Review of 'Spatial and temporal variations in glacier surface roughness during melting season, as observed at August-one, Qilian mountains, China' by J Liu, R Chen, and C Han

The study by J Liu et al. demonstrates the first use of an automated photogrammetric apparatus to monitor surface roughness at a daily timescale for an ice cap in China. The authors supplement these observations from a single site with meteorological records from nearby, as well as manual photogrammetric measurements at a variety of locations across the ice cap during the course of the ablation season. The authors thus investigate spatial and temporal variations of surface roughness during the ablation season, as well as linkages to surface energy balance. From the automated roughness measurements they find that roughness is temporally variable and highly modified by precipitation, with both rain and snow precipitation leading to a reduction in roughness. From the manual measurements, they find that the seasonal firm/ice transition zone corresponds to the maximal surface roughness at any point, while ice or snow surfaces both exhibit lower surface roughness. The authors also suggest a link to the importance of turbulent fluxes in the whole energy balance.

The target of spatially-extensive surface roughness measurements is a novel development, and useful to understand roughness variations. While the general patterns of seasonal and spatial variability are very likely to be accurate and form a nice story, the authors seem to have some fundamental misunderstandings about surface roughness metrics and their meaning. In addition, the methods are not entirely clear, results are given to an unrealistic and misleading precision (also without any uncertainty assessment), and although the written English is generally correct, the writing style is particularly abrupt. Consequently, although the authors have painted a nice picture of the spatiotemporal evolution of surface roughness at August-one ice cap, the manuscript needs substantial revisions before it should be considered for publication in The Cryospehere.

Major points:

Fundamental misunderstanding of surface roughness. The authors seems to confuse Z_o and topographic surface roughness, which are not the same: while approaches have linked the two, the aerodynamic roughness length is not simply a topographic parameter, and efforts to assess Z_o based on topographic parameters need to be validated with micrometeorological measurements. Furthermore, the authors' effort to produce a grid-based estimate of surface roughness is only applicable for the case of isotropic roughness, which is not the case for ice surfaces.

Lack of clarity with regards to several methods. The authors mention two specific efforts to estimate Z_o from topographic profiles: Lettau (1969) and Munro (1989). It is not clear which is actually used in this study, now how it was applied to the gridded height data. In addition, numerous details of the energy balance model used are missing, while the authors may have accidentally disregarded conduction of heat.

Unrealistic precision, no uncertainty of Z_o estimates or energy balance. The accuracy of Z_o is provided relative to control and check points on photogrammetric frames, and is reported to the tenth of a millimetre. However, it is unlikely that the actual measured positions of their control and check points are known to this accuracy. Furthermore, the surface height models produced by the structure-from-motion processing appear to be oversampled by a factor of 10x in each dimension, relative to the reported point densities. Finally, no assessment of uncertainty has been conducted for the Z_o estimates or the energy balance calculations.

No evidence given of cryoconite, but of red algae. This may be a misunderstanding of some sort, but the authors refer multiple times to the development of cryoconite and its effect on surface roughness, a phenomenon that would certainly explain some of the surface roughness dynamics that they observe. However, the first time cryoconite is mentioned is with regards to Figure 2, but Figure 2 does not provide any evidence (to my eye) of cryoconite – rather, red snow algae is clearly evident. This gives some concern of a basic misinterpretation of results.

Some grammar improvement needed, also some changes to the writing style are needed, as it is not currently suitable for TC.

Detailed comments:

L1-2…during 'the' melting season

L18. Zo was calculated from this data – you need to say how. Manual measurements of what type? Micrometeorology? Profiles of elevation difference?

L37-63. It is apparent from this section that the authors misunderstand several key concepts relating to Zo and turbulent heat transfer more generally; I suggest a careful read of Smith et al (2016) for a review of the differences. First, Z_o may be commonly called 'surface roughness' but its full title is the 'aerodynamic roughness length' (for momentum transfer/heat transfer). In any case, it emerges in the bulk aerodynamic approach as a constant of integration that results from the interaction of the boundary layer with the surface. It is a meteorological term (not a topographical term) that is influenced by both properties of the boundary layer and the surface. One can determine an effective surface roughness 'directly' from eddy covariance measurements (and less directly from wind towers), but it is highly variable in time primarily because the boundary layer is often highly variable. The variability of the boundary layer leads to a different fetch over which the layer is interacting with the surface topography. The microtopographic roughness (which you have calculated) is thus a very good indication of Z_o, but the relationship is not direct or linear, as the energy balance is controlled not just by surface topography at an individual location, but is variably influenced by its surroundings (e.g. Steiner et al, 2019). Thus, it is difficult to trust the values of Z_o produced by this study, as they are not validated by wind tower or eddy covariance observations (which actually resolve Z_o). However, microtopographic roughness metrics are a very strong proxy for Z_o (e.g. Nield et al, 2013), so I have much more confidence in the temporal and spatial variability presented by the authors. However, I think they need to very carefully reframe their introduction to conform with established theory.

L42. Please provide references indicating that microtopographic Z_o is more accurate than wind profile or EC measurements. I don't know how one can claim this, as those methods are the 'ground truth' of Z_o at a site.

