The Cryosphere Discuss., 7, 789–839, 2013
www.the-cryosphere-discuss.net/7/789/2013/
doi:10.5194/tcd-7-789-2013

The Cryosphere Discussions
Open Access

Uncertainties and re-analysis of glacier mass balance measurements

M. Zemp[1], E. Thibert[2], M. Huss[3], D. Stumm[4], C. Rolstad Denby[5], C. Nuth[6], S. U. Nussbaumer[1], G. Moholdt[7], A. Mercer[8], C. Mayer[9], P. C. Joerg[1], P. Jansson[8], B. Hynek[10], A. Fischer[11], H. Escher-Vetter[9], H. Elvehøy[12], and L. M. Andreassen[12]

[1]Department of Geography, University of Zurich (UZH), Zurich, Switzerland
[2]IRSTEA, UR ETGR Erosion Torrentielle Neige et Avalanches, Saint-Martin-d'Hères, France
[3]Department of Geosciences, University of Fribourg (UFR), Fribourg, Switzerland
[4]International Centre for Integrated Mountain Development (ICIMOD), Kathmandu, Nepal
[5]Department of Mathematical Sciences and Technology, The Norwegian University of Life Sciences (UMB), Ås, Norway
[6]Department of Geosciences, University of Oslo (UiO), Oslo, Norway
[7]Scripps Institution of Oceanography, University of California, San Diego, USA
[8]Department of Physical Geography and Quaternary Geology, Stockholm University, Stockholm, Sweden
[9]Commission for Geodesy and Glaciology, Bavarian Academy of Sciences and Humanities, Munich, Germany
[10]Zentralanstalt für Meteorologie und Geodynamik (ZAMG), Vienna, Austria

[11]Institute for Interdisciplinary Mountain Research, Austrian Academy of Sciences, Innsbruck, Austria
[12]Norwegian Water Resources and Energy Directorate (NVE), Oslo, Norway

Received: 1 February 2013 – Accepted: 19 February 2013 – Published: 4 March 2013

Correspondence to: M. Zemp (michael.zemp@geo.uzh.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Glacier-wide mass balance has been measured for more than sixty years and is widely used as an indicator of climate change and to assess the glacier contribution to runoff and sea level rise. Until present, comprehensive uncertainty assessments have rarely been carried out and mass balance data have often been applied using rough error estimation or without error considerations. In this study, we propose a framework for re-analyzing glacier mass balance series including conceptual and statistical toolsets for assessment of random and systematic errors as well as for validation and calibration (if necessary) of the glaciological with the geodetic balance results. We demonstrate the usefulness and limitations of the proposed scheme drawing on an analysis that comprises over 50 recording periods for a dozen glaciers and we make recommendations to investigators and users of glacier mass balance data. Reanalysis of glacier mass balance series needs to become a standard procedure for every monitoring programme to improve data quality and provide thorough uncertainty estimates.

1 Introduction

Changes in glacier mass are a key element of glacier monitoring, providing important information for assessing climatic changes, water resources, and sea level rise. The most extensive dataset of glacier-wide in-situ mass balance measurements covers the past six decades (WGMS, 2012; and earlier volumes) and is widely used to assess global glacier changes (e.g. Cogley, 2009) and related consequences of regional runoff (e.g. Weber et al., 2010) and global sea level rise (e.g. Kaser et al., 2006). However, the majority of these data series consists of just a few observation years and most results are reported without uncertainties (Zemp et al., 2009).

There are a dozen mass balance programmes with continuous time series dating back to 1960 or earlier (Zemp et al., 2009). Combined with decadal geodetic surveys, these long-term glaciological mass balance series provide a unique opportunity for a re-analysis of the mass balance series and a quantitative assessment of the related uncertainties. Earlier works found both agreement (e.g. Funk et al., 1997) and disagreement (e.g. Østrem and Haakensen, 1999) between the mass balance results from the two methods. Recent studies have carried out extensive homogenization and uncertainty assessments for reanalysis of mass balance series (e.g. Thibert et al., 2008; Rolstad et al., 2009; Huss et al., 2009; Koblet et al., 2010; Fischer, 2010, 2011; Zemp et al., 2010; Nuth and Kääb, 2011; Andreassen et al., 2012). However, there are not yet guidelines available for a standard process and a direct comparison of the findings from the above studies is challenging.

In the summer of 2012, a workshop organized by the World Glacier Monitoring Service (http://www.wgms.ch) in collaboration with Stockholm University was held on "Measurement and Uncertainty Assessment of Glacier Mass Balance" at the Tarfala Research Station in northern Sweden (Nussbaumer et al., 2012). The workshop built upon results and experience of earlier workshops at Tarfala in Sweden (Geografiska Annaler, 1999) and Skeikampen in Norway (IGS, 2009) and brought together a group of experts currently working on these issues. Its major goals were to discuss methods and related uncertainties of glaciological and geodetic mass balance measurements as well as to find a consensus on best practices mainly for validation but also for the homogenization and calibration of (long-term) observation series.

The present paper is a joint outcome of that workshop and aims at providing best practices for re-analysis of mass balance series. First, we provide a brief review of observation methods, related uncertainties, and reanalysing procedures for observation series. Second, we present the corresponding results from a selected number of glaciers with long-term mass balance programmes and discuss these in light of the proposed re-analysis scheme. Finally we conclude with recommendations for data producers and state implications for data users.

2 Theoretical background

2.1 Terminology and components of glacier mass balance

A common language and terminology is a basic requirement for developing any best practice. In this work, the terminology (in English), formulations, and units of measurement follow the "Glossary of Glacier Mass Balance and Related Terms" by Cogley et al. (2011). In terms of the uncertainty assessment, we mainly differentiate between random (i.e. unpredictable fluctuation in the readings of a measurement) and systematic (i.e. bias pushing results always in the same direction) errors (i.e. disagreements between measured and true value).

