University of Potsdam | Institute of Environmental Science and Geography
Karl-Liebknecht-Str. 24-25 | 14476 Potsdam-Golm | GERMANY

Etienne Berthier

Editor at *The Cryosphere*

via email

Faculty of Science
Institute of Environmental Science and Geography

Monika Pfau
email: monika.pfau@uni-potsdam.de
Date: 2023-07-19

Dear Dr. Berthier,
Dear Editors,

thank you for giving us the opportunity to further revise our manuscript tc-2022-194 entitled **"Cast shadows reveal changes in glacier surface elevation"**. We addressed all comments and questions from the editor and the two reviewers in a point-by-point reply letter, and are happy to present a revised version of our manuscript.
In our reply letter, comments from the editor (EC) and reviewers (RC1 and RC2) are blue, and our answers are in black font. References to specific lines refer to the revised manuscript.

We look forward to hearing from your decision in due course.

Yours sincerely,

Monika Pfau
with co-authors Georg Veh and Wolfgang Schwanghart

From the editorial board, we received the following comments:

EC1: Both reviewers recognized the large amount of work made to improve the manuscript. One is now fully satisfied with the revised text whereas the other reviewer still challenge the use of View Finder Panorama as the best DEM for the Baltoro case study. A sensitivity analysis, describing the difference of dh on Baltoro when using SRTM1 instead of VFP (as nicely down for Aletsch) should help to reconcile both views and strengthen further your study. (I wondered whether you and the reviewer used the same version of SRTM1. If you want me to ask his version of SRTM1 do not hesitate to contact me directly by email.)

EC1A1: We acknowledge that the use of the DEM from Viewfinder Panorama (VFP) has attracted considerable criticism from one of the reviewers, fuelling a debate that we had not anticipated. Indeed, it is difficult to show whether the void-filled SRTM or the VFP better represents Mitre Peak near Baltoro Glacier because the interpolation method remains unknown in both cases. However, the reviewer's assessment was based solely on a visual assessment of a hillshade, and the conclusions drawn remain subjective and conjectural. For example, the reviewer argued that *"in view of this, it seems obvious that even if underlying filling data are unknown, there SRTM 1" can be seen as a better candidate for this study which weakens the rebuttal. The assessment of peak elevation is clearly affected by large uncertainty and unconvincing I believe to conclude."* This assessment is free from any quantitative support that SRTM-1 has a better interpolation algorithm or smaller uncertainties than VFP. We had clearly pointed out in the manuscript and in the reply letter that there is no independent high-resolution DEM for this peak to quantify possible offsets between VFP or SRTM from a reference surface. The 8-m HMA-DEM also features extensive voids at Mitre Peak.
The difference between VFP and SRTM suggests that the peak in VFP is 154 m higher than in SRTM. A profile drawn across Mitre Peak in both DEMs shows that the elevation in steep

