Response to RC2. The authors provide responses to the reviewer comments in italic serif typeface throughout whereas the reviewers comments remain in normal sans serif typeface.

GENERAL

First of all, I would like to apologize for the time it has taken for me to provide this review. This was an oversight on my part and in no way reflects a lack of interest in this paper or the subject.

This manuscript by Medley et al. discusses a new modeling framework that uses a combination of the MERRA-2 reanalysis and the Community Firn Model to arrive at estimates of SMB and its individual components over the Greenland and Antarctic Ice Sheets for the period 1980-2019.

It is an important and welcome contribution to the existing regional climate models and firn models available for estimating ice-sheet mass balance by means of altimetry or the input-output method. It is encouraging that more of these models become available, and MERRA-2 combined with CFM is presented as a comprehensive modeling framework.

There are three main concerns with this manuscript, which I hope can be addressed by the authors in a thorough revision. But I am sorry to say that this probably entails a substantial effort. In fact, it would have been better from the start to divide this work into a paper that evaluates MERRA-2 over ice sheets, and a paper that deals with the CFM for elevation change and runoff estimates.

We thank the reviewer for taking the time to present a thorough review of our submitted manuscript. We appreciate that the reviewer sees value in our work, and we have worked to satisfy all the reviewer's concerns to improve the quality of the research and manuscript. We think the reviewer's insight has largely improved the manuscript, and we thank them accordingly.

1) Evaluation of MERRA-2 against independent observations.

In a global sense, MERRA-2 has been evaluated in a large paper by Gelaro et al. (2017). However to my knowledge, the performance of MERRA-2 over Greenland and Antarctica has not been assessed against observations, like it has been done extensively for e.g. MAR (Fettweis et al., 2013, Agosta et al., 2019), HIRHAM5 (Mottram et al., 2017), RACMO (Van Wessem et al., 2018, Noël et al., 2018) or CESM for climate variables ánd individual SMB components (temperature, snowfall, surface melt, albedo, cloud cover, firn temperature, runoff, etc.). This paper does present SMB and its components but they are not verified, evaluated, or compared against existing estimates. This is worrisome, since MERRA-2 serves as the input for CFM which computes the surface elevation changes needed to correct for ICESat and other altimeters. Thus, the authors will have to make sure that the reader can be confident about the MERRA-2 output over the ice sheets.

The reviewer is entirely correct that we must ensure the quality of the forcing data to understand its strengths and limitations in representing polar climate conditions. As the reviewer points out, a full analysis of multiple variables over both ice sheets is an extensive undertaking; thus, we present SMB analysis of MERRA-2, but also state that MERRA-2 has been extensively evaluated over the ice sheets and have added additional citations which should provide the reader more confidence in the use of MERRA-2. Input for the CFM is net accumulation, melt, rain, and temperature. We provide an evaluation of SMB as best we can replicate recent SMB model intercomparison exercises for each ice sheet presented in this journal (i.e., Mottram et al., 2021 and Fettweis et al., 2020). We provide citations *regarding temperature evaluations from MERRA-2 that show its capability in providing realistic air temperatures for both ice sheets.*

2) Uncertainty analysis

The paper lacks uncertainty analyses. Uncertainty intervals of the SMB components are a standard deviation of the sample of annual values. But there is no estimate of the uncertainty of annual values, established for example by independent evaluation against observations (see 1). For the volume and height changes, no uncertainty estimates are provided. Most effort in presenting numbers is not in the numbers themselves, but in quantifying the confidence that we put in them. This is also crucial for establishing an uncertainty estimate of mass loss determined by altimetry. A robust uncertainty analysis must be added to the paper.

Quite a number of processes could be identified that give an uncertainty. Regarding the surface elevations changes from CFM, probably the most important ones are the uncertainty in SMB, snowfall, surface melt, and the assumption that the RCI period is representative for the long-term climate (see comments by reviewer 1).

Adding an uncertainty estimate allows for a fair judgement of the numbers presented in this paper, and for an comparison with existing literature (see also point 3), like for example the volume change estimates from Kuipers Munneke et al., which at first sight appear to be larger over Greenland than from MERRA-2.

The reviewer brings up an excellent point regarding the importance of also providing uncertainties. We have added a section detailing our uncertainty analysis, which includes insights into the comparisons with observations for SMB (and how it translates into the model uncertainty). We have also generated perturbation ensemble CFM runs in which both climate variables and firn model parameters as perturbed at several locations across the ice sheet to evaluate the impact on FAC and thickness change through time. These uncertainties provide the reader realistic bounds on our knowledge of the firn evolution stemming from uncertainties in both the forcing, firn model parameters, as well as choice of RCI, which as the reviewer rightly points out helps with intercomparison of two models.