The mass balance of a glacier is defined as the sum of all accumulation (acc) and ablation (abl) processes (see Fig. 2 in Cogley et al., 2011) and a distinction can be made between surface (sfc), internal (int), and basal (bas) balances. Based on the conservation of mass within a column of square cross section extending in the vertical direction through the glacier, the mass-balance rate of the column is

$$\dot{m} = \dot{\mathrm{acc}}_{\mathrm{sfc}} + \dot{\mathrm{abl}}_{\mathrm{sfc}} + \dot{\mathrm{acc}}_{\mathrm{int}} + \dot{\mathrm{abl}}_{\mathrm{int}} + \dot{\mathrm{acc}}_{\mathrm{bas}} + \dot{\mathrm{abl}}_{\mathrm{bas}} + \frac{q_{\mathrm{in}} + q_{\mathrm{out}}}{\mathrm{d}S}, \tag{1}$$

with q referring to the flow of ice into or out of the column with fixed horizontal dimension, $\mathrm{d}S = \mathrm{d}x\mathrm{d}y$.

The point mass balance cumulated over one year a (or more generally: over the span of time from t to t_1) is linked to the mass balance rate b by

$$b_a = m(t_a) - m(t_0) = \int_{t_0}^{t_a} \dot{m}(t)\mathrm{d}t. \tag{2}$$

TCD
7, 789–839, 2013

To obtain the specific (i.e. glacier-wide) mass balance, the point balances are integrated over the glacier mean area S over the same time span:

$$B_a = \frac{1}{S} \cdot \int_S b_a \mathrm{d}s. \tag{3}$$

Note that the mass balance components of a floating glacier tongue or ice shelf are not included here as they are not considered to be suitable for glaciological mass balance measurements (cf. Kaser et al., 2003).

2.2 Glaciological observation method

The principal steps of the direct glaciological observation method include the measurement of ablation and/or accumulation at individual points as well as the interpolation between the measurement points and extrapolation to inaccessible regions of the glacier. The method was described in detail by Østrem and Brugman (1991) as well as summarized by Kaser et al. (2003) with particular attention to low latitude glaciers. The basic principles of the glaciological method are widely accepted and have not changed much since the onset of the earliest glacier mass balance programmes. However, the detailed implementation does vary between different glaciers and observers. As such, the absolute number and density of stake and snow pit observations varies from glacier to glacier and through time (e.g. Fountain and Vecchia, 1999). Another typical inconsistency is the deviation from the traditional contour line method as proposed by Østrem and Brugman (1991), for the spatial integration of point observations. Often, statistical interpolation schemes are used instead (e.g. Lliboutry, 1974; Jansson, 1999) or observed mass balance gradients are applied to the glacier hypsometry (e.g. Funk et al., 1997). The direct measurements are typically carried out seasonally or annually and cover the components of the surface mass balance. At some glaciers the measurements are performed in monthly (at some inner-tropical glaciers) or even at daily resolution (at few points during summer seasons). Observers at some cold

TCD
7, 789–839, 2013

or poly-thermal glaciers account for internal accumulation too. The results are usually reported as conventional balances for the mass balance year referring to the floating-date, the fixed-date, or the stratigraphic time system, and as specific mass balance in the unit meter water equivalent per year ($\mathrm{m\,w.e.\,a^{-1}}$). Equilibrium line altitudes, accumulation area ratios, and mass balance gradients are usually calculated from mass balance distribution with elevation (ranges).

There are three main types of errors in the glaciological method: the field measurements at point locations and the spatial averaging of these results over the entire glacier. The field measurements are subject to errors in (i) point observations which is essentially a height determination error (e.g. due to measurement precision; tilt, sinking and floating of ablation stakes; tilt of snow probings and difficulties in identifying last year's surface in the snow pack, e.g. due to ice lenses), (ii) density measurements and associated assumptions (with errors expected to be larger for snow and firn than for ice), (iii) superimposed ice which can be measured but its spatial variability is often not well captured by the stake network (e.g. Schytt, 1949; Wright et al., 2007) and due to (iv) flux divergence which is irrelevant to the specific balance (cf. Cuffey and Paterson, 2010) unless the samplings between divergence and convergence zones is unbalanced (Vallon, 1968). Error sources related to the spatial averaging of the point measurements include (v) the local representativeness of the point measurements (i.e. the ability of the observational network to capture the spatial variability of the surface balance), (vi) the method (e.g. contour, profile, kriging) used for interpolation between the point observation and for extrapolation to unmeasured regions (e.g. Hock and Jensen, 1999; Escher-Vetter et al., 2009), (vii) the under-sampling of inaccessible or difficult glacier areas with potentially different surface balances such as those due to crevasses, debris covers, steep slopes, avalanche zones (e.g. Østrem and Haakensen, 1999). Few studies attempted to quantify all these errors including Thibert et al. (2008), Huss et al., (2009), Fischer (2010), Zemp et al., (2010), Hynek et al., (2013), and references therein.

TCD
7, 789–839, 2013

In addition, and common to all mass balance series is (viii) the issue of the glacier area changing over time: typically, the glacier area of the most recent geodetic survey is typically used as a constant for the calculation of the specific glaciological balances for the years up until the next geodetic survey. Especially for large relative changes, this requires a recalculation of these annual "reference-surface" balances with updated glacier areas (and elevation bins) for every year in order to provide "conventional" balances (cf. Elsberg et al., 2001; Huss et al., 2012). This can be done by adjusting the surface area and recalculating the specific balance for each elevation bin of the glacier.

