



Supplement of

Modeling the impacts of climate trends and lake formation on the retreat of a tropical Andean glacier (1962–2020)

Tal Y. Shutkin et al.

Correspondence to: Tal Y. Shutkin (shutkin.3@osu.edu)

The copyright of individual parts of the supplement might differ from the article licence.

Sect. S1.1 DEM Coregistration

The 1962 DEM was coregistered to that of 2008 using ESRI ArcPro georeferencing tools. Six control points were identified through visual comparison of stable features in the DEMs and used to calculate a first-order transformation of the 1962 DEM, aligning it with the later epoch. Stable terrain was then selected by masking out pixels covered by ice in the 1962 imagery and those above 5000 m altitude (Fig. S1). The latter adjustment accounts for systematic error along ice-free ridge lines due to DEM resolution differences. Stable terrain residuals were then calculated by subtracting the elevation values of the masked DEMs. To identify further biases between the elevation datasets, residuals were compared against topographic variables including aspect, slope, altitude, length (the distance across the long axis of the DEMs), and width (the distance across the short axis of the DEMs) (Fig. S2). Linear regressions showed that the length variable had the most significant impact on residuals, explaining 15% of the variability in residuals ($p < 0.0001$). This indicates a systematic tilt over the short axis of the 1962 DEM with respect to that of 2008. We therefore apply a Z-dimension correction to the 1962 DEM using the parameters of the linear regression, then recalculate elevation model residuals over stable terrain. After the second adjustment, R-squared values for all explanatory variables are reduced below 0.10 ($p < 0.0001$), indicating that any systematic bias in glacier volume change would be negligible (Fig. S2). After coregistering the 1962 DEM to that of 2008, residual error between stable ground pixels of the two DEMs is significantly reduced. The original mean error with a one standard deviation window falls from 33 ± 29 m to 0 ± 12 m. We further compare our stable ground residuals to those generated using the common Nuth and Kääb, (2011) algorithm, finding that our methods are more robust (lower standard deviation of residuals) than those achieved using the standard approach (Fig. S3). As seen in Fig. S3, residuals over stable ground are normally distributed with a mean of approximately 0 m and a standard deviation of about 12 m. Systematic (e.g., aspect-related) bias is minimized, though may be evident locally near the southeast corner of the model domain. Note that this localized issue was persistent across coregistration methods, including when using the Nuth and Kääb (2011) algorithm. The large positive residual situated off glacier is due to an artifact in the 1962 DEM wherein a peak is represented as a plateau. This peak is located above 5000 m, however, so did not bias the DEM coregistration process.

Having minimized systematic error through DEM coregistration, we follow Hugonnet et al., (2022) to evaluate random error over stable terrain and infer uncertainty in elevation change over the 2008 glacier area. Random error is quantified by considering both the heteroscedasticity and spatial correlation of error and is assumed to derive predominantly from error or artifacts in the 1962 dataset. Heteroscedasticity is evaluated across gradients of DEM slope and curvature, calculated using methods from Horn (1981) and Zevenbergen and Thorne (1987), respectively. The spatial correlation of error is estimated by an empirical variogram using Dowd's estimator (Dowd, 1984). The uncertainty in elevation change within the 2008 glacier boundary is then calculated as the average pairwise product of pixel uncertainties times the spatial correlation of error between each two pixels (Hugonnet et al., 2022 eqs. 17-19).

Implementing these methods using the xDEM Python package (xDEM contributors, 2024), we arrive at a mean elevation change of -22.61 ± 0.81 m across the glacier (Fig. S4). Maintaining the density assumption of 900 kg m^{-3} as used throughout the manuscript, this translates to a geodetic mass balance of $-442 \pm 16 \text{ mm w.e. a}^{-1}$. The change in elevation across the entire DEM including stable and unstable terrain is shown in Fig. S4. We consider the level of uncertainty arrived at using these methods to be acceptable for the purpose of our study. Since most error over stable ground (and therefore elevation change uncertainty) is attributable to artifacts and errors in the 1962 DEM, we also consider the acceptable uncertainty range as testament to the adequacy of our 1962 dataset.

As an additional quality assurance step, we identify the presence of outlying values (95th percentile) in the difference of DEMs that may impact the geodetic mass balance calculation (Piermattei et al., 2024). The outlier detection procedure followed three steps. First, elevation change data within the 2008 glacier boundary were binned according to their positioning in 50 altitude bins (~16 m) according to the 2008 DEM. Second, pixel values for surface height change were compared to the mean value of each altitude bin. Pixel values with z-score absolute values greater than 1.96 (two-tailed 95th percentile) were considered to be outliers resulting from DEM or coregistration errors and were replaced by the mean value from the appropriate altitude bin. Finally, the resulting map of change in surface height was used to recalculate the specific (area averaged) mass balance across the entire glacier surface.

After removing outliers, the new SMB was calculated to be -435 mm w.e. per year, a 1.5% positive change from the original figure. Maps depicting the original DEM difference, altitude-binned averages, and detected outliers are included below (Fig. S5). We note that the difference in specific mass balance is within the uncertainty window estimated in the previous step and that it is difficult to distinguish between outliers caused by map artifacts versus extreme natural phenomena. Furthermore, previous glaciological studies using outlier detection and gap filling have operated with very different data constraints. For example, Piermattei et al. (2024) use ASTER and TanDEM-X data, both of which are known to contain artifacts and data voids in mountainous regions. Geodetic mass balance estimation using these global datasets therefore may require outlier correction. In our case using local datasets without issues such as cloud cover, this requirement is less apparent. Given this ambiguity and the negligible impact of outlier correction on the resulting geodetic mass balance estimation when compared to the overall uncertainty, we proceed with the original difference of DEMs and its associated specific mass balance.

