

First I want to thank for the detailed, careful and very helpful comments in this revision!

I'd like to reply on some of the general comments as well as to some specific remarks:

The referee states that a quantitative assessment is missing. This is a very good point, and I am looking forward to add numbers on:

- height and extent of seasonal snow cover
- estimate of crevasse volume
- analysis of the spatial distribution on density changes in the surface layer

The detailed measurements on the amount of basal melt will be finished at the end of the mass balance season, and will be published soon. In any case, it is very difficult to extrapolate that data for the total glacier area and to the past. I can give a number, but I will add a quite high error bar.

A number of studies were carried out to verify the accuracy of the DEMs, which are cited in the paper (Würländer, Abermann, 3 Diploma thesis of geodesists). The accuracy of the DEMs was verified with DGPS and tachymetric measurements. The results of these studies show that local errors related to photogrammetry of low contrast images caused by fresh snow or shadows are much higher than random errors. Flat areas are mapped more accurately than steep ones (Bollmann et al, accepted), the choice of the processing algorithm is extremely important. We made a huge effort to verify our geodetic network in the last times, and annual surveys were carried out. The results are published in the cited papers. Including all this material in this paper would possibly be too much.

The referee stated that the article is in between a teaching book, a statistical analysis and a detailed study on specific glaciers.

The intention of the article was the following: We had seen a difference between the results of the geodetic and the direct method published by Geist (2003). The difference is too large to explain it with measurement errors. So the first step is to investigate if this problem occurs in the last 10 years only or also before. To prepare the field, clear definitions and process descriptions are included (teaching book). The aim was to investigate specific glaciers in different periods (advance and retreat) to make clear if the difference was occurring only the last decade during strong retreat.

It was not the aim to investigate the accuracy of the direct data, a number of measurements were performed to be able to that, and will be summarized soon, but in a separate article.

A detailed study for Hintereisferner is in preparation (with homogenized mass balance data and very accurate LIDAR DEMs, lots of measured data on w.e. of seasonal snow cover). The article in discussion here should set the frame for the detailed study and answer the question if data of Hintereisferner is completely different from other glaciers or if this is a general problem.

I used a misleading formulation of the result. I did not mean to state that 'the geodetic and direct mass balance differ systematically', but that 'the geodetic and direct mass balance may differ systematically for specific glaciers'. According to my results, the differences can compensate for a larger sample of glaciers, but I can also imagine samples or climate change conditions, for example in case of strong retreat of all glaciers in the sample, in that these differences do not compensate.

The homogenization process was not carried out for other glaciers for specific reasons summarized below in the replies to the specific comments: the homogenization should improve the accuracy, quality or error estimate of a data set. This is not practicable/necessary/useful for all glaciers analyzed in this study. Here comes a list:
HEF: already homogenized

JAM: Same observer since 1988, no new maps - > these data is homogenous!

SSK: Same observer since start of the series, very careful analysis - > these data is homogenous!

VF: The Munich group has already reanalyzed the first years of measurements (Procedure unpublished).

HK: The total series was done by the same observer, who published the results. He is ~90 years old, and will possibly not do a reanalysis. He has the stake and pit data, but I am not sure if any other person will be able to reconstruct the mass balance pattern without him.

KWF: The first observers, Queck and Schimpp, passed away. Not all the original stake and pit data and mass balance maps are available for the years before ~1964. After that, the measurements were done by one person and the analysis was made by a second person. These persons did not change, so this second period is homogenous. I don't see a possibility to homogenize the first period – not before and not after this study!

Detailed comments

Page 1152, Line 5: giving (only) the mean ANNUAL mass balance (e.g., -0.5 m w.e. a-1) and the mean CUMULATIVE difference (e.g., -0.7 m w.e.) is somewhat misleading. For reasons of comparability you should also give the mean ANNUAL difference.

Yes, a very helpful suggestion.

P1152, L8: the given accuracy for LiDAR of 0.002 m w.e. seems rather a theoretical value than what is reported in glacier applications (see also your own comments on P1163). Indicate if these values refer to vertical or horizontal accuracy.