Assuming a linear area change over a period of record (PoR) covering N years the annual "conventional" balance B_{glac} of an elevation bin e is calculated for each year t as

$$B_{\text{glac}.e.t} = B_{\text{ref}.e.t} \cdot \frac{S_{e.0}}{S_{e.t}}, \tag{4}$$

where

$$S_{e.t} = S_{e.0} + \frac{t}{N} \cdot (S_{e.N} - S_{e.0}), \tag{5}$$

with elevation bin areas $S_{e.0}$ and $S_{e.N}$ from the first and the second geodetic survey, respectively.

For the entire glacier, the "conventional" balance is now regularly computed as the area-weighted balance sum of all (E) elevation bins:

$$B_{\text{glac}.a.t} = \frac{\sum_1^E B_{\text{glac}.e.t} S_{e.t}}{S_t}, \tag{6}$$

For glaciers with strong non-linear area changes, the (normalized) front variation series might be used to weight the interannual area changes. Complex balance gradients or strong changes in surface elevations might need to be addressed by re-integrating the point observations, such as using a distributed mass balance model (e.g. Huss et al., 2012).

Observation principles were mainly developed on and for mid-latitude glaciers on the Northern Hemisphere which are mainly changing by winter accumulation and summer ablation. The practicability of these principles and the relative importance of the error sources listed above might hence be limited to the seasonal analysis of high-altitude and high-latitude glaciers where any season can be the accumulation season (Chinn, 1985), low-latitude glaciers where ablation occurs throughout the year and multiple accumulation seasons exist (Kaser and Osmaston, 2002), and monsoonal glaciers of the Himalaya where accumulation and ablation occur at about the same time (Ageta and Fujita, 1996). As a consequence, annual balances might not refer to exactly the same observation period. As such, published results from the Northern Hemisphere, typically covering the observation period from 1 October to 30 September, might in fact have an overlap of six months with corresponding observation periods from the Southern Hemisphere (i.e. 1 April to 31 March). However, cumulative annual balances integrate these seasonal complexities and as such might not be relevant to the following comparison with decadal geodetic balances

In summary, the glaciological method measures the surface mass balance components, sometimes accounting for internal accumulation, and is subject to the following error classes: field measurements at point locations, spatial integration over the entire glacier, and reference area changes over time.

2.3 Geodetic observation method

The geodetic observation method determines volume change by repeated mapping and subsequent differencing of glacier surface elevations. Common methods include ground surveys, e.g. using theodolites (e.g. Lang and Patzelt, 1971) or global navigation satellite systems (e.g. Hagen et al., 2005), photogrammetry (e.g. Finsterwalder and Rentsch, 1981), and laser-altimetry (e.g. Arendt et al., 2002; Joerg et al., 2012). The most ideal geodetic survey results in a digital elevation model (DEM) that covers the entire glacier, minimizing potential errors from extrapolation (see for example, Arendt et al., 2002; Berthier et al., 2010). The methodological description below focuses on DEM differencing that samples the entire glacier surface and does not consider extrapolation errors. It also assumes that all elevation data are referenced to the same datum and projection.

Volume changes derived by differencing DEMs can be expressed by the following equation:

$$\Delta V = r^2 \sum_{k=1}^{K} \Delta h_k, \tag{7}$$

where K is the number of pixels covering the glacier at the maximum extent, Δh_k is the elevation difference of the two grids at pixel k, and r is the pixel size. Geodetic surveys are ideally carried out at the end of the ablation season, simultaneously with the glaciological survey, and preferably repeated about every decade. A time separation of about one decade accentuates the detection of a climatic signal and reduces the impact of short-term elevation fluctuations due to seasonal and annual meteorological processes. The results of the geodetic method thus refer to the time span between two surveys and are reported as volume change in the unit cubic meter ice equivalent (Eq. 7). For a comparison with the glaciological balance, the volume change is converted into the specific geodetic balance over the PoR in the unit meter water equivalent (m w.e.):

$$B_{\mathrm{geod.PoR}} = \frac{\Delta V}{\overline{S}} \cdot \frac{\overline{\rho}}{\rho_{\mathrm{water}}}, \tag{8}$$

where $\overline{\rho}$ is the average density of ΔV and $\overline{S}$ is the average glacier area of the two surveys at time $t0$ and $t1$ assuming a linear change through time:

$$\overline{S} = \frac{S_{t0} + S_{t1}}{2}. \tag{9}$$

TCD
7, 789–839, 2013

Commonly, the geodetic balance is obtained using a density assumption of the volume gained or lost (see Sect. 2.4). By assuming that the potential bed elevation changes are negligible, the geodetic mass balance covers all components of the surface, internal, and basal balances.

Sources of potential errors in elevation data can be categorized into sighting and plotting processes. Sighting includes errors that are related to the measurement process and originate from the platform, the sensor and the interference with the atmosphere. Plotting includes errors that relate to the analogue (e.g. map) or digital (e.g. DEM) representation of the sighting results including geo-referencing, projection, co-registration, and sampling density. In the case of historical surveys, uncertainties are introduced by the digitization of analogue contour maps for differencing with digital elevation models (DEMs) from newer surveys (cf. Koblet et al., 2010). Physical modelling of these errors is only possible with full information on sighting and plotting processes (e.g. Thibert et al., 2008; Joerg et al., 2012) which is often not available.

Alternatively, statistical approaches can be used to assess combined DEM errors by using the population of DEM differences over non-glacier terrain (assuming it is stable). In contrast to the physical error modelling, this approach incorporates all known and unknown error sources except errors that are spatially consistent in both DEMs. A principal bias in elevation differences is included from misalignment of the DEMs that are differenced. This misalignment translates into a bias in the derived elevation changes and is directly related to the combined slope and aspect distribution of a glacier. Therefore, we recommend to perform 3-D co-registration of the DEMs. An analytical relationship and simple solution to DEM misalignment is presented in Nuth and Kääb (2011), and the procedure is explained briefly in Supplement A.

