I want to thank Emmanuel Thibert for his helpful discussion. I think his method demonstrated in his J Glac Paper on Sarenne glacier is an important and valuable tool for estimating accuracies. Nevertheless, the application of statistical tools is one of a number of approaches to a problem. I did not include a statistical analysis for several reasons:

- Error estimates for the DEMs of the 1969 and 1998 Austrian glacier inventory DEMs were performed by Würländer et al. I have actually no reason to doubt their results.
- For Hintereisferner and Kesselwandferner, annual tachymetric surveys allow an estimation of the accuracy of the DEMs. But unfortunately, the tachymetric survey, the acquisition of the DEM and the mass balance measurements did not take place on exactly the same day. Thus, interpolation is necessary. In any case, I would trust the comparison of these data more than a statistical analysis, when no estimate of a 'true' value is available.
- I am not a geodesist, and have only marginal experience in compiling DEMs with even today's and especially the last centuries methods. I would prefer to have done such an analysis by a professional, and will ask a project partner of us if he would be interested in such a task.
- Glaciological as well as climatological problems often are based on small data samples. To generate for example white noise, the number of samples should be 10 000 – a number I am far away from in my study. As long as this number of member of the statistical ensemble is not reached, the data will be biased, not really random, like throwing dice for 10 times: although the cube might be perfectly symmetric, and might be able to allow an equal distribution for the numbers 1 to 6 after definitely playing more than 100 times. Since we have here a multiparameter problem, the number of data samples must be even higher. One solution would be to do statistics on a point level. But here the problems comes up, that the elevation of one point is not independent from the surrounding, and that for example tie points or other points with higher priority in the specific algorithms used for processing the DEMs show higher priority than others. Processing a DEM is a very complex task, and a number of papers and teaching books are focussing on statistical properties of the DEMs. One of the easiest parts of literature is e.g. Geostatistics for Engineers and Earth Scientists. We do not have information on the algorithms used for the processing of the DEMs, and not all the original data to reprocess the DEMs. So I decided to use the DEMs, and cite the error analysis done by the geodesists which did the analysis.
- I am not fully confident that the errors of photogrammetry are fully random for this small sample of glaciers. The DEMs are not fully independent from each other, and in case of oversaturation or shadows, the person doing the analysis fills the gaps.

Applying the same kind of analysis would certainly lead to more optimistic conclusion in...
your analysis (hopefully), establishing that delta.b is due to the natural scattering of the data in relation to random error. This requires quantifying errors for each comparison. Therefore I would suggest using the delta.b/N parameter when comparing the 2 methods. delta.b/N is more intelligible as it can be used whatever the covered period, and can be compared to annual random error in b(direct), which is typically 0.2 m w.e./yr.

I did not intend to draw pessimistic conclusions but rather show some potential mechanisms we should maybe not neglect. But I agree that I could add the deltab/N!

Other minor comments follow:

P. 1152 Line 5: you should perhaps give also the average of the absolute value of delta.b. Yes!
P. 1153 Line 15: Also should be highlighted that geodetic balance integrate the overall natural spatial variability of the balance at the glacier surface, which is not accessible from the direct method due to an impossible exhaustive sampling. Yes! This is an important point.

Line 23: the combination was proposed because the geodetic method seems less affected by systematic errors, and these ones are very difficult to detect and quantify with a good confidence level (Thibert et al., 2008, J. Glaciol., 54(186)) I will include that, and I think there is certainly a point in doing that. The cost for it is to lose the control data – again a question of philosophy.

P. 1177 and 1178 Table 3-4: 2 additional columns could be included: - 1rst additional column indicating delta.b/N - second column indicating the annual random errors in b(direct) and b(geodetic), this last one being reduced to an annual amount from: sigma(geodetic)/(N)1/2. Yes I can include that.

Interactive comment on The Cryosphere Discuss., 4, 1151, 2010.

C881