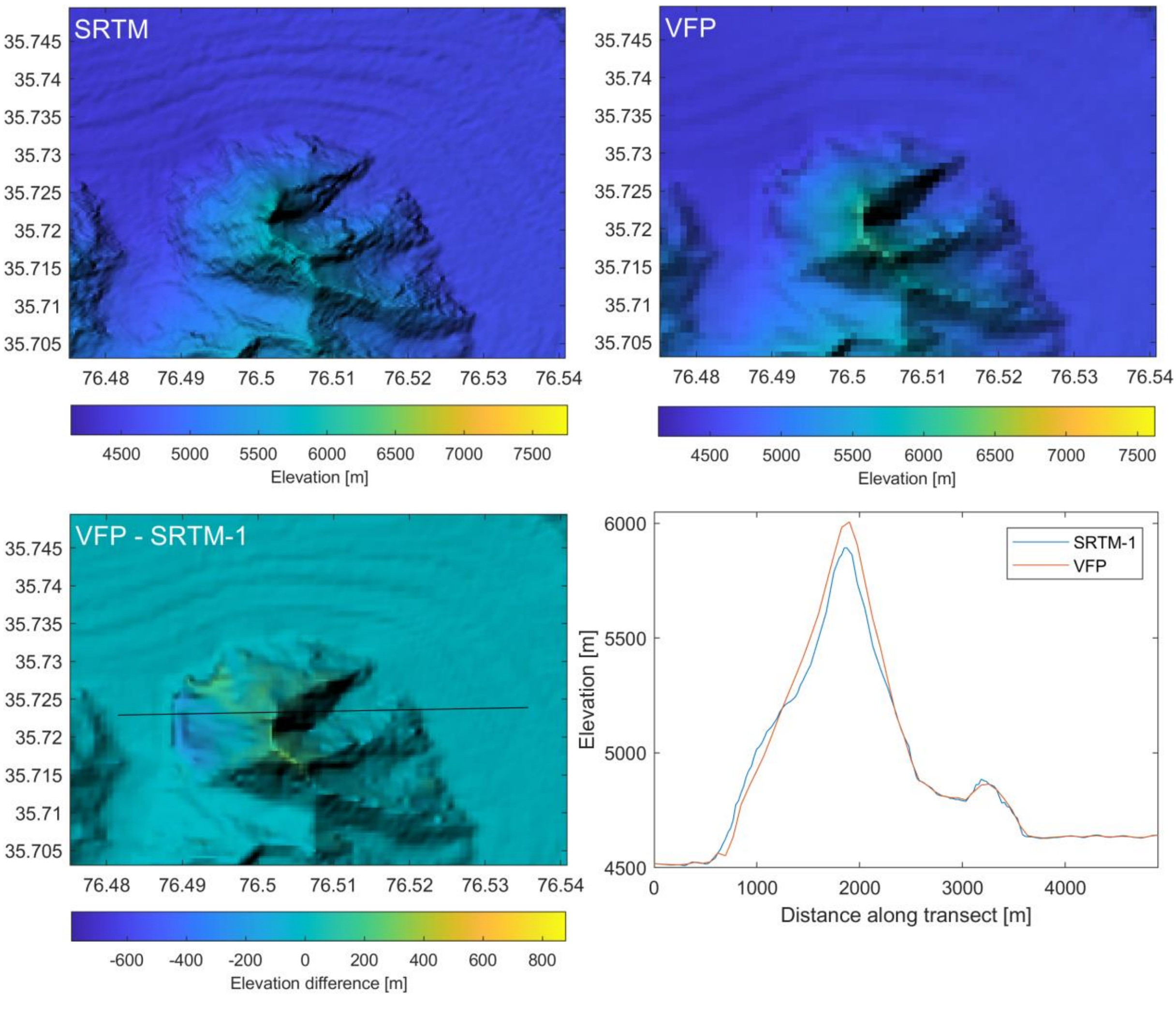


Figure 1: Comparison of elevation from two DEMs at Mitre Peak, adjacent to Baltoro Glacier, Pakistan. *Upper panels show color-coded elevation values at Mitre Peak in SRTM and VFP DEM draped over a hillshade. Lower panels show that the difference between the two DEMs can be several hundreds of meters higher. The black line in the lower left panel shows the location of the transect in the lower right panel.*

topography is higher in VFP, while elevations on the flat glacier surface are identical (Figure 1). Geoid heights (differences between the geoid and ellipsoidal elevations) are ~22 m in this region. Thus, we exclude the possibility that both DEMs have different vertical datums. We thus conclude that only the shape and height of Mitre Peak can cause differences in modelled shadows.

In Figure 2, we compare the shadows from SRTM, acquired in 2000, and VFP. Accordingly,

