

Response to Reviewer Comments

Manuscript: egusphere-2025-5795

Journal: The Cryosphere

Authors: Spencer et al.

We thank both reviewers for their thorough and constructive comments. The manuscript has been substantially revised in response. Major changes include: (1) a fully open-source co-registration pipeline replacing the ArcPy-based approach; (2) a revised uncertainty propagation methodology using heteroscedastic error modelling and Rolstad et al. (2009) spatial propagation via xDEM; (3) incorporation of a 1975 glacier inventory as the reference outline for 1954-based epochs; (4) addition of Pléiades 2025 and ASTER 2008/2018 DEMs to improve temporal coverage; and (5) replacement of the E-Divisive changepoint analysis with Sen's slope and Mann–Kendall tests for monotonic trend detection. All responses below are numbered to correspond to the reviewers' original comments. Manuscript line numbers refer to the revised version.

Reviewer 1 (Robert McNabb)

Comment received: 10 April 2026

General Comments

RC1.1

I think some care needs to be taken with the vertical datum used by each of your two DEMs, especially the 1954 topographic map DEM. The SRTM that you are using is provided as the height above the EGM96 geoid, while the 1954 map is most likely the same as what is reported by Fariás-Barahona et al. (2020), orthometric height (m a.s.l.) measured from tide gauges along the coast. I think there's a good chance of large differences between the vertical datum of the topographic maps and the EGM96 geoid, which would likely show up as differences (errors) on the off-glacier areas of your DEMs, even after co-registration. It might be comparatively minor, but I think it's at least worth mentioning in the data description. That said, I think you might also want to mask your "source" DEM for the co-registration - it looks as though there are areas outside of the DEM that are still being included for the co-registration, which may also have an impact here.

Response:

We thank the reviewer for raising this point. Source DEMs were already being masked during co-registration; we have clarified this in the revised methods section. We have also conducted a thorough review of the vertical datums across all DEMs. The SRTM DEM is provided as heights above the EGM96 geoid (orthometric). The 1954 IGM DEM is based on Chilean topographic maps using the PSAD56 datum, which approximates orthometric heights at this latitude and is consistent with EGM96 to within approximately 1 m at 37°S — a difference we consider negligible relative to the overall uncertainty of the IGM DEM (stable-terrain NMAD of ~26 m). The UAV-derived DEMs and Pléiades 2025 DEM were generated in WGS84 ellipsoidal heights; we now apply a pixel-wise EGM96 geoid undulation correction ($N \approx 22\text{--}24$ m at 37°S, 71°W)

using the pyproj library prior to co-registration and differencing, converting all ellipsoidal DEMs to orthometric heights consistent with SRTM. The ASTER DEMs provided by R. Hugonnet were co-registered to TanDEM-X (ellipsoidal) and have also been geoid-corrected using the same procedure. These corrections and their spatial variation across the study area are now described in the revised §3.x.

RC1.2

I don't think the approach taken in §3.5, to convert the off-glacier NMAD value to m w.e. and use this as the uncertainty for the mass balance is correct. Because the mass balance is computed as an aggregation of the elevation change, the uncertainty of the volume change needs to be calculated by propagating the uncertainty of a single elevation change pixel through the spatial aggregation — see, for example, Hugonnet et al. 2022 or Rolstad et al. 2009. Because you are using python for some aspects of this, you might want to look at the xDEM python package (<https://xdem.readthedocs.io>), which includes tools for estimating the DEM uncertainty and propagating that uncertainty to volume changes.

Response:

We agree with this comment and thank the reviewer for the specific methodological guidance. The previous approach was incorrect and has been fully revised. We now implement a two-step uncertainty propagation pipeline following Hugonnet et al. (2022) and Rolstad et al. (2009) using the xDEM Python package.

In the first step, we infer a spatially variable per-pixel elevation error map $\sigma(x,y)$ from stable off-glacier terrain using the xDEM function 'infer_heteroscedasticity_from_stable'. This function models error as a function of terrain slope and maximum curvature via a two-dimensional binning approach, capturing the well-documented dependence of DEM error on terrain ruggedness (Hugonnet et al. 2022). Where insufficient stable terrain pixels were available for heteroscedastic modelling (e.g. the CerroBlanco UAV DEM, which has only ~1,200 stable pixels), a spatially uniform error equal to the stable-terrain NMAD was used as a fallback.

