Comments on "Uncertainties and re-analysis of glacier mass balance measurements", by M. Zemp et al., *The Cryosphere Discussions*, 7, 789-839 (2013)

Graham Cogley, April 2013

General Comments

This paper presents the results of studies presented at and arising from another in the roughly decadal series of "Tarfala workshops" on the measurement and assessment of glacier mass balance. It offers a thorough and in some ways innovative framework for estimating and presenting errors, with the somewhat discouraging result that typical measurement programmes can yield uncertainties even greater than estimated in earlier studies. Nevertheless the framework will be valuable as a starting point for efforts to improve the measurements and to understand them better. A crucial part of the framework is the emphasis on "reanalysis", with a meaning not quite the same as that in meteorology; it is the process of re-examining existing measurements and if necessary (as determined by a careful comparison of glaciological time series with multi-annual geodetic measurements) adjusting them so that the two kinds of measurement give comparable answers.

Although I have some concerns about points of detail, I think this paper is to be welcomed as an important step forward. I am especially pleased to see it because it addresses one of the six tasks set by the IACS Working Group on Mass-balance Terminology and Methods in 2008. Only one of those tasks was completed, and while the resulting *Glossary* was a very worthwhile product the Working Group left a good deal of unfinished business behind it. This paper fills what is probably the most serious of the gaps, and does so very satisfactorily on the whole.

It is rather long, and below I have suggested one possible way to shorten it. The people most likely to use it will, unavoidably, find it rather difficult to follow, but there is no question that the practice of mass-balance measurement needs to be moved onto a more rigorous footing.

Substantive Comments

P789

Title — It is not necessary to hyphenate "reanalysis". I suggest following the meteorologists (and the *Glossary of Glacier Mass Balance*) and treating it as one word. To avoid possible ambiguity, however, its meaning in glaciology should be defined in the Introduction.

P791

L4 — "Until the present". But "Until recently" would be more accurate (see for example the cited papers by Thibert et al. 2008 and Husset al. 2009).

L18 — "measurements by the glaciological method". A long-standing problem with the WGMS archive is the failure of observers to contribute their geodetic measurements.

L25 — "decadal": I feel discomfort at the consistent use in the paper of "decadal", which has an exact meaning, in the sense of "multi-annual" or even "long-term". Either of these alternatives would be preferable.

P794

L1 — "specific (i.e. glacier-wide) mass balance": "specific" does not mean "glacier-wide". It means "per unit area". All that is needed here is "To obtain the glacier-wide mass balance", but the rest of the text needs checking for this.

L4-6 — This very awkward sentence ends up hinting that one ought not to measure the mass balance of floating ice. Say something like "Floating glacier tongues and ice shelves are not considered here, because their mass balance is often dominated by frontal and basal terms that are not addressed by the glaciological method."

P795

L2 — "conventional balances": do not introduce this technical term without defining it. More simply, just delete "conventional" and leave the definition where it is now on P796.

