Response to review

We would like to thank the reviewer again for their time and further comments, and are glad that they've found the revisions generally satisfactory. We will discuss their outstanding queries in more detail below, but will address two general themes first.

The remaining reviewer concerns seem to have been triggered by the new figure 8 in the revised manuscript showing the change in surface mass balance caused by the RCP4.5 climate forcing, and that the anomaly projected by FAMOUS does not have a similar pattern to the anomaly projected by the MAR RCM. We would first like to stress that what is being shown here is not the *absolute* SMB projected for the RCP4.5 forcing in each model, but the change compared to the 1980-1999 SMB base state shown in figure 6. This stress may be unnecessary, but it's not entirely clear from the review comments that they are talking about figure 8 as an anomaly plot, so we thought it best to be clear. Had we plotted the absolute SMBs for the RCP4.5 forcing for the two models (figure RR1) it would have much the same pattern as figure 6. Viewing the SMB in this absolute fashion shows a much more expected pattern of surface mass balance, but we do think it more useful to show the details of the anomaly compared to the baseline in terms of understanding how the climate of our model changes and the sensitivity of the SMB.

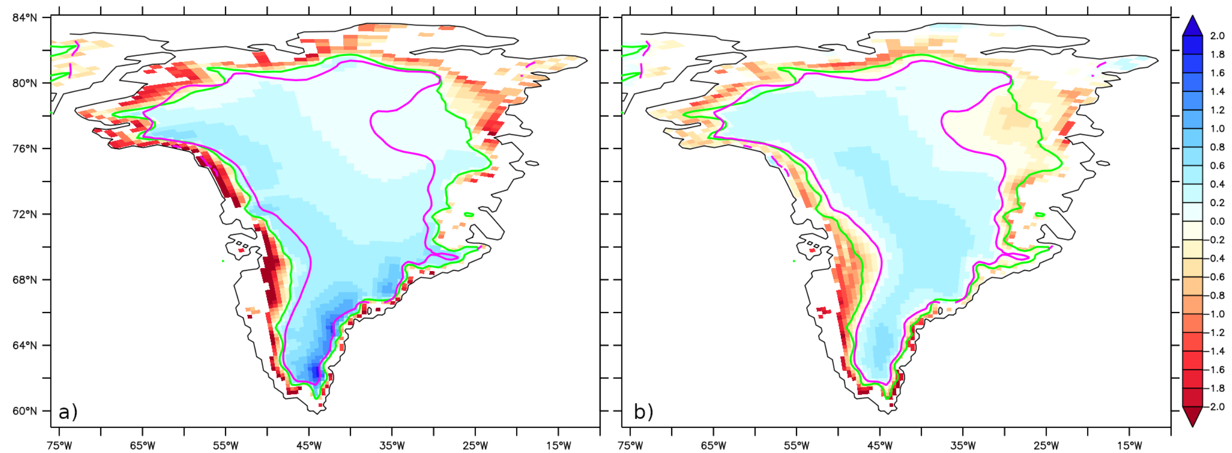

Figure RR1. Annual average surface mass balance (m/yr LWE) for the surface of Greenland under RCP4.5 in a) MAR (2080-2099 from a transient simulation forced by MIROC5 SSTs and atmospheric boundary conditions: minimum-4.02, maximum 1.98) and b) FAMOUS-ice (forced by constant 2080-2099 MIROC5 SSTs: minimum=-4.16; maximum=0.80). To visualise the distribution on sub-gridscale tiles, FAMOUS-ice results have been trilinearly mapped to the same topography as used in MAR. The green contour shows the MAR equilibrium line, the pink is the FAMOUS-ice equilibrium line.

We will proceed under the assumption that the reviewer did in fact recognise that figure 8 (and supplementary figures A12-A16) show only the changes of FAMOUS and MAR induced by RCP4.5 climate forcing compared to each model's late 20th century baseline, and that they are primarily unhappy that the physics response to climate change in FAMOUS do not closely match those simulated by MAR.

As stated previously main goal of this paper is a technical description, which we think worth publishing for its own sake. We agree that understanding the physics producing SMB and its sensitivity to climate change is important to understand, but it's not the primary aim of this paper - for FAMOUS-ice that is addressed in Gregory et al. (2020), and is being investigated further in a current project using large ensembles of parameter-perturbed model configurations. The climate results shown in our paper are from a single member of those ensembles, and whilst it is intended to be broadly representative of the behaviour of FAMOUS-ice it is not the only realisation of climate and climate change that the model can produce. We have tried to stress this more in the latest revision (eg page 7, line 29; page 14, line 9; page 24, line 18).

