Reviewer #2

Review of: “Runoff from Greenland’s firn area – why do MODIS, RCMs and a firn model disagree?”

By Horst Machguth et al.

Summary

This paper examines and compares meltwater runoff limits in Greenland (i) derived from MODIS imagery and (ii) predicted by two regional climate models (RCMs), RACMO and MAR. The authors find that in general the runoff limits predicted by RACMO are lower than he observations, and those from MAR are higher than observed. The variability in the MAR limits is more closely aligned with the observations, while the RACMO results show comparatively little interannual variability. The authors attribute much of the difference between the models to differences in the meltwater schemes (bucket schemes). The higher runoff limits in MAR leads to higher predicted runoff volume (up to 29% along the KAN-U transect) than in RACMO.

In general, I found this paper to be scientifically sound and well written. It will make a quality contribution to our understanding of meltwater processes on ice sheets (and limitations in modeling them. There are several issues that should be addressed prior to publication, but I expect that most of these should be quite tenable. I have split my comments into “general comments” and “specific and line by line comments”; note that the line-by-line section includes both small comments (e.g., typos) and larger questions I had as I read the paper.

We thank the reviewer for their feedback. We reply to all comments below.

General Comments

1. My primary comment is that the discussion (particularly Section 5.2) needs to be a bit more substantive and based on the results presented. One of the chief takeaways of the paper seems to be that the formulation of the bucket scheme, but this section only sparsely uses the results to support their claim that the bucket scheme is responsible for the differences. For example, is there information that I can glean from figure 5 to support this?

In part this concern comes from the fact that the comparison uses two RCMs, and there is much more in RCMs than the bucket scheme. For example, are RACMO and MAR predicting similar amounts of snowfall and similar winter temperatures to each other, which could change the melt dynamics? The authors also present on albedo differences between the two, but are there other differences in the terms in the energy balance (e.g. calculation of the turbulent fluxes, downwelling longwave, etc.) that could be different between the two? What about how the models handle heat transfer, especially with phase change and associated latent heat?

I don’t think that the authors need to do a ton of additional analysis to this end, as I agree that the formulation of the bucket scheme is likely to make a difference. But, I do think it would be appropriate to include a paragraph or two about what other factors might contribute to the differences between the model, and why the bucket scheme formulation is the most important one.

This comment is similar to the overall critique of reviewer #1. We refer to our answer there, please also see our response to the major comment number 4 of reviewer #3. Differences in melt input, caused by differences in the surface energy balance, would obviously trigger differences in meltwater percolation. This was partially considered in the manuscript (mainly Fig. 5 left and right columns) but differences in melt input have now given more attention. The figure shows that both RCMs have similar melt. Albedo is also similar but there are differences in 2012 albedo directly above the RACMO-FDM runoff limit. This sudden change in RACMO-FDM albedo is discussed in the manuscript. Nevertheless, Fig. 5 shows that the main difference is not in melt or albedo, but in a major disparity in refreezing, which happens in the subsurface.

Apart from the revisions already mentioned to reviewer #1 we have more clearly emphasized the above argumentation.

2. This could be my ignorance, but throughout the paper the authors seem to use RCM and firn model somewhat interchangeably. I've previously operated thinking that the RCM is an atmospheric model, which is coupled to a subsurface snow/firn model. A bit of language (in section 2.2 perhaps?) clarifying this may be useful.

Agreed, we have revised the text to use a clearer terminology

3. The implications of the paper seem to be mostly limited to a vague paragraph at the end of section 5. Would it be a reasonable amount of work to include the total GrIS runoff from both MAR and RACMO and discuss the uncertainty in runoff with a more detailed discussion of the implications on our understanding of GrIS SMB?

We agree that the importance of runoff in Greenland's firn area for the ice sheet's total mass balance needs to be estimated. However, we would like this study to focus on exploring the reasons behind the models disagreement and quantifying the relevance of firn area runoff to total mass balance is beyond the scope of our study. However, we now highlight the study by *Glaude et al.* (2024) in the introduction and discussion as their work shows that there is very substantial uncertainty in simulations of future mass balance of the Greenland firn region. In particular, they demonstrate large differences in RCM simulated future melt extent which leads to pronounced uncertainties in future projections of Greenland mass balance.

4. Clarity: mostly the paper is well written and clear, but there are a number of instances (especially in the discussion) that were not written clearly. I note some of these in my specific comments below. My recommendation is to try to avoid writing in passive voice.

We thank the reviewer for pointing out these issues which we have corrected as described below.

5. Model settings: can you provide more detail about the firn model settings that were used? How were the firn columns spun up? Are model settings the same to the extent possible, e.g. surface snow density, etc.? Do both models use similar surface energy balance schemes?

