Reply to the reviewers comments

Dear Reviewers

Please find below the point-to-point response to your reviews. They were all highly appreciated as they were very constructive and helped to improve the paper considerably. I have indicated in my reply for which points I have revised the manuscript, otherwise I explained the reasons for having the manuscript unchanged. As all reviewers were unhappy with the partly confusing description of the experimental set-up, I have rewritten this section completely and added three tables with the details of the processing. This new text and the tables are attached to this reply to allow you an evaluation of the new description. I hope you are satisfied with the reply and look forward to the further discussion.

With thanks for your work and best regards, Frank

Dear Reviewers

In my previous author comment I made two statements about the work of other persons that were not justified. I apologize for these statements and have removed/replaced them with this new version.

Frank Paul

Referee #1 (R. Giesen)

Received and published: 23 July 2010

General comments

This paper addresses an important issue in the response of glaciers to a climate change and is thereby an contribution of great interest to the cryospheric community. As clearly outlined in the `Background' section, changes in the surface mass balance due to the dynamical behaviour of a glacier are not always taken into account in mass balance studies. Based on a set of experiments with a mass balance model, this study assesses whether this is justified. Even though the (quantitative) results may only apply to the European Alps, the author illustrates that dynamic changes of a glacier can have a significant effect on the mass balance and therefore cannot be disregarded in long-term mass balance studies. The paper is well-written, concise and has a clear structure. I recommend this paper for publication in The Cryosphere, provided that the points below are addressed.

At some points, the author should be more careful in his statements, especially when the dynamical behaviour of a glacier is concerned. Of these, two points need particular attention:
1. Steady-state conditions. The author correctly states that steady-state glaciers are a theoretical concept, due to the continuously fluctuating climate. While a glacier in an equilibrium-state would have a stationary front, this does not imply that glaciers showing only small length changes for a number of consecutive years can be called stationary. Only a forcing with a period considerably longer than the response time can lead to a near-steady state; this does not apply to the climatic fluctuations observed in the real world. When periods with a near-stationary front do occur, they most likely indicate either a slowdown of a negative (or positive) trend due to some years with more positive (negative) mass balance or a change from advance to retreat or vice versa. This concept can be compared to a harmonic oscillation: at the point where the amplitude is largest, the movement is small, but this is certainly not an equilibrium- state! The author assumes steady-state conditions for the two model periods because the front was more or less stationary. This assumption is very likely not correct and should be avoided throughout the paper.

> I fully agree that the description of the steady-state assumption was not clear enough and have revised it. In principal, I would argue that glacier extents at the end of the Little Ice Age (LIA) around 1850 do also refer to a dynamic equilibrium with the climatic conditions as the forcing was sufficiently long (400 years with some fluctuations), even for the largest glacier (Gr. Aletsch) with an approximated response time of about 80 years. This was likely not the case for the 1970s, where only the more fastly reacting mountain and small valley glaciers might have reached a new dynamic equilibrium and hence a steady-state extent.
>
> In regard to the past 150 years, I would prefer not to compare the observed glacier changes with a harmonic oscillation. In my opinion there is a long-term (centennial) trend of glacier retreat over this period that is mostly caused by increasing temperatures. Whenever the continuously adjusting glacier extent comes in balance with mean climatic conditions over a decadal period (e.g. 1890s, 1920s and 1960s in the Alps), the observed fluctuations depend more on short-term variations, e.g. precipitation amounts. This is also how I interpret the results of this study, the 1970s glacier extents do indeed represent steady-state extents for most glaciers (of course not for the largest ones).

The author uses the steady-state assumption to tune the precipitation towards zero mass balance for all glaciers. From the model description it is not clear to me whether this is done for all model runs or only when determining the mass balance sensitivity. Except perhaps for the sensitivity experiments, zero mass balance is not required for the comparison between the two periods and tuning of the precipitation is therefore not necessary.

> Yes, indeed. The precipitation tuning has only been done for the sensitivity experiments (as described on page 746, L15). For the purpose of this analysis it is not required that also the extent is in a dynamic equilibrium, but it is helpful to justify the balanced-budget assumption (revised now). As only differences in mass balance are calculated for each glacier, possible further changes of the extent (as required for the largest glaciers in the 1970s period) do have a minor impact on the sensitivity.

I would suggest to keep the tuning to a minimum, as this might affect the results, especially when correction factors are large. One representative set of gridded air temperature and precipitation could be chosen to use in all model runs; modelled mass balance can perhaps be validated with mass balance measurements that are likely available for some of the glaciers in the 1970s and 1980s?

The model results can't be compared to 'real' mass balances as the entire set-up is fully synthetic. However, the model itself has been validated in the cited previous studies (see also comments to reviewer 2).

2. Change of surface elevation and extent. Although the experiments performed allow for a separation of the effects of surface lowering and area change, which is very interesting indeed, the author should be careful in describing them as two independent processes. Changes in glacier length and area are a response to changes in surface height, i.e. glacier thickening or thinning and are therefore dynamically coupled. Further questions and remarks are specified below.

Yes, I fully agree that this is the case for the comparably healthy glaciers in the considered period. It does, however, not apply for the currently observed collapsing and disintegrating glacier tongues where the thinning leads to rock outcrops that strongly enhance glacier melt and lead to separation from tributaries. I have adjusted the related statements.

Specific comments

738_3: What is small?

In this context it means that it can be measured but it is not recognizable otherwise (e.g. a 0.1 degree temperature increase over a period of 10 years).

738_23-25: Although glacier changes are easily linked to climate change, their records are relatively short and sparse compared to other proxies. Stating that glaciers are `the best' indicators is in my opinion too bold, especially outside the glaciological community many researchers will not agree.

It is a widely accepted statement from the IPCC reports and many other high level documents related to the monitoring of climate change indicators (e.g. glaciers and icecaps are in the list of Essential Climate Variables). Apart from measured temperatures, glacier changes are the only 'three star indicator' for climate change and the observation that ice melts when temperatures increase is still striking, also for the public.

740_17-19: What is exactly meant by an `ongoing forcing'? In my opinion, this could be anything from a constant change with respect to the initial condition to highly variable conditions. It seems that the author here refers to a linearly increasing temperature?

A linearly increasing trend with some natural variability is maybe the best description, i.e. different from a step change (added).

740_21-22: Did this one degree temperature increase occur as a step change around 1985 or gradually in the period after 1985? This is not clear from the text.

It was very close to a step change (added).

742_15-18: The assumption that the 1980s DEM is also representative for the 1970s is logical, changes in surface elevation were probably not dramatic in this period. These are the best data available and the argument of steady-state conditions is not needed in this respect.

Yes I agree and have replaced it with balanced budgets.

743_7-8: Please indicate which percentage of the data was reset to the DEM25.

This varies from glacier to glacier but is between 30 and 60% of the area (added).

743_16-17: If I understand it correctly, all experiments use the same climate, except perhaps for the mass balance sensitivity experiments. Why not use present-day air temperature in the experiments instead of searching for mean annual values that give near-zero mass balance in 1850?

