

The manuscript ‘Extreme Galcier Melt in the Central Tibetan Plateau during the Summer of 2022: Detection and Mechanisms’ by Zhu and co-authors present a detailed analysis of the energy-balance of Sangqu Glacier using the EBFM model. The authors further apply a statistical approach to determine whether the summer 2022 was extreme in terms of glacier melt and discuss the underlying atmospheric mechanisms leading to the high recorded melt rates.

This is a good study presenting a detailed process-oriented approach to better understand the links between glacier and atmosphere. The model used was applied with care and extensively validated against in situ measurements. Despite this, I still had a number of points to raise that I feel would need to be addressed before this manuscript can be considered for publication.

Response: We are grateful for the reviewer’s insightful suggestions and acknowledge the remaining points that need to be addressed. We fully agree with the reviewer that these points are essential for improving the clarity and quality of our manuscript. We carefully addressed all the comments raised by the reviewer and provide a point-by-point response below.

Thank you again for your time and expertise.

#### **General comments:**

**Comment 1:** Statistical analysis with very short time series: I found it hard to follow the description of the new framework used to identify these extreme events from time-series of just a few years. This methodological section goes straight into technicalities, without explaining the general principle, making it difficult to understand what the authors are actually after. Furthermore, given the extremely short focus period, I have serious doubts about the value of such an approach. Along these lines, I fail to understand why the authors followed what seems to be a purely statistical approach without leveraging the energy-balance model that they validated at this site and could be expanded in time.

Response: We sincerely thank the reviewer for these insightful comments. They have helped us substantially improve the clarity and rigor of our manuscript.

**Section 4.4 has been rewritten to clarify the logic of the dual-threshold framework, and a brief justification for using a statistical approach instead of model extension has been added at the end of that section.**

#### **1. Methodological description and conceptual framing**

We fully agree that the original description of the dual-threshold framework in Section 4.4 was too abrupt and lacked a clear conceptual introduction. In the revised manuscript, we have completely rewritten this section. We now begin by stating the

core idea explicitly: to translate the conventional percentile-based definition of an extreme event (which requires long records) into a parametric form ( $\mu - k\sigma$ ) that can be calibrated from short training periods. By establishing a mathematical equivalence between the two definitions within the training period, we obtain a range of valid multipliers  $[k_{\min}, k_{\max}]$ , which yields both a relaxed threshold and a high-confidence threshold. This design serves two purposes: (i) it provides a parametric rule applicable to the test period, and (ii) the interval between the two thresholds explicitly quantifies the uncertainty arising from the shortness of the training series. We believe this revision clarifies the logical chain of our methodology and addresses the reviewer's concern about the intelligibility of the approach.

## **2. Value of the statistical approach with short time series**

The reviewer raises a critical question about the value of such a method given the limited observation length. We fully acknowledge that, in an ideal setting, long observational records are preferable. However, over 90% of glaciers on the Tibetan Plateau have mass balance records shorter than ten years, making traditional methods (e.g., percentile thresholds, generalized extreme value distributions) inapplicable. Our dual-threshold framework is specifically designed for this common but challenging scenario. It transforms a non-parametric definition into a parametric rule, and—importantly—the interval between the relaxed and strict thresholds provides a quantitative measure of the uncertainty introduced by the short record. This allows, for the first time, a statistically defensible identification of extreme events in short-term glacier mass balance series. We have now emphasized this point more clearly in the revised manuscript (Section 4.4).

## **3. Why not extend the energy-balance model instead?**

We appreciate this thoughtful question. Indeed, our energy-balance model (EBFM) was carefully validated and could theoretically be extended in time to identify extreme melt events. However, we see two main reasons for developing a separate statistical framework. First, the EBFM requires high-quality, high-frequency meteorological forcing (radiation, humidity, wind, etc.). For years prior to 2020 in our study region, such observations are either unavailable or subject to large uncertainties, and using

reanalysis-only forcing would introduce uncontrolled errors that could compromise the detection of extremes. Second, and more importantly, the two approaches serve complementary purposes. The EBFM provides a process-based understanding of *how* extreme melt occurs, while the statistical framework offers a simple, transferable tool that relies solely on annual mass balance observations. This makes it readily applicable to other data-scarce glaciers across the Tibetan Plateau and beyond, where high-quality meteorological data are often lacking. We have added a brief clarification of this rationale in the revised manuscript (Section 4.4).

**Comment 2:** Parallel with extreme glacier mass losses in other regions of the world: This study is very much focused on the Tibetan Plateau, and I find it a bit disappointing that it does not draw more parallels with recent glacier extreme events that have been reported in other regions of the world, such as Canada, Switzerland, the Andes or Svalbard. Some mechanisms described here are also described for these other regions and it would be interesting to insist on how these events relate or differ.

**Response:** We thank the reviewer for this constructive suggestion. We agree that situating our findings within the broader context of recent global glacier extreme events significantly strengthens the discussion.

**Section 6.2 has been substantially expanded to systematically compare extreme glacier melt events across different mountain ranges worldwide, highlighting common drivers and regional differences.**

Following this suggestion, we have substantially expanded Section 6.2 (formerly titled “Comparison with extreme glacier mass loss events in other regions during 2021–2022”) to include a systematic comparison with glacier mass loss extremes reported in diverse mountain regions worldwide, including the Swiss Alps (Van Tiel et al., 2026), Austria (Voordendag et al., 2023), Svalbard (Bernard and Friedt, 2025), the Pyrenees (Izagirre et al., 2024), the Tien Shan (Xu et al., 2024), western Canada and the United States (Menounos et al., 2025), and New Zealand (Vargo et al., 2020).

In the revised discussion, we first synthesize common characteristics across these regions, such as record ablation rates, rapid glacier degradation, and the critical role of albedo feedback. We then examine shared driving mechanisms, including anomalous

large-scale atmospheric circulation (e.g., westerly jet shifts), summer heatwaves superimposed on winter snow deficits, and energy balance anomalies characterized by enhanced incoming shortwave radiation and sensible heat flux. We also highlight regional differences, such as the role of early spring melt onset on the Tibetan Plateau versus the importance of liquid precipitation events in the Arctic, and the varying contributions of anthropogenic forcing across regions. This expanded comparison not only demonstrates that the extreme melt observed on Sangqu Glacier is part of a broader global phenomenon but also helps to contextualize the specific mechanisms operating in the central Tibetan Plateau.