We have added additional detail and references to publications describing the models.

6. The figures are creative and well made, but I generally found that soft colors were hard to see and that text (e.g., legends, axis labels) is too small. (There are also specific comments below that I made while reading the paper.)

We have modified axis fonts and colors, were possible.

Specific and Line by line comments

L31 paragraph: It may be worth mentioning here that the ELA and runoff limit vary from year to year (or calling it a "zone" to encompass that variability?), to differentiate between shifts in ELA and runoff limit that change the long term mean SMB.

Agreed, added.

46: "oppose" – I would choose a different word here. Intercompare? Or just "compare"? "analyze the differences between the runoff limits…"?

Agreed, done.

47: remove period after 2021

Done

Section 2.2 (related to my general comment 2 above): Can you provide slightly more detail/clarification about the differences between RACMO2.3p2 and IMAU-FDM, especially in the context of how you use them in this paper? My previous impression was that IMAU- FDM was coupled to RACMO as the subsurface scheme – is this not the case? Here, is the only reason you are bringing in the online version of IMAU-FDM to evaluate model physics and outputs that are not provided from the coupled model, or are you running IMAU-FDM as well for comparison? (Table 1 clarifies this somewhat, but I think it would be helpful to clarify the text slightly as well.)

We have added more detail, as also mentioned in our reply to point 2 above. Furthermore, the description of the firn models and modules in Section 3.2 has been extended.

83: "agree to" -> "agree with"

Done.

90: I get the gist of what you are doing as described here; however, is it possible to use Figure 1 to help illustrate the polygons?

Showing all polygons directly in Fig. 1 will make the figure unreadable and the polygons would be small. However, we added a supplementary figure to illustrate the concept.

Section 3.2.2 – regarding the IMAU bucket scheme – can you provide more specific detail about how it handles ice lenses and slabs, as you do for MAR? (this will help a lot with clarity of discussion section.)

Agreed, we have rewritten this paragraph in order to provide more detail.

151: Please clarify: Is the annual maximum runoff limit the highest elevation where runoff occurs for each year? Given the definition of runoff limit provided in the introduction, why is the maxY!"# not simply where runoff is greater than zero? Also, would it make more sense to consider a maxY!"# threshold as a percentage of the annual snowfall than just a value?

On the ice sheet, the runoff limit is indeed defined as the elevation where runoff starts. However, this definition cannot be directly applied to the RCMs. The RCMs simulation of runoff is fundamentally different from the actual processes and there can be numerical issues where tiny amounts of runoff always occur. In discussion with RCM modellers we decided to use a fixed threshold. Using a threshold that is a function of the amount of annual snowfall would make it very difficult to understand RCM runoff positions. Furthermore, the larger the amount of annual snowfall, the less likely near surface lateral runoff and instead aquifers will form, which are not the topic of this study.

183: can you speculate why the approach does not work well in that terrain?

Yes, we added a brief discussion in the appendix.

Figure 2: The caption and accompanying text needs more detail (text especially) of how to read the various lines. It may help to label the pale lines "IMAU-FDM daily" and "MAR daily". I initially missed the fact that you say where the transects are at the end of the caption. Perhaps include that detail in the first section of the caption, and make the flowlines bolder in figure 1? Figure 1 legend could also include the flowlines. Also, the axis labels on this figure are too small, and the figure would be easier to view if the colors were bolder. I don't quite understand the lat/lon labeled at the tops of the figures, as this is a transect and not a point?

Thank you for this feedback, of which some refers to Fig. 1, if we understand correctly? We have addressed these issues and hope that the new figure is more readable. The coordinates are provided as a rough indication to where the transect is located. This allows the reader to immediately understand whether a transect is in the far north or south. We have mentioned this now and also reduced the coordinate precision to one digit.

200: What difference are you referring to specifically here?

We have removed this sentence.

Figure 3: This is a neat figure. However, the text size throughout and the soft colors make it difficult to read. In panel a, is it possible to darken the flowlines as well?

Done.

Figure 3/Line 200 – It was surprising to see RACMO1k and IMAU-FDM differing so much. I wrote this comment before seeing the text about this later in the paper. I don't think you need to add more in this paper, but it is surprising to see and I hope you will investigate further in the future.

Thank you for the comment.

Figure 4: perhaps include the region along with the panel label at the top to add clarity, i.e. a: NW and b: K-transect

Done.

204: can you quantify this variability with a simple correlation?

Unfortunately, we do not fully understand the reviewer's comment. There are two variabilities mentioned on this line, (1) variability with intensity of the melt season and (2) differences in interannual variability between MODIS and MAR. Both were based on qualitative assessment, which we have now replaced with a quantitative assessment.

208: "shorter in the north than in the south" – is this robust, or just the case for the areas you picked?

