

Response to reviewers:

Dear reviewers, we would like to thank you for your constructive comments on our manuscript. We tried to address most of the concerns you raised, and to apply corrections where appropriate to improve our manuscript.

To facilitate readability, our responses to reviewers are displayed in blue and modifications in the manuscript are highlighted in red. These suggested changes, together with additional minor corrections, are also displayed in red in the attached revised manuscript.

Reviewer #1:

This paper presents an updated version of the polar version of the RACMO2 regional climate model (RACMO2p2), evaluated over the Greenland ice sheet against various observational datasets. Model updates include changes to the concentration of impurities assumed to be deposited onto the snowpack, the grain size of snow that has experienced meltwater refreezing, the albedo of superimposed ice, and the saltation coefficient for drifting snow. These changes generally result in an improved agreement between modeled and observed atmospheric variables, radiative fluxes and SMB. Some biases persist and can be corrected by future improvements to model physics and parameterizations and/or downscaling of model outputs.

General Comments

The paper is well written and the presentation is generally clear. The paper is not novel in the sense of presenting new model physics or parameterizations, but the changes to RACMO2 that are presented seem to have a significant impact on representation of the Greenland ice sheet surface and the agreement between modeled and observed SMB. A detailed validation of this updated version of the model has been conducted. The paper is therefore likely to be of interest to the cryospheric and climate modeling communities, and is important in providing details about and validation of a new version of RACMO that will be used for future studies of Greenland mass balance. I feel that the paper should be published in the Cryosphere after some relatively minor concerns are addressed below.

One general concern is the use of net surface energy balance to indicate energy available for melting. If the snow temperature is below 0°C, this energy must first be used to warm the snowpack before contributing to melting. In the ablation zone during summer where temperatures are close to freezing, most of the net energy goes to melting, but some of it must go into warming the snow/ice. The authors should revise the manuscript to refer to the net energy balance rather than melt energy, or explain why they can assume that the net energy balance can be considered melt energy.

In RACMO2, the skin temperature (T_{skin}) of snow and ice is derived by closing the surface energy budget (SEB), using the linearized dependencies of all fluxes to T_{skin} and further assuming, as a first approximate, that no melt occurs at the surface ($M = 0$). If the obtained T_{skin} exceeds the melting point, T_{skin} is set to 0°C; all fluxes are then recalculated and the melt energy flux (M) is estimated by closing the SEB using Eq. 1. Therefore, snow melt is possible even though the upper snow layer remains below the melting point. In that case, the generated meltwater refreezes directly in this upper snow layer, rapidly warming it to 0°C. This is now clarified in the manuscript: “In RACMO2, the skin temperature (T_{skin}) of snow and ice is derived by closing the surface energy budget (SEB), using the linearized dependencies of all fluxes to T_{skin} and further assuming, as a first approximate, that no melt occurs at the surface ($M = 0$). If the obtained T_{skin} exceeds the melting point, T_{skin} is set to 0°C; all fluxes are then recalculated and the melt energy flux ($M > 0$) is estimated by closing the SEB in Eq. 1., assuming that no solar radiation can directly penetrate the snow or ice interface.” As the term “melt energy” is commonly used in multiple publications, e.g. Van den Broeke et al. (2017), Van Wessem (2013, 2017), Ettema et al. (2010), and to remain consistent with previously published RACMO2 papers, we decided to keep M to refer to melt energy.

Regarding comparisons between RACMO2.3p1 and observations for some of the plots, it would be useful to see how much RACMO2.3p2 improves on the previous version with respect to meteorological variables. Adding RACMO2.3p1 to Figs. 3, 4, and 5 could potentially make them difficult to interpret, but I think that the authors should at least provide some statistics with regard to the meteorological comparison (corresponding to statistics for Figs. 3 and 4). It would be nice to see the corresponding figures for RACMO2.3p1 in the supplementary material as well. For RACMO2.3p1 outputs corresponding to Fig. 5, there is a similar figure in Noël et al. (2015) as the authors mention, and I think the tables provide enough information to understand the improvement.

As RACMO2.3p1 only covers the period 1958-2015, a direct comparison with RACMO2.3p2 (including 2016) cannot be conducted without discarding a substantial amount of data. Therefore, we decided to keep Figs. 3 and 4 as are and included similar figures for RACMO2.3p1 (1958-2015) in a Supplementary Material (see Figs. S5 and S10). These figures also list statistics for the radiative fluxes and meteorological data. We included a new table (Table 1) summarizing the main statistics in the revised manuscript together with the following sentences in Section 3.1 and 3.2, respectively. “Table 1 and Fig. S5 compare the agreement of RACMO2.3p2 and version 2.3p1 with in situ measurements. We find an overall improvement in the updated model version, showing a smaller bias and RMSE as well as an increased variance explained. Notably, the remaining negative bias in 2-m temperature (Fig. S5a) and the systematic dry bias (Fig. S5b) in RACMO2.3p1 have almost vanished in the updated model version (Figs. 3a and b).”

