

1. References are inadequate. The citations in the first few paragraphs of introduction appears to prefer popular recent papers that might show up at the top in a quick web research. Some basic references on the topics discussed are absent conspicuously.

More references have been discussed and added to the introduction. In detail:

- We expanded the discussion of satellite laser and radar altimetry, which was previously missing from the introduction.
- Acknowledged recent advances in proxy-based downscaling of geodetic mass balance, incorporating Banerjee A et al., 2023

2. I am unsure about the claim of introducing a new method (L70, Sect3, ...). Many papers, including those cited here, used albedo and other proxies to obtain mass-balance estimates. We agree with the reviewer that the use of albedo and other remote-sensing proxies to infer glacier mass balance has a long history. In the revised manuscript, we have removed any implication that the approach is conceptually “new” in the sense of introducing a previously unknown proxy–mass-balance relationship.

Our intention was instead to highlight the specific combination and scaling strategy used in this study:

- Calibrating glacier-wide MODIS summer albedo anomalies against in-situ mass-balance anomalies on a regional scale
- Linking these anomalies to long-term geodetic mass balance from Hugonnet et al. (2021) to derive absolute annual mass-balance estimates for thousands of glaciers,
- Applying this framework consistently at regional scale across the Alps, Scandinavia, and Svalbard.

This approach applied to all glaciers in the region, including those with no in-situ data, and the conversion of anomalies into absolute values using decadal geodetic constraints is what distinguishes our implementation from prior single-glacier or small-region albedo-based studies.

3. Assuming a constant gradient $\frac{db}{d\alpha}$ for a given region needs very strong justification.

Please compute the gradient for each of the WGMS glaciers using annual observations of glaciological balance and mean albedo. Then plot the distributions separately for each of the regions. If these distributions are not sharply peaked ones - you cannot make this assumption. LOOCV may not be the best method to check robustness here. The best-fit gradient, say for the Alps was computed with 641 observations. So is not going to vary drastically if you omit ~20 of these observations. As you can see the percentage spread in the gradient (L210) increase from 4%, to 17% and 32% as the number of data points change from 641 to 355 and 112.

You may see from Table 2, the error metrics actually suggest a constant global gradient value shows essentially the same performance as the regional values. This is most likely suggesting the inadequacy of the metrics in testing the performance of the model.

It may better to find a glacier-specific value of the gradient following Banerjee et al (2023) as that does not require in-situ data. Incidentally, they report quite different gradient values of ~12 and ~18 mwe for Argintere and Saint-Sorlin Glaciers in the Alps (Table S7) for the proxy minimum albedo.

In response to this comment, we performed additional analyses to 1) evaluate the variability of glacier-specific gradients, 2) reassess the robustness of regional gradients using an alternative cross-validation strategy, and 3) test the feasibility of deriving glacier-specific gradients from geodetic data alone.

1) Glacier-specific gradients from glaciological data

We first computed individual mass-balance–albedo anomaly gradients for all WGMS glaciers with continuous annual observations over the 2000–2019 period. For each glacier, the gradient was estimated by linear regression between annual glaciological mass-balance anomalies and corresponding summer albedo anomalies. The resulting regional statistics are:

European Alps (21 glaciers):

mean gradient = 11.13 m w.e., standard deviation = 2.57 m w.e.

Scandinavia (11 glaciers):

mean gradient = 15.86 m w.e., standard deviation = 6.02 m w.e.

Svalbard (2 glaciers):

gradients = 4.50 and 5.85 m w.e.

These results show that, for the European Alps and (to a lesser degree) Scandinavia, the glacier-specific gradients exhibit reasonably low variability around a regional mean, supporting the use of a regional representative value. In Svalbard, the limited number of long observational records prevents a meaningful statistical characterization

2) Replacement of LOOCV with 3-fold cross-validation

We agree with the reviewer that leave-one-out cross-validation (LOOCV) may not be optimal given the large number of annual observations per glacier and the relatively small number of independent glaciers. We therefore replaced LOOCV with a 3-fold cross-validation at the glacier level.

For each region, the glaciers were randomly divided into three subsets. In each iteration, gradients were estimated using two-thirds of the glaciers and evaluated on the remaining subset, ensuring that validation is performed on glaciers not used to estimate the gradient.

