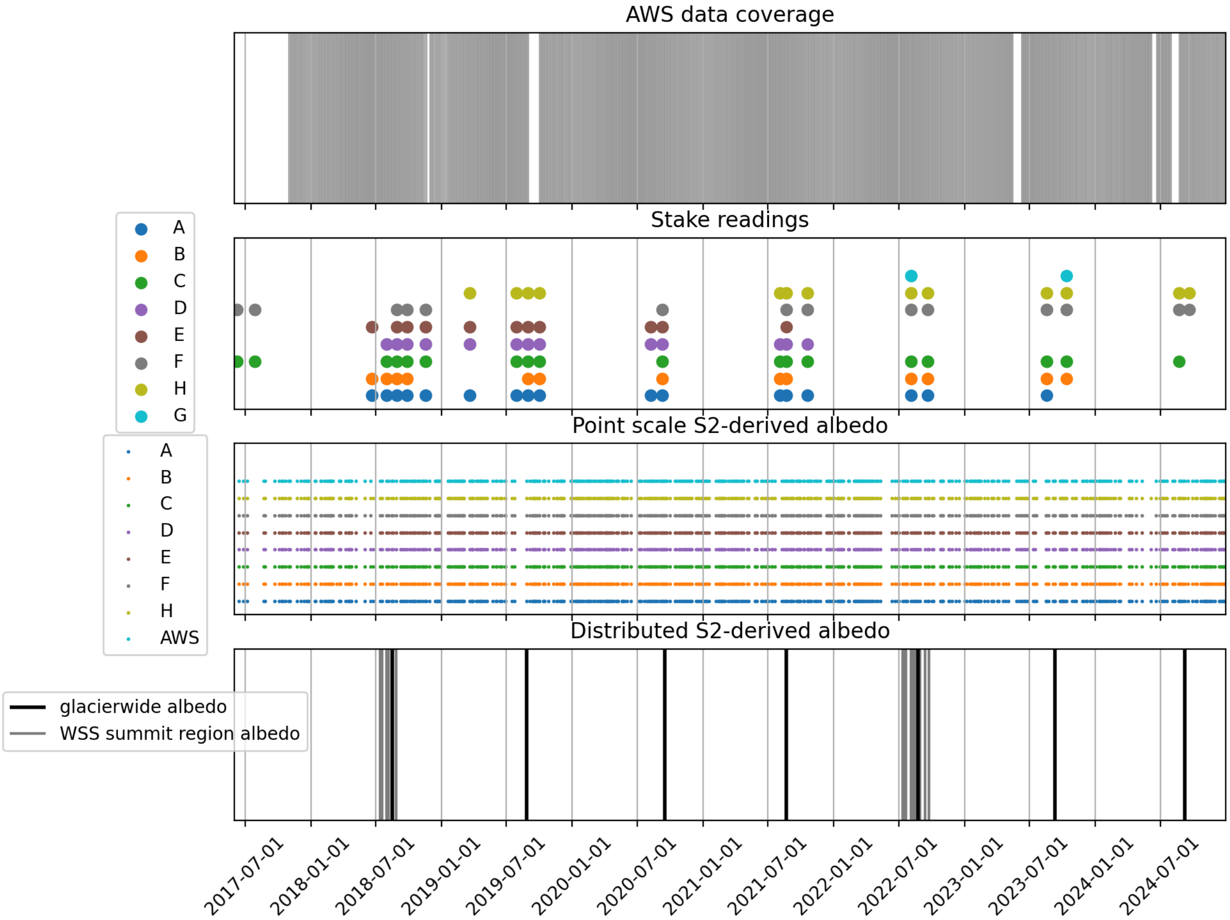
RC1:

Thank you very much for the encouraging comments! We truly appreciate the constructive feedback and the time spent on the review! Please see below for responses to specific comments (our responses are in blue text).