This was a qualitative assessment, we agree it would need to be undermined with statistics. As it is not of particular relevance in the context of this study, we have removed the statement.

212: It May be useful in some cases when talking about a specific RCM to use e.g. Y$%& or &%'!"#

Yes, this has been changed

217: "appears more step-wise": Is this just an artifact of the gridding?

We removed this statement as our analysis does not demonstrate this statement quantitatively.

219: This looks like it is only true for MAR?

The challenge is the different temporal resolution of the available data: MAR at 1 day, RACMO-FDM at 12 days. Unfortunately, the latter are unavailable at 1-day resolution. We now briefly comment on this in the text.

Figure 6: the direction of the y-axis is opposite what I would expect (I would expect time to proceed downward in the direction of reading) – so I recommend switching that or noting this in the caption.

We added a comment in the figure caption.

Figure 6: It is not clear to me why the 10m firn temperature at the lower elevations much cooler than at the runoff limit (e.g., Figure 6g). Why does the melt cause 10m temperature at runoff limit to increase to near the freezing point, whereas at lower elevations this does not occur? I suspect this is due to there not being firn at all (figure 7), but it would be good to clarify here. (Perhaps label the figure T()# instead of T*+!, 10 m, as there isn't actually firn there?

This is correct, there is no firn at elevations much lower than the runoff limit. Consequently, meltwater runs off without releasing any latent heat in the subsurface. We have relabelled the figure as suggested.

245: Can you explain the odd (discontinuous in distance) 10-m temperatures in MAR? I would expect a (more or less) monotonic change as distance increases.

The snow layer of each pixel is evolving independently of the other ones from the initialisation. Moreover, the snow pack vertical discretisation could be very different following the considered pixel. There are thin snow layers close the surface but thick layers (of several meters) at depth. Therefore, the 10 m temperature shown here is the temperature of the snow layer including the 10 m depth. For some pixels, it could be the temperature of a unique ice layer of 15 m thick or to a 1 m thick snow layer close to 10m depth. To reduce this problem of spatial discontinuity (identified in this paper) in the MAR snow pack, a very light spatial smoothing/filtering is now applied over the 1st layers (the deeper layer) of each pixel in the latest version (3.14.1) of MAR (this study uses 3.14.0). We have commented on this issue in the text.

250: does this mean that in IMAU-FDM, ELA and maxY&'$ are effectively the same thing?

This could be the case, but we have not investigated the ELA as simulated by the RCMs.

266: "MAR simulates runoff between the two runoff limits" – not exactly clear what you mean by this. I think you mean the additional runoff simulated by MAR above the IMAU-FDM runoff limit and below the MAR limit, but perhaps there is a clearer way of stating this.

Thank you for pointing this out, we have tried to clarify the wording.

273: “the larger the difference in total runoff simulated by the two RCMs” – even with the below paragraph I think is phrased as too strong of a claim, as I don’t see anything in this regression based on the total runoff in each. I think the claim needs to be a bit more nuanced, along the lines of “the amount of MAR melt above the IMAU-FDM runoff limit increases exponentially as a function of the MAR melt below the IMAU-FDM runoff limit. Assuming the MAR and IMAU total melt below the IMAU-FDM runoff limit are similar, this implies that the difference in predicted melt between MAR and IMAU increases in high-melt seasons”. (or rearrange the section a bit so that the comparison of common area is not at the end.

The total runoff in RACMO-FDM is represented in the regression by MAR runoff below the runoff limit of RACMO-FDM. The RCMs’ similarities in “ablation area runoff” are discussed in the following paragraph. We have changed the wording in line with the suggestion of the reviewer.

279: is this maxY-$%.?

Yes, this information was missing, updated.

285: “16 % out of which almost four fifths” (and previous sentence) – these statistics are hard to follow – i.e., 80% of 16% - can you make it a bit clearer for the reader (perhaps adding the actual volumes would do this?)

Adding volumes can lead to more confusion as the analysis focuses on a transect where runoff has, in a strict sense, the unit m^2. We have reworded to clarify the message.

288: “regardless of fundamental differences between runoff processes detected from remote sensing and their simulations.” – clarify the sentence structure here.

Done.

306: Aren’t bucket schemes instant (all routing occurs within a model time step)? That could be mentioned here instead of “RCM vertical routing is much faster”.

Correct, done.

311: It is not clear to me how what is described in this paragraph is lack of inertia – can you elaborate what you mean?

The word “inertia” does not add to clarity. In line with the above comment, we also use “instantaneous” here.

