Dear Etienne Berthier,

thank you very much for your feedback.

We are very sorry that referee #2 did not have the clean version of our manuscript and we understand very well that only with the tracked-changes document, editing must have been time-consuming and annoying. Unfortunately, we do not have a reason for this technical error and hope that this time everyone will have all the revised versions of the manuscript.

During this revision period, we have formulated our point-by-point responses in more detail and separately for each referee.

We intensively discussed all the feedback we received from the two referees and you. Besides minor changes, we revised the introduction and discussion in order to contextualize our methods and results and hope that those changes meet your expectations.

All co-authors and two independent readers proof-read the manuscript.

If any further questions arise, feel free to contact us. Looking very much forward to your final decision.

Kind regards,

Joschka Geissler and Co-Authors

Dear Ben Pelto,

your positive feedback on our revised manuscript is highly appreciated. The improvements of the presented method (e.g. transferability) and also textual quality were only possible thanks to your valuable feedback during the interactive discussion. We agree that those changes greatly improved our manuscript itself and increased its benefit for the scientific community. In our now published version of the manuscript, we have further revised and improved the contextualization in the scientific context and our conclusion to emphasize this benefit even more.

Apparently, there was a technical problem with the revision, so that referee #2 did not receive the clean version of the revised manuscript. However, we hope you received all uploaded versions, even though your line-numbers also correspond to the authors-tracked-changes-document.

Best regards,

Joschka Geissler and Co-Authors

Point-by-Point answer to your feedback

L215 Change "intense glaciological surveys" to "detailed", "rigorous", or "thorough".

Sentence has been rewritten. Now L197

L543 Perhaps only recommend adjusting the density of firn, the density of ice is generally assumed to be unchanged between regions.

Thank you for the good advice. This simplifies this recommendation. New (L450):

"We recommend, however, adjusting the firn density values to the corresponding region."

L648 I would avoid stating that this is the main disadvantage of photogrammetry since the associated biases and errors were not investigated here, get right to the point.

We now avoid this more general conclusion and are now more specific on where we see the potential of our correction. New (L436):

"This correction as well as our density function allow a spatially explicit, retrospective determination and correction of geodetic mass balances and is transferable to other glaciers. The results show that the glaciological method can be greatly complemented with photogrammetric analyses, increasing the accuracy of the glaciological mass balance series, revealing regions of anomalous mass balance conditions, and allowing estimates of the imbalance between mass balance and ice dynamics."

Dear Referee,

thank you for your detailed review of our manuscript. If we understand you correctly, the clean version of our manuscript was not provided to you and we totally understand that doing a review only with the "authors tracked changes" was challenging. Considering the high number of changes made, this must have been annoying. We have no explanation why the clean version of the manuscript was not provided to you and are sorry for this inconvenience. We hope that you will receive all documents after this review. If not, we kindly ask you to refer to TC and ask for all uploaded documents.

Please note that our authors-tracked-changes document was created using the MS Word "Compare documents" tool. This is also suggested to use by TC: https://www.the-cryosphere.net/submission.html#manuscriptcomposition. Changing the resulting layout etc is – at least to my knowledge – not possible. For this review, we improved our authors-tracked-changes document considering your advice: References-changes are now highlighted and old versions of changed figures are now shown in black and white.

We also provide, as requested, a point-by-point answer to both referees separately.

In addition to your comments,

- the manuscript was proofread by all co-authors.
- all confusing sentences that we found were improved.
- we revised our introduction and parts discussion for a better contextualization of our work.

Concerning your comment on our Author Comments (AC1 and AC2), we just want to clarify that the authors are aware that a major revision takes its time. We incorporated most of the referee comments during the first revision of the manuscript and thus took the time needed. The authors again thank the referees and editor for their valuable feedback during the review process that greatly improved the quality of our publication.

Sincerely,

Joschka Geissler and Co-authors.

Point-by-point-answer:

Title: The authors write: 'Thank you for your valuable suggestion, we discussed your feedback and adapted our title accordingly'. They did not adapt the title accordingly. The title is changed, but not as suggested, it is strange that they then write accordingly. Such an example of adaptation is confusing and not very convincing.
First of all, to explain this confusing comment: During the last revision, we have discussed your comment on our title intensively. We first decided to change our title according to your suggestions but ended up not to do so. Unfortunately, we forgot to change our comment accordingly. We are very sorry for this confusion.

