The Cryosphere Discussions

Interactive comment on "A minimal model for reconstructing interannual mass balance variability of glaciers in the European Alps" *by* B. Marzeion et al.

B. Marzeion et al.

ben.marzeion@uibk.ac.at

Received and published: 5 January 2012

We would like to thank Ian Howat for obtaining the reviews, and we would like to thank the reviewers for providing detailed, constructive and helpful comments on our manuscript. We were able to address all their points (as detailed below), and their suggestions have lead to significant improvements of our manuscript.

General Remark

A private discussion with Regine Hock concerning our approach to include liquid

precipitation into the positive side of the mass balance equation (page 2804, lines 14-18 of the discussion paper) has lead us to change the set up of the mass balance model. This change had only a very minor quantitative, and no qualitative impact on the results of our study, but it has caused some changes to the manuscript that were done not in response to the two reviews (but with relevance to some of the points raised by the reviewers, see below).

As is mentioned in the discussion paper, the performance of our model (measured by rmse) suffers if only solid precipitation is taken into account in the mass balance equation. If a pure reconstruction of the mass balance variability is the objective, it is therefore reasonable to include liquid precipitation into the mass balance – even though it is not simple to understand, or even counterintuitive, why this should improve model performance. But apparently, the amount of liquid precipitation includes information that is relevant to the mass balance anomalies, additionally to the information that is included in the temperatures. There are several reasonable explanations how liquid precipitation may influence the mass balance (e.g., directly by percolation and refreezing, which would also change the subsurface energy balance, or by changing the ablation via altering the surface energy balance, or it may include relevant information on more indirect influences, such as cloudiness), but from our model it is not possible to conclude which is the most important influence. (Note also that it is probably reasonable to assume that measurements of total precipitation have a better quality and are more readily available than measurements of solid precipitation.) This information is lost when only solid precipitation is taken into account.

However, we do not present a pure reconstruction of the mass balance anomalies in our study, but we also attribute mass balance variability to temperature and precipitation variability. Including liquid precipitation into the positive side of the mass balance equation implies that the negative side of the mass balance equation (i.e.,

$\sum_{i=0}^{12} -\mu(max(0, T_i)))$ has to include instant runoff of liquid precipitation onto the glacier, additional to melting of glacier ice. Therefore, our parameter μ has to be bigger than it would be if only melting of ice was concerned, which may lead to an overestimation of the temperature sensitivity relative to the sensitivity on precipitation.

We therefore recalculated all the reconstructions presented in our manuscript, using only solid precipitation, and accepting a slightly higher rmse (11 mm in the mean). This changed the estimated parameter values for a and μ in such a way that temperature sensitivity essentially remained the same, but precipitation sensitivity was slightly increased. We also changed the text of the manuscript accordingly, and included a discussion along these lines in the discussion section.

Finally, we would like to thank Regine Hock for initiating this discussion.

Response to Anonymous Referee 1

Major concerns

1. **Comment:** Glacier terminus: In their simple model, each glacier is attributed an elevation to which the temperature is extrapolated. The authors assume – without any discussion – that the elevation of the glacier terminus is indicative for the entire glacier. Why? Is the elevation of the glacier terminus really the variable that best integrates glacier geometry, and is representative for the entire glacier? A lowlying glacier terminus could be explained by high accumulation, as well as a large high-elevation accumulation area. The authors definitively need to provide more details on why the glacier terminus altitude was chosen as the connection

of the glacier with climatological variables and not, for example, the median elevation, the ELA, the elevation range, or the glacier area.
Response: The glacier terminus height was chosen as the reference point because it provides a temperature threshold: If temperatures at the terminus are below freezing, it is reasonable to assume that no melt occurs at the entire glacier surface. Our model then assumes that as temperatures increase above freezing at the terminus height, the melt at the glacier surface will increase linearly with temperature. Additionally, our choice to use the terminus height as the reference point also has to be seen under the light of data availability: the median elevation and ELA are not known for all the glaciers in the Alps (note also that the ELA will vary considerably from year to year), and we do not see a direct link between glacier area and the specific mass balance we model.
We added a sentence to the manuscript describing this reasoning.