It is well-known that there is considerable uncertainty in the projection of future regional climate between models. Changes in clouds, and subsequently on radiative fluxes, make the single biggest contribution to that uncertainty, and since shortwave forcing is such a significant influence on GrIS surface conditions, it is not surprising that two climate models might produce quite different projections of the sensitivity of surface shortwave and melt for the GrIS, even with similar local SST boundary conditions. FAMOUS is constrained only by SSTs and is free to develop its own large-scale atmospheric circulation, whilst MAR is additionally constrained by local MIROC5 atmospheric conditions. RCMs such as MAR may be considered the "gold-standard" for simulations of the recent climate for which they have been calibrated, but that's quite a different thing from assuming that they will also have the correct cloud or climate sensitivity, something that is simply not known.

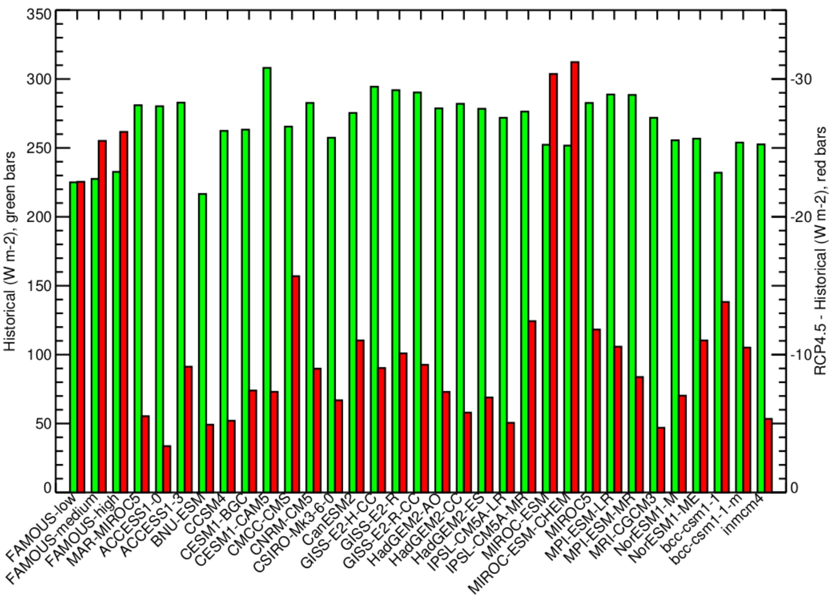

Figure RR2. Area integrated June-July-August average surface downwelling shortwave for GrIS from a range of CMIP5 and other models. Green bars: at the end of the 20th century from historical simulations; red bars: the anomaly between simulations with RCP4.5 greenhouse gas forcing and the historical

Figures RR2 and RR3 show the range of shortwave fluxes simulated for summer conditions in GrIS across a selection of climate models, mostly participants in CMIP5. The FAMOUS-ice example we have used in this paper is "FAMOUS-medium", and the MAR simulation we show in the paper for reference is labelled "MAR-MIROC5". The red anomaly bars (right axis) in RR2 shows that there is a wide variety of reductions in GrIS downwelling shortwave under RCP4.5 conditions across these models, and whilst these configurations of FAMOUS-ice do have a relatively large reduction it is not the largest. MAR is at the very other end of the range, with a very small reduction in downwelling shortwave. There is no obvious relationship between the magnitude of the flux in the historical and the response to climate change. A similar plot of shortwave absorbed at the surface of GrIS shows an even wider range of responses, with some models absorbing more radiation in RCP4.5 than in the historical, and some less. FAMOUS is again near one extreme of this range and MAR near the other, but both have shortwave responses that look reasonable when considering a broad range of climate models.

