Response to 'Comment on esurf-2021-18' from Jan Henrik Blöthe

We thank Dr. Blöthe for his very constructive and helpful comments which will clearly help to improve our manuscript by pointing out weaknesses, but also suggesting the respective improvements which we will be necessary to be included for the next version of the manuscript. We try our best to provide as many details as possible to better reply to the reviewer. Please find all the details below.

Below are our responses to the reviewer's comments, with their initial comments in black, our responses in blue, and quotes from the manuscript italicized.

[1] Overall, the manuscript is well written, but sadly lacks rigorous identification of previous work and the work conducted here. Especially in chapter 3 (Methods and data), the authors need to put more effort into making it very clear, which information has been obtained from earlier studies and how the data produced in the framework of this manuscript was produced.

We thank you for pointing out the disadvantage of our manuscript and we will make it clear in the next version.

[2] A very basic and important point that the authors fail to include in their work is a thorough error assessment of all measurements they conducted. This is neither addressed in the methods section, where it remains rather unclear how exactly mapping has been conducted, nor are the errors associated with mapping (based on visual interpretation?) in remotely sensed imagery discussed later in the text. For

Sorry, we did not calculate the error on the visual interpretation and bare ground automatic extraction. Such errors are usually estimated by ±0.5 pixel uncertainty along the boundary shape, we will add the error estimation in the next version of the manuscript.

[3] Chapter 3.2: Here the authors mainly present details of earlier studies that quantified glacier mass balances changes of the HLG. I would recommend to include these background information into the description of the study area (Chapter 2) and shift the focus of Chapter 3.2 to the method applied here.

Thank you for your advice, we will move that background information to Chapter 2 or others.

[4] From the technical description in L158-164, I take that the authors used three digital elevation models (DEMs), namely "the TopoDEM (1966), the Shuttle Radar Topography Mission (SRTM) 30m DEM (2000) and the ASTER-DEM difference between 2000 and 2016". While I am familiar with the latter two (the ASTER DEM data is from Brun et al. 2017, include reference here), I am not aware of the TopoDEM (1966). If this refers to the DEMs calculated in Cao et al. (2019), this needs to be indicated here. Moreover, from the technical details outlined here, it remains unclear to me whether the authors only analyzed five profile lines, or calculated full DEMs of difference and used five profile lines to visualize the data. Please elaborate this in more detail and outline, in the case data was only analyzed along five profile lines, how these are encompassing the full variability of surface changes on the glacier tongue.

TopoDEM (1966) is not calculated by Cao et al. (2019), but has been reported by Liu et al. (2010)

and Zhang et al. (2010). We calculated full DEMs and used five profile lines to visualize the data. We will add more details to make this information clearer.

These profile lines were selected to try to cover every type of slope and keep a certain distance between the transverse profiles so that the extracted results are better for comparison. To present more surface changes information, we will add the dh/dt maps in the supplementary files in the revised manuscript.

[5] Furthermore, in L165-174 (still Chapter 3.2), the results of earlier studies that quantified ice flow dynamics are presented. In the final sentence the authors describe a comparison of "long-term ice flow velocity changes […], based on three-periods results of 1982-1983 (in situ observed), 2007-2011 and 2014-2018 (SAR satellite derived)." Where does this data come from? Is this an analysis done by the authors, or does it refer to the studies presented before? As the dates mentioned here do not match the time spans given for the studies cited above, I would ask the authors to either clearly indicate the provenance of these data.

The *1982-1983 (in situ observed)* data are from L165-168 '*the earliest observations of surface velocity of the HLG were during 1982-1991*'. Here we used the data from the earliest period (1982-1983).

The *2007-2011 and 2014-2018 (SAR satellite derived)* data are from L170-171 '*acquired by PALSAR-1/2 satellites from 2007 to 2018, Liu et al. (2019a)*'. Here we used the data from the two period (2007-2011 and 2014-2018).

We have marked the source of the data in parentheses, if this is still not clear enough, we will consider adding a scanned picture of the 1982-1983 published velocity map and the gridded velocity fields shared by Liu et al. (2019a) in the supplementary files in the revised manuscript.

