General comments

This manuscript describes the method used to generate a new dataset of frontal ablation for a sample of Greenland's tidewater glaciers. The dataset and workflow will be a great asset to the community and I am confident that the products will enable lots of good science. Overall, the method appears sound but I think the pre-processing of the terminus delineations and the estimate of ice thickness changes need to be improved to avoid erroneous changes in glacier volume. The majority of the manuscript is used to describe the method for calculating mass changes owing to terminus advance and retreat – I think more attention needs to be devoted to describing and documenting the dataset itself, and to quantifying the impact of various choices through the processing chain. In particular, the contribution of advance/retreat to frontal ablation should be quantified for all glaciers, individually and combined, at a range of temporal scales. I think the error metric for terminus mass change needs to be improved, particularly through the inclusion of thickness errors.

Specific comments

<u>Data product:</u>

The errors in D and TMC should also be provided, so that users can see how they contribute to the error in F.

I think it would be helpful if the time-series for each glacier were combined into a single file, perhaps a NetCDF. This would make it much easier for potential users to amalgamate the measurements from different glaciers.

A few issues that I assume are minor:

- For Akullersuup Sermia, the last two entries for TMC and F are empty, but D and Max_Error_F are provided. Is that expected?
- Time period: Again for Akullersuup Sermia, data are available for the period 1988 to 2018. In the manuscript, the abstract states 1987-2020, the caption of Figure 9 states 1987-2015 but line 361 of the submitted version states 1987-2018. Do different glaciers have different temporal coverage, with some (a few or most?) extending to 2020? If the glaciers do have different time periods, it would be easier for users if you at least make all of the time-series the same length by padding with NaNs as necessary.

<u>Data sources:</u>

Initially, I found this section a little confusing. It reads somewhat like an abstract for the methods, giving a very brief overview but without much detail, some of which is repeated again in the methods. The purpose of this section is just to tell the reader what data you used – giving them some identifier or name – so that you can refer to them unambiguously in your methods. With that in mind, I wonder if this section might be better structured as a series of bullet points where you clearly state the parameter (terminus position, elevation, ice velocity, discharge) and what data source(s) were used, then bring Supplementary Table into the main text at this point. I would recommend adding the time-period to the table (and possibly to the text describing the data).

<u>Methods – fjord geometry and terminus positions</u>

Line 169: lower/upper fjord wall is hard to understand. Should it be southern/northern? If so, then is the code able to deal with glaciers flowing north to south or south to north? Could you use true left or right instead – that would be independent of glacier orientation, but it would entail bringing in the flow direction data earlier, so might not be worth the additional effort.

Fjord wall delineation and polygon creation: overall, this approach for creating polygons seems very robust, but I have a few queries and have spotted a few areas where the algorithm appears to struggle:

1. For ice streams like Jakobshavn Isbrae/Sermeq Kujalleq, where there is no clear lateral topographic boundary in the vicinity of the present-day terminus, how did you determine where to place the fjord walls?
2. As glaciers thin, they can reduce in width. Depending on the date of the image used to digitise your fjord walls, you might be slightly (and it probably is only very slightly) overestimating or underestimating the change in area owing to terminus position change in years far before or after the image was acquired. I think it would probably be more effort than it's worth to quantify the impact of this because the associated error is likely smaller than other errors, but I suggest that this effect is acknowledged somewhere.
3. Definition of terminus width when filtering polygons: As I understand it, the purpose of this step is avoid making long extrapolations from the digitised terminus to the digitised fjord wall. Just browsing some of the terminus polygons, I noticed some that still had what I judged to be quite long extrapolations. For example, Jakobshavn Isbrae/Sermeq Kujalleq on 2002-05-30 has a ~4.5 km completely straight line and 2004-10-03 has a ~6 km straight line, which I assume are extrapolations. I assume this is occurring because the fjord widens a lot as the terminus has retreated, so even short digitisation along wide sections of fjord, could be as long or longer than the minimum fjord width further seaward. One alternative approach that comes to mind is to compare the length of the digitisation to the distance to the nearest fjord wall – if the remaining distance is only a small fraction of the length, then extrapolation may be ok. Actually, if one end of the digitisation intersects the fjord wall, then you could even compare the straightline distances between the unfinished end point and both fjord walls to determine whether extrapolation is justifiable. I deliberately chose this glacier because I thought it would be tricky, as the width of the fjord changes a lot and the glacier has bifurcated, so these might be isolated issues.
4. Linking broken digitisations: I can see that you remove multi-part delineations in stage 1 of your filtering, but I found some examples (e.g. Jakobshavn Isbrae/Sermeq Kujalleq on 2005-03-03) where it looked like linking digitisation segments was done in the wrong order, resulting in the northern part of the digitisation being extrapolated to the southern fjord wall, creating lots of self-intersections and probably a very large apparent terminus advance. It may be worth checking that stage 1 is working, or perhaps throw in some checks for large gaps between segments would help to get the order correct. Or at least check for large gaps after linking everything would help you to discard this type of failure.
5. There are some terminus digitisations that appear to go outside of the fjord walls (e.g. Jakobshavn Isbrae on 2015-03-04). I assume this is because the walls were digitised before all of the terminus traces are available. Is this likely to be consequential for the measurement of frontal ablation? If not, then I think it would be acceptable to retain or exclude these as you see fit, rather than modify the algorithm or redraw the fjord walls.