“Compared to the previous model version (Table 1), changes in the cloud scheme have significantly improved the representation of LWd (Figs. 4c and S10c), showing a reduced negative bias and RMSE. These modifications have also somewhat decreased the positive bias in SWd (Fig. 4a), relative to RACMO2.3p1 (Fig. S10a). In addition, LWu is notably improved in RACMO2.3p2: the remaining negative bias in LWu has almost vanished (Figs. 4d and S10d). This can be partly explained by the much better resolved 2-m temperature in RACMO2.3p2.”

In Section 3.3 and Tables 1-5, the signs of biases for upward and downward fluxes and interpretation in the text are confusing and sometimes inconsistent. It seems that the biases are generally considered with respect to the absolute value of fluxes (e.g. a negative bias for an upward flux is an underestimate of the upward flux) but this is not always the case. The calculation of net fluxes is also inconsistent. Mostly, the upward flux bias is subtracted from the downward flux bias, but not always. The authors should make sure that the signs for all biases are consistent, and that the text interprets the direction of the biases correctly. The authors could use the same conventions for both fluxes and biases, but this should be made clear in the figure captions and in the text, to remind the reader that a positive bias for an upward flux is an underestimate of the upward flux. For example, if SWu (Obs) is -70.9, and the RACMO2.3p2 bias (RACMO2.3p2 – obs) is 4.5, the authors can indicate that the magnitude of SWu is underestimated by 4.5 in RACMO2.3p2. Alternately the authors could define upward fluxes as being positive, which is more intuitive, and Eq. 1 could be changed so that upward fluxes are subtracted rather than added.

As suggested, we decided to adopt the convention that all SEB fluxes are set positive. We corrected equation 1, values in the main text and in the tables accordingly. We also noticed a small bug in the script calculating statistics at site S10; this has been corrected in the text and associated table. We also made sure that LWn and SWn biases are correctly defined, and clarified when necessary if fluxes are underestimated.

The section on the Northeast Greenland ice stream seems a bit out of place with respect to other sections. The authors should provide some more background at the beginning of Section 4.2 discussing how the ice discharge measurements can be used to evaluate SMB. The discharge measurements should also be mentioned in Section 2.5. Also, the authors should provide a rationale for their assumption of equilibrium between SMB and discharge for the northeast Greenland basin for the 1958-2015 period. The better agreement with discharge measurements suggests an improvement to SMB, but doesn't necessarily prove that SMB is accurate. This should be clarified in Section 4.2.

We clarified as follows, in Section 2.5: “In addition, we compare modelled SMB with annual glacial ice discharge (D) retrieved from the combined Zachariae Isstrøm and Nioghalvfjærdsbrae glacier catchments in northeast Greenland (1975-2015; yellow line in Fig. 6a), presented in Mouginit et al. (2015).”.

In Section 4.2: “For northeast Greenland's two main glaciers, Zachariae Isstrøm and Nioghalvfjærdsbrae (79N glacier; yellow line in Fig. 6a), solid ice discharge (D) estimates are available for the period 1975-2015 (Mouginit et al., 2015). Assuming that this glacier catchment draining ~12% of the GrIS area remained in approximate balance until ~2000 (Mouginit et al., 2015), i.e. D = SMB, measurements of D at the grounding line of these marine-terminating glaciers can be used to evaluate modelled SMB.”.

Followed by: “However, it is important to note that, while good agreement is obtained between averaged SMB and D before 2001, suggesting a glacier catchment in approximate balance as in Mouginit et al. (2015), this does not necessarily confirm that spatial and temporal variability of northeast Greenland SMB is accurately resolved by the model.”.

Specific Comments

P. 1, Line 10: Be more specific here. How are the patterns “better resolved”?

We reformulated as: “[...] better resolves the spatial patterns and temporal variability of SMB compared with the previous [...]”.

P. 1, Line 13: “future climate scenario projections” is unclear. Do the authors mean “projections of GrIS climate and SMB in response to future climate scenarios”?

We reformulated accordingly. Thank you.

P. 2, Line 19: Change “model simulations” to “RCM simulations”

Thank you.

P. 2, Lines 25-26: This phrase is confusing. Also, I don't believe the Box (2013) approach used data assimilation. Suggest revising “and data assimilation ... accumulation measurements...” to read: “and reconstruction of SMB obtained by combining RCM outputs with temperature and ice core accumulation measurements...”