The resulting gradients and performance metrics are:

- **European Alps**

$$\frac{db}{d\alpha} = 10.78, 10.29, 10.74 \text{ m w. e.}$$

$$\text{MBE} = -0.045 \text{ m w.e.}, \text{MAE} = 0.35 \text{ m w.e.}, \text{RMSE} = 0.45 \text{ m w.e.}, R^2 = 0.60$$

- **Scandinavia**

$$\frac{db}{d\alpha} = 14.1, 12.0, 12.5 \text{ m w. e.}$$

$$\text{MBE} = -0.06 \text{ m w.e.}, \text{MAE} = 0.61 \text{ m w.e.}, \text{RMSE} = 0.80 \text{ m w.e.}, R^2 = 0.43$$

- **Svalbard**

$$\frac{db}{d\alpha} = 4.3, 6.9, 5.5 \text{ m w. e.}$$

$$\text{MBE} = -0.18 \text{ m w.e.}, \text{MAE} = 0.33 \text{ m w.e.}, \text{RMSE} = 0.43 \text{ m w.e.}, R^2 = 0.31$$

The consistency of the gradients across folds, together with stable error metrics, supports the robustness of using regional gradients derived from independent glaciers.

3) Glacier-specific gradients from geodetic data

Following the reviewer's suggestion and the approach of Banerjee et al. (2023), we attempted to derive glacier-specific gradients without in-situ data, using the 5-year geodetic mass-balance estimates (2000–2004, 2005–2009, 2010–2014, 2015–2019) from Hugonnet et al. (2021). This yields only four points per glacier for regression.

For WGMS glaciers in the European Alps, the resulting gradients span a very wide range, including negative and unrealistically large values (e.g., -9.5 to $+33.6$ m w.e.). The full list is provided here below:

RGI60-11.00002	6.407520
RGI60-11.00080	12.528509
RGI60-11.00106	10.825690
RGI60-11.00116	-4.415460
RGI60-11.00190	-0.935892
RGI60-11.00251	33.582806
RGI60-11.00289	11.569489
RGI60-11.00593	3.120621
RGI60-11.00597	2.658384
RGI60-11.00603	5.566993
RGI60-11.00647	1.608230
RGI60-11.00719	2.509030
RGI60-11.00781	4.015523
RGI60-11.00787	4.768256
RGI60-11.00804	5.595338
RGI60-11.00843	8.360002
RGI60-11.00897	2.824240
RGI60-11.01238	11.237468
RGI60-11.01450	-0.032161
RGI60-11.01509	6.511951
RGI60-11.01704	-2.057295
RGI60-11.01776	2.108883
RGI60-11.01834	-0.468828
RGI60-11.01863	2.711169
RGI60-11.01876	-0.485078
RGI60-11.01987	3.770623
RGI60-11.02072	-3.171651
RGI60-11.02249	1.514736
RGI60-11.02679	-6.779121
RGI60-11.02704	-4.272475
RGI60-11.02746	-5.794192
RGI60-11.02764	-5.097536
RGI60-11.02766	3.229170
RGI60-11.02773	-2.305258
RGI60-11.02774	0.020193
RGI60-11.02801	0.164339
RGI60-11.03084	-1.142738
RGI60-11.03638	-5.569017
RGI60-11.03648	-9.530868
RGI60-11.03651	10.561447
RGI60-11.03671	-0.914088
RGI60-11.03674	-8.190754

This large dispersion reflects the combined effect of the small number of points available for regression and the noise associated with 5-year geodetic estimates.

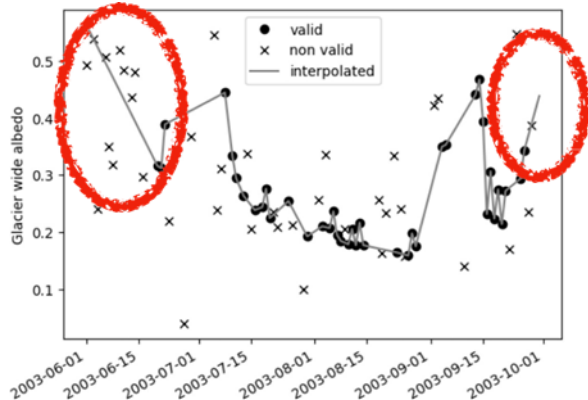
While we acknowledge that some glacier-to-glacier variability in the mass-balance–albedo relationship exists, our additional analyses demonstrate that regional gradients derived from in-situ observations represent a more robust compromise.