If 3-D co-registration is successful, then the bias of the co-registration ($\overline{\mathrm{d}h}$ in Eq. A2) is removed and there remains the uncertainty of the vertical co-registration adjustment(s). One approach to estimate this potential unremoved vertical error is to introduce additional elevation datasets, either as a control or as a part of a time series. When three or more datasets are available, co-registration can be performed between each of these, and the summation of the 3-D co-registration vectors returns the residual error remaining within the series. Studies have shown that remaining vertical errors can reach magnitudes of at least 1–3 m (Nuth et al., 2012; Berthier et al., 2012) and should be included in the uncertainty assessment. Additional systematic errors in geodetic volume changes and balances can originate from changing reference areas (e.g. due to frontal fluctuations or ice divide migrations) and from glacier regions uncovered by the geodetic survey(s). It is therefore important to keep the glacier masks (and areas) consistent within both glaciological and geodetic analyses.

In addition to the errors related to the DEM co-registration, an uncertainty exists mainly related to the combined precision of the geodetic acquisition systems. For our statistical approach, the standard deviation of the elevation differences on stable terrain indicates the uncertainty of the DEM differences for individual pixels. The standard error indicates an uncertainty when spatially averaging the data such as for estimating glacier-wide changes. In the standard error, random errors are reduced by the number of independent measurements (i.e. pixels) and thus spatial auto-correlation commonly present in elevation data (e.g. Schiefer et al., 2007) must be accounted for (Etzelmüller et al., 2000). A method to determine the uncertainty related to the spatial auto-correlation based on semi-variogram analysis (σ_S in Eqs. B2 and B3) is described in Rolstad et al. (2009), and is summarized briefly in Supplement B. As discussed in Rolstad et al. (2009) there may be more than one scale of spatial correlation related to the derivation of the DEMs. It must be emphasized that it is generally the largest correlation scale that has the greatest impact on the spatially averaged uncertainty. A final consideration for statistical uncertainty analysis is whether the bedrock terrain surrounding the glacier is representative of the glacier surface. This depends upon the elevation acquisition technique (for example, in photogrammetry, low visible contrasting glacier surfaces may have larger random errors than high contrasting bedrock surfaces), the slope distribution of the surrounding topography versus glacier topography (Kääb et al., 2012), and/or if the differenced elevation data are of varying resolutions (Paul, 2008).

TCD
7, 789–839, 2013

In summary, the geodetic method measures the elevation differences integrating changes from all components of surface, internal, and basal balances. The result is subject to sighting and plotting errors which are best addressed by assessing the integrative errors of elevation differences over stable terrain surrounding the glacier. A 3-D co-registration of the DEMs – prior to elevation differencing – reduces the vertical elevation bias between the DEMs and needs to be followed by estimation of the uncertainties related to the remaining elevation error and to the spatial auto-correlation in the elevation differences.

2.4 Generic differences between glaciological and geodetic mass balance

A direct comparison of glaciological and geodetic balances requires accounting for survey differences (i.e. in time system and reference areas) as well as for generic differences between the glaciological and the geodetic balances (i.e. internal and basal balances). The corrections related to the survey differences need to account for ablation and accumulation between the glaciological and the geodetic surveys and for common reference areas with regard to ice divides and glacier boundary definitions Accounting for the generic differences basically means to quantify (if possible) the following mass balance components and related uncertainties: internal ablation (incl. heat conversion from changes in gravitational potential energy) internal accumulation, basal ablation (incl. ice motion, geothermal heat, and basal melt due to basal water flow), and basal accumulation.

The geodetic method measures changes in glacier volume and thus must be transformed from ice equivalent to water equivalent units (Eq. 8) for survey comparisons. Glacier elevation changes are a combined result of changes in the surface mass balance and the flux divergence at a point (Cuffey and Paterson, 2010). Below the ELA, changes are either ice ablation or emergence resulting in a density conversion equivalent to that of ice. In cases with known (observable) firn line changes, the density conversion over the area of firn coverage change can be approximated by an average density of firn and ice over those pixels (Sapiano et al., 1998). In areas with permanent firn cover, the density conversion depends on the relative contributions of surface and dynamical components to the elevation change and is commonly between 500 and 900 $\mathrm{kg\,m^{-2}}$. Special cases occur when a change in elevation results solely from firn compaction/expansion leading to volume changes with no associated mass change or in cases of increasing/decreasing elevations and firn compaction/expansion with depth, respectively when the mass conversion can be larger than the density of ice. Unless firn pack changes are carefully investigated and/or known, a first approximation is using a glacier-wide average density together with a plausible uncertainty range which includes Sorge's law (cf. Bader, 1954) as upper bound, such as $\overline{\rho} = 850 \pm 60\,\mathrm{kg\,m^{-3}}$ (cf. Sapiano et al., 1998; Huss, 2013), to be accounted for in the overall error budget. If biases are suspect, then sensitivity tests can help to determine the magnitude of bias potential in these density assumptions (e.g. Moholdt et al., 2010; Kääb et al., 2012; Nuth et al., 2012).

In summary, the comparison of glaciological and geodetic balances requires accounting for survey differences in time system and reference areas, for internal and basal mass balance components, and for errors related to the density conversion of the geodetic balances.