Basically for three reasons: The purely synthetical set-up has several advantages when analysing the results, the value is easy to find (two iterations), and the mean mass balances from the reference run is not two far away from realistic values (+/-1 m w.e.). With present-day temperatures, mean mass balance in 1850 would have been around -1 to -3 m w.e. and the observed effects of a change in extent and surface elevation would likely not have been that clear.

744_1-3: Is this value of 0.5 based on measurements? If so, please mention this and provide a reference. Does this factor also include attenuation in the cloud-free atmosphere or is this accounted for in SRAD?

The value of 0.5 is a roughly approximated mean value for the Alps and derived from NOAA/AVHHR. The likely best reference for this has been added (Meerkötter et al., 2004).

The attenuation in the cloud-free atmosphere is considered by SRAD for a standard mid-latitude atmospheric profile.

744_22: What is the annual range/amplitude of the temperature cycle?

The annual range is 7.5 degrees (added).

744_23-24: Is the lapse rate based on measurements? How large is the uncertainty and how does this affect the results?

No, it is a widely used mean value for the Alps reflecting summer conditions. The impacts of the uncertainty in the lapse rate on the mass balance have been modelled by Machguth et al. (2008). There is no effect of a small change in lapse rate in this study as only *differences* in modelled mass balance are analysed, i.e. the same lapse rate has to be applied in both model runs.

745_1: The calculation of the longwave radiation is not described, which parameterizations were used?

The longwave radiation (in and out) is parameterised as described by Klok and Oerlemans (2002) and Oerlemans (1991, 1992).

745-746: I read through the `Experiments' section a couple of times and still had difficulties understanding how the experiments were exactly configured. The author attempts to supply a structure by using letters and numbers, but I would suggest to provide the overview in a slightly different way. Instead of distinguishing between runs A and B and the experiments numbered with letters E, the individual model runs could be named more precisely (and given in a table) and then the difference between two model runs can be referred to. E.g.:
A = 1850 DEM and 1850 extent
B = 1850 DEM and 1973 extent
C = 1980s DEM and 1850 extent
D1 = 1980s DEM and 1973 extent
D2 = 1980s DEM and 1973 extent+split up of glaciers
Then experiment E1 would become A-C (or C-A?), and so forth.

Thank you for this suggestion. I have tested and discussed with my colleagues several combinations of the coding but all of them had disadvantages either. The idea behind the use of a different coding for the model runs (A, B) and the experiments (E1-E4) was to clearly distinguish between the results of the mass balance model and the different combinations of zonal averaging. In this regard I should have numbered E4a and E4b (the sensitivity experiments) with A_{sens} and B_{sens} (and the reference runs as A_{ref} and B_{ref}). As the experiments do always refer to two steps (i.e. spatial averaging and subtraction from a reference run) and Fig. 5b showed the difference between E3b and E3a, some confusion was introduced. To resolve this and have the experimental set ups more transparent (incl. the tuning issue) I have now created three tables describing the model runs, the zonal averaging and the calculated differences. I hope everything is (more) clear now.

746_10-26: The sensitivity experiments need additional clarification as well. Do I understand it right that in these two runs the corresponding DEM and extent are used? In the terminology described above, configurations A and D1?

Yes for both questions. In the revised version this is hopefully clear.

Then it seems that also different climates are used for the two runs, with a one degree warmer climate for D1. This was probably done to obtain zero mass balance for both geometries.

In principal yes, the 1 degree higher temperature for your D1 configuration gives a mean mass balance of +0.1 m (considering all entities > 0.1 km^2).

But now it becomes impossible to distinguish between changes in mass balance sensitivity because of the different geometry and changes due to a different initial climate state.

For the climate I think this is the only way to calculate sensitivities, starting with a zero balance for both extents, increasing the temperature by 1 degree and analyse the results. To be comparable with the result for the 1850 extent, mass balances for the sensitivity experiment with the 1985 DEM & 1973 extent have been averaged over the related 1850 extents.

By performing two additional runs (geometry A and one degree warmer climate, geometry D1 and reference climate) these effects could be separated. However, considering the uncertainty in the mass balance distribution for both periods, the paper would benefit from a focus on the other experiments (E1-E3), where the actual mass balance distribution is of less importance. To achieve this, the mass balance sensitivity calculations could be omitted altogether or reduced to one geometry and climate.

As I found the decrease in mass balance sensitivity for smaller glaciers rather interesting, I would prefer to keep these experiments as well. I hope with the new tables the experimental set-up is sufficiently clear.

746_15-18: Are large correction factors needed to obtain zero mass balance? Please quantify this by giving the range of correction factors.

Correction factors were between -74 and +65% of the initial precipitation values (added).

747_4-5: How are these standard deviations computed, from the glacier's total mass balance or from the gridpoint mass balances? The same question applies to the mean values.

At first, from all gridpoint mass balances a mean value (arithmetic) is computed for each glacier. Standard deviations do then refer to the variability displayed in the respective sample, e.g. the variability of mean values in Fig. 3a or the variability of differences in case of the other figures. The mean value shown in the legends do refer to the area-weighted mean, i.e. it is a count of cell values rather than of glacier entities. As I prefer to have the analysis performed with the arithmetic means, I have calculated these as well (they are slightly different) and revised the text accordingly.

748_10-13: It would be very interesting to separate these two opposing effects, which could be done by performing a model run with the 1980s DEM, but using the SRAD output from the 1850 DEM.

Yes, I fully agree. But based on the reviewers comments I have the impression that this would result in an even more complex model set-up. I would thus prefer to get this paper understandable first and perform further interesting investigations later on.

748_23-27: Considering the largely overestimated ablation on UAG mentioned before (lines 2-4), the results of the model experiments are probably not realistic for this glacier. Drawing conclusions about the effect of extent changes for UAG can therefore better be avoided.

Yes, but this basically refers to the mass balance values in individual years. With a focus on differences, the initial offset does only play a minir role. Acdtually, the most heavily debris-covered glaciers (UAG and OAG) are not the most extreme ones in the comnparisons.

749_16-20: How can you be sure that the artificial tuning towards a zero mass balance has not affected the mass balance sensitivity? The glaciers that are far from zero mass balance apparently are most sensitive.

This does only apply for the glaciers with a negative mass balance. Those with a positive 'far from zero mass balance' have actually the lowest sensitivity.

This could be realistic, but the precipitation correction is also largest for these glaciers and this may be reflected in the mass balance sensitivity.

No, the amount of precipitation correction (either positive or negative) is not well correlated with the mass balance sensitivity.

The uncertainty introduced by the precipitation correction could be quantified by calculating the sensitivities for the situation without tuning, which could for example be added to Fig. 7a. However, as mentioned before, I would prefer to have the mass balance sensitivity discussion reduced altogether.

I have thus not added much on this discussion in the revised paper. When starting from a mass balance that is different from zero, the sensitivity is not much different but the results become much less comparable (among individual glaciers and between the two dates). In my opinion the mass balance sensitivity can only be assessed when starting from a zero balance.

