effectively contributes to mass gain until the firn is saturated with liquid water or until the surface consists of bare ice. But rainfall has a minor effect and this first sentence may be clearer with "snowfall", so we will correct as suggested.

l. 18: However, if air temperatures exceed -> However, model simulations suggest that, if atmospheric warming exceeds

⇒ This will be corrected as suggested.

l. 22: can be conducive of -> can be conducive to

⇒ This will be corrected as suggested.

l. 31: sea level -> sea level change

⇒ This will be corrected as suggested.

l. 36: in particular with regard to firn saturation by meltwater and subsequent runoff -> in particular with regard to firn saturation by meltwater and subsequent ponding

⇒ This will be corrected as suggested.

l. 87: Eq. 1: At what level is the warming specified? For this expression to be accurate, the warming should be weighted over the atmospheric column in which the water vapour resides. This is different for Eq. 2 (l. 92), where near-surface warming can be used.

⇒ This is near-surface warming. This is an approximation that works because the troposphere is warmed relatively uniformly from the surface to ~300 hPa (Donat-Magnin et al. 2021). This will be added to the text.

l. 103: surface -> near-surface (I assume)

⇒ This will be corrected as suggested.

l. 108: " The m parameter is introduced to avoid unrealistically high melt rates" Can you provide a value, and how is it determined? Ah, it is provided in l. 120, but is given in kg m-1 s-1. Can you provide a more intuitive value, i.e. what this means per year?

⇒ We will add "1.80×10^{-4} kg m^{-2} s^{-1}, i.e., 15.5 mm/day" (this is also 5.7 m/year).

l. 109-112: This is unclear, please clarify.

⇒ We will better explain this part in the revised manuscript. To reconstruct the annual melt rate (or SMB minus runoff) at a given location and on a given year, an option is to start from the annual melt rate in another simulation and to correct this value to account for the air temperature difference (ΔT) between the two simulations. Given internal climate variability and uncertainty in the exponential fits, there is no reason to only use

1986 from simulation A to reconstruct 1986 in simulation B. So we reconstruct 1986 in simulation B twenty times, from years between 1976 and 1995 (and different ΔT values for every reconstruction) in simulation A, then we take the average of these 20 reconstructions.

We process similarly when we extend a simulation back in time: for example, 1950 (or any other year before 1980) is reconstructed 20 times from individual years in the corresponding MAR simulation that covers 1980–1999, and then we take the average.

This removes a part of the internal climate variability in the melt or SMB pattern of a given year (e.g., the influence of a strong synoptic event at a given year) but keeps the internal climate variability of air temperature in the reconstructed timeseries.

l. 180: typo "ooverestimated"

⇒ This will be corrected as suggested.

l.231: " Our projections over the grounded ice until 2100 agree quite well with previous estimates of sea level contribution reported by the IPCC for the three scenarios". Can these numbers be compared, i.e. did these IPCC assessments also report estimates based on SMB-only?

⇒ Yes, the IPCC emulated SMB values from the AR5 to the SSP scenarios (IPCC-AR6-WG1 Tab. 9.3). The values are already reported in our Tab. 4 which is pointed to at the end of the quoted sentence.

Section 3.2: It would be informative to provide the modelled atmospheric warming of the various models in addition to the SMB and runoff. Main motivation is that some models project negative SMB over the grounded ice sheet in the 22nd century, which must require a very strong warming, as well as significant rainfall. For these extremes, how well does the initial assumption hold that rainfall remains small?

⇒ This is a good suggestion. We will add maps of near surface warming at 2200 for the two groups of models shown in Fig. 9.

⇒ The assumption that rainfall still has a negligible role in much warmer climate is probably good given that:
  - Melt rates increase much more than rainfall rates in warmer climate (Tab. 2 of Donat-Magnin et al., 2021 and Fig. 2 of Kittel et al., 2021).
  - Melting is much more efficient than rainfall at saturating the firn with liquid water because melting removes snow/firn and converts it to liquid water while rainfall just adds liquid water (Appendix B of Donat-Magnin et al., 2021).

l. 293: lead -> led

⇒ This will be corrected as suggested.

l. 294: Georges -> George

⇒ This will be corrected as suggested.

l. 323: "...while it can take more than a decade to saturate it if melt rates are just above the threshold..." Theoretically, it can take an almost infinite amount of time if the threshold is just passed.

⇒ Yes, this is exactly what we meant, we will replace “more than a decade” with “an infinite amount of time”.

l. 363: criteria -> criterion

⇒ Thank you, this will be corrected.