We reformulated accordingly.

P. 2, Lines 27-28: Add reference detailing CESM future simulations (Vizcaíno et al., 2014): Vizcaíno, M., Lipscomb, W. H., Sacks, W. J., and van den Broeke M.: Greenland surface mass balance as simulated by the Community Earth System Model. Part II: Twenty-first-century changes, *Journal of Climate*, 27, 215-226, doi: 10.1175/JCLI-D-12-00588.1, 2014. The authors might consider citing this study that presents a simulation of SMB from the GEOS-5 model: Cullather, R. I., Nowicki, S. I., Zhao, B., and Suarez, M. J.: Evaluation of the surface representation of the Greenland ice sheet in a general circulation model, *Journal of Climate*, 27, 4835-4856, doi: 10.1175/JCLI-D-13-00635.1, 2014.

Thank you, reformulated as follows: “Vizcaíno et al. (2013, 2014) and Cullather et al. (2014) respectively used the Community Earth System Model (CESM) at 1° resolution (~100 km) and the Goddard Earth Observing System model version 5 (GEOS-5) at 0.5° resolution (~50 km) to estimate recent and future mass losses of the GrIS.”

P. 2 Line 49 – P. 3 Line 1: There are likely also improvements that could be made regardless of resolution, that a high-resolution simulation could not fix. Perhaps mention this also.

To clarify this, we inserted the following sentence: “Important modelling challenges and limitations still need to be addressed in RACMO2 regardless of the spatial resolution used: e.g. cloud representation (Van Tricht et al., 2016), surface albedo and turbulent heat fluxes (Section 6).”

P. 3, Line 56: Could the authors provide a reference for RACMO2.3p1?

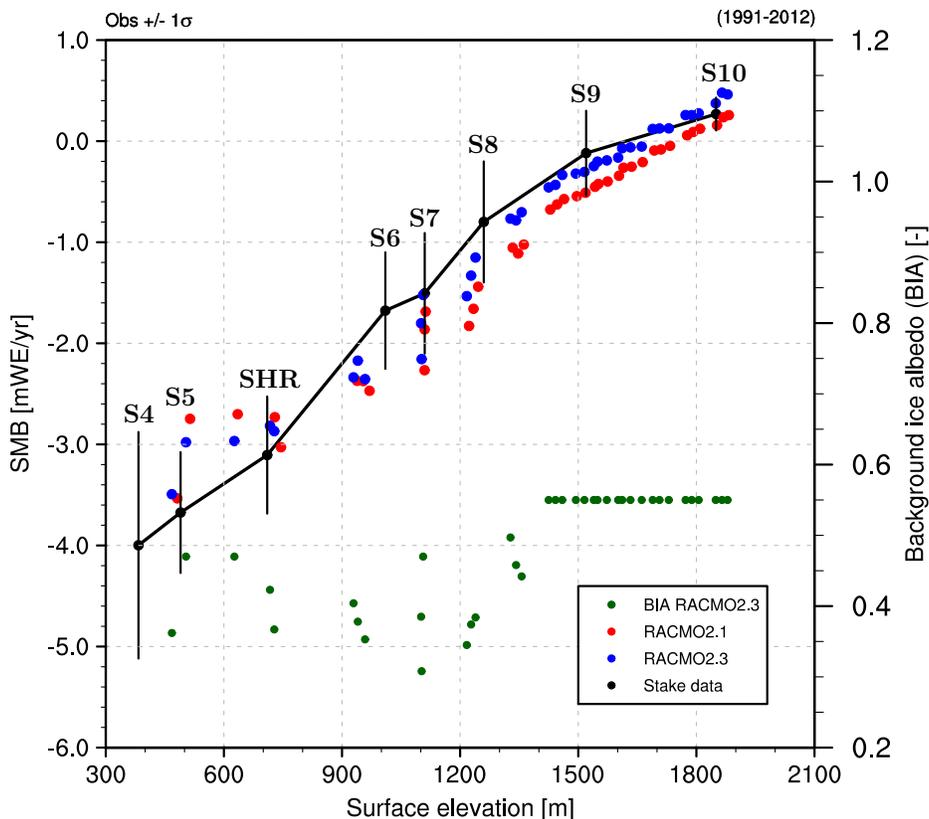
We added a reference to Noël et al. (2015).

P. 4, Lines 82-84: As noted above, the net energy absorbed by the snowpack must be used to raise the surface temperature to the melting point before it can be used for melting. “M” should therefore be changed to “Enet” and this sentence should be revised accordingly.

See our previous response to General comments.