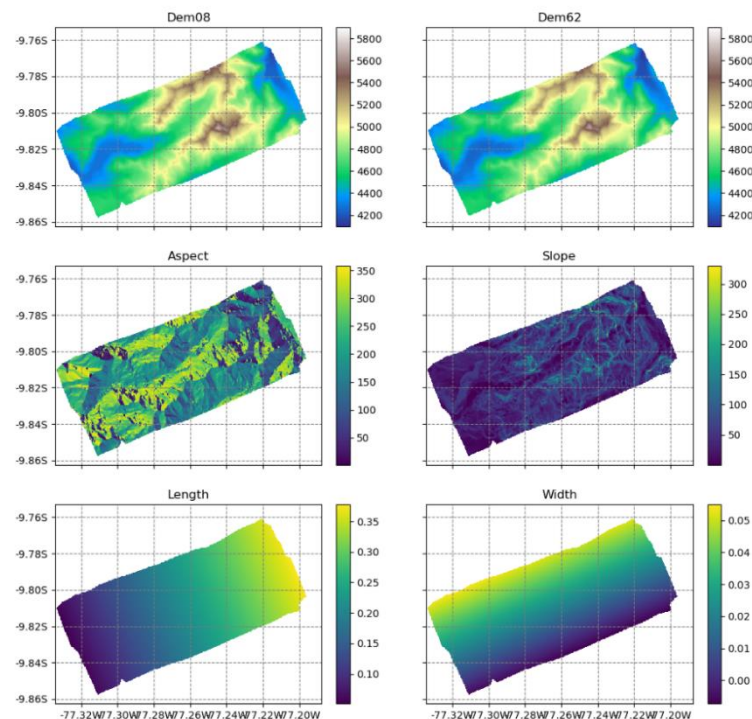


Figure S1: Input variables for DEM coregistration. Note than length and width area calculated using indices and have no units.

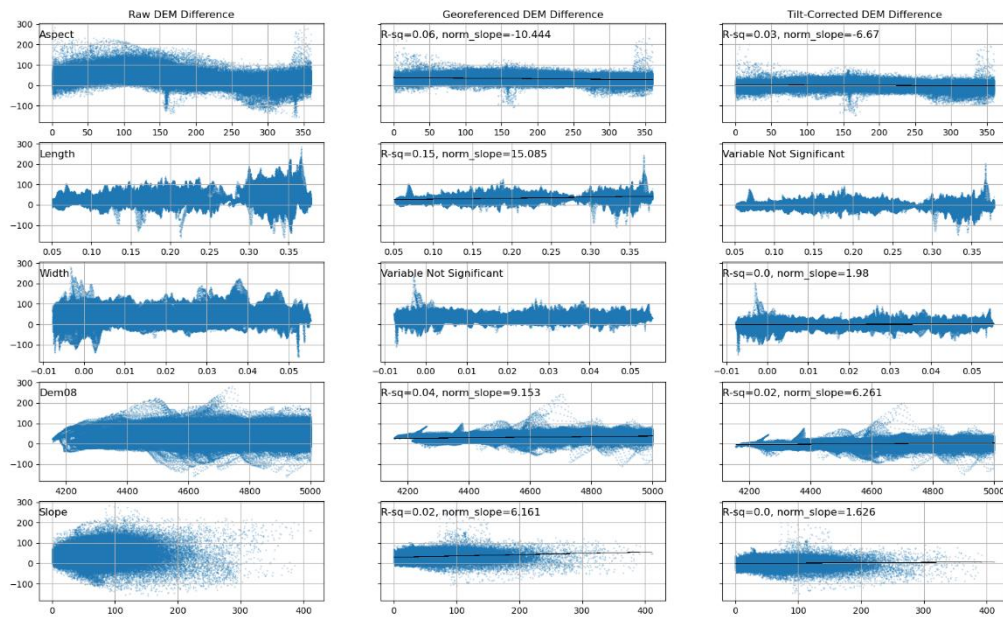


Figure S2: Matrix plot of coregistration variables (rows) and coregistration stages (columns). Variable names are listed in the first column. Each point on a given subplot corresponds to a single stable-ground pixel pair. The x-axes indicate the magnitude of a given variable (e.g. 0-360° for aspect) and the y-axes indicate the residual error in meters. For variables with significant R-squared values relating the magnitude and residual error, the R-squared and normalized slope (norm_slope) are included on the subplot. The normalized slope considers the change in residual error predicted for a unit change in the coregistration variable.

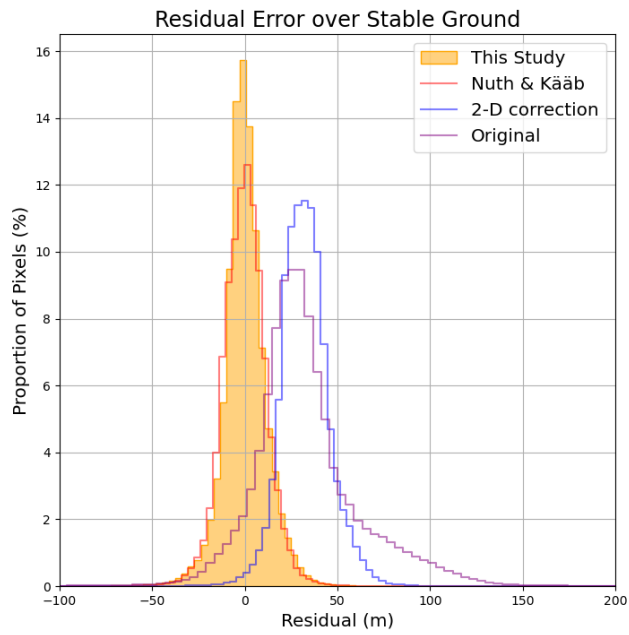


Figure S3: DEM coregistration stable ground residuals. The three stages are shown: original DEM difference prior to georeferencing (purple); the DEM difference after georeferencing in two dimensions (blue); and the DEM difference after applying a linear correction to the elevation values along the length of the georeferenced DEM (solid yellow). Residuals achieved using the Nuth and Kääb (2011) algorithm (red) are shown for comparison.

