**Anonymous Reviewer #2**

We thank the reviewer for the constructive comments. In our letter, we highlight the comments from the reviewer in orange. Our responses are in black font, and our planned corrections are highlighted in bold font.

**R2C1:** The authors of the manuscript present a promising new way of extracting glacier elevation change from the changing extent of shadows cast from direct sunlight. The manuscript clearly portrays the potential of this approach, and important discussion is made on the potential of scaling this approach to regional or global scope. I find the validation analyses slightly unconvincing, however; a general agreement between this approach and one "reference" classical DEM comparison of questionable quality is found at Grosser Aletschgletscher, which is used as the main argument for the use of this approach. From a critical point of view, one general agreement is not enough to prove the potential of this method. If the authors either had another convincing case-study, or put less emphasis on this one agreement in the text, it would not be as large of an issue. I also do not consider absolute proof necessary for this study to be a novel contribution and an interesting read. I have comments on the text and some details of the analysis, but I generally find it well written and not in need of significant revision. I therefore suggest minor revisions of this work, and I look forward to be able to share this with my colleagues when published in its revised state!

**R2A1:** We thank the reviewer for emphasizing the value of our work and encouraging us to revise a few minor points. We see our study as a proof of concept, making researchers aware of this untapped resource to use shadows for quantifying local changes in glacier thickness. Our approach hinges on the availability of freely available data, which may bias our attention to glaciers with a high number of DEMs and satellite images. In our revised manuscript, we will study in detail the effect of DEM resolution at Baltoro Glacier (see our reply **R1A2** to Reviewer 1) and also compare our trends in glacier elevation change to those from Hugonnet et al. (2021). These revisions will allow for a better balance in representing the different glaciers and data sources in our study.

## General comments:

**R2C2:** In the end of the discussion, you mention the potential to automate and therefore efficiently scale the analysis. This is what I think is the biggest take-home message, and I suggest that it should be reflected in the abstract. Doing this manually on single glaciers is only so impactful, but the fact that this approach utilises globally available data, with relatively small automation challenges, speaks to the massive potential of this approach!

**R2A2:** We thank the reviewer for this positive appraisal and recognition of our work. In our revised manuscript, we will stress the idea of the reviewer: **"We conclude that our approach has the potential to complement existing or future in situ measuring networks anywhere**

**on Earth where mountains shade parts of adjacent glaciers. We thus enrich glaciological and geodetic assessment with a new method that helps quantifying glacier elevation changes especially at high altitudes with limited access."**

**R2C3:** I was initially worried about the use of terrain-corrected Landsat images, but with some handkerchief mathematics, I could calm myself down. I suggest modifying the reasoning that I present here and putting it in writing, as it strengthens your choice of data. My worry was that the images will be skewed depending on what DEM is used by USGS to correct them.

Optimally, one should run the correction with the same DEM as the one you use for shadow simulation, which would of course add significant computational overhead to the approach. If that is not done, there will be an error that is dependent on the DEM difference and the incidence angle (the relative angle between the pixel coordinate and the satellite at the time of acquisition). Landsat 8 orbits at an altitude of 705 km and its OLI instrument has a swath width of 185 km according to Wikipedia, meaning that the incidence angle varies between 0 and about 7.5 degrees. This means that the DEM bias that is required to shift the image by one pixel is around 230 m, which is extremely unlikely. Therefore, it should not matter too much, meaning that a potentially incorrect terrain correction is fine. If this is brought up in the text, it would save others having to go down the same rabbit hole as me!

**R2A3:** Ideally, our workflow should make use of images that are all co-registered to a terrain-corrected master image. Co-registered images will increase the precision of our analysis and avoid systematic offsets. **In our revised manuscript, we will discuss the effects that may arise if images with more than a pixel offset will enter our workflow.** We agree with the reviewer that an additional source of uncertainty would be introduced if different DEMs were used for terrain correction for different scenes covering the region. However, according to the Landsat L1TP processing description (https://www.usgs.gov/landsat-missions/landsat-levels-processing), the elevation data used for relief displacement remain the same throughout the Landsat time series.