-Lea Hartl on behalf of the authors

The study focuses on quantifying the evolution of albedo in the accumulation zone of the WSS (Austria). By combining in situ measurements with satellite observations, the authors demonstrate a good agreement between these approaches, enabling a detailed analysis of the spatiotemporal variability of albedo. On this glacier, as on many other Alpine glaciers, accumulation zones are progressively shrinking, exposing not only firn from previous years but also an increasing proportion of exposed bare ice at the surface, leading to a significant decrease in albedo. Based on seven years of observations, the authors highlight a pronounced decline in albedo within the accumulation zone, where ice exposure at the surface has become more frequent, particularly since 2022. Due to the positive feedback of albedo, this phenomenon further enhances glacier melt, a process the authors quantified in this study. Despite its critical role in glacier mass balance evolution, this mechanism remains relatively understudied, particularly in terms of its spatial variability. Results presented in this study thus provide valuable insights for the glaciological community.
I thoroughly enjoyed reading this paper, which is clear, well-written, and well-structured, presenting novel and valuable findings. I particularly appreciated the discussion section, especially sections 4.3 and 4.4. The introduction is also well-written and clear; however, the list of cited references is often not exhaustive. It would be beneficial to either expand the references or, at a minimum, indicate their non-exhaustive nature using "e.g." (e.g., lines 1, 20, 26, 31, 36...). The methodology section is well-structured and engaging, but data availability (e.g., time period, resolution, etc.) remains somewhat unclear at times. Given that numerous measurements and observations were conducted over different periods and at varying temporal and spatial resolutions, a more visual representation (such as a figure or table) summarizing the measurement periods, resolution, uncertainties, etc., could significantly improve readability. Finally, before publication, I believe it would be valuable to clarify the (not major) points outlined below.

We will address the points mentioned above (expand reference list, indicate that it is not exhaustive) and add a data overview figure (example figure showing data type and availability through time below).



Incertitudes

It seems important to better clarify and quantify uncertainties, particularly to improve the discussion and comparison between the two methods.

1) Regarding the AWS observations, I appreciate the effort to quantify uncertainty presented in the appendix, but some questions remain: What is the sensor accuracy (line 137)?

The uncertainty in daily albedo due to sensor accuracy is 14%, derived from the approximate uncertainties for daily values at mid-latitudes given in the sensor manual. We added this to the revised manuscript in parenthesis. The sensitivity of the sensors is 12.81x10^-6 and 12.83x10^-6 V/ (W/m2) for the up and down facing sensor, respectively. While this is an important metrological metric, the error computation that follows from this and other factors is complex and we therefore follow the guidance of the manual.

Where does the 14% uncertainty come from (e.g., Figure 4 and S5)?

Section 2.3. states: “Based on standard error propagation (root sum of the squares) for the ratio of up- and downwelling radiation, we assume an uncertainty of 14% for the daily albedo”. The up- and downfacing sensors each have an uncertainty of 10%. 10^2 + 10^2 = 200. The root of 200 is 14.1.

The uncertainty related to surface roughness could be mentioned (e.g. line 131)

Added a note on this in the suggested line with additional references.

2) The S2 images provide high-resolution spatially distributed albedo data, which is highly relevant. However, uncertainties associated with this method are not quantified and should be reported in the study (e.g., Figure 4). Additionally, for this method, it is unclear whether and how the solar zenith angle at the time of image acquisition is accounted for in the albedo calculation. Could this have a significant impact on the albedo from S2?

Thus, the comparison between AWS and S2 approaches (e.g., bias, RMSE, Section 3.1.2) could be further discussed in relation to their respective uncertainties.

Quantifying uncertainties in S2 and similar earth observation data is highly complex and subject of ongoing work in the EO and metrology community (e.g. Mittaz et al 2019). Challenges arise from the numerous sources of uncertainty and processing steps that take place between the collection of raw at-sensor telemetry (L0) and derivation of calibrated top of atmosphere radiance (L1), and bottom of atmosphere reflectance (L2), which involve radiometric, geometric, and atmospheric corrections, and principles of radiative transfer. We refer to Gorroño et al (2017, 2024) for an overview of uncertainty contributors combined in accordance with the GUM. The uncertainty related to solar zenith angle is assumed to be negligible compared to other sources of uncertainty by Gorroño et al (2024). Gorroño et al (2017) provide further context on this. We note that at present there is no universal, standard way of quantifying or reporting uncertainties in L2 products. We agree with the reviewer that this would certainly be desirable and that it is beneficial to attempt rough uncertainty estimates even though exact quantification is beyond the scope. We propose an approach based on values reported in Gorroño et al (2024) for the S2 L2A product per spectral band (Table 2 in Gorroño et al, 2024). We use the values generated with the MCM approach for the Amazon scene. For the bands used in our albedo computation, these are:

B2: 9.75%; B4: 7.77%; B8: 1.12%; B11: 1.89%; B12: 3.30%

The sum of the squares for these values is 171.15 and the root thereof is 13. Accounting for the correlation between the bands (Gorroño et al, 2024, Fig 5) adds 3%, resulting in an estimated uncertainty of 16% We have included this uncertainty estimate in Fig. 4 and expanded the discussion of this issue (including comments on assumptions made in this estimate) in the revised manuscript.