We believe this revision substantially enhances the global relevance of our study and directly addresses the reviewer's suggestion. The added content is presented in Section 6.2 of the revised manuscript.

**Comment 3:** Differences in mass balance measurement periods: The mass balance records are reported to end in July in 2022, compared to October in the previous years, which bodes the question of how comparable these measurements are, especially considering that 2022 is pointed out as the extreme year of the time series.

**Response:** We thank the reviewer for raising this important issue. We fully acknowledge that the 2022 mass balance record ends in July, whereas the previous years' records extend through October, and we agree that this discrepancy warrants careful consideration regarding the comparability of these measurements.

**Model parameter optimization, seasonal-scale validation (updated Fig. 2d), and cross-validation with the neighboring Xiao Dongkemadi Glacier (updated Fig. 2h) have been added in Section 4.2 to address the incomplete 2022 observational record.**

The shortened observational period in 2022 was unfortunately unavoidable due to the COVID-19 outbreak in Tibet during that summer, which prevented field access after July. To ensure that our identification of 2021–2022 as an extreme mass loss year is robust despite this limitation, we took the following three complementary approaches:

#### **1. Model parameter optimization**

Following the suggestion of Reviewer 1, we applied the NSGA-II multi-objective

optimization algorithm to better constrain the key parameters of the energy-balance model (EBFM) using all available in situ observations. This optimization improved the model's ability to simulate the full seasonal cycle of mass balance, particularly during the ablation season, thereby enhancing the reliability of the simulated mass balance for the July–October period in 2022 when direct observations were missing.

## **2. Additional seasonal-scale validation**

To further evaluate the model's performance at sub-annual timescales, we incorporated the observed seasonal mass balance from the summer of 2021 (June–September) into our validation framework. This allowed us to directly assess the model's ability to capture ablation-season mass balance—the period most relevant to the 2022 extreme event. The comparison shows good agreement, with a root-mean-square error of 329 mm w.e. at the seasonal scale (see updated Fig. 2d). This additional validation strengthens our confidence in the model's capability to simulate mass balance dynamics during the ablation season.

## **3. Cross-validation with a neighboring glacier**

We collected independent mass balance observations from Xiao Dongkemadi Glacier, located approximately 90 km from Sangqu Glacier, for the complete 2021–2022 balance year. The close agreement between the simulated annual mass balance at Sangqu Glacier (–919 mm w.e.) and the observed value at Xiao Dongkemadi Glacier (–881 mm w.e.) provides strong independent evidence that the 2021–2022 mass loss at Sangqu Glacier was indeed exceptional, and that the partial observational record captures the overall extreme nature of the year.

Together, these three strategies—model-based reconstruction constrained by optimized parameters, seasonal-scale model validation using summer 2021 observations, and independent verification from a neighboring glacier—substantially mitigate the concern arising from the incomplete observational record in 2022. We have incorporated these improvements into the revised manuscript, with the updated model validation now reflected in Figure 2.

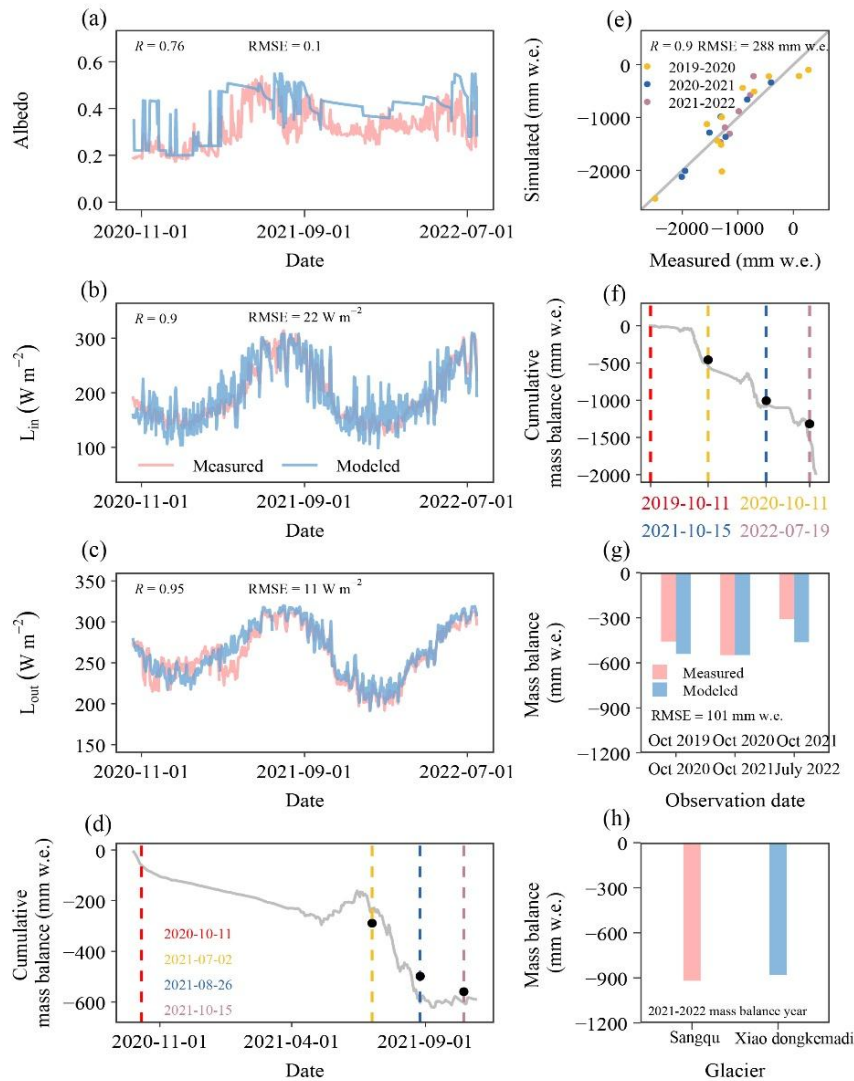


Figure 2. Comparison between observations and simulations of (a) albedo, (b)  $L_{in}$ , (c)  $L_{out}$ , (d) glacier mass balance from June to September 2021, (e) point mass balance at different elevations, (f) accumulation mass balance, and (g) glacier-wide mass balance for the periods 2019/20, 2020/21, and from October 2021 to July 2022. (h) Comparison of simulated glacier mass balance for Sangqu Glacier in 2021/22 with observed mass balance for Xiao Dongkemadi Glacier.