P. 4, Line 87: Add “net” before “sensible and latent turbulent heat fluxes” for clarity. Done.

P. 5, Lines 109-111: The corrections that have been made also can affect the ablation zone, though they probably have less of an impact there. Were similar biases found in the ablation zone previously?
 From Fig. 9 in Noël et al. (2015) (see below), you can clearly see that compared to observations along the K-transect (black dots) SMB in RACMO2.3p1 (blue) is overestimated in the ablation zone, due to too low melt (ablation) rates.



P. 5, Lines 116-118: If possible, can the authors provide evidence that supports decreasing the size of refrozen snow grains?

We could not find a proper reference that validates our assumption. However, Kuipers Munneke et al. (2011) suggest that the grain size of refrozen snow should be ascribed a value larger or equal to 1000 μm . We therefore chose to set the grain size of refrozen snow to 1000 μm : this value minimizes the positive melt bias in the model across the GrIS percolation zone.

P. 5, Lines 136-137: Can the authors be a bit more specific about the levels or height at which upper atmosphere nudging is applied?

See our response to Reviewer #2.

P. 5, Lines 139-141: Provide a few more details about this. What are the “best” profiles and how are they derived?

We clarified this by reformulating as: “The model has about 40 active [...] 1957 using estimates of temperature and density profiles derived from the offline [...] (Ligtenberg et al., 2011). These profiles are obtained by repeatedly running IMAU-FDM over 1960-1979 forced by the outputs of the previous RACMO2.3p1 climate simulation until the firn column reaches an equilibrium.”

P. 6, 149-151: Are fractional areas of ice vs. tundra allowed in a RACMO grid box? If so, it would be useful to have this information here.

To clarify this, we inserted the following sentence: “In RACMO2, a grid-cell with an ice fraction ≥ 0.5 is considered fully ice-covered.”

P. 6, Line 156: Please specify the version number. Is this version 5 or version 6?

We now mention version 5 in the revised manuscript.

P. 6, Lines 157-160: These sentences are a bit unclear. I think the authors are saying that MODIS values for bare ice albedo below 0.3 are replaced by a value of 0.3, and MODIS values above 0.55 are replaced with 0.55. Any grid cells without a valid MODIS estimate are assigned a value of 0.55. Please clarify.

Your interpretation is right, except for non-valid MODIS estimates which are set to 0.47. For clarity, we reformulated as: "In RACMO2, minimum ice albedo is set to 0.30 for [...] , and a maximum value of 0.55 [...].".

P. 7, Lines 201-202: Are these biases statistically significant? It might be useful for the reader to have this information.

We calculated the significance of the biases in Figs. 3-4 by comparing the mean bias listed in each scatter plot with 2 standard deviations of the biases (95% interval). None of the biases were larger than 2 standard deviations, so that these biases are determined as insignificant. As the biases are relatively small and negligible compared to the absolute value of the meteorological data and radiative fluxes, we deem that this is not necessary to mention in the manuscript.

P. 7, Lines 209-210: Can the authors be sure that the LWd underestimation leads to the LWu underestimation?

We reformulated as: "The negative biases in LWd and 2-m temperature partly lead to LWu underestimation of [...].".

P. 8, Line 214: Can the authors elaborate here? Is there a difference because of heterogeneity in fresh snow distribution leading to differences between the model estimate and local measurements?

We elaborated as follows: "In RACMO2, precipitation falls vertically, i.e. no horizontal transport is allowed, and is assumed to be instantly deposited at the surface. Consequently, the spatial distribution of summer snow patches may be locally inaccurate, resulting in large albedo discrepancies when compared to point albedo measurements.".

P. 8, Lines 229-231: It's a bit unclear that the values in parentheses are biases and not absolute magnitude of the quantities. Clarify here and where applicable elsewhere in the text, e.g. "... between overestimated SWn (bias of 16.2 W m⁻²)".

Done.

P. 8, Line 231: Make clear whether SWu is over- or underestimated. I believe it's underestimated. (See general comments.)

Done.

P. 8, Lines 236-237: I believe the newest MCD43A3 product includes a correction for sensor deterioration, but if v5 is used here, this still applies. Perhaps clarify with "underestimated surface albedo for the MCD43A3 v5 product".

We reformulated accordingly.

P. 8, Lines 238-239: Again this is confusing because of sign conventions. If the signs of the biases follow the conventions, the net bias should be -23.9 W m⁻² and not 0.3 W m⁻².

This is now corrected. See also our response to General comments.

P. 9, Line 252: Also add reference to Table 2 here.

Done.

P. 9, Line 253: Clarify under- vs. overestimated, use "~4 W m⁻²" to indicate that the value is approximate.

Done. See also our response to General comments.