**R2C4:** References are sorted inconsistently in the text. Sometimes they are sorted by year and sometimes alphabetically. This should be made consistent in accordance to the journal standard (which I believe is by year).

**R2A4**: The journal's guidelines for authors suggest that "in terms of in-text citations, the order can be based on relevance, as well as chronological or alphabetical listing, depending on the author's preference". **We decided that we will sort in-text references by year in the revised manuscript.**

**R2C5:** This comment requires no changes to the manuscript; it is simply a suggestion for the future. Filtering of the geodetic lines as shown in Figure 3e could easily be done programmatically. If all lines are given an index before cutting them, then the cut line that is furthest away from the sun for each index would be the line to keep, whereas all others are erroneous.

**R2A5:** We thank the reviewer for this helpful suggestion. We consider implementing this idea when applying our approach to a larger scale.

**R2C6:** I am not a Bayesian wizard, and therefore needed help from my colleagues to understand the statistics that are performed in this manuscript. I was told that this is a good approach as it is simple and understandable for anyone that is into the field. For the others (like me), however, could you keep that in mind for Section 3.2? For example, it is not argued for why this is preferable over a simpler optimization method (e.g. least squares). I now understand that this approach simply keeps the outgoing uncertainty high when provided with few data, whereas a regular regression would result in one misleading line that may be completely wrong. An easy argument for your approach is that you have few data points, and therefore cannot use regular optimization tools. But this was not clear to me before I had spent a few hours figuring out what is going on.

**R2A6:** A key motivation for using multi-level models is that we can obtain parameters (e.g. intercepts and slopes) for elevation change of different glaciers within one single model. Thus we can learn both the parameters for all glaciers and the variation between parameters by using a set of individual glaciers that are part of a larger population of glaciers. Multi-level models are advantageous for datasets with a different number of observations in each group, or in our case glacier, in which one glacier might have hundreds of bearing lines (e.g. Great Aletsch Glacier) and others might have fewer data (e.g. Gulkana Glacier). The hierarchical model structure avoids over-fitting parameters for glaciers with many bearing lines and generally improves inference for groups with few data points.

Bayesian inference estimates all model parameters based on the data (the likelihood) and prior information about those parameters. If a glacier had few bearing lines in some years, the posterior distribution of parameter estimates will be broad and vice versa. The annual trends in glacier elevation change depend entirely on the available data, and the posterior distributions reflect the uncertainties accordingly. In essence, our model could also estimate the trends for two or one or even no years of mapped bearing lines, a situation that is difficult to deal with in ordinary least squares regression. This estimate would correspond to the mean trend derived from all glacier shadows included in our analysis. The more bearing lines we map for a new, unobserved year at a particular glacier, the more specific our estimated trends will be for that glacier if its trend deviates from the global mean of the entire population in our dataset.

**R2C7:** Also, in your equations, I see no factor for elevation. Elevation change is highly correlated with elevation (i.e hypsometric gradient), so different points at different elevation are expected to have different values. Can you argue in the text for why this is hopefully not the main reason for the spread you present later in the Results chapter? Because if it is, I would suggest you update the model accordingly.

**R2A7**: We are unsure how to interpret the reviewer's comment. The elevation of the glacier at a given bearing line is not formally part of our model. We estimate $\delta h$ by calculating $\delta l$, the change in the length of bearing lines through modelled and mapped shadows, using a trigonometric relationship, see Fig. 1d in our previous manuscript. What we understand from the reviewer's comment is that values of $\delta l$ might correlate with their associated elevation on the glacier. In other words, as we go from lower elevations to higher elevations along the rim of shadow, the reviewer assumes that values of $\delta l$ might increase, if positively correlated, or decrease, if negatively correlated with glacier elevation. However, the shape of the shadow depends on the shape of the adjacent mountain, not the glacier, so that we expect to find a high variance of $\delta l$, even at a constant elevation of the glacier. The example of Mitre Peak (see Figure **R1A2** in the reply to Reviewer 1) underlines this concept: although Baltoro Glacier is flat (little change in elevation in the shadow covered area), bearing lines through that shadow through that shadow will have large variance.