3 Conceptual framework for the re-analysis of glaciological and geodetic mass balance series

The glaciological method is able to capture the spatial and temporal variability of the glacier mass balance even with only a small sample of observation points (e.g. Lliboutry, 1974; Fountain and Vecchia, 1999) but is sensitive to systematic errors which accumulate linearly with the number of seasonal or annual measurements (Thibert et al., 2008). Hence, the ideal way to assess random and systematic errors is to combine the glaciological method with decadal geodetic surveys (Hoinkes 1970; Haeberli, 1998). In the following we present a comprehensive scheme for the entire re-analysis process including six principal steps (Fig. 1). The observation steps

TCD
7, 789–839, 2013

include measurements and documentation of glacier mass balance which are subject to methodological and observer-related artefacts. The aim of the homogenization is to reduce these artefacts whereas the uncertainty assessment is concerned with the estimation of remaining systematic (σ) and random (ε) errors. Validation compares the glaciological with the geodetic balance. In the event of significant differences, the iteration steps are designed to identify and quantify the corresponding error sources. Should a large bias of an unknown origin be revealed, the glaciological balance is calibrated to the geodetic balance.

3.1 Observations

For the glaciological method, observations at stakes and pits are carried out in seasonal or annual field surveys and later inter- and extrapolated to derive glacier-wide mass balance. Over the years the observational set up is subject to various changes such as in the stake and pit network, in the observers, in inter- and extrapolation methods, and in glacier extent. Similar inconsistencies are often present in geodetic data series. Due to the decadal intervals the individual surveys are usually carried out with different sensor and platform techniques, by different operators and analysts, and using different software packages and interpretation approaches. For later re-analysis it is important that the related meta-data are stored and made available with the observational results.

3.2 Homogenization

As a consequence of the heterogeneity of the observations, both the glaciological and the geodetic data series need to be homogenized independently. The aim of this step is use available (meta-)data, to homogenize the data series so that the annual (decadal) results are comparable for the entire glaciological (geodetic) series. Typical issues for the glaciological method are the change in inter- and extrapolation approaches (e.g. from contour line to altitude profile method), the use of different glacier catchments, or the annual (non-)adjustment of changing glacier extents (cf. Eqs. 4–6) For the geodetic method, the main task is to ensure that the digital terrain models from the different surveys are appropriately co-registered and that there is sufficient stable terrain surrounding the glacier or independent elevation data to quantify the uncertainties of spatially averaged elevation differences (as described in Sect. 2.3). In cases where earlier surveys resulted in topographic maps (with a focus on horizontal accuracy) it might be necessary to reprocess the original survey data (cf. Koblet et al., 2010). Homogenization aims at identifying and removing biases and quantifies related uncertainties. The procedures are strongly related to available (meta-)data from and process understanding about the glacier considered. Examples of detailed homogenization exercises are for example found in Huss et al. (2009), Fischer (2010), and Koblet et al. (2010).

3.3 Uncertainty Assessment

The aim of this third step is to estimate remaining systematic and random errors related to the homogenized glaciological and geodetic data series as well as to the generic differences between the two balances. Therefore, the uncertainties related to the lists of potential error sources above (Sects.s 2.2–2.4) need to be estimated and cumulated for time periods between consecutive geodetic surveys. The resulting parameters can be summarized as follows:

For each PoR covering N years, the mean annual glaciological balance $\overline{B_{\text{glac}.a}}$ is calculated as

$$\overline{B_{\text{glac}.a}} = \frac{1}{N}\sum_{t=1}^{n} B_{\text{glac}.a.t}. \quad (10)$$

Estimates of systematic (ε) and random (σ) errors related to the field measurement at point location (point), to the spatial integration (spatial), and to glacier area changing over time (ref) are described in Sect. 2.2.

The related total systematic error is expressed as the sums of individual sources (which can be of positive or negative signs) and years divided by the number of years

N of the PoR:

$$\overline{\varepsilon_{\text{glac.total}.a}} = \frac{\varepsilon_{\text{glac.total.PoR}}}{N} = \frac{\sum_{t=1}^{N} (\varepsilon_{\text{glac.point}.t} + \varepsilon_{\text{glac.spatial}.t} + \varepsilon_{\text{glac.ref}.t})}{N}, \tag{11}$$

Whereas the related total random error cumulates the individual sources and years according to the law of error propagation assuming they are not correlated:

$$\overline{\sigma_{\text{glac.total}.a}} = \frac{\sigma_{\text{glac.total.PoR}}}{\sqrt{N}} = \frac{\sqrt{\sum_{t=1}^{N} \sigma^2_{\text{glac.point}.t} + \sigma^2_{\text{glac.spatial}.t} + \sigma^2_{\text{glac.ref}.t}}}{\sqrt{N}}, \tag{12}$$

For reasons of comparability, the geodetic balance is also expressed as mean annual values

$$\overline{B_{\text{geod}.a}} = \frac{B_{\text{geod.PoR}}}{N} \tag{13}$$

together with estimates for systematic (ε) and random (σ) errors related to the combined DEM uncertainty remaining after co-registration and considering the spatial autocorrelation of elevation differences (DEM), see Sect. 2.3. The mean annual systematic error is expressed as:

$$\overline{\varepsilon_{\text{geod.total}.a}} = \frac{\varepsilon_{\text{geod.total.PoR}}}{N} = \frac{\varepsilon_{\text{geod.DEM.PoR}}}{N} \tag{14}$$

and is reduced to zero after successful 3-D co-registration (see Sect. 2.3).