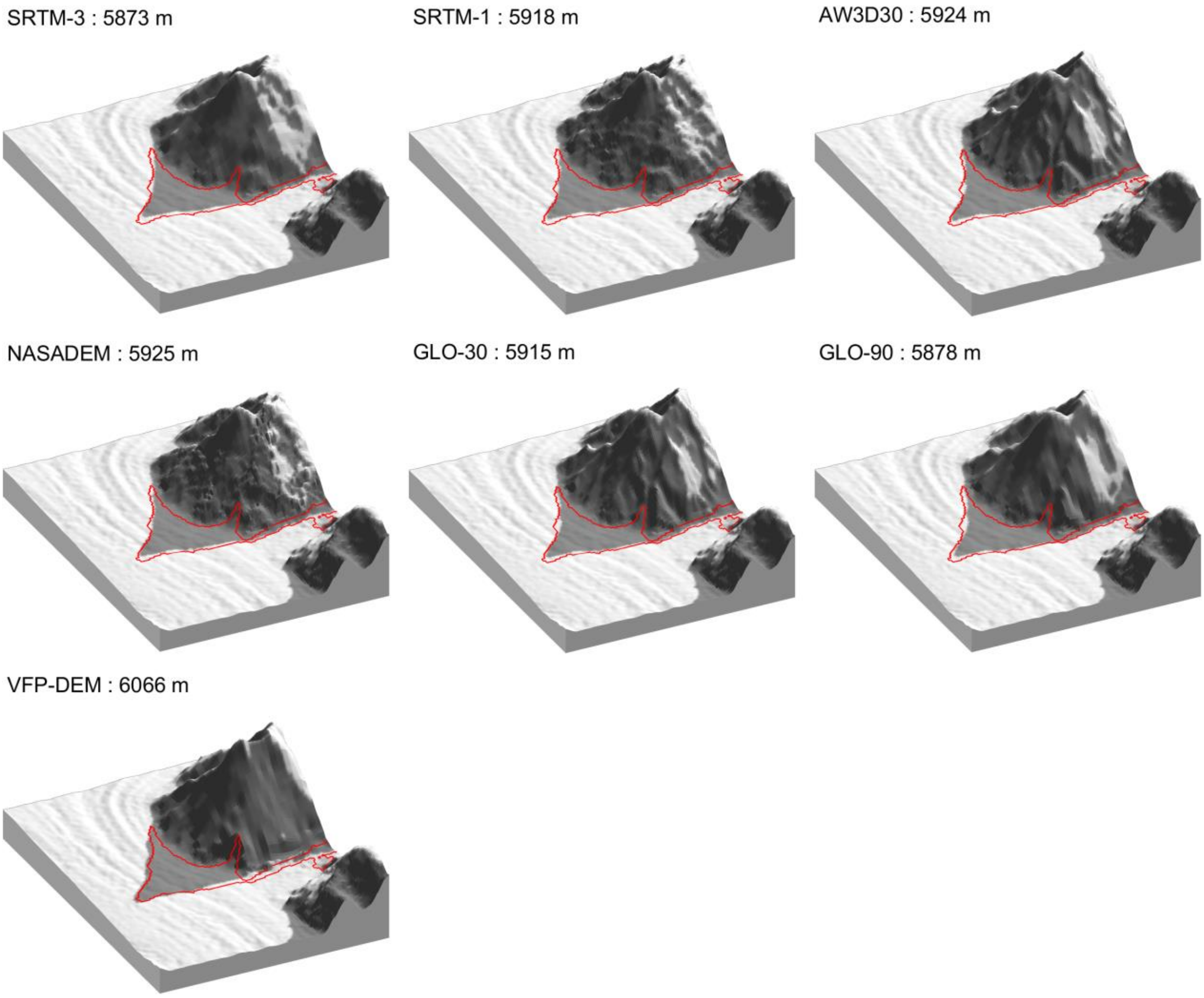


Figure 2: Hillshades including modelled shadows of the Mitre Peak adjacent to Baltoro Glacier, Pakistan. *We used the SRTMGL1 and replaced the Mitre Peak with different DEMs. Shadows were calculated with an azimuth of 151.9° and a sun elevation angle of 29.5°. These values refer to the sun position during the acquisition time (Jan 24, 2000) of the Landsat image from which shadows were mapped (red outline). Visual comparison shows that the SRTM+VFP creates the best match between modelled and actual cast shadows, whereas there are pronounced offsets between actual shadows and those derived from other DEMs.*

the modelled shadow from VFP is slightly longer than the mapped shadow from Landsat imagery, while that from SRTM is too short (Figure 2), consistent with the differences in peak elevations. VFP is based on SRTM, but our comparison clearly shows that the VFP of the Mitre Peak has not been calculated by simple void filling of SRTM-3 data, but we rather speculate that other data were taken into account (e.g. Russian topographic maps as in the Hispar Muztagh, Karakorum, http://viewfinderpanoramas.org/elevmisquotes.html#asia).

