Answer to Community Comment

In the following, we provide a point-by-point Author Response (AR) to any of the Community Comments (CC) obtained for the manuscript that was under discussion. When presenting suggestions for how the manuscript text could be revised, text is presented in *italics*.

CC1: Sect 3.1: The part of the third paragraph on assigning a date to each DEM is hard to follow. In the example discussed and in Fig 2, why do you omit the DoDs of 2009-2018 and 2008-2018 (blue) and 2008-2015 (red)? The three blue DEMs (2009, 2018 and 2023) seem to be producing inconsistent thinning estimates: '09-'18 ∼ 0.9, '18-'24 ∼ 0.9, but '09-'24 ∼ 1.8?

AR1: We rephrased the mentioned paragraph as follows aiming for more clarity:

However, for some of these DEMs the actual acquisition dates were not traceable and only the year the DEM was known, thus the exact period of the corresponding ice volume change observations could not be resolved. To address this drawback, we used the consistently reported dates of the SNFI DEMs to assign the missing dates to the swisstopo DEMs, since both datasets are based on the same aerial image acquisition flights (see Section 2). Therefore, when only one DEM is available for a given year, the date of the SNFI DEM is also assigned to the swisstopo DEM. In years with multiple DEMs with different acquisition dates, the mean acquisition date of the SNFI DEMs for that year is taken and assigned to all swisstopo DEMs in the respective year. The standard deviation in the acquisition date indicates the temporal spread of the DEMs within that year. Ice volume changes with lower standard deviations in acquisition dates are favoured to reduce uncertainties and improve the temporal constraint on the time period of ice volume changes based on swisstopo DEMs (GLAMOS, 2024d).

The ice volume changes mentioned in the comment are missing because either the DoD had a void ratio larger than 50% or because the spread of DEMs within one year, which determines the uncertainty in the date to be assigned, was too high.

Fig. 2b shows the geodetic ice volume change and not mass balance. Therefore, we understand that it could be source of confusion. Consequently, we now change Fig. 2b and show geodetic mass balance instead of volume change. This allows for better comparability and should promote better clarity.

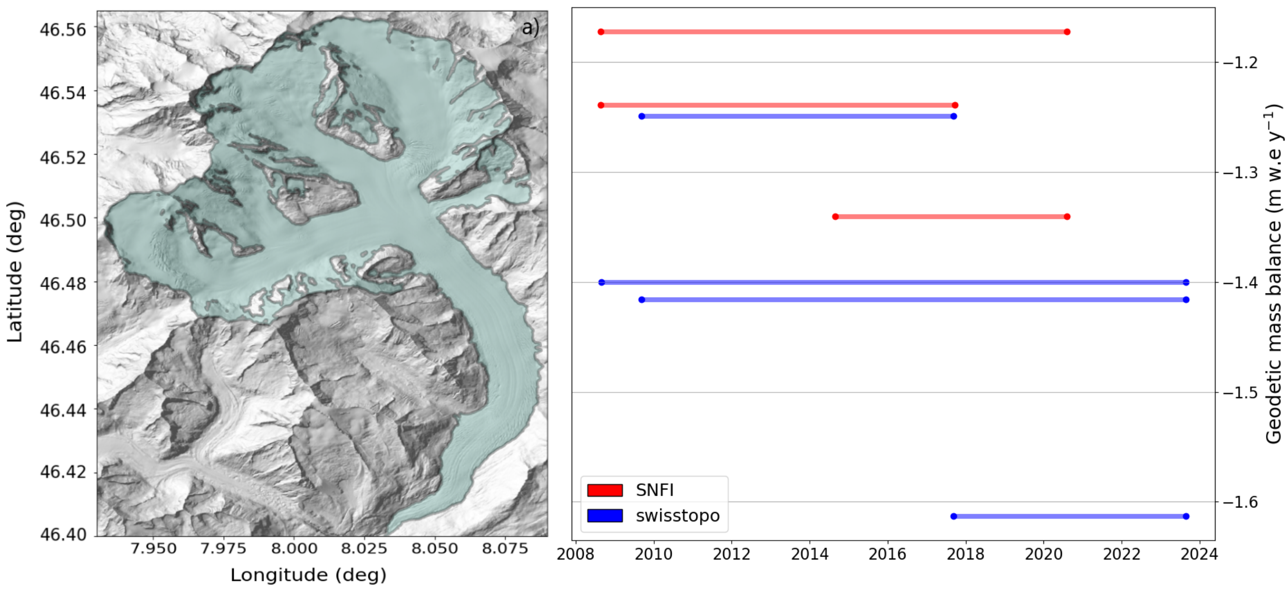

Figure 1: Temporal coverage of the geodetic mass balance estimates available for Grosser Aletschgletscher. a) Overview of Grosser Aletschgletscher with glacier outline from the SGI2016. b) Geodetic mass balance based on the SNFI DEMs (red) and swisstopo DEMs (blue).