P796

L1-8	This discussion of conventional and reference-surface balances needs more clarity. Firstly, most measurements are actually mixtures of the two kinds, unless the area adopted in any one year is extrapolated somehow from the shrinkage rate determined from the two most recent geodetic surveys (or, less commonly, measured every year). Secondly and probably more importantly, the reference surface has not only a particular area but also a particular distribution of elevations; with modern GNSS technology, most point measurements nowadays are probably "conventional" already, so that the actual measurements are more like mixtures than ever. I suggest clarifying explicitly that the analysis in the paper is focussed on conventional balances. A reference-surface balance can only be obtained by correcting both to the reference area and to the reference surface elevation. This introduces further elements of uncertainty, and the uncertainty in the conventional balance is difficult enough to manage anyway. (This is not to dispute the importance of the later analysis aimed at ensuring that comparisons of glaciological and geodetic balances are comparisons of like with like.)
P797	
L5-11	One could add to this list glaciers on which most accumulation is by avalanching (e.g. in the Himalaya) and most ablation is by sublimation (e.g. the highest glaciers in Bolivia or Tibet). But see my comment below on section 2.2 as a whole.
L12-15	I would delete this intrusive sentence. The point it makes is entirely valid, but the following sentence does not apply to it and it is not of fundamental importance.
P794-797	
Section 2.2	This comment applies in fact to the whole of section 2. This paper will be a long and rather tough read even for those with an immediate technical interest in the subject, so I wonder whether it could be made more accessible by reducing the amount of technical material which is simply taken from earlier sources (principally the *Glossary*, although that source was mainly a codification of still earlier work). Some of the material is essential for the setting out of notation and basic ideas, but all the same I would urge the authors to try to reduce section 2 by about a third to a half.
P800	
L12	Standard error: this should be defined carefully here, as the standard deviation divided by the square root of the number of independent items of information in the sample. Problems arise later from assuming that this number is equal to the number of items in the sample.
P802	
L16-17	This, and some of the material from the previous paragraph, ought to be in section 2.2.
P803	
L2	"homogenization" should be defined here (or perhaps earlier at P792 L19).
P805	
General	Review eqs. 11 to 15 to determine whether all are needed, given that the later analysis is all in terms of cumulative balances.
L16	This is not correct. The number of degrees of freedom in the geodetic measurement is exactly one for present purposes, and the division should be by the number of years (not the root). The quantity $1/N$ is simply a unit conversion factor. The error in a 16-year geodetic balance, when expressed as a rate per year, is very different when divided by 4 instead of 16.
P807	
L2-3	The common habit of comparing confidence regions to see if they "overlap" is not strictly correct. If the uncertainties of the two balances are normally distributed then the probability of their being "different" (drawn from different populations) depends on the difference between them and on their variances. A Student's-*t* test is needed, and ideally also an *F* test for equality of the variances.
L23	Explain briefly where 1.96, and 1.64 below, come from.
P810	

L3 What is an "interim" calibration? This adjective should probably be deleted.

L5 "a significant difference". Make things easier for the reader by using "bias" only as a "within-method" descriptor and using "difference" only for between-method differences.

P816

L2-3 The non-additivity of the reanalysis procedure explains the very odd result that it might advise you to correct some parts of a long glaciological record but to refrain from correcting the entire record. It is essential that the paper offer some clear guidance about what to do when this happens. As it stands, the paper implies that there are two decades of the Storglaciären record for which there are two "equal best" approximations to the truth.

P818

L5 The Cogley and Adams estimate was –200, not –300 (plus an unknown additional amount to allow for neglect of internal accumulation).

L9-10 Why do these numbers not add up, even when 325 and 140 are added "in quadrature"?

Stylistic Comments

P791

L6 "without consideration of errors".

L14 "… quality, including reliable uncertainty estimates".

L16 "subject" rather than "element".

L21-22 "most … consist of".

P792

L1 "for quantitative assessment …". The remark about reanalysis is unnecessary.

L8 "for standardization of the process".

L18 Change "as well as" to "and", and check the rest of the text for over-frequent use of "as well as". Comma after "practices". Insert "of glaciological series" after "validation".

L20 Change "providing" to "suggesting" or "proposing".

L22 "reanalysis", not "reanalysing".

L23 Delete the unnecessary "the corresponding".

L26 Change "state" to "summarize".

P793

L4 "of Cogley".

L5 "We differentiate between …".

L9 "values".

L15-17 "ds" might be better than "dS".

L18 "over one year b_a". Delete the colon.

L19 "from t_0 to t_1".

P794

L8 Omit "direct", which is distracting. Change "include" to "are", and check the text for later uses of "include" followed by complete lists.

L14 "since the earliest measurements."

L16 "The number and density". Delete most or all other uses of "as such"; usually it doesn't mean anything.

L18 "variation" rather than "inconsistency".

L24-25 "at monthly (on some …". "at a few points".

P795

L1 Do not hyphenate "polythermal".

L7-9 "main sources of error". Add the third source after "glacier": ", and changes of glacier area."

L9-10 "errors in (i) height determination (e.g. due to imprecise measurement; …".

L12-13 "density measurement errors and …"

L14 "superimposed ice, which is difficult to measure and of which the spatial variability is often not …".

