Response to reviews, edited from the GMDD author comment to accompany resubmitted ms.

We would like to thank both reviewers for their thoughtful comments on our paper. We were glad to see that both thought a revised manuscript would make a timely and important contribution to the field as a full GMD paper. We agree that additional analysis and content along the lines they suggest would improve our model description, and here note in detail changes we have made in response to each point raised.

As written, our paper was intended primarily as a technical description of how our novel FAMOUS-to-ice-sheet coupling has been implemented, with a focus on the downscaling techniques and enough illustrative detail in the results to show that our coupled model simulations could be scientifically plausible even with the low native resolution of the FAMOUS atmosphere. We are pleased to see that Reviewer 1 thinks we have succeeded in this aim, at least for a modern climate state (first blue sentence below). Given this intention, we are not keen on having this paper include extensive analysis and evaluation of any individual climate simulation where that would not speak to our primary technical focus. Certain issues that significantly affect the setup of a simulation with a coupled climate-ice model (for instance spinup/initialisation techniques, which are still the subject of much ongoing research) we considered to be out of scope for this paper. It's clearly not sensible to suggest that a modelling technique should be described without giving some evidence that its results are likely to be fit for some stated purpose if applied appropriately, and the results and analysis we have done are framed with this in mind. We wish to demonstrate the capabilities of the coupling techniques in the model and describe typical results, rather than evaluate the large-scale climate of the FAMOUS atmosphere or the general behaviour of the Glimmer ice sheet (both of which have been done in previous studies) under any particular set of boundary conditions.

Below, the reviewers' comments are in blue italics and our responses to each follow in black. Line numbers given in our responses refer to the track-changed version of the resubmitted manuscript, line numbers in reviewer comments refer to the original discussion paper.

Reviewer 1

The presented analysis demonstrates that FAMOUS-ice in combination with an ice sheet model is capable to simulate a reasonably realistic GrIS under present day climate conditions. The surface mass balance, however, appears to be extended to too high elevations and albedo seems to be overestimated in the ice sheet's interior. Given that other ESMs of similar design and resolution exhibit a qualitatively better skill (e.g. Kapsch et al., 2020; Fettweis et al., 2020), the analysis of the model should answer the question, whether these shortcomings are the result of a biased climate or of a misrepresentation of specific processes at the ice sheet's surface.

It is true of course that the simulation of surface mass balance components in FAMOUS-ice contains biases. We would note that the comparison with the ESMs contributing to GrSMBMIP (Fettweis et al., 2020) is perhaps too stringent a test for FAMOUS-ice - both CESM2 and MPI-ESM are much more physically complex and computationally expensive CMIP6-class models run at significantly higher resolution (CESM2's GrSMBIP submission was from a 1 degree model (37 gridboxes for every 1 in FAMOUS) and 32 vertical levels (almost 3x as many as FAMOUS) and MPI-ESM submitted results from an even higher resolution model - T127, 95 vertical levels.

Regardless of what is used as a comparison or its computational cost, FAMOUS-ice must be shown to stand on its own merits as suitable for its intended purpose, which we think it does - it seems Reviewer one agrees (their first sentence). Their question of whether the SMB, albedo and other GrIS biases that we do have are rooted in the large-scale climate or our downscaled surface modelling is a good one though, and very relevant to what we want to address. This is now addressed in a substantial new paragraph in the Discussion (page 23).

Fig. 2 indicates that the downward shortwave radiation is strongly biased towards lower values while downward longwave radiation does not exhibit any clear bias. As argued by the authors, this may indicate an overestimated cloud cover, which should be substantiated by decomposing long wave radiation into atmospheric emissivity and near surface temperature. A potentially biased cloud cover should be reflected in atmospheric emissivity. Near surface temperature influences both, downward longwave radiation and turbulent heat flux and is an important driver of the SMB, which should be analysed as well. Given that the spatial distribution of surface mass balance and albedo from FAMOUS-ice differ strongly from respective fields simulated *by MAR, I would propose to additionally show selected surface exchange characteristics (Fig. 2) as 2-D fields in comparison to MAR. It is important to not only analyse the spatial characteristics of the simulated SMB but to also evaluate the model's response to climate variations. The model is explicitly designed to simulate fundamental climate change and consequently this paper should allow to assess the model's skill to respond to climate change. This could be accomplished by using a transient ocean forcing from a 21st century climate projection.*