70

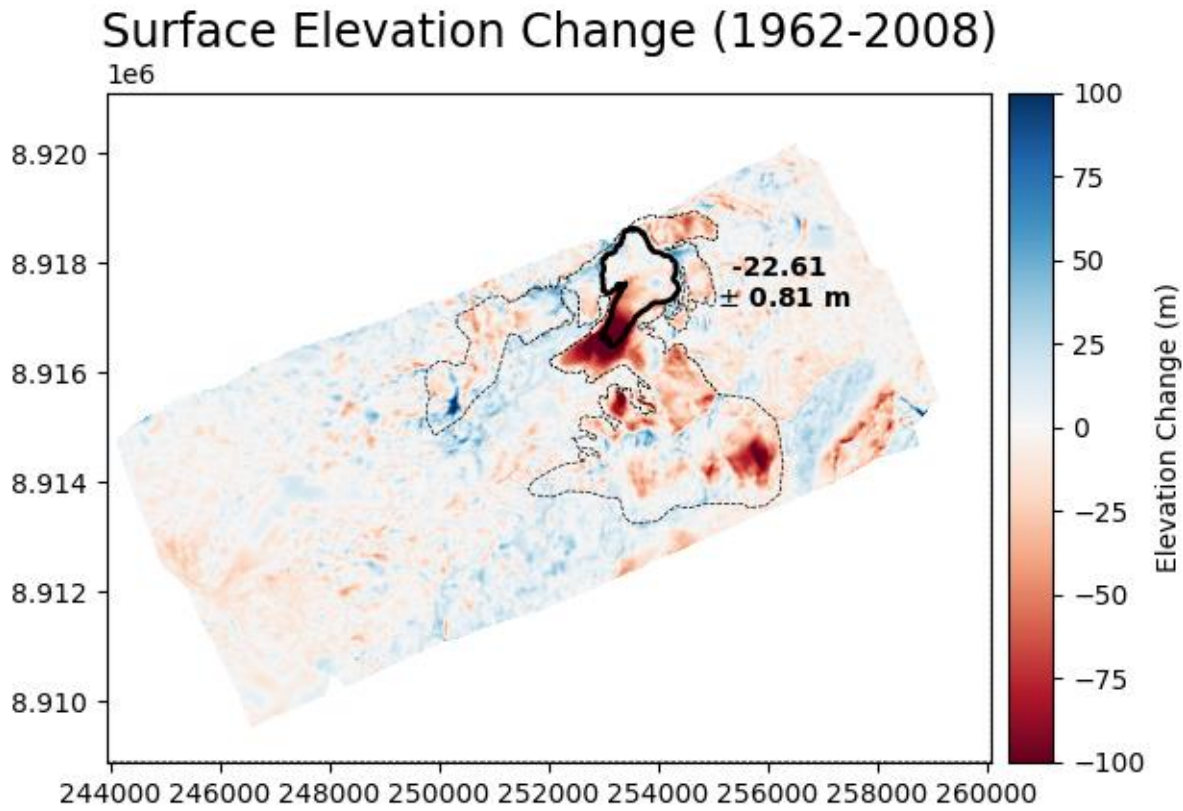


Figure S4: Difference of DEMs (1962-2008) shows significant surface height change across the glacier ablation zone. The 2008 glacier boundary is outlined in bold and unstable terrain (including the 1962 glacier boundary) is delineated by a dashed line. Note that additional terrain above 5000 m in 2008 was also considered to be unstable.

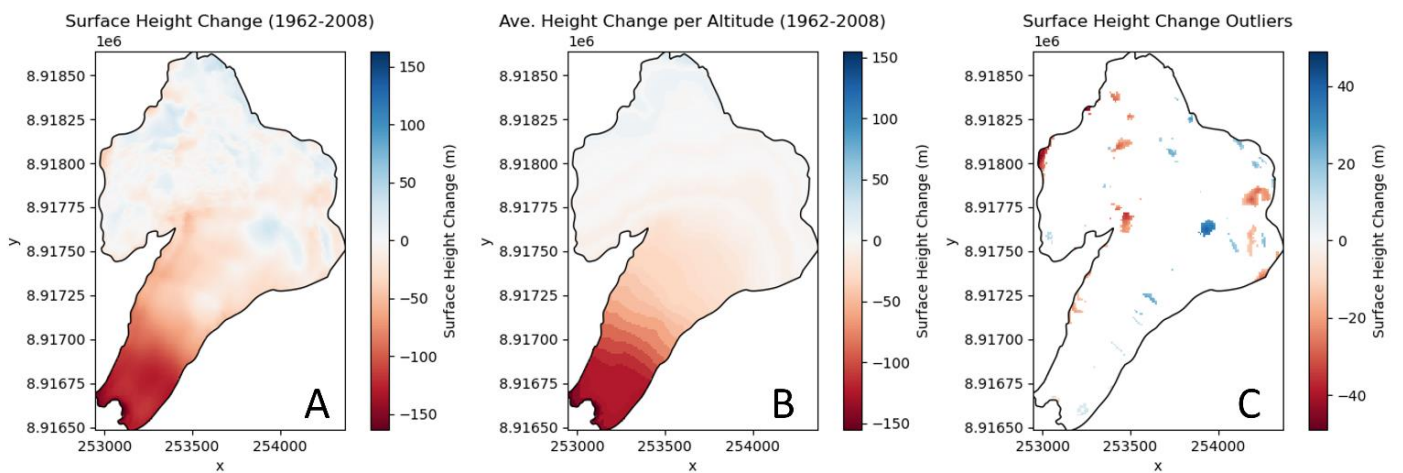


Fig S5: Outlier detection began from the map of elevation difference (a), which was then binned into 50 elevation bands (b). Pixel values exceeding 2σ deviation from the elevation band mean were considered outliers (c).