In the second step, the spatial correlation structure of elevation errors was characterised by fitting a spherical variogram model to the empirical variogram of off-glacier elevation differences, computed using the Dowd estimator. The uncertainty in mean glacier elevation change $\sigma_{\bar{dh}}$ was then propagated following Rolstad et al. (2009, Eq. 14), which accounts for spatial autocorrelation: $\sigma_{\bar{dh}} = \sigma_{\text{pixel}} \times \sqrt{(A_{\text{corr}} / A_{\text{glacier}})}$, where $A_{\text{corr}} = \pi L^2 / 2$ is the area of spatial correlation and L is the variogram range. Mass balance uncertainty is then derived as $\sigma_{\text{MB}} = \sigma_{\bar{dh}} \times \rho_{\text{ice}} / \rho_{\text{water}} / \Delta t$. This approach substantially reduces uncertainty estimates compared to the previous NMAD-only method. For example, for Glaciar Nevado (Pléiades 2000–2025), uncertainty decreased from ± 0.23 to ± 0.10 m w.e. yr^{-1} , yielding a signal-to-noise ratio of 7.2. The revised methodology is described in full in §3.5 and the code is available in Notebook 03 of the updated GitHub repository.

RC1.3

Were you able to digitize the glacier outlines from the topographic map? I know that not every topographic map will include glaciers as a distinct feature, but given that other studies such as Farias-Barahona et al. (2020) were able to include the glacier outlines digitized from similar maps, I figure it's at least worth asking about. If glaciers were not identified on the map used, this should be mentioned in the appropriate section as an explanation for why it is not included. If it is possible to include the 1954 outlines, I think it's worth re-doing at least part of the analysis by comparing the mass changes computed with these outlines.

Response:

Glacier outlines are not clearly delineated as a distinct feature on the 1954 IGM topographic map, precluding direct digitization of 1954 glacier extents. We have added a sentence to §3.x explaining this. However, we have obtained a 1975 glacier inventory for the Nevados de Chillán complex digitized from Landsat imagery, which provides the earliest available outlines and substantially improves the analysis. The 1975 inventory is now used as the reference outline for all 1954-based epochs, replacing the 2000 outlines used in the submitted version. The 1975 inventory records a total complex area of 16.02 km², compared to 2.91 km² in 2000, representing a loss of 81.8%. Mass balance for 1954-based epochs is now calculated using a time-weighted mean area approach following Cogley et al. (2011), with the 1975 inventory serving as the start-of-period area estimate (the closest available outline to the 1977 midpoint of the 1954–2000 observation period). The use of the 1975 rather than 1954 outlines, and its implications for mass balance estimates, is discussed in §3.x.

Minor and Specific Comments

I. 101–102

I would not characterize these as "images" - the IGM DEM you describe is based on topographic maps which were generated from many images, but what you describe is not using those images. Similarly, the SRTM DEM is not an "image", but rather a DEM.

Response:

Corrected throughout. The IGM 1954 dataset is now described as a topographic map-derived DEM and SRTM is described as a DEM, not an image. Table 1 has been updated accordingly.

I. 145

Why would manual flying impact the accuracy of the DEM?

Response:

Manual flight paths result in suboptimal image overlap (departing from the recommended ~70–80% forward and side overlap), which reduces the number of common tie points available for Structure-from-Motion photogrammetric reconstruction. This decreases DEM accuracy in areas of reduced overlap. The higher stable-terrain NMAD of the CerroBlanco UAV DEM (~12 m) relative to the LasTermas DEM (~5 m) after co-registration is consistent with this explanation, as

the CerroBlanco survey was conducted under more challenging terrain and flight conditions. We have added this clarification to §3.x.

I. 162

Should this be two files, if both points and lines were digitized?