CC2: Sect 3.2.1: The melt factor is not defined with an equation. What is "debris-coverage" - the fraction? the total area? How is it kept constant as the hypsometry evolved? How do you, if at all, validate the performance of the subdebris-melt module?

AR2: We rephrased the subsection "Debris cover" as follows to promote more clarity. For details, we refer to AR16 of Reviewer Response #2

CC3: Sect 3.2.2: Why not validate the performance of Eqs. 5 and 6, along with a glacier-specific c, using the observed geodetic balance and area-change data, wherever these are available? Please see the limitation of these method pointed out in the comments in the attached pdf. How are the issues of a different scaling exponent, stagnation of tongue, etc., known for the debris-covered glaciers taken care of in this evolution model?

AR3: In our opinion, the above-mentioned validation approach would be beyond the scope of the study. After all, we clearly acknowledge in the manuscript that this approach is rather simplistic and allows us to roughly estimate a glacier-specific area correction for every considered glacier for the few years after the inventory date. This is not a study about a glacier evolution model. The estimated area changes have a minor effect on the results, as is now shown with the added uncertainty assessment but we still wanted to integrate this factor relying on a simple and straight-forward approach. In the revised manuscript, this was expanded though in response to other reviewer comments. For changes in the manuscript we refer to AR4 of Reviewer Response #2

CC4: Sect. 3.2.3: Why do you consider only the difference in various geodetic estimates as the only source of error? What happens to the glaciers with only one or two geodetic data points? Is the final prediction uncertainty comparable with the spread in the geodetic data, or is it much less than that? If it is the latter, then how can you explain that?

AR4: We performed a new uncertainty assessment, including various sources of uncertainty. We refer to AR1 in the answer to review #1 for what this comment is concerned, and for the respective changes to the revised manuscript.

CC5: Sect 3.2.4: You seem to extrapolate only cprec . However, it is not apparent from the heading or the first several sentences of the section. Also, the rationale for doing this is not given. The best-fit values/range of cprec remained hidden for me even as I searched both the papers (this and the JoG one). Therefore, I do not have any clue if an RMSE of 0.37 is good or bad. Moreover, the consequent uncertainty in the computed winter and annual balance is not discussed. There is a fair chance that this could be very significant.

AR5: We apologize for the lack of clarity caused by our chosen section heading. We suggest rewording the heading of the section to *Extrapolation of the precipitation correction factor* to avoid confusion. The precipitation correction factor has to be extrapolated because of lacking SCAF observations for the remaining glaciers (see also Sect. 2 for the availability of SCAF observations). We suggest rewording as follows to avoid any misunderstanding:

Therefore, the parameters derived for the 87 glaciers must be extrapolated to the remaining Swiss glaciers that lack SCAF observations.

We also added the following sentence, giving the range and average values for c_{prec}, so that the RMSE of 0.37 is better put into context.

This model is employed to extrapolate the precipitation correction factor c_{prec} *to approximately 1300 glaciers. The resulting* c_{prec} *values range from 1.1 to 3.5, with a Swiss-wide area-weighted average of 1.8.*

Regarding the last part of the comment, the uncertainty introduced by the extrapolation is accounted for the new version of the uncertainty assessment. We refer to AR1 in the answer to review #1.

CC6: Sect 3.3: The logic behind the choice of the 10 glaciers is not explained (L250).

AR6: We apologize for not providing full insights. The selection of the ten glaciers is motivated by the intersection of glaciers that are part of the GLAMOS monitoring programme (and thus have detailed in situ seasonal mass balance observations available) and that have been calibrated with SCAF observations (i.e. not extrapolated c_{prec}). In the revised manuscript, we reworded as follows:

Seasonal glacier-wide mass balance is validated on a subset of ten glaciers (Fig. 4) to evaluate the performance of the calibration relying on geodetic ice volume changes and SCAF observations (see Sect. 3.2.3). These ten glaciers were selected as they are part of the GLAMOS long-term monitoring programme and have SCAF observations available for calibration.

CC7: There is no validation attempted for the daily-scale product. It is hard to imagine that constants like c_{prec}, which are derived using a seasonal-scale mass balance, will capture the variability on a daily scale. While the model resolution is daily, robust estimates may require some averaging.

Strangely enough, the MAD is given as %-age for shorter time scales, while the actual values are given in the seasonal and annual scales. The scatter in fig 5 does not scale with the value, and the noise appears additive (as opposed to multiplicative). The biases are omitted for the subseasonal case as well. In fact, In all three cases, the scatter plots show that there is a significant variation in MAD and bias from one glacier to another. Its consequence for regional to catchment scale estimates remains to be assessed.