749_27: I assume that a 150 m upward shift of the equilibrium line corresponds to a one degree increase in temperature?

Yes, indeed.

750_24-751_2: This estimate is based on many assumptions and does not add any new information obtained from the results in this paper, I would suggest to omit it.

I agree that this is a different story and have removed it.

Would it be possible to say something about the effect of ongoing and future changes in glacier extent on the (measured) mass balance, based on the results from this paper?

No, the only thing that can already be seen is based on other evidence and you suggested above to omit this.

751_16-24: Does this imply that the results of this paper are mainly valid for rather healthy glaciers?

Yes, currently in particular the larger glaciers lose mass without changing their extent too much, i.e. the down-wasting is self-accelerating (i.e. mass balance - altitude feedback).

751_25-28: More negative than a constant extent and elevation, I assume? Do mass balances always become more negative? Is this not dependent on the change in geometry?

Yes, when 'todays' extent and elevation is used for glaciers that have been larger in the past, the geometric effect is compensated and mass balance values become comparable in climatic terms (of course, in this case you do not need to compare mass balances). With the larger extent in the past, the same climate as today will automatically give a more negative mass balance. So measured mass balance in itself is not a good proxy to compare climate conditions when the geometric properties have changed (i.e. mass balances provide thickness change rather than climate conditions).

753_7-9: It is still not clear to me whether precipitation has been adjusted for all runs or only for part of the experiments. Reading this sentence, it seems that precipitation was always adjusted, but in other sections the author refers to `untuned' model runs. This should be more clearly stated in Section 4, preferably in a table. Furthermore, the terms should be used consistently throughout the paper.

Yes, I agree. This is a misleading statement as the tuning does indeed only refer to the sensitivity experiments. I have revised the statement accordingly.

Figs. 4 and 5: By his choice of the colouring in the figures, the author aims to stress the spatial variability in the results. However, the main conclusions are drawn on the differences between the various experiments, but with the chosen colour scales, comparison between figures is difficult.

I agree, this looks like a mismatch. But please recognize that I only compare the mean values for the entire region across the experiments and not the values for individual glaciers. As described in 747-6, the only purpose of the figures is to illustrate the spatial variability. I have now written this more clearly in the beginning of section 5.

I would suggest not to derive the colour scaling from the mean and standard deviation per experiment, but to use the same colouring for the four experiments. The mean value and the standard deviation can be presented in a table. This would ease the comparison of the four maps, while the spatial variability would still be visible. Moreover, when the colour interval is the same for all maps, the spatial variability can be compared between maps.

By using the same colour table for all figures the differences among glaciers become invisible for some of the experiments. The values for individual glaciers should NOT be compared from map to map. This can be highly misleading as also the glacier extents are different (1850 for 3b, 4a and 6; 1973 for 4b and 5a/b). I have added this important issue in the beginning of section 5.

Technical corrections

All technical corrections have been considered.

Figs. 2-5: The font size in the legends is very small and hard to read, please use a larger font.

The size is now at 300 dpi for both the figure and the legend. In the final version it should be well readible (I will check the proofs).

Fig. 6: In the legend, the mass balance sensitivities are positive instead of negative.

This results from the subtraction of the respective model run (A3) from the (tuned) zero mass balance run (A2).

Fig. 7: The font size of the labels and axes titles is rather small, I would suggest using a larger font. In panel b, the axes labels are positive instead of negative. Furthermore, I would advise to use the same range for both axes (0.2-0.9).

For Fig. 7b it is already at the limit of the available space, the range at both axes is now the same.

Additional references:

Meerkötter, R., König, C., Bissolli, P., Gesell, G., and Mannstein, H.: A 14-year European Cloud Climatology from NOAA//AVHRR data in comparison to surface observations. Geophys. Res. Lett., 31, L15103, doi:10.1029/2004GL020098, 2004.

Machguth, H., Purves, R., Oerlemans, J., Hoelzle, M. and Paul, F. (2008): Exploring uncertainty in glacier mass balance modelling with Monte Carlo Simulation. The Cryosphere, 2, 191-204.

Anonymous Referee #2

Interactive comment on "The influence of changes in glacier extent and surface elevation on modeled mass balance" by F. Paul (Received and published: 5 August 2010)

This paper analyzes the impact of the surface geometry used in glacier modelling studies on the obtained surface mass balance. For a study site with several dozens of glaciers in the Swiss Alps mass balances over the geometries of 1850 and 1985 are calculated using a model, and sensitivities are analyzed. The author reaches the conclusion that a large percentage of the glacier reaction to climate change is 'hidden' in the geometric adjustment. The question which is addressed in this paper is important and worth investigating. Whereas I strongly agree with the statement that 'the climatic interpretation of mass balance data is rather complex', I have several, partly substantive, objections against the methodologies used. In particular, I do not agree with the main conclusions that are drawn from the experiments.

> I hope my response helps to clarify most of them. It seems that most of the objections of them were based on a misunderstanding.

I have two major, substantive points, several general comments on assumptions and methods applied in this paper, and a number of specific comments.

1. The main conclusion of the paper is that 50-70% of the glacier reaction to climate change are 'hidden' in the geometric adjustment. This number is presented as a general statement, which seems to be valid for all glacier mass balance studies dealing with long-term reconstructions. In my opinion this is highly problematic as it transmits the message to glaciologists measuring mass balance in the field: 'Whatever you measure, you will only capture a small part of the climatic signal'.

> The latter is in my opinion indeed the case. The mass balances as measured in the field (annual or seasonal) do provide mass change in the year of measurement, process understanding and local data for model calibration / validation rather than a climatic signal (also because climate is defined over a much loger time scale). The climate signal derived from a long-term series needs to consider both the dynamic response (adjustment of the geometry) and the measured mass balance.

Therefore, I ask the author to carefully put into context his conclusions: The number 50-70% is not valid in general! Based on the analysis presented in this paper only inferences on the period 1850-1985 in the Alps with the use of corresponding DEMs for mass balance modelling can be drawn.

> I fully agree to this statement and have stressed the reference period in the revised paper.

Mass balance studies, for example in the scope of long-term monitoring programs, are often based on aerial photographs that are available only over the last few decades. In the context of these studies, e.g. between the 1960s and present, the number of 50-70% that is 'hidden' in the geometric adjustment would be misleading. The effect necessarily converges to zero the smaller the changes (or the time period) between the DEMs are. Thus, I expect the 'hidden' percent- age to be much (!) smaller in the case of e.g. a reconstruction since the 1960s.

> Of course!

The author should be very clear about the fact that this study (and the presented numbers) is specific to the DEMs 1850-1973/85 and the study site. Mass balance reconstructions over other periods, using other DEMs will show significantly different percentages of glacier reaction that is 'hidden' in the geometric adjustment.

> I fully agree! But please keep in mind that I have never talked about a different reference period or that these numbers apply in general. Throughout the entire paper I only refered to the 1850 - 1973/85 period.