For the modern simulation originally addressed in our paper, there is now more discussion of the downscaled surface fluxes, including turbulent fluxes, starting on page 7, line 29, and 2D spatial distributions of a number of relevant fields are now included as Supplementary material. Downwelling longwave is now plotted for JJA for easier comparison to the shortwave, and a compensation between the two radiative components can simply be seen for the FAMOUS-ice bias wrt MAR - see also Reviewer 2's comments on this. The scope of our analysis is somewhat limited by the availability of comparable diagnostics from the model simulations in question. The original FAMOUS-ice simulations used to illustrate this paper did not output usable diagnostics of sensible heat flux, so new ones have been run to provide this field. By necessity these were conducted on a different computing platform and are not perfect reproductions of the earlier simulations. For this reason, some of the minor detail in the FAMOUS-ice figures has changed in this revision, as have the FAMOUS-ice data in Table 1. The 5-letter code (xotzb) used in the title to denote the precise configuration of FAMOUS used has also changed to match.

The sensitivity of the coupling scheme to climate change in FAMOUS is a matter raised by both reviewers. This is clearly an important topic given the intended use of the coupled model, however given the caveats towards general simulation evaluation given in our preamble we are wary of analysing the model climate too far down this path in this paper. The SMB response will be controlled to a large degree by the background climate response of FAMOUS to climate warming, which has been studied previously in the literature. We note that Gregory et al. (The Cryosphere, 2020) already explores in some depth the response of an ensemble of the coupled FAMOUS-ice Glimmer system to a range of climate change forcings, so our additions will take that paper as context and provide complementary detail. Consequently, taking one future climate scenario as illustrative, we have created a new subsection (7.2, page 17) looking at how some of the SMB component fields change under a warmer climate, and added new data to Table 1 of integrated SMB and its components (page 17).

Furthermore, I found the manuscript at times hard to read,as essential information with respect to the snow pack schemes are only found in referenced literature:Page 10: It would be helpful to briefly summarize which processes influence the grain-size of snow and to generally characterize the relationship between snow albedo and climate forcing, and to give typical grain sizes of new snow and old/wet snow. (Alternatively, a figure illustrating albedo as a function of influencing factors could be beneficial.) This is particularly desirable as Marshall (1989) is not easily available and as Gardner and Sharp (2010) indicates a sensitivity to grain size (0.05 /mm) towards greater grain sizes, which seems to be considerably smaller than the values used here

Influences on the grain-size of snow and what happens in our snow model are now described on page 9, line 31, and a new figure 3 shows how our resultant broadband snow albedo varies as a function of grain-size. Compared to the original parameters in this version of the snow scheme in JULES, our model has an increased sensitivity to grain size (new figure 3). The changes were made for the visible part of the two-stream radiation scheme in FAMOUS. Snow grain size is limited to 2000μm, which leads to an associated minimum in albedo. These adjustments were made to provide a more realistic representation of seasonal snow albedo evolution (e.g., Stroeve (2013) https://doi.org/10.1016/j.rse.2013.07.023). Additional observational studies were referenced, as was work on albedo optimisation in two-stream radiation code in ECHAM4 (Roesch(2012), 10.1029/2001JD000809). The tuning parameter for which values are given in the main text, $\Delta a_{\mathrm{snow,visible}}$, is not directly used as the sensitivity of albedo to the grain size, but as part of a larger formulation. It is thus not directly comparable to the 0.05 /mm inferred from Gardner and Sharp (2010) by the reviewer. As can be seen from the orange line in the new figure 3, the albedo sensitivity to grain size that results from our formulation is in fact similar to this 0.05 /mm estimate, although the gradient is not constant across the full range of grainsize.

I also missed a separate short paragraph on how the snow pack evolves over time:How does density change if snow accumulates on top, what happens if the first layer melts, how and when are densities of the snow column prognostically calculated, what is the typical depth of compacted snow?

There is now a paragraph on page 9 (starts line 7) describing the evolution and relayering of the snow pack model as mass is added or removed.