The accuracy of the LIDAR data is now analysed in detail in the now accepted paper of Bollmann et al., 'Potential of airborne laser scanning for geomorphologic feature and process detection and quantifications in high alpine mountains', Zeitschrift für Geomorphologie 54, Supplementary Issue "Laser Scanning Applications in Geomorphology" (this issue): xx-yy" The vertical accuracy of the ALS data is given as +/-0.07 m for the test site (including Hintereisferner). So my suggestion would be to use this number and cite the paper of Bollmann. This error can of course be larger in case of erroneous geocoding of the LIDAR data as result of missing tie points on the ground, but this might be stated later in the paper.

P1152, L19: I agree with your main conclusion that geodetic and direct mass balance data are "complementary", but not that they "differ systematically" (see general comments above).

A agree with that, my suggestion is to replace that formulation with 'can differ (systematically?) for specific glaciers under specific types of climate forcing'

P1153, L5-6: direct glaciological (not "or"); geodetic (or volumetric)

ok

P1153, L17-19: ...well, a complete WORLDWIDE dataset of direct glaciological methods would not show a bias either. And every real-world dataset might have some method-dependent bias...

I suggest 'A data set representative for the world's glaciers'... or skip that 'teaching book part'

P1153, L20-25: I think one should strictly differentiate between the use of geodetic methods for

(i) the validation and calibration of direct glaciological data series at specific glaciers (e.g., by using photogrammetry or LiDAR) and

(ii) assessing the representativeness

of the few observation series for their entire mountain range (e.g., by using

DEM differencing from SRTM and National DEMs for thousands of glaciers, cf. Paul and Haeberli 2009, GRL). Concepts, methods, accuracy requirements, and interpretations will differ strongly between the two tasks.

I think a third point iii) the calculation of mass balance with the geodetic method, will become increasingly important in the next decades. My suggestion is to replace that part of text with a bullet-point list with these 3 applications and add your last sentence
P1154, L1: the higher (temporal) resolution of the direct glaciological method does not necessarily lead to greater accuracy! It might actually be easier to quantify the accuracy of geodetic data.

We quantified the accuracy of geodetic data with field measurements for Hintereisferner and Kesselwandferner, I can add the results to the paper. The error characteristics of JAM, VF, SSK and HK should be similar.

I think it is very important to include ice free area in the DEMS, this is most suitable for an accuracy check.

P1154, L5-9: A complete uncertainty assessment of direct glaciological and geodetic mass balances needs to consider much more potential sources of errors than just density assumptions and extrapolation of point measurements. For a comprehensive list of potential stochastic and systematic error sources see Thibert et al. (2008, JG), Huss et al. (2009, AG), and Zemp et al. (2010, TC).

I could include a list of these errors additionally to the citations already given. I am not sure if a statistical analysis as done by Thibert (2008) results in a better error estimate than the comparison with point data measured in field.

P1154, L20-21: The calibration (i.e., adjustment) of the direct (cumulative annual) glaciol. mass balance to the geodetic (decadal) mass balance does reduce the systematic error but not the stochastic (i.e., random) one. Improve terminology and concept of the uncertainty assessment.

Yes, I will try to add a section with definitions.

P1155, L4: Without giving the information on total change (i.e., signal) and corresponding time period, you cannot judge a mass balance difference (i.e., noise) to be "enormous". Also it should be noted, that the comparison by Cogley (2009, AG) does ignore any systematic uncertainties in both methods (e.g., differences in survey dates) which might be OK for statistical analysis of large samples but certainly not for the interpretation of differences at individual glaciers.

This is an important information and will be added.

P1156, L20-26: The density (e.g., of 850 kg m⁻³) used for conversion of the geodetic volume changes to mass balance is usually based on the assumption of a constant density profile in the accumulation area (under steady-state conditions for glaciers with constant accumulation rate and no melting in that zone; cf. Sorge 1935 and Bader 1954). I do not really understand how your concept of the "surface layer" does fit into this concept? Please define your term "surface layer" and explain your assumptions (and potential deviations from Sorges law) including corresponding uncertainty estimates.

Ok, I can add information on the Sorges data and its relation to the density profiles in the firn layer measured on Kesselwandferner in the 1980s together with my concept of surface layer.