P. 9, Lines 269-270: According to most of the previous calculations of net flux, these terms don't compensate. There is underestimated downward flux and overestimated upward flux, so the net flux is underestimated.

Both Swd and SWu are underestimated. Thank you.

P. 9, Line 271: Add reference to Table 4.

Done.

P. 9, Lines 272-273: Again, here the biases have been added rather than subtracted to get the net flux, in contrast with calculations for other sections.

This is now corrected.

P. 10, Line 289: Here the SWu bias, shown as positive in Table 1, is referred to as negative, which would make sense if conventions are followed everywhere, but is not consistent with earlier discussion (e.g. p. 9, line 264, where a positive bias for SWu is referred to as an overestimation).

At line 289, we mention site S6 showing a SWu underestimate of 3.8 W/m², while line 264 refers to site S9 close to the equilibrium line where SWu is overestimated by 3.5 W/m². We clarified this by replacing “Here, [...]” by “At site S6, [...]”.

P. 11, Line 316: The increase in refreezing is attributed to an increase in precipitation, but along the west coast, there is a decrease in precipitation in some areas. Perhaps another factor could be persistence of snow cover as a result of reduced melting.

Thank you. We reformulated as follows: “[...] a result of combined enhanced precipitation and reduced summer melt (delaying the disappearance of the seasonal snow cover), that increased [...] (Fig. 7f).”.

P. 12, Lines 353-354: Show numbers for both model versions for comparison.

The numbers previously shown were incorrect, we revised accordingly based on bias and RMSE listed in Table 6: “[...] shows a decreased bias from 606 mm w.e. in RACMO2.3p1 to 424 mm w.e. in version 2.3p2, and reduced RMSE from -133 mm w.e. to -54 mm w.e., and an increased R2 from 0.92 to 0.97.”

P. 12, Lines 358-362: Is this correction applied to the values in Fig. 9?

This correction is applied to both RACMO2 versions in Figs. 9a and c. It is now mentioned in the figure caption as: “At S10, modelled SMB is estimated as the difference between total precipitation and melt.”.

P. 13, Lines 402-405: Provide some numbers to illustrate that the new version performs as well as the previous version.

We included the following numbers: “[...] as well as the previous version, i.e. bias of 1.20 m w.e. yr-1 and RMSE of 0.47 m w.e. yr-1 (Noël et al., 2016), although [...]”.

P. 13, Lines 419-420: What are the new values for RMSE, bias, and error at QAS_L?

We calculated a model (RACMO2.3p2) RMSE and bias of 2.35 m w.e. and 2.21 m w.e. at QAS_L, respectively. This is now mentioned in the revised manuscript.

Tables 1-5: The term ME is used here for melt energy, but the term M is used in the text. These should be consistent. As noted in the general comments I believe this should really refer to the net energy balance. Captions for Tables 2-5 can be reduced to “Same as Table 1 for Station...”

Please, see our previous response in the General comments. We replaced ME in Tables 2-6 by M.

Figure 11: The red points in (a) indicate something different from (b) and (c). I feel that the authors should include RACMO2.3p2 for (a), (b) and (c), and use the same color scheme. A different color could be used to show the measurements from QAS_L.

We judge that showing RACMO2.3p1 data in Fig. 11b and c would make the scatter plots unclear and confusing. A similar comparison using RACMO2.3p1 data is already conducted in Noël et al. (2016). We now display QAS_L data in orange and we modified the caption accordingly.

Units for statistics should be the same for all figures if possible, and should correspond to the units in the text.

We deem that units should remain mm w.e. in Fig. 11a as bias and RMSE are relatively small. However, we think that m w.e. is more adequate for Figs. 11 b and c due to the larger bias and RMSE (2 orders of magnitude). Therefore, we decided not to change the units in the revised manuscript.

Also, I believe the third line of caption: “version 2.3 (red)” should read “version 2.3p1 (red)”.

Done.

Technical Corrections

P. 1, Line 10: Change “than the previous model version” to “compared with the previous model version”. OK.

P. 2, Line 29: Change “to explicitly resolve” to “of explicitly resolving”. OK.

P. 2, Line 33: Change “evaluate and improve” to “evaluating and improving”. OK.

P. 2, Line 45:”. If the authors are still referring to previous versions of the model, change to “is underestimated” to “was underestimated”.

We kept “is underestimated” as this is still valid in version 2.3p2.

P. 2, Line 50: Change “near-kilometre” to “near-kilometre-scale” We replaced by “near-kilometre scale”.

P. 3, Line 55: Change “all over Greenland.” to “across the GrIS.” OK.

P. 5, Line 132: Add “an” before “11 km horizontal...” OK.

P. 5, Lines 134-135: Mention that the model domain is shown in Fig. 1 to make clear what Fig. 1 is showing. We reformulated as: “[...] 6-hourly basis over the model domain shown in Fig. 1.”