2. **Comment:** Determination of SSC parameters: It is not completely clear to me how C_T and C_P were determined. For example, why are positive values for C_T, and negative value for C_P found for individual months? A positive C_T indicates that higher air temperature leads to more positive mass balance. Can this be explained physically?
Response: On page 2810, line 1 of the discussion paper we explicitly state that the SSC parameters were determined by multiple linear regression. As we also state explicitly on page 2810, lines 21-24, that multiple linear regression is problematic, because the positive values for C_T and negative values for C_P are physically not meaningfull.

3. **Comment:** Comparison to the SSC model: A comparison of the model skill with the less parsimonous SSC model is performed and it is found that the SSC model performs worse although 24 parameters have been calibrated compared to only two in the present approach. I asked myself, if this comparison is 'allowed'. As the authors state (page 2810, line 5), the SSC model is calibrated differently than

in its original presentation (Oerlemans and Reichert, 2000). What is the impact of the simplified approach to calibrate the SSC parameters chosen here on the uncertainty in the parameter values?

Response: In footnote 4 (page 2811) we state that the limitations of the SSC model get less significant the more independent data from the glacier are available. As we state on page 2810, lines 9-11, we use our way to obtain the SSC parameters as a straw man for models trying to achieve better performance by including more parameters. Our objective is not to criticize the SSC model as it was published in Oerlemans and Reichert (2000) (we state this on page 2810, lines 4-8), but to illustrate that the apparent increase in model performance when more parameters are included is deceptive, and that cross-validating the model detects this (see also our response to the specific comment regarding page 2811, lines 14-19).

4. **Comment:** Inner-alpine glaciers: An intriguing result of the analysis is that inner-alpine glaciers (receiving less precipitation) are more sensitive to precipitation changes than glaciers with high accumulation at the flanks of the Alps. However, small and very small glaciers (they are mostly characterized by high precipitation amounts, and are often situated at low elevation) are known to show a strong dependence on precipitation variability (see e.g. Kuhn, 1995, ZGG). This contradicts the findings presented here. I would expect glaciers with very high accumulation rates to be sensitive to changes in these, as a few percent precipitation decrease would result in a significant loss in total accumulation in m w.e. a[1]. The question that needs to be addressed by the authors is whether the results based on their model are significant (and can be explained), or their results are artefacts of a model that was not constrained with separate measurements for accumulation and ablation.

Response: As mentioned above (General Remark), we now use only solid precipitation in the positive side of the mass balance. In the sense that the HISTALP

C1656

precipitation data are based on measurements, we now are distinguishing between accumulation and ablation explicitly in our model. This has lead to a generally increased influence of precipitation on the mass balances, as can be seen in fig. 10 and 11 (revised manuscript, fig. 11 and 12 of the discussion paper), in line with the argument of the reviewer. However, we found that in our results, the relative influences of precipitation and temperature are uncorrelated with the area of the glacier.

Glaciers with very high accumulation rates typically extend to low elevations, implying that their termini experience above-freezing temperatures for relatively many months of the years. This also implies that temperature variability has more time to influence the mass balance than it has for a glacier terminating at higher altitudes. We therefore think that our results are not necessarily contradictive, and not artifacts of the model. We explain this reasoning in sect. 3.3. Note also that the glaciers considered in Kuhn (1995) are not found among the lowest terminating glaciers considered in our study, but rather in the lower third (compare their elevation with fig. 12 of the discussion paper, fig. 11 in the revised manuscript), and are probably not selected representatively.

5. **Comment:** Table 1: When looking at the values of $a_\mathrm{optimized}$ and $\mu_\mathrm{optimized}$ for the individual glaciers quite some spread emerges. Could the author explain possible reasons for these differences? e.g. a obtained for Sarennes differs by almost a factor of 2 from Wurtenkees. The parameter μ is almost double for Careser compared to Gries (both glaciers are in a relatively similar climatological setting). Furthermore, is there an explanation why the rmse is relatively low (around 200-300 mm ww.e.) for glaciers in the Eastern Alps, and above 400 mm w.e. for glaciers outside of Austria. Is this related to (i) the quality of the HISTALP data base, (ii) to the model setting that works best for the climatological conditions in the Eastern Alps, (iii) to the quality of the mass balance measurements, or (iv) length of the data series?