P1157, equation (2): Note that your way of calculating the specific geodetic mass balance (i.e. dividing the volume change by the larger area) is different than most approaches in the literature. Traditionally, the mass balance (B_{geo}) changes are divided by the area AVERAGE of both survey years. With the introduction of GIS-based raster analysis, b_{geo} is sometimes calculated as the average thickness change of all rastercells. All three methods might lead to somewhat different results which might need to be discussed.

I refer to Hoinkes, 1970; this method was used for all volume change calculation also in the Austrian glacier inventories. I can explain that and give a short estimate of the resulting difference.

P1157, L26: How are the glaciological mass balances in Austria adjusted from the

floating-date measurement to the fixed-date system?

This is explained in detail in Fischer and Markl, 2009, ZGG. I can add that. The accumulation and ablation occurring between the survey date and 30 Sep is corrected for the total glacier area for HEF KWF and JAM. VF and is surveyed on 1 Oct.

P1158, L20-22: Your point regarding the issue of the projection of mass balances is certainly interesting. Please detail further. This applies, however, to both the glaciological and the geodetic mass balance and should not introduce a methodological uncertainty.

Yes. I can refer to an J Glac 1962 with several published discussions on that topic. They summarize in 'yes its not correct, but practicable'

P1159, L11, and further down: Did the firn cover reduce more than the ice cover?

Yes.

For the conversion of the geodetic volume changes to mass balances it is not just the areal extension of the firn that matters. The density of the volume change is determined by the three-dimensional quantity of melted/newly formed snow, firn, and ice between the two surveys. You should try to quantify the related uncertainties.

Yes, so this actually is my surface layer concept. I can present better numbers than actually given in the paper, but previously I tried to keep the paper short ...

P1159, L12-14: In Fischer (2009, GPC) you nicely show for Hintereisferener that the homogenization of the mass balance series is an essential step towards the comparability of (annual) mass balances within a glaciological data series. In my view, this should be done for the other glaciers and also for the geodetic mass balances too before any uncertainty assessment and further comparison!

As stated in the reply to the general remarks, this is not practicable/necessary/useful for all glaciers analyzed in this study. Here comes a list:

HEF: already homogenized

JAM: Same observer since 1988, no new maps - > these data is homogenous!

SSK: Same observer since start of the series, very careful analysis - > these data is homogenous!

VF: The Munich group has already reanalyzed the first years of measurements (Procedure unpublished).

HK: The total series was done by the same observer, who published the results. He is ~90 years old, and will possibly not do a reanalysis. He has the stake and pit data, but I am not sure if any other person will be able to reconstruct the mass balance pattern without him.

KWF: The first observers, Queck and Schimpp, passed away. Not all the original stake and pit data and mass balance maps are available for the years before ~1964. After that, the measurements were done by one person and the analysis was made by a second person. These persons did not change, so this second period is homogenous. I don't see a possibility to homogenize the first period – not before and not after this study!

P1159, L14-19: You correctly mention the uncertainty of horizontal (and potential other) shifts of DEMs due to changing datum and projection. Such systematic horizontal shifts may have a major impact on the thickness and volume change analysis and need to be quantified (e.g., Koblet et al. 2010, TC) and corrected before a comparison with the direct glaciological mass balance. Kääb (2005, UZH) for instance presents an analytical approach to analyse and quantify such horizontal shifts in DEMs.

This was done by 3 Diploma theses of students of Geodesy with two summers of field campaigns with DGPS in Inner Ötztal supervised by Gerd Merkel, FH München. Since I recently was forced to remove such grey literature from my citation list, I did not cite these studies. But I will really love to do that, and I can give a list of accuracies of fix points (its in the order of cm to mm)! These measurements used a number of permanent GPS stations, and were accurate enough to find a so far undetected error at the permanent station of Patscherkofel maintained by the Bundesamt of Eich und Vermessungswesen.

P1160, Chapter on Results: Make sure that your values of mass balances and differences

between the methods are comparable with respect of units (see also comments above, P1152).

Yes, thank you very much for this helpful comment.

P1161, L25: Note that above you used a value of 850 (not 900) kg m⁻³ for converting geodetic volume changes to mass balances. Try to be consistent.

I should add 'for glacier ice': The mean density of the glacier even in areas where only bare ice (no firn) occurs is lower than 917 kg/m³ when one third of the glacier volume is air and not ice.