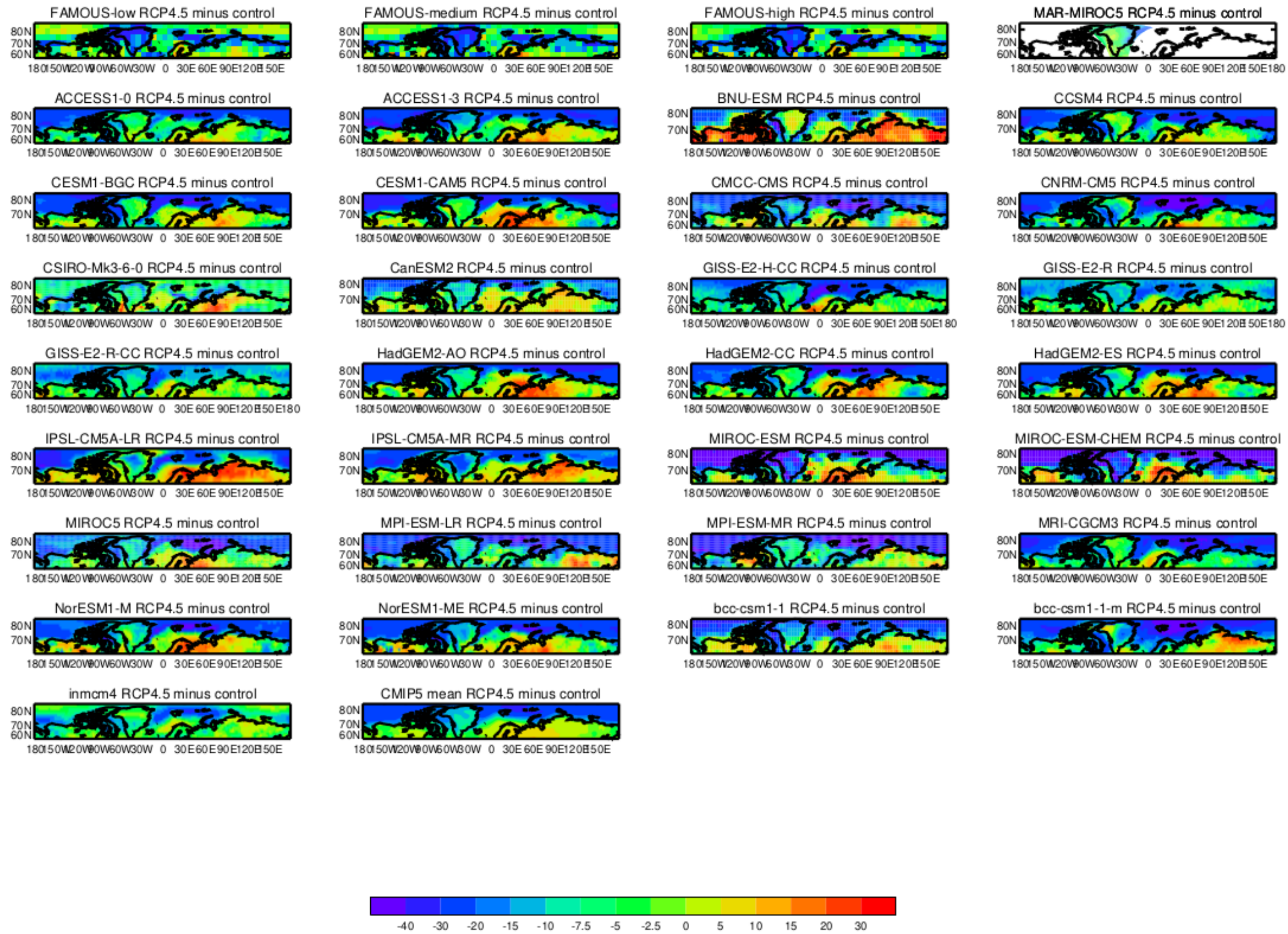

Figure RR3. Changes in June-July-August average surface downwelling shortwave (W/m-2) for GrIS from a range of CMIP5 and other models.

RR3 shows that patterns of shortwave radiation changes across these models also show a wide variety, and not just over Greenland. We have noted where the local radiative responses of FAMOUS-ice and MAR sit in the context of this broad range of model in the revised paper (page 19, line 1).

In general then, whilst we don't disagree with the need to understand the physics behind the sensitivity of climate processes and surface mass balance in our model, we are satisfied that our illustrative example is within the range of behaviours generally displayed by other GCMs and RCMs, even if the pattern of SMB change in the example we used here do not match that in MAR. It may be instructive to investigate why that mismatch exists, but it is not of itself an indication that our model or our coupling schemes are not useful for their intended purpose. That being the case, and given that the primary aim of this paper is to describe the technical aspects of the FAMOUS and how it is coupled to the ice sheet model, we don't think that going into significant extra detail of the climate behaviour of our example simulation in this GMD paper is justified - especially in the light of ongoing, large ensemble tuning work with this model that will be published in its own right when complete (noted on page 24, line 18).

