ANONYMOUS REFEREE #2

General

This is an interesting paper that investigates the response of the Greenland ice sheet firn layer to idealized positive and negative temperature perturbations on century time scales. The paper is well and concisely written, and the figures are of excellent quality. My comment are therefore relatively minor and should be fairly easy to address.

We thank the referee for their careful review of the manuscript and constructive feedback. We have made changes based on the suggestions and believe the writing to be much improved. In particular, we appreciate the feedback on the description of other firn models compared to SNOWPACK. We thank the referee for their input and hope that our modifications have addressed their concerns. Please find below our responses to comments in blue text, and specific changes made to the text in ***blue, bold, italic font.*** Line numbers refer to the original manuscript.

Major comments

l. 71: "SNOWPACK uses near-surface atmospheric conditions as input". This is confusing; I assume more input parameters are required to calculate the surface mass and energy balance at the upper boundary of the firn model? Such as surface radiation and mass fluxes? Upon further reading this is specified in l. 92, but please correct here to avoid confusion.

We have changed this to ***SNOWPACK uses a meteorological dataset as input*** to improve the accuracy of this statement.

l.77: "Unlike semi-empirical firn models used in similar studies (e.g., Kuipers Munneke et al., 2015), SNOWPACK does not use a positive degree day parameterization for calculating melt." As it is written here, this is incorrect. The "semi-empirical" refers to the way snow densification is parameterised and does not refer to the way these models are forced. The study of Kuipers Munneke et al (2015) cited here used an idealized/simplified melt forcing simply to facilitate interpretation of the development of firn aquifers. Other studies by that same author use realistic surface mass and energy balance to drive the firn model, i.e. equivalent to what has been done here. There remains an important difference though. In the model setting used here, SNOWPACK calculates its own surface energy balance, whereas in most other studies the firn models are directly forced by surface skin temperature from regional climate models, being the result of the closure of the surface energy balance. Although not critical for this study, it is good to clearly separate these different approaches.

We appreciate this feedback and acknowledge that our text as written is misleading. In our effort to distinguish SNOWPACK from other models, we have incorrectly made assumptions about other work. We have changed this sentence to the following:

While many other firn models rely on surface skin temperature from the atmospheric forcing to calculate melt (e.g., Steger et al., 2017; Medley et al., 2022), SNOWPACK does not take this approach.

Section 2.3. The model coverage of Greenland is at relatively low resolution, and points that do not meet the spin up criteria are removed. As a result, only 1724 locations remain, an order of magnitude less than when the ice sheet would have been resolved at e.g. 10 x 10 km grid cells. This low resolution is also apparent from e.g. Fig. 1. Could you comment at how that potentially influences the (ice sheet integrated) results, and also regionally especially in regions with strong climate gradients? And can you indicate in Fig. 1 the outline of the contemporary ice sheet, so that it becomes clear which parts of the ablation zone have been removed from the analysis?

This is a great point and something that we have considered but had not added to the text. Most of the grid cells excluded with the spin up conditions are in the lowest elevations of the ablation zone. Based on our results, we expect very little change in firn air content to occur in those areas simply because there is no firn there, but rather exposed bare ice. The original Figure 7f is a great example of how very minimal changes occur in firn air content where the initial temperature and melt rates are already very high. As such, excluding the very margins is likely to have very little impact on our results since the margins contain little to no firn. Even 1 °C of cooling is not likely to have a substantial impact there (Fig. 7f). We have made the following changes in Section 3.1 to help convey the point that the very edges have no firn:

The largest changes in both warming and cooling experiments occur along the margins of the ice sheet but slightly inland from the edge of the model domain since firn is not present at the lowest elevations.

We appreciate the suggestion for adding the ice sheet outline and have done so for the original Figures 1, 2, 3, and 9. In the captions, we have also added the following text to describe the outline: ***The thin black line represents the ice sheet outline.*** As an example of this figure modification, below is the modified Figure 1 and its caption.