**R2C8:** Currently, there is no consideration of elevation change uncertainty between the DEM comparisons in Section 3.3. The scanned topographic map and the SwissALTI3D DEM of Grosser Aletschgletscher for example may have substantial offsets, which are usually quantified from neighbouring stable terrain. Since the North America DEMs have gone through peer-review, they must have an associated uncertainty in the publication. For Grosser Aletschgletscher, I suggest validating that differences on stable terrain are close enough to zero and putting it in writing, or in a new figure. Without considering this source of error, I would not trust the validation from Switzerland, and therefore cannot trust the statement of similar trends on L230.

**R2A8:** We agree with the reviewer that we did not assess the uncertainties of the comparison of the topographic maps and the SwissALTI3D DEM. Unfortunately, we do not have access to the georeferenced topographic data but need to rely on manual measurements on the web-interface of swisstopo.admin.ch. However, in our revisions, we will address this point and measure points of stable terrain to exclude that measurement errors or changes in vertical datum of the maps have caused a trend which is absent in reality.

**R2C9:** Please validate the date of the ArcticDEM product you used over Gulkana Glacier. The ArcticDEM explorer only shows strips from 2011–2016 in that location, while you claim that your DEM is from 2009.

**R2A9:** We confirm that our strip is from June 16, 2009, and the file we used is: SETSM_WV01_20090616_10200100079A2600_1020010007D06000_seg1_2m_v3.0.tif

**Specific comments:**

**R2C10:** L18: Mention that the thinning pertains to the points of measurement, not the entire glacier. Otherwise, people might cite this number incorrectly.

**R2A10:** We agree that this and the previous statement could be misleading. We will revise this phrase to: **"We validated the relative differences with in situ geodetic measurements of glacier elevation change where these shadows occurred. We find that shadow-derived glacier elevation changes are consistent with independent photogrammetric and geodetic surveys in shadowed areas. Accordingly, a shadow cast on Baltoro Glacier**

**(Karakoram, Pakistan) suggests slight local increases in elevation between 1987 and 2020, while shadows on Great Aletsch Glacier (Switzerland) point to the most negative thinning rates of about 1 m per year. Our estimates of glacier elevation change are tied to the occurrence of mountain shadows, and may help complement field campaigns in regions that are difficult to access."**

**R2C11:** L36: Is 141,000 km³ for mountain glaciers or total land ice?

**R2A11:** The value is for all glaciers in the Randolph Glacier Inventory, and we used the estimate from Millan et al. (2022).

**R2C12:** L37: Change "cover" to "covers"; magnitudes are in singular.

**R2A12:** km³ (cubic kilometres) are plural, so our statement was correct.

**R2C13:** L40: Change "elevations" to "elevation"; magnitudes are in singular.

R2A13: We will change our wording accordingly.

**R2C14:** L41: Change "(ICESat)" to "(e.g. ICESat)".

**R2A14: We will change our wording accordingly.**

**R2C15:** L44: There are more large-scale long-term studies (e.g. Belart et al., 2020; Geyman et al., 2022; Mannerfelt et al., 2022).

**R2A15:** We thank the reviewer for making us aware of these studies. **We will cite them in the revised manuscript.**

**R2C16:** L48: All Landsat programme products do not have a resolution of 30 m. Only most bands of the TM, ETM+ and OLI instruments (see the Landsat program wikipedia article).