Gorroño, J., Fomferra, N., Peters, M., Gascon, F., Underwood, C. I., Fox, N. P., Kirches, G., & Brockmann, C. (2017). A Radiometric Uncertainty Tool for the Sentinel 2 Mission. *Remote Sensing*, *9*(2), 178. https://doi.org/10.3390/rs9020178

J. Gorroño, L. Guanter, L. Valentin Graf and F. Gascon, "A Framework for the Estimation of Uncertainties and Spectral Error Correlation in Sentinel-2 Level-2A Data Products," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-13, 2024, Art no. 5634613, doi: 10.1109/TGRS.2024.3435021.

Mittaz, J., Merchant, C. J., & Woolliams, E. R. (2019). Applying principles of metrology to historical Earth observations from satellites. *Metrologia*, *56*(3), 032002.

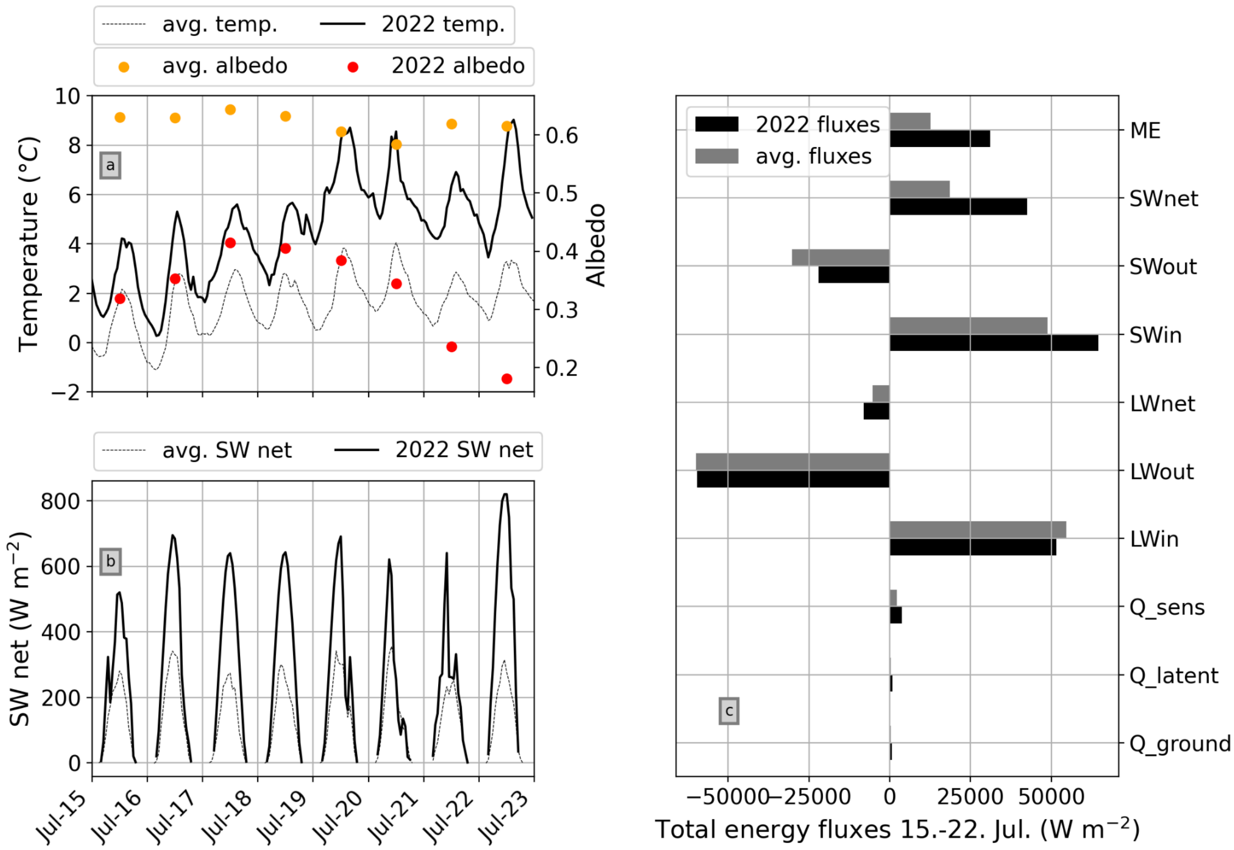
Spatial variability and albedo

1) The study of the spatial variability of albedo is both interesting and innovative and could be better illustrated. For example, lines 243–244 and Section 3.2 could be accompanied by a spatially distributed albedo map, as this information is not clearly visible in Figure 5.

Figures 8 and 12 show spatially distributed albedo as contour plots. Section 3.2 addresses time series of albedo at point locations, so a single albedo map would not convey the same information. We are happy to include additional maps of spatially distributed albedo if that would be beneficial but would ask for a clarification of what sort of time frame should be considered in maps that would accompany section 3.2.