2. Figure 5b caused me to put a large question mark behind the main conclusions presented in this paper: Whereas the average change according to experiment 3a is +0.36 m w.e. (this number is discussed throughout the paper), the average change for experiment 3b is only +0.03 m w.e. Moreover, according to experiment 3b the large majority of the glaciers is even below the arithmetic average. They show a negative sign of the mass balance change when using the smaller (1973) instead of the larger (1850) glacier extent. The average slightly above zero is obviously due to a few glaciers with extremely positive changes (up to +2.9 m w.e.!). Thus, Figure 5b, which only differs from Fig 5a / experiment 3a by the consideration of glacier entities, presents results

that lead to completely reversed conclusions! If I believe the results of experiment 3b instead of 3a the conclusion would be that glacier wastage is even accelerated by the geometric adjustment. I am not saying that this interpretation is correct, but it indicates that this kind of model results can be highly ambiguous!

Please apologize for this confusion. It seems that the description in the text and the figure caption is misleading. The experiment 3b as depicted in Fig. 5b does only show the CHANGES of the mass balance as calculated in exp. 3a (Fig. 5a) due to a different consideration of entities. These are NOT differences in mass balance compared to the reference model run, but compared to the results of exp. 3a. In this regard you are fully correct that the separation of comparably small tributaries have only a minor influence on the modelled mass balance of the main glacier, but when both glaciers are similar in size (e.g. LGG) the changes are more substantial for both parts. I have revised the description strongly.

Some examples: GAG experiences a mass balance change of about +0.3 m w.e. in experiment 3a. In experiment 3b, however, this value drops to –0.1 m w.e. (which is hardly recognizable – even misleading – in the figures because the colour scales are not the same; it should be made consistent!). The only cause for this change in sign is that one minor tributary has split off. It seems that this tributary was virtually unconnected with the main glacier in 1850. Similar shifts occur for all larger glaciers in the study area. The important difference can therefore only be due to severe problems in the modelling procedure. To me, this is an indication that model applied here might not be suitable to answer the questions posed.

As you might see from the above, the reason for the unclear changes are due to a poor description of the experimental set-up by the author rather than 'due to severe problems in the modelling procedure'. Actually, this is also the reason why the colour schemes need to be different. First of all you do not see the changes when one colour table (e g.with the largest range) is used for all figures, secondly the physical meaning of the values is lost (now they are all in classes of 1/2 standard deviation), and finally you might feel obliged to compare the values between the figures directly, which is misleading in the case of Fig. 5a and 5b. For these reasons I would prefer to keep the colour schemes as they are.

As pointed out in the review of R. Giesen, it is not clear which results are based on a tuned, and which on an untuned run, which might explain some of the large differences between the glaciers in experiment 3b. The author only discusses the results of experiment 3a; why are the results of experiment 3b with contradictory results not addressed in the paper? If there is a reason for this, it should be provided. I personally consider experiment 3b to be better suited to solve the questions posed in this paper because it is closer to 'reality' and reveals potentially erroneous mass balance distribution.

I hope this misunderstanding is clarified now. There was no discussion of exp. 3b in the main text as this was only a small 'side-kick' of the modelling (i.e. the zones for averaging the values have been changed) that have little additional scientific implications. However, I have now added a comment also on this point in the discussion. Only the starting mass balance for the sensitivity studies was tuned with precipitation amounts.

General comments on assumptions and methods:

* I miss the link to reality in the modelling study. Normally, every model needs to be validated using some kind of field data before drawing any conclusions from the results. Validation is not present in this study! Model results cannot just be justified by comparing them to parameterizations developed for mountain range scale estimates. This missing validation is surprising, as for several glaciers within the study region a wide range of high-quality field data, partly covering almost the entire period of interest would be available. As long as the model can not (more or less) correctly reproduce reality (and you know that it does!) it cannot be applied to calculate sensitivity to climate and geometry change. As R. Giesen in her review, I therefore also ask the author to present a validation of the model results before interpreting them.

As explained in the paper, the modelling study uses a synthetical set-up for the meteorological forcing to obtain reproducible results and to have a focus on the native effects for a clear interpretation of the results. How should the result of such a synthetical set up be validated? When you introduce 'reality' (whatever this is) to such a model, the native effects will be disturbed and you will not be able to distinguish clearly between cause and effect. The link to reality in this experimental study (it is a model) is hence rather limited.

I fully agree that a model needs to be validated to trust in the obtained results. For this reason the study was performed with a model that has been forced with 'real' data and validated with field measurements in several previous studies (e.g. Machguth et al. 2006b, Paul et al., 2008 and 2009). After more than 20 years of distributed mass balance modelling and dozens of papers that have shown how this works, I would like to go here one step further and simply apply such a model to tackle some of the more interesting problems. Please also consider that I analyse mass balance differences. The aim is to find out how the mass balance changes for specific changes of the forcing or the glacier geometry, rather than to model real mass balance values for a specific year. Such a relative comparison of modelled values is largely independent from the used model.

I agree that there is one glacier in the study region (Grosser Aletsch) that has long-term measurements of accumulation and runoff as well as mass balance measurements for shorter time periods. There is another glacier nearby (Rhone) that also has long-term runoff data and mass balance data. However, the model used here is not applied to individual years, but uses a synthetical set-up of meteorologic data. In this regard a direct comparison with these measurements is not possible.

I think that also an unvalidated model can 'be applied to calculate sensitivity to climate and geometry change'. There is a seminal study by Oerlemans and Hoogendoorn (1989) that already illustrated more than 20 years ago how this works for a 'fictious glacier at the Sonnblick observatory'. In the study presented here this approach is transferred to three dimensions using two different DEMs and glacier extents from two points in time to provide the zones for spatial averaging of the modelled mass balance.

* The steady-state assumption is discussed in detail in the review by R. Giesen. I completely agree with the points that are well presented in this review, and I will not repeat them. This issue should absolutely be addressed by the author in a revised version of the manuscript.

I fully agree that the steady-state terminology used to describe balanced-budget mass balance is misleading in regard to the temporal time scales involved (decades vs. one year) and have corrected it.

* The tuning to a zero mass balance using precipitation as a tuning variable bears problems for several reasons (see also review of R. Giesen). In my opinion, the assumption that all glaciers have a zero mass balance in 1850 and 1973 (which seems to be quite central to the paper) cannot be justified. It is know from different studies (e.g. Paul and Haeberli, 2008) that glaciers show largely differing rates of mass loss for a similar climatic forcing. Why should they all have exactly the same mass balance in one certain year that is interpreted as a steady-state? This assumption seems to be in contradiction to observational evidence the author has published himself. The glaciers have strongly differing response times of up to one century (e.g. GAG). Whereas some small glaciers might be close, most larger glaciers are probably quite far from a steady-state, thus favouring different mass balances.

