Review of
Improved representation of the contemporary Greenland ice sheet firn layer by IMAU-FDM v1.2G
by Brils et al.

Reviewer: C. Max Stevens

This paper details updates to the commonly used IMAU firn densification model. It demonstrates the new model's improvements by comparing model outputs from the new and old IMAU models using several metrics. The paper is clearly written and well organized. The science and ideas presented are well founded. The paper will be a good contribution to *Geophysical Model Developments*, and I am happy to recommend its publication after the authors address several questions I have.

General Comments:

1. My impression after reading the paper was that the authors focus almost entirely on the improvements made in the new model and do not discuss uncertainties and limitations of the model. Given that firn models are often cited as the (or one of the) largest contributors of uncertainty to altimetry-derived mass balance estimates and the fact that numerous research groups use IMAU-FDM, I think it would be useful to include a section discussing the uncertainties in the model's outputs. Additionally, beyond just providing uncertainty bounds (e.g. +/-X%) I think it would be useful in that section to include discussion of why these uncertainties persist – is it due to missing representations of physical processes in the model, the way the model is calibrated, a propagation of uncertainties in the forcing data, or something else?

2. I was surprised to see the firn aquifer site chosen as a test site because the authors state explicitly (line 209) that their model does not simulate standing water on ice layers, which is what is occurring in the firn aquifer zone. Why should one expect the model to perform well at a site where it is not configured to simulate the observed conditions?

3. I am curious about why the authors only use cores from dry sites (line 153) to derive the densification model – this seems fine, but then they use the equation derived for dry sites to simulate firn compaction in the wet-firn zone. (Or, line 153 states, "MO corrects for the dry compaction rate" – is there an additional factor added to correct for wet compaction rate?) Why is using a dry-firn equation for wet firn a valid thing to do? Is there good reason to believe that an equation developed for dry-firn compaction should simulate wet-firn compaction equally well? Why not use cores from wet-firn sites to make a more general MO?

4. I would like more detail about the latent heat source term L in equation 5. How is that implemented in practice? What is the numerical scheme you are using? How do you determine what L is? E.g., are you using an enthalpy solving method, solving the heat equation for temperature and then making a correction for layers where there is liquid water with temperature below freezing?

5. I would like more discussion of why you chose surface density to be a function of the previous year's temperature. Wouldn't you expect surface density to change on short time scales based on local conditions (e.g. warm snow event, cold snow event, strong wind saltation)? You are running your model at very high temporal resolution, which to me means that you think your equations adequately describe physical processes well enough to predict how the density evolves at sub-daily timescales. But, it seems that you are saying that you can only predict your boundary condition at annual timescales – why should I believe that the 3-hour resolution density profiles (or elevation change predictions) are correct when the surface boundary is not calculated with that resolution? Please provide some discussion on this apparent paradox. Is the reality that our knowledge of what determines surface density is at present deficient, and so the best course of action is to not introduce additional model uncertainty by using an equation that does not work well?

6. "old" vs. "new" settings: several times in reading the paper I got confused reading about old vs new. Your new model is v1.2G – what was the old one called – v1.1G? I suggest referring to them throughout the paper by version number rather than "old" and "new", and reference your equations when needed. An example of where confusion arises: on line 122, are you saying that "old" and "new" only refer to which density parameterization you are using, or does it refer to the updated densification equation also?

7. Equation 3: Does $\dot{b}$ evolve through time based on changing accumulation rates through time, or is it constant for a given site based on the spin up climate? If the latter, why? Imagine a site that has warmed since 1980; why should the densification rate of snow that fell in e.g. 2010 be dependent on the accumulation rate from decades earlier when the climate was different? Arthern et al. (2010) formulated their model equation using a steady-accumulation assumption (see their Appendix B); is it appropriate to use your densification equation with variable accumulation, either climatic variability or changing climate? Why not just formulate your model in terms of the stress?

Line by line comments:

Introduction – your introduction is very much about melt, but the rest of your paper is more about the firn densification process and your new equations to simulate that (and your meltwater scheme is unchanged from previous versions). Additionally, your pilot application (section 4) is about altimetry and not melt processes. I suggest editing the introduction to talk a bit more broadly about firn and specifically include a bit more information about altimetry. As it is currently, I read the first paragraph and thought the paper was going to be about the firn's decreased ability to retain meltwater and threshold behavior associated with that; it isn't until the 4th paragraph that you get to topics that are more specifically addressed in your paper.

Line 15: firn doesn't "represent" the transitional stage, it **is** the transitional stage

L22: 'collapses' – choose a different word to differentiate between collapse meaning is damaged – perhaps something like "until at some point the pore space is insufficient to accommodate melt and the system is fundamentally changed"

L24-25: much more extensive than what?

L29: do you mean reduces *by* a factor of 1-4 (also, why the large range on that factor?)

L30: just say “at least 44%”; “no less than” can confuse the reader

L56: describes → describe

L58: resulted → result

L69: I believe that SUMup specifies that each core used should be cited with its original publication to give full credit to those to worked to generate those data, rather than broadly just citing the SUMup papers. I have seen this done by including a supplement listing the cores used and the appropriate citation for each.