Response:

The reviewer is correct. The digitization produced two shapefiles: one containing spot height points and one containing contour polylines. The contour shapefile was converted to a point file by extracting vertices (using the "Feature Vertices to Points" tool in ArcGIS or the equivalent in QGIS). The resulting vertices shapefile was then merged with the spot heights shapefile into a single combined point file for interpolation. The manuscript has been corrected to reflect this two-stage workflow.

I. 163

What is the density/spacing of the vertices for the polylines? Is this consistent throughout the area digitized?

Response:

The mean distance between adjacent vertices is 36.4 ± 21.5 m (standard deviation), computed as the mean nearest-neighbor distance for contour vertices converted to points. Spacing is therefore variable rather than perfectly uniform, reflecting the curvature of contour lines — denser in areas of complex topography and sparser on gentle slopes. This information has been added to §3.x.

I. 165

Why IDW, rather than a method like Topo to Raster?

Response:

Topo to Raster implements the ANUDEM algorithm, which uses the river network to enforce hydrological consistency in the interpolated surface. While ANUDEM is generally superior to IDW for hydrological modelling applications, its primary advantage — the river network correction — is not appropriate for our use case. The streams in the glaciated zone of Nevados de Chillán vary substantially in shape and position as the glaciers retreat, meaning that the river network assumed by ANUDEM would not be representative of the 1954 landscape being reconstructed. Furthermore, rivers in the immediate vicinity of the glaciers are narrow and produce minimal topographic expression in the contour data, such that the hydrological correction would have negligible effect on the resulting DEM. IDW is therefore appropriate here, and we note that it is consistent with the approach used by Farías-Barahona et al. (2020) for similar topographic map-derived DEMs in the Chilean Andes. We have added a brief justification to §3.x.

I. 171–179

I don't think this is a correct characterization of the approach taken by both Dussaillant et al. (2019) or Hugonnet et al. (2021). Both of those approaches used many ASTER DEMs, but not to create a single composite DEM. There are issues with ASTER DEMs over steep (north-facing) slopes or low-contrast areas such as snow and ice, but unknown/unquantified signal penetration in the SRTM DEM also poses an issue. I am also not sure that characterizing the SRTM C-band DEM as "comparing images taken seconds apart" is correct, given that it is a mosaic of acquisitions over an 11-day period (Rabus et al., 2003).

Response:

We apologise for the inaccurate characterisation and have revised §3.x accordingly. The text previously confused individual ASTER stereo DEMs with the ASTER GDEM composite product; we have corrected this distinction. Both Dussaillant et al. (2019) and Hugonnet et al. (2021) used individual ASTER stereo acquisitions rather than a composite, and we now describe this correctly. The description of SRTM as "comparing images taken seconds apart" has been removed and replaced with an accurate description of the 11-day acquisition period (Rabus et al., 2003). We have also added discussion of known ASTER limitations over steep north-facing slopes and snow/ice surfaces and acknowledged the potential for unquantified C-band signal penetration in SRTM (Berthier et al., 2018). Regarding the SRTM penetration bias specifically: SRTM was acquired in February 2000 at the end of the austral summer ablation season, when glacier surfaces in the ablation zone consist predominantly of wet ice into which C-band penetration is negligible (<1 m; Rignot et al., 2001). Any residual penetration into firn in the accumulation zone would cause a systematic underestimation of mass loss, and our estimates should therefore be considered conservative in this respect. This is consistent with the approach of Farías-Barahona et al. (2020) for nearby Chilean glaciers.

I. 186

Neither Shean et al. 2023 nor Berthier et al. 2024 are included in your references.

Response:

Both references have been added to the reference list.

I. 200, 201, 258

Use MB rather than Mass–Balance; include a citation for the chosen snow/ice density value; Glaciar Nevado.

Response:

Corrected throughout: "MB" replaces "Mass–Balance"; Huss (2013) is now cited for the ice density value of 850 kg m⁻³; "Glaciar Nevado" is used consistently.

Fig. 3

Showing this as a stretched raster, rather than classified, would make it easier to see the local variation in elevation difference. I would also show the elevation difference of off-glacier areas, though this could be done with some transparency.