**Comment 4:** Calibration and validation of the energy-balance model over the same time period: if I understand correctly the model was calibrated (for the precipitation gradient) and validated against the same data. This makes me confused as to why there seems to be a bias between the simulated and measured point mass balances in Fig. 3a? And if this statement is correct, this is a limitation that would need to be raised in the discussion.

**Response:** We thank the reviewer for raising this important methodological point. We fully understand the concern regarding the potential overlap between model calibration and validation data, as well as the observed bias in point-scale mass balance shown in Fig. 2e (formerly Fig. 3a). We address these issues in the revised manuscript through

the following improvements and clarifications.

**The separation of calibration and validation datasets has been clarified in Section 4.2, and a systematic discussion of multi-scale model uncertainties has been added in Section 6.4.**

### **1. Clarification of calibration and validation data separation**

We acknowledge that in the original submission, the description of model calibration and validation was not sufficiently explicit. To address this, we have revised Section 4.2 to clearly distinguish between calibration and validation datasets. Specifically:

**Calibration:** The NSGA-II multi-objective optimization algorithm was applied using glacier-wide annual mass balance observations from the periods 2019–2020 and 2020–2021, as well as daily albedo measurements from the AWS, to constrain key parameters (e.g., snow/ice albedo, turbulent exchange coefficients).

**Validation:** Model performance was evaluated using independent datasets that were not used in calibration, including:

- **Point-scale mass balance** at different elevations (Fig. 2e);
- **Seasonal mass balance** for June–September 2021 (Fig. 2d);
- **Cross-comparison** with Xiao Dongkemadi Glacier observations for the 2021–2022 balance year.

This separation ensures that the validation results provide an unbiased assessment of model performance.

### **2. Comprehensive explanation of model simulation uncertainties**

The reviewer correctly notes the discrepancies between simulated and observed mass balance, particularly at the point and seasonal scales. We acknowledge these as inherent limitations of distributed glacier modeling under data-scarce conditions. In the revised manuscript (Section 6.4, *Research Limitations*), we now provide a systematic discussion of the sources of model uncertainty, including:

**Spatial representativeness of meteorological forcing:** Extrapolation from a single AWS using simplified lapse rates and gradients introduces spatial heterogeneity that cannot be fully captured.

**Parameterization uncertainty:** The model is most sensitive to snow/ice albedo and temperature lapse rate, as quantified in our sensitivity analysis (Tables S1–S6). Calibration of these parameters is constrained by limited observations.

**Scale mismatch:** Point-scale stake measurements represent discrete locations, while the model simulates grid-cell averages; local variations in snow distribution and microtopography contribute to point-scale scatter.

**Observational uncertainty:** In situ mass balance measurements themselves carry uncertainties from density assumptions, stake height readings, and spatial interpolation.

Importantly, despite these uncertainties, the model achieves substantially higher accuracy at the **glacier-wide scale** (RMSE of 101 mm w.e.; Figs. 2f and 2g), which is the primary scale of interest for our extreme event analysis. The close agreement between the simulated annual mass balance for 2021–2022 (–919 mm w.e.) and independent observations from the neighboring Xiao Dongkemadi Glacier (–881 mm w.e.) further supports the reliability of our conclusions. We have added a balanced discussion of these scale-dependent uncertainties in the revised manuscript.

### **3. Additional seasonal-scale validation**

To further strengthen confidence in the model’s ability to capture ablation-season dynamics—the period most critical for the 2022 extreme event—we incorporated the observed seasonal mass balance for June–September 2021 as an additional validation metric. The comparison shows good agreement, with an RMSE of 329 mm w.e. at the seasonal scale (see updated Fig. 2d).

### **4. Cross-validation with a neighboring glacier**

As described in response to Comment 3, we also collected independent mass balance observations from Xiao Dongkemadi Glacier for the 2021–2022 balance year. The close agreement between the simulated annual mass balance at Sangqu Glacier (–919 mm w.e.) and the observed value at Xiao Dongkemadi Glacier (–881 mm w.e.) provides additional independent evidence that the model performs reliably under the extreme conditions of 2022.

We believe these clarifications and improvements—particularly the clear separation of calibration and validation datasets, the systematic discussion of model uncertainties,

and the addition of multiple independent validation metrics—address the reviewer’s concerns and strengthen the overall credibility of our model-based analysis. All corresponding updates have been incorporated into the revised manuscript, primarily in Sections 4.2, 6.4, and Figure 2.

**Comment 5:** Limited details on the heat transfer in the ice: my expectation is that this glacier is cold-ice, which would have potentially a strong influence on the energy budget. However there is very limited description of how the heat transfer into the ice was calculated. What did you take for the heat transfer coefficients? Was a sensitivity analysis performed? Was there any validation data available?

**Response:** We thank the reviewer for raising this important point regarding the subsurface heat flux calculation. We agree that the original manuscript lacked sufficient detail on this aspect, and we have substantially expanded the description of the model setup and validation in the revised manuscript.

**The calculation of subsurface heat flux has been expanded in Section 4.1, validation against ice temperature measurements has been added in Section 4.2 (Fig. S3), and the glacier’s thermal regime description has been supplemented in Section 2.**

## **1. Calculation of subsurface heat flux**

In our energy-balance model, the subsurface heat flux ( $Q_G$ ) is computed by solving the one-dimensional heat conduction equation through the snow/firn/ice column, following the approach of Van Pelt et al. (2012). The effective thermal conductivity ( $\kappa$ ) is parameterized as a function of density based on Östin and Andersson (1991) for snow and Sturm et al. (1997) for firn and ice. The heat capacity ( $c_p$ ) is treated as a function of temperature following Yen (1981).

Effective thermal conductivity ( $\kappa$ ) is parameterized as a function of density using the formulation of Sturm et al. (1997):

$$\kappa = 0.138 - 1.01 \times 10^{-3} \rho + 3.233 \times 10^{-6} \rho^2 \quad (1)$$

Specific heat capacity ( $c_p$ ) follows Yen (1981):

$$c_p = p \cdot (152.2 + 7.122 T) \quad (2)$$

The model resolves vertical profiles of temperature, density, and liquid water

content, allowing for the simulation of heat conduction, meltwater percolation, and refreezing processes.