2) The interpretation of the relationship between albedo and SMB is valuable; however, the effect of temperature is not discussed. Years with low albedo values (e.g., 2022–2024) are also among the warmest, making it difficult to disentangle the impact of albedo feedback on melt from that of high temperatures. This aspect could be further explored or at least mentioned. Additionally, temperature data are reported in Figure 7b, but this panel is not discussed when analyzing the summer of 2022.

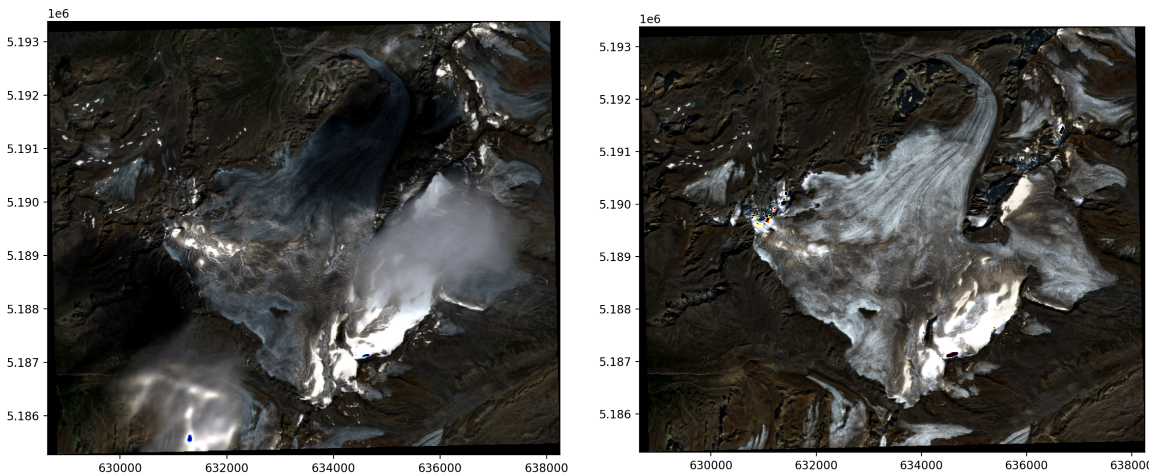
Section 3.3 on summer 2022 does include comments on the temperature data shown in Fig 7b. We have included a reference to panel b in the text to make this more clear. We also added a paragraph and an additional figure to section 3.4.1 showing the energy fluxes during the 2022 heatwave compared to average 2018-2024 conditions. The net shortwave component of the energy balance was the main factor contributing to increased energy available for melt (ME) during the heatwave. The turbulent fluxes (which are related to air temperature) were elevated during the heatwave compared to average conditions but their contribution to overall ME is much lower than that of shortwave radiation (as can be expected for mid-latitude glaciers in summer). We quantify the contributions in the revised manuscript.



Additional figure showing the 2022 heat wave and average 2018 to 2024 conditions during the same time of year and associated energy fluxes in panel c: ME: Energy available for melt. Q_sense: Sensible heat flux. Q_latent: Latent heat flux. Q_ground: Ground heat flux.