**R2A16:** We agree and will add **"a moderate spatial resolution of 30 m in the visible to shortwave infrared electromagnetic spectrum".**

**R2C17:** L49: Change "mapped in Landsat" to "mapped from Landsat"

**R2A17: We will change our wording accordingly.**

**R2C18:** L51: Change "help reveal" to "be used to estimate" or similar. Outlines themselves cannot reveal elevation change. The first reference uses area for a rough estimation of volume change, and the second uses them to crop their DEM differencing.

**R2A18: We will change our wording accordingly.**

**R2C19:** L64: The DEM abbreviation is already introduced in the abstract, and the DEM abbreviation is used in L49.

**R2A19: We will delete "Digital Elevation Models" to shorten the text.**

**R2C20:** L81: Change "long(decadal)" to "decadal"

**R2A20: We will change this statement accordingly.**

**R2C21:** L86: The line starting with "High and steep" is very short, and could be rephrased or combined with another.

**R2A21:** We strive to keep our wording simple and would therefore like to keep this sentence unchanged.

**R2C22:** L87: Specify that "Concordia" is the name of a mountain.

**R2A22:** In our manuscript, we had written that Concordia (~4,500 m a.s.l.) is the confluence of Baltoro and Godwin-Austen Glacier. **We will add this location also in Figure 2.**

**R2C23:** L89: Change "Gulkana" to "Gulkana glacier" to be consistent with L90 and to help the reader. I know that the word "glacier" occurs in the beginning of the sentence, but the wording is still not obvious to me.

**R2A23: We will change our wording accordingly.**

**R2C24:** L105: Did you validate the georeferencing? Sometimes they can be off if not enough GCPs exist. Please just mention that this was (hopefully) not a problem.

**R2A24:** We could not find any inconsistencies and will add to our revised manuscript: **"L1TP images offer high radiometric and geodetic accuracy by using ground control points and correcting for topographic displacement using regional DEMs (https://www.usgs.gov/landsat-missions/landsat-levels-processing#L1TP). We could not find any notable offsets between successive images."**

**R2C25:** L109: Which ArcticDEM product was used? The mosaic or an individual strip? If an individual strip, which id?

**R2A25: We used individual strips, and will add the individual IDs to Table A2.**

**R2C26:** L112: Do you mean the swissALT3D DEM 2019 version? The acquisition year over Grosser Aletschgletscher is 2017–2018 as is correctly mentioned later.

**R2A26:** We are unsure about the reviewer's comment.

**R2C27:** L122: It would be nice to know a bit more about why the green band was chosen. For example, did it have the highest difference between shaded and unshaded locations? It is mentioned that "shadows appear dark" in the green band, but not "darker" than the other bands.

**R2A27:** In our revised manuscript, we will add: **"Snow, firn, and ice have minimal absorption in the blue-green range, whereas red and infrared light is strongly absorbed on these surfaces. This trait enhances contrast at the interface of glaciated surfaces and shaded, colder areas with increasing wavelength. Incoming and reflected electromagnetic wavelength in the green band is also less affected by the Rayleigh scattering in the atmosphere compared to the blue band that has a shorter wavelength. The green band therefore offers good compromise between contrast and surface reflectance measured at the sensor".**

**R2C28:** L125: Does the algorithm in SAGA account for the Earth's curvature? The effect is in the order of 1 m in elevation per 3 km of horizontal distance, so this would matter somewhat at these scales. Please just mention if it is already accounted for, or mention that future improvements can be made by accounting for it.

**R2A28:** Olaf Conrad, the core developer of SAGA, emailed us that the algorithm 'Analytical hillshading' does not account for Earth's curvature. Yet we expect, if at all, a minor role of Earth's curvature in our results. The longest bearing line that we modelled had a length $d$ of 308.73 m (Gulkana 2019). We can calculate the deviation $a$ of a plane from the spherical surface (radius $r$) with increasing distance $d$ from the point of contact (plane to sphere) using the theorem of Pythagoras: $a = \sqrt{d^2 + r^2} - r$. For $d$ = 0.30873 km and $r$ = 6371 km, i.e. the mean radius of the Earth, we obtain a deviation $a$ = 0.0007480318 km (i.e. 7.4 mm) from Earth's surface. In our analysis, however, we use only the much shorter segment of the bearing line that extends beyond the modelled shadow, and therefore obtain even smaller values for $a$.

