Review of 'Debris cover and the thinning of Kennicott Glacier, Alaska, Part B' by Leif Anderson et al., under consideration for The Cryosphere

Part B of the Anderson et al trilogy aims to combine empirical relationships of surface properties and melt rates, based on the field measurements presented in Part A, with remote sensing observations of different surface features (particularly ice cliffs) in order to arrive at distributed estimates of melt rates for the period of observations. The analysis then uses these distributed melt values to address whether or not the identified melt hotspots can explain the thinning patterns evident at Kennicott. This is an important question as the glaciological community is trying to disentangle the influence of surface mass balance and ice dynamics for debris covered glacier evolution, and this is the first time the question has been addressed in Alaska, where surface debris is prevalent.

As such, this represents an important contribution to a current topic of research, and provides an answer to that question for Kennicott Glacier – these hot spots do have an important effect on the surface mass balance, but it is not plausible that they compensate for the overall melt reduction due to surface debris. The study could be more systematic to provide a definitive answer, and I have some criticisms regarding the empirical relationships presented in determining the distributed estimates of melt, as well as the difference in temporal scales between long-term elevation change (52 years) relative to single-year field measurements. However, although I have quite a few comments, no changes along these lines are likely to change the conclusions of the study. Rather, my principal concern is that the separation of this analysis from both Parts A and C reduces the strength and presentation of the entire analysis, while also leading to repetition of text and figures, as well as cross references. There are certainly gaps in the analysis because aspects have been included in Parts A or C rather than here, and some restructuring across the three manuscripts might improve the readability of all three. This is a choice for the authors and Editor to contemplate, but my opinion is that some consolidation would be beneficial.

Main comments:

My principal question in reading this manuscript was whether it should be standalone or integrated with Part A (and possibly C, which I have not reviewed). I appreciate that this is a difficult decision, and that Parts A, B, and C combined represent a substantial body of work. There are certainly some advantages to be considered for maintaining separation between the manuscripts, but my impression is that A and B are both weaker manuscripts separated from one another. The field measurements were clearly collected for the purpose of deriving distributed melt rate estimates, which leaves Part A without a compelling conclusion or discussion. At the same time, Part B requires frequent reference to Part A, or blind faith of the part of the reader with regards to the methods and results of the field data. As a consequence, there is also quite a bit of repeated material to cover (for example, content from 4 figures in Part A is also displayed in Part B). My instinct when reviewing Part A was that much of that material could be more meaningful if integrated with Part B, in supplementary information if not in the main text, and I now think that would greatly improve the readability of this manuscript.

I like the ice cliff delineation method in its simplicity (although some details need to be clarified, below), but its transferability is not very clear. Often when developing/proposing a new method it is necessary to see how robust the method is, but in this case the method has clearly been developed specifically to map cliffs in this particular scene, in order to apply the empirical melt relationship. As such, it is a relatively small part of the story in deriving the distributed melt estimates. At the same time, the maps of supraglacial ponds are not integrated very well into the story, while the exclusion of supraglacial streams (also mapped from satellite imagery) is a bit

The study aims to address the role of melt hotspots in explaining thinning rates. For this to be a definitive analysis in this regard, I feel like this needs to be done in a more systematic manner, whereas the present analysis seems to focus exclusively on ice cliffs. The distribution of supraglacial streams seems to be a major gap in the analysis of this part of the study, especially as the properties of supraglacial streams are assessed in Part C. Streams are mentioned throughout the background as hot spots and possible factors contributing to melt, but are then completely neglected in the methods and discussion. I understand that the role of surface streams cannot be assessed quantitiavely in this manuscript, but neither can the role of ponds. Similarly, although internal ablation is usually regarded as negligible in this type of specific mass balance assessment, there have been suggestions that this is a non-negligible term for extensively debris-covered glaciers. It is exceedingly unlikely that this mechanism could lead to the debris-cover anomaly, but for completeness I think it should be considered numerically along with the cliffs and ponds.