## **2. Sensitivity analysis**

To assess the sensitivity of glacier-wide mass balance to subsurface heat transfer, we perturbed the effective thermal conductivity ( $\kappa$ ) and specific heat capacity ( $c_p$ ) by  $\pm 10\%$  while keeping all other parameters at their calibrated values. The glacier-wide annual mass balance has a sensitivity of  $-9.9$  mm w.e.  $a^{-1}$  to effective thermal conductivity and  $+10.5$  mm w.e.  $a^{-1}$  to heat capacity. These values are substantially smaller than the sensitivities to surface energy balance parameters (e.g., snow albedo:  $123.2$  mm w.e.  $a^{-1}$ ; temperature lapse rate:  $-141.4$  mm w.e.  $a^{-1}$ ), indicating that subsurface heat transfer has a relatively small impact on the overall mass balance. Detailed sensitivity results are provided in Tables S1–S6.

## **3. Validation against in-situ ice temperature measurements**

To evaluate the model's ability to represent subsurface heat transfer, we compared simulated and observed ice temperatures at depths of 3 m and 5 m within the glacier. The simulations show strong agreement with observations, yielding correlation coefficients of 0.94 and 0.96, respectively. These results have been added to the revised manuscript and are now presented in Figure S3. This validation provides confidence that the model captures the thermal regime of the glacier adequately, even though the glacier is likely polythermal or cold-based in its upper reaches.

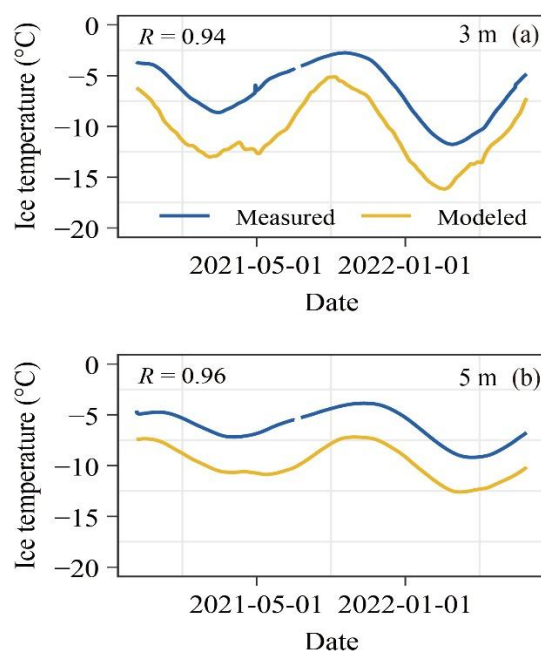


Figure S3. Comparison of observed and simulated ice temperatures at (a) 3 m and (b) 5 m.

#### 4. Additional clarification on glacier thermal regime

Based on the measured ice temperatures and the observed presence of meltwater percolation and refreezing in the ablation zone, we consider Sangqu Glacier to be polythermal, with cold ice in the upper accumulation zone and temperate ice in the ablation zone. This thermal regime is consistent with observations on neighboring Xiao Dongkemadi Glacier, where cold ice temperatures and superimposed ice formation have been documented (Fujita et al., 1996).

#### Summary of revisions

In summary, we have: (1) completely rewritten Section 4.4 to clarify the dual-threshold framework; (2) substantially expanded Section 6.2 with global comparisons; (3) added multi-strategy validation (parameter optimization, seasonal-scale validation, and neighboring glacier cross-validation) to address the incomplete 2022 record; (4) systematically discussed model uncertainties in Section 6.4; and (5) provided detailed descriptions of subsurface heat transfer calculations with new sensitivity analyses and ice temperature validation. All corresponding updates have been incorporated into the revised manuscript.

**Line-by-line comments:**

**Comment 6:** L54: This feels a bit like an oversell. While the study itself is interesting, I didn't find anything absolutely groundbreaking there.

**Response:** Yes, we revised this sentence. In lines 46 to 49.

“This study proposes a framework for identifying extreme mass loss events in short-term glacier mass balance records and provides insights into glacier mass balance responses to extreme weather and climate events.”

**Comment 7:** L61: I'm not convinced that wildfires fit in the category of 'weather and climate events'.

References would be welcome here.

**Response:** Yes, we revised this sentence, removed “wildfire,” and added references.

Please see lines 55 to 57.

“Extreme glacier melt events directly reflect glacier mass balance responses to extreme weather and climate events, such as heatwaves, droughts (Zhu et al., 2024a; Xu et al., 2024; Zhu et al., 2026; Van Tiel et al., 2026).”

**Comment 8:** L62: Could you be a bit more specific here? Give some numbers?

**Response:** Thank you for the suggestion. We have revised the sentence to include specific temporal evidence from Xiao Dongkemadi Glacier, the longest-monitored glacier in the central Tibetan Plateau with observations since 1989. On this glacier, extreme mass loss events were concentrated in 2000–2022 compared to 1989–2000 (Yao et al., 2012; Zhu et al., 2026a). Please see lines 57 to 61.

The revised text is as follows: “Across the Tibetan Plateau and surrounding regions, evidence from Xiao Dongkemadi Glacier, which holds the longest mass balance record in the central Tibetan Plateau with observations since 1989, indicates that extreme mass loss events have been concentrated in 2000–2022, compared to 1989–2000 (Yao et al., 2012).”

**Comment 9:** L62, 195: Should be 'the Tibetan Plateau'

**Response:** Yes, we changed it to “the Tibetan Plateau” throughout the text.

**Comment 10:** L101-103: See general comment, this does not sound realistic.

**Response:** Thank you for this critical comment. We agree that the original statement may have overstated the feasibility. To address this, we have thoroughly revised Section 4.4 to present our dual-threshold framework with greater precision. The revised section now includes: (1) a clear rationale linking the percentile and mean–standard deviation approaches; (2) a detailed description of the calibration and validation procedures using long-term glacier records; and (3) an explicit discussion of the framework's limitations

when applied to short-term data. We believe this revision provides a more realistic and transparent account of our methodological contribution.

Correspondingly, we have also revised the introductory statement to reflect this more nuanced perspective, which now reads: “Consequently, developing and testing frameworks—such as the dual-threshold approach proposed in this study—for identifying extremes from short-term observations (under ten years) is essential for advancing glacier research in this region.” Please see lines 94 to 97.