We conclude that both DEMs are suitable for approximating the geometry of the shadow at Mitre Peak, but both have their limitations. This is not surprising given the steep topography of Mitre Peak, whose gradient is difficult to represent in DEMs with 30 or 90 m cell resolution (Figure 3).

Figure 3: Steep topography at Mitre Peak. Photograph by By Anne Dirkse (www.annedirkse.com) - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=34911367

EC2: I also agree with his point that the comparison to Hugonnet et al. rate of elevation change should be performed only on the common pixels, hence at the edge of the cast shadow and not over the entire shadowed glacier area.

EC2A2: We agree and asked Romain Hugonnet to provide the same summary statistics only for the pixels covering the area between the edges of the smallest and the largest shadow mapped in satellite images. We opted for an area rather than a single line along the edge of the shadow, given that the shadow extents in satellite images varies in our study period

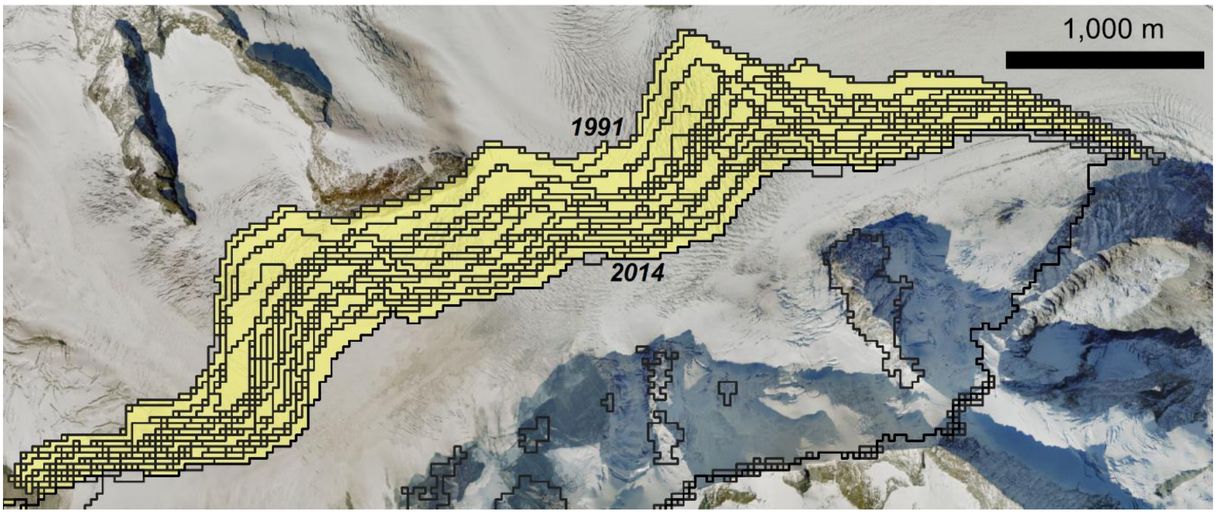


Figure 3: Area between the smallest/ shortest (2014) and the largest/ longest (1991) shadow cast from Dreieckhorn onto Great Aletsch Glacier (Switzerland). *Lines within the yellow area are all other 21 shadow outlines obtained between 1987 and 2019. Image in the background is a Bing Aerial basemap available through QGIS.*

(Figure 4). The example of the Great Aletsch Glacier shows how small the spacing of the shadow edges is in our study period, so that it is difficult to decide on exactly one of them. Choosing the area between the largest/ longest and the smallest/ shortest shadow better reflects the variance in glacier elevation change through time and provides a good compromise for comparing our data with those of Hugonnet et al. (2021).
Nevertheless, we did not find any substantial differences between the earlier and our updated analysis after running the models with the modified datasets (except that we had confused the trend at Gulkana West with that of Gulkana East in Figure 4). Thus, the revised analysis confirms the generally good agreement between our data and those of Hugonnet et al. (2021). Drivers for some of the diverging trends, as observed at Gulkana East, are now discussed in more detail in the discussion.