C1657

Response: While the values of the parameter μ can be described quite accurately as temperature sensitivity, the parameter a is a bit more complicated to understand, as it may represent both a precipitation lapse rate, aeolian transport of snow, or avalanching affecting the mass balance of a glacier (page 1806, lines 15-16). Moreover, the climatic setting of Sarennes and Wurtenkees is actually quite different, with Wurtenkees (according to HISTALP data) receiving approx. 50 % more precipitation annually (both in the liquid and solid fractions) than Sarennes. Similarly, according to HISTALP Gh. Careser experiences significantly colder temperatures during summer than Griesgletscher, which explains the necessitiy of a higher temperature sensitivity.

It is not quite clear why the rmse appears to be bigger in the Western Alps. It could indeed be caused by lower HISTALP quality in the Western Alps (although we do not find any evidence suggesting this), or the model setup working better in the Eastern Alps (note however that this is a circular argument, and we do not see any reason why the general approach should work better in the East than in the West). We don't think there is reason to believe there is a generally different quality of mass balance measurements, and we can exclude the influence of the length of the time series (see fig. 6 in the discussion paper, fig. 5 in the revised manuscript).

6. **Comment:** General focus of the paper: The authors decide to focus their paper on the presentation and the validation of their model, and to provide only one exemplary application. This application is very interesting though, and I expected some more discussion of it. Rather than just focussing on the sometimes a bit lengthy presentation and validation of the model, I would suggest to emphasize the application and the interpretation of the results.
Response: We agree that there is much more in the results of the presented reconstruction than is currently presented in the manuscript. Note however that the manuscript is already quite long (41 pages in Discussion format). The analysis

of the results should be made just as thorough as the validation of the model, and if we were to include it in this manuscript, the paper would become very hard to digest. Presentation of analysis of the results is already under way in other manuscripts.

Detailed comments

- **Comment:** page 2801, line 3: Just a general remark: I am not sure whether a study that only relies on the modelling of annual mass balance values, can claim to *investigate* the interaction between glaciers and atmosphere.
  Response: We are not quite sure to what the reviewer refers to here: We actually do not use the term *investigate* anywhere in our manuscript.

- **Comment:** page 2802, line 27: What other kinds of analysis (that are not performed in this paper) would this data set also allow for?
  Response: E.g., identifying spatially coherent patterns of MB variability, identification of modes of atmospheric variability governing MB variability, etc.

- **Comment:** page 2803, line 1-13: This part could be shortend or even be omitted to save space.
  Response: We think this is the appropriate place to introduce the terminology used in the paper, as it may be confusing if the terms are introduced "on the fly" within the manuscript.

- **Comment:** page 2804, line 23: This statement is actually not understandable at this point of the paper – it rather appears puzzling. Either provide an explanation here, or move it completely to the discussion at the end of the paper.
  Response: Text removed following the change in the setup of the model (as described above).

- **Comment:** page 2805, line 4-5: "the amount, ... ". Unclear. What is done exactly?
  Response: Also this part of the text was removed (together with fig. 1, which in the new setup of the model is far less relevant).

- **Comment:** page 2806, line 3: Isn't rather Cogley (2005) meant here?
  Response: Cogley (2009) is actually the correct reference here, as we refer to the individually trained model here. Lateron in the manuscript, when all the WGI-XF-glaciers are modeled, we get the information from Cogley (2005).

- **Comment:** page 2806, line 21: I understand how one parameter (e.g. μ) can be minimized using eq. 3. But now, the equation contains two unknowns (a and μ). What is the procedure to minimize both of them? Equation 3 does not apply then anymore.
  Response: This is achieved by applying a nonlinear least-square solving algorithm (Levenberg, 1944).