**Comment 11:** L110: remind here the period to which this anomaly applies.

**Response:** Yes, we revised this sentence to: “Hua et al. (2023) demonstrated that a 6.45 °C temperature anomaly, observed in central China over the 21-day period from 6 to 26 August 2022, broke historical records and established a new threshold for defining future record-breaking daily maximum temperatures in the region.” Please see lines 103 to 106.

**Comment 12:** L153: It would be useful to indicate the mean value of annual precipitation at this site. I expect this to be a very dry site, and it would be useful to have these numbers in mind to put the energy budget into context.

**Response:** Yes, we have added the mean annual precipitation.

Please see lines 151 to 155.

The revised sentence reads: “Based on the combined in situ and ERA5-corrected precipitation data, the mean annual precipitation at the Sangqu Glacier site was 510.5 mm for the period October 2019–September 2022. Over 70% of this precipitation occurs during June–September, coinciding with the warmest season (Zhu et al., 2024a).

**Comment 13:** L157: Are there any available geodetic mass balance estimates for this glacier?

**Response:** Yes, based on the data provided by Dussaillant et al. (2025), we added the geodetic mass balance data for Sangqu Glacier from 1957 to 2024. Please see lines 153 to 155.

“Multi-temporal digital elevation model differencing and in situ stake measurements reveal continuous mass loss in this region since the 1990s (Yao et al., 2012; King et al., 2023). Geodetic mass balance estimates from Dussaillant et al. (2025) indicate that Sangqu Glacier lost  $\sim 515.4$  mm w.e.  $\text{yr}^{-1}$  over the period 1957–2024, consistent with the regional mass loss pattern.”

**Comment 14:** L161: Photo should be singular Figure 1: Panel d is hard to make sense of. It would make sense to show a DEM hillshade or satellite image of the glacier, or at the very least use a different colorscale.

**Response:** Thank you for these suggestions. We have updated Figure 1. We believe the

revised figure is now clearer and easier to interpret.

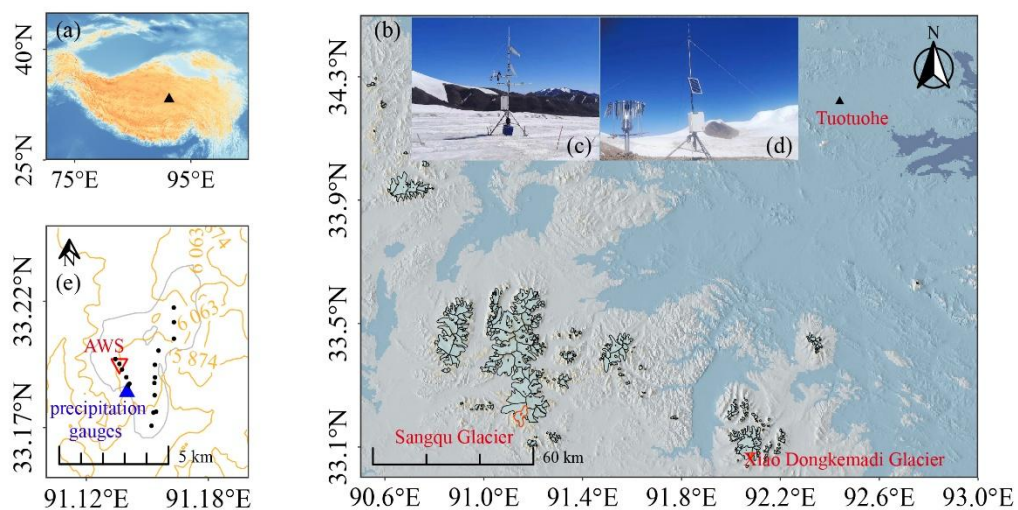


Figure 1. Overview of the study area. (a) and (b) Location of the study area on TP. (c) Photos of AWS and precipitation gauge on Sangqu Glacier (Photo credit: Fei Zhu). (e) Stakes on glacier, AWS, and precipitation gauges are located on Sangqu Glacier.

**Comment 15:** L163: Please indicate the location of these precipitation gauges.

**Response:** Yes, we revised Figure 1 to make the position of the rain gauge more prominent.

**Comment 16:** L171: greater -> more

**Response:** Yes, we have revised it.

**Comment 17:** L174: More details would be welcome here, as well as a picture appended to Fig. 1. Was this precipitation gauge heated? Shielded?

**Response:** Thank you for this suggestion. We have updated Figure 1 to include a photo of the precipitation gauge, which is now shown in the revised panel (d). Regarding the gauge specifications, it is a Geonor T-200B precipitation gauge equipped with a double-alter shield to minimize wind-induced undercatch. These details have also been clarified in Section 3.1.

**Comment 18:** L210: 'where' does not fit here

**Response:** Yes, we deleted "Where" and made revisions elsewhere in the main text.

**Comment 19:** L241-244: I don't get this. Why do these time series need to be reconstructed if they already exist?

**Response:** Thank you for this question. The mass balance observations at Sangqu

Glacier began in October 2019, whereas the automatic weather station (AWS) was not installed until October 2020. Consequently, to align the meteorological forcing with the full mass balance observation period (October 2019–September 2022), we reconstructed the missing meteorological variables—specifically air temperature, relative humidity, wind speed, and incoming shortwave radiation—using a random forest method that combines ERA5 reanalysis with AWS observations during the overlapping period. Please see lines 262 to 265.

Comment 20: L247: What is a ‘three-fold cross validation’? And a 10-fold?

**Response:** We thank the reviewer for the valuable question. Five-fold cross-validation is a widely used technique to assess the generalization performance of a model. In our study, we employed it to evaluate the predictive skill of the Random Forest model used for reconstructing near-surface meteorological variables (air temperature, relative humidity, wind speed, and incoming shortwave radiation) at the glacier site.

Specifically, the training dataset was randomly divided into five equal subsets. In each of the five rounds, four subsets were used for model training and the remaining one for validation. This process was repeated three times (three repetitions) with different random partitions to reduce the variability associated with a single data split. The final model performance was then averaged across all validation rounds, yielding robust estimates of predictive accuracy (e.g.,  $R^2$  values of 0.99, 0.96, 0.94, and 0.96 for the four variables at the hourly scale). This approach ensures that the reconstructed meteorological forcing used to drive the energy-mass balance model is reliable, especially for periods lacking in situ observations (e.g., August–September 2022).

