Answer to Reviewer 2

In the following, we provide a point-by-point Author Response (AR) to any of the Reviewer Comments (RC) obtained for the manuscript that was under discussion. When presenting suggestions for how the manuscript text could be revised (*italic text* in quotation marks), the line numbers refer to the revised manuscript.

RC1: Terminology: For the process that the stake melts out of the ice due to ablation and hence increasing its length of the "free end", the author use the term (vertical) displacement (L52 and others). This causes some confusion as displacement is usually connected to ice flow, which is not intended to be detected in the presented method and apart from tilting the ablation stake has no direct influence here. I suggest using the term surface elevation change (or something equivalent) instead.

AR1: This is a good point, and we realize that our wording could cause confusion. Also taking into account the comment of Reviewer #1, we suggest to replace “stake displacement” with “stake emergence”. We believe that “stake emergence” is better suited than “surface elevation change” since the latter wording is often used in studies of geodetic mass balance to indicate the surface elevation change at a given position, i.e. at a set of fixed coordinates. In our case, instead, we are measuring the surface mass balance of a point that is advected downstream via the glacier’s ice flow. We note that our stake emergence is not to be confused with the concept of “emergence velocity”, which is the difference between the local elevation change and mass balance. To avoid this latter confusion, we added the following sentence to the manuscript:

L XXX: *“To automatically read the stake, a computer-vision algorithm is used to derive the stake emergence between two images. Note that the stake emergence is the vertical movement of the stake out of the ice and is not to be confused with the concept of “emergence velocity”, i.e. the* difference between the local elevation change and mass balance*.”*

RC2: Preselection of images: In the method section the authors describe that images are taken every 20 minutes. I wonder if the method analyses all images or if a preselection using which criteria was applied.

AR2: Correct, the images are acquired every 20 minutes when there is enough daylight while during night, the camera does not acquire images. For the method’s application, we do not perform any image preselection, i.e. we processes all of them. To better clarify this, we suggest rephrasing L80 as follows:

Line XXX: “*To automatically read the stake, all images acquired during the season are processed with a computer-vision algorithm that derives the stake emergence between pairs of subsequent images.*”

RC3: Threshold for extreme melt event (L185-196): In Tab. 1 the authors show that there is an altitude difference of almost 800 m between the stations used for deriving daily ablation values. This difference might largely explain the spread in the mass balance anomaly in Fig. 8. However, the altitudinal distribution of the stations is skewed to lower altitudes and thus taking the mean of the stations for defining the threshold should be reconsidered. I think the median is more significant than

the mean, although the number of stations is low. Speaking of which, the number of observations presumably might change over the ablation period, as higher stations experience a longer snow cover. There should be a note how the number of observations affects the interpretation of the mass balance anomaly.

AR3: We disagree with the first part of this comment, or we do not exactly understand it. In Figure 8, we cannot identify a trend by which the spread in the mass balance anomaly would be larger for lower elevations. To better show this, Figure A (here below) shows the standard deviation of the anomaly of every station against its elevation. If we understand the Reviewer's comment correctly, this figure should show a trend towards higher standard deviation for lower elevations, which is apparently not the case. To us, this observation makes sense because the anomaly we compute (see Sect. 3.3) is related to the average mass balance at that given station. This means that any difference in average mass balance (which, we agree, is directly determined by the station's elevation) does not play a role anymore.

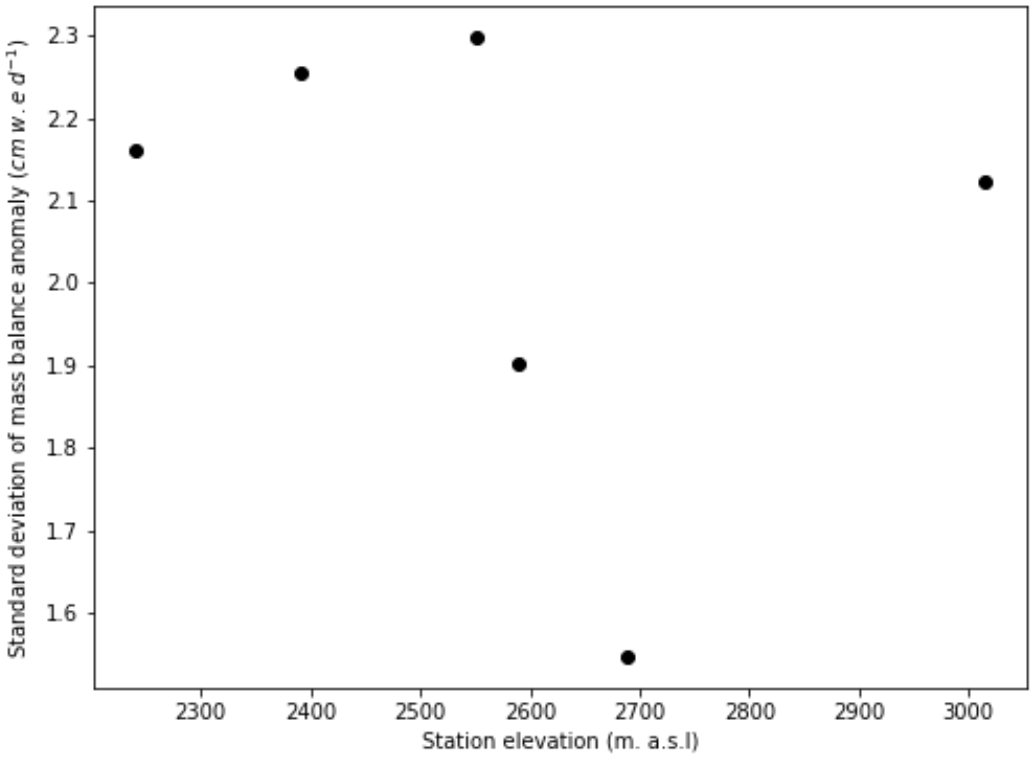

Figure A: Standard deviation of the mass balance anomaly against elevation.