P. 7, Line 191: Change “of 23 AWS” to “from 23 AWS” OK.

P. 7, Line 192: Change “output” to “outputs”. OK.

P. 7, Line 194: Add “and” after “10-m wind speed,” OK.

P. 7, Line 201: change “with a small negative bias” to “with the model exhibiting a small negative bias” for clarity. OK.

P. 8, Line 231 and Line 239: Change “too low cloud cover” to “underestimated cloud cover”. OK.

P. 8, Line 241: The van den Broeke (2008) reference seems to be missing from the reference list. This reference is actually Smeets and van den Broeke (2008). This has been corrected.

P. 9, Line 258: Change “too large SHF” to “SHF to be overestimated”. OK.

P. 10, Lines 295 – 298: The language could be improved here. Suggested revision: “In Section 3, we discussed the overall good ability of RACMO2.3p2 to reproduce the contemporary climate of the GrIS, which is essential for estimating realistic SMB patterns. Here we compare SMB from RACMO2.3p2 and RACMO2.3p1 over the GrIS. For further evaluation, we focus on three regions where there are large differences in SMB between the two versions.” Thank you, we revised accordingly.

P. 11, Line 318: Change “very GrIS margins” to “extreme margins of the GrIS”. OK.

P. 12, Lines 360-362: This sentence is a bit wordy... suggest changing “decreasing the bias...” to “decreasing the bias by 260 mm w. e. yr-1 to -40 mm w. e. yr-1 and the RMSE by 200 mm w. e. yr-1 to 210 mm w. e. yr-1.” OK, thank you.

P. 12, Line 378: Change “3 months” to “3 month”. OK.

P. 14, Line 426: Add “an” before “11 km resolution”. OK.

P. 14, Line 434: change “narrow ablation zones” to “the narrow ablation zone”. OK.

P. 14, Line 444: Change “to capture” to “to capturing”. OK.

P. 15, Line 455: Change “cryoconites” to “cryoconite”. OK.

P. 15, Line 467: Change “proves to accurately capture” to “accurately captures”. OK.

Figure 2: Although it is not necessary since the caption provides a description, a legend on Fig. 2a would be useful for the reader. Done.

Figure 5: The dashed lines on the legend are hard to distinguish from solid lines. Although relatively small, we judge that Fig. 5 is readable as is. We did not modify this figure in the revised manuscript.

Figure 9: The black and blue colors are a bit hard to distinguish. Can the blue color be made slightly brighter? In caption, remove “the” in “for the S5”. We decided to display observed data in light gray, and kept dark blue for RACMO2.3p2 data to remain consistent with other figures. Thank you.

Figure 10: The yellow line is difficult to see. The color could be made slightly darker. Add “and” after “a) daily snow albedo”. As in Fig. 9, observational data are now displayed in light gray.

Fix references to follow format for The Cryosphere. We hope that the references are now suitable.

Reviewer #2:

This is a very nice well-written and detailed paper describing the revised RACMO2.3p2 model and the effect of improved model tuning on the Greenland ice sheet surface mass balance estimates. This kind of paper is extremely important for users of SMB data to read and digest in order to understand the likely biases and uncertainties within model output and the thorough analysis, while not really presenting much novel scientific research is an important addition to the scientific canon. It is very well structured and easy to read and the authors are to be congratulated on a thorough overview. That said I have some issues, which I feel should be addressed before final publication.

Points for consideration

In section 2.3 Model Updates, it is noted that there have been some changes to the cloud scheme but these are not discussed in much detail and it is not clear how the large adjustments to the l_{crit} in mixed phase and ice clouds were reached. This is a pretty serious adjustment of the model as many of the other model parametrizations in the radiation scheme are likely to be tuned to these kind of values, possibly giving erroneous results or different compensating biases. However little detail is given as to how or why the particular values for these adjustments were chosen. Nor are the effects of this adjustment alone described in any detail – for instance on lines 398 to 400 reference is made to a precipitation bias in the SE compared to observations, but it is not clear if this bias is reduced or increased from RACMO2.3p1 and if this is a result of the cloud scheme changes or for example the change in topography caused by moving to the GIMP DEM. More information on how this change in cloud parametrizations has altered precipitation in particular would be helpful, as the differences in the topography shown in figure 2 seem also to be related to the change in distribution of precipitation, at least in some locations as shown in figure 7.

We agree that the cloud scheme updates, i.e. tuning of the critical cloud content (l_{crit}), is not described in sufficient detail. The goal of this tuning was to obtain a uniform increase in precipitation over Greenland in order to minimize the inland dry bias observed in RACMO2.3p1.