EC3: L235. It is unconventional to acknowledge a colleague in the main text. This belongs to the acknowledgement section (where you need to correct "help"). Hence you could reformulate L235 and rather cite a personal communication for a peculiar data extraction.

EC3A3: We reformulated this sentence to (L236-238): "*We used time series of glacier elevation change extracted along simplified outlines of glacier shadows (Fig. 2), provided as summary statistics on mean glacier elevation change between 2000 and 2019 by Romain Hugonnet (pers. comm., 2023) (Fig. 6)*".

EC4: Figure 5. The legend needs to explain which results are from DEMs and which are from cast shadows.

EC4A4: We added a legend to this figure accordingly.

From Reviewer #1 (R1), we received the following comments:

R1C1: L45: All these new references are from work based on aerial and terrestrial photographs. I suggest to move the reference before the "Corona and Hexagon" satellite names, or to rephrase the sentence to fit the references better.

R1A1: We agree and rewrote the sentence to read as (L43-45): *"These appraisals are largely constrained to the past two decades (Belart et al. 2020; Geyman et al. 2022; Mannerfelt et al. 2022), with few exceptions such as Corona and Hexagon missions, which provided one-time stereo image pairs between the 1960s and 1970s (Lovell et al. 2018; Dehecq et al. 2020)."*

R1C2: L120: It would be nice with a second reference to your appendix table with the ArcticDEM ID here.

R1A2: We added a reference to Table A2 accordingly.

R1C3: L126: This is the first time SRTM-3 is mentioned. Could you clarify in a parenthesis what the difference is to SRTM-1? (It's stated further down, but the reader is introduced to the abbreviation here).

R1A3: In our revised manuscript, we have deleted the paragraph where this sentence was located. We have also deleted the entry regarding SRTM-3 in Table A2.

R1C4: L145: I believe the second "shadow" in this sentence should be in plural.

R1A4: We used the plural instead.

R1C5: L234-235: "[...] shows a sinusoidal up and down." sounds like a word is missing. Perhaps "[...] shows a sinusoidal variation up and down throughout the seasons."?

R1A5: We changed this phrase so that it reads (L235-236): "… time series of glacier elevations show seasonal variations".

R1C6: L265: Please change "[...] 10 times that of [...]" to "[...] 10 times more negative than [...]" or similar to lower ambiguity about this statement (it's unclear in which direction it is 10 times different).

R1A6: We changed this statement accordingly.

R1C7: L418: If you wrote "less negative" instead of "lower" the sentence would read better in my opinion. -1.42 is lower than -1.08, but the magnitude is the opposite (as I presume you allude to). I find it less confusing to write "less negative".

R1A7: We agree and changed the wording accordingly.

R1C8: L488: Romain is indeed a "hepful" person! (Little typo in "help")

R1A8: We changed "hep" to "help" accordingly.

From Reviewer #2 (R2), we received the following comments:

R2C1: R1C2/R1A2: With respect to my comment on the use of VFP vs SRTM 1'', I agree that the SRTM 1'' has void filled and is not more explicit than VFP about the source of data in this area. It is also true that both mountains casting shadows over Baltoro are those affected by voids. The DoD in Figure R1C1 reveals very well the extent of those voids, while also demonstrating that data used to fill either VFP or SRTM 1'' are very different. That being said,