Removal of delineations with fjord walls: I was surprised to see that you remove the PROMICE delineation in Figure 3. I couldn't find the corresponding part of the text that described that step, but it looks like an important choice. I don't know how many delineations that step removes, so perhaps it makes little difference, but wouldn't it be straightforward to include delineations of that type and just clip them when they intersect the fjord wall (which you are doing anyway)?

Use of velocity to remove traces that have advanced too far: I really approve of this step, but I noticed some digitisations that had advanced unrealistically far, so I wonder if further cleaning is required. For example, at Jakobshavn Isbrae, there is a ~12 km advance along the southern fjord wall from 2007-03-08 to 2007-04-05, which seems unrealistic to me. The centreline does not advance so much, which is presumably why the algorithm retains the latter trace. I wonder if setting a maximum rate of change in area for each glacier, based on the long-term average, width-averaged near-terminus velocity would help to identify and remove these traces?

- A minor point on this step: I found the use of "midpoint" and "centreline" a bit confusing. Presumably, the intersection of a central flowline with a trace is not always the midpoint of the trace?

Line 186/7: As written it sounds like a trace on, say, the 30th of the previous month would not be used even if the next month did not have a trace until, say, the 10th. Is that right? I'm also not sure I follow this logic, because it implies that delineations closest to the 1st of each month would somehow be more accurate than other traces. Why not just say that the accuracy of the measurements is too low to resolve short-term (<1 month) changes in terminus position? If that is the intended meaning here.

Line 191/2: Can you clarify what is meant by "subsequent temporal averaging"? It sounds like averaging of terminus positions, but I'm not sure that is intended.

Line 235: "quality control and temporal filtering". Do these refer to the steps described above? If so, I can't see where temporal filtering comes in.

Line 247: "nearest point on the polygon boundary". Just looking at one polygon boundary (Jakobshavn), the points are spaced quite far apart. Since you connect the terminus delineation with polygon points, not polygon edges, this means the terminus doesn't always go to the closest part of the fjord. I wonder if it would be worth making your polygon point spacing smaller, to improve that connection. I have found combining arclength and interparc quite helpful for this kind of thing:

Line 249: Somewhere further up I mention some traces that have quite long extrapolations to the fjord edges and suggest a comparison like this to reduce those. Since you already have it implemented, it may be worth trying with some different thresholds to reduce the extrapolation distance.

Methods – surface elevation change/ice thickness

Line 127 (data sources): "glacier specific surface elevation change rates" – without checking the dataset from Khan (2017), it's not clear what this means. Are they scalar values for each of the study glaciers representing the average rate of change over that period, or are they time-series of elevation change rates representing rates over certain (ir)regular epochs? Are they observations of elevation change, or based on fits through the data? Are they averaged across the whole basin, or some smaller area near the terminus? If so, did you do that here or is that how the data come from Khan (2017)? If you do retain this section describing the data sources, then I think it would be worth clarifying these points.

Line 128 (data sources): “latest ArcticDEM image...”: Perhaps some of this becomes clear in the methods, but I’m having trouble working out how it can be a single ‘latest’ image, that also covers the most advanced and retreated position, and that also covered the full extent of the tidewater glacier. Presumably you needed at least two images per glacier and picked images acquired closest-in-time to when each glacier was at its most advanced or retreated? The same goes the AeroDEM. I think you should clarify which ArcticDEM dataset was used here – the strips or the mosaics, and at what resolution? I think ArcticDEM begins in 2007, so if I understand correctly, you have an 18-year period (1988-2006) with no surface elevation observations – at least, that is how it appears from the short statement in the data sources section, which is one of the reasons I suggest rethinking that section.

Line 255: Careful with “ice surface elevation” vs “ice surface elevation change”. As written, it’s not clear which you have from Khan (2017).

Use of ArcticDEM and AeroDEM: I have some concerns about the direct use of these datasets. The ArcticDEM is an excellent product, but I do not think that the strips are georeferenced sufficiently well nor erroneous height measurements sufficiently well removed to use the strips as-is in a straight DEM differencing. I am not familiar with the AeroDEM, but these may suffer from similar imperfections. In this case, the thickness changes may be a small fraction of the volume and mass change of the glacier, given the small integration area, but that should be quantified – if you applied no thickness changes, how would that affect your measurements? Nevertheless, I don’t think it is justifiable use the raw strips and you should either perform some corrections on the strips to remove erroneous data, remove vertical biases and ensure proper alignment, or you should use some alternative elevation or elevation change dataset.