P1163, L15: Are you sure that seasonal snow cover alone can explain the misfit? You may use the survey dates of both methods, the aerial images (on which the DEMs are based), and meteo data in order to check and quantify this effect.

I can add measured snow data for HEF, KWF and JAM. The thickness of snow cover is very difficult to calculate from met data, especially when wind drift is possible. I can add a citation for that, and explain why I rather stick to 'observed szenarios'. The aim of this paper is not to evaluate the direct mass balance for these specific glacier, but to quantify errors which could occur by applying the geodetic method by

i) comparison to measured mass balance and
II case studies

P1162, L21-22: I can only agree with this first statement and encourage you to do so. I would suggest that you first introduce a list of potential uncertainties and a concept on how to quantify these stochastic and systematic errors for the Austrian glaciers.

I think one paper is quite short to do an overall accuracy assessment of the direct method, so I referred to previously done work for the direct method.

P1162, L22-23: Show why the estimates by Fountain and Vecchia (1999, GA) and Kuhn (1999, GA) should be valid for your glacier sample.

The aim of this study is not an error assessment of the direct method, and I don't think that this is possible to include, together with the other data suggested to include above (This would be a teaching book then). I can certainly do that, but in a different paper. The estimate of Kuhn should be relevant for Hintereisferner, since this is the glacier described in the study. The number of pits and stakes /km² on the other glaciers is comparable to that of Hintereisferner, so that the estimate of Kuhn should be valid. I could just refer on that citation, and drop Fountain and Vecchia.

P1162, L25-27: Why should the accuracy from Svartisen Ice Cap (by Rolstad 2009, JG) apply to the geodetic methods used in your glacier sample?

This is not the intention to say, it is just a summary of what other people found out. I could of course skip that reference.

P1163, Accuracy of DEMs: I like this chapter, but the estimates for stochastic uncertainties seem partly to be too optimistic and you are somewhat mixing horizontal and vertical accuracies (and precisions?!). Furthermore, systematic errors (between DEMs) should be analysed and discussed with respect to their influence on glacier thickness and volume changes.

I could cite a lot of literature on the specific DEMs. But the main point is that I am not sure how to separate the effects of digitizing the DEMs (in dependence from the algorithms, pixel size, position of pixels.. used) from the original DEM. I think this is a very important task, but has to be done by professionals, not glaciologists. I will ask our Partner at FH Munich if he would see a possibility to do that (and publish the results). I think it will be very difficult to get the funding for such a one or two years project. What I can do, is to add some results already published for photogrammetrically derived maps in the surrounding of the test glacier. Additionally I might find out some geodetic literature done in non-glacierized areas.

P1164, L5-8: see Kääb (2005, UZH) for a detailed discussion and analytical solution of this issue.

Yes, I can give the reference and a short description.

P1164, L10-15: I would suggest quantifying the influence of these uncertainties on the specific geodetic mass balances for each period of comparison.

Yes, very helpful suggestion.

P1164, Chapter on Seasonal Snow Cover: you should quantify the influence of these uncertainties on the specific geodetic mass balances for each period of comparison. See also comments above related to P1163, L15.

I will give measured data as far as available.

P1165, L4: above you mention a density of ice of 917 kg m⁻³... Try to be consistent.

Yes

P1165, L21-23: refreezing of melt water (i.e., superimposed ice and internal accumulation) is different from the issue of density changes and, hence, should rather be treated in separate sections/chapters.

Yes, very helpful suggestion.

P1165, L26: the potential influence of (changes in) crevassed areas on the density assumption, and finally on the specific mass balance, is interesting but needs to be quantified. You may assume a maximum depth of crevasses (e.g., 30 m) and use areal extent and crevasse frequency for a first-order quantification.

Yes, I will do that.

P1166: I agree that density changes of the "surface layer" is a potential source of error. You should, however, also discuss three-dimensional changes in that "surface layer". Is it plausible that your "worst case" estimates apply for all survey periods? Make sure that the values given are comparable to specific mass balances.

Yes, very helpful suggestion.

P1167, L11-19: What is the influence of these local (worst case) estimates on the specific mass balance?

I will add an area analysis.