Response:

We have updated Fig. 3 to use a stretched (continuous) colour ramp rather than a classified scheme, which better conveys local variability in elevation change within individual glaciers. Regarding showing off-glacier elevation differences: we are cautious about displaying off-glacier elevation change in this volcanic setting, as apparent elevation changes in ice-free terrain may reflect volcanic uplift, subsidence, or other non-glaciological processes rather than measurement error. To avoid ambiguity, we have retained the glacier-masked display but reduced the opacity of the off-glacier basemap to provide topographic context. We note this limitation in the figure caption.

I. 290

What is the spatial scale over which the differences are autocorrelated?

Response:

Variogram ranges (correlation lengths) are now reported for each DEM pair in the revised §3.5 and in supplementary Table S1. Ranges vary substantially by DEM type: 882 m for the IGM 1954 (reflecting large-scale systematic errors in digitised topographic maps), 187–280 m for ASTER DEMs (consistent with pixel-scale stereo matching errors), 505–637 m for the UAV DEMs, and 187 m for Pléiades. These values are physically plausible and consistent with previously reported correlation lengths for similar DEM types (Hugonnet et al. 2022).

Fig. 4

Change the labels in the legend to be a line, rather than the outline of a rectangle, using matplotlib.lines.Line2D.

Response:

Updated. All legend handles in Fig. 4 (and other figures using glacier outline overlays) now use matplotlib.lines.Line2D, consistent with the reviewer's suggestion.

I. 300–304

How does the % area change compare to the area-averaged elevation difference for glaciers in different sub-complexes? Did you also compare the area change to the area-averaged elevation/mass change for each glacier?

Response:

We have added a per-glacier analysis comparing area change (2000–2019) to mean elevation change (2000–2025, Pléiades) for each individual glacier in the new §4.x and Fig. X. Spearman correlation between area change (%) and mean elevation change rate was $\rho = -0.018$ ($p = 0.94$, n

= 20 glaciers), indicating no significant relationship between area retreat and surface lowering rate at the individual glacier scale. This decoupling is discussed in §4.x and likely reflects the influence of volcanic geothermal heating on basal melt, which drives ice loss independently of surface area. A hypsometric analysis of elevation change by 50 m elevation band is also presented in Fig. X.

I. 360 and Figs 6–9

Is this meant to be January and February? In Fig. 6-9, it seems as though the trends are only shown if they are significant, and there is no trend shown in Fig. 7a-c. Additionally, include the p-value. Show vertical changepoints on all plots and include the year in the legend. Also show the slope pre- and post-change.

Response:

Following the comments of Reviewer 2 (see RC2 response below), we have removed the E-Divisive changepoint analysis from the hydroclimate section entirely and replaced it with Sen's slope trend lines and Mann–Kendall significance tests. Segmented regression was applied to test for acceleration in trends; it did not converge on a statistically robust breakpoint for any series, confirming that hydroclimatic change is better characterised as monotonic rather than abrupt. Pre- and post-2000 slopes are now shown on all panels where trends are significant, and p-values are reported in the legend for all panels. The analysis has been extended to include March (DJFM rather than DJF) as the most snow-free month and the most glacier melt-relevant for the study area. Non-significant panels are now clearly labelled as such.

I. 400

Why would the single year show a larger magnitude than multiple years?

Response:

We have added an explanation to §4.x. Single-year mass balance estimates are more susceptible to inter-annual variability in snowfall and temperature, which can produce anomalously high or low values relative to multi-year averages that smooth out this variability. This is a limitation of point-in-time geodetic estimates and is now explicitly acknowledged.

I. 407–420

I think you should also compare to Hugonnet et al. (2021) per-glacier estimates and WGMS annual mass change estimates for the same glaciers.

Response:

We have added direct numerical comparisons to Hugonnet et al. (2021) per-glacier mass balance estimates and WGMS annual mass change data in the revised §4.x. Furthermore, as noted in our response to Reviewer 2, we have obtained bias-corrected 30 m ASTER DEMs for 2008 and 2018 from R. Hugonnet directly and integrated them into our analysis pipeline, providing intermediate

epochs between the 2000 SRTM baseline and the 2024/2025 surveys and substantially strengthening our engagement with the Hugonnet et al. (2021) dataset.