Sect. S1.2 Ground Penetrating Radar Survey and Validation

A detailed map of 2008 ice thickness was derived by subtracting the 2014 basal topography from the 2008 LiDAR Dem surface. This is presented in Fig. S6. We validate these derived thicknesses against a minimal GPR survey from 2009.

Only 17 georeferenced point measurements were provided by the 2009 GPR survey report, which was used for validation purposes only. As elaborated upon in the main text, the points show general consistency with the subsequent 2014 survey (Fig. S7a). The 2009 GPR points span from the bottom of the glacier in the southwest towards the center of the glacier in the northeast of Fig. S7b. The points are located approximately along the centerline of the glacier and are each in proximity to multiple measurements from the subsequent GPR survey. There is no apparent relationship between the mean distance from the 2009 data points to their respective nearest neighbors and the resulting difference between measured and derived thickness. There does, however, appear to be a slight spatial bias, with derived thicknesses being more likely to underestimate the 2009 measurements at lower elevations. The significant outlier where the derived thickness is approximately 28 m thinner than observed occurs at a discontinuity in the 2009 survey, suggesting that the technician may have needed to navigate an obstacle which may have produced abnormalities in the ice thickness profile or potentially caused an error in measurement or radargram interpretation.

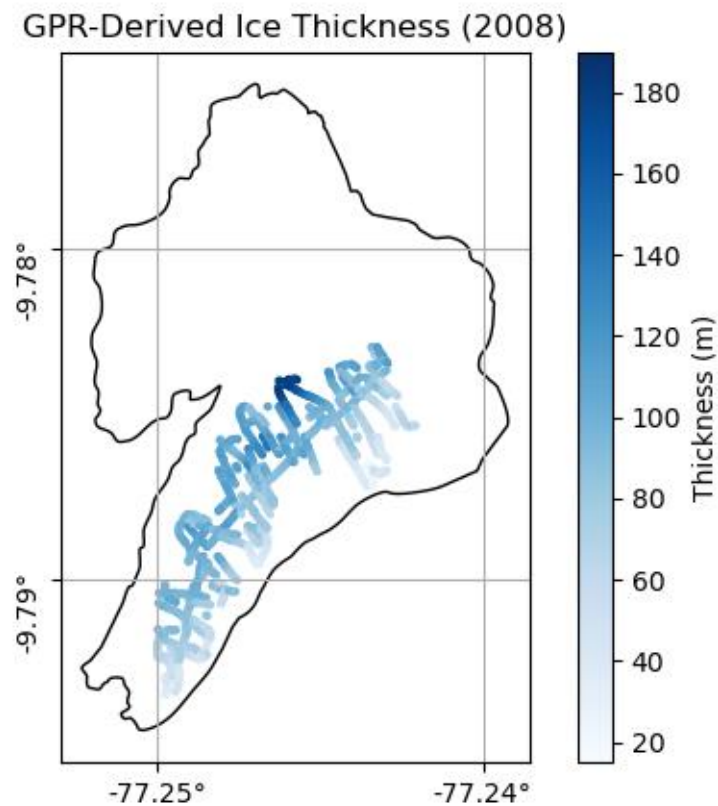


Figure S6: Ice thickness map representing the year 2008. Thickness in 2008 was derived by subtracting basal topography as measured in the 2014 GPR survey from the 2008 LiDAR DEM surface heights.