We have clarified this in the revised manuscript (Section 4.1) to improve methodological transparency. Please see lines 281 to 285.

“To minimize overfitting and obtain robust estimates of model performance, we applied three repetitions of five-fold cross-validation, where the training data were repeatedly split into five subsets, each serving in turn as the validation set.”

Comment 21: L251-252: Did you consider looking into seasonal bias corrections?

**Response:** We thank the reviewer for raising this important question regarding seasonal bias correction. Please see lines 285 to 289.

Our response is as follows:

**The downscaling framework inherently captures seasonal signals.**

In the current version, we adopted a multivariate Random Forest approach that uses not only ERA5 2 m temperature but also high-level atmospheric variables (temperature, humidity, wind components, and geopotential height at 300 hPa and 500 hPa) as predictors. These variables carry strong seasonal cycles, allowing the model to implicitly learn and represent seasonal variations during training.

**Validation for the ablation season shows no substantial seasonal bias.**

To specifically assess model performance during the key study period (summer

ablation season), we evaluated the agreement between downscaled and observed meteorological variables for June–September (Fig S8). The  $R^2$  values for air temperature, relative humidity, wind speed, and incoming shortwave radiation during the ablation season are 0.96, 0.94, 0.90, and 0.95, respectively. These results indicate that the downscaled forcing does not exhibit systematic seasonal bias during the summer heatwave period.

**Explicit seasonal bias correction was not applied, but we acknowledge its value.**

Although we did not perform explicit seasonal bias correction (e.g., seasonally stratified downscaling or post-hoc adjustment), the evidence above supports that the current downscaled data are reliable for analyzing anomalies during extreme events. We agree with the reviewer that explicit correction represents a valuable improvement and have added this as a limitation in Section 6.4, highlighting it as a direction for future work.

We appreciate the reviewer’s insightful comment, which has helped us strengthen the methodological rigor of our study.

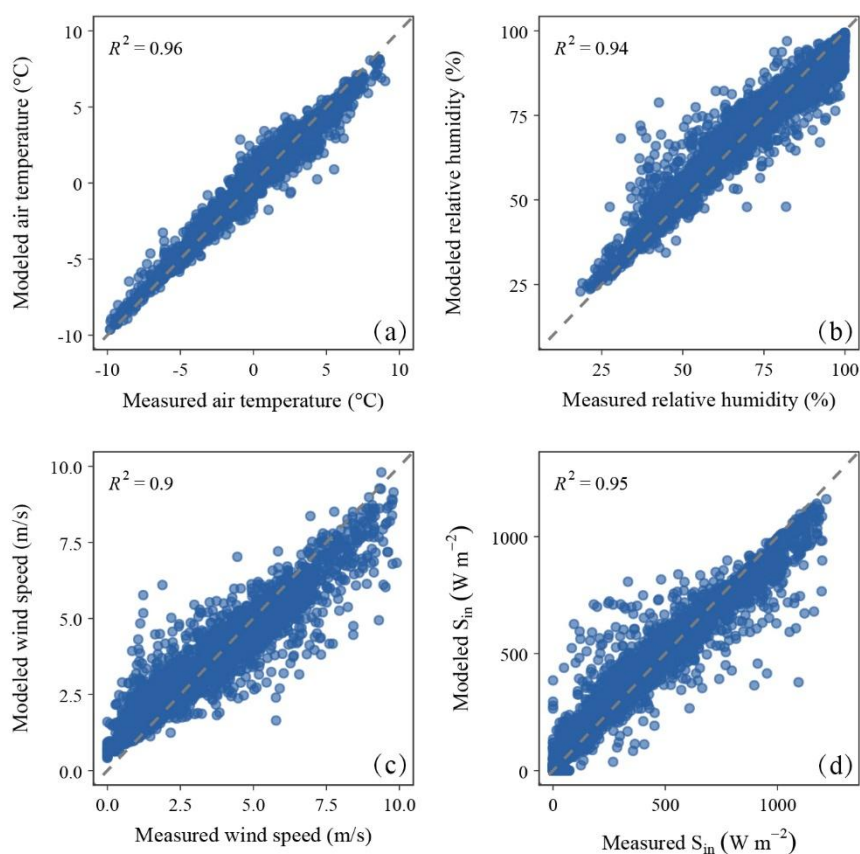


Figure S1. Comparison of observed and ERA5 downscaled (a) air temperature, (b) relative humidity, (c) wind speed, and (d) incoming shortwave radiation at the daily scale during the ablation season (June–September).

**Comment 22:** L270: How was this lapse-rate determines. It has been shown that the lapse rate on glaciers can strongly vary from the environmental lapse rate.

**Response:** We thank the reviewer for raising this important point. We fully agree that the lapse rate on glaciers often differs substantially from the environmental lapse rate and can exhibit strong spatiotemporal variability across different glaciers and seasons. Please see lines 300 to 308.

In this study, we did not assume a fixed or environmental lapse rate. Instead, the lapse rate was treated as a key parameter and determined through model calibration. Specifically, using in-situ observations from the automatic weather station (AWS) on the glacier surface and the energy-mass balance model (EBFM), we employed a multi-objective optimization algorithm (NSGA-II) to calibrate the lapse rate together with other major parameters, ensuring that the simulated glacier mass balance matched the observations as closely as possible.

We also conducted sensitivity tests for the temperature lapse rate (see Sect. 4.2 and Supplementary Table S3). The results show that when the lapse rate varies within  $\pm 0.1$  °C/1000 m of the calibrated value, the corresponding change in annual glacier-wide mass balance is approximately  $\pm 141.4$  mm w.e. Although this sensitivity is relatively high, this uncertainty is within the overall mass balance uncertainty (264.3 mm w.e.) and does not alter the main conclusions regarding the extreme melt event and its driving mechanisms. We have presented these sensitivity results in the main text.

Comment 23: L270: ‘the parameter’ -> ‘a calibrated parameter’

**Response:** Yes, we made revisions.

Comment 14: L271: Where does the albedo come from? Can the albedo at the AWS location be considered representative of the whole glacier.