L47. 'Direct measurement' is strange nomenclature; microtopographic approaches, including the Lettau (1969) approach, are anything but direct.

L52. Rees and Arnold (2006) is also sensible to mention here.

L55. Other examples of this approach are Rounce et al (2015), Quincey et al (2017), and Miles et al (2017).

L56. The photogrammetric approaches need validation, as the relationship between topographic roughness and aerodynamic roughness length is also affected by local meteorology (Nield et al, 2013).

L74/Figure1. Both panels need a scale. The political map of China is irrelevant to the current study; of more relevance are dominant weather patterns and elevation, including areas outside China's claimed border. Furthermore there is no need to depict the South China Sea, which results in a very poor use of space. What is the polygon within China? It is not identified in the figure or caption. Please provide information about the image of August-one in panel (b) – date, satellite, etc. Red and green are poor choices of color for icons in panel (b), as many people cannot distinguish between these two colors.

L79. Please provide sensor specifications and measurement uncertainties for the AWS.

L81. The sensor measures relative surface height; it does not measure mass balance. Also in L104

L83. There was a windbreak fence installed on the glacier?

L94. How were the positions of the control and check points measured? You report accuracies relative to these positions of less than 1 mm, but I am not convinced that you could locate the control point positions to a higher accuracy than this. Also, how was the frame structure anchored?

L102. Did you choose the daily best-exposed sets of photos manually or automatically? For days with multiple very clear photo sets, was there strong agreement in derived Z_o or a consistent diurnal variation?

L112. Does the August-one ice cap have an accumulation area?

L120. How are the seven pairs of convergent photos arranged? Do you use all 14 photos to produce the DEM and orthoimage? Did you ever carry out the manual photogrammetry at the automatic site?

L124/Figure 4. Panels b and c are switched relative to the text, which led to some confusion about the numbers of check points and control points. I see no evidence of cryoconite in the image, but of red algae which is commonly found on melting snow.

L135. The standard reference for this processing workflow as applied to glaciers is Westoby et al (2012). Also, this approach has already been applied to estimate surface roughness of glacier surfaces: Quincey et al (2017), Miles et al (2017), Steiner et al (2019).

L141-146. This content belongs in the background. Note that Lettau (1969) was the first such effort (of which I am aware). It is also worth noting the extensive review of microtopographic metrics by Nield et al (2013).

L145. Munro (1989) is probably the appropriate first reference here, as is Brock et al (2006).

L161. The method described (based on the standard deviation of detrended elevation) is precisely the Munro (1989) method.

L172-3. Averaging over cardinal directions is only meaningful for surfaces that are isotropic. However, the literature has repeatedly shown that melting ice is strongly anisotropic, as the direction of wind strongly dictates the pattern of melt, and feeds back via roughness. So this 'averaging all cardinal profiles' is entirely unsuited to your study site, unless you can demonstrate that the ice surface is indeed isotropic in terms of roughness, which would be highly surprising.

L174. Some things are not entirely clear to me about your method. First, do you use all profiles in each cardinal direction? Second, it is not clear if you have implemented the exact Lettau approach or the Munro approximation in your 'all profiles' approach. Third, such an implementation (all profiles averaged, for either Lettau or Munro) has already been implemented and tested for a glacier surface. Please see Miles et al, (2017).

L179-181. The surface energy balance presented is not quite accurate for a 'melting' glacier, but for a 'temperate' glacier. Do you have any evidence that the August-one ice cap is temperate? If not, there also needs to be a term for heat conduction.

L191. My impression is that you use your calculated Z_o value for the bulk aerodynamic approach. How do you integrate your 3-hourly (half-day) Z_o values with your model? At what timescale is the model run? What uncertainty does the input meteorology have, and what uncertainty does this produce for your results?

L204. Is an environmental lapse rate entirely appropriate for this site? Do you have lapse rate measurements?

L205. How confident are you that the AWS measurements are broadly representative of the entire ice cap? Do you have evidence to back up this claim?

L210-211. I am not sure how you get seventeen (17), as you have 4 control points and 3 check points. Similarly, I do not understand what the 31 manual photography pairs are – please explain.

L210-219. This entire section is an amalgamation of bullet points; please rewrite to conform to style for The Cryosphere.

L212 and L216. The reported point densities do not justify a resolution of 0.1mm, but of 1mm. These DEMs are 100x oversampled.

L213. The average georeferenced error is greater than 1mm for half of the control points, and nearly all check points. However, I am also not certain how precisely you could have measured the location of the control and check points. Please provide details and uncertainty.

L225. Yes, but part of this is also the difference of your survey design. For the automatic measurements, the camera is moving linearly, and the density of tie-points is much higher in the foreground compared to the background. For the manual method, although the survey design is not clear, more photos were taken and I presume that they surrounded the target area. This type of survey would be expected to provide a much more robust elevation model.

L228. Rees and Arnold (2006) did indeed suggest millimetre vertical accuracy. The also suggest a fetch length of 3-6 meters as relevant for the majority of energy balance situations, which is considerably larger than your domain.