4. You acknowledge that annual geodetic data of Hugonnet et al. (2021), is not useful. However, you use it for some kind of validation in Sect 4.2. Moreover, the ‘validation’ in Sect 4.2 does not appear reasonable. The ‘MODIS’ values are derived by a linear transformation of the albedo, where the linear transformation is obtained by fitting the latter to the ‘WGMS’ values. So ‘MODIS’ and ‘WGMS’ values must have a correlation, which of course will not go away completely by adding a random shift (I presume it is the decadal \mathbf{B} for each glacier that you add to get ‘MODIS+DEM’).

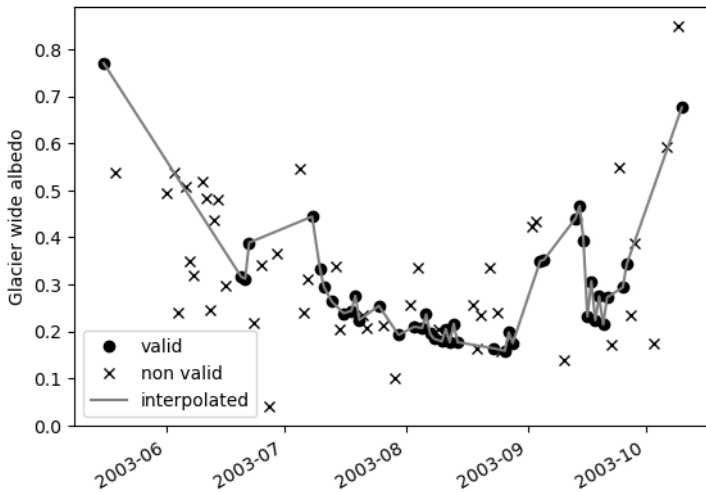
We would like to clarify that Section 4.2 does not aim to provide an independent validation of the proposed method. The formal validation of the albedo-anomaly approach is performed in Section 4.1, where annual mass balance estimates are directly compared against independent WGMS glaciological measurements using standard error metrics.

The purpose of Section 4.2 is instead to quantify the incremental value of incorporating MODIS albedo information into the geodetic mass-balance dataset of Hugonnet et al. (2021). It is well established that the Hugonnet et al. dataset yields its most robust results when averaged over multi-annual or decadal periods, whereas annual values alone are dominated by noise. In Section 4.2, we therefore do not interpret the comparison as a strict validation exercise. Rather, the comparison is designed to illustrate how much the agreement with glaciological observations improves when annual geodetic estimates are augmented with MODIS-derived albedo anomalies, relative to using geodetic data alone. In this sense, Section 4.2 should be interpreted as a demonstration of added skill rather than a validation.

5. Please explain the choice of albedo for 1Jun and 30Sep clearly. For example, in Fig 3 how do you decided on the first and the last points? The mean albedo of ~ 0.31 shown in Fig. 4 for year 2003 is larger compared to most of the valid observations (black circles) and is dependent strongly on the choice of the end points



As specified in section 3.1 “If the start or end of the time series contained fewer than 70% valid pixels, we used the average albedo from the nearest preceding (for the start) or subsequent (for the end) day that met the threshold, enabling interpolation to be performed across the entire period.”. To clarify this point we substituted figure 3 including also the nearest preceding and subsequent valid points:



6. I am surprised that the authors do not attempt an error analysis. The uncertainty in the minimum glacier-wide albedo due to systematic measurement errors, and cloud cover needs to be estimated, and propagated appropriately into the mass balance estimates. The effect of the endpoints (see point 4 above) on the mean albedo needs to be included in the error analysis as well. Similarly, the glacier-specific spread in the gradient should also be estimated (see point 2 above) and incorporated. The glaciological mass balance estimates typically have significant errors as well, and will contribute to the errors in your computation. It is easy to take into account the above sources of error by adding appropriate noise to the inputs in a Monte-Carlo. Be mindful about the well-studied systematic bias between geodetic and in-situ mass balance measurements.