You are fully correct that a glacier with a steady-state extent does not need to have a balanced-budget (i.e. zero) mass balance and vice versa as both responses act on different time scales. And I also fully agree that the extent of the glaciers in the 1970s must not necessarily be a consequence of the same climatic conditions as for example the response time of Grosser Aletsch glacier (about 80 years) is much larger than for most of the other glaciers in the study region. However, for the extent at the end of the Little Ice Age (LIA) the climatic signal is of centennial nature and the reconstructed short term fluctuations during this period might be dominated by changes in precipitation rather than by temperature. So assuming a similar temperature forcing for all glaciers in the region as a cause for their extents at the end of the LIA is well justified in my opinion.

This is different for the extents observed for many glaciers during the 1970s. As cooler climatic conditions lasted for about one or two decades, most of the fastly reacting mountain glaciers adjusted to these conditions in that period with stationarity or a slight advance and it can be assumed that their new extent refers to the same climatic conditions. However, this is of course not the case for Grosser Aletsch or Unteraar glacier with their much longer response times. Their mass balance change due to the geometric change is thus likely too small.

Please consider that the model run starts with the climatic conditions that give a mean zero mass balance (area weighted averaging) for all glaciers in the study region and not for each individual one. These temperature conditions result in a wide spread of individual mass balances as shown in Fig. 3b. So there is no contradiction with your expectations or my previous studies in regard to this variability for an individual year. The zero-mass balance tuning is a theoretical concept to have the same starting conditions for all glaciers. This tuning is required as the mass balance sensitivity does depend on the hypsography of the glacier and thus on the starting conditions for the ELA (as explained in the paper). Small unconsidered changes of the geometry do only have a minor influence on the sensitivity (again: this is a difference between two modelled mass balances, when the geometry changes both values change but not their difference).

* The main point of the paper is that the interpretation of glacier mass balance series calculated over changing glacier geometry bears some problems (I agree with that!). The background given in the first chapter is long and quite comprehensive, and illustrated with many references. I do, however, not understand, why the author never mentions the concept of analyzing point mass balances instead of glacier-wide area-averaged quantities. The analysis of long-term mass balance series that refer to one measurement point on the glacier surface is a long established concept that is well recognized in the glaciological community. For the interpretation of climatic trends and variations this concept is almost unaffected by the problems with glacier-wide mass balance series, which is discussed here. I know that point mass balances are not the topic of the paper, but I think that this concept merits to be shortly discussed as well, because it provides relatively simple solutions to most of the biases highlighted in this study.

I have a slightly different opinion on the use of point mass balances for detecting or interpreting climate trends. Such data need to be detrended to be comparable and might then show the variability of the remaining signal (second principal component) but the dominant climate trend is lost. Not using such data here is related to their difficult interpretation and that they are not comparable from glacier to glacier without statistical manipulations. So the parameter that is comparable on a global scale is the area-averaged mass balance and there is some need to separate the effects of climate forcing on measured mass balance from other (partly self-accelerating) effects (downwasting, disintegration, etc.). As a first step, this study analyses how large the effect of the geometric change (extent and elevation) on modelled mass balance can be.

* The statistical significance of the results needs to be tested. I suggest that error bars should be given for all numbers. Based on statistical tests it should be checked whether the resulting effects are different from zero. Only considering mean values can be problematic as they might strongly be biased by outliers or unrealistic results from individual glaciers. Analyzing area-weighted quantities instead of arithmetic means, or calculating the effect for individual glaciers separately might also be a solution.

For the latter point I think the opposite is the case. When I consider the largest glaciers (e.g Aletsch and Unteraar) only as individual glaciers rather than with their entire area, these potential outliers do have much less weight than the other way round. Moreover, the number of glaciers (about 60) should be sufficient to reduce the influence of outliers. However, the mean values displayed in the figures with the colour-coded half standard deviations do actually refer to area-weighted values. So for the revised paper I have also calculated the arithmetic means which I found more useful for the purpose of this study. As they are slightly different than the area-weighted means I have revised also the text, including the conclusions. To better illustrate the significance of the mean values, I have added now standard deviations to all given difference values.

* The meaning of the term ‘hidden’ which is often used to illustrate the conclusions needs additional attention. What does it mean? How should it be interpreted? Let me consider the most negative mass balance year in the Alps, 2003, with measured balances of almost –3 m w.e. If 50-70% are ‘hidden’ in the geometric adjustment does this mean that the mass balance in 2003 would have been –6 to –8 m w.e. if the 1850 DEM would have been used to calculate surface mass balance with the same climate?

When the DEM AND the extent of 1850 is used, the mass balance with the 2003 climate would be between -4.5 and -5.1 m w.e. However, as the 50-70% refers to the period 1850 to 1973/85 and glaciers are about 20% smaller in 2003 (compared to 1985), the effect is likely even higher. So your interpretation of the term ‘hidden’ is basically correct. The message is that the interpretation of the mean mass balance from two epochs for a specific glacier in climatic terms can be misleading when the extent has strongly changed.

I doubt that the author wants to imply this result is realistic, but it is a possible conclusion based on the percent-number given in the paper.

Regarding the percentages, this is indeed the result of the modelling (but I do not know from which calculation the -6 to -8 m comes).

And what about positive mass balance years? My point here is that it is crucially important to tell the reader what this percentage of 'hidden glacier reaction' means. Maybe providing a percentage is the wrong type of measure for quantifying the effect.

I think a given percentage of an expected signal is the best measure to describe this. Please be sure that I here refer to a longer time period and not to individual years. Of course, the same principle applies to advancing glaciers: a part of the climatic response is 'hidden' in the grow of the surface area and can't be measured in the field. The total effect can only be assessed between two steady-state extents and not in-between. As I have written in the introduction, this study looks at the 'maximum possible effect'. When you look at shorter periods (e.g. 1850 to 1920 might work as well), the effect is of course smaller.

Specific comments are provided below:

* p739, l5: This statement represents the common understanding in the scientific community (many more, also much older citation could be provided here). However, in my understanding it is in direct contradiction with the main conclusions of this paper. If 50-70% of the glacier reaction to climate change (see Abstract, page 738, l13) is 'hidden' in the geometric adjustment, then mass balance can barely be interpreted as the 'direct and undelayed reaction' to climate change. This contradiction needs to be solved in some way.

You might consider to carefully check what I have written (exactly for the reason that you mention here): 'While glacier mass balance can be ... reaction to the annual atmospheric conditions'. This is of course different from climate change. So mass balance measurements are indicators of glacier volume change (at least when carefully calibrated with the geodetic method after some years) rather than of climate change but . The often found mixture with a climatic interpretation is maybe due to the fact that mass balance can be modelled by using time series of data from meteorological stations and to name these climate data.

* p740, l29: Can the authors provide a short definition of 'down-wasting'. How can this process be easily distinguished from 'active retreat'. Is there really an observational proof that all Alpine glaciers are now affected by 'down-wasting'? Is there evidence that they are increasingly affected throughout the last century?