I also recommend to improve the structure of the paper by a further break down of sub-sections. Separate subsections might cover initial conditions, upper and lower boundary conditions of the snow pack as seen by the land surface model, and boundary conditions at the ice surface as seen from the atmosphere. Finally the analyzed experiments, together with the MIROC forcing and the MAR and RACMO simulation should be introduced in a separate section, possibly before section 4

We have no objection to reconsidering the subdivision of material in the paper. However, as noted already, we would like the focus to remain on a technical description of the methods used and the general capabilities of the model rather than a detailed evaluation of one particular climate realisation. Some of the suggestions made in this comment would not fit with our technical focus - for instance, description of initial conditions for the simulation, and on reflection we have decided to keep the structure much as it was.

p. 5, l. 8: Is it possible to quantify the computational cost of the new model version in comparison to the old version?

Computational expense of this version of FAMOUS are now in the Discussion, page 24, line 23.

p. 5, l. 17: "We will describe later..." Please refer to the respective section.

done (here and elsewhere)

p. 7, l. 9: high-latitude -> high-altitude?

either would fit, I think, for this example. By "high-latitude" we meant latitudes near the poles

p. 7, l. 11: Is the sub-surface lapse rate in line with the coarse resolution subsurface temperature distribution?

That's a good question. This is noted on page 7, line 20

p. 11, l. 12: (fig:4) -> (Fig. 2) ?

yes, that's a typo.

Fig. 3, 5,6, : As the colorbar does not cover the full range of values, maybe include the range of values in the figure caption.

good idea, this has been done for all figures with spatial distributions in

p. 12-14: does the spatial representation of SMB, albedo improve qualitatively if the number of elevation classes is increased?

This is discussed now in the Discussion on page 23, line 29

p. 12, l. 9: It should be stressed that FAMOUS and MAR are consistently forced with the climate model output while RACMO is forced by reanalysis data. Also please specify the ERA forcing used and the period which was analysed.

this has been done

p. 14, Tab. 1: Please specify which ERA forcing is used here, Fettweis 2013 only has RACMO(ECMWF), and the RACMO(ERA) experiment that was used in this paper should be added to this table

I'm actually confused by the notation in Fettweis 2013, which seems to me to use $\mathrm{RACMO_{ECMWF}}$ and $\mathrm{RACMO_{ERA-Interim}}$ in the same sense in places. Our Table 1 has been changed to use the same notation as the table in Fettweis 2013, but have used the same citation for ERA-Interim for both MAR(ERA_interim) and RACMO(ECMWF).

Reviewer 2

You argue that Helsen et al., 2012 developed an empirical parameterization to translate global climate fields to usable boundary conditions. However, they state that their method is not usable for models where ablation areas are not resolved.

this reference has been removed

How does FAMOUS-ice prevent snow from growing infinitely thick? Does it include a firn-to-ice densification scheme, or do you cap the snow above a certain thickness?

The snow scheme does include snowpack compaction and densification processes, although over time this alone would not prevent infinite columns of ice-density "snow" from forming where local accumulation is greater than ablation. In FAMOUS-ice we use the coupling scheme as outlined in section 8.1, whereby annual changes in snow mass are removed from FAMOUS every year and given to/taken from the mass of the ice sheet model via the SMB term. This is all clarified in the new paragraph on the evolution and relayering of the snowpack on page 9.

For downscaling of the temperature and (incoming?) longwave radiation from the atmosphere to the elevation classes you use an empirically tuned method, where you assume that the near-surface climate gradients are similar to the climate gradients *in the lower atmosphere. Is this a valid assumption? For example, the temperature decreases with elevation along the GrIS near-surface, while the lower atmosphere has a temperature increase with height due to inversion.*

The surface temperature and downwelling longwave are downscaled with elevation using a fixed lapse rate. This is commonly done in this sort of downscaling (see references in the original paper), although it is undeniably a rather simplistic method and cannot reflect local conditions such as inversions that deviate from its cooling-with-height paradigm. The value used for this lapse rate is sometimes treated as a tunable parameter in order to achieve a realistic surface temperature distribution on the ice sheet. Used, as we do, to reflect subgrid-scale surface variation in a single land model gridbox, this near-surface-like scaling is more justified than trying to use the state of the free atmosphere column, especially in a lower-resolution model like FAMOUS where the vertical resolution in the atmosphere is rather low. There is some new clarification of our approach to this on page 6, line 14.

I am not convinced you improve on the elevation class implementation analysed by Sellevold et al., 2019, as you are not comparing the same metric - that is, you do not subtract the grid-cell average from the same grid-box elevation class simulated value.