We will respond now to individual aspects of the review. The reviewers' comments are in blue italics and our responses to each follow in black. Line/figure numbers given in our responses refer to the track-changed version of the resubmitted manuscript, line and figure numbers in reviewer comments refer to the revision of the discussion document that they reviewed.

Individual comments

The new Fig. 8 exhibits a poor comparison to the SMB from the regional model MAR. Even though I am aware that model evaluation on the basis of other models is problematic, I would in this context consider MAR to be some kind of gold standard. The disagreement is substantial and I have doubts that the model is reliable, especially as FAMOUS was using the same ocean forcing as used in MAR. I am still not sure if here the climate model or the surface module is the problem, and this question should be answered.

We've outlined our views on the reliability of model physics in FAMOUS compared to other models in the preamble above, and the current revision of the paper already contains a discussion of whether model biases are down to the background climate in FAMOUS or our approaches to downscaling and coupling. With regard to the particular pattern of the SMB anomaly simulated by the two models under climate change, we would note two additional points.

One is that there is additional variation in this pattern in FAMOUS with regard to the degree of climate change. In the paper we used a MIROC5 RCP4.5 forcing since we had directly comparable simulation data available from both FAMOUS and MAR. Forcing our FAMOUS-ice configuration with stronger RCP8.5 SSTs (also from MIROC5) produces an SMB anomaly field with a stronger pattern of increased melt at the margins of the ice sheet (figure RR4,b), looking more like the MAR pattern (see also Fettweis et al., 2013) compared to the relatively ubiquitous melt anomaly in our FAMOUS-ice example under RCP4.5. The variation in regional change responses to climate change across models and scenarios is considerable. In general, the fixed ice sheet mask of an atmosphere-only model like MAR can diagnose high negative SMBs at the ice sheet margin that may not be realistic, since the ice is modelled as a feature that can never disappear however much it is forced to melt. Although the ice extent changes in our illustrative simulation are not large at the point in the RCP4.5 simulation where we diagnosed the SMB anomaly in figure 8, it remains true that care must be taken interpreting details at the ice margins of models with different ice sheet masks.

We would further note that FAMOUS-ice is intended for long-term simulations of coupled climate-ice evolution. In an atmosphere-ice coupled model like FAMOUS-ice, the area integrated mass balance is often more important than its gradients, since ice dynamics will respond to determine the shape of the ice sheet. Table 1 shows that the integrated SMB component magnitudes are not so different from MAR's, and Gregory et al. (2020) show that the area-integrated magnitudes of SMB and its components in FAMOUS-ice are similar to RCMs over a wide range of 21st century climate forcings, so from that point of view we are happy that our model is reliable and useful for its intended purpose.

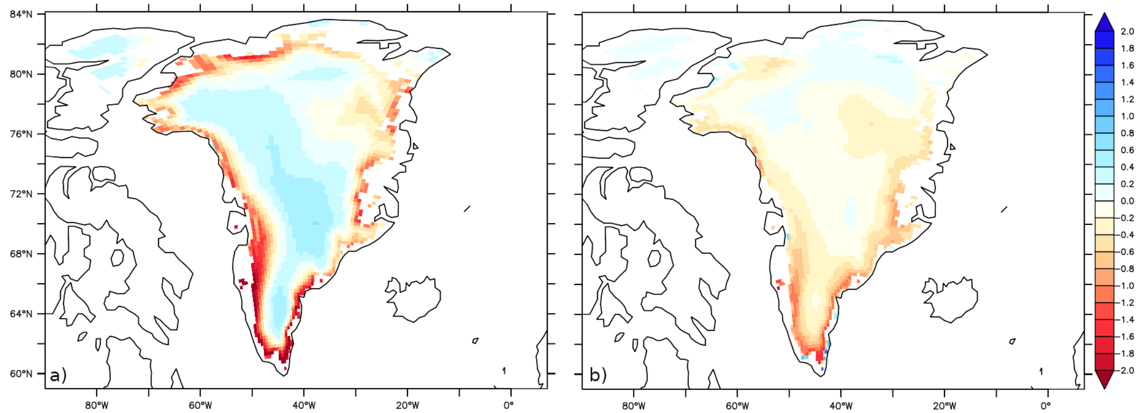

Figure RR4. a) Surface mass balance (m/yr LWE) for the surface of Greenland for MIROC5 SSTs from years 2080-2099 under RCP8.5 in FAMOUS-ice b) Change in this SMB compared to the 1980-1999 baseline in figure 6

If the climate model is the problem (it seems there is too much cloud cover), it would be helpful to also show the near surface temperatures. But neither in this manuscript nor in Gregory et al 2020 I have found an evaluation of the 2m air temperature distribution. (I take it that the model is generally slightly too warm over Greenland).