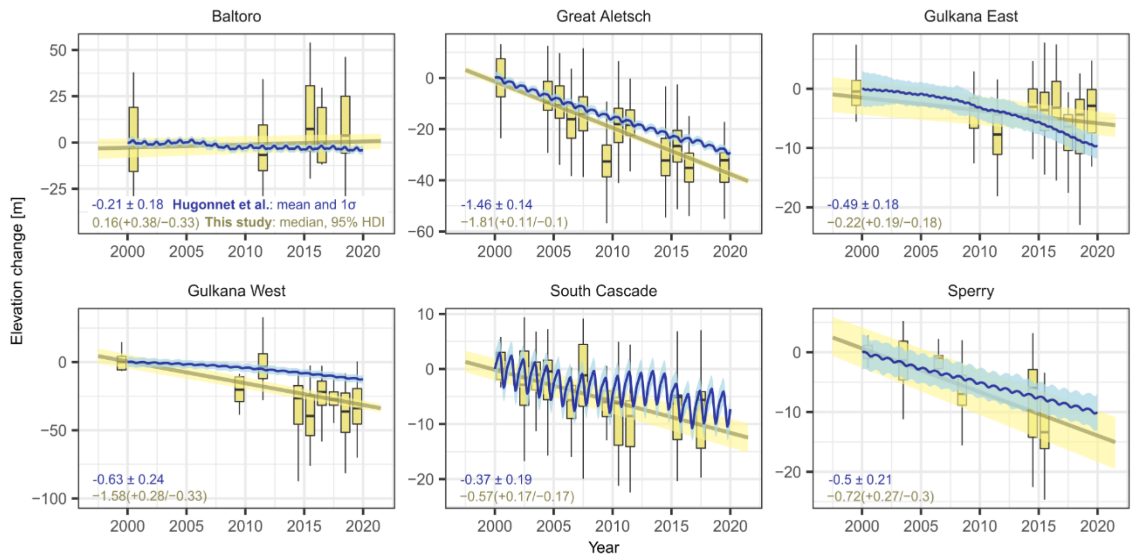
the sole appearance of hillshade DEMs from VFP or SRTM 1'' arguably suggests that SRTM 1'' conveys substantially better resolution and details which I believe should justify its use over VFP regardless. It also suggests that the voids in SRTM 1'' are not simply interpolated in this instance as the hillshade reveals topographic details that an interpolation may hardly achieve. Following on this, the authors now cite Fig. 6 in Liu et al. 2019 to support higher accuracy of VFP in steep terrain. For some reason, the reference provided in the rebuttal is wrong (namely, Liu, Xiaodong; He, Pengcheng; Chen, Weizhu; Gao, Jianfeng (2019): Improving Multi-Task Deep Neural Networks via Knowledge Distillation for Natural Language Understanding. Available online at https://arxiv.org/pdf/1904.09482.). The reference provided in the revised paper is correct. Nevertheless, I am confused by this argument. It is true that Fig 6 in Liu et al. 2019 suggests MAE of SRTM 3'' is greater than VFP, yet MAE of SRTM 1'' is reported as substantially better regardless of slope. I am left puzzled by this result not making the author reconsider the use of VFP, since they construe it as invalidating the use of SRTM 3''.
As the authors explain themselves, DEMs of higher resolution might better preserve the distinct shape of the mountains. In view of this, it seems obvious that even if underlying filling data are unknown, there SRTM 1'' can be seen as a better candidate for this study which weakens the rebuttal. The assessment of peak elevation is clearly affected by large uncertainty and unconvincing I believe to conclude. That being said, I find that SRTM 1'' resolves Mitre Peak with an elevation of 5994m above EGM96 which is not as different from the estimated elevation of the peak as the author suggest. In fact, it is unclear why the authors report a height of only 5904m from SRTM 1'' in Fig R1A2 which undermines again the rebuttal of this comment.
Overall, I remain unconvinced by this answer and still believe the relevance of using VFP is weak and unfounded.

R2A1: We note that the discussion about the shape and height of Mitre Peak has taken on an intensity that we had not anticipated. First of all, we agree with the referee that SRTM void filling is not simply interpolation. In fact, version 3.0 has been filled with ASTER GDEM version 2.0. This, however, may not necessarily improve the SRTM DEM in terms of accuracy as the ASTER GDEM has severe issues in mountainous terrain (see http://www.viewfinderpanoramas.org/reviews.html#aster). Second, we cannot reconstruct the elevation of 5994 m of Mitre Peak in SRTM 1", which the reviewer had mentioned. We obtained the DEM from opentopography (www.opentopography.org) and the DEM has the vertical datum EGM96. In our reply to the editor (EC1A1 and Figure 1), we showed that different vertical datums cannot cause the possible offset between the SRTM and VFP because glacier elevations in both DEMs are the same. In our comparison of peak elevations, we use the unprojected DEMs so that we can also exclude that gray-value interpolation during reprojection causes lowering of the elevations. De Ferranti (the producer of the VFP) replaced missing values with data obtained from topographic maps. The higher values of the Mitre Peak in the VFP data suggests, that this data is eventually more accurate than any of the DEMs.
The main manuscript now contains a section in which we explain the comparison of shadows retrieved from the different DEMs as well as Figure 2 shown above (Figure 10 in the revised manuscript).