It is unclear if the subdebris melt estimates were included in the validation. If so, what is the corresponding MAD?

AR7: Regarding the temporal scale of the results, i.e. whether the daily-scale model output can be validated, we refer to AR7 in the answer to review #1, where the same point was raised. Regarding the MAD of the daily scale validation, it is given in percentage simply because we compare different period lengths and thus different magnitudes of mass balance. For example, a same error of 0.1 m w.e. over a 10-day period could otherwise not be confronted to a 100-day period. Therefore, it is difficult to put such a bias into perspective, the reason because it is omitted. For the annual and seasonal scale, however, the temporal scale is fixed, and we therefore provided absolute values as suggested.

As already mentioned, we cannot expect that our method will have the same performance for all glaciers. Since a similar point was mentioned in CC3, we refer to AR3 for details and related changes to the manuscript.

Yes, the validation includes the effect of supraglacial debris. We added the following sentence:

Note that for model validation, the effect of supraglacial debris and changes in glacier area were included as described in the respective sections.

CC8: The limitation of the uncertainty analysis is another factor that, in my view, weakens the paper. Only using the spread of the geodetic balance values – that too without taking into account their uncertainty – is done without any serious justification. The reference to the previous JoG paper actually increases the confusion – I did not find any clear computation of the uncertainty in mass balance there (see comment in attached pdf). The MAD is not the same as the uncertainty.

The Fig. 5 of the JoG paper may suggest SCAF has less scatter for the larger glaciers like Aletsch, which is not unexpected. However, it also had the largest MAD for annual balance, which was about twice as large as the quoted mean. The Fig. 4 of the present paper also suggests a possibly larger scatter of winter balance for some of the smaller glaciers. However, the scatter for some of the smaller and larger glaciers is similar for the annual balance. Trends like this need to be investigated carefully to avoid uncertainties and biases present in the fitted model.

AR8: We acknowledge the incomplete uncertainty estimation in the submitted version (see also replies to both reviewers), and we invested considerable effort to perform an improved uncertainty assessment. Again, we refer to AR1 in the answer to review #1 for details on this. Of course, it is expected that the method has different performances on different glaciers, and as discussed in Cremona et al. (2025), the method is likely to be challenged more on small glaciers than on large glaciers. However, after checking the individual MADs and biases, no significant trend attributable to glacier size became evident. We refer to AR2 of Reviewer Response #2 for changes to the manuscript.

CC9: In fact, the biases on individual glaciers, including some of the large ones, may be significantly higher than the mean quoted in the text (L257). The biases are not discussed much except giving a mean bias, which is going to be small because of the oscillating sign. Could you check the mean absolute bias? A clear systematic bias also shows up in Fig5 , with the model overestimating melt wherever the observed melt is less (more negative) than -2 m/y. This bias and its effect on the model output, which are completely ignored, are likely to be significant, particularly on the extreme years, and demand a thorough analysis.

AR9: Unfortunately, we do not exactly understand the reviewer's concern here. In the text both the bias and the mean absolute difference (i.e. accounting for the oscillating sign) are shown and discussed. Furthermore, both the results shown in Figs 4 and 5 do not indicate any systematic skew towards higher model misfits at very negative mass balances in our opinion. As all results are clearly presented for the inspection of the reader, we would not see what additional analysis could be performed here. Of course, we acknowledge that for some glaciers, the disagreement can be substantial. For example, for Plaine Morte and Basodino, there's a model error of more than 1 m w.e. in one individual year (see Fig. 4b). But we feel that it is not possible to specifically discuss every individual data point, given that the oveall statistics are shown and are favourable.

CC10: It was not demonstrated if that the subdebris melt and glacier-geometry evolution, two pieces that were added to the model here, actually improves/changes the estimates. How much they increase the uncertainties needs to be considered as well, given the known limitations of the specific schemes used (see attached pdf).

AR10: Although the subdebris melt and the glacier-geometry evolution are methods used for first-order approximation, they are conceptually important components of the system. Even though they may not be expected to fully resolve all physical processes in full detail for every glacier, they will provide better estimates than not considering those effects at all. We however agree that some clarification and better description of those approaches was needed and have implemented this in the revised manuscript. Please refer to AR4 of Reviewer Response #2.

CC11: It may be true that the set of independent geodetic mass balances for a glacier may have a large spread. However, that may not help in getting an accurate measure of uncertainty, as these data sets may have different error bars, and sometimes there may not be enough measurements. Another alternate approach, which we had taken in Banerjee et al., 2022, JoG, is to add appropriate noise in a Monte Carlo and compute the mass balance for each case to generate a large ensemble to obtain the error bar.