3) Furthermore, regarding the results presented in Figure 12, it seems important to mention that the satellite images for 2023 and 2024 were acquired in September, a period when the glacier's albedo is likely not at its lowest (see Figure 2). Could the comparison of glacier-wide albedo between years be somewhat biased by the late acquisition dates in these two years (e.g., line 343)? This point could be addressed.

The images were selected for minimum snow cover and all suitable (cloud free) images were considered. Both 2023 and 2024 had relatively long ablation seasons that extended into October. In 2024, the lowest albedo at the AWS was measured on September 8 and the S2 image in Fig 12 is from Sep. 7 (see Fig 3 and Table 2). The image and the AWS minimum coincide very closely. In 2023, the lowest albedo value at the AWS occurred on August 23. A small snowfall event then brightened the surface. This snow melted again in the following days across most of the glacier, although some snow remained at the AWS location. The S2 image from August 24 is partially affected by clouds and cloud shadows (left image in the figure below). In the S2 image from September 10 (shown in Fig. 12 and below on the right) the remaining seasonal snow cover has retreated further compared to the Aug. 24 image, although a minimal amount of snow from the summer snowfall in late August remains at the AWS location. The exposed bare ice surface appears visually brighter in some areas of the Sep. 10 image but we do not believe this indicates a bias due to the time of year.