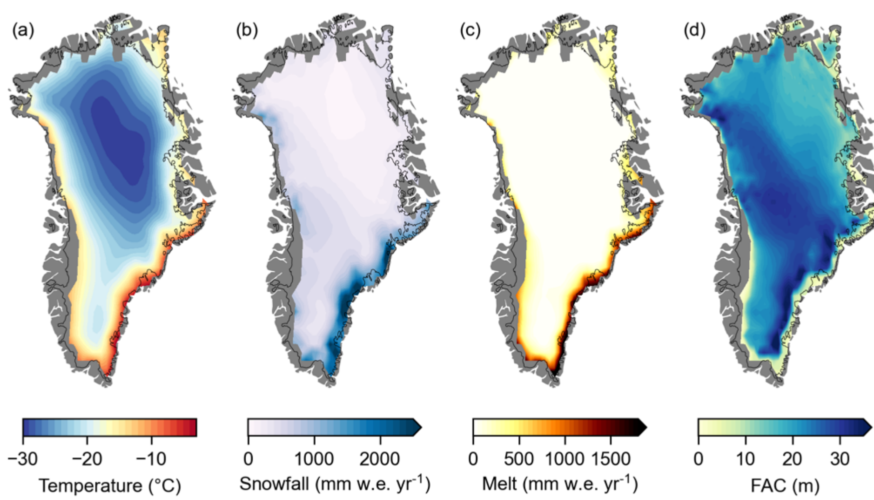

Figure 1: Greenland Ice Sheet mean annual (a) temperature, (b) snowfall, (c) melt, and (d) firn air content (FAC) from the 100-yr control climate (i.e., the 1991 climate) prior to the warming and cooling experiments. The thin black line represents the ice sheet outline.

As for the spatial resolution, this is something we cannot change if we choose to continue using MERRA-2. Our results are limited by the native resolution of our chosen forcing. We selected MERRA-2 because it has been used before in a study with a more realistic model framework (Thompson-Munson et al., 2023). A finer resolution could offer more detail in the steeper sloped marginal areas (e.g., the southeast), and it would be interesting (though beyond the scope of this work) to use an alternative atmospheric forcing such as RACMO.

Same section and results section: you select a single year as a baseline for all your experiments. Am I correct that Figs. 1a, b, c then represent the climate of 1991? Apart from the absence of a trend in firn air content for that year (is this averaged over the ice sheet, how does it hold regionally?), can you also provide information on the temperature/accumulation/melt of that particular year relative to climatology? It appears from Fig. 1 that melt is rather low along the western marginal ice sheet (although I realise that most of the ablation zone is not part of the model domain).

Yes, that is correct that this represents the 1991 climate. We have added ***(i.e., the 1991 climate)*** to the Figure 1 caption and added ***(i.e., for the 1991 climate)*** to line 121 to clarify this point. The addition of the ice sheet outline should help demonstrate that yes, the melt on the western margin is low because the ablation zone is excluded with the spin-up procedure. However, we have also updated the figure in response to another review, and it now shows some of the melt along the western margin—though not within the ablation zone. Please see the modified Figure 1 in the above response.

As for 1991 relative to the climatology, we would point you to Thompson-Munson et al. (2023) Figure A1 which shows time series of MERRA-2 variables used to force SNOWPACK from 1980 through 2020.

Forcing experiments: if I understand correctly, only 2 m air temperature is varied, but other variables that normally are closely connected to air temperature (incoming longwave radiation, specific humidity, rainfall fraction) are left unchanged? If so, please mention this specifically to avoid confusion.

We have specified “**2-m**” in line 117 and in added the following at line 118:

Note that we do not alter any other forcing variables and we use “air temperature” to mean 2-m air temperature throughout the rest of the manuscript.

Fig. 7: What causes the fluctuations with frequency of about 1/decade in panels a and c?

These fluctuations are a result of the output frequency of the model combined with leap days. They are not “real” signals within the data but rather just a result of the 14-day model output frequency being affected by an extra calendar day every 4 years. We double checked this in a few examples by changing the output frequency to hourly and then resampling to a 14-day frequency. We chose to output data on a coarser temporal resolution than the model is run at in order to keep file sizes manageable. Also, to explain this point in the manuscript, we added the following to the original Figure 7’s caption:

Also note that the approximately decadal oscillations in the time series are a result of leap years affecting the 14-day model output frequency and are not real signals within the data.

Minor and/or textual comments

l. 20: asymmetric -> temporally asymmetric (to distinguish it from spatially asymmetric, which I thought was meant when first reading this sentence)

We have changed this here and in the abstract in line 17.

l. 25: drives meltwater runoff -> drives enhanced meltwater runoff

Done.

l. 31: limit -> limit and/or delay (in the case of aquifers)

Done.

l. 32 and following discussion: Firn air content and density are presented here as two separate characteristics, but it would seem to me that they are one-to-one coupled? Or am I overlooking something?

This is a great point that another reviewer also brought up (see response to Referee #2, line 116). We have made the following changes to this paragraph:

“such as air content and density” → ***like air content***
“Changes in firn air content and firn density” → ***Changes in firn air content***

We kept one distinction between density and firn air content because we refer to specific papers that looked at one or the other variable.

l. 139: “slightly inland from the ice edge”. As large parts of ablation zone are not considered, the ice edge can be >100 km away. Perhaps better use “slightly inland from the equilibrium line”.

We have changed this to:

155 ***slightly inland from the edge of the model domain since firn is not present at the lowest elevations.***