We discussed this comment again and would like to adapt our title following your suggestions. We agree, that the new version of the title tends less for misinterpreting the title that our methods are only be applicable to the Ötztal. Additionally, this new title adds the word 'mass balances', that represent a major part of our results.

New title: Analyzing glacier retreat and mass balances using aerial and UAV photogrammetry in the Ötztal Alps, Austria.

Line 35: 'The availability of high -resolution multi-temporal digital aerial imagery for most of the glaciers in the Alps will provide a more comprehensive and detailed analysis of climate change- induced glacier retreat.' Is it really the availability of images that provide .. I would rather say 'the availability of …. images provide opportunities for a more comprehensive and detailed analysis of glacier retreat.'
L23: We agree with your suggestion and added 'opportunities'. New: The availability of (…) imagery (…) provides opportunities for a more comprehensive and detailed analysis of climate change induced glacier retreat.

Line 67: the authors here write on airborne laser scanning and digital photogrammetry as standard methods, but here one could also differ between data sources (aerial and satellite imagery and lidar) and methods. One can also combine various data sources to obtain geodetic mass balance so this section could been improved in my opinion

This section was improved and rewritten. New version:

L 38: *"For retrieving geodetic mass balance data, different remote sensing methods exist, varying in the platform (e.g. satellite, airplane, UAV) and sensor (e.g. Laser Scanner, Optic Camera, Radar) used. Their specific benefits and limitations have been analyzed and discussed in different studies (Baltsavias et al., 2001; Bamber and Rivera, 2007; Kääb, 2005; Pellikka and Rees, 2010)."*

Line 68. Present tense on published data, e.g. are, but here you could also add newer literature, this is 20 years old.
This sentence was completely revised. We proofread our manuscript and checked for tenses.

Line 70. But the methods can be combined so don't understand this reasoning, Belart et al (2019) combines several DEMs.
We completely revised our introduction and hope that we now clarified the contextualization of our manuscript. See also comment on line 82.

Line 77. It is common to apply a correction to compare geodetic and glaciological, this is explained in Zemp et al. (2013) and many other papers, e.g. Andreassen et al (2016) (https://tc.copernicus.org/articles/10/535/2016/)
L73: We completely revised our introduction and hope that we now clarified the contextualization of our manuscript. See also comment below.

Line 82-83. I still don't understand what is unique by this, using meteorological data for correction and using drone data is not new. See point above. Maybe it is just clumsy writing, but I expect more of the revised version.
L73: We completely revised our introduction and hope that we now clarified the contextualization of our manuscript. Geodetic mass balance extrapolation is commonly performed by e.g. using field data or meteorological data (e.g. using degree-day models). We use, as you know, an additional UAV-survey that provides us an altitude-dependent correction-function. By then combining our simple degree-day approach to this correction

function we receive an altitude-dependent degree-day function. This then allows an altitude-specific correction of our geodetic mass balances that is also applicable to 2015 and 2009, by incorporating the positive degree day sums of the respective correction periods. Such a combination of methods has, to our knowledge, not been used in earlier studies.

Line 84. Here you write this calibration but the line before does not talk about calibration.
We completely revised our introduction. This sentence does not exist anymore.

Line 91. Here you refer to glaciated, in the abstract you refer to glacierized. This is not the same. Glaciated is often referred to as covered by glaciers in the past.
L94: We changed this according to your suggestion to 'glacierized'.

Line 137. I would say the uncertainty can be higher than 1 cm for stake readings when surface is uneven (ice).

L 131: For the Vernagtferner, stake reading conditions are relatively good. (Ice)surface is relatively flat and thus the mean error is typically about 1 cm. However, this is the mean error and not the maximum, that can of cause be higher.

Line 153. Why separate ablation area and accumulation are, is it not ablation and accumulation on the entire glacier?
L144: We discussed this comment. Unfortunately, we do not completely understand what you mean. We think, your comment is on this sentence:

While ablation varies between 0 and up to 4.5 m w.e. a-1 in the ablation area, accumulation only varies between 0 and about 0.3-0.4 m w.e. a-1 in the accumulation area.

With this sentence, we want to show that the variability of the specific mass balance is higher within the ablation area then in the accumulation area.