The relationships between elevation and debris thickness, and between elevation and ice cliff backwasting, are very weak. In neither case does elevation appear to be a primary control of the property. How different would your results be if you interpolated spatially (within flowlines or lithological units for example)? Did you consider alternative methods to provide a distributed debris thickness estimate? For example, one could consider the radiance measured by satellite thermal infrared imagery as related to debris thickness, and use your measurements to constrain this relationship locally to upscale in space much more meaningfully. For ice cliffs, justification of a fit to elevation needs to be more explicit in this Part of the study.

The conversion of backwasting rates into ice cliff melt misses a slope correction for the cliff area, which results in an underestimation of melt (see comment on L190). This raises a difficult question that has not been carefully considered yet for debris-covered glaciers, which is that the real surface area can be 10-20% higher than the planimetric area. For this study comparing geodetic thinning observations with estimated melt, that is an important aspect to consider, as melt occurs relative to the real surface area. This effect is especially pronounced for ice cliffs, but is also crucial for the 'background' melt rate if the glacier surface is highly variable.

It is too bad that more recent geodetic difference data were not included in the analysis. At present it is not clear how the long-term thinning rate relates directly to the 2011 observations. Would it be possible to use the ArcticDEM datasets to derive a recent-period thinning pattern? This would be more meaningful for a comparison to the contemporary distribution of surface features. The long-term perspective is still useful for understanding dynamic changes, but comparing 50-year lowering rates to one-year melt patterns does not provide a definitive answer, especially for a clearly changing system.

Some rewriting is needed for readability and presentation standards in The Cryosphere. Although the ideas are well developed, some sections of the paper read as bullet points and/or the word choice has not been considered carefully, leading to some of my comments below.

Detailed comments:

L23. Presumably this 19% of melt is for the debris-covered area?

L25. Just a comment, for you to adopt or disregard: The literature has tended to use 'ponds' for these features as they are much smaller than supraglacial lakes on, e.g., the Greenland or Antarctic Ice Sheets.

L25. It would be nice to have the %areal coverage numbers here in the abstract, not just 'doubled'.

L27-27. This wording isn't very clear. By average melt rates you seem to mean the average for all surfaces, but as worded this seems to refer only to sub-debris melt. In the latter half of the sentence, do you mean that the overall melt relationship still follows an Ostrem-type relationship, even after accounting for cliffs and ponds?

L34. These are broken sentences. The first half needs a reference even if it is now well understood.

L95. Has this happened progressively in the past 82 years, or primarily over some later period?

L98. Why is 'partially' here?

L108. Which spectral bands of the image do you use?

L113. This is true in certain conditions, depending both on debris lithology and meteorology. Such conditions may be prevalent for Alaskan supraglacial debris and melt seasons, but it is important to think whether such a method is transferable.

L118. This is semantics perhaps, but I would argue that this whole workflow is your 'cliff detection method' (the name of step 2). It is a bit strange to have a 'cliff detection method' as the main step of a 'cliff detection workflow'.

L118. Is the histogram stretch just a linear min/max stretch? Is this histogram stretch applied to the image globally or locally within a patch? What spectral band(s) are used in either approach? I suppose that you also start with a debris outline (and glacier outline), which is important to acknowledge for a remote sensing approach.

L123. Is the saturation stretch part of your step 1, or a separate part of step 2?

L125. I see that the size of the moving window is included later as one of your parameters in the MC optimization. It might be good to give the reader a warning that that is the direction the methods are going, and that you will start with a description of the implementation of each approach first.

L126-7. I would certainly not say that adaptive thresholding is insensitive to changes in debris cover and illumination, but it may be less sensitive. Ice cliffs are not uniform in surface character, and can appear both brighter and darker than the surrounding topography in different circumstances. It may be that these nuances are not so evident in the lower portion of Kennicott Glacier, but two particular cases would pose a major challenge for the ABT approach: 1) a population of ice cliffs with variable surface character (debris-free vs covered with fines) which will increase the spectral variance of the cliff population; 2) otherwise dark (potentially wet) debris. This is discussed later in the manuscript, but for a presentation of a new method, I think the accuracy and appropriate application needs some further advertisement/warnings.