I refer here to down-wasting as a minor change in glacier extent (or retreat) that is accompanied by a strong decrease in elevation of the tongue. This is basically observed for nearly all larger valley glaciers and some rather flat other types in the Alps (see images in Paul et al., 2007 or Paul and Haeberli, 2008). For typical mountain glaciers (with short response times) the thickness change is restricted to the terminus and accompanied by strong retreat, so that there is no down-wasting here. I have made no studies about the behaviour over the last century, but assume that the currently observed down-wasting is a consequence of the sudden increase in temperature by about one degree (step-change) in the mid-1980s. This was too fast and too strong to get the 'normal' dynamical reaction.

* p741, l9: Why is the interpretation 'less clear than in previous periods'? Also previous periods were characterized by mass balances below (or also above) the long-term average for more than just one year leading to length and area changes that are not completely incomparable to today. In order to be able to make this statement geometrical changes in 10-20 year periods throughout the entire century would need to be considered. To make this statements it is not sufficient to compare observations from the last (extreme) decades to averaged changes over the last century.

I agree with your statement but the difference here is in the temperature record and the studies that have used length changes to model changes in temperature (e.g. Oerlemans, 2005). When analysing long-term homogenized time series from Swiss climate stations (starting in 1864), you will see that temperature is mostly oscillating around a slowly increasing mean value. After 1980, the fluctuations are around a much higher mean value, i.e. in a different state of the climate. In consequence, we see fastly collapsing and disintegrating glacier tongues which is likely unique over the past 150 years. Reconstructing temperatures from length changes of Alpine glaciers is currently not advisable.

* p742, l10: It should be clearly stated which topographic maps were used here. To which year do these maps refer to? Exactly to the year 1850?

No, this strongly depends on the map sheet considered and is for a longer period of about two decades. However, the extents that have been digitized are not only based on these map sheets but also on field evidence (i.e. LIA moraines). Most of the glaciers in the Alps and in the study region reached the here used extent around 1850.

Furthermore, it is not clear whether 100 m contour lines that were directly digitized from the maps, or reconstructed isolines were used for establishing the DEMs. My main point is that some estimates of map accuracy should be provided. This would allow a quantification of uncertainty in the final results arising from DEM uncertainty (which is probably considerable for the 1850 map). The accuracy of the 1850 maps could, for example, be assessed by drawing a comparison to the up-to-date DEM in non-glacierized areas.

The contour lines used here are a mixture of reconstructed isolines and those given on the earliest topographic maps (e.g. regarding their curvature). You cannot get the accuracy of the isolines on glaciers by comparing elevations outside of glaciers with modern maps, mostly due to the different methods used in obtaining these lines, the different responsible surveyors, and other effects (e.g. the datum has also changed several times). Moreover, as described in the results section of the study, the change in surface elevation has only a minor impact on the change in modelled mass balance and DEM accuracy (in particular in the accumulation region with its comparably small overall changes) has thus a negligible overall effect.

* p743, l7: Where exactly was the 1850 DEM reset to the 1985 DEM? 'Towards the accumulation area ...' is not precise enough. The author should either state an elevation boundary or a percentage of the glacier surface for which this correction was applied. Is there any estimate what additional uncertainty this correction causes in the modelling results? It is probable that surface elevation in large parts of the accumulation area has significantly decreased throughout the last century. The impact on the overall mass balance can not just be neglected without verifying its importance.

Actually you have to neglect it as there are no means to better quantify them. I have not analysed the regions of replacement in detail, but they largely refer to the accumulation region of all glaciers (i.e. c. 30-70% of the total area, but strongly varying from glacier to glacier). I have added this information to the revised paper. But as described above, when a 200-300 m elevation change has a small overall effect on the mass balance, the effect of an unconsidered 20-30 m change is likely even much smaller. So this is really good news! Please also consider that there is in general little melt in the accumulation area of a glacier and thus little impact on the mass balance when neglecting this uncertainty.

* p743, l15: Some explanation might be of benefit here why the study does not rely on 'real' daily meteorological data, but on a rough approximation of the annual cycle based on a cosine function. I assume that the sensitivities could be quite a bit different when using a realistic weather variability throughout the year.

As explained above, the reason for using this synthetical set-up is to obtain reproducible results with a clear answer on cause and effect as well as data availability. When a 'real' forcing is used, information on mean daily temperature, precipitation and global radiation for c. 1850 in the region of Aletsch glacier are needed. Any ideas? Neither the mass balances itself nor the sensitivities will change much when the artificial set-up is replaced with a more realistic forcing that has the same annual means, as the mass balance basically results from a combination of the different factors over the full time period and for each grid cell of the used DEM. Even for rather different input data sets on a daily scale (e.g. a climate station and RCM output) the differences in the modelled mass balance are small (Paul and Kotlarski, 2010).

* p743, l18: The paper does not state anywhere which meteorological data were used. Do the data originate from measurements at a weather station?

No, as described in section 2 I used the precipitation grid by Schwarb et al. 2001, a constant cloud cover of 0.5 over the year and the potential global radiation as modelled by SRAD for each day of the synthetic year. The mean annual temperature and amplitude was found by an iterative adjustment to get a zero balance over the entire region for the 1850 extent and DEM.

* p744, l24: Is this temperature gradient representative for the study region, i.e. does this value originate from comparison of different weather stations?

This is a standard value from the literature that was used in many previous studies.

* p745, l1: Can the author provide a reference that confirms that turbulent heat fluxes are independent from wind speed? To my knowledge turbulent heat fluxes depend strongly on wind speed.

In this section the MODEL is described. So in the model turbulent fluxes do not depend on wind speed (of course they do in reality!). A rough approximation of reality is included by having the exchange coefficient dependent on the elevation distance from a mean ELA.

* p745, l9: Also the values of the other model parameters should be provided here.

They have been listed in the cited previous studies but I have added them now also in the revised paper.

* p745, l14: Do 'net balances' refer to area-averaged annual balances here? Provide a short definition.
Yes, indeed. They are calculated as described in the same sentence.

* p746, l5: Here and elsewhere: Why 'us'? The paper is written by only one author.
Yes, this is correct. I think the more polite form is to use 'we' (e.g. Kääb, 2005). I have now removed all we/us sentences.

* p747, l2-10: This paragraph sounds rather like a Figure caption and should be shortened.
This section has now been revised and better explained.

* p747, l18 - 748, l4: The obtained mass balances in the study region for the same climatic forcing vary by 3.6 m w.e.! This is higher than the difference between the most negative and the most positive mass balance year in the Alps since the beginning of the measurements. Thus, this extreme variability is barely realistic, and physical interpretation of the data, as performed in this paragraph seems to be impossible to me. The interpretation of very positive mass balance at the Northern Alpine rim should be revised: Are there such mass gains in reality, or in the model? If it was in reality, then these glaciers would now be advancing quickly; if it was in the model precipitation is much too high there, or calculated global radiation to low.
You are fully correct, there is a large spread in the modelled mass balances for the initial model run (now between -1.3 and +1.7 m w.e., see Fig. 7a). Please consider that most of the glaciers in this sample have never been measured and some of them have very specific charcteristics. Excluding one extreme glacier, the range of values is between -1.3 and +1.3 m w.e. When assuming that the model calculates ablation rather accurate (apart from UAG and OAG where unconsidered debris-cover is responsible for too negative mass balances), the differences result from wrong accumulation amounts. This means that both input data and unconsidered processes in the model (snow redistribution by wind and avalanches) are responsible for the extreme values. The wide range of values do also nicely illustrate why precipitation is used in most models as a tuning factor: its real amount is largely unknown (both: what comes down at a specific location and what remains there) and it has a large impact on the mass balance. The cited study by Machguth et al. (2009) gives additional comments on the related regional precipitation variability. Indeed, the used precipitation data set by Schwarb et al. (2001) needs some refinement at high elevation sites.