- **Comment:** page 2807, line 9: Can some details about the auto-correlation lag time be given? I have troubles understanding what has been done here exactly.
  Response: The time series of the mass balances of some of the glaciers have an auto-correlation. If only one value was removed from the time series during the cross validation, because of the auto-correlation the remaining values would not be completely independent from the removed value. In order to ensure complete independence, additional values have to be removed from the time series, the number depending on the length of the auto-correlation.
  We added a brief explanation to the text.

- **Comment:** page 2810, line 1: The parameters of the SSC model were obtained based on multiple regression. I do not completely understand the procedure. Can the annual mass balance anomaly really be used to obtain a temperature sensitivity for e.g. the winter months? Some more details about the justification

of the method applied here to determine SSC parameters might be helpful.
  Response: It is of course possible to do this, but it is indeed questionable and we state this in the text (see also reply to comment 2 above).

- **Comment:** page 2811, line 14-19: This paragraph is important, but I do not completely get it. What exactly is $r_{\text{SSC,fitted}}$?
  Response: $r_{\text{SSC,fitted}}$ is the correlation one would calculate if one omits an independent validation such as the cross validation. I.e., the difference between $r_{\text{SSC,fitted}}$ and r_{SSC} is the spurious model performance that is detected by the cross validation.
  We added a brief explanation to the text.

- **Comment:** page 2815, line 13: How significant is the correlation of glacier terminus altitude and precipitation / temperature dependence of the glaciers? When looking at Fig. 12 it looks as the statement is only valid for very few data points, and for the vast majority of the glaciers no elevation dependence at all emerges.
  Response: The correlations are not particularly strong, but they are highly significant ($p < 0.01$). We added the actual values of the correlations to the text.

- **Comment:** page 2818, line 18: It might be helpful to shortly describe what reference-surface balances are, rather than just providing a reference. Many readers might not be familiar with this term.
  Response: We added an explanation to the text.

- **Comment:** page 2820, line 9: This paragraph is useful to discuss the value of the model results, and addresses a very important question: How well are mean mass balances reproduced. I suggest providing even more discussion. For example, what can be learnt from the bias shown in Figure 13? How can it be interpreted? e.g. the small biases for the glaciers in the Eastern Alps, and the huge bias for Aletschgletscher? What does a positive bias mean? Model too negative/positive?

Response: We added an explanation to the text (positive bias means model too positive). The discussion around the model bias here would be repetitive of the discussion on page 2819 lines 21-24, since – as the similar correlation values indicate – the reconstructions we compare with could just as well be considered "truth", and the comparison with the other reconstructions does not imply other caveats than those mentioned there.

References

Levenberg, K., 1944: A method for the solution of certain problems in least squares, *Quarterly of Applied Mathematics*, **2**, 164–168.

Response to Mauri Pelto

General comment:

Comment: At present the paper requires several points of clarification, and a more detailed verification on Hintereisferner in order for the reader to be able to assess the potential with confidence. I do not disagree with the particular model choices or the methods of validation. I simply need a better explanation of choices and presentation of validation on Hintereisferner and in production of a reasonable balance gradient. The detailed discussion of the model to the entire Alps is not useful without a better verification of the method.

Response: We have the impression that there is a slight misunderstanding here: We do not evaluate the model on Hintereisferner exclusively, but on 39 different glaciers in the Alps, of which we get robust parameter estimates for 15. On these 15 glaciers we base the construction of the mean model. Perhaps this misunderstanding is rooted in fig. 10 of the discussion paper, where we illustrate the validation for Hintereisferner.

But note that the exact same figure is provided for all 15 robust mass balance reconstructions of the individually trained, and the mean model, in the supplementary material.

Our method does not rely on mass balance gradients, but uses the terminus of the glacier as a reference point in order to calculate specific reference-surface mass balances. It is also not possible to calculate mass balance gradients with our model. Instead, we rely on the cross-validation procedure, which is explained in the text extensively.

We think that we were able to clarify both these points with the changes made to the manuscript, as detailed below.