**R2C29:** L128: Specify that it's RGI version 6.

**R2A29: We will revise this statement to "Randolph Glacier Inventory (RGI) V6.0".**

**R2C30:** L187: Are you entirely sure that the maps represent these exact years? Swiss maps are very often asynchronously updated over time, and can have large (tens of metres) differences to what one would expect (c.f. Fig. 10 in Mannerfelt et al., 2022).

**R2A30:** We used the years at which the topographic maps were available first in the online interface of swisstopo (link). As stated above, we will aim to detect possible offsets between SwissALTI3d data and the maps on stable terrain. The paper by Mannerfelt et al. (2022) provides useful backgrounds and details for this analysis.

**R2C31:** L207: The swissALTI3D DEM is a mix of Lidar at low elevation and aerial photogrammetry at high elevation (see the technical information on swisstopo's website). At Grosser Aletschgletscher, it is therefore most likely based on photogrammetry.

**R2A31:** Yes, you are right. Information by swisstopo (link) shows that the area of the Great Aletsch glacier was last updated in 2019. The update was based on photogrammetric aerial imagery from 2016-2018. We will update the description of the data in a revised version of the manuscript.

**R2C33:** L211: Repeat what the line spacing is. 45 lines is not informative unless the spacing is mentioned.

**R2A33:** In line 129, we had written that we use a regular spacing of 30 m between the bearing lines, and will add this information here again.

**R2C34:** L231: The topographic maps may have a high resolution, but it is far from guaranteed to have a high accurancy and precision. I would rephrase this from "high-resolution DEMs" to "reference DEM comparison".

**R2A34: We will change this statement accordingly.**

**R2C35:** L244: Is the substantial variance found before or after considering the elevation dependent elevation change that is observed on almost every glacier in the world? If they are at relatively similar elevation, then say that: "... in spite of their similar elevation ...".

**R2A35:** We will change our wording to: **"The elevation changes obtained from bearing lines have substantial variance in a given year despite covering a small range in elevation along the glacier**

**R2C36:** L252: Where is the date of the GLO-90 DEM information from? On the associated OpenTopography website it says 2011–2015

**R2A36:** Yes, you are right. The data were acquired through the TanDEM-X mission between 2011 and 2015. We will correct this mistake in a revised version of the manuscript.

**R2C37:** L277: Right now, you only show a difference in the horizontal variability that is associated with different DEMs, not the effect on calculated elevation change. My Figure 7 suggestion below would solve this problem.

**R2A37:** We would like to refer the reviewer to our reply **R2A46**.

**R2C38:** L294: The effects of SAR penetration would lead to a potential negative bias (longer shadows), no? But since this bias will be consistent between years, one would just have to subtract the shadow-derived elevation change at the year of acquisition of the DEM. So this is arguably a very simple fix. Could this be put in words, assuming you agree with me?

**R2A38:** We agree, and will add: **"Our approach does not account for any offsets in elevation due to the SAR signal penetrating into snow. Yet this offset can be treated as constant when drawing bearing lines through shadows, given that the input DEM (SRTM) remains unchanged in our analysis."**

**R2C39:** L302: "Precision" (spread) AND "accuracy" (bias). There seems to be both consistent and inconsistent errors in the shadow-maps derived from poor DEMs.

**R2A39: We will add "accuracy" accordingly.**

**R2C40:** L309: Please elaborate on why you think winter months are better suited for this method. Lower solar angles, I presume? What about the contrast between shaded and unshaded terrain?