To that end, we carried out a series of sensitivity experiments, i.e. one-year simulations, to test changes in spatial distribution as well as in the time scale of precipitation formation. From these experiments, we found a linear relationship between the critical cloud content (l_{crit}) for mixed and ice clouds, the vertical integrated cloud content (liquid and ice water paths that also affect the SEB through changes in cloud optical thickness), and the integrated precipitation over Greenland. Increasing l_{crit} for mixed (2x) and ice clouds (5x) uniformly enhances precipitation, except for some high accumulation regions in southeast Greenland (see Section 5.1 and Fig. 11a).

These new settings were then tested for a longer period. This approximately cancelled the dry bias observed in RACMO2.3p1. This also led to larger but realistic vertical integrated cloud content and did not strongly affect the SEB and surface climate of the GrIS. Note that l_{crit} is also strongly adjusted in the ECMWF physics compared to commonly used values in the literature. For example, for snow clouds modelled by a cloud parcel model, Lin et al. (1983) set l_{crit} to $1 \cdot 10^{-3} \text{ kg kg}^{-1}$, while the ECMWF physics uses a value of $0.3 \cdot 10^{-4} \text{ kg kg}^{-1}$. The justification for this is provided at p. 90 of ECMWF-IFS (2008): “A lower value (of l_{crit}) is appropriate for a GCM sized grid box (unless subgrid cloud variability is explicitly taken into account)”. As RACMO2 employs much smaller grid boxes than GCMs, higher values of l_{crit} , typically $\sim 10^{-4} \text{ kg kg}^{-1}$, are not unreasonable.

In brief, changes in precipitation shown in Fig. 7a are mostly driven by the tuning of l_{crit} . Locally, large changes in mountainous areas, i.e. notably in southeast Greenland, are due to changes in the topography. Likewise, tuned l_{crit} is the main driver of changes in SWd and LWd through changes in cloud optical thickness, which of course slightly impact the other SEB components. However, these are even more affected by other changes applied in the new model version, i.e. changes in ice and snow albedos. To clarify this, we added the following sentences: “The values of l_{crit} adopted in RACMO2 were obtained

after conducting a series of sensitivity experiments, i.e. one-year simulations, to test the dependence of precipitation formation efficiency, spatial distribution and cloud moisture content on l_{crit} and other cloud tuning parameters. From these experiments, we found a linear relationship between l_{crit} for mixed and ice clouds, the vertical integrated cloud content, i.e. liquid and ice water paths that also affect the SEB through changes in cloud optical thickness, and the integrated precipitation over Greenland. These new settings were then tested for a longer period and proved to almost cancel the dry bias observed in RACMO2.3p1 (see Section 5.1). This led to larger but realistic vertical integrated cloud content and did not strongly affect the SEB and surface climate of the GrIS. For instance, the induced changes of surface downward shortwave and longwave radiation are only about -4 W m^{-2} and 7 W m^{-2} , respectively, peaking in central Greenland. While the obtained increase in l_{crit} is relatively large, especially for ice clouds, it is important to note that it is also strongly adjusted in the original ECMWF physics compared to commonly used values in the literature: e.g. Lin et al. (1983) set l_{crit} to $1 \cdot 10^{-3} \text{ kg kg}^{-1}$ for ice clouds, while the ECMWF physics, tuned for GCM sized grid cells, uses a value of $0.3 \cdot 10^{-4} \text{ kg kg}^{-1}$ (ECMWF-IFS, 2008). As l_{crit} depends on model grid resolution, i.e. GCMs running at lower spatial resolution require lower value of l_{crit} (ECMWF-IFS, 2008), the use of a larger l_{crit} for e.g. ice clouds ($1.5 \cdot 10^{-4} \text{ kg kg}^{-1}$) in RACMO2 is deemed reasonable. In addition, this value remains well within the range of values previously presented in the literature (Lin et al., 1983).”

On a similar theme, I note that the small improvement in LWd and SWd on the K-transect is reasonably attributed to the change in the cloud scheme. It would be very interesting to see if this improvement is consistent across Greenland at stations other than the K-transect. There is some reason to believe that western Greenland is often modelled well but in other regions RCMs do a less good job of reproducing observed climate variables, possibly due to biases in cloud schemes as well as the complex topography in other areas. As there is now a fairly large amount of data available from Promice stations it would be nice to see some geographical spread in the figures presented in figure 5 and tables 1-5, perhaps limited to maybe 3-4 extra stations in north, south and east Greenland to determine if the positive results from western Greenland are replicated elsewhere.