AR11: We agree and have therefore updated our uncertainty assessment. We refer again to AR1 in the answer to review #1.

CC12: Another related question: could you add the modelled mean balance for Aletsch in Fig. 2b, with an uncertainty band, and see how it compares with the spread of the input geodetic values, which vary by almost 100%. Since this is the basic input calculation are based on, the uncertainties in output should be comparable to this spread. If the procedure yields a lesser spread, the questions would be why and how. That's where the uncertainty band in Fig6 may require a careful revisiting.

AR12: We appreciate the suggestion, however, we believe this addition would go beyond the intended scope of Fig. 2. The purpose of this figure is primarily to illustrate the availability of the input data rather than to perform a quantitative comparison between modelled and observed balances. Such comparisons are already presented and discussed in the validation section, where model outputs are compared against mass balance observations. Also, we revised Fig. 2 as it was source of confusion, and we thus refer to AR1 for the changes.

CC13: The model uncertainty as a function of time scale of prediction, starting from days to decades, would be a good addition to the paper.

AR13: We agree that this would be a good addition to the manuscript and accordingly added this. We refer to AR1 in the answer to review #1 for changes to the manuscript.

CC14: The discussions can be more mindful of the model assumptions, and the inferences have to be substantiated using the model output. For example, to describe the modelled variability, you refer to things like avalanching, wind-driven redistribution, Saharan dust, etc., none of which are present in the model! Also, while you consider the role of P and T, you do not look into radiation and SCAF to understand the variability of observed mass balance, particularly that of the annual and summer balance. According to me, some really interesting features in our output (e.g. see the comments to Fig. 10, in the annotated PDF) are overlooked. I get the feeling the powerful, detailed data set of forcing and response that you produce can be exploited a bit more.

AR14: We agree and adjusted and/or added several paragraphs in the Discussion section touching upon these points. We refer to the two Reviewer Responses for specific changes to the manuscript.

CC15: The writing and referencing leave scope for improvements. I have pointed out several instances of complicated sentences or incomplete information in the annotated PDF attached, which should help illustrate the point.

AR15: Many thanks for this effort - highly appreciated. Considering these suggestions, the text was rephrased and adjusted throughout the paper.

CC16: The authors may want to look beyond a set of papers familiar to them and look harder for the most appropriate references. While I could not check all the references, I did find a few surprising inclusions. Here are a few examples:

L19 van Tiel et al. (2025), which deals with buffering of runoff in an extreme year, may not be the most suitable reference for general properties of runoff from glacierised catchments. I remember referring to the excellent paper by Hock, Jansson, and Braun (2005) while discussing runoff variability.

L22 From the titles of the three references cited, while not totally unrelated, they do not seem to be the most appropriate ones on the topic of " increasing interest in accurate monitoring of glacier mass balance and runoff on the regional scale".

L25 "Dussaillant et al., 2018; Denzinger et al., 2021; ... " how were these papers chosen, when you want to introduce something as basic as geodetic mass balance? While these are interesting global and local scale studies, they may not serve as a basic introduction to the method and are also not the most relevant ones for your study area, as far as my limited understanding of the topic goes.

L225 Cremona 2025 JoG does not really show that as far as I can tell. See the comment in the annotated PDF.

AR16: Thanks for this critical re-assessment of the chosen references. We have carefully gone through the text again and added/exchanged references as suggested.

CC17: Sect 4.4 A quantitative comparison with only one of the previous reconstructions (where some of you are coauthors) is a missed opportunity I feel. Can you not compare with the other existing studies? Are there other reconstructions available for the Swiss glaciers – at least at annual scale, or for specific glaciers or regions – beyond van Tiel 2025? Apart from van Tiel 2025, you only mention Dussaillant 2025, but do not compare with their results. As it stands, the claim made in L355 is unsubstantiated. Are there not other reconstructions, say, based on remote-sensing proxies like snowline etc.? They are many in the Himalayan literature that I am more familiar with. I am half sure they may be there Swiss Alps. In general, there may be a need to connect better with the existing literature and work by other groups, through improved referencing, and by incorporating results from such studies.

AR17: No, to our knowledge, there are no other comparable reconstructions at the mountain-range scale in the region and at a (sub-)annual resolution. Of course, several studies purely based on the geodetic method are available providing long-term mass changes (e.g. Sommer et al., 2020; Hugonnet et al., 2021) but these do not add useful information in the present context. We now, however, explicitly include a comparison with The GlaMBIE Team (2025), relying on a similar approach as discussed in Dusaillant et al. (2025) and van Tiel et al. (2025). We refer to AR3 of Reviewer Response #2 for changes to the manuscript.

References