[6] Chapter 3.3: This is a very brief description of how slope movement was quantified, given that large parts of the results and discussion build upon this data. While manual tracking of tie points in repeated imagery can be considered a fairly robust technique, I would recommend that the authors at least try to quantify the error associated with this tracking of tie points. This can be easily achieved by tracking stable surfaces in the vicinity of the slope failures. Furthermore, may I suggest to include the individual vectors for manually tracked tie points in Fig. S2? Last but not least, let me point out that there are multiple software solutions that allow for automated tracking in consecutive imagery, which would result in full 2D velocity fields that might enable a much deeper insight into the mechanisms of the failures (and reveal local variability).

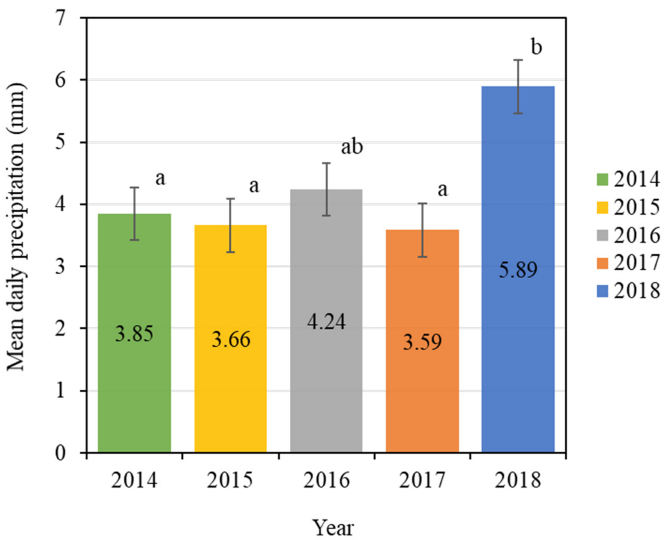
This is a good suggestion, and we will try to quantify the error in this part in the next version of the manuscript. We will consider including the individual vectors for manually tracked tie points in Fig. S2 if necessary.

Indeed, there are much software that can automatically track in consecutive imagery, and we have tried it. However, because the vegetation on the Type B slope is well-established and fewer tie points can be identified, the automatic tracking effect is not good, we therefore chose manual tracking with higher accuracy.

[7] Section 5.1.1: Here the authors discuss the possible preparatory and triggering factors for the

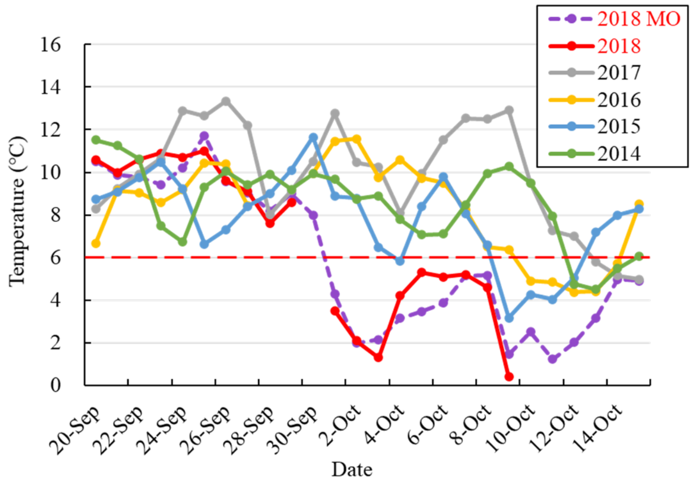
rock fall they observe during their study period. The authors might want to further elaborate how the exceedance of a precipitation threshold of 60 mm for a single day in June 2018 is connected to the triggering of a rock fall in October 2018. This is mainly referring to L362-65 and Tab. 3, where the authors argue for a precipitation intensity anomaly, which I find hard to follow given the data. Yes, 2018 has seen one day with more than 60 mm of daily rainfall, i.e. 61 mm (L364). Without giving more details on the exact precipitation values for 2016 that also saw three days with more than 40 mm per day, I do not think this can be seen as an indication for a precipitation intensity anomaly. In my view, it would be worth to look at the antecedent rainfall in the five to ten days before the failure and compare antecedent rainfall statistics between years, especially in a setting with very few days without rainfall. Furthermore, in section 5.1.1 the authors argue for a potential triggering by frost action. It is my feeling that also this remains speculative, as the cold interval the authors refer to here (01-09 October 2018; L369, Fig. S4) happened at least 5 days before the failure that the authors date to 15 October 2018. However, the temperature data for 2018 is unavailable for the days following 09 October, making this link questionable.