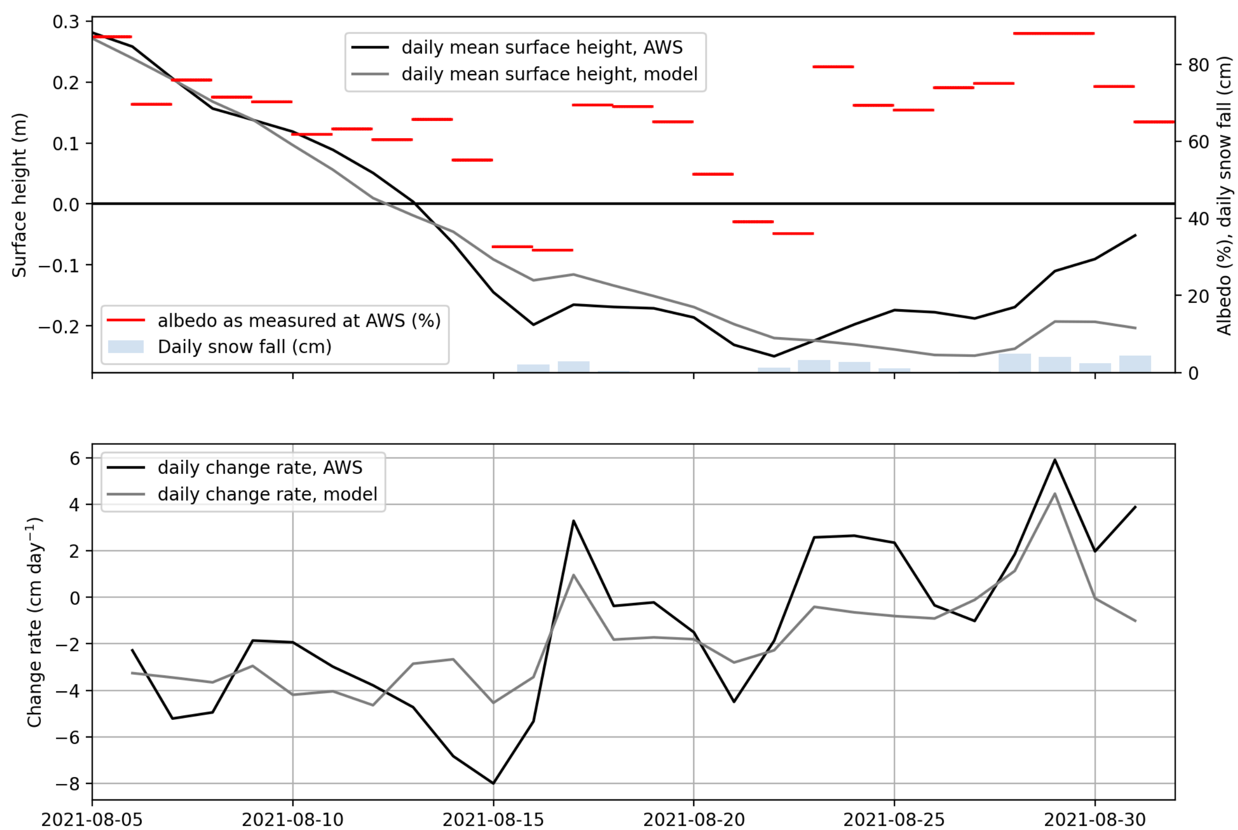
Simulations

Although the authors mention that the model validation exercise seems to be outside the scope of the study, the fact that they quantify the impact of albedo on melt using this method and highlight it in the results (e.g., lines 324-327, Figure 11), and also in the discussion, conclusion, and abstract makes it, in my opinion, difficult to avoid an precise evaluation of the model. If this precise quantification is a result the authors wish to emphasize, I strongly encourage them to properly evaluate the model. Otherwise, I suggest they focus solely on relative comparisons from sensitivity tests.

Thank you for this feedback. Reviewer 2 had similar comments and we have expanded and restructured the section explaining the sensitivity experiments to provide more information and quantitative comparisons of modeled and observed surface height change. The supplementary material pertaining to this has also been expanded. Possibilities for model evaluation are limited by the available observational data, mainly the surface height change information derived from the SR50 records. There is a lot of noise in this data set and small changes in particular cannot always be reliably extracted. We have explained the limitations of the SR50 in more detail and provide comparative statistics on modeled and observed surface height change for an example period with relatively good data quality in the revised manuscript. We would like to keep the absolute melt values from the different model runs as part of the results but will reduce focus on these towards more frequent usage of relative values in the revised manuscript.

Some suggestions for model evaluation : The simulated vs. observed snow-to-ice transition could be quantified using a delta day (line 201) ; Fig 9 could be evaluated using the SR50 (although the SR50 was used to force the model with snowfall, there is no precipitation during the period presented); ice temperature sensor measurements could be compared with the model simulations.

We have added a delta day quantification and additional statistics for the evaluation period detailed in section 2.5. We also provide a quantitative comparison of the cumulative modeled and observed surface height change for the period shown in Fig. 9. The quality of the SR50 data is not ideal during this period, which we also explain in this section. The "model evaluation" figure in the supplement has been adapted to show a comparison of daily surface height change.



Updated supplementary figure showing measured and modeled surface height change, and observed albedo and snowfall during the evaluation period (upper panel). The lower panel shows modeled and observed daily change rates.

Additionally, some clarifications on the simulation parameterizations seem important, particularly regarding model calibration and initial conditions (e.g., in the appendix). Finally, is the surface in the simulations in Figure 11 still ice? This should be specified. If so, is an albedo of 0.6 realistic for ice?

Following this comment and similar points by reviewer 2 we have restructured section 2.5

(sensitivity experiments) to more clearly explain what was done. Regarding Fig. 11: we assume an ice surface covered by a minimal amount of snow, like would be the case after a small summer snowfall. This is now clarified in Sec. 2.5, along with the main assumptions made regarding initial conditions. We will prepare a code repo with the relevant config files to accompany the revised manuscript.

Line by line comments:

Lines 27-31: long sentence difficult to understand. Please reformulate.

Split into multiple sentences as follows: “At local and regional scales, ice albedo depends on meteorological factors like solar elevation, cloudiness, and radiation budget (Volery et al., 2025) and on surface roughness (Irvine-Fynn et al., 2014). Additionally, the presence of liquid water on the ice surface and the characteristics of the pore space and the weathering crust impact albedo (e.g. Dadic et al., 2013; Traversa and Di Mauro, 2024). Light absorbing impurities of organic and inorganic origin, including carbon, algae, and dust, can lead to albedo reductions and darker glacier surfaces (e.g. Oerlemans et al., 2009; Di Mauro et al., 2017; Goelles and Bøggild, 2017).”

Line 80 to 88: What is the measurement period for the AWS (in relation to the above comment) and other sensors, and what is the temporal resolution? The same applies for the ice temperature sensors and the camera.

The measurement period for the AWS (including all associated sensors) is from the date of installation to present with some minor gaps due to power supply failures and other technical issues. Data are logged as ten minute averages, as stated. The measurement period for the camera is also from the stated date of installation to present. The camera takes a picture every two hours during daylight hours. We have added a note to this section clarifying this.

Line 83: Ice temperature sensor: Is it used in this study?