L230/Figure3. It is not clear what this chart shows – the y axis is labelled 'Differences', but is this RMSE, MAD, or…? Please clarify.

L231-4/Figure 4. Same problem and Figure 3. Should be merged with Figure 3 as a second panel.

L238. No description of profile analysis is included in the manuscript, only of a DEM analysis. Please provide more detail.

L239. Do you have an estimate of the uncertainty of these Z_o values?

L242-254. Listing a narrative as bullet points in the results is not particularly aesthetic, and this section should be rewritten as a paragraph. More importantly, this section mixes results and interpretations. Please present the observations, then interpret them.

L254/Figure 5. Z_o values are more commonly presented on a logarithmic scale, as even a factor of 2 makes little difference in the turbulent fluxes, whereas a factor of 10 can be a considerable difference regardless of value. This is, in part, due to the bulk aerodynamic approach. Also, it would be very nice to include a set of panels depicting the surface at different parts of this record (high and low values, for example).

L258. One order of magnitude is not a particularly large variation of Z_o.

L260. I have not yet seen evidence of cryoconite holes; the image in Figure 2 is unconvincing. Also applies to L280

L263,276,277. I see no need to include p-values here.

L274. Was there no accumulation in this year?

L283/Figure 6. Is there a reason that the lines are shown with different styles? For comparison, it would be good for all 4 panels to have the same y-axis limits.

L288-302. Somewhere in this section there should be a reference to Figure 7.

L310,324,353,339,345. The use of sub-headings here just breaks up the text.

L320/Figure 7. In panel (a), please use a logarithmic scale for Z_o. Is panel (c) showing net solar radiation, or downwelling – not specified. The y-axis upper limit in panel (d) should be 100%. In general, all time-series look smoothed. Please provide details of exactly what is shown. In the caption, please be sure to provide the year.

L330/Figure 8. What is the uncertainty of each of these values quantities?

L335. If latent heat and sensible heat account for so little of the energy balance, how much impact does a variation of Z_o from 0.25 mm to 2.5 mm have on the total energy balance?

L343. The 'visible smoothing' is not clear to me from Figure 7. Please explain where you see this.

L349-350. As turbulent fluxes matter very little for your energy balance, the match is not due to the calculated Z_o.

L358-380/Fig10 and 11. I do not think this analysis is very well grounded in theory. First of all, as the turbulent fluxes depend on Z_o, you are comparing a quantity to a modified version of itself in Figure 10d and Figure 11. In fact, this exactly corresponds to the shape of the fit in bulk aerodynamic theory (which you have used to relate Z_o to the turbulent fluxes). So on one hand, none of this section is unexpected, but nor does it provide any novel insight. On the other hand, if you intend to examine the potential feedbacks between energy balance and surface roughness, that would be very interesting, but would require the use of a lagged correlation (in which case your variables would be independent).

L389. Again, Westoby et al (2012) is probably an even more appropriate reference here.

L391. I disagree with this because your survey setup is entirely different for the manual and automatic methods. See my comment with regards to L225.

L400. I believe you are referring to the glacier terminus. Please replace ‘terminal’ with ‘terminus’ throughout the manuscript.

L403. This is the very interesting result of your study: following the zone of maximum roughness as it migrates upglacier. But a key question is how important are turbulent fluxes in this zone? Perhaps they are relatively unimportant everywhere else, but in this transition zone you have maximum Z_o and the zone also migrates across much of the glacier, highlighting the importance of transient surface characteristics.

L429. Please be careful and consistent with the terminology that you use. In this study you have examined topographic roughness and the aerodynamic roughness length (which are not quite the same thing, see Smith et al, 2016).

L431. I do not think this is a meaningful result, see my comment on L358-380. This also applies to L439.

L434. What do you mean by ‘heavy-loading glacier’? I have not heard the term before.

L437. The link between cryoconite holes and surface roughness is indeed important, and you should make this link explicit earlier. However, your manuscript has not presented any clear evidence of the cryoconite development process occurring at your site.

L440. I do not understand what you are referring to here, with regards to quantitative vs qualitative research. Please explain more clearly what you are implying.

L456. What type of studies? Please make some concrete suggestions; at present this discussion and conclusion makes very little contribution to the field.

L470 and L472. Duplicate reference.

References mentioned in the review

Miles et al. (2017). Highly variable aerodynamic roughness length (z0) for a hummocky debris-covered glacier. *Journal of Geophysical Research: Atmospheres*

Nield et al. (2013). Estimating aerodynamic roughness over complex surface terrain. *Journal of Geophysical Research: Atmospheres*

Quincey et al., (2016). Evaluating morphological estimates of the aerodynamic roughness of debris-covered glacier ice. *Earth Surface Processes and Landforms*

Rounce et al., (2015). Debris-covered glacier energy balance model for Imja-Lhotse Shar Glacier in the Everest region of Nepal. *The Cryosphere*

Steiner et al., (2018). The Importance of Turbulent Fluxes in the Surface Energy Balance of a Debris-Covered Glacier in the Himalayas. *Frontiers in Earth Science*

Westoby et al., (2012). ‘Structure-from-motion’ photogrammetry : A low-cost, effective tool for geoscience applications. *Geomorphology*