Thank you for your comment! We can understand your opinion that the precipitation in 2018 is not considered abnormal. We will revise the description of this part in the new version of the manuscript. And we will revise the study time to five to ten days before the rockfall occurred (**5-20 Oct 2018**) and will add the below figure in the manuscript.



The mean daily precipitation between **5 and 20 Oct 2018** is 5.89 mm which is significantly higher than the 2014, 2015 and 2017, although not significant compared with 2016, the number is still high.

As for the temperature data after 9th October 2018, we obtained and checked the daily temperature data of four periods (2 am, 8 am, 2 pm, 8 pm) by manually observation at the 3000m observation station at Mt. Gongga, and calculated and plotted the daily mean temperature curve, which has been added to Fig. S4 (the purple dash line in the below figure). We will revise it in the manuscript.



Note: MO means manual observation.

[8] L35-40: Here the authors gather five modes of response to slope failure, though I have the feeling that mode 5 "paraglacial debris cones and valley fills" is rather the results, i.e. deposit of the processes listed in 1-4.

We thank you for pointing out this error and we will correct it in the next version.

[9] L67-68: Not really relevant at this point, as the tourism activity at the site is detailed in L109-117.

We will remove it.

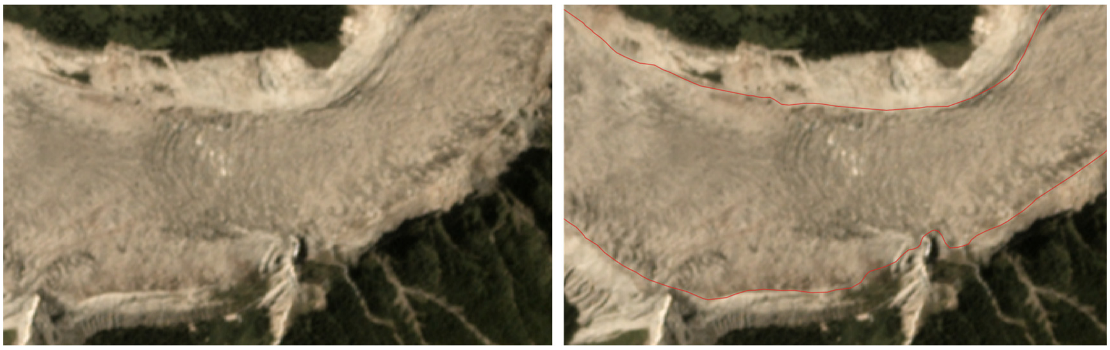
[10] L101-02: In L152-53, this information is from Zhang et al. 2010, please add reference here

Thank you, we will add it.

[11] L124: How was the glacial area mapped exactly? As the HLG is debris-covered, a precise delineation between debris-covered ice and the debris-covered surroundings is not trivial. Please elaborate in detail, how mapping of glacier extent was conducted and what the associated uncertainty of this mapping was.

Due to the complex terrain of HLG ice surface-numerous ice cliffs and crevasses- as well as the high slope on both sides of the glacier, it is very difficult to use GPS to measure the glacier boundary on the ground. Therefore, at present, the use of medium- and high-resolution remote sensing images has become the most effective method of glacier boundary extraction. Although the boundary of the debris covered glacier is not as distinct as the clean glacier, it is still visible. As the mentioned above comment [2] the errors are usually estimated by ±0.5 pixels.

The following figure takes PL image as an example:



[12] L124-26: Also for the paraglacial slope failures (PSFs), it is not clear how exactly these were mapped. In L121-24 it is described that using the NDVI, vegetation covered areas were excluded. But how exactly was the mapping of the PSFs achieved in remote sensing imagery. What were the criteria for mapping? Was this mapping field-evidence based and if so, when was field-work conducted?