P1167, Chapter 4.2 Comparison of the results to published data: What is the purpose of this comparison?

To answer the question if the Austrian glacier behave different from the rest of the world or not

And what does it show that we do not already know from the work by Cogley (2009, AG)?

I could not see the behaviour of specific glaciers in the work of Cogley, since I could not separate the glaciers in the 'point cloud'

Any basic corrections (e.g., common density assumption) that are required for such a "statistical" comparison of the raw data? I think you should better motivate the discussion.

ok

As such you nicely show that the difference between the two methods (a) are overall not systematic, (b) do overall not depend on the length of the survey period, and (c) can be systematic for specific glaciers.

ok.

P1168, L10-12: What is the (statistical) threshold for the decision that the data are "in accordance"?

Of course this numbers must be seen together with a mean mass balance of the specific glaciers. So the 'accordance' sentence should be replaced with the numbers.

P1169, L1-12: the overall stochastic uncertainty has to be calculated according to the law of error propagation whereas systematic uncertainties are to be cumulated.

Yes, very good suggestion.

P1169, L5: Please clarify why the (stochastic) uncertainty of the density assumption is supposed to be 10% of the geodetic mass balance?

The density of the surface layer can be between about 0.7 kg/m³ and 0.9 kg/m³ (I can give numbers and spatial analysis). The 10% are the minimum and maximum value resulting from that. I will provide an analysis.

P1170, Conclusions: in my view, only a comprehensive uncertainty assessment (including corrections for systematic uncertainties and error bars for stochastic uncertainty)

will allow to directly compare the glaciological with the geodetic mass balances and provide a statistical threshold of their (no) accordance. At present, the final remarks and numbers given seem to be rather arbitrary than thorough conclusions from the results and discussion.

A very detailed paper analysing Hintereisferner data is in preparation. The numbers given in here are by now way arbitrary, but based on mean and extreme values of observed data. I did not add these tables and data, because it lengthens the paper to about 1/3, but it is no problem to do that.

P1170, Acknowledgements: I think it would be appropriate to acknowledge the work of the large number of glaciologists that have contributed to this great datasets over many decades.

Yes, I am sorry, it was my intention to do that

P1172, 7-9: I believe the paper by Funk et al. was published in 1997 (not 1996), and the time period covered in the title is 1961-94 (not 1961-95).

ok

P1174, L22-25: note that meanwhile the revised paper is published in The Cryosphere, 4, 345-357, 2010.

ok.

P1176, Table 2: when two months are given for one year – does this refer to two DEMs or to just one produced out of data from two surveys? How did you correct for the differences in survey dates, also when compared to the date of the field survey?

The 1979 DEM was compiled from photogrammetric data and survey data at different dates (It will be best to give the citation). For the 1969 DEMs, we did not find out when which flights were carried out. The LIDAR data contains several pixels from the second survey, but it is impossible to find which pixels are affected without reprocessing the data which is not available for us, but governmental property. Since I will add point data anyway, I can give them for both dates.

P1177, Table 3: give information on density assumption (bgeo) in table caption. Also, any corrections applied to the mass balances (e.g., adjustments to fixed-time system?) should be clearly stated.

ok.

P1178, Table 4: the values for Storglaciären are in the wrong column: bdirect should be in bgeo and vice versa. Note that for bgeo we changed our density assumption from the TCD paper (825 kg m⁻³) to the TC paper (860 kg m⁻³). Please indicate in the table caption if you recalculated the bgeo from other publications with your density assumption of 850 kg m⁻³.

ok.

Tables: All Tables 1-4 are relevant and should remain in the paper. However, Table 2 might be converted into a figure showing the timing of the different survey types per glacier on a time axis.

ok.

Figures: Reduce the number of figures. I would suggest to keep Figs. 3, 4, 5, 15, 16; reduce the Figs. 6, 10, 11, (13) which basically show the same issue; and reduce the “teaching book” Figs. 1, 2, 8, 9, 12, 14 to only the most relevant issues.

ok.

P1185, Fig 7: it is hard to compare bdirect to bgeo in this figure – vertical bars of bgeo and corresponding cumulative bdirect might be more appropriate and would allow to include systematic and stochastic error bars.

ok.