- A related but straightforward question: Are the AeroDEM and ArcticDEM relative to the same geoid?

Line 269: “the SCR is” – can you clarify where the SCR comes from in this case if no Khan (2017) data are available.

Line 276: can I check that you ensured that the ice surface and bed topography data were relative to the same geoid?

Line 279: this point is only really relevant to Jakobshavn Isbrae up to 2003 - when calculating the volume of floating ice, how did you determine the ice thickness?

In this section, you draw on a range of datasets with very different spatial resolutions. Can you clarify which datasets were up or downsampled and to what resolution, in order to perform these steps?

<u>Methods – uncertainty quantification</u>

Line 316: statement re ice density uncertainty is repeated on line 314

Line 326: a 30 meter error for the delineated part seems acceptable, but some of the polygons contain long segments that were extrapolated, not delineated. I think that it would be worth retaining information about the length of the extrapolations and incorporating that into the error estimate somehow. Alternatively, you could erode your polygons by 1 or more pixels and calculate the resulting change in area.

Equation 9 and preceding text: shouldn't this just be the root-sum-square of the mass uncertainties at t_1 and t_2? Where the mass uncertainty is the root-sum-square of the uncertainties in the length, width and thickness estimates at each time.

- As written, the thickness uncertainties are not described (at least not here). They should account for the uncertainties in the reference surface elevation, surface elevation change between the reference time and measurement time, and bed elevation.

Results

Paragraph starting line 360: it would be nice to see a comparison of the long-term average discharge and long-term average frontal ablation for all of the glaciers combined and individually, either/or as numbers of a graphic. There are some glaciers listed here that don't appear in the top contributors list of Mankoff et al. (2020), so the inclusion of mass change owing to terminus advance/retreat is clearly affecting the picture of solid freshwater export from the ice sheet, but it would be great to put numbers on that. This would really support the statement on line 374 that include terminus changes increases frontal ablation – I'm sure it does, but you should use your dataset to demonstrate it.

Line 373: does this imply that including terminus position change does not affect frontal ablation on longer time-scales? As with the above point, including the numbers somewhere in the main text will demonstrate it either way.

Line 397: "values which have to be summed…" – since you have a time-series of terminus area, you could directly calculate the difference between some appropriate terminus at the start and end of each decade, then you wouldn't have to accumulate the errors through time. I appreciate that is not the 'product' being published here, but it would be a fairer comparison. Writing this made me realise that the comparison with Kochtitzky et al. (2023) is also quite difficult because termini can move quite quickly over short times, so there is no such thing as a decadal average terminus position change – it will always be a comparison of snapshots at two times, regardless of how far apart they are. With that in mind, it might just be worth comparison your termini with theirs, as that is likely the largest source of error.

Line 408: I can't quite work out what you mean by "spatial variability between observations" in this context. Is this to aid comparison between glaciers at some point? Can you rephrase this?

Line 404: "any tidewater glacier" in Greenland or anywhere globally?

In this list starting on 405: don't you also need surface elevation change data and discharge?

Conclusion

Line 420: it is good to see a statistic here, but from the data presented in the manuscript, it is hard to evaluate how representative or whether it is an extreme case that occurred at one time on one glacier. It would be great if you could quantify how much the inclusion of terminus position change affects frontal ablation for each glacier and all glaciers combined over seasonal, annual and longer time-scales.

Technical corrections

Line 165: comma after “Here”

Line 184 and line 194: geolocation of georeferencing?

Line 188: 1986-present. Check correct because the data and abstract show 1988-2018/2020?

Line 381: odd use of brackets in the citation

Line 383: remove “however” and replace comma with semi colon (I think…)

Line 442: check the unit on “terminus mass change” in the glossary and elsewhere

Figure 4 caption: “m_2 is to the right of the normal n” – it is when viewed from that perspective, but can you be certain that is always how it is being implemented in the code when you have glaciers flowing in different directions? Would the central flowline be a useful reference point from which to determine direction here?

Figure 8 and equivalent panels in the SI: I find the lines in this panel very hard to read at 100%. Can you make them thicker and produce the plot at the intended display size, if you aren’t already? The max error text is also impossible to read at 100 %.

Figure 9: check for font size consistency

Figure 9: are the circles representing 14-20 % identical? I cannot tell the difference.

Figure 10: Perhaps it’s personal preference, but I think this plot would be clearer if you (1) included the difference between the estimates and (2) sorted the glaciers according to that difference (or sort by frontal ablation magnitude).

Figure 10: I find the bars quite hard to see – could you make them wider by using up some of the white space between glaciers?

Figure S4: This may be more than a technical correction. I was surprised to see a glacier included that did not include discharge, given the contribution of discharge to frontal ablation. I would recommend removing this glacier and others like it, or perhaps including them separately somehow.

Benjamin Davison