**Response:** We thank the reviewer for raising these two important questions regarding the source and spatial representativeness of albedo. Our response is as follows:

### **(1) Albedo data source**

At the automatic weather station located at 5,700 m a.s.l. on Sangqu Glacier, we installed radiation sensors to continuously measure incoming shortwave radiation ( $S_{in}$ ) and outgoing shortwave radiation ( $S_{out}$ ). Surface albedo was directly derived as the ratio  $S_{out} / S_{in}$  at the AWS site. This observed albedo time series was used for model calibration and validation. Please see lines 364 to 370.

### **(2) Spatial distribution of albedo in the model**

In the energy-mass balance model (EBFM), albedo is not assigned using the point observation alone. Instead, we employed physically based parameterization schemes: snow albedo is parameterized following Oerlemans and Knap (1998) and Bougamont et al. (2005) as a function of temperature and humidity, while ice albedo follows Mölg et al. (2008) based on dewpoint temperature. These schemes, combined with distributed inputs (topography, solar geometry), allow albedo to be calculated for each grid cell across the glacier, thus capturing spatial heterogeneity. Please see lines 227 to 230.

### **(3) Calibration and validation**

Key parameters in the albedo schemes (e.g., snow albedo decay coefficients, reference ice albedo) were calibrated using the AWS-observed albedo together with multi-objective optimization (NSGA-II) to ensure the best agreement between simulated and observed glacier-wide mass balance. Fig. 2a in the main text compares the simulated daily albedo with observations at the AWS site. The good agreement ( $R = 0.76$ ,  $RMSE = 0.1$ ) confirms the reliability of the parameterization at the point scale.

#### **(4) Limitations of point-scale albedo and model handling**

We fully acknowledge that a single-point albedo measurement cannot fully represent the spatial albedo distribution across the entire glacier. This is precisely why we adopted a distributed modeling approach: the calibrated albedo parameters are applied to all grid cells, and the model accounts for topographic and radiative effects to simulate spatial variability. To quantify the uncertainty associated with albedo parameterization, we performed sensitivity tests (Sect. 4.2 and Table S2). The sensitivities of annual glacier-wide mass balance to snow albedo and ice albedo parameters are 123.2 mm w.e. and 126.3 mm w.e., respectively, and these uncertainties are included in the total mass balance uncertainty (264.3 mm w.e.). They do not affect our main conclusions regarding the identification and mechanisms of the extreme melt event.

Comment 25: L281: What emissivity value was used?

**Response:** We thank the reviewer for this detailed question. In the calculation of outgoing longwave radiation, we employed the Stefan-Boltzmann law:  $L_{out} = \epsilon\sigma T_s^4$ , where  $\epsilon$  is the surface emissivity,  $\sigma$  is the Stefan-Boltzmann constant, and  $T_s$  is the surface temperature. Following the parameterization scheme of van Pelt et al. (2012) used in their Arctic glacier energy-mass balance model, we set  $\epsilon = 1$ . Please see lines 219 to 220.

To validate this parameterization, we compared the simulated  $L_{out}$  with AWS observations (Fig. 2c in the main text). The good agreement ( $R = 0.95$ ,  $RMSE = 11 \text{ W m}^{-2}$ ) confirms the applicability of this emissivity value in our study.

Comment 26: L281: What was used for the surface roughness and why?

**Response:** Regarding the parameterization of sensible and latent heat fluxes, we clarify as follows:

In this study, we did not employ the conventional aerodynamic method that relies on surface roughness for calculating sensible heat flux ( $Q_H$ ) and latent heat flux ( $Q_L$ ). Instead, we adopted the parameterization scheme of Klok and Oerlemans (2002), which uses bulk transfer coefficients to directly compute the turbulent fluxes, thereby avoiding explicit dependence on surface roughness length and stability correction functions. Please see lines 220 to 226.

We acknowledge that the bulk transfer coefficient method simplifies the representation of near-surface turbulent processes and may not fully capture stability effects. To assess the associated uncertainty, we conducted sensitivity tests for the

exchange coefficients (Sect. 4.2 and Table S2), which show that the impact on annual mass balance is within acceptable limits. Moreover, the validation of energy balance components against AWS observations (Fig. 2) supports the adequacy of this parameterization for our study.

**Comment 27:** L286-288: I fail to see the relevance here, and in this entire paragraph – everything should be calculated relative to the AWS elevation?

**Response:** We deleted content that was not very relevant to this paper.

**Comment 28:** Figure 2: What does the color scale correspond to?

**Response:** We redrew Figure 2 and used a new color scheme.

**Comment 29:** L326: How were the glacier-wide values obtained? Since your model does not seem to account for any type of snow redistribution process, I wonder if a validation against a glacier-wide mass balance extrapolated from stake data is useful here.

**Response:** We thank the reviewer for these two interconnected and insightful questions. Our responses are as follows:

### **(1) How glacier-wide mass balance values were obtained**

The glacier-wide mass balance in this study was simulated using the distributed energy-mass balance model (EBFM). The glacier was discretized into  $100\text{ m} \times 100\text{ m}$  grid cells, and energy and mass balances (including melt, snowfall, sublimation/evaporation, and refreezing) computed for each cell independently. The glacier-wide mass balance was then obtained by area-weighted averaging. Please see lines 265–268 and lines 189–191, respectively.

### **(2) Snow redistribution processes and the validity of validation against glacier-wide mass balance**

We fully agree with the reviewer that snow redistribution processes (e.g., wind-driven snow transport, avalanches) can significantly influence the spatial distribution of mass balance on many glaciers, especially in complex terrain or regions with strong winds. We acknowledge that the current EBFM model does not explicitly account for snow redistribution, which is indeed an important simplification.

Nevertheless, we consider the validation against glacier-wide mass balance derived from stake measurements to be useful for the following reasons:

Glacier-wide mass balance is a core target for distributed model validation. Distributed models are designed to simulate spatial patterns, but validation against the observed glacier-wide mass balance remains a critical test of overall model performance. In our study, we validated the model against both glacier-wide mass balance (Fig. 2f, 2g) and point-scale mass balance at different elevations (Fig. 2e). The

point-scale comparison provides additional constraint on the elevation-dependent performance of the model.