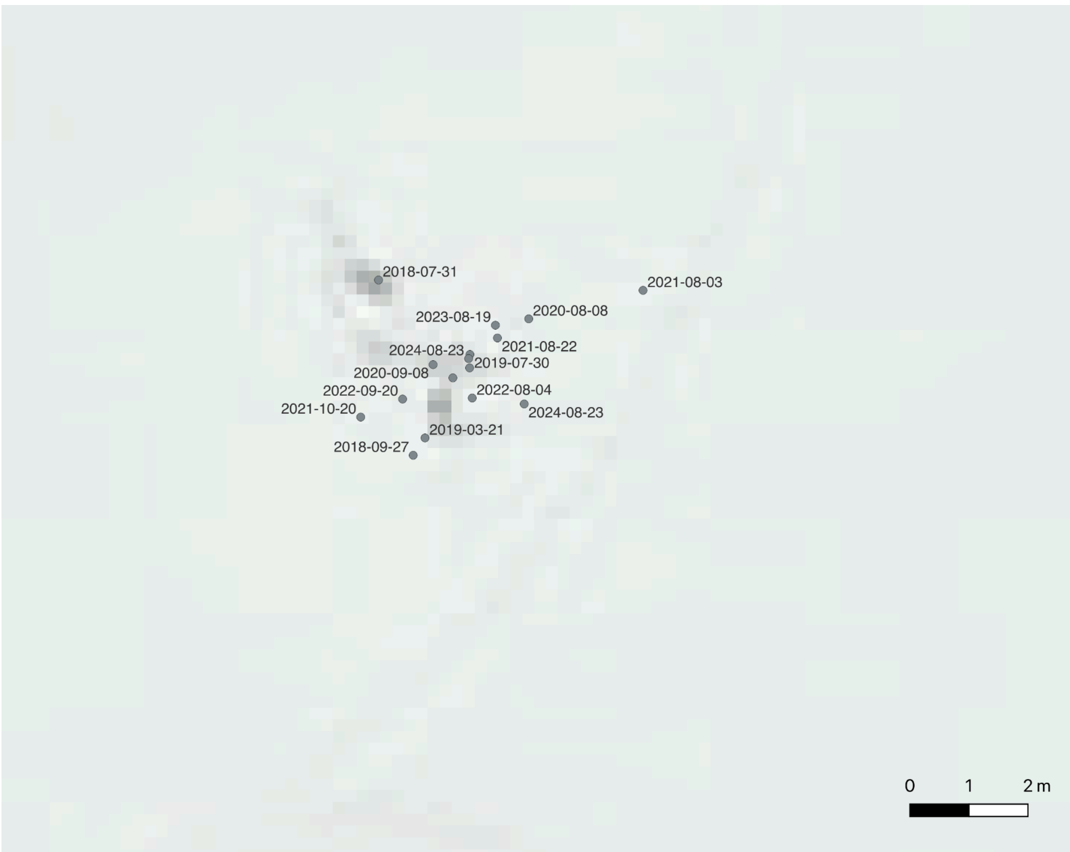
No. The information from the thermistor strings informs the initial ice temperature assumptions made in the energy balance model in a general way but this sensor is not essential for the study. We mention it here for completeness along with the other components of the AWS. We can remove this if it is confusing or seems unnecessary.

Line 85-86: It took me some time to understand that we are in the accumulation zone, but with exposed ice (which is uncommon for an accumulation area). A clearer description here could help, especially since Figure 1 shows only snow.

Added the following description and changed figure 1 to show an image with exposed ice. “The location of the AWS is in the highest region of the glacier within the (former) accumulation zone. Multi-year firn is no longer present around the AWS and bare ice is exposed if the seasonal snow cover does not persist through the summer.”

Line 105: "Ice flow is not apparent" – Please specify: "Ice flow is lower than..."

We would prefer to keep the current phrasing. We do not have a precise way to quantify “lower than”. In the GNSS coordinates gathered during site visits, no systematic shift of the position of the AWS or stakes over time is apparent. The coordinates form a “cloud” without an apparent direction of movement. Hence, any movement due to ice flow would be less than the locational uncertainty of the points. The image below shows the GNSS coordinates of the AWS as recorded during visits.



