The authors attempt to develop a method for obtaining glacier-scale annual mass-balance estimates that exploits the known correlation between albedo and glaciological mass balance. However, in my opinion there are serious limitations in the design of the study, the method proposed, and validation strategy adopted.

Major comments

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.

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.

3. Assuming a constant gradient $\dfrac{db}{d\alpha}$ 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.

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 the 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 $\bar{B}$ for each glacier that you add to get 'MODIS+DEM').

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
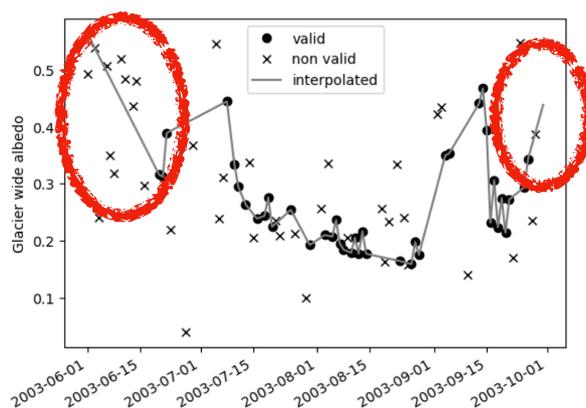
2003 is larger compared to most of the valid observations (black circles) and is dependent strongly on the choice of the end points.

6. I am surprise 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.

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.

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.

Minor comments
L10: the approach is not 'new'
L14: provide bias along with RMSE
L15: What controls the regional variability in model performance
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**)
L74: rationale behind choosing the study area?
L90: explain 'dynamic mass balance'
Fig5 show three different subfigures for the three regions.

References:
Banerjee A, Singh U, Sheth C. Disaggregating geodetic glacier mass balance to annual scale using remote-sensing proxies. Journal of Glaciology. 2023;69(276):683-692. doi:10.1017/jog.2022.89