For the second part of the comment, related to the number of stations, we reworded as follows to note that the number of observations influences the interpretation of the mass balance anomaly:

LXXX: "*The 85% quantile corresponds to –2.55 cm w.e d−1, and the event is classified as extreme when the mean anomaly across the stations exceeds this threshold. Because the stations have different operation periods (Tab. 2), the mean anomaly may be computed from a different number of stations and consequently be affected.*"

RC4: L32-34: I do not agree with these two sentences. A number of studies assess reasonable short term mass balance variations from geodetic measurements (e.g., Klug et al., 2018; Zeller et al., 2022; Beraud et al., 2022; Vincent et al., 2021). Consider rephrasing or omitting.

AR4: Thanks for the specification. We suggest rephrasing as follows:

LXXX: "*Interpreting geodetic mass balances at short time scales can be challenging, though, since the results are sensibly affected by the choice of the volume-to-mass conversion factor (Huss, 2013). Recent studies have nevertheless been using results from geodetic studies to gain insights on short-term mass balance variations (e.g., Klug et al., 2018; Vincent et al., 2021; Zeller et al., 2022; Beraud et al., 2022).*"

RC5: L77: Please rephrase and consider the width of the tape as well.

AR5: We suggest rephrasing as follows:

LXXX: "*The station setup consists of a camera and an aluminum stake that is marked with tapes of different colors. The tapes have a width of 2 cm and are placed 2 cm apart (Fig. 2).*"

RC6: Eq. 1: Please explain how you determine the stake inclination. Is it measured during the field visits or derived from the images or…?

AR6: The stake inclination is detected automatically by the algorithm, i.e. the angle between the red line in Figure 2 and the vertical axis). We clarified this with the following wording:

LXXX: "*Here, α is the stake inclination with respect to the vertical axis, i.e. the angle between the red line in Figure 2 and the vertical axis. The inclination is derived automatically by the algorithm […]*"

RC7: L257: Consider depicting these periods also in Fig. 8.

AR7: We thank the Reviewer for the suggestion. However, since Figure 8 shows the four different years, we think that only depicting heat waves in 2022 could be misleading. We thus prefer not to show this information in the Figure.

RC8: L305-307: Please better explain how winter snow accumulation impacts melt anomalies between individual stations. Winter snow accumulation might have an influence on the length of the ablation season, but how does it alter the ablation anomaly during core summer, when winter snow has melted since long?

AR8: The reviewer's question seems to be triggered by a misunderstanding: we agree that the winter snow accumulation has no direct effect on the summer melt rates (an indirect effect could be the albedo of the ice, which is arguably different for locations in which the snow cover is present over a longer period of time). Our point, however, was not the presence of the snow as such, but rather the different snow amounts, which indicate a difference in local meteorology. To avoid the possible misunderstanding, we reworded the sentence as follows:

L XXX: "*The reason for this is likely to be the different local meteorological forcing: while Findelgletscher and Rhonegletscher are influenced by southerly weather patterns, Glacier de la Plaine Morte experiences weather that is more similar to the Northern flanks of the Alps. This difference in meteorological forcing is corroborated by major differences in winter snow accumulation at the stations: whereas only 1-2m of snow are typically present in April on the tongue of Findel- and Rhonegletscher, more than 4m are often recorded on Glacier de la Plaine Morte […]"*

New References

Beraud, L., Cusicanqui, D., Rabatel, A., Brun, F., Vincent, C., and Six, D.: Glacier-wide seasonal and annual geodetic mass balances from Pléiades stereo images: application to the Glacier d'Argentière, French Alps, Journal of Glaciology, 1–13, https://doi.org/10.1017/jog.2022.79, 2022.

Klug, C., Bollmann, E., Galos, S. P., Nicholson, L., Prinz, R., Rieg, L., Sailer, R., Stötter, J., and Kaser, G.: Geodetic reanalysis of annual glaciological mass balances (2001-2011) of Hintereisferner, Austria, The Cryosphere, 12, 833–849, https://doi.org/10.5194/tc-12-833-2018, 2018.

Vincent, C., Cusicanqui, D., Jourdain, B., Laarman, O., Six, D., Gilbert, A., Walpersdorf, A., Rabatel, A., Piard, L., Gimbert, F., Gagliardini, O., Peyaud, V., Arnaud, L., Thibert, E., Brun, F., and Nanni, U.: Geodetic point surface mass balances: a new approach to determine point surface mass balances on glaciers from remote sensing measurements, The Cryosphere, 15, 1259–1276, https://doi.org/10.5194/tc-15-1259-2021, 2021.

Zeller, L., McGrath, D., Sass, L., O'Neel, S., McNeil, C., and Baker, E.: Beyond glacierwide mass balances: parsing seasonal elevation change into spatially resolved patterns of accumulation and ablation at Wolverine Glacier, Alaska, Journal of Glaciology, 1–16, https://doi.org/10.1017/jog.2022.46, 2022.