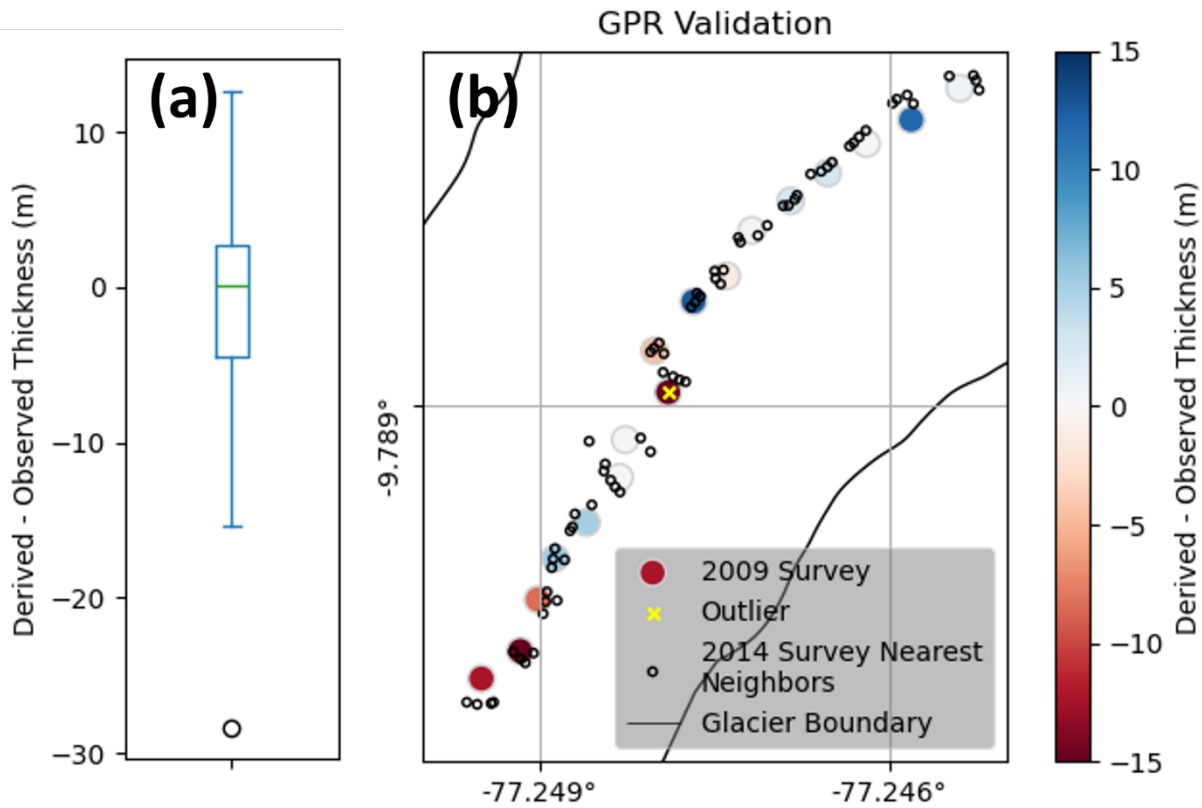
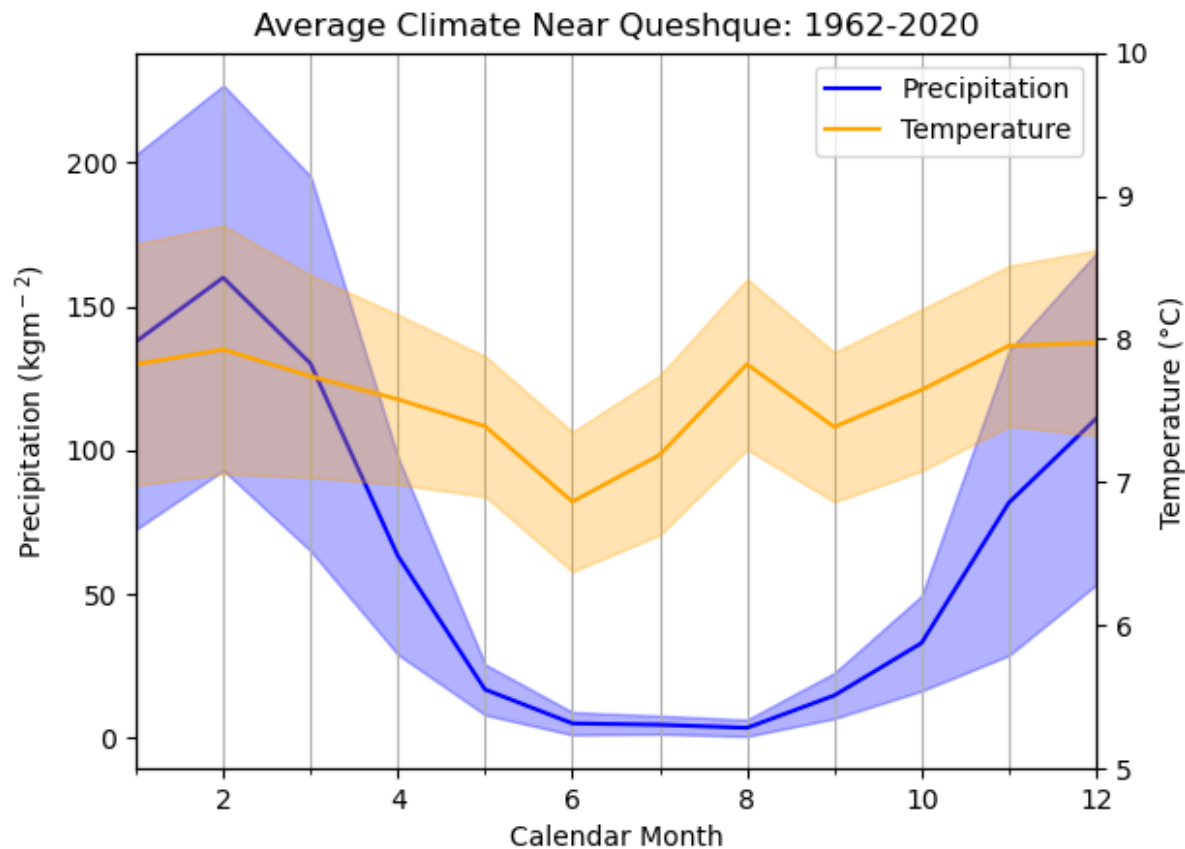


Figure S7: (a) Box and whiskers plot indicating the variability in derived 2008 ice thickness error as compared to measurements from 2009, indicating the presence of a single negative outlier. (b) A map illustrating the GPR validation process. Validation data from the 2009 GPR survey are depicted by colored circles. Deeper red indicates that the four nearest neighboring points (small white circles) from the 2014 GPR survey increasingly underestimate the 2009 observed thickness. Deeper blue represents the opposite. The single outlier is identified by a yellow X and the edges of the glacier tongue are drawn in thin black lines, indicating that the validation survey ran roughly along the glacier's central flowline.