[13] L130-34: Again, it remains unclear at this point how PSF boundaries were extracted, what validation means in this respect and based on which criteria manual correction was done.

As mentioned before, the current method does have the disadvantages you indicate, but it can delineate the expansion of exposed slopes in most area. We will fully discuss the uncertainty of the delineation boundary in the next version of the manuscript. Manual correction of PSF is mainly based on the visual interpretation of the RS image, eliminating the pattern and obvious mis-extraction (relying on empirical judgment).

[14] L138: I would suggest to state this in more detail. At this point, it is unclear, whether the "mean quality of 0.01 m" refers to the position accuracy of the RTK UAV, or to the SFM output. Furthermore, "+1 ppm (RMS) in XY" is neither clear in this regard. In order to allow the reader to follow and to judge the quality of the data used in this study, I suggest to explain in detail, how the UAV images were processed and what the residual mismatch of their geolocation is.

The GNSS quality information are required from DJI official website (https://www.dji.com/uk/phantom-4-rtk/info#specs)

Horizontal 1 cm + 1 ppm （RMS）
1 ppm means the error has a 1mm increase for every 1 km of movement from the aircraft (we will add this sentence in the next version)

UAV measurement is a very mature technology, albeit a new one; we will describe our UAV mappings in the revised manuscript with some more details.

[15] L142-43: I take it that the authors co-registered and orthorectified the UAV-derived orthomosaics with the PALSAR DEM, or is the sentence correct that individual UAV images were used for this? May I ask the authors to elaborate this a bit more, as this is cionfucing? IN line XX you write that these are already orthorectified?

We use PL images and ALOS PALSAR DEM to co-register and orthorectify the UAV images which are synthesized by ContextCapture Center Master Software based on UAV photographs. Due to the steep slope of HLG, DEM correction is required for the synthesized images. We will add more detail in this sentence, thank you.

[16] L156: The authors compiled a data set here, but I fail to see where this data set is included in the manuscript? Is there a figure or table that shows this compilation?

We used the dataset in Figure 3 for ice thinning (elevation change) analysis. Both L156-160 illustrate the data used in Figure 3, and we will add some details to make it clear.

[17] L174: In the very brief section following this heading, the "outline change rate" is not mentioned nor explained how this is quantified. Instead, the authors quantify the rate of headscarp erosion. Consider rephrasing the heading here.

Ok, thanks for mention it, we will change the heading to "Slope movement and headscarp erosion rate".

[18] L178-180: May I suggest to elaborate more clearly how the outlines of failures were used to calculate a mean annual retreat rate for headscarps?

We calculated the mean distance between the two outlines where the headscarp erosion occurred at the head of the slope. It can be obtained by using the Average Nearest Neighbor tool in ArcGIS. We will add this sentence in the next version.

[19] L220-23: The time spans given here in the text are not the same as in Fig. 4. What is the rate between 2000 and 2019?

What I want to express here is that the growth rate of EBGA has increased from ~0.01 km^2 a^{-1} at the beginning (1990-2000) to 0.1 km^2 a^{-1} now (2019-2020).

We will change the time period to every ten years for comparison.[20] L224: Which area are you referring to here?

Sorry to confuse you, but what I want to say here is the area of EBGA.

[21] L225-26: Where is the data for the lower frequency of PSFs?

Sorry, we will add it in the next version.

[22] L227: Where does the knowledge of slope material come from? Did you map out slope material distribution during field work?

We will remove this. The classification of the three typical styles of paraglacial slope failures are mostly based on geomorphological characters. And sorry, we didn't map out the slope material distribution.

[23] L239-44: This description of the rock fall (PSF type A) needs to be refined. In L240 it is stated that the rock fall occurred on a south-west facing slope, while in L244 it is stated that the mass detached from a steep north-facing slope? Is it that the general topography is south-facing and the nice of the detachment faces north? The picture in Fig. 5b, however, does look like the source is also facing the glacier – please clarify.

Sorry, we will revise it. It this paper, we uniformly use the south- and north-facing slope to express the position of the slope.

[24] L241-42: This is not precise enough. How can a deposit of a rock fall be “450m in height fro the glacier surface” and at the same time, “cover a height of 380 m”?