L16	Delete “due to”. “to the glacier-wide balance”
L17	“sampling”.
L22	“observations”.
L26	“The few studies which have attempted … include Thibert”.
P796	
L2	Delete “typically”.
L6	“bands” would be more appropriate than “bins”, here and frequently later in the text.
L16	Insert “The time *t* is zero in the year of the first survey.”
L18	Delete “balance”.
L19	This use of subscript a for “all” leads to overloading; it already stands for “annual”. Perhaps use g instead?
L20	“strongly”.
L22	“large” rather than “strong”. “elevation”.
P797	
L1	“such as by using”.
L3	“glaciers in” (not “on”).
L4	“which mainly change”.
L12	Delete “As such,”. Also at L16.
L16	“reduce” or “eliminate” rather than “integrate”.
L24	Delete “subsequent”.
L28	Delete “most”.
P798	
L3	Change “that samples” to “over”, which will clarify that the only possible subject of the verb “does not consider” is “description”.
L14	“seasonal and interannual”.
L17-18	“the geodetic volume change is converted into the specific mass balance”.
P799	
L1	“by making an assumption about the density of the volume”.
L2	“If it is true that the change of bed elevation is negligible”.
L7	“interference of the atmosphere”.
L10-12	This sentence is superfluous.
L23	“solution for”.
P800	
L6	“regions not covered”.
L8	“both within and between”.
L25-26	“in photogrammetry, glacier surfaces with low visual contrast … than high-contrast bedrock surfaces”.
L28	“whether” rather than “if”.
P801	
L13-15	“for ablation … surveys”: the text is garbled here. The second “for” clause should say something like “and to ensure that the same areas are being compared.”.
L17-19	Expand the two “incl.”s. Comma after “energy)”.
L25	“emergence, so that the appropriate density is that of ice”.
P802	
L1	“the appropriate density depends”.
L3	Superscript –3 (not –2).
L4-5	“expansion”: how can firn expand? Perhaps this is an attempt to refer to depth hoar?
L6	Delete “respectively”.
L7-8	“is to use”.
L8-10	Delete “which includes Sorge’s law (cf. Bader, 1954) as upper bound” (not relevant in this context), and end the sentence at “Huss, 2013).”

L11 “suspected”. “potential magnitude of bias”.
P803
L1-8 It would be helpful to number the six steps or to give their subsection numbers.
L16 “sensors and platforms,”.
L22-23 “is to use available data and metadata to”.
L24 “so that the entire glaciological series is comparable with the geodetic measurement”.
P804
L2 Change “digital terrain models” to “DEMs”.
L4 “glacier, or other independent elevation data, to”.
L13-14 Delete “remaining”. Change “related to” to “in”. “as well as in”.
L17 Delete “consecutive”. Change “parameters” to “variables”.
P805
L3 Change “Whereas” to “However”.
L5 Enclose the three summands on the right in brackets. Add a period at the end.
L6-7 “as a mean annual rate”.
P806
L5 Change “attributed” to “attached” or “assigned”. “attributed” implies causation.
L5-7 “estimates, if they are known, from both the glaciological and geodetic balances results in”. I do not understand “makes sense in case of a later calibration”.
L9-10 Subscript “total” should not be italic in eqs. 16 and 17.
L15 “the purpose of this step”.
L24 Delete “(decadal) value of the”.
P807
L1 “dispersion” or “scatter” rather than “scattering”.
L7 Delete either “comparison” or “test”.
L8 There is nothing wrong with the verb “cumulated” (although “summed” would be more usual), but as an adjective, as here, “cumulative” is preferable.
L13 “of the two methods”.
L19 “the dispersion of the data”.
L22 “The more consistent the two methods, the closer … .
P808
L2 “of making”.
L4 “probability of”.
L5 “that mistaken rejection of H is twice as likely,”. Delete “under this condition”.
L10 “a type II error”.
L14 “such that”. “For type-I risks α of 5% and 10%”.
L16 “designated for calibration” (or “flagged”). “of maintaining”.
P809
L19 “Equation (25) indicates”: not really. How the number of years enters the equation should be explained briefly.
L15 “Once a systematic difference between the two methods is detected with high confidence,”.
L17 Delete “in”.
L22 Delete “so far”.
P810
L2-3 “discrepancies require”.
L12-13 “invoking”, not “inferring”.
L16 “of the origin of the difference, the divergence”. “remaining”, and the parenthesis, are confusing and unhelpful.
P811