I. 418

You could also discuss the value of longer-term local studies like this one — these are comparatively rare worldwide, but especially in the Andes.

Response:

We have added a paragraph to the discussion (§4.x) emphasizing the scientific value of long-term, high-resolution local studies in the context of Andean glaciology, where such records remain rare despite the importance of glacier meltwater for downstream communities.

I. 433–434

Why the disagreement [with previous studies on summertime precipitation]?

Response:

The discrepancy between our finding of no significant trend in summertime (DJFM) precipitation and the regional decreases reported by Garreaud et al. (2020) and Rubio-Álvarez & McPhee (2010) likely reflects differences in spatial scale and record length. Both referenced studies analyzed regional reanalysis or multi-station datasets spanning broader domains and longer periods, whereas our analysis is based on two local stations with records from 1980. At the station scale, DJFM precipitation exhibits substantially higher interannual variability than annual totals due to ENSO forcing, reducing statistical power to detect trends over ~40-year records. We have added a sentence acknowledging that our local finding does not contradict regional-scale drying, which is primarily driven by winter and spring precipitation reductions rather than summer precipitation (Garreaud et al., 2020).

I. 448

Can you report the surface area change here?

Response:

We have added glacier surface area values to this sentence: the complex declined from 16.02 km² in 1975 to 2.91 km² in 2000 and 1.90 km² in 2019, representing an 88% reduction over the study period.

I. 518

I have to think this would be somewhat negligible compared to the uncertainties in the SRTM, which are typically reported as being at least 5 m or so.

Response:

We agree with the reviewer's assessment. The potential contribution of volcanic uplift to apparent ice surface lowering is negligible relative to the dominant SRTM uncertainty (± 5 –16 m; Rodriguez et al., 2006), and geodetic uplift rates at active volcanic systems rarely exceed a few centimeters per year during non-eruptive periods. We have simplified the text at l. 518 to acknowledge this source of uncertainty briefly without overstating its magnitude.

Reviewer 2 (Liam Taylor)

Comment received: 7 May 2026

General Comments

RC2.1

The large number of datasets used in this study means that presenting a coherent story about the analysis is not an easy task. I think that the authors could benefit from understanding what their analysis can and cannot say. Most notably, care should be given to how the polygons (from the year 2000 or 2019) are used in analysis of 1954 datasets – I note, for instance, that the 2000 polygons were used to derive stable terrain, but this would inevitably include ice melted from 1954-2000 (and thus, the terrain would not be 'stable'). As glacier area change also cannot be included in the analysis, is it more accurate to report the overall figures as ice elevation change rather than mass balance?

Response:

We thank the reviewer for this important methodological concern. To address it, we have incorporated the 1975 glacier inventory as an additional mask during co-registration, constructing a conservative union of the 1975, 2000, and 2019 glacier outlines. This substantially reduces the inclusion of formerly glaciated terrain as "stable" ground during co-registration of the 1954 IGM DEM. We acknowledge that terrain glaciated in 1954 but deglaciated by 1975 remains unmasked; however, given the magnitude of area loss between 1975 and 2000 (81.8%), we consider the 1975 inventory to be a reasonable conservative proxy. The use of the 1975 inventory also allows us to report results as specific mass balance (using a time-weighted mean area approach following Cogley et al. 2011) rather than elevation change alone, which we consider more physically meaningful. We are open to the reviewer's feedback on whether this addition is considered sufficient to justify reporting mass balance rather than elevation change.

RC2.2

I'm also not sure that the change point analysis is the appropriate statistical test for a time series where you are not expecting to see a system change, but rather a linear trend. Care with regard to this analysis would improve confidence in the study.

Response:

We agree with the reviewer and have removed the E-Divisive changepoint analysis entirely. As the reviewer notes, distributional changepoint methods assume sudden regime shifts, which is not appropriate for gradual anthropogenically-driven climate trends subject to ENSO variability.

Hydroclimatic trends are now characterised solely using Sen's slope estimator and Mann–Kendall significance tests, which are appropriate for monotonic trends in non-normally distributed time series. To confirm the absence of a statistically robust breakpoint, we additionally applied segmented (piecewise linear) regression; this did not converge on a significant breakpoint for any series, confirming that the observed changes are better described as continuous monotonic acceleration than abrupt regime shifts. Pre- and post-2000 Sen's slope estimates are reported to enable comparison with the glacier mass balance record, which has its baseline at 2000.