The corresponding mean annual random error is estimated as

$$\overline{\sigma_{\text{geod.total}.a}} = \frac{\sigma_{\text{geod.total.PoR}}}{\sqrt{N}} = \frac{\sqrt{\sigma^2_{\text{geod.DEM.PoR}}}}{\sqrt{N}} = \frac{\sqrt{\sigma^2_{\text{coreg}+}\sigma^2_{\text{autocorr}}}}{\sqrt{N}} \tag{15}$$

TCD
7, 789–839, 2013

and integrates uncertainties related to the remaining elevation error (σ_{coreg}) and to the spatial auto-correlation in the elevation differences (σ_{autocorr}) as root sum of squares

For a direct comparison, both balances need to be corrected for systematic errors. In addition, the error estimates related to density conversion (dc) and survey differences (sd) are attributed to the geodetic balance. Deducting internal and basal balance estimates from the geodetic balance results in comparing surface balances and makes sense in case of a later calibration. The resulting corrected balances and their random errors are expressed as:

$$\overline{B_{\text{glac.corr}.a}} = \overline{B_{\text{glac}.a}} + \overline{\varepsilon_{\text{glac}.total.a}}, \text{ with} \tag{16}$$

$$\overline{\sigma_{\text{glac.corr}.a}} = \overline{\sigma_{\text{glac}.total.a}}, \text{ and} \tag{17}$$

$$\overline{B_{\text{geod.corr}.a}} = \overline{B_{\text{geod}.a}} + \overline{\varepsilon_{\text{geod.total}.a}} + \overline{\varepsilon_{\text{sd}.a}} - \overline{B_{\text{int}.a}} - \overline{B_{\text{bas}.a}} \tag{18}$$

with

$$\overline{\sigma_{\text{geod.corr}.a}} = \sqrt{\overline{\sigma_{\text{geod.total}.a}}^2 + \overline{\sigma_{\text{dc}.a}}^2 + \overline{\sigma_{\text{sd}.a}}^2 + \overline{\sigma_{\text{int}.a}}^2 + \overline{\sigma_{\text{bas}.a}}^2}. \tag{19}$$

In summary, this step is to provide estimates for both systematic and random errors for the glaciological and for the geodetic balances as well as for the generic differences between the two methods. Both balances are corrected for systematic errors in such a way that the results from the two methods can be validated against each other taking into account a possible calibration.

3.4 Validation

The corrected glaciological and geodetic balance series can be compared directly after having completed the three steps above. For this purpose, the corrected glaciological balances are cumulated over the time span between two geodetic surveys and then validated against the corresponding (decadal) value of the geodetic balance (cf. Eq. 20). The first check is to discern whether the discrepancy between the two methods can

TCD
7, 789–839, 2013

be explained by their natural scattering: if the random uncertainties of the two methods overlap, the corresponding difference is not statistically significant and the two data series cannot be considered as incoherent. A second intent of this test is to detect remaining systematic errors which may not be physically assessed or calculable for applying corrections.

Adopting conventional error risk (e.g. confidence intervals), the following statistical comparison test supports decisions concerning whether to accept the null-hypothesis H: the cumulated glaciological balance is not statistically different from the geodetic balance. We define the discrepancy Δ over the PoR as the difference between the cumulative glaciological and the geodetic balances, both corrected for identified systematic errors:

$$\Delta_{\text{PoR}} = B_{\text{glac.corr.PoR}} - B_{\text{geod.corr.PoR}} \tag{20}$$

The common variance of both methods is defined as the sum of both random uncertainties, cumulated over the PoR covering N years, following the law of error combination assuming that they are uncorrelated:

$$\sigma_{\text{common.PoR}} = \sqrt{\sigma^2_{\text{glac.corr.PoR}} + \sigma^2_{\text{geod.corr.PoR}}}, \text{ with} \tag{21}$$

$$\sigma_{\text{PoR}} = \overline{\sigma_a}\sqrt{N}, \tag{22}$$

and represents the total scattering of the data.

Finally, we can define the reduced discrepancy

$$\delta = \frac{\Delta_{\text{PoR}}}{\sigma_{\text{common.PoR}}}. \tag{23}$$

The more consistent the result of the two methods, relative to the total uncertainty, the closer δ is to zero. Working with a 95 % confidence interval (i.e. 1.96× sigma which corresponds to the often used 2× sigma uncertainty), we can accept the hypothesis H

TCD
7, 789–839, 2013

(i.e. $\Delta_{\text{PoR}} = 0$) if $-1.96 < \delta < 1.96$. Under this condition, there is a probability of $\alpha = 5\,\%$ making a wrong decision and rejecting H although the results of the two methods are actually equal (i.e. error of type I, false alarm). Alternatively using a 90 % confidence interval (i.e. 10 % probability for an error of type I), we can accept H if $1.64 < \delta < 1.64$. This means that the rejection of H is easier and more series qualify for calibration under this condition

In search of potential systematic errors in the observations, the substantial power of the statistical test is given by the ability to reject H when it is actually false and a significant difference εreally exists. If the test outcome is to accept H in that case, a second type error is therefore committed. This second type of risk, whose probability is denoted β, depends on the adopted risk α, and ε, and is given by:

$$\beta = F\left(u_\alpha - \frac{\varepsilon}{\sigma_{\text{common.PoR}}}\right) - F\left(-u_\alpha - \frac{\varepsilon}{\sigma_{\text{common.PoR}}}\right), \tag{24}$$

where F denotes the cumulative distribution function of the standard (zero-mean, unit-variance) normal distribution, and u_α is such as $F(u_\alpha) = \alpha$. For 5 and 10 % type I risks α, u_α equals 1.96 and 1.64, respectively. Under higher type I risk α (more series being appointed for a calibration), the risk β to maintain an incorrect glaciological series (not recalibrate when the series is actually erroneous) is naturally expected to decrease. This second type error risk can be calculated for each mass balance series assuming that the discrepancy ε corresponds to the measured difference Δ_{PoR}.

When the common variance of both methods is given, it is possible to estimate the lowest bias $\varepsilon_{\text{limit}}$ which is detectable. This detection limit can be calculated as:

$$\varepsilon_{\text{limit.PoR}} = \left(u_{1-\alpha/2} + u_{1-\beta}\right)\sqrt{\sigma^2_{\text{glac.PoR}} + \sigma^2_{\text{geod.PoR}}}, \tag{25}$$

where again u_γ is given by the cumulative distribution function of the standard normal distribution as $F(u_\gamma) = \gamma$. For $\alpha = \beta = 10\,\%$ admissible errors, $(u_{1-\alpha/2} + u_{1-\beta})$ is equal to 29, so that the detectable error is a little less than 3 times the common variance.