* p748, l15: This statement should be verified using a statistical test! The difference between positive and negative effects is quite large. Is the mean really significantly different from zero? If the difference is not statistically significant the author should revise the related statements.
The value is likely not significantly different from zero as one standard deviation (0.05) brings the arithmetic mean (-0.02) to a positive value. I have now discussed this in more detail as the value depends slightly on glacier size (i.e. it becomes significantly different from 0 for larger glaciers). Standard deviations are now added to all mean difference values.

* p748, l22: Here and elsewhere: It is not clearly stated how the mean 'for all glaciers' is calculated. I assume it is the arithmetical mean. It would be interesting to also provide the area-weighted mean, which is less affected by the large number of small glaciers with potentially much higher uncertainties in the modelling.
It was actually the area-weighted mean and I have recalculated all values as I prefer to use the arithmetical means (see also above). Both values are now discuseed in the revised paper.

* p748, l29: 'special topographic conditions': Provide an example.
It is due to the loss of strongly shaded parts as mentioned in the sentence before (revised now).

* p749, l20: It all also correlates well with the prescribed precipitation distribution pattern (Fig 2a). Glaciers with initially little precipitation (towards the border of the study area) need higher correction factors, and thus more precipitation, leading to a higher sensitivity. Is this interpretation consistent with the explanations provided in the paper?
In principal a higher sensitivity would be related to higher precipitation amounts. However, after tuning precipitation, all glaciers might get a rather similar total precipitation amount and the sensitivity should thus also be very similar. Actually, the amount of the required tuning is not related to the sensitivity.

* p750, l9: Are these two numbers really comparable? The estimate for the long- term mean mass balance refers to the hydrological mass balance, the calculated sensitivity is based on an unchanged geometry. In my opinion, this comparison could only be performed if the long-term mean mass balances would be based on constant glacier geometries. Moreover, the estimates referenced here cover the entire Alps, and not only to the study region. Also for this reason comparison is not possible.

The calculation of the sensitivity is based on a current geometry rather than an 'unchanged' geometry, i.e. in reality also the geometry adjusts in response to this change. Actually, the reason for this study was to find out what the contribution of the geometric adjustment is. The comparison with values from other parts of the Alps should only illustrate that this is really an alpine-wide signal and that the results obtained here are valid for the entire Alps.

* p751, l1: This conclusion should be formulated more clearly, or the estimate be completely removed. It is difficult to understand.

It is a simple calculation: 40%-25% = 15%, 15%/1% per year = 15 years (= 'about ten more years of retreat' in the text). But this part has been removed now based on the comments from reviewer 1.

* p752, l7: According Nemec et al. (2009) no hydrological balances were calculated. The result originates from 'a gradual linear interpolation' of the calculated mass balance for two different geometries. Hydrological mass balances obtained over dynamically changing ice extent and surface topography might be different from this value.

Yes, indeed, but likely not that much. Actually, the calculation of a mean hydrological balance from the volume loss derived by a geodetic survey (DEM differencing) is also done by dividing the volume loss by the mean area (S1+S2)/2 over the entire period.

* p753, l7: The author should be more careful with interpreting the model results. I assume the model is not accurate enough to allow inferring on effects of snow redistribution etc. Only if all other processes (melting, albedo variation, initial precipitation amounts) are absolutely 'correct' the importance of snow redistribution processes can be interpreted.

I principally agree with this comment. However, I have written 'and/or' and cited earlier studies that demonstrated this effect. It can also be deduced from the partly very high precipitatiion tuning factors for individual glaciers (up to +/-70%).

Additional references:

Kääb, A. 2005. Combination of SRTM3 and repeat ASTER data for deriving alpine glacier flow velocities in the Bhutan Himalaya. Remote Sensing of Environment. 94 (4), 463-474.

Huss, M., Hock, R., Bauder, A. and Funk, M., 2010. 100-year glacier mass changes in the Swiss Alps linked to the Atlantic Multidecadal Oscillation. Geophyiscal Research Letters, 37, L10501.

Paul, F. and Kotlarski, S. 2010. Forcing a distributed mass balance model with the regional climate model REMO, Part II: Downscaling strategy and results for two Swiss glaciers. Journal of Climate, 23 (6), 1607-1620.

Short comment (M. Pelto)

Interactive comment on "The influence of changes in glacier extent and surface elevation on modeled mass balance" by F. Paul
M. Pelto
mauri.pelto@nichols.edu
Received and published: 21 June 2010

Paul (2010) presents a critical analysis of the impact of changing glacier extent and surface level on the climatic signal form glacier mass balance records. For those of us undertaking long term mass balance programs the resultant climate signal from glaciers with much reduced size is obviously biased. To correct this bias is a matter of import that Paul (2010) quantifies and provides us with a methodology for assessment. This important paper in the introduction brilliantly lays out the problem. The suggestions below primarily focus on a clearer presentation of the various experiment setups and better illustrating the impacts of each experiment by focusing on one or two example glaciers. There are four experiments run, each with two scenarios. This results in eight outputs.

I found myself going back over the experimental description section too many times. I recommend using a table that lays out the model parameters in addition to the text, easier to reference.

> I have added now three tables that describe the different model runs, zonal averaging experiments and the calculated differences. I hope the experimental set-up is more clear now.

Create a sub-section for each experiment in the text as well. It would be of particular use to examine one or two glacier in detail as example of how each experiment affected each of those glaciers. The quantitative results for the glacier response to each experiment can be summarized in a table and the qualitative discussion for the glacier just placed in each experiment sub-section, possibly Upper Grindelwald and/or Baltscheider could be used.

> The reason for only describing the general picture rather than exemplifying this with individual glaciers is at least two fold.
> - the idea is not to model the effects for individual glaciers (this can be done better with a more sophisticated approach like in Nemec et al. (2009)), but to have a look at the mean signal over a larger region and its spatial variability
> - when illustrating the effects additionally for individual glaciers, the general statements will be much less clear and the model description likely even more confusing (e.g. there is likely no glacier that behave always like the mean)
> I would thus strongly prefer to have the presentation of the results focused on the mean values (see also reply to the comments by the other reviewers)

The figures already display the response of each glacier, but the discussion does not focus on any specific glacier consistently. By consistently focusing on a glacier the reader will better understand qualitatively and quantitatively the impact of changing glacier extent on the mass balance climate signal.