We will revise the “deposit” to “rockfall area” or others.

We will change the sentence to “*The rockfall area is 283 m in width, 200 m in length, covers a 2D area of ~47,000 m^2, with a vertical height of 380 m and a slope length of 472 m.*”

[25] L258-59: Please indicate what magnitude is considered small here, as to me “each with a mean area of 750 m2” is not clear.

All the small rockfalls found were all smaller than 4,000 m^2, with a mean area of 750 m^2, much smaller than the large rockfalls (47,000 m^2) that occurred in 2018. We will describe it more clear in the revision.

[26] L262-63: Please try to be consistent in labelling the processes: “Sediment-mantled slopes slide and collapse” vs. “Sediment-mantled slope slide and collapse”. In my view, both seem a bit clumsy – you might want to rephrase.

Thank you for pointing out that, we will be consistent in labelling the processes: “*Sediment-mantled slopes slide and collapse*”. “*Sediment-mantled slopes”* is a term in paraglacial geomorphology, please see “Colin K. Ballantyne, Paraglacial geomorphology, Quaternary Science Reviews, Volume 21, Issues 18–19, 2002, https://doi.org/10.1016/S0277-3791(02)00005-7”

[27] L263-66: The numbers given here are surprisingly round and do not match the sums of the individual numbers mentioned in Tab. 2. Also, in Tab. 2, there are no errors associated with the numbers given for the area. Please include a statement on the precision of these estimates in the text. This also applies to L310.

The four Type B slope areas in Table 2 add up to 371673m^2, which is approximately 370,000 in L264. We will add "~" in the next version.

As for the number “297,000 m^2” in L265 is the total increase area from 1990 to 2020 calculated though satellite images, which is not the one calculated from UAV imagery (2016-2019).

We will add more details about the use of the data in Table 2 so that audiences can understand. Thank you for pointing out.

[28] L271-75: This comes as a surprise here and rather belongs to the discussion.

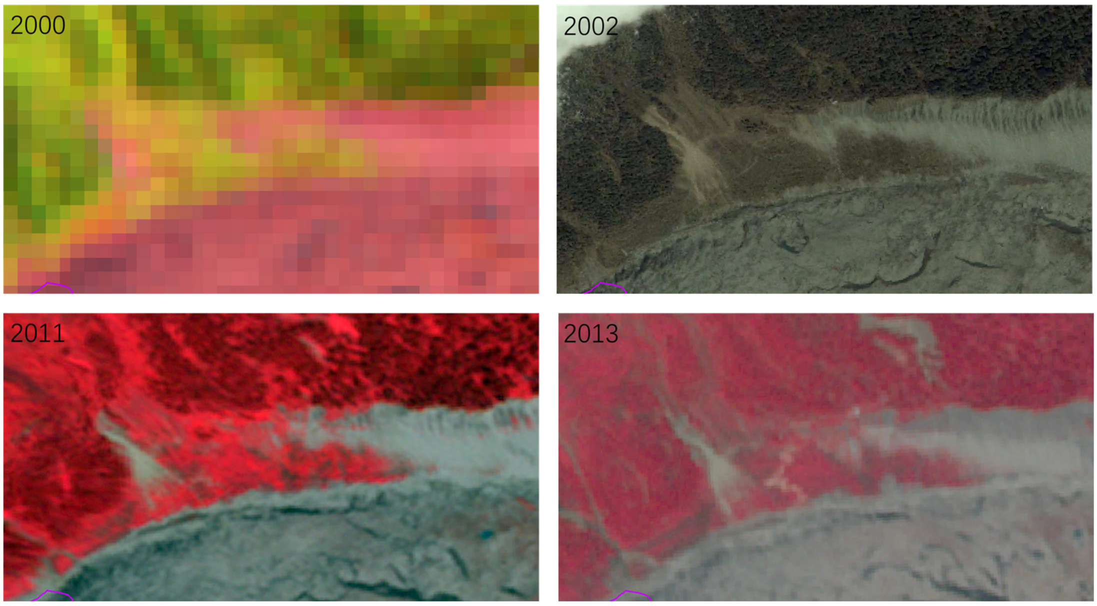
Thanks for mention, we will correct it.