R2C2: I am unconvinced that the comparison with Hugonnet et al (2021) should account for all pixels inside the shadow. Shouldn't it rather be only those pixels mapped along the edge of the shadow instead?

R2A2: We agree and asked Romain Hugonnet for the same summary statistics only for the pixels at the edges between the largest and smallest mapped shadow. We assume that these shadows are the endmembers of the variance in shadows derived from Landsat images in our study period. Re-running our analysis using the new data did not change our findings (Figure 5); however, the trends at Great Aletsch Glacier agree better between the two methods. With the exception of one year on the Great Aletsch and Gulkana East glaciers, the Gaussian process regression models of Hugonnet et al. (2021) overlap with our data (interquartile ranges of the boxes), indicating good agreement between the two methods.

Previous Figure 6



Revised Figure 6

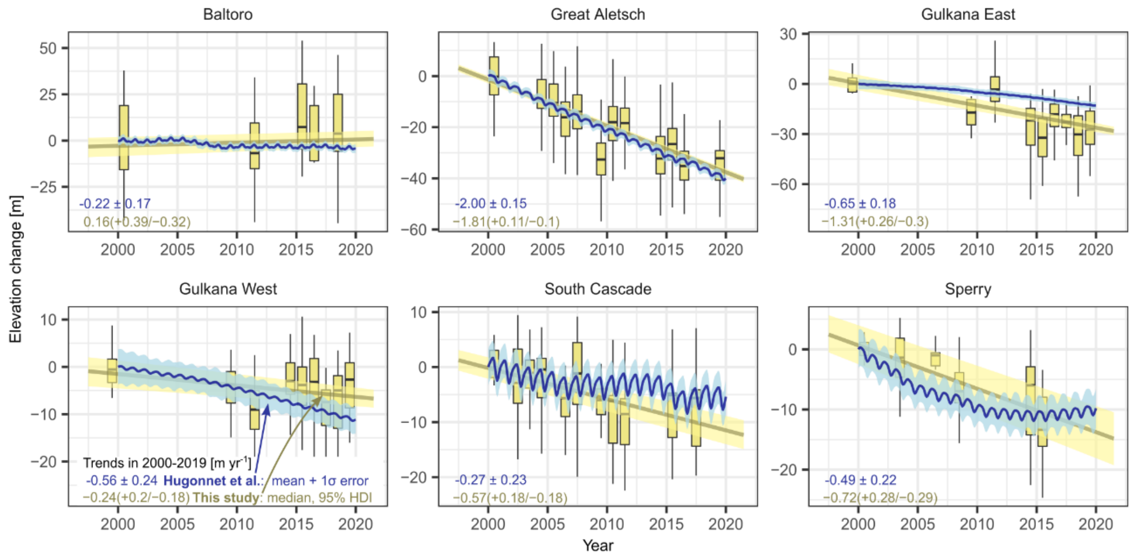


Figure 4: Comparison of the original Figure 6 (top panels) and revised Figure 6 (lower panels). *The trends calculated from simplified shadow outlines (top) largely agree with those obtained from the area between the shortest and longest shadow on each glacier. Note that we had confused Gulkana East with Gulkana West in our previous submission. We also improved the legend.*

R2C3: As per my comment above, I don't understand why all pixels in shadow are used since the method can only capture changes on the edge of the shadow.

R2A3: Please see the answer **R2A2** above.

R2C4: Further to that, the response provided in R1A4 "suggest slight local increases" is now "suggests no change" in the revised manuscript. This is a rather significant change in conclusion which further underpins what I still believe is a limitation of the method, and at least deserves more explanation as to why the inference has changed given I understand the authors did not change the DEM. In effect, it appears the data for Baltoro in revised Figure 4 have been modified compared to the original Figure 4 which led to a different statistical model. This requires clarification.