3) Embedding in existing literature

While I believe that MERRA-2/CFM is a very valuable addition to the suite of models that are currently available to estimate SMB and firn changes over the ice sheets, the literature from those existing models is largely ignored in this paper. It would be insightful if the numbers in this study are compared to existing estimates of SMB and firn volume change (MAR, RACMO, CESM, HIRHAM, ...). Also it would be fair to refer to previous work into, e.g., quantifying seasonal cycles of firn processes (Ligtenberg et al., 2012), or volume loss over Greenland.

We appreciate the importance of model intercomparison, which is why we compared our results to those from two recent SMB model intercomparison exercises to show that our SMB estimates fall within the range of several different models (rather than comparing to individual models; Lines 490-499). While an intercomparison with other models is an interesting analysis, it falls outside the scope of this paper, which is to present the new results from MERRA-2 and CFM. Future work will involve a comparison of volume change products.

Not referencing the work of Ligtenberg et al., 2012 was an oversight on our end, and we have now included it in our paper. If the reviewer could be more specific regarding omitted literature on "volume loss over Greenland," we would greatly appreciate it. Because this paper is not generating model simulations of SMB, but rather using it from an existing global model, we didn't extensively compare to existing papers that describe various SMB model updates and improvements. The focus of the work is rather on the CFM modelling, which is new. We do, however, add references to some of the recent SMB model work as suggested by the reviewer to point readers to additional resources if they are interested. We believe that the additions will better contextualize the existing work, and how this effort builds on an existing strong research community.

SPECIFIC COMMENTS

Page 1 line 10: suggest to include the time frame of the simulations in the abstract: "new simulations of firn processes (1980-2019)"

Done.

Page 5 line 126: how is temperature initialized for locations with SMB<0?

The initial ice column is isothermal, set at the mean annual temperature. We have clarified the sentence.

Page 7 line 191: why was the functional form of equations (11) and (12) chosen to be an Arrhenius-style rate equation? Is there a reason to suppose that the misfit between modeled and observed densification rates should take this form?

This was based off the functional form of the densification equations 9-10. Densification follows an Arrhenius type dependence on temperature whereas the overburden or accumulation rate is assumed to be linearly dependent (exponent equal to one). We took on the functional form from the source equation rather than adding our own expectation of the functional form because it makes interpretation of the impact on the densification parameters more straightforward rather than adding a new form to the calibration on top. We have discussed this choice and its implications more without the text. We explored a linear form on the calibration, but the model fit was worse than using the functional form of the densification equations.

Page 7 line 211: it is interesting to see that E_0 is found to be zero. It implies that the correction factor R_0 reduces to b^{beta_0}, i.e. with no dependence on temperature. For the reader, it would be interesting to show this result rather than the current figure 4. The present figure 4 is of limited interest, since the correction procedures forces the data points to the 1:1 line by design. It would be more interesting to show the validity and behavior as an Arrhenius plot:

R_0 = b^(beta_0) exp(-E0/RT)

ln(R_0) = beta_0 ln(b) - E_0 (1/RT)

With E_0 = 0, plotting ln(R_0) against ln(b) would show beta_0 as the gradient and 0 as the intercept of a linear plot.

For R_1 it is more difficult. A normal Arrhenius plot would look like

ln(R_1) = (-E_1/R) * (1/T) + beta_1 * ln(b) which is a linear y = mx + d with 1/T as x and beta_1*ln(b) as a constant. However, the latter is not constant but depends on b. Perhaps it is possible to select a subset of ∼20 firn cores with a narrow range in b so that a constant value for (beta_1*ln(b) can be assumed, and ln(R_1) can be plotted against 1/T as a linear function.

Thank you for the suggestion to expand more on how the calibration procedure impacts the original densification rates. We plot R_0 and R_1 as a contour plot with temperatures on the x-axis and accumulation rates on the y-axis, and colors represent the calibration coefficients. The reviewer is indeed correct that there for the first stage of densification, E_0 = 0 implies the calibration does not have any dependence on the temperature. For Stage two, there is temperature dependence. We also note a recent paper that independently came to near similar results using a different technique (Verjans et al., 2020), which we reference now as well.

Page 9 line 270: I can understand that you have looked only at the significantly related predictors, and not at their physical interpretation. But I think that mean northward wind speed is an awkward predictor especially in the Greenland context. Could the performance of the model be retained while dropping V_0?

Yes. Use of northward wind only was an oversight (and something that the other reviewer noted as well), so we no longer use northward wind in our surface density calculation.