We agree that surface air temperature would be a useful field to show, and have now re-run our FAMOUS-ice example to produce this diagnostic. Since FAMOUS-ice uses an surface energy model to calculate snow melt the surface air temperature is not directly correlated with melt, of course. Surface air temperatures for the 1980-1999 reference period for MAR and our example FAMOUS-ice simulation are shown in new figure A7, with the response to RCP4.5 climate change in figure A12. A7 shows that FAMOUS is generally a little warmer than MAR, particularly at height. Since the low resolution of FAMOUS cannot support tight horizontal gradients this is not very surprising, although the specified lapse rate between the elevation tiles will play a role here too. As is already shown, the shortwave radiation absorbed by GrIS in FAMOUS-ice and MAR is rather similar so this is probably not the reason for a warm surface bias. The climate change response in air temperature is quite different to MAR, which we would expect from differences in the response of the cloud and shortwave we've already talked about. In A12 MAR shows a general warming everywhere on the ice sheet, whilst FAMOUS-ice warms more strongly with height. This may indicate a change of atmospheric lapse rate in FAMOUS in the warmer climate (the lapse rate between tiles is fixed), but this pattern is also correlated with the change in net shortwave absorbed at the surface. The net shortwave change is determined by the overlap between a region of reduction in albedo in the north and east, and the change in downward shortwave (figure A16,b) controlled by the cloud response over GrIS in FAMOUS-ice. Some discussion of this has been added to the revised paper (page 7, line 35; page 18, line 1).

Furthermore I wonder if the disagreement in shortwave radiation is related to the coarse resolution rather than to a "wrong" background climate (in that case downscaling of the atmospheric transmissivity might be considered). Is the shortwave radiation also biased outside of Greenland? How would Figure 2 look if the native FAMOUS resolution was used and MAR/RACMO was also remapped conservatively to the same coarse resolution?

In development we didn't find that the downwelling shortwave varied with height in FAMOUS in a manner that suggested a significant scaling with tile elevation would be physically justified (already noted on page 7). FAMOUS data in figure 2 is plotted from the FAMOUS tiles rather than interpolated values and we are not convinced that we would learn we do not already know about the shortwave biases in FAMOUS by remapping and rebinning everything onto the coarse resolution grid of the FAMOUS atmosphere. Figure RR3 (using the native grid for all models, including FAMOUS) shows a wide variety of shortwave responses to climate change across CMIP5-era models, and we do not think that our FAMOUS example is uniquely different in any part of the northern hemisphere in this context. We consider analysing climate biases outside of GrIS beyond the scope of the current work.

How does figure 2 look for the 2080-2099 period?