We employed a multi-objective optimization algorithm (NSGA-II) to calibrate the model using both point-scale (stake) and glacier-scale (total mass balance) observations. Although snow redistribution is not explicitly represented, the calibration using point observations can indirectly capture its average effect on the spatial distribution. The limitations of this omission have been added in lines 840–842.

**Comment 30:** L368-369: missing references here.

**Response:** Yes, we added references. Please see lines 599–600.

“This finding is further substantiated by two independent datasets (Hugonnet et al., 2021; Dussaillant et al.; 2025).”

**Comment 31:** L388: How many are there?

**Response:** We rewrote this sentence. Please see lines 575–577.

“To evaluate the robustness of our proposed dual-threshold framework across different climatic regimes and record lengths, we applied it to five glaciers with long-term mass balance observations on the Tibetan Plateau (Fig. 5 and Fig. S4).”

**Comment 32:** L432: I don’t think I’ve seen any mention of the spatial resolution at which the model was run? Were snow redistribution processes (from avalanches and wind) accounted for in any way?

**Response:** We thank the reviewer for raising these two points. Regarding the spatial resolution, the distributed energy-mass balance model (EBFM) was run at a grid spacing of 100 m. This information has been added to the revised manuscript in Sect. 4.1. Please see lines 265–268.

Regarding snow redistribution processes (e.g., wind-driven transport, avalanches), we confirm that the current EBFM model (van Pelt et al., 2012) does not explicitly account for these processes. We have added a clear statement acknowledging this simplification in Sect. 6.4 (Research Limitations) of the revised manuscript, along with a discussion of the associated uncertainties. We appreciate the reviewer’s thorough reading, which has helped us better clarify the model’s applicability and limitations.

**Comment 33:** Figure 4: It would be helpful to plot the net energy balance on top of this bar plot. Also, were there any altitudinal variations?

**Response:** We thank the reviewer for this constructive suggestion. We have revised Figure 4 by adding the net energy balance (the sum of net shortwave radiation, net longwave radiation, sensible heat flux, latent heat flux, and glacier heat flux) as an overlay on the bar plot, while retaining the individual energy balance components. The revised figure better illustrates the contribution of each component to the melt energy.

Regarding altitudinal variations, we agree with the reviewer that energy balance components exhibit clear elevation-dependent patterns. Accordingly, we have added a supplementary figure (Fig. S9) showing the altitudinal distribution of melt energy across different elevation bands. A brief summary of these results has been included in Sect. 5.1 of the revised manuscript.

We appreciate the reviewer's careful reading, which has helped us better present the spatial heterogeneity of the glacier surface energy balance.

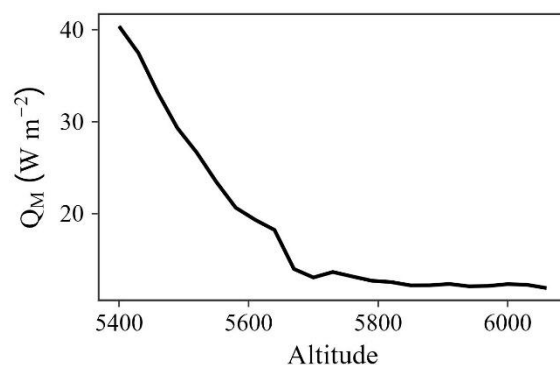


Figure S9. Melt energy versus elevation for the 2021–2022 mass balance year.

**Comment 34:** L522: why were these glaciers chosen specifically?

**Response:** We thank the reviewer for this question. The selection of these glaciers was based on the following considerations:

**Climatic representativeness:** The Tibetan Plateau and its surrounding regions are characterized by harsh environmental conditions, making long-term continuous glacier mass balance observations extremely challenging. The selected glaciers are located in the central (Xiao Dongkemadi Glacier), northern (Qiyi Glacier, Ningchan No. 1 Glacier), and western (Guliya Glacier, Anglong Glacier) Tibetan Plateau, covering the major climatic regimes of the region—from the relatively humid central plateau to the hyper-arid western plateau. This distribution allows us to capture the diversity of glacier responses to extreme events under different climatic backgrounds.

**Spatial coverage and comparative analysis:** By selecting glaciers distributed across different regions of the plateau, we aim to evaluate the transferability and robustness of our dual-threshold framework under varying climatic conditions and glacier types, thereby enhancing the generalizability of the methodological approach.

We have added a corresponding explanation in Sect. 5.2 of the revised manuscript to clarify the selection criteria and the climatic representativeness of these glaciers. Please see lines 262–265. We appreciate the reviewer's careful reading, which has helped us better articulate the rationale behind our study design.

**Comment 35:** L541-542: references missing.

**Response:** Yes, we have added the relevant references. Please see lines 599–606.

**Comment 36:** L738: this whole paragraph sounds like a repetition of what was stated above.

**Response:** We thank the reviewer for pointing this out. We agree that this paragraph largely repeated content presented earlier. Accordingly, we have removed this paragraph from the revised manuscript.

In addition, we have substantially rewritten Sect. 6.1 (Atmospheric circulation) to provide a more focused analysis of the circulation anomalies associated with the 2022 extreme melt event and their driving mechanisms, avoiding redundancy with the preceding sections. The revised content can be found in Sect. 6.1 of the updated manuscript.

We appreciate the reviewer's careful reading, which has helped us improve the conciseness and logical flow of the manuscript.

**Comment 37:** L755: I am missing a discussion of the limitations from the energy-balance model and its application.

**Response:** We thank the reviewer for raising this important point. We agree that a thorough discussion of the limitations of the energy-balance model and its application is essential for a transparent presentation of the study.

Accordingly, we have added a new Section 6.2 (Model limitations and uncertainties) in the revised manuscript, where we systematically discuss the main limitations of the energy-mass balance model (EBFM) used in this study, including:

- Uncertainties associated with spatial extrapolation of meteorological variables from a single AWS.
- Simplified parameterization schemes for key parameters (e.g., albedo, turbulent exchange coefficients).
- The absence of snow redistribution processes (wind-driven transport, avalanches).
- The assumption of fixed glacier geometry and its implications for long-term simulations.
- The impact of sparse observational data on model calibration and validation.

We also assess how these limitations affect the main conclusions of the study and outline directions for future improvements. Please refer to Sect. 6.2 of the revised manuscript for details.

We appreciate the reviewer's valuable suggestion, which has significantly improved the transparency and completeness of the discussion on model applicability.