Line 139 and throughout the document: "Low" and "very low" refer to specific values (i.e., 0.2 and 0.4) as indicated here. These terms are used throughout the document, sometimes with quotation marks and sometimes without. Conversely, "low albedo" is sometimes mentioned without explicitly referring to these values, making the text harder to follow. Please ensure that quotation marks are consistently used when referring to these specific values, or alternatively, use a uniform notation (e.g., alb < 0.2).

We will go through the manuscript to ensure consistency.

Line 145: Could you provide further details, such as the spatial resolution of the S2 images, the number of images, and the period covered?

We have added the spatial resolution (10x10m) and clarified that we use S2 data for the 2018-2024 period, i.e. the period of record of the AWS and stake data. In section 2.4.2 we additionally specified that we select one image per year to assess minimum snow cover conditions.

Line 182: "Albedo as input" – do you mean albedo from the AWS?

In this case yes, although any kind of albedo data could be used. We specify this in the revised manuscript.

Line 216: Unclear where the value of 0.3 comes from. Could you clarify?

This refers to data as measured by the AWS. We have added references to the respective table and figure. We also removed unfortunate typos in the table, which likely caused this confusion (sorry).

Line 234, 238: "Generally coincide or occur" – This statement could be quantified (e.g., using delta day) to add more weight to the comparison.

We added a note on this in the revised manuscript. In 2018, S2-derived "low albedo" was observed five days prior to "low albedo" conditions at the AWS. Otherwise the S2-derived low albedo periods are within the AWS low albedo periods. In line 238, the statement is followed by a description of cases when there are discrepancies with examples. The stakes are not expected to have low albedo at exactly the same dates due to varying snow melt patterns. The irregular nature of the S2 time series makes it challenging to meaningfully interpret shifts of a few days - these may be due to snow melting earlier or later at one stake compared to the next, or to differing availability of S2 imagery.

Line 248 and throughout the paper: Be consistent with units: mm w.e. should be accompanied by a time period (e.g., mm w.e. yr^{-1}) (e.g., line 248, line 323, line 396, etc.).

Added time periods where appropriate. In cases where it is not a year or a day, the period is stated in the text (e.g. “X mm w.e. over five days”).

Line 246 to 251: To give more weight to the different ablation values, you could mention the percentage they represent relative to the mean, especially for 2022.

We would prefer to keep the focus on absolute values rather than percentages of a mean. The stake data exhibit high year-to-year variability, which would not be well represented by averaging over all years. The large variability is - in our opinion - the key characteristic of the stake data set from this location and more relevant for the future development of the glacier than the deviation of single years from an average value.

Line 252: Specify that this refers to albedo from S2.

Added this clarification.

Line 419: While I find this discussion relevant and convincing, wouldn't the primary effect of darkening at the glacier scale be more related to the expansion of the accumulation zone?

The darkening is due to both the loss of snow/firn area and unusually dark ice surfaces in these years. We clarified this in the revised text and added a reference to the figure showing albedo histograms: "At the glacier scale, 2022 and 2024 stand out as very low albedo years. This is due to the almost complete loss of snow and firn in these years and a high fraction of bare ice surfaces with "very low" albedo (Fig. 14). Possible explanations for the darkened ice surface in these particular years include….."

Line 459-460: Also, the firn.

Changed to "firn and ice surface conditions"

Figures and Sup. Mat.

General: Many figures should be larger because they are sometimes difficult to read.

We have adjusted font sizes and figure design to improve this.

Figure 1b: Glacier outlines and 50m contours are barely visible (green and blue lines).

Changed the color and increased the linewidth.

Figure 5: Difficult to read. Consider splitting it into two panels: one with AWS-in situ and AWS-S2 (to show the comparison) and another with S2 at the stakes (to show spatial variability).

We have split the figure as suggested.

Figure 7: It is difficult to differentiate the colors specific to each stake. Add (a) and (b) directly on the figure.

Panel labels (a, b) are included in the figure. We have adjusted the colors to hopefully make them easier to differentiate.

Figure 12h: The contours are hard to see, especially the green ones. Could they be made larger or changed to a different color?

We changed the color and kept only one set of contours to make the figure cleaner and easier to interpret.

Sup. Mat.: The order of references to the supplementary material is not always chronological (e.g., Line 90: referenced as 2 in the supplementary material but should be 1). Additionally, some references to the supplementary material could be more specific (e.g., Line 186: which figure does this correspond to?).

We revised the structure of the supplement to be in chronological order and have added more specific references to the supplement in the main manuscript.