Key comments

- **Comment:** 2804-4: Why is the temperature at the terminus used in the equation? This is partly raised because equation (4) requires an adjustment to the terminus anyway. Why not use the ELA where we know mass balance=0.
  Response: We do not know the ELA for all glaciers in the Alps; additionally, the ELA will vary considerably from year to year. Given these limitation, the glacier terminus provides a natural temperature threshold: If temperatures at the terminus are below freezing, it is reasonable to assume that no melt occurs at the entire glacier surface. Our model then assumes that as temperatures increase above freezing at the terminus height, the melt at the glacier surface will increase linearly with temperature.
  We added a sentence to the manuscript describing this reasoning. (See also response to the first reviewer's first comment.)

- **Comment:** 2804-24: Where is the evidence that this method improves the esti-

mate of accumulation?

Response: Because of the change in model setup, this sentence was removed here and is no longer directly relevant. However, we still discuss this in the discussion section: The evidence comes from a reduction of the rmse if liquid precipitation is included in the positive side of the mass balance equation (see discussion section of the revised manuscript).

- **Comment:** 2807-13: The scaling functions $a_{\text{optimized,cross}}$ and $\mu_{\text{optimized,cross}}$ should produce a reasonable balance gradient that can be compared to mean gradients for the Alps from Greis, Hintereisferner, Vernagtferner etc. If this cannot be done than the model output results cannot be robust. It seems vital to devote a figure to a reconstructed balance gradient.

  Response: Since our method is based on a reference point, it is not possible to calculate mass balance gradients. But note that the results from the cross validation measure how well the model performs outside the availability of measured mass balances. Therefore, the cross validation on all the glaciers shows exactly what the reviewer asks for: that the parameter estimate and reconstructed mass balances are robust if enough mass balance measurements are available for training of the individual model, and that the mean model results are robust as well.

- **Comment:** 2813-11: Figure 10 is offered as an exemplary case for testing the model. This figure is not convincing as constructed for this purpose. The time span used far exceeds that for which glacier mass balance data exist. The time span presented is too long for detailed examination of accuracy and potential bias of the model. The contention is that the model provides reasonable annual results. This needs to be carefully examined with respect to the 50 year long Hintereisferner annual mass balance record. This graph should focus just on the 1953-2003 period. After this verification provides a view of the model output versus observation in a detailed manner, than Figure 10 is fine for illustrating the

C1664

long term reconstruction and the impact of temperature and precipitation on the above, as is done in part in Figure 13. The skill scores and correlation coefficient reported are quite good, but do not illustrate the annual variations. For example is mass balance using the model underestimating the mass balance more during years with particularly low mass balances as Figure 3 would suggest, or for Hintereisferner is the situation different.

Response: Figure 10 is shown exclusively to illustrate the long term reconstruction and the impact of temperature and precipitation. The measures of model performance, such as bias, correlation, variance etc. should not be based on visual inspection. The entire sections 2.2 and 2.4 are devoted to the verification the reviewer asks for. Specifically, the skill scores and correlation coefficients reported in table 1 are based on the annual variations shown in fig. 10 (and the supplementary material), for Hintereisferner as for all the other glaciers included in our analysis. It is true that the model has a reduced variance compared to the observation, which is explicitly shown in fig. 4 of the discussion paper (fig. 3 in the revised manuscript).

- **Comment:** 2816-18: It is concluded that "Doing so increased the mean rmse of the model, indicating that our model's implicit distinction between liquid and solid precipitation over the glacier surface is more accurate than the more global, i.e. less glacier-specific, estimate of the solid fraction of monthly precipitation contained in the HISTALP data." This may be the case, but the case is not made strongly here. How much did the mean RSME improve? How accurate is the HISTALP solid fraction? Is this the HISTALP solid fraction determined for terminus elevations?

  Response: We rewrote this discussion in the revised version of the manuscript, following the change in the model setup to include only solid precipitation. Because of this, it is also far less relevant in the new version (we could have omitted it entirely, but we thought it is of interest anyway), but we also included the accu-

C1665

rate number in the manuscript: The rmse is reduced by 11 mm in the mean.

Specific comments

- **Comment:** 2801-9: Surface mass balance is more closely related to the atmospheric forcing than changes in glacier length.
  Response: We agree, and re-wrote to clarify.