315: “This feedback mechanism, by which ice slabs thicken, is challenging to mimic through a relatively instantaneous bucket scheme.” I don’t necessarily agree with this – in the model domain, once there is meltwater percolating to the slab all snow/firn above the slab is temperate. If the slab is below freezing, some of the meltwater can refreeze (according to the heat transfer scheme in the model). Why would an instantaneous bucket scheme prevent the model from simulating this process correctly, albeit in a single time step rather than over some timescale? If the slab is being buried in IMAU-FDM and not in reality, a simpler explanation to me is that there is too much snowfall and not enough melt in RACMO at KAN-U. Wouldn’t this be consistent with your findings that IMAU-FDM has lower-elevation maxY&'$?

Our statement was misleading or incomplete. The issues lie in the RCMs limiting meltwater storage in firn to irreducible water content. No slush formation is allowed, which by definition is meltwater exceeding the irreducible water content. The ice slabs thicken primarily by accretion of superimposed ice on top of the slabs, which requires slush to persist on top of the slabs for longer time periods. This cannot be simulated by the RCMs at the moment. Too much snowfall could also cause the effect of the slab getting buried, but this is not the problem here. We have modified the text for the message to become clearer.

316: "In particular, both RCMs do not permit any slush formation and even thick ice layers must remain "permeable" for meltwater to be routed vertically." I am not sure what your point is with this sentence – can you expand a bit to clarify what you are saying?

If the RCMs would treat layers of the density of ice (or any other threshold, such as density of pore close-off) as impermeable, then this would also require to have percolating meltwater ponding on top of such an impermeable layer. The ponding water would completely fill the pore space of the overlying porous layers which would mean that slush is formed in an RCM's firn layer. Both RCMs do not form slush because any water exceeding the irreducible water content is instantaneously routed downwards.

We have reworded the criticised sentence to make the statement clearer in line with the above explanation.

330: "in the absence of pore space, even moderate amounts of melt will run o-" – isn't this what would be expected in reality? If there is no pore space available to store the water shouldn't all water run off?

Yes, this behaviour is intended. The question is more whether the pore space, as simulated by RACMO-FDM, corresponds to the spatial distribution of pore space in reality. We changed the wording to make this clearer.

332 – 335: in general this paragraph is hard to follow, in part because (as I noted earlier) the description of the bucket scheme in IMAU-FDM is not fully described. Here it is implied that any meltwater is allowed to pass through ice slabs – is that correct?

Yes, this is correct. As explained above, no water can pond on any ice layer regardless of who thick that layer is. Hence the water passes through ice layers. However, it does this without any interaction with the ice. For any interaction of percolating meltwater with a layer in the firn model, a layer needs to contain pore space and existing water content in the layer needs to be below the maximum irreducible water content. We have added text to clarify this behaviour.

336: "pronounced step change in surface albedo in 2012" Is this shown anywhere? Is this in time or space? Not exactly clear what you mean here.

Yes, this is shown in Figure 5c. We have added a reference to the figure for clarity.

337: Why does higher albedo reduce likelihood of "meltwater percolating to the bottom of the firn where it would run off"? Or are you trying to say that the higher albedo above that step change reduces melt and thereby reduces the volume of meltwater percolating to the bottom of the firn and running off?

Yes, this is what we are trying to say. We changed the wording to make the message clearer.

340: on one hand/on the other – colloquial phrase

We have adapted the wording.

358: “As the primary cause we identify the discrepancies between the two maxYrcm”: this could be rephrased to be clearer, e.g. the primary cause is the fact that MAR consistently predicts a higher runoff limit (and thereby a larger area producing runoff) than IMAU-FDM”

Done

359: why is this surprising? If both predict similar runoff in the ablation zone, wouldn’t the one with higher runoff limit be expected to produce more runoff?

Our message was unclear. The “surprising” refers to the fact that melt is rather small in the vicinity of the runoff limit as compared to the ablation area. Hence, one could expect that differences in modelled ablation area runoff dominate total difference in runoff. As this is unclear, we have reworded the sentence, avoiding the word “surprising”.

393: “strongly later water flux” – can you clarify what this means, or perhaps a typo (layered?)?

Thank you for pointing this out, it should read “strongly lateral”, changed.

Figure A4: this is pretty wild to see this much disagreement – I hope you’ll continue work like the present paper to figure out what is going on with the models here.

Thank you for the comment. Yes, the disagreement is substantial. It is challenging to simulate the interaction of firn and meltwater. We hope our work motivates further research on this topic.

References

Glaude, Q., Noel, B., Olesen, M., Van den Broeke, M., van de Berg, W. J., Mottram, R., et al. (2024). A factor two difference in 21st-century Greenland Ice Sheet surface mass balance projections from three regional climate models under a strong warming scenario (SSP5-8.5). *Geophysical Research Letters,* **51**(22). https://doi.org/10.1029/2024gl111902