[29] L276-78: How was the error of 0.04 cm d-1 quantified?

From L149 “*mean quality of 0.15 m in XY*”, error automatically calculated by ground control points during the UAV registration procedure.

[30] L283: How do the authors know that the detectable slope movement began around 2000? Has this been published, or did the authors run additional analysis beyond the 2016-2019 UAV surveys that have been described in the text. Same applies for L296.

According to the historical remote sensing images (2000 TM, 2002 Google earth, 2011, 2013 RapidEye, please see the below figure), the deformation of the slope began to appear in 2000. Same applies for L296.



[31] L286: “the landslide has fallen” sounds as if it was a vertical movement that was quantified? Until now, it is my understanding that the authors quantified 2D horizontal displacement.

Yes, the landslide was quantified only in 2D horizontal directions. The slide rate could be estimated based on the surface slope of the terrain. Thanks for mention, we will correct the wording.

[32] L400-02: This comes as a surprise as a) in L361-62 the authors argue that 2018 has seen relatively low temperature. Furthermore, Fig. 9 suggests that 2017 has a data gap in temperature readings.

Thanks for mention, we will correct it.

[33] L430-38: I am not sure how this is connected to the data and topic presented in this study? Consider removing paragraph or elaborate the connection to paraglacial slope adjustment.

Thank you for your suggestion! We decided to add more detail to make the link between glacier velocity and hillslope processes clearer.

The slowing of the glacier velocity and the glacier thinning corroborate each other. On the one hand, the thinning of the glacier slows down the glacier velocity; on the other hand, when the glacier velocity slows down, the ice flux transported from upstream to downstream is reduced, which accelerates the glacier thinning and ultimately leads to the slope sliding.

[34] L444-46: Is this supposed to be a general statement? Also, check grammar.

Thanks for mention, we will correct it.

[35] L448-50: It is hard to see how the deposition of debris onto the glacier is directly affecting climate. As the authors try to outline in Figure 8, debris cover generated by slope failure might have

an effect on glacier downwasting, which in turn can have a tiny effect on the climate.

Yes, as show in Figure 8, debris cover generated by slope failure might have an indirectly effect the climate.

L456: We will change the sentence to "…which in turn affected the surface energy balance for melting and therefore downwasting rate of the glacier, which will thus influence the rate of runoff generation and its contribution to the sea level rise from glacierised catchments."

[36] L464-70: While McColl and Davies (2013) showed that also failures of similar magnitude as B2 in the present study can deform glacier-ice at the rate of mm/yr (assuming a min. average thickness of ~1 m), the evidence presented here is limited and the discussion of this aspect remains too surficial. It might be a way forward, to present the displacement data in more detail and quantify the supposed narrowing and squeezing effect that is mentioned here to back this aspect with data.

As reply in comment [6] above, we will include the individual vectors for manually tracked tie points in Fig.S2 according to your suggestion.

[37] L474-75: What data is this statement based on? Is there a study that found this increase in debris-flows and flash-floods that could be cited here?

Yes, there is a study can be cited here: "Lu, R., and Gao, S. (1992). Debris flow in the ice tongue area of Hailuogou Glacier on the Eastern slope of Mt. Gongga. Journal of Glaciology and Geocryology, 73-80". However, there is no data and literature to support the increase in frequency of debris-flows and flashfloods, we will remove "*increased frequency of*", and cite the literature in the text, thanks for mention.

[38] L482: In L195-96 the authors state that between 2016-2019 the glacier terminus retreated by >150 m with a rate of ~52 m yr-1. Which is true?

I'm sorry, it's 450 m (15 m a^{-1}) in L482. We will check all the numbers in the next version.

[40] L485-86: In your data, type A landslides are limited to one rock fall. Did you find evidence for debris avalanches as well?