To provide a regional evaluation of meteorological and radiation data, we decided to add individual scatterplots for four different sectors of the ice sheet (NE [Figs. S3-4], NW [Figs. S5-6], SE [Figs. S7-8], SW [Figs. S9-10]). These scatterplots are similar to Figs. 3 and 4. They list statistics for the four regions, respectively, and clearly stress that RACMO2.3p2 agrees as well with observations in these four sectors. See also our response to the General comments 2) of reviewer #1. We also added the following sentences in the revised manuscript; in Section 3.1: “To provide some regional insight on the model performance, Table S1 and Figs. S1-S4 compare modelled meteorological data from RACMO2.3p2 with AWS measurements (green dots in Fig. 1) clustered in four sectors of the GrIS, i.e. NW, NE, SW and SE, respectively. These sectors correspond to the four quadrants delimited by longitude 40°W and latitude 70°N , respectively. These regional scatter plots unambiguously show that RACMO2.3p2 performs as good in each of these four sectors of the GrIS.”.

In Section 3.2: “Clustering AWS measurements within four sectors of the GrIS (Figs. S6-S9 and Table S1), RACMO2.3p2 shows good and equivalent agreement in NW, NE, SW and SE Greenland, respectively.”.

Upper atmosphere relaxation is mentioned on line 136-7 but no details are given. I would like to see this expanded with details on which fields are nudged and at which levels in the atmosphere as this is important for interpreting the atmospheric model output.

We elaborated as follows: “To better capture SMB interannual variability in this new model version, upper atmosphere relaxation (UAR or nudging) of temperature and wind fields is applied every 6 hours for model atmospheric levels above 600 hPa, i.e. $\sim 4 \text{ km a.s.l.}$ (Van de Berg and Medley [2016]). UAR is not applied to atmospheric humidity fields in order not to alter clouds and precipitation formation in RACMO2. ”.

The authors acknowledge that boundary forcing is important for results (line 20) but the differences between results from ERA40 and ERA-Interim forced years are not explored at all. It would be

helpful to have a time series of SMB and the components for the full 1958-2015 averaged over the full ice sheet for the full period. This would show if, for example, there is a marked change in precipitation or melt potentially resulting from the switch in boundary forcing in 1979 is visible in more detail.

Time series of RACMO2.3p1 have been already published in e.g. Van den Broeke et al. (2016) [11 km] and Noël et al. (2017) [1 km] and show no abrupt changes (larger than the interannual variability) between 1978 and 1979, i.e. when the re-analysis forcing switches from ERA-40 to ERA-Interim. The same also applies to RACMO2.3p2 simulation.

It would also give a better sense of the interannual and decadal scale variability in SMB of Greenland. Plotting these with model version p1 masked with the same ice mask would also allow us to assess the differences in SMB over the full ice sheet that result from the improvements introduced here.

As discussed in Section 5.1, trends and time series cannot be directly derived from the new RACMO2.3p2 version as “correct” precipitation and underestimated runoff, i.e. due to unresolved high melt rates over low-lying marginal outlet glaciers and narrow ablation zone, lead to overestimated GrIS SMB at 11 km resolution, highlighting the need for further statistical downscaling. In addition, relevant climatological averages (1958-2016) for the main SMB components are already listed in Sections 5.1 [11 km] and 5.2 [1 km]. At the discretion of the editor, we are happy to include time series of annual downscaled SMB components from RACMO2.3p2 at 1 km.

I am not quite clear if the improvements to the snow module are part of the online RACMO model or the offline firn model – I assume the former, but please clarify this in sections 2.1 and 2.3.

RACMO2 (climate model including a snow/firn module) and FDM-IMAU (firn model) are two different models. As mentioned, IMAU-FDM is run offline and is forced by the climate data of RACMO2. In Section 2.1 and 2.3, we only refer to RACMO2 settings and updates and do not discuss FDM-IMAU. FDM-IMAU is only mentioned in Section 2.1 to clarify how a snowpack initialization has been obtained for the 1st of September 1957, date at which the RACMO2.3p2 simulation starts. FDM-IMAU simulations forced by RACMO2.3p2 will be discussed in a forthcoming paper. See also our response to reviewer #1.

“The model has about 40 active [...] 1957 using estimates of temperature and density profiles derived by the offline [...] (Ligtenberg et al., 2011). These profiles are obtained by repeatedly running IMAU-FDM over 1960-1979 forced by the outputs of the previous RACMO2.3p1 climate simulation until the firn column reaches an equilibrium.”

On line 216 you note that AWS data is sometimes spurious, Ryan et al 2017 in GRL also showed that the siting of stations (for very good reasons!) also leads to spurious underestimation of albedo – this should probably be referenced. [This reference has been included in the revised manuscript.](#) Thank you.