L133. Some of this content is Results

L142. Was this 3% disagreement in total area? Can you provide a dice score for the two independent outlines? It is very possible to derive largely different cliff distributions but arrive with the same area.

L144-149. If I understand correctly, there are thus 6 parameters for the ABT and 5 for the SED implementations, with 4 shared between the two. How did this occur in practical terms? 2500 runs for each implementation, or 2500 runs used the same values for the shared parameters? More importantly, it is worth noting that with 5 parameters, 2500 runs results in an effective sampling of ~4.8x in each parameter (4.8^5 ~ 2500).

L152. 'The origin' is a bit ambiguous here, as it is true for figure 5, with the x-axis as the negation of the true positive rate. So really this is ranked by distance from (1,0) in your optimization space, correct?

L153. Why did you choose to reduce the FP rate (at the expense of TP) from the optimal parameter set? Can you please provide a dice coefficient for this parameter set for each approach?

L157. Process observation (2) actually refers to melt rates, rather than backwasting rates.

L158. The influence of lakes was noted earlier by Brun et al (2016) and Miles et al (2016), among others.

L163. In Part A it is clear that while elevation is a principal driver of debris thickness variability, there is considerable heterogeneity within any elevation bin. As your field measurements of debris thickness could not encompass the entire study area (that would not have been feasible), do you think they sufficiently characterise the unmeasured area (particularly the NW of the domain)? Have you tested the importance of debris thickness heterogeneity in your overall melt estimates? The subdebris melt relationship is not linear, so melt calculated with a mean thickness may not accurately approximate the mean melt rate.

L167. It is worth noting that you neglect internal ablation as well as other thermokarst processes (ponds, streams) in this computation for practical reasons.

L176. Please provide a goodness-of-fit for this empirical equation.

L180. It is interesting that as formulated, b_ice is the measured clean ice melt rate near the top of the study area, rather than the lapsed melt rate for each debris point. This is much more practical, but ignores the real melt suppression by the debris as a shortcut to a rate. In any case, I presume that the equation (as in Anderson and Anderson 2016) is based on the measurements presented in Part A? Please provide a goodness-of-fit measure.

L185. I am sceptical of this linear fit given the spread of observations in Part A – a goodness of fit would be expected to be very low. If elevation is a secondary control, what might you presume is a primary control for the difference in backwasting rates?

L187. The similarity of backwasting rates for cliffs with/without lakes may be due to the observation type and period. Ponds and streams tend to incise thermoerosional notches, which can later collapse, thus enhancing the seasonal mass losses but not affecting what one would observe from the top of the cliff over a month or two. This is not a criticism of your work, it is just worth noting that this nonobservation doesn't mean a melt enhancement is not occurring.

L190. The correction in backwasting rates for cliff slope is correct, but it is also necessary to correct the cliff area from planimetric to surface area in order to correctly estimate melt from these inclined features, as melt occurs perpendicular to the surface. Thus

A_i =A/cos(theta)

With theta=40 degrees, this is a factor of 1.3 to all of the cliff-related melt calculations.

L192. I think it's reasonable to apply a constant slope for the melt calculation, but why is this the 'most likely case'? Can you please provide some supporting information as supplemental figures, etc?

L201. 'Most likely' is superfluous here; it is an estimate. It's nice that you provide bounds!

L202. This 'best estimate' is using the parameter values already given in the text, correct?

L213. I suppose you use the lapse rate (per timestamp? Hourly? Daily? Mean?) between the two stations for this estimation? It is notable that this lapse rate approach corresponded poorly to your on-glacier temperature observations (in the debris-covered area, Part A) – how do you think such an approach would correspond to the on-ice calculations?

L218. Did you attempt to digitize ice cliffs in 1957 as well? The mention of ponds (and long-term change) is quite sudden, and should maybe be better integrated with the text.

L219. 'insure' should be 'ensure'

L243. Should this be between 520 and 620 m? The fractional area is more meaningful than total area for understanding the cliff distribution.