115 **Figure S8: Climatology near Queshque Glacier during the period 1962 through 2020. Mean temperature (2 m) and precipitation are depicted with 1 σ bounds as shaded regions. Climatology data combine CRU (Harris et al., 2014; New et al., 2002) and PISCO (Aybar et al., 2020; Huerta et al., 2023) products. Note that the values shown here are subsequently adjusted by the precipitation factor and temperature bias within the mass balance model (see section 3.5.1).**

Sect. S1.4 Surface Velocity Validation

120 Mean elevation-binned surface velocities were compared to observations representing the year 2018 using RMSE and MAE metrics (Table S1). Results indicate models 2, 3, and 4 are most representative of realistic local ice dynamics.

Model Number	Model Temp Bias (°C)	RMSE (ma ⁻¹)	MAE (ma ⁻¹)
1	-6.5	9.8	4.9
2	-7.0	3.5	2.5
3	-7.5	2.7	2.2
4	-8.0	3.8	3.2
5	-8.5	5.6	5.2
6	-9.0	7.6	6.9

Table S1: RMSE and MAE values of elevation band flowline surface velocities versus observations from Millan et al. (2022). Model temperature biases are included for reference.

Sect. S1.5 Glacier Terminus Positions

125 Glacier terminus positions were mapped from aerial, Landsat 8, and Sentinel 2 imagery to construct a timeseries spanning 1962-2023 (Fig. S9). We used the Zonal Statistics as Table tool in ArcGIS Pro to extract mean and standard deviation elevations from

the 2008 DEM where it intersects with mapped termini. The resulting values are presented in Fig. S10. Note, however, that that these calculated elevations reflect the surface height in 2008, not the actual terminus elevation of a given year. For this reason, standard deviations for years prior to 2008 (ice-free in the 2008 DEM) are much greater than for years where ice is present in the elevation dataset. This reflects that glacier ice has a lower surface roughness than its surroundings.

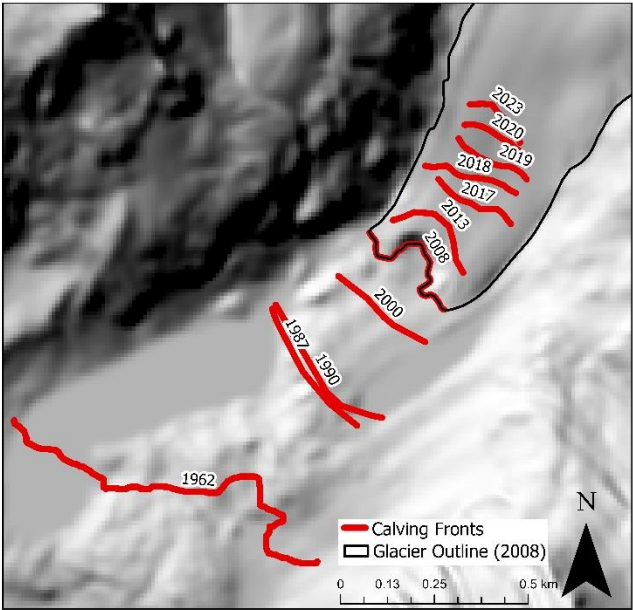


Fig. S9: The mapped glacier termini overlaid on the 2008 DEM hillshade.

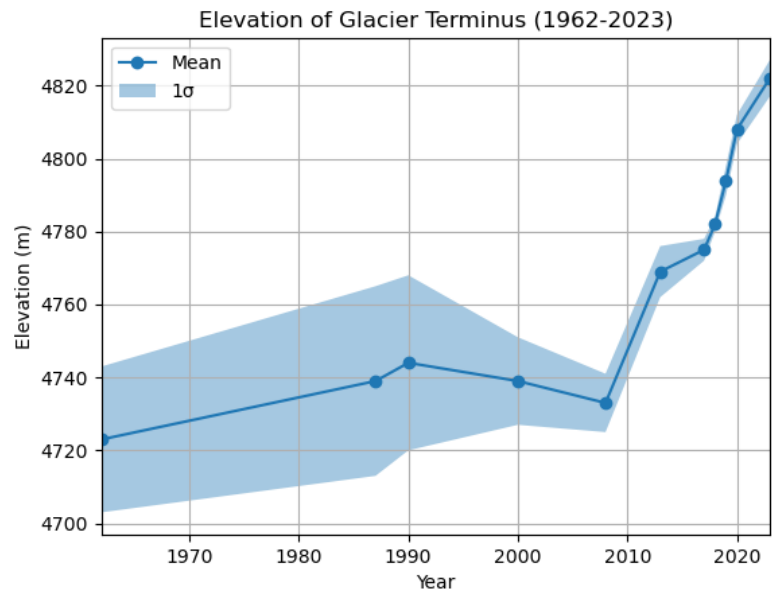


Fig. S10: The glacier terminus elevations for each observation year derived from the map in Fig. S9 by averaging the DEM elevations where they overlap the mapped terminus of a given year. The shaded region represents one standard deviation in the DEM values.