Line 204. It is common to use average glacier area, e.g. Zemp et al (2013) that you refer to. This will impact the results and you need to recalculate it to compare with glaciological balance.
L190: Please note that our geodetic mass balances were derived following the exact same workflow as it is shown in Zemp et. al 2013:

1) *Deriving the volume change (see our equation 1) is equally performed in Zemp et al 2013 using the maximum extent of the glacier:*

$$\Delta V = r^2 \sum_{k=1}^{K} \Delta h_k, \quad (4)$$

where K is the number of pixels covering the glacier at the maximum extent, Δh_k is the elevation difference of the two grids at pixel k, and r is the pixel size. Geodetic surveys are ideally carried out at the end of the ablation season, simultaneously with the glaciological survey, and preferably

2) *We use the average glacier area during volume-to-mass-conversion (see our equation 2) as it was proposed in Zemp et al. 2013:*

$$B_{\text{geod.PoR}} = \frac{\Delta V}{\overline{S}} \cdot \frac{\overline{\rho}}{\rho_{\text{water}}}, \quad (5)$$

where $\overline{\rho}$ is the average density of ΔV, assuming no change in bulk glacier density over the balance period, and $\overline{S}$ is the average glacier area of the two surveys at time *t0* and *t1* assuming a linear change through time as

$$\overline{S} = \frac{S_{t0} + S_{t1}}{2}. \quad (6)$$

Line 213. The glaciological data is mentioned in 3.2. you could define ELA in chpter 3.2. on line 144 and write how it is calculated from the equilibrium line. And remove 'For the Vernagtferner, the altitude of the equilibrium line altitude (ELA) is known from intense glaciologic surveys on an annual basis (BAdW, 2019).'
L149: Thank you for your advice. We moved the explanation of the ELA data to chapter 3.2.The sentence For the Vernagtferner… was removed.

Line 214, you mean the mean ELA for these three periods? Rewrite: 'The observed ELA of Vernagtferner varied between xxxx and yyyy in the study period, the mean ELA was … for the period ….
L197: Thank you for your suggestion. We do not think that giving mean and range information makes much sense in this case since we only have two periods and one mean value for the overall period. However, we added 'averaged over the entire period' so that the reader clearly understands that the ELA for the period 2009-2018 is the mean ELA of the respective period. New sentence:

"In contrast to most of the glaciers within our study area, the ELA (Sect. 3.2.) of the Vernagtferner is known and lies at 3217 m.a.s.l. for the period 2009-2015, 3278 m.a.s.l. for the period 2015–2018 and, averaged over the entire study period, at 3237 m.a.s.l. for 2009-2018."

The variation from 3217 to 3278 to 3237 is very small so how much will the calculation impact the result? Could comment in the text.
L197: Thank you for this comment. Please note, that because of the area-height-distribution of the Vernagtferner, those small changes of the ELA have a large influence on the glacier mass balance and on the AAR. To underline this, we derived a rough estimate of the "AAR". We derived the area above the respective ELAs and divided those areas through the total area of the Vernagtferner:

ELA [m.a.s.l.]	AAR [%]
3217	29.3
3237	24.8
3278	14.8

For other glaciers that are thinner and longer, a change of 60 m of the ELA will not make a big difference, but for the Vernagtferner this will change a lot. In our opinion, this also underlines the great advantage of integrating the ELA for the volume-to-mass conversion (thus the density assumption).

We added this information to the discussion:

" For the Vernagtferner, ELA increased 61 m between our two study periods 2009 – 2015 and 2015 – 2018. This relatively small change in elevation changed the AAR of this glacier by about 15% and thus has an influence on the bulk density that should not be neglected."

Line 224. The data should be homogenised, not only an error estimation. A bit confusing writing.
Thank you for your comment. We hope the new version of this paragraph is clearer:

L205: "To allow a comparison of the geodetic and glaciologic mass balances, both datasets were reanalyzed independently according to the Steps 1 to 4 in Zemp et al. (2013): Datasets were homogenized (Sect. 3.2 and 4.1.1.), annual glaciological mass balances were accumulated to the periods 09/2009-09/2015, 09/2015-09/2018, and 09/2009-09/2018, mean annual mass balances of the respective periods (Sect. 3.2. and 4.1.2) as well as systematic and random errors for all geodetic datasets were derived (Nuth and Kääb, 2011)."

Line 229. Calibration is done when needed so you should comment in the result if this is needed – after comparing the results. so state her that you homogenised the data and quantified the errors.
L209. As we argue in the ms, systematic errors are of great interest to our paper so calibration of those would not make sense. As a consequence, we think that the following sentence, that is now in our manuscript, is important to clarify to the reader why no full reanalysis was performed.