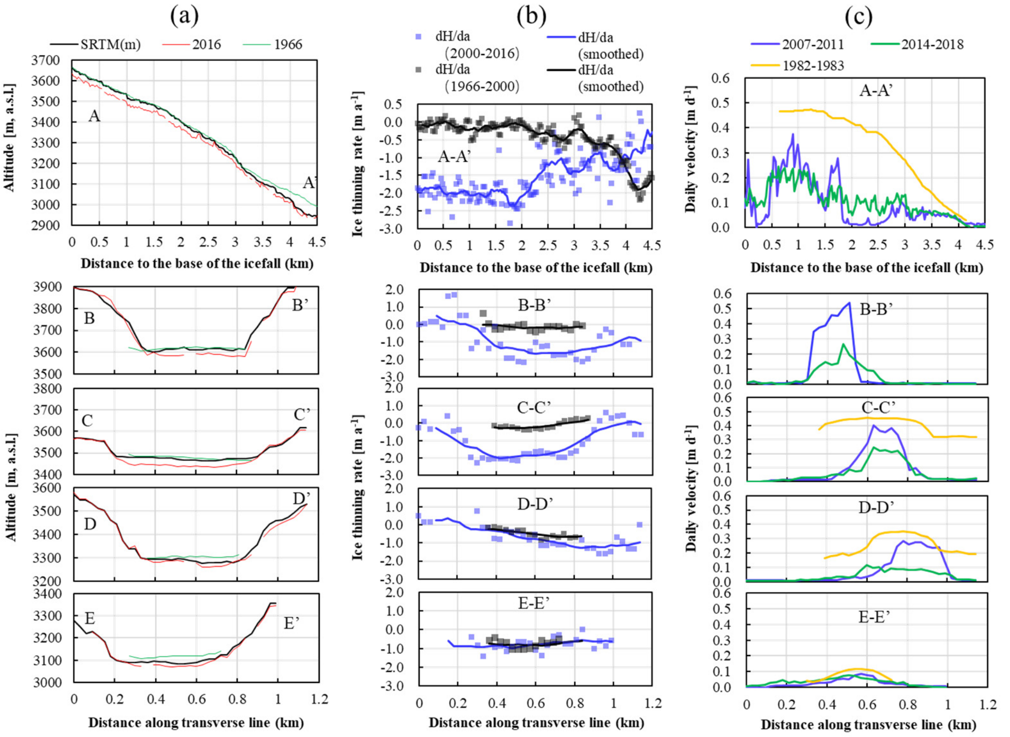
No, we didn't find any debris avalanches.

[41] Figure 1: Might I suggest to add the glacier outline for the years 1982/83, as in the text the authors have quantified (or cited) the displacement values for this period (Fig. 3).

Ok, we will add it.

[42] Figure 3: Is there a reason for the black line being thinner in D-D` of the first column? Also, in the caption, replace annual thinning rate with average annual thinning rate, as these have been quantified over multiple years, right? What remains unclear is the provenance of the data shown in the right column. Has this data been calculated from remote sensing data by the authors? If not, add the references to the data to the caption.

Thank you for your careful review. We have corrected the D-D 'line as shown in the following picture and will correct the caption and add the reference.



[43] Figure 4: Change "(white line)" to "(blue line)" in the caption, as this is showing the glacier outline, judging from the legend.

Thanks for mention, we will correct it.

[44] Figure 6: Might I suggest to label the vertical axis of the last column with "displacement velocity"? Slide velocity implies a vertical component, but these are horizontal displacement values, right? Furthermore, in the caption, what does the reference to Qiao Liu et al. imply? Has this data been published before? I cannot find this reference in the reference list.

Thanks for mention, we will correct the vertical axis to "displacement velocity". This data has not been published, it was requested by the ESD journal to include the source in the maps and photos. The UAV images were taken by Qiao (Corresponding author of this manuscript) et al and therefore need to be noted here.

[45] Figure 8: May I suggest to give this figure a more detailed caption that explains the arrows and the reciprocal effects? Also, why does Type C have a larger arrow than Type A and B, what exactly does "remarkable glacial debuttressing" imply? Also check grammar of text box PSF Type A.

Ok, we will add more detail on the caption. The three arrows of type A-C on the right end are just for the sake of aesthetics, nothing else.

[46] Figure 9: Here it should be indicated over which time window the moving average for temperature and precipitation has been calculated. Furthermore, the reason for and length of the temperature data gap in 2017 should be mentioned in the caption. Also, are Tem_smooth and Pre_smooth appropriate abbreviations for a running mean of these variables?

Thanks for mention, we will correct it.