Sect. S1.6 Mass Balance Gradient Sensitivity Experiment

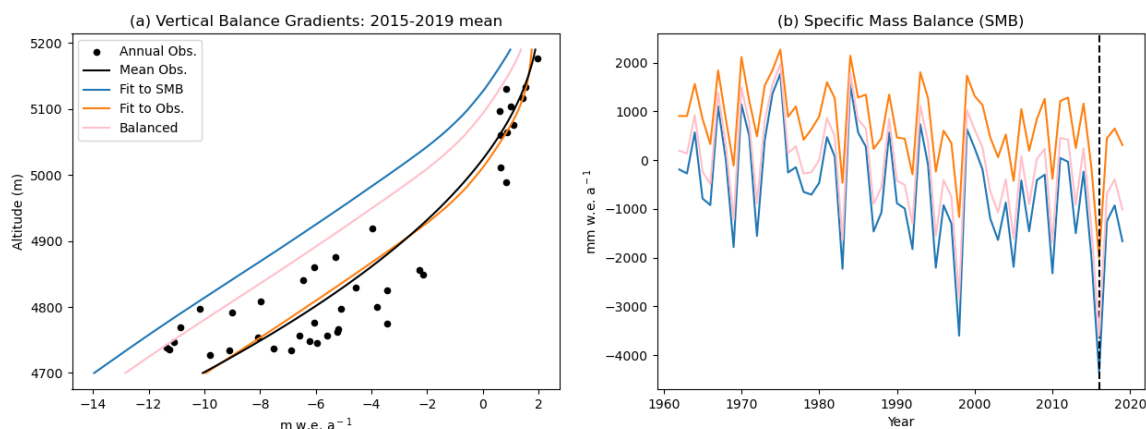


Fig. S11: (a) The observed and modeled vertical balance gradients (mass balance profiles) fit to the observed SMB (blue), the observed mass balance profile (orange), and to a glacier in long-term (1962–2008) climatic equilibrium (pink). (b) Annual SMB from the same models, 1962–2020. The dashed line represents the strong El-Niño year of 2016, which is the only year that observed ablation rates (points left of pink curve in Fig. 4a) would produce negative mass balance across the glacier.

Sect. S1.7 Climatological Trend Analysis

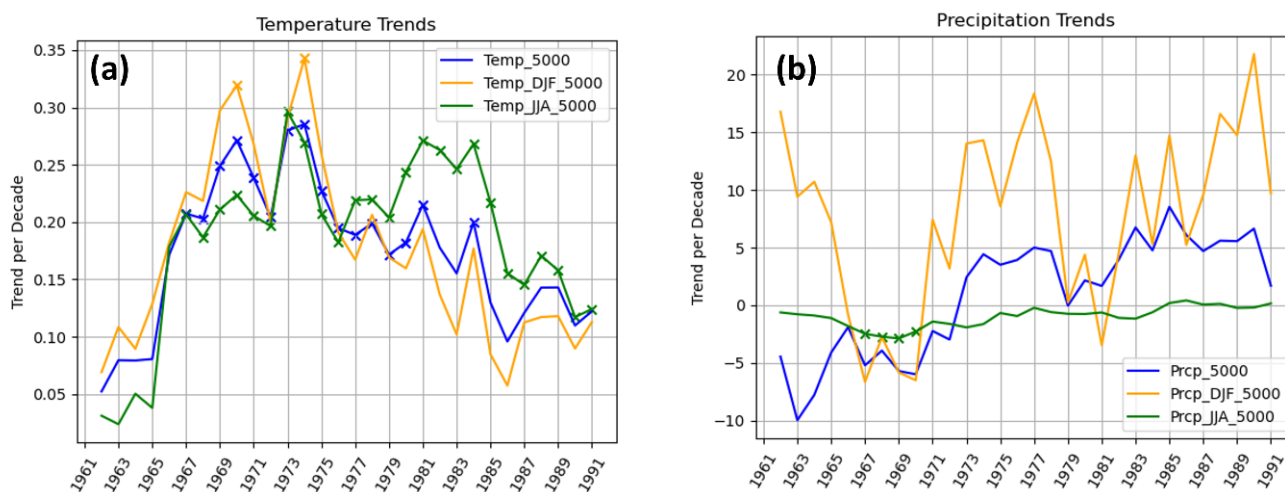


Figure S12: Trends per decade in the extended PISCO dataset for temperature (°C decade⁻¹) (a) and precipitation (mm decade⁻¹) (b) at the assumed altitude of 5000 meters within the grid cell containing the Queshque Glacier centroid. Trends are calculated in running 30-year periods for annual values, wet season (DJF) and dry season (JJA). Significant trends are marked with crosses and correspond to the 30-year period beginning at the given year.

