Review of

Uncertainties and reanalysis of glacier mass balance measurements By Zemp et al, The Cryosphere Discussions

The paper gives an overview of glaciological and geodetic mass-balance measurements and provides a framework for re-analyzing long-term time-series based on the glaciological method. It is overall a valuable contribution.

Major concerns

- 1.) My main comment concerns the recommended procedure that in case of discrepancies between glaciological and geodetic results, the glaciological results per default are the ones to be corrected to match the geodetic results (e.g. line 9, page 791; page 812, line 19). I think this general statement/recommendation is not justified. It includes the implicit assumption that the geodetic balance is the 'truth' to which the glaciological balance needs to be adjusted. This may be the case in many of the examples in the literature but in general it is conceivable that a geodetic balance is more off than a glaciological balance. Whether or not the glaciological balance should be corrected to the geodetic balance (or vice versa) has to depend on the uncertainties of each of the balances.
- 2.) Second, the thresholds given on page 822, lines 25ff for the decision whether or not a reanalysis is necessary are problematic; and I don't think they should be included here. An absolute number cannot be given because the threshold bias must depend on the uncertainties in both balances. The absolute difference has to be compared to the uncertainties in both balances and also to the absolute balances. For example, a reanalysis of a glacier is not necessarily justified if the uncertainties in the geodetic balance are larger than the differences between methods no matter whether or not they exceed the given thresholds.

General comments

- 3.) The paper is somewhat difficult to read mostly because it is too long; the authors should try to shorten it in order to make it more readable to a broader audience. The paper combines a review-type overview with a framework/recommendations. Since the former is incomplete and not fully developed anyway, and the latter is the more innovative and main point of the paper, I would suggest that the review-type information is cut to a minimum and repetitions avoided. To shorten, I also suggest that
 - a) trivial equations (like for example, how an arithmetic mean is computed or a deviation from the mean) are deleted
 - b) the small summaries after each subchapter are deleted (they are not very informative and may rather be integrated into the conclusions).
 - c) page 813: add the half page of references to the Table and remove here.
- 4.) Due to the extensive review-type parts of the paper, the essential part, i.e. the framework and recommendations get a bit lost. Chapter 3 could benefit from some re-writing. Instead of starting with review-type information which seems somewhat off-topic I would suggest that the authors give a short overview of their framework how to reanalyze rather than more text-book like descriptions about stakes and pits etc. This would also help to understand the following subchapters. Chapter 3.3. starts with 'The aim of this

- third step is ...' but in fact I missed what the first 2 steps are. Why are 'Observations' part of the steps in the re-analysis? They are the input, but what are the actual steps you do? Maybe a flow diagram of the steps of re-analysis would help. Figure 1 is unfortunately not very helpful in that respect. Basically the essence of the 'new' framework gets lost in excessive review-type information and extreme detail (e.g. page 17). What I miss is an overview (text and figure) of the main steps of the re-analysis.
- 5.) The paper deals exclusively with usually smaller, land-terminating glaciers that have a long-term glaciological program. Hence it is rather limited in scope, considering that a huge fraction of the world's mountain glaciers are marine- or lake-terminating glaciers and many other methods have emerged to quantify mountain glacier mass changes on various scales. This is no problem for the paper but the authors should clearly state this limitation and make clear that this is about land-terminating glaciers, generally small in size, with a glaciological mass balance program.
- 6.) Terminology: The authors claim to follow Cogley et al., 2011, however, they do not in a few of instances:
 - a) The term 'specific' is used incorrectly. It is not the glacier-wide balance but the balance in units per m2 (e.g. page 794 and 795, line 16).
 - b) conventional/reference balances (see below)
- 7.) Those familiar with calibrating and validating models may be somewhat confused by the use of the terms calibration and validation in this context (although their usage here is technically correct). It may be useful to define what you mean; and / or use re-analyze instead of calibrate. Currently the paper uses both terms for the same process, and it may be better to stick consistently to 're-analyze' as also indicated in the title.
- 8.) Acronyms should be avoided as much as possible. PoR is not necessary and only makes the paper (even harder than it is anyway) to read. In many cases it can just be omitted or replaced by balance period or just period.

Detailed comments

- 1.) page 793, line 11: Delete 'processes' → 'sum of all components of accumulation and ablation ..."
 - line 13 is only true for a land-terminating glacier. This needs to be included.
- 2.) Page 794: non-inclusion of floating glacier tongues or ice shelfs are mentioned but not lake- or marine terminating glaciers (which often have a grounded tongue).
- 3.) Page 794, line 10. Replace 'inaccessible' by 'unmeasured'. Often, regions are accessible but still not measured and one needs to extrapolate.
- 4.) Page 795, line 1: add reference(s)
- 5.) Page 795, line 3: conventional balance: add a definition or at least a reference since not everybody is familiar with this concept
- 6.) Page 795, line 14: What about internal accumulation?
- 7.) Page 796: the discussion about conventional and reference surface balance is incomplete and not correct: the reference-surface balance does not only use a constant area but also a constant hypsometry, i.e. measured point balances first need to be extrapolated to the previous elevation before integrating over the previous ice extent.
 - For the sake of shortening this paper, the entire discussion about these 2 types of balances

- may be deleted (or shortened), since it does not really relate to the re-analysis framework discussed here.
- 8.) Page 797: page 3: add 'land-terminating glaciers'
- 9.) Page 797 last line: something is wrong in the English
- 10.) Page 798: 'the volume change is converted into a specific mass change by:
- 11.) Page 798 line 22: why is the assumption of linear change necessary. Don't think that is true; it rather implies an assumption about the geometry of the bed topography.
- 12.) Page 799: There is excessive detail and emphasis on sighting and plotting, (and too much detail also on page 800), however, I lack a more detailed mentioning and discussion about density assumptions in converting volume changes into mass changes, since this is an aspect with hugely diverging assumptions in the literature.
- 13.) Page 801, line 23: replace 'surface' by 'climatic-basal' mass balance
- 14.) Page 802-803: lots of repetition of problems that were mentioned earlier. Can be shortened a lot and better be replaced by a brief overview of the main steps of the framework.
- 15.) Page 806: what do you mean by validation?
- 16.) Page 807, line 14: delete 'the PoR covering' (can be done in most cases where PoR pops up).
- 17.) Page 811, line 9: replace 'calibrated' by 'reanalyzed' (also elsewhere).
- 18.) Page 811, line 15: This seems arbitrary. What is the physical rationale that any bias would exclusively come from errors in the summer balance rather than the winter balance. Ablation tends to be far more homogenous (and simply elevation dependent) due to the large air temperature dependency, and therefore can relatively well be captured by a few measurements. However, snow accumulation varies greatly within tens of meters and rarely are sufficient probings available to capture the spatial distribution. Hence, one would expect the opposite.
- 19.) Page 812, line 11, 16. Here both are used: re-analyzed and calibration. This is confusing.
- 20.) Chapter 5.4 is weak. It mixes text-book like statements about basic theory (like VII) with recommendations; hence the (more important) recommendations are diluted. Make this chapter more concise and focus on the recommendations and their rationale.
- 21.) Conclusions are too general and should be more concise. Maybe reiterate your main points about the framework.
- 22.) Same symbol is used for density and reduced discrepancy