- **Comment:** 2801-18: A better reference than the weak Roe and O'Neal, 2009, should be used.
  Response: We don't agree that the reference is weak, but included another reference nevertheless in order to substantiate.

- **Comment:** 2801-26: Is the long term data series for mass balance from Sarennes warrant inclusion here?
  Response: It is, and it is already included as are many other long-term measurements (see suppl. material, and our response to the general comment above.)

- **Comment:** 2803-17: reword to "The model is established for a glacier..."
  Response: Text changed accordingly.

- **Comment:** 2805-5: The amount is proportional to what temperatures specifically?
  Response: Removed an no longer relevant in the revised manuscript.

- **Comment:** 2805-20: Why is mse insensitive to Tmelt? This is a crucial point to identify.
  Response: As is shown in fig. 2 (fig. 1 of the revised manuscript), the mse is quite sensitive to T_{melt}. What we state here is that the mse is insensitive to the *optimization* of T_{melt} for each individual glacier. The reason is that the optimized values for T_{melt} are close to zero anyway.

C1666

- **Comment:** 2807-11: spelling- interval
  Response: Text changed accordingly.

- **Comment:** 2807-16: good use of cross validation
  Response: Thank you!

- **Comment:** 2808-11: How many were rejected and what was the basis for rejection, just the number of values?
  Response: Of the 39 glaciers, 22 were rejected because their parameter estimate was not robust (see fig. 5 in the discussion paper), and two were rejected based on meta data (see footnote 3). We added the absolute numbers to the text.

- **Comment:** 2809-9: Why does Sarennes for example have such a high a-optimized?
  Response: The meaning of the value of a is hard to judge. As we state in the introduction (page 2806, lines 15-17), it may correspond to a precipitation lapse rate, aeolian transport of snow, avalanching, or any combination of these (and probably other factors as well). Our model is designed having the most basic ingredients of a mass balance in mind (i.e., accumulation and ablation), but otherwise it is trained in order to get the most out of the available information (measured mass balances, temperature, and precipitation). We then use the cross validation in order to determine whether this was successful - which means that we can rely on the model even if we cannot possibly understand the physical meaning of each parameter value.

- **Comment:** 2809-Section 2.3: I am not clear on why the SSC model is used here, given the adjustments described that limits its advantages from the original application.
  Response: As we state in the text (page 2810, line 9) we use the SSC model as a straw man for other modeling approaches that rely on increased numbers

C1667

of parameters in order to "enhance" model performance. Our point here is not to criticize the original Oerlemans & Reichert (2000) approach, but to illustrate the danger of not validating the model independently. Because of this, we explicitly state that the problem of lacking independent validation becomes smaller the greater the amount of independent data is - but it does not go away.

- **Comment:** 2811-1: The reduced effectiveness due to more model parameters is obvious, however, what if those parameters are more accurately known as they are in the original Oerlemans and Reichert, (2000).
  Response: It should in fact be obvious. However, there are numerous mass balance models published (some of which we cite) that increase the number of parameters without checking whether it actually improves the model. The reason that we are discussing the cross validation, and the SSC straw man model to such an extent is that it provides a powerful illustration of the fallacies rooted in lack of rigor in model validation. Note also that in footnote 4 we state that if the parameters are more accurately known, this is less of an issue.

- **Comment:** 2813-2: Split this sentence into two parts.
  Response: Text changed accordingly.

- **Comment:** 2815-13: The support of the statements in this paragraph are weak.
  Response: We added the correlation values to the text in order to strengthen our argument.

- **Comment:** 2820-10: Oerlemans in his Minimal Glacier Model (2008) noted that it is remarkable that the sensitivity of glaciers to temperature change can be estimated by just two parameters: the mean slope and the atmospheric temperature lapse rate. Contrast this with your statement at the beginning of this paragraph
  Response: We find it hard to see a connection to Oerlemans' statement here, since we neither consider mean slope nor lapse rate.

- **Comment:** Figure 5: Is this necessary given Fig. 6-8 and Table 1?
  Response: We think it is necessary, as it is the basis of the decision which glaciers to include and which to exclude from the final set (see also our response to your comment 2808-11 above.)