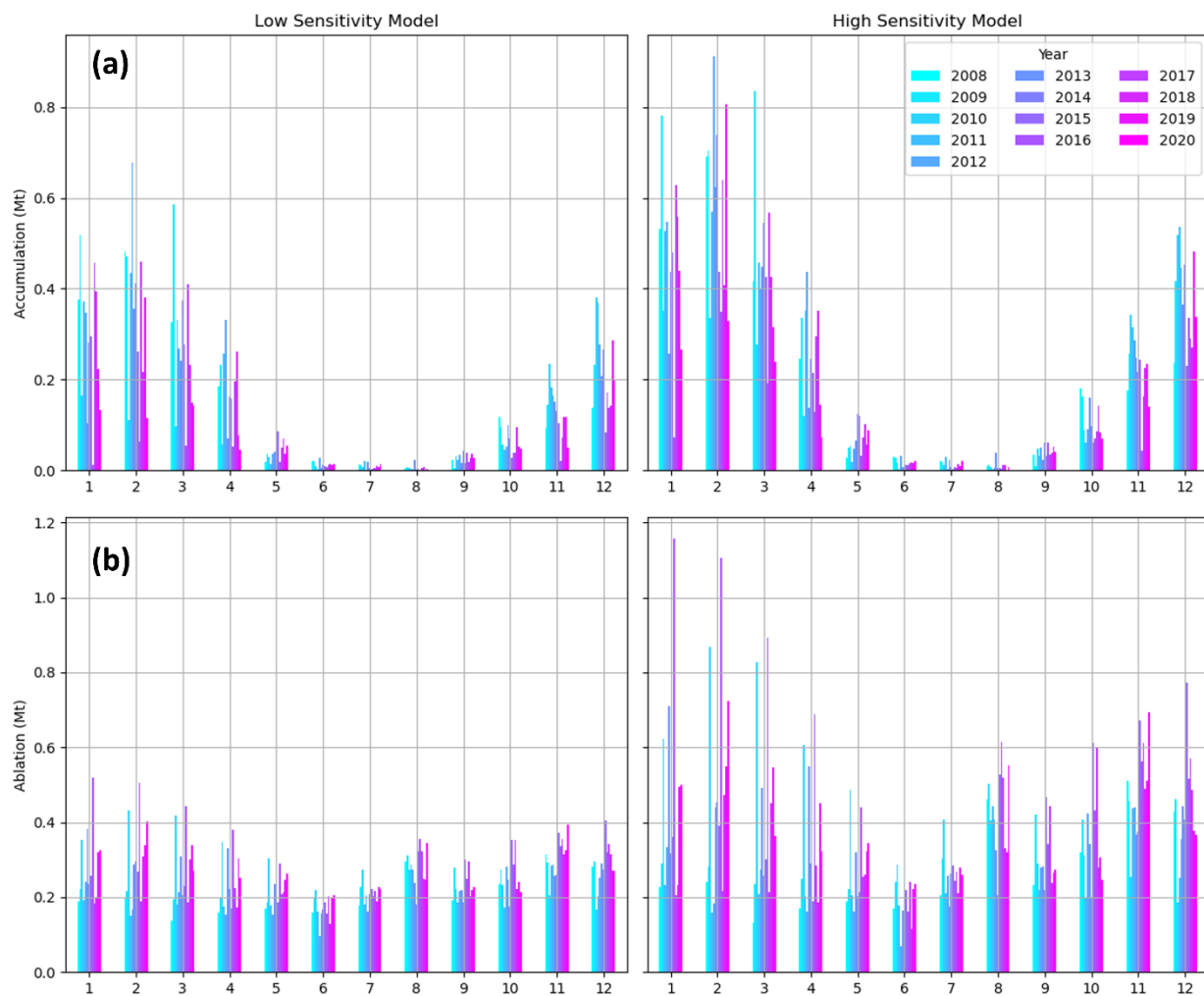


Figure S13: Model accumulation (a) and ablation (b) for each month of each year (2008-2020). The low (left) and high (right) sensitivity models correspond to model numbers 2 and 5, respectively.

References

- Aybar, C., Fernández, C., Huerta, A., Lavado, W., Vega, F., and Felipe-Obando, O.: Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day, *Hydrol. Sci. J.*, 65, 770–785, <https://doi.org/10.1080/02626667.2019.1649411>, 2020.
- 160 Dowd, P. A.: The Variogram and Kriging: Robust and Resistant Estimators, in: *Geostatistics for Natural Resources Characterization: Part 1*, edited by: Verly, G., David, M., Journel, A. G., and Marechal, A., Springer Netherlands, Dordrecht, 91–106, https://doi.org/10.1007/978-94-009-3699-7_6, 1984.
- Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, *Int. J. Climatol.*, 34, 623–642, <https://doi.org/10.1002/joc.3711>, 2014.
- 165 Horn, B. K. P.: Hill shading and the reflectance map, *Proc. IEEE*, 69, 14–47, <https://doi.org/10.1109/PROC.1981.11918>, 1981.
- Huerta, A., Aybar, C., Imfeld, N., Correa, K., Felipe-Obando, O., Rau, P., Drenkhan, F., and Lavado-Casimiro, W.: High-resolution grids of daily air temperature for Peru - the new PISCOT v1.2 dataset, *Sci. Data*, 10, 847, <https://doi.org/10.1038/s41597-023-02777-w>, 2023.
- 170 Hugonnet, R., Brun, F., Berthier, E., Dehecq, A., Mannerfelt, E. S., Eckert, N., and Farinotti, D.: Uncertainty Analysis of Digital Elevation Models by Spatial Inference From Stable Terrain, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 15, 6456–6472, <https://doi.org/10.1109/JSTARS.2022.3188922>, 2022.
- New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, *Clim. Res.*, 21, 1–25, <https://doi.org/10.3354/cr021001>, 2002.
- 175 Nuth, C. and Kääb, A.: Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change, *The Cryosphere*, 5, 271–290, <https://doi.org/10.5194/tc-5-271-2011>, 2011.
- Piermattei, L., Zemp, M., Sommer, C., Brun, F., Braun, M. H., Andreassen, L. M., Belart, J. M. C., Berthier, E., Bhattacharya, A., Boehm Vock, L., Bolch, T., Dehecq, A., Dussailant, I., Falaschi, D., Florentine, C., Floricioiu, D., Ginzler, C., Guillet, G., Hugonnet, R., Huss, M., Kääb, A., King, O., Klug, C., Knuth, F., Krieger, L., La Frenierre, J., McNabb, R., McNeil, C., Prinz, R., Sass, L., Seehaus, T., Shean, D., Treichler, D., Wendt, A., and Yang, R.: Observing glacier elevation changes from spaceborne optical and radar sensors – an inter-comparison experiment using ASTER and TanDEM-X data, *The Cryosphere*, 18, 3195–3230, <https://doi.org/10.5194/tc-18-3195-2024>, 2024.
- 180 xDEM contributors: xDEM, <https://doi.org/10.5281/zenodo.11492983>, 2024.
- Zevenbergen, L. W. and Thorne, C. R.: Quantitative analysis of land surface topography, *Earth Surf. Process. Landf.*, 12, 47–56, <https://doi.org/10.1002/esp.3290120107>, 1987.