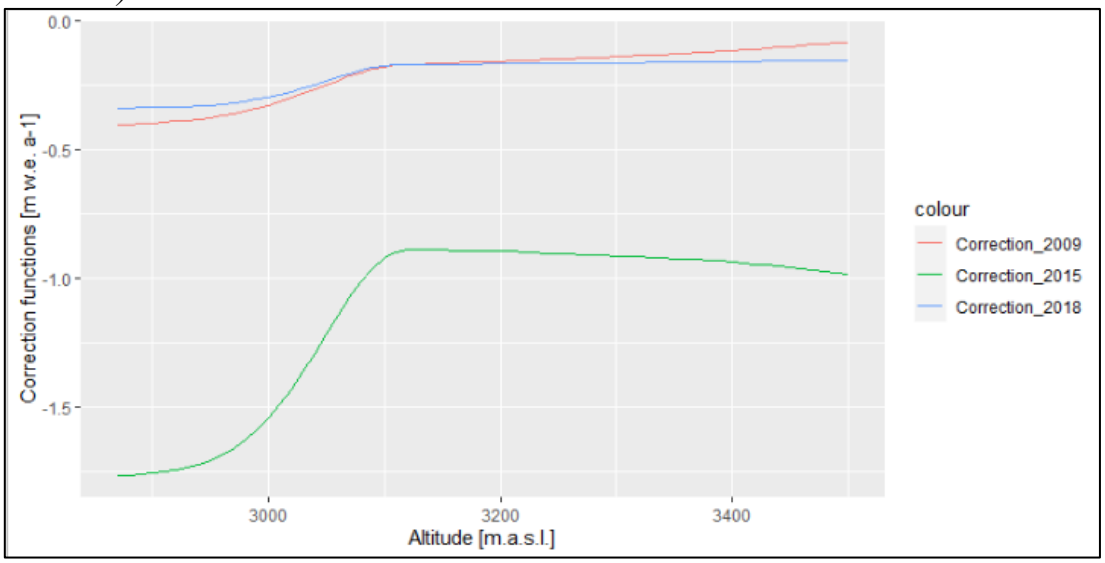
"Because one main objective of this paper was to analyze systematic differences between the two methods, iterative adjustment and calibration of the data (Step 5-6, Zemp et al. (2013)) was not performed."

Line 259. This method -> be specific, write DDF method if this is what you mean. Details on your work should be in the paper.
L228:. Changed from "this method" to "DDF method". Details on our work can be found in our Paper, Sect 4.2.

Line 262. Do you really calculate the geodetic mass balance for the correction periods, do you now calculate a correction? In table 2 you refer to correction parameters, I would liked to see the result.
We calculate the geodetic mass balance for the correction periods and add/substract those from the original geodetic mass balance as shown in table 2.

1) *We compute the surface elevation change between our UAV-survey and the aerial survey 2018 for all altitudinal bands and multiply those values with our density assumption for each altitudinal bands. The resulting correction function ($B_{geod,e,t=corr}$) can be interpreted as the geodetic mass balance of this month in 2018 (Figure 3).*

2) *By dividing this $B_{geod,e,t=corr}$ through the positive degree day sum (again for each elevation change this value changes due to the vertical lapse rate) and the number of days we derive our DDF_e . This degree day function [m w.e. a-1/(°C*d)] can be interpreted as the (geodetic) mass balance change that occurs per day per positive degree for each individual elevation band. Please note, that this function is only valid in the ablation period of the hydrologic year. Since all our correction periods are between August and September, this is valid. (Equation 4)*
3) *By multiplying this DDF_e – function with the individual (altitude-dependent) positive degree day sums and the number of days of the respective correction period, we derive the geodetic mass balance of each individual correction period (Equation 5 and plot below).*



4) *Those individual correction functions are then added/subtracted from the original geodetic mass balance of the entire study period. (See table 2)*
E.g. $B_{geod,e,09-18} = B_{geod,e,09/2009-09/2018} - B_{geod,e,t=1} + B_{geod,e,t=3}$

As you can see in the plot above, the corrections for the 27 days in 2015 will have the greatest influence on the geodetic mass balances. This is because this correction period is the longest and has the highest PDD.
For the geodetic mass balance 09/2009 – 09/2018, where the two correction functions of 2009 and 2018 of the plot above are subtracted from each other, the correction will only slightly change the geodetic mass balances.
Our comparison with the glaciologic data (Figure 8) shows that those corrections result in geodetic mass balances approaching to the glaciologic mass balances. This is the case for all periods, even though meteorologic conditions have been different between the three years.