We have now implemented a comprehensive uncertainty analysis, explicitly quantifying and propagating uncertainties associated with all terms entering Eq. (1): from Eq. (1), annual glacier mass balance is expressed as:

$$B_i = (\alpha_i - \bar{\alpha}) \frac{db}{d\alpha} + \bar{B},$$

where each term is affected by uncertainty. We explicitly quantify and propagate the following 1- σ uncertainty components:

1) Uncertainty associated with glacier-wide summer albedo anomaly ($\alpha_i - \bar{\alpha}$)

The uncertainty on annual glacier-wide summer albedo, accounts for measurement noise, glacier size and sensitivity to cloud-cover filtering. We estimate σ_{α_i} , i.e. the uncertainty associated with glacier wide summer albedo anomaly for year i using a Monte Carlo approach with 500 realizations. In each realization:

- the fraction of valid pixels required to retain daily albedo values is randomly varied between 60 % and 80 %, capturing uncertainty related to cloud filtering;
- for each retained daily albedo value, Gaussian noise is added to represent systematic uncertainty in MODIS albedo retrievals. The standard deviation of the gaussian noise is set to $0.05/\sqrt{n}$ where n is the number of MODIS pixels covering the glacier (Daveze et al., 2018).
- the annual anomaly is then computed as $\alpha_i - \bar{\alpha} = \alpha_i - \frac{1}{N} \sum_{i=1}^N \alpha_i$, with $N = 20$ years (2000–2019).

The standard deviation of the resulting ensemble values is taken as σ_{α_i} .

2) Uncertainty associated with the mass-balance–albedo gradient

The uncertainty on the gradient $\frac{db}{d\alpha}$, denoted σ_g , is estimated as the standard deviation of gradient values obtained from the 3-fold cross-validation experiment introduced in response to Comment 3. This approach captures the glacier-to-glacier variability of the albedo–mass-balance relationship within each region and avoids reliance on individual time-series fits.

3) Uncertainty associated with the geodetic mass-balance reference

The uncertainty on the long-term geodetic mass-balance reference, $\sigma_{\bar{B}}$, is taken directly from the Hugonnet et al. (2021) dataset, which provides glacier-specific uncertainty estimates.

Propagation to annual mass-balance uncertainty:

Once σ_{α_i} , σ_g , and $\sigma_{\bar{B}}$ are defined, the 1- σ uncertainty on annual mass balance, σ_{B_i} , is computed by error propagation:

$$\sigma_{B_i} = \sqrt{\sigma_{\bar{B}}^2 + \left[(\alpha_i - \bar{\alpha}) \frac{db}{d\alpha} \right]^2 \left(\frac{\sigma_{\alpha_i}^2}{(\alpha_i - \bar{\alpha})^2} + \frac{\sigma_g^2}{\left(\frac{db}{d\alpha} \right)^2} \right)}.$$

7. It is not unexpected that your method do not perform well in Scandinavia (L296), as there is a large number of points with +ve balance (Fig 7). It is known that albedo-mass-balance-correlation do not work well for years with near zero to positive balance (eg, Banerjee et al, 2023). So I am not sure about your argument that it is because the DEM data is bad there.

You seem to miss this point about the saturation in albedo proxies in Sect 5.3, where you highlight how good they are for highly negative balance years.

We agree that the reduced performance of albedo-based approaches during years with near-zero or positive mass balance represents an important additional explanation for the lower correlation observed in Scandinavia. At the same time, we note that geodetic uncertainty also plays a role. The RMSE associated with the Hugonnet et al. (2021) geodetic mass-balance estimates is higher in Scandinavia than in the European Alps and Svalbard (Table 3). Because the long-term geodetic mean mass balance enters Eq. (1) as an anchoring term, its uncertainty propagates into the reconstructed annual mass-balance estimates, as explicitly shown by the error-propagation framework introduced in response to the previous comment. We therefore interpret the lower performance in Scandinavia as the result of both (i) reduced sensitivity of albedo to mass-balance variability in near-equilibrium years and (ii) higher uncertainty in the geodetic reference.

Regarding the discussion of proxy saturation in Sect. 5.3, we clarify that this section refers specifically to the comparison between the albedo-anomaly method and the snowline altitude (SLA) proxy. In very negative mass-balance years or for lower altitude glaciers, the SLA can reach the maximum glacier elevation once seasonal snow has fully melted, after which further mass loss is no longer reflected in SLA. In contrast, glacier-wide albedo integrated over the entire ablation season remains responsive to these conditions: persistent low albedo values provide a clear signal of enhanced melt. This distinction is most relevant for strongly negative balance years and does not contradict the reduced performance of albedo proxies during near-zero or positive balance conditions.