L245. Importantly, this is of reported values.

L247. N is too small to note any meaningful correlation between debris thickness and ice cliff coverage, unless you have a physical mechanism to implicate.

L255 and 265. I am still confused about the ice cliff melt rate distribution, for 2 reasons. First, it appears that at high elevations in the study area, it appears that your modelled ice cliff melt rates are lower than the modelled subdebris melt rates. This is not plausible (or there would not be cliffs!). Second, the ice cliff melt rates in this region are also lower than for clean ice, which should be an approximate lower bound for ice cliff melt at all elevations: are nearly bare ice, but with surface debris well below the critical thickness (thus enhancing melt relative to bare ice, if we can neglect increased shading).

L257. These rates correspond to a mean cliff enhancement factor of 1.72 relative to the mean subdebris melt rate, which I suppose is lower than anywhere else due to the thin debris.

L259. 'Dominates' is a strange term here. Certainly the reduced melt rate due to debris thickness is apparent, and debris thickness differences are more important than the difference in cliff density.

L265. This appears to be a typo – the cliff melt rates are an order of magnitude lower than under debris?

L271. This is again very disjointed to the rest of the analysis. Also, the low lake coverage in the upper ZMT makes a lot of sense as this area has steeper surface slopes in 2009 (Fig 2).

L282. I appreciate consideration of the applicability and extension of this method to other sites/scenes. I think the biggest challenge for application to other scenes is that the tested parameter sets produced extremely variable results, and would need optimisation for every new site and image.

There are also seasonality patterns to consider- cliffs often retain snow longer than the debris surface, for example.

L324. The variability in observed backwasting rate is considerably stronger than any bias due to the observation location – the question is really where the mean lies.

L340. The potential distal effect of these features is conceptually well understood to be via internal ablation along englacial conduits, but is not possible to validate at present. See Benn et al (2001, 2012, 2017), Sakai et al (2002).

L344. This section/paragraph feels orphaned. It is worthwhile to note that even accounting for the hypsometric distribution of cliffs, the spatial pattern still emulated Ostrem's curve (just with different effective thicknesses), suggesting that this concept might be useful as a proxy for the altitudinal SMB pattern even where cliffs account for 40% of melt – just not directly comparable to stake measurements.

L353. 'counter' - should be singular

L356. 'trend' should be pattern

Table 3. I don't think that Buri and Pellicciotti (2018) is the most appropriate study for Lirung Glacier for this purpose. Why the comment on EB below the table?

Figure 2. It is not clear what the bars are in the upper left – is this the domain with supraglacial debris?

Figure 3. This is a very nice conceptual summary! Can you include a pond or stream?

Figure 4. In the caption for 'c', there is a reference to a 'black line' which corresponds to the 'solid' line I think.

Figure 5. This is a very nice summary of the method. Can you reproduce the same for the Sobel method to be included in the Supplementary Information?

Figure 6. Panel (b) does not depict the bare ice area outlines as in (c).

Figure 7. Nice depiction of the optimization. No colorscale is shown, though, and due to the different axis ranges, it is difficult to visualize the lowest Euclidean distance.

Figure 9. See my comments in the text on line 255. Surely the lowest ice cliff melt rate (here 2.9cm/d) should correspond in space to the highest sub-debris melt rate (5.8 cm/d) – at the highest elevations. But then, the cliff melt rate should not be lower that the subdebris melt rate – that makes little sense. This suggests to me that the linear parameterization of ice cliff melt with elevation may not be appropriate.

Figure 10. Nice summary. Can you include a depiction of the cliff-only melt rates vs elevation in panel (a)? It is interesting that the cliff portion of melt is highest high in the ZMT, but still makes little difference in the mean melt rate profile. Also, it would be very meaningful to complete a version of panel (a) for the min and max melt parameterizations. Effectively these estimates are generous uncertainty bounds for your results.

Figure 11. Do you have a depiction of supraglacial streams (density or otherwise) to complete the picture?

Figure 12. This is orphaned from the discussion and seems like an odd figure to close on.