Fig 4. Why not have 2009 and 2018 outlines of all glaciers on this figure. The font is not very readable.
Thank you for your suggestions. We added the outline of 2009 for the Hintereisferner (and the respective survey date). Font size was increased. We do not want to add further outlines of other glaciers because we want the reader also to focus on the gradient of surface elevation change around the glacier tongues. If we would add further outlines, they would partly overlay those gradients.

Fig. 5. The font not very readable. Which glacier tongue in the center-right, add name of ID.
Changed to "second glacier tongue from the right" and added a white dotted circle to clarify the position of the glacier tongue of interest. In our opinion, the font should not be bigger, and its readability is good. Please note, that all glacier tongues, that are visualized with outlines belong to the same glacier, the Hochjochferner. Thus, they all have the same ID even though they are no longer spatially connected (because of glacier retreat).

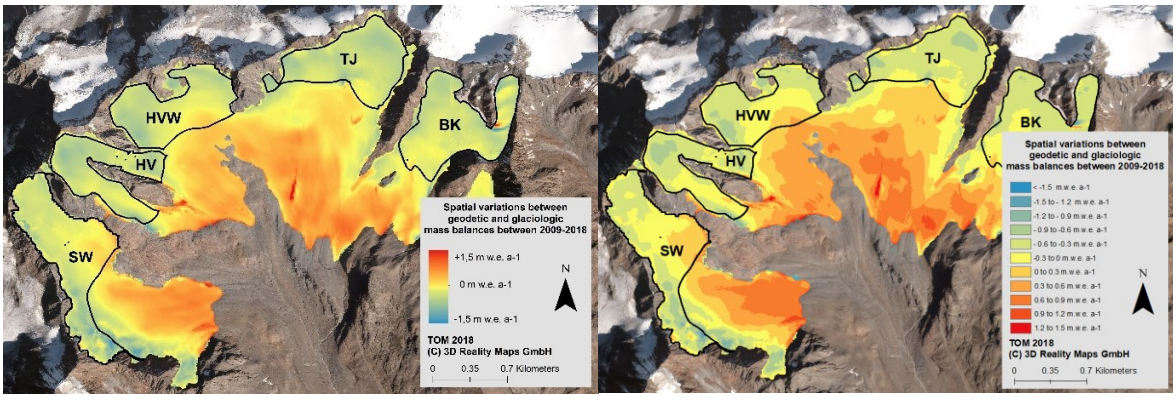
Fig. 6. This figure text is very short.
Instead of 'topography' we now enumerate the different plots b-d.

Fig. 9. Is this the corrected values?
Yes, see Sect 4.2., where the variation function was described (Equation 7). The variation function is the difference between the corrected geodetic and the glaciologic MB. Clarified in the caption.

Fig. 10. What is the scale used in this figure, it is difficult to see the different categories. Is it uneven colour scale such as rainbow map (https://www.nature.com/articles/s41467-020-19160-7)? It could be better to have the scale on the figure or use discrete classes, e.g. 8 or 10 classes. The table could be taken out and showed in a table.
Thank you for this valuable comment on Figure 10. We tried a visualisation using discrete classes (Range per class: 0.3 m w.e. a-1; see below). However, we feel like this visualisation would reduce the amount of information that the reader could get out of this figure. And, more importantly, the gradient of smaller features, for instance of debris-covered areas are important to see, since they are part of our discussion.

Concerning the choice of colours-scale. A suitable colour-scale must have three colours in order to be able to distinguish between the three important cases $B_{geod}>B_{glac}$, $B_{geod}<B_{glac}$ and $B_{geod}=B_{glac}$. Thus, any other choice of colour would not improve the scale. In our opinion, a linear scale, such as the colour-scale chosen, is appropriate.

To increase understandability, we removed "Low" and "High" from the legend and added the "0-value" to underline the linearity of the colour scale.



We agree that the table is better seperated from the figure and applied the changes in the manuscript.

Line 454. Here you talk on height changes but the figure shows changes in m w.e.

Thank you for finding this mistake. Corrected accordingly.