We agree that the metric used in Sellevold et al. is more stringent than the vertical profiles shown in our original figures. We have now reproduced the Sellevold metric for selected downscaled fluxes on the ice sheet, and included it in the supplementary material (figure A6). This supports our claim that all the gradients act in the same sense in FAMOUS-ice as in the RCM used in Sellevold et al., although there are differences in the sensitivity of certain quantities to the elevation difference between each tile and the gridbox mean.

Below 1000 m, the latent heat flux is positive in FAMOUS (Fig. 2), while negative in MAR and RACMO. First, it should be indicated whether the positive values means energy loss or gain at the surface. I am assuming the latter, as MAR and RACMO simulate climatological mean sublimation at lower elevations. Can you explain why FAMOUS simulates climatological mean deposition at these elevations?

Thanks for this - there was a sign error in converting the FAMOUS-ice sublimation diagnostic into latent heat in this plot. Latent heat fluxes on Greenland in FAMOUS-ice are generally very small - this is noted in some additional discussion on page 8, line 2, and shown in the new supplementary figures.

You argue that the shortwave down is too low because of a high biased cloud cover. I am surprised to see that this is not compensated for by an overestimation of longwave down. Is this due to the atmosphere being warmer in FAMOUS?

Figure 2 now shows JJA longwave for more straightforward comparison with JJA shortwave, and some compensation can indeed be seen in the differences between FAMOUS-ice and MAR. There is no such longwave/shortwave compensation when comparing RACMO and FAMOUS-ice, but RACMO here is forced by the ERA-Interim reanalysis, so is subject to a very different background climate and this comparison is not necessarily to be expected.

The simulation of albedo shows clear differences, particularly at higher elevations,when compared to MAR. You argue that the trigger for the lowered albedo at the higher elevations is a warm temperature bias at higher altitudes. Including a (supplementary) figure of April/May near-surface temperature or snow grain size compared to MAR/observations would likely strengthen this argument. Further, is it possible to assess whether the albedo is triggered too much and/or too early in the melt season?

Further detailed analysis of different parts of the albedo parameterisation have shown that snow grain influences are not producing the widespread low summer albedos in these simulations. Instead, they come from the bare-ice values that are used when very dense firn appears at the surface, which seem to be used too readily. A similar feedback exists, with warmer temperatures lowering the albedo further via the bare-ice melt-pooling parameterisation and excessive melt potentially contributing to refreeze and increased densification of the snowpack underneath which may be exposed to the surface later. The change in attribution of the low albedo from snow-grain processes to the bare-ice parameterisation is noted in a new paragraph on page 13.

Also, if the atmospheric temperatures are too high at high altitudes, the melt energy contribution from the sensible heat flux should also be higher. As you don't show the sensible heat flux, I am left to wonder how the simulation of the sensible heat flux is.

Sensible heat fluxes are very low in FAMOUS-ice. This is noted now on page 8, line 2, and shown in a supplementary figure.

How well is rain simulated over the ice sheet? As the ice sheet is warm at higher altitudes, is there an overestimation of rainfall at higher altitudes? Could this also contribute to a lower albedo?

Rainfall distributions are now shown in a supplementary figure, and there is some discussion of precipitation bias in the Discussion on page 23, line 17. The model parameterisations mean that rain cannot affect the albedo calculation, as explained now on page 13, line 13.

The presented surface mass balance shows a realistic distribution with ablation at lower elevations and accumulation at higher elevations. However, the ablation areas are biased large and the high accumulation areas in the Southeast and Northwest show too little accumulation. As you don't present a transient climate simulation, I think the possible effect of these biases in transient simulations deserves a paragraph in the discussion section.

Reviewer 1 also requested some consideration of how the model responds to climate change (See above) - there is now a new subsection (7.2) discussing SMB sensitivity, and number of supplementary figures showing the spatial distribution of changes to key fields under climate change forcing.

It would be easier to understand some of the thickness anomalies in Fig. 8 if you showed the ice velocities. Consider adding it to supplementary materials.

this has been done, figure A15

I found many occurrences of double parentheses with citations. If latex is used, use(text before citation) or to avoid.

these have been fixed

P. 11, l. 12: replace fig:4 with figure 4.

fixed

P. 19, l. 2: "a model"

fixed

P. 19, l 2: "allow for coupled"

I think this sentence is actually grammatically OK as it stands?

Fig. 7, caption: replace x in: 10x km

fixed. Thanks!