L13-15	I find this quite surprising, and cannot see why it is OK to blame the summer balance "by default". One thinks of G. Østrem's difficulties with the probing of incredibly deep snow on Ålfotbreen (1999 Tarfala volume). A short justification of this assumption is needed.
P812	
L12	"databases in which they are stored".
P813	
L7	"and an additional nine".
L25	"measurements on Storglaciären have".
P814	
L13	Comma after "(2004)".
L18	"differences from".
L23	"quantification of whether". (The condition here is that expressed in German by *ob* rather than *wenn*).
P815	
L13-14	"level", not "interval".
L19	Delete "which".
L26	"years".
P816	
L11	Delete the comma. "difference", not "bias".
L14	"example of".
L22	Change "source (*s*)" to "source".
L24	"centred".
L25	"shifted to agree with the mean".
P817	
L8-9	Delete "As in the case of SIL,".
L11	Do not hyphenate "elevation distribution".
L16	"the conceptual framework described above".
L27	"with a mean of –454".
L28	"locations is estimated to be".
P818	
L8	"has estimated random".
L9	"which are attributed to".
L15	"have an average". More generally, "attribute with" is not English; you can only say "attribute to", and even then it is not the right verb for the present purpose. There are several instances to be corrected in the following text.
L20-21	"to be 3 mm".
L21	"correlations".
L22	Delete "so far".
L23	"is estimated to be 74".
L27	"include", not "account".
P819	
L8	">10% contribution from basal melt".
L10	"Some of the proposed corrections" (e.g. the Reijmer-Hock correction is not "very large").
L16	"500 mm" (not "m"!).
L18	"homogenized, with the aim of reducing". Period at end of sentence.
L27	"226 mm w.e. a^{-1}".
P820	
L3-4	Change the three "points" to "PoRs".
L22	"indication of".
L24	"weak" or "poor" rather than "low" as a qualifier of "ability". See also L29.
L25	"than".

L27	"are two or more times the value".
P821	
L9	"on average".
L16	"when the stress on α risk set loose" does not make sense.
P822	
L4	"Over the course of the PoR".
L8	"to surpass the random-error sum by enough". "with a useful confidence level" would improve on "significant".
L10	"scatter". End the sentence at "observations"; the rest is repetitious.
L25-27	Do not omit "w.e." after "mm" (three instances).
P823	
L20	"Reanalysis of mass-balance series, especially of long series, based".
L23-24	"identification of a need to calibrate a … implies large biases".
P824	
L3	"briefly": see substantive comment on P794-797 Section 2.2.
L8	"a dozen European".
L10	Delete "Thereby,".
L20	I would change "homogenization" to "record".
Table 1 L2-3	This sentence is unclear to me.
Table2 L7-9	"(i.e., a type-I risk)". "the type-II risk".
Figure 1	L2: Italicize *n*, and preferably make it upper case for consistency with usage in the text. L5: Delete "/with".
Figure 2	L4: "elevation distribution" (insert space). Explain why, in panel a, the mean difference of the annual calibrated and (interpolated) geodetic balances is not zero.
Figure 3	L2: "black diagonal line".
Figure 4	L2: "(horizontal axis)". "arbitrarily". L3: "vertical axis".
Figure 5	L6: "to become lower".

Supplementary Information

Appendix A

Para1 L2-3	"a shaded-relief image of".
Para1 L4	Vivid though it is, the slave-master metaphor is superfluous to this discussion, and somewhat distracting. (Q: Who decides which is the master? A: It doesn't matter. [Or does it? See below.])
Para1 L12	Change "softwares" to "programs".
Para1 L12	The "parameter" *c* is not an unknown. It can be calculated from information in the two DEMs. Indeed, eq. A1 might be more informative if *c* were on the left-hand side. There is, however, an underlying question that the text does not answer: from which DEM do the terrain slope and aspect come? As I understand the procedure, it is iterative because you have to choose one DEM or the other as the source of this information.
Para2 L1	"The solution of the co-registration problem can fail on flat …". It is not an equation, unless an eq. A3 is provided to describe the translation of Para1 algebraically.
Para2 L2	Surely "greater than" should be "less than"?

Appendix B

General	This Appendix is fine, but it should note that the semivariogram from kriging is not the only way to estimate spatial autocorrelation in a gridded dataset.