Line 461. 'All glaciers' are written two times in this sentence, remove one occurrence.
Again, thank you for finding this typo. The diligence with which you read the document is even more impressive considering that you only had the authors tracked changes version available. Thanks you very much for that.

Line 518. It is not easy to see the dead ice body from figure 4, even if it is marked. I suggest to show it in a subset figure, e.g. add a frame so it is possible to see it.
We are not able to extract the exact outline of the dead-ice body only with the data we have. Thus, we would like to remain with only referring to its existence and the surface elevation changes that obviously occurred because of the existing dead ice body. We checked the manuscript for misleading statements concerning this dead-ice body and removed the square from figure 4.

Line 538. It is the resulting geodetic mb that must be considered with caution.
L439: This was badly written. Thank you for pointing this out. New version of this sentence: Thus, the density assumption and the derived geodetic mass balances for this period must be considered with caution.

Line 540. Why do you not use it for the other glaciers then, or try to estimate it with the data you have on ELA and orthophotos or retrieve ELA using satellite images?
Within this paper we developed and tested a robust method to correct geodetic mass balances temporally. We compared our results to glaciologic data spatially and quantitively and derived (uncorrected) geodetic mass balances of 23 different glaciers. We showed that there is great potential if all existing photogrammetric data sets were analyzed accordingly. In our opinion, applying our method on other glaciers, including determining the altitude of the ELA from satellite imagery etc and collecting meteorological data for the respective glaciers would be beyond the scope of this ms.

Line 636. How does your method compare to other methods such as mass balance modelling being used to correct for acquisition dates?
We contextualized our correction method as well as our density-function and thus added a comparison with other studies in the introduction (see below) and discussion:

Density function:
"The presented altitude-related density function is easily applicable and need low computational effort and therefore greatly complements other existing density conversion factors, that account for e.g. firn compaction processes or rely on classification methods or modelling (Pelto et al., 2019; Reeh, 2008)."

Correction function:
"Such corrections can for instance be applied by using a simple degree-day model (Belart et al., 2019) or field measurements (Fischer et al., 2011). These methods, however, are either not suitable for retrospective corrections where no field data was collected or do not account for the spatially distributed, glacier specific accumulation and ablation patterns of each glacier (Huss et al., 2009)."

Line 645. What about Lidar surveys? And aerial imagery with poorer contrast (e.g. ice caps)?

This was a very important Feedback in our opinion and that we incorporated within our discussion. The presented method to adjust geodetic mass balances is not limited to UAV and/or photogrammetry. It can be used with many other geodetic survey setup. We specified this in our discussion:

L473: "The required geodetic survey is neither limited to a platform, nor to a sensor. Thus, all geodetic survey configurations that allow the determination of geodetic glacier mass balances from DSM differences are suited for deriving the presented correction function. We used a dedicated photogrammetry-based UAV survey, flexible and low-cost, that enabled a correction with high spatial resolution. We used a dedicated photogrammetry-based UAV survey, flexible and low-cost, that enabled a correction with high spatial resolution."

Line 648. But this was only done for one glacier. Using UAV in addition to aerial imagery for all these glaciers is not manageable. And you need to show how you can use information from one glacier to the others to make this argument valid.
See L467, where we explain how our method can be transferred to other glaciers:

"For transferability to other glaciers, individual correction functions must be determined for each glacier individually, as these are directly related to glacier-specific slope gradients, orientation, the height of the glacier tongue, and area-height distributions (Fig. 11). To determine such individual correction functions, temperature information (for the correction period as well as all periods to which the correction is applied) and an additional geodetic survey is needed to estimate the surface elevation changes during the correction period. The length of the correction period was one month within this study, however we do not expect poorer results if varying this period by 1-2 weeks"

Please also note, that with the increasing range and flight time of UAV, especially fixed-wing UAV are suitable to survey large areas and thus many glaciers could be corrected accordingly and that as soon as one correction function is derived, our method allows the retrospective correction of multiple time periods for each glacier.

Appendix or supplementary material in the end of ms. As it is only one table you could have it in the manuscript itself as ordinary table.
Applied. Please note that we also doublechecked Figure 11 as well as the table 4 for equivalent writing of all glacier names.

Table text above the table. Are this corrected values? You need to add more info.
L395, added: Geodetic mass balances were derived by using a fixed density factor.

Use dot (.), not comma (,) as decimal operator throughout in paper
Checked the manuscript again.