8. One cannot agree with your conclusion that “To validate the method, we compared our estimates against 1108 in-situ measurements from 76 glaciers across the three regions”. You have calibrated your method using these 1108 measurements, and cannot use the same data for validation.

We have introduced an independent validation strategy based on a 3-fold cross-validation at the glacier level, as described in detail in our response to Comment 3. In this revised framework, WGMS glaciers within each region are randomly divided into three subsets. For each fold, gradients are estimated

using two-thirds of the glaciers, while the remaining one-third, not used in the calibration step, are used exclusively to evaluate model performance. By iterating this procedure, each glacier is evaluated using a gradient derived from other glaciers, ensuring independence between calibration and validation datasets. We believe that this revised validation strategy addresses the reviewer's concern and provides a more rigorous and transparent assessment of the method's performance while still making effective use of the available in-situ observations.

Minor comments

L10: the approach is not 'new'

We removed "new" from the sentence (see response to comment 2)

L14: provide bias along with RMSE

Ok

L15: What controls the regional variability in model performance

Regional variability in model performance is primarily controlled by a combination of climatic regime, glacier characteristics, and data availability, which jointly influence the strength and consistency of the relationship between glacier-wide summer albedo anomalies and annual mass balance.

First, climatic conditions strongly affect the sensitivity of mass balance to albedo variability. In the European Alps, where glaciers are predominantly temperate, variations in snow persistence and surface albedo closely track interannual mass-balance anomalies, resulting in higher explained variance. In contrast, Scandinavia glaciers are characterized by higher winter accumulation and frequent years with near-zero or positive mass balance reduce the sensitivity of annual mass balance to summer albedo. In Svalbard, cold and polythermal glacier conditions, shorter melt seasons, and the frequent persistence of snow cover can limit the dynamic range of albedo variability, weakening the anomaly–mass-balance relationship.

Second, regional differences in the quality and representativeness of calibration data play a role. The European Alps benefit from a relatively dense and long-term network of in-situ mass-balance observations, enabling a more robust estimation of the albedo–mass-balance gradient. In Scandinavia and especially in Svalbard, the smaller number of glaciers with long continuous records increases uncertainty in the gradient estimation and reduces model performance.

Finally, the reliability of annual geodetic mass-balance information used as a long-term reference varies regionally, contributing differently to the overall uncertainty propagation. Regions where geodetic estimates are more uncertain tend to exhibit reduced performance even after incorporating albedo information.

L18: "Glaciers are sensitive indicators of climate change" – better avoid such generic and beaten-to-death sentence (eg, <https://www.un-glaciers.org/en/key-messages>)

The sentence has been reformulated: "Glaciers respond sensitively to atmospheric and radiative forcing, and their annual mass balance provides a critical metric for understanding their response to changing climatic conditions."

L74: rationale behind choosing the study area?

The chosen study regions (European Alps, Scandinavia, and Svalbard) were selected for two main reasons, which we have now clarified in the Introduction and further detailed in Section 2.

First, these regions encompass a wide range of glaciological and climatic settings, including temperate mountain glaciers (European Alps), maritime and subpolar glaciers (Scandinavia), and cold to polythermal Arctic glaciers (Svalbard). This diversity makes them well suited for assessing the applicability and limitations of an albedo-based approach across contrasting accumulation and ablation regimes, and climatic conditions at regional scale.

Second, all three regions benefit from relatively dense and long-term in-situ glacier mass-balance observations coordinated by WGMS. This availability of high-quality glaciological data provides a robust basis for calibrating the albedo–mass balance relationship and for evaluating estimation

performances.

L90: explain ‘dynamic mass balance’

By “dynamic mass balance” we refer to mass-balance conditions that exhibit strong interannual variability. High winter accumulation, frequent snowfall events, and strong year-to-year variability in atmospheric forcing can lead to rapid switches between more positive and negative annual mass balance.

Fig5 show three different subfigures for the three regions.

Fig. 5 has been now substituted with this:



Figure 5: Glacier-wide summer albedo annual anomalies versus annual glacier mass-balance anomalies from WGMS observations for glaciers in the European Alps (a), Scandinavia (b), and Svalbard (c). Dashed lines show the linear relationships obtained from the 3-fold cross-validation, where the mass-balance–albedo anomaly gradients ($\frac{db}{d\alpha}$) are estimated iteratively using two-thirds of the glaciers in each region. The spread between dashed lines reflects the variability of the gradient across the three cross-validation folds.