> No, likely not. It will be much more confusing as individual glaciers will not follow the mean trend in all experiments. This might be an exercise for a later study when I have the basic results of this study on 'safe grounds'. From the reviewers comments I have the impression that the experimental set-up is rather difficult to follow and I do not want to make it even more complicated at this stage.

We will in most cases be applying such a corrective model to a single glacier.

740-6: Should mention the non-steady state possibility (Paul et al., 2004).

> Due to my misleading use of the steady-state terminology I have revised this section now.

740-19: Can be either Pelto (2006) or Pelto (2010), the latter is in the references.

> I have changed it to Pelto (2010).

745-3: “..a mean ELA of 2900 m is used for the entire model domain extent.” At what m step in the model is this ELA applied?

There is only one mean ELA of 2900 m used for the entire model domain. The DEM has decimeter elevation steps.

746-18: Are the correction factors for various changes in precipitation?

No, they vary from glacier to glacier but the precipitation map (Fig. 2a) is constant.

747-27: Why the more positive values in the central basin?

Good question! It might be due to an overestimation of precipitation amounts in this region on the Schwarb et al. (2001) map.

750-25: Should be emphasized that despite the shrinking glacier area that should be increasing mass balance that the increasing negative balances indicate a non-steady state response to current climate. Maybe this only applies to some of the glaciers, but this would be important to identify as well.

As this has already been described on P740, L20 and I have not investigated the impacts of the past 20 years, I would prefer not to repeat this here (it requires a broader context).

752-15: The enhanced sensitivity are for glaciers that end up lacking an accumulation zone and will disappear (Pelto, 2010), versus the reduced sensitivity for those with a protected and avalanche fed accumulation zone that are difficult to eliminate (Hoffman and Fountain, 2005).

Yes, indeed, for some alpine type glaciers these effects can be observed as well. I have added the latter and a further reference.

New section 4 and Tables

4. Experiments

To obtain comparable results, a rather synthetical set-up is used for the experiments. Three modeling steps can be distinguished: (1) application of the mass balance model using different mean temperatures and the DEMs from 1850 and 1985 as an input (cf. Table 1), (2) zonal averaging with the 1850 and 1973 extents for calculation of a mean mass balance per glacier (cf. Table 2), and (3) calculation of differences to a reference data set (cf. Table 3). For step (1), model run A1 serves as a reference for the 1850 glacier geometry and B1 for the 1973/1985 conditions. For both runs also a tuning run is performed (A2, B2) that has adjusted precipitation amounts to obtain a zero mass balance for each glacier as a starting point for the sensitivity runs (A3, B3). They use one degree higher mean annual air temperatures but the same tuned precipitation grid as in A2 and B2 (Table 1). One additional run (B1T) was performed that is based on B1 but uses a one degree lower temperature.

Table 1 near here

For the tuning, the precipitation sensitivity (change in mean mass balance due to a 10% lower precipitation) was calculated for each glacier and modeled mass balances were divided by this sensitivity to derive the related correction factors. The precipitation grid (Fig. 2b) was corrected with these glacier specific factors (ranging from -74% to +65%) and new mass balances were calculated. After the second iteration modeled mass balances were close to zero (+/-0.02 m w.e.). The zero balance tuning is required to eliminate the influence from the initial balance on the calculated sensitivity. For example, a reduced sensitivity could be possible for a glacier that has already a negative balance with its ELA located in the steeper accumulation region. In this case, the size change of the ablation region is only small for a given ELA shift. The precipitation sensitivity was not re-calculated starting from a zero balance.

Based on the model runs from step (1), different combinations of zonal averaging (listed in Table 2) were performed in step (2). The reference runs AR (BR) are based on the initial runs A1 (B1) and use the respective extent from 1850 (1973) for zonal averaging of mass balance values. The sensitivity experiments AS (BS) are based on the sensitivity runs A3 (B3) and also use the respective extent from 1850 (1973) for zonal averaging. To determine the influence of glacier extent and surface elevation change on mass balance, the model runs A1 and B1T were combined with different extents. In Experiment 1 (E1) mean mass balance values are calculated from model run B1T and the 1850 extents, in E2 mean values are calculated from model run A1 and the 1973 extents, and in E3 mean values are calculated from run B1T but with the 1973 extents. Due to glacier split after 1850, the 1973 extents are considered in two different ways: in E3a all parts that belonged to the former 1850 extent are treated as one entity and in E3b all entities in 1973 are considered separately.

Table 2 near here

In the final step (3) the differences of modeled mean mass balances are calculated (Table 3). Difference 1 (D1) provides the change in mass balance due to a change of the DEM (i.e. surface elevation), D2 due to a change of the extent, and D3 due to changes in both the DEM and glacier extent. The difference D4 results from the different assignment of glacier zones in E3b and is calculated as the difference to D3 to enhance the changes. With these differences it is possible to separately assess the change in mass balance due to surface lowering, decreased glacier extent and a combination of both adjustments. The differences D5 and D6 refer to the sensitivity model runs A3 and B3.

Table 3 near here

Tables

Run	DEM	Temp.	Mean mb	Mb. range	Comment	Figure
A1	1850	15	0	-1.3 - +1.7	initial run 1850	3b (7a)
A2	1850	15	0	+/-0.1	tuned to 0 mb	-
A3	1850	16	0.64	0.35-0.87	temp. sens. 1850	6 (7a/b)
B1	1985	16	0	-2 - +1.5	initial run 1973	-
B1T	1985	15	+0.42	-1.3 - +1.8	for exp. E1 and E3a/b	-
B2	1985	16	0	+/-0.1	tuned to 0 mb	-
B3	1985	17	0.58	0.26-0.84	temp. sens. 1973	(7b)

Table 1: Overview of the performed model runs. DEM: used digital elevation model, Temp.: used temperature, Mean mb: mean (area weighted) mass balance for the entire region, Mb. range: range of individual values, Figures: the results are presented in the indicated figures with the number in brackets refering to the scatter plots, exp.: experimet, temp. sens.: temperature sensitivity.

Code	Temp.	Run	DEM	Extent	Comment
AR	15	A1	1850	1850	Reference experiment with the 1850 DEM
AS	15+1	A3	1850	1850	Temperature sensitivity for 1850
BR	16	B1	1985	1973	Reference experiment with the 1985 DEM
BS	16+1	B3	1985	1973	Temperature sensitivity for 1973
E1	15	B1T	1985	1850	Change of DEM
E2	15	A1	1850	1973	Change of extent
E3a	15	B1T	1985	1973	Change of DEM and extent (1850 entities)
E3b	15	B1T	1985	1973	Change of DEM and extent (1973 entities)

Table 2: Overview of the glacier extents used for spatial averaging of the calculated mass balance from the model runs listed in Table 1 (column 'Run').

Difference	Input 1	Input 2	Figure
D1	A1	E1	4a
D2	A1	E2	4b
D3	A1	E3a	5a
D4	D3	E3b	5b
D5	A2	A3	6
D6	B2	B3	-

Table 3: Overview on the calculated differences. For each difference input 2 is subtracted from input 1.