**R2A40:** Yes, mainly because of sun angles (and calculation is not influenced by the climate change related shift of starting point of melting season as well as summer heat correlated differences in melting rates (but by snowfall). On top of that we thereby wanted to make sure that the selected season of the year of every Landsat scene acquisition matches the time when SRTM was acquired.

**R2C41:** L330: Steep topography is arguably not a requirement; just stable topography. Indeed, steep slopes increase the potential time of day at which shadows can be created, so it works better with steeper topography. But the approach still works in shallow topography; just at fewer hours of the day! On e.g. ice caps, there is no stable topography to cast shadows, so the approach fails. So steep topography and high latitudes are preferable, but stable topography is the only theoretical requirement.

**R2A41:** We agree and will rewrite this statement to: **"Our method can be applied globally, but is restricted to those glaciers that are surrounded by stable topography. Ideal environments for our approach are glaciers close to steep topography in high latitudes, producing cast shadows long enough to infer differences in bearing lines".**

**R2C42:** Figure 1: I just want to add that I think this is a great figure that explains the concept simply and artfully! Great job!

**R2A42:** We appreciate the kind words of the reviewer.

**R2C43:** Figure 2: The Randolph Glacier Inventory is mentioned without abbreviating it, contrasting it to Figure 3.

**R2A43: We will introduce the abbreviation in Fig. 2 and in the caption of Fig. 3, will delete the full name and use the abbreviation only.**

**R2C44:** Figure 3: The RGI abbreviation is introduced multiple times. Only RGI is necessary here.

**R2A44: We will reduce the redundant abbreviations.**

**R2C45:** Figure 5: It is unclear where the grey bubbles come from. Are these from other satellite images than Landsat? Please clarify this in the text.

**R2A45:** All bubbles are from other sources than Landsat. We will change the caption to: **"Grey bubbles are historical data (Siegfriedkarte) obtained before the Landsat period".**

**R2C46:** Figure 7: I recommend to add another axis label on top where you show calculated height change assuming the solar parameters you decided. This would solve my issue in L277 that you have actually not shown the variable DEM quality effect on elevation change.

**R2A46:** This is a very good idea. We will add a second x-axis to map horizontal values to vertical offsets given the chosen azimuth and sun elevation.

**References:**

Belart JMC, Magnússon E, Berthier E, Gunnlaugsson AT, Pálsson F, Aðalgeirsdóttir G, Jóhannesson T, Thorsteinsson T, Björnsson H. 2020. Mass Balance of 14 Icelandic Glaciers, 1945–2017: Spatial Variations and Links With Climate. Frontiers in Earth Science 8:163. doi: 10.3389/feart.2020.00163.

Geyman EC, van Pelt WJJ, Maloof AC, Faste Aas H, Kohler J. 2022. Historical glacier change on Svalbard predicts doubling of mass loss by 2100. Nature 601:374–395. doi:10.1038/s41586-021-04314-4.

Mannerfelt ES, Dehecq A, Hugonnet R, Hodel E, Huss M, Bauder A, Farinotti D. 2022. Halving of Swiss glacier volume since 1931 observed from terrestrial image photogrammetry. The Cryosphere Discussions :1–32doi:10.5194/tc-2022-14.

**REFERENCES**

Hugonnet, Romain; McNabb, Robert; Berthier, Etienne; Menounos, Brian; Nuth, Christopher; Girod, Luc et al. (2021): Accelerated global glacier mass loss in the early twenty-first century. In *Nature* 592 (7856), pp. 726–731. DOI: 10.1038/s41586-021-03436-z.

Millan, R.; Mouginot, J.; Rabatel, A.; Morlighem, M. (2022): Ice velocity and thickness of the world's glaciers. In *Nature geoscience* 15 (2), pp. 124–129. DOI: 10.1038/s41561-021-00885-z.

U.S. Geological Survey (2023): Landsat Levels of Processing. Available online at https://www.usgs.gov/landsat-missions/landsat-levels-processing, updated on 2/13/2023, checked on 2/13/2023.