Page 10 line 297: I am struggling with the fact that no meltwater-related output from MERRA-2 was used in this manuscript, only snowfall. Is MERRA-2 not designed to provide this output? Is it of insufficient quality? Was it forgotten? Regarding runoff, was the runoff output from MERRA-2 used, or from the firn model? This remains unclear. In any case, is the runoff from the firn model comparable and similar to the runoff from MERRA-2?

MERRA-2 simulates liquid water processes, including melt flux, infiltration, refreezing, and runoff; however, actual melt flux variables were not saved within a data collection. Only runoff is available. We use the runoff provided by the firn model because the snowpack physics used over the ice sheets was derived for mountain snowpack environments, not polar ice sheet conditions and have made that clear in Section 2.2, and we also make it clear when we discuss how SMB is calculated in Section 3.2 (i.e., we specifically mention that runoff from the CFM simulations is used). We provide a comparison of runoff derived from the CFM with that from MERRA-2 as well.

Page 10 line 300: The degree-day modeling approach is confusing, and quite rough. The approach by Van den Broeke et al. (2010) was devised only in cases where daily mean temperatures temperatures are available, to compensate for the lack of representation of the daily cycle. I have tried to understand why such low T_0 are required to get a good match at an annual scale. I guess this is only possible if MERRA-2 has a strong cold bias, or a very poor representation of the daily temperature cycle, or if the degree-day approach actually fails over the entire range of melt-temperature combinations on both ice sheets. I would like the authors to comment on this. More generally, I wonder if surface melt in MERRA-2 is poorly represented, or if it was forgotten or otherwise impossible to obtain this output variable.

We have reworded much of the section on degree-day modeling to help make clear how the actual melt rates were calculated. There is unfortunately no way to recover surface melt from MERRA-2, which is why we use the degree-day approach. We are completely transparent that this is a major weakness of the work, especially when considering the Greenland Ice Sheet. We instead use the degree-day approach as a best approximation. We agree that we should have more discussion into the implications for such an *approach, as well as the derivation, and we also compare plots of the derived melt with the training datasets. It's important to note that the reasons could be many in number why the numbers needed are less than 273K, including either MERRA-2 model deficiencies (as the reviewer points out) but also MAR model deficiencies as well. We have added context and refined the Section 2.2.1 on degree-day modeling, including reiteration of the model deficiencies and its role on firn volume change.*

Page 13 line 382: a map showing the absolute change in FAC would be interesting here.

We have added a map of FAC change.

Page 14, section 3.2: To my opinion, it is important to emphasize that the surface melt estimates from Greenland and Antarctica are not independent numbers based on the outcomes of a physical model. Rather, these are the results of a degree-day model calibrated to observations (Antarctica) and another model (MAR, Greenland). Especially for Antarctica, it is presented (lines 427 and further) as if the melt estimates are an independent result of MERRA-2 and CFM, whereas in reality the degree-day method is tuned to reproduce the numbers by Trusel et al. over a part of the RCI. Page 15,

We appreciate the concern of the reviewer. We have made note that only snowfall, evaporation, and rainfall come from MERRA-2, melt is forced to match output from other work, and runoff is derived via the CFM using that calibrated melt model.

line 444: RFI -> RCI (?)

Yes. We have corrected this typo.

===

Agosta et al., 2019. Estimation of the Antarctic surface mass balance using the climate model MAR (1979-2015) and identification of dominant processes, The Cryosphere

Fettweis et al., 2013. Estimating the Greenland ice sheet surface mass balance contribution to future sea level rise using the regional atmospheric climate model MAR, The Cryosphere

Gelaro et al., 2017. The modern-era retrospective analysis for research and applications , version 2 (MERRA-2), Journal of Climate

Kuipers Munneke et al., 2015. Elevation change of the Greenland Ice Sheet due to surface mass balance and firn processes, 1960–2014, The Cryosphere

Ligtenberg et al., 2012. Quantifying the breathing of the Antarctic Ice Sheet, GRL

Mottram et al., 2017. Surface mass balance of the Greenland icesheet in the regional climate model HIRHAM5 : Present state and future prospects. Low Temperature Science. Series A. Physical Science.

Noël et al., 2018. Modelling the climate and surface mass balance of polar ice sheets using RACMO2 - Part 1 : Greenland (1958-2016), The Cryosphere

Trusel et al., 2013. Satellite-based estimates of Antarctic surface meltwater fluxes, GRL

Van den Broeke et al., 2010. Temperature thresholds for degree-day modelling of Greenland ice sheet melt rates, GRL

Van Wessem et al., 2018. Modelling the climate and surface mass balance of polar ice sheets using RACMO2, part 2: Antarctica (1979-2016), The Cryosphere