L90: remove sentence here about subsurface radiation – you say that again later, which is a more appropriate place for that information

Figure 1 – caption says purple circles, but they are green. Text later in paper (L325) says green. I suggest switching to purple to differentiate between sites and temperature measurements.

L108/Eq. 1: Fausto et al. (2018) concluded that using a value of 315 kg m^{-3} is better than using a parameterization equation. You do mention this later in your paper, but I think it would be better to move some of that text to this point in the paper and reference the results/sensitivity tests in Section 3.1.

Table 2/Figure 3/Line 155 – You state the R^2 value is 3e-3 – perhaps I am misinterpreting what you are doing, but that number to me indicates a very poor fit. Perhaps expand in your text what you are regressing? You should also specify in Table 2 that σ is the variable you are using to notate standard error. And, what is the standard error you are calculating – does this imply you have some normal distribution you are looking at? I am not sure what that statistic is telling me or how you are calculating it.

L159: use caution with word “below” – do you mean deeper than that density horizon in the firn, or densities less than 550? (likewise “above”).

L168: I am confused – you say first on 168 that L includes subsurface absorption of radiation, but then on the next line you say you ignore it.

L202 – citation formatting error

L209: You state that the latter assumption is thought to be valid; what about the first (standing water)?

219: missing a v for snowdrift – should be v_{snd} to be consistent with others

L245: Can you explain the 3-minute time stepping in more detail? I am not sure what you mean here. Do you run the model at 3-minute resolution and then just save the results every 3 hours?

256: “here the FAC is calculated” Do you mean that Equation 9 calculates the FAC over the entirety of the firn column?

258: FAC<15 melt – is this an observation from you model results? Say so.

L261-265: I am confused – the numbers you write in this paragraph are different than those shown in Figure 5.

Figure 5 and L264: define what ‘Bias’ means specifically, i.e. how are you calculating it, and what is it actually a measure of?

L261-265: You should be more consistent delineating ‘FAC in dry’ vs ‘low FAC’ either use wet/dry or high/low, and state your thresholds

263: why did it switch from underestimation to overestimation?

L267: here is another example of confusion with old/new – in this case you are just referring to the old/new densification equations?

Eq. 10: Isn’t this the RMSE? Why are you referring to it as cost function and calling it Phi, and then pivoting in line 271 and calling it RMSE (and then again using Phi at line 273)?

L271-276: Is the improvement of v1.2G over the “old” model, as demonstrated in Figures 5 and 6, due to the new densification equation, or due to the improvement of the surface density parameterization?

L293: I would expect that a cold bias would be more due to incorrect handling of meltwater (e.g. not enough refreezing, so not warm enough) rather than due to a cold bias in the RACMO forcing, especially given that cold sites have a warm bias. Can you comment on this? Regarding the warm bias, does this mean RACMO is biased warm in the colder areas, despite you saying that RACMO is biased cold in the previous sentence?

L295/Figure 8: Indeed, the new model does slightly better than the old, but it appears to be an incremental improvement, and the bigger issue is that there is a still a substantial misfit compared to the observations (more than 5 degrees at 2 m), especially at Summit in the summer. Please provide discussion on this misfit, and the implications it has on your other results (i.e., a temperature difference of 5 degrees will substantially alter the densification rate predicted by your densification equations).

301: Why do you not include shallower observations of temperature, which are available for DYE-2? Wouldn’t this provide an additional metric to test how well the model is capturing the latent heat release due to refreezing?

301: How much of the difference in modeled temperature is due to the (a) new conductivity parameterization, (b) the new surface density, and (c) the new densification equation?

319: Figure 9 shows that your model is predicting penetration depth that is too shallow by a factor of 3 – but your text focuses on the improvement of the new model over the old. Please add text describing (or hypothesizing) the remaining deficiencies in the model that cause it to fail to predict the penetration depth accurately.

335: Did you fit trendline to the 1970 to 2000 time series to find the m/yr increase, or are you just differencing the 2000 and 1970 values, and then dividing by the number of years?

338: "Nevertheless, the individual velocity components being very different" – sentence structure issue – "are very different"?

Figure 10: Your model predicts that the elevation has lowered in the last decade, but surface elevation measurements from Summit (Hawley et al., 2020; sorry to self-reference but it is the dataset I am most familiar with) show that the elevation at Summit increased at 0.019 m/year from 2008 to 2018. Can you explain your model's inconsistency with the observations?

Figure 10: I am curious about the sudden decrease in elevation in 2019-2020, which appears to be related to the low accumulation in 2019. Is that decrease consistent with altimetry data?

340: Given that v_{snow} and v_{fc} offset each other, is your new model effectively the same from an altimetry standpoint as the old model? By this I mean: is the new formulation a better mathematical representation of physical processes occurring in nature? Or, does 'new' differ from 'old' just in that you added several cores to the calibration and changed the surface density, so the calibration coefficients are different? Does this mean that in using your densification equation, the surface density must be prescribed as you do with Equation 2? If so, can your equation be used to model firn in a location where Equation 2 is not valid?

References:

Hawley, R. L., Neumann, T. A., Stevens, C. M., Brunt, K. M., & Sutterly, T. C. (2020). Greenland Ice Sheet elevation change: Direct observation of process and attribution at summit. *Geophysical Research Letters*, *47*, e2020GL088864. https:// doi.org/10.1029/2020GL088864