RC2.3

There are also opportunities that could improve the study. The analysis to 2024 includes 11 glaciers due to practical considerations with the UAV survey. However, you could make use of secondary datasets – e.g. Hugonnet et al. (2021) – to report elevation thinning to 2020 for all glaciers. I pose this suggestion only as an idea for the authors to consider; the study is complete without this extra analysis.

Response:

We thank the reviewer for this suggestion. We contacted R. Hugonnet directly and he generously shared his bias-corrected 30 m ASTER DEMs for our study tile. After quality filtering (stereo correlation threshold ≥ 60 following Hugonnet et al. 2021, spatial outlier removal vs SRTM, snow bias assessment, and a minimum 30% glacier coverage threshold), two scenes were retained: March 2008 and March 2018. The 2018 ASTER result (-0.506 m w.e. yr^{-1} for the whole complex) matches our Pléiades 2025 result almost exactly, providing strong independent validation of both datasets. We additionally include a Pléiades 2025 stereo DEM built by co-author Dr J. Berkhoff (FAU Erlangen-Nürnberg) that covers the full complex at 1 m native resolution, extending the full-complex record to 2025. Together these additions substantially increase temporal resolution between the 2000 SRTM baseline and the 2024 UAV surveys.

Specific Comments

I. 21–22

Report change from 1954–2000 and 2000–2024 to demonstrate that there is a clear acceleration between the two time periods, rather than the overlapping window.

Response:

Revised. The abstract and introduction now report non-overlapping periods (1954–2000 and 2000–2024/2025) to clearly demonstrate acceleration in mass loss, rather than the overlapping window used previously.

I. 36

Suggested rephrase to "The relative contribution of glacier melt is..."

Response:

Agreed, sentence rephrased accordingly.

I. 39

What do "small" and "significant" mean in this context? Quantifying will help to sell the importance of glaciers in this region.

Response:

We have replaced the qualitative descriptors with quantitative values in the revised text, reporting specific glacier area, meltwater contribution estimates, and downstream population exposure where data are available.

I. 42, 49, 50, 73, 81

"glaciated" > "glacierized"; repetition of "Fig. 1B"; provide values in km².

Response:

"Glacierized" used throughout (I. 42 and I. 534). Repeated "Fig. 1B" reference deleted (I. 49). Glacier areas now reported in km² at I. 50, 73, and 81.

Figure 2

Why not also produce mass balance calculations for 2000–2024 for the glaciers in the UAV survey? Note: this Fig references 18 glaciers in the UAV survey, but in the abstract this was 11 glaciers?

Response:

Mass balance for 2000–2024 was computed for the UAV-surveyed glaciers in the submitted version (reported in §4.x and Fig. 5); Figure 2 is a methodological flowchart and did not include this step, which was an omission. We have updated Figure 2 to include the 2000–2024 mass balance calculation step explicitly. The reference to 18 glaciers in Figure 2 was a labelling error from an earlier version of the analysis. The UAV survey covered 11 glaciers: 1 in the Cerro Blanco subcomplex (Glaciar Nevado, CL108116004) and 10 in the Las Termas subcomplex, based on the DGA 2000 glacier inventory. This is consistent with the abstract and has been corrected in Figure 2 and throughout the manuscript.

I. 157

Did you perform any vertical transformation or georeferencing to ensure that the PSAD56 datum was comparable to the vertical SRTM and UAV-based datums?

Response:

Yes — see response to RC1.1 above. We now apply pixel-wise EGM96 geoid corrections to all ellipsoidal DEMs (UAV, Pléiades, ASTER) prior to co-registration and differencing. The PSAD56 datum of the IGM DEM approximates EGM96 to within ~1 m at 37°S, which we consider negligible given the IGM DEM's overall uncertainty. This is described in the revised §3.x.

I. 175

While true, have you considered and accounted for the ice penetration bias that might be present in SRTM data (see Gardelle et al. 2017)?