TCD
7, 789–839, 2013

Equation (25) indicates how the threshold of bias detection is lowered for longer time series. Adapted for annual values of systematic and random errors in the glaciological mass balance measurements, Eq. (25) becomes:

$$\varepsilon_{\text{limit.PoR}} = N\varepsilon_{\text{limit.}a} = \left(u_{1-\alpha/2} + u_{1-\beta}\right)\sqrt{N\sigma^2_{\text{glac.}a} + N\sigma^2_{\text{geod.}a}} \tag{26}$$

so that the detectable annual bias $\varepsilon_{\text{limit.}a}$ is given by:

$$\varepsilon_{\text{limit.}a} = \left(u_{1-\alpha/2} + u_{1-\beta}\right)\sqrt{\frac{\sigma^2_{\text{glac.}a}}{N} + \frac{\sigma^2_{\text{geod.}a}}{N}}. \tag{27}$$

Since annual uncertainties σ_{a} do not depend on N, the detectable systematic error is lowered as the PoR increases, and it decreases as $1/\sqrt{N}$ for long time series (see Fig. 5).

In summary, this step tests whether the unexplained bias between the glaciological and the geodetic methods is significant and provides estimates for the detectable bias. Corresponding calculation examples (for Eqs. 20–27) are given in Supplement Table C.

3.5 Iteration

Once a systematic error is detected with a high confidence level, a first step is to locate the corresponding error source by going back to the homogenization process and/or in the uncertainty assessment. The statistical exercise above thus helps to identify the survey period with the greatest discrepancies. Re-evaluating the available metadata for each potential source of error might raise issues which were not considered in the first round and might lead to a new homogenization effort for one or both methods. Re-evaluating the uncertainty assessment might reveal that uncertainties were over- or underestimated, or were not considered so far. However, any homogenization of the observations should be well supported by measurements or process understanding and not just for enforcing a match of the observations. Unexplained discrepancy requires interim calibration and further research.

3.6 Calibration

If a significant bias cannot be reduced with available (meta-)data and methods in the steps above one can take the decision to calibrate the glaciological balances – which are most sensitive to systematic error accumulation – with the geodetic results. The aim of the calibration is to maintain the relative seasonal/annual variability of the glaciological method but to adjust it to the absolute (decadal) values of the geodetic method. Procedures for calibration of mass balance series are described by Thibert and Vincent (2009) and by Huss et al. (2009) using statistical variance analysis and distributed mass balance modelling, respectively. Here we propose a simple approach without inferring the statistical linear model by Lliboutry (1974) or its expansion to unsteady state climate conditions (Eckert et al., 2011) and without the need for a numerical mass balance model.

Unless there is a clear hint of the origin of the remaining bias (i.e. not covered by the quantitative uncertainty assessment), the divergence from the geodetic balance is corrected in a first step by calibrating the annual glaciological balances as follows:

Over a PoR of N years, for which both glaciological and geodetic balances are available and homogenized, we calculate the mean annual glaciological balance $\overline{B_{\text{glac.corr.}a}}$ (see Eq. 16).

For each year t of the PoR, the centred glaciological balance β_t is calculated as the deviation from the mean:

$$\beta_t = B_{\text{glac.corr.}a.t} - \overline{B_{\text{glac.corr.}a}} \quad (28)$$

TCD
7, 789–839, 2013

Over the PoR it results that

$$\sum_{t=1}^{N} \beta_t = 0. \tag{29}$$

Over the same PoR, the mean annual geodetic balance $\overline{B_{\text{geod.corr.}a}}$ is calculated (see Eq. 18).

For each year of the PoR, the calibrated annual balance $B_{\text{cal.}t}$ is defined as

$$B_{\text{cal.}t} = \beta_t + \overline{B_{\text{geod.corr.}a}} \tag{30}$$

in which the mean comes from the geodetic and the year-to-year deviation from the glaciological balance.

For any year n within the PoR, the cumulative calibrated balance is

$$B_{\text{cal.}n} = \sum_{t=1}^{n} B_{\text{cal.}t} = n \cdot \overline{B_{\text{geod.corr.}a}} + \sum_{t=1}^{n} \beta_t. \tag{31}$$

For the last year of the PoR ($n = N$), the cumulative calibrated balance equals the product of N and the corrected annual geodetic balance because of Eq. (29)

In a second step, the seasonal balances are calibrated. Unless there is a clear hint that the bias comes from the spring observations, the winter balance B_w remains untouched as it is usually independent from the annual survey:

$$B_{\text{cal.w}} = B_{\text{glac.w}} \tag{32}$$

and the bias in the annual balance B_a is fully attributed to the summer balance B_s:

$$B_{\text{cal.s}} = B_{\text{cal.a}} - B_{\text{cal.w}}. \tag{33}$$

Thirdly, the balances of the elevation bins are adjusted to fit the calibrated annual (or seasonal) values. For each elevation bin e of each year t of the PoR, the centred elevation bin balance $\beta_{e.t}$ is calculated as the deviation from the un-calibrated annual glaciological balance:

$$\beta_{e.t} = B_{\mathrm{glac}.e.t} - B_{\mathrm{glac.corr}.a.t}. \tag{34}$$

Then, the calibrated elevation bin balance is defined as

$$B_{\mathrm{cal}.e.t} = \beta_{e.t} + B_{\mathrm{cal}.t}. \tag{35}$$

This approach basically shifts the glaciological balance profile (i.e. balance versus elevation) to fit the calibrated specific balance and, hence, maintains the balance gradient as long as the resolution of the elevation bins is high enough.