R2A4: We thank the reviewer for bringing this change in our last response letter to attention. Following our initial analysis, we had identified a few bearing lines that had lengths far outside the expected frequency distribution, and excluded them from the analysis. In addition, some bearing lines were wrongly classified as lines connecting modelled shadows with mapped shadows whereas they were actually connecting one of those shadows with the outline of the Randolph Glacier Inventory outside the area of the second shadow. We had made those adjustments to the dataset for all considered years. To this end, these changes led to different results compared our very first manuscript version.

R2C5: Finally, the authors dispute my statement that Hugonnet et al. (2021) shows the trend at Baltoro is "unambiguously negative". They argue that on a pixel basis, the uncertainty exceeds the elevation change. Although that latter is correct, it cannot be construed into defeating my argument since Figure R1C4(a) reveals the wide-spread negative value. Suggesting that uncertainty on a pixel basis applied to the overall budget, or even a subset of pixels, is erroneous. It corresponds to confusing standard deviation applying to a single measurement compared to standard error applying to the average. For this argument to be valid, it would imply that Huggonnet et al (2021) data are biased, which is not to be the case by design. Suggesting that the authors data are more accurate on a local scale can only be supported by a rigorous analysis of the distribution of values from pixels along the shadow with the standard error being considered in view of the sample size.

R2A5: We thank the reviewer for clarifying their former statement. We believe that there is a misunderstanding in the interpretation of the phrase "unambiguously negative". We interpreted this phrase only such that every measurement (cell value) of glacier elevation change, i.e. the mean **and** the error associated with that mean estimate), in the data from Hugonnet et al. (2021) needs to be smaller than zero to be "unambiguously negative". In showing the measurements for every raster cell, we showed that the trend in mean elevation change in any cell is both positive and negative, assuming one standard deviation error, regardless whether this cell was extracted at the edge of the shadow or not. In addition, some individual pixels even had positive mean values of glacier elevation change. We did not talk about the "overall budget" of the observed changes from Hugonnet et al. (2021), nor did we conclude that their methods or results are “biased”. Indeed, our manuscript acknowledges their elaborate method and the high value of their dataset.

R2C6: Thank you for addressing separately both areas of Gulkana. Nonetheless, I am sceptical again that the proposed method suggests a rate of thinning seven times (not 10 times as written in Section 4) faster in the West (Ogive mountain) compared to East (Icefall Peak), namely -1.58 compared to -0.22 m a-1. Again, Figure R1C4(b) provides key insight to compare glacier change in both areas. Although the authors signal this discrepancy, they unfortunately come short of providing a convincing explanation or stressing what I believe could be exemplifying the limitation on the method.

R2A6: We are unsure where these values of glacier elevation change at Gulkana Glacier come from as we could not find those in our revised text or figures. In any case, we agree that there are differences, both in the method and in the results between our appraisal and that of Hugonnet et al. (2021), let alone any other appraisal using remote sensing or field work. In our manuscript, we had written: *“One reason for the discrepancy between the two datasets may be the rigorous filtering of outliers in the dataset of Hugonnet et al. (2021), whereas our method maintains the elevation changes of all bearing lines, regardless of their distances from the mean or median”*. In the discussion, we now point at the advantage of using more informed priors from other glaciers or studies to reconcile the posterior trends, even if the physical or methodological drivers of the underlying trends remain unknown (L425-432): *“In any case, our Bayesian framework objectively propagates these errors and uncertainties. One promising avenue for future research is to use more informed priors based on previous research on glacier elevation change (Hugonnet et al. 2021). Narrower and stronger priors may reduce the width of our posterior trends on glacier elevation changes that we currently observe at Sperry Glacier, for example (Fig. 4). They might also offer a better compromise to balance some of the differences within our data (e.g. between Gulkana East and West), and also between our data and data from previous research. One of these examples may be the outstanding trend at Gulkana West (Fig. 6), where the physical causes and methodological differences between our appraisals and that of Hugonnet et al. (2021) remain to be determined.”*