Response:

See response to RC1 I. 171–179 above. C-band penetration into wet ice is negligible (<1 m; Rignot et al., 2001), and SRTM was acquired during the late austral summer ablation season (February 2000) when glacier surfaces are predominantly wet ice. Any residual penetration into firn would cause a conservative (underestimated) mass loss signal. This limitation is now acknowledged in the revised manuscript.

I. 189, 195

Is there a risk that by using polygons from the year 2000 to co-register a DEM from 1954, that you are unintentionally adding in unstable terrain? Do you have glacier boundaries from 1954?

Response:

See RC2.1 response above. We do not have 1954 glacier boundaries (outlines are not identifiable as a distinct feature on the IGM topographic map) but now incorporate the 1975 inventory to substantially reduce this issue. The 1975 outlines are used in the glacier union mask for co-registration and as the reference area for mass balance calculation in 1954-based epochs.

I. 211

Where do these control points come from?

Response:

Ground control points used in UAV photogrammetric processing were collected using differential GNSS (RTK) in the field during the 2024 survey campaigns. We have added a sentence to §3.x clarifying the source and number of GCPs used for each UAV survey.

I. 221

It is not necessary to divide by the number of years for the error estimates, as you aren't assessing change over time of the stable terrain, but rather identifying what the difference of this (non-changing) terrain is and using that as the error.

Response:

Agreed. The previous approach of annualising the NMAD was incorrect and has been removed. We now report stable-terrain NMAD as a direct measure of elevation error, and propagate this through the spatial aggregation using the Rolstad et al. (2009) equation as described in RC1.2. The error is not divided by the number of years.

I. 231

Why not use the average temperatures from the dataset rather than manually deriving from minimum and maximum values?

Response:

We have updated the temperature analysis to use daily mean temperatures computed directly from the dataset, rather than averaging minimum and maximum values. The revised results are consistent with the submitted version, confirming that this change does not materially affect our conclusions. The methods text has been updated accordingly in §3.6.

I. 287

Topographic variables. Which statistical test did you use here?*

Response:

Corrected to "topographic variables" throughout. We now report both Spearman rank correlation (primary, as our elevation change distributions are non-normal) and Pearson correlation for all glacier-level and pixel-level topographic analyses. The statistical test is now explicitly stated in §3.x and §4.x.

Fig. 5

Is the bar missing for Las Termas in the 2000 outlines? Better to plot 1954–2000, then 2000–2024, then 1954–2024.

Response:

The missing bar was a plotting error and has been corrected. The figure has been reordered to show 1954–2000, 2000–2024/2025, and 1954–2024/2025 sequentially, which more clearly demonstrates the acceleration in mass loss between the earlier and later periods.

I. 409, 425, 430

Also Hugonnet et al. (2019); can you describe (quantify) these trends and similarities to your data?; topographic variables.

Response:

Hugonnet et al. (2019) reference added. Trends and similarities with our mass balance results are now quantified in §4.x. "Geographic" corrected to "topographic" throughout.

I. 445–454

This section about peak water having passed could be strengthened with further reference to previous studies that have quantified this effect in the region – e.g. McCarthy et al. (2022) or Ragettli et al. (2016).

Response:

We have added references to McCarthy et al. (2022) and Ragettli et al. (2016), which provide regional context for peak water in central Chile, and have expanded the discussion of peak water to situate our streamflow findings within this broader literature.

I. 465

Is it not possible to identify this from UAV images?

Response:

Distinguishing debris-covered ice from recently deglaciated terrain in UAV imagery is challenging at this site due to similar spectral and textural properties. We are developing machine learning approaches for automated classification in an upcoming study, and have noted this in the revised manuscript.

I. 475

Given that glacier area change isn't reflected in the calculations, is it more accurate to report glacier thinning estimates than mass balance change estimates?

Response:

See RC2.1 response above. The incorporation of the 1975 inventory and the use of a time-weighted mean area approach following Cogley et al. (2011) substantially addresses this concern. Glacier area change is now explicitly accounted for in the mass balance calculation for all epochs, and we consider reporting as specific mass balance (m w.e. yr⁻¹) to be appropriate given these revisions.