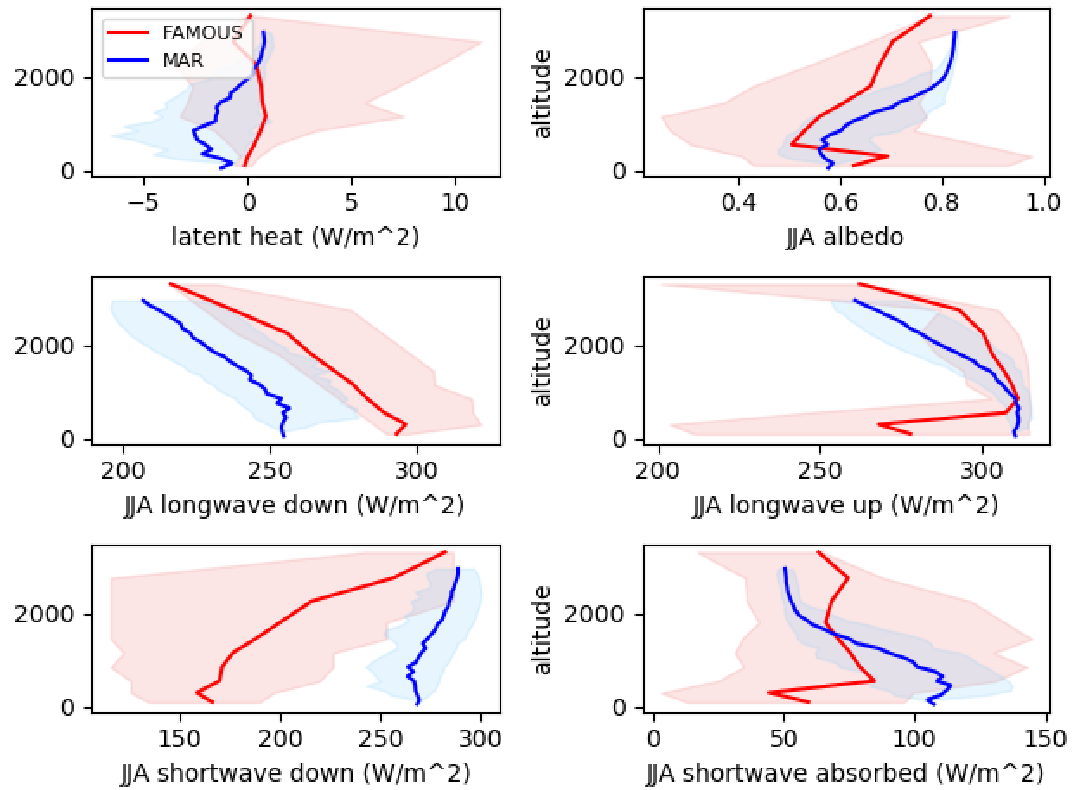

Figure RR5. Profiles of surface exchange characteristics taken from every surface modelled on the Greenland ice sheet binned by surface elevation. Red: FAMOUS-ice; blue: MAR (2080-2099). FAMOUS-ice and MAR are forced by the MIROC5 climate (2080-2099). Latent heat is an annual average, negative values imply sublimation. Shading represents the full range of values found in each time-averaged elevation bin.

For reference here it is, although we feel that we have already discussed whether it is necessary to analyse further the differences in climate change response between FAMOUS-ice and MAR in the context of this paper, so it has not been added to the current revision. The variation of albedo with height matches better in the models for this period, the fundamental biases in downwelling long and shortwave remain. Since the albedoes match better, it thus follows that the shortwave absorbed will now differ more between the two models.

In any case it is worrying that the albedo does not exhibit the typical spatial pattern. After tuning it is likely that albedo compensates biases in the climate, and in consequence the sensitivity of the surface scheme would be biased, too (good results for the wrong reasons) which could also explain the poor agreement in Fig. 8.

It's not clear whether this comment refers to the albedo of the 20th century state in FAMOUS-ice or the albedo response under climate change. Assuming we're talking about the former (and referring to our responses above for whether we think agreement a necessary thing to achieve for what we show in figure 8) the paper does already discuss these matters. The albedo tuning is indeed a partial compensation for the downwelling shortwave bias, but it doesn't necessarily follow that the general climate sensitivity of the model will be biassed in a particular direction as a result of this nor, as we have seen above, that it causes the model to have a climate response that is exceptional either in pattern or magnitude in the context of a broader range of climate models.

The modern-day simulation of this FAMOUS-ice configuration does have a lower albedo at the margin of the GrIS than MAR (figure 4), so some of the margin is already near the lowest albedos that ice can realistically have. Under a moderate climate change scenario like RCP4.5 it is thus possible for MAR to simulate larger albedo reductions of the lowest-lying parts of the ice margin. This will undoubtedly contribute to differences in the pattern of SMB simulated in these models, although changes in downwelling shortwave and other energy components are also important. However, this difference in potential albedo response only affects the very edge of the ice and does not necessarily dominate the SMB response of a model over longer periods of time (when prolonged large negative SMB near the ice margin should remove those areas from the ice sheet and coupled flow dynamics become important) or under a stronger climate forcing which would affect larger areas of ice (figure RR3). A note discussing this has been added to the text (page 18, line 6).

Finally I would like to point out that Kapsch et al, 2021 also present Greenland SMBs based on T31 resolution simulations, which seem to be fine; and from my own experience, PDD schemes are likewise able to produce a reasonable SMB response based on lapse rate corrected T31-temperatures (if insolation is similar to today's Greenland and temperatures are reasonable).

I don't think we would dispute this, and would argue that we have shown FAMOUS-ice also produces reasonable simulations of present-day Greenland SMB, and can produce plausible future projections. The assumption made in PDD schemes of a climate- and geographically constant degree-day factor, even though it might give a reasonable SMB simulation for observed climate, introduces a potentially large and unquantifiable systematic uncertainty in modelling other climates, whereas an energy balance approach such as the one we (and Kapsch et al.) use should be more generally applicable.