Finally new values for ELA and AAR can be derived conventionally from the calibrated balances of the elevation bins.

Re-analyzed mass balance series and derived parameters need to be flagged accordingly in any databases stored. This can be done by linking both glaciological and geodetic mass balance series through a lookup table including information on the re-analysis status (e.g. not re-analysed, homogenized only, validated but no calibration needed, validated and calibrated) and provide reference to related publications.

The calibration of glaciological mass balance series implies a bias of an unknown origin which might change over time (e.g. when a polythermal glacier becomes temperate). Note that the approach proposed here does not change the original stake and pit measurements but fits the glacier-wide results to the geodetic balance (see Sect. 5.2). This allows for reproducibility and later re-analysis exercises when new information about potential error sources or a new geodetic DEM becomes available.

4 Selected glaciers with long-term observation programmes

The analysis following in Sect. 5 is based on selected glaciers with long-term measurements including both glaciological and geodetic surveys and with available information

TCD
7, 789–839, 2013

for estimating related uncertainties. In general, the re-analysis steps were carried out according to the best practice as explained in Sects. 2 and 3 with individual deviations where more/less information was available for a more/less sophisticated approach.

An overview of the glaciers and PoRs used is provided in Table 1. The analyzed dataset consists of a total of 46 PoRs from 12 glaciers including: 38 decadal PoRs with an average time span of 11 yr, ranging from 4 to 32 yr, as well as 8 overall periods for glaciers with more than one PoR, and additional 9 PoRs with alternative calculations for Storglaciären (cf. Sect. 5.1). Lower and upper parts of Goldbergkees and Wurtenkees are analyzed separately due to the disintegration of the glacier before the analyzed PoRs (1998–2009 and 1998–2006). Details about glaciological and geodetic surveys and related uncertainty assessments are found in the following publications: Hynek et al. (2013) for Kleinfleisskees (FLK), Goldbergkees lower part (GLP) and upper part (GUP), Wurtenkees lower part (WLP) and upper part (WUP); Fischer and Markl (2009), Fischer (2010) for Hintereisferner (HEF), Jamtalferner (JAM) and Kesselwandferner (KWF); Moser et al. (1986), Reinwarth and Rentsch (1994), Reinwarth and Escher-Vetter (1999) for Vernagtferner (VER); Huss et al. (2009) for Gries (GRS) and Silvretta (SIL); Eckert et al. (2011), Thibert et al. (2008) for Sarennes (SAR); Geist et al. (2005), Elvehøy et al. (2009), Haug et al. (2009), Kjøllmoen et al. (2011, and earlier issues) for Engabreen (ENG); and Holmlund (1996), Albrecht et al. (2000), Holmlund et al. (2005), Koblet et al. (2010), Zemp et al. (2010) for Storglaciären (STO). All glaciological and most geodetic mass balance results are made available through the World Glacier Monitoring Service and published in WGMS (2012, and earlier volumes).

5 Results and discussion

5.1 Re-analysis of glacier mass balance series: the example of Storglaciären

Glaciological mass balance measurements have been carried out without interruption since 1945/46 together with aerial surveys at approximately decadal intervals

TCD
7, 789–839, 2013

(Holmlund et al., 2005). The resulting vertical photographs have been used to produce topographic maps which are described in detail by Holmlund (1996). However, the volume change assessment derived from digitizing these maps had been challenged by inaccuracies in maps and methodologies and revealed large discrepancies as compared to the glaciological balances over the same periods (Albrecht et al., 2000). Koblet et al. (2010) reprocessed diapositives of the original aerial photographs and produced a homogenized dataset of DEMs and a related uncertainty estimate. Based on these new DEMs, Zemp et al. (2010) re-analysed the glaciological and geodetic mass balance series of Storglaciären including a detailed uncertainty assessment. Their main conclusions were that both the new geodetic and the glaciological balances (between 1959 and 1999) fit well as long as systematic corrections for internal accumulation, as proposed by Schneider and Jansson (2004) are ignored. The conceptual framework introduced above for re-analyzing mass balance series now allows these conclusions to be reproduced and quantified.

The parameters required for a statistical decision in the event that the cumulative glaciological balance significantly differs from the geodetic balance (i.e. rejection of H) are shown in Table 2 as explained in Sect. 3.4. For each PoR, the cumulative glaciological balance is corrected for systematic errors as well as for generic differences to the geodetic balance and given together with the random uncertainties. The geodetic balances, also corrected for systematic errors, are given with their random uncertainties. The cumulative discrepancy shows the difference between the two balances and is put into context with the random uncertainties (through the common variance). This results in the reduced discrepancy which allows statistical quantifying if the two balances fit or not as shown in the following examples:

The results including the old DEMs (from Albrecht et al., 2000) for the periods 1969–80 and 1980–90 both have reduced discrepancies far beyond the 90 % and 95 % confidence intervals and, hence, show that the glaciological are significantly different from the geodetic balances (i.e. H to be rejected). Interestingly, there is no such discrepancy for the overall PoR (1959–90) which is because the two strongly erroneous decades have cumulated discrepancies of opposite signs. This nicely demonstrates the importance of testing both the entire PoR as well as the individual decadal intervals.

Comparing the geodetic results from the homogenized DEMs (from Koblet et al., 2010) with the glaciological findings shows the improvements in the periods 1969–80 and 1980–90 with much smaller cumulated discrepancies. H is now clearly accepted for both periods. However, the additional decade (1990–99) reveals significant differences between the methods in spite of a cumulated discrepancy similar to those of the other accepted periods. Here, the reason is the better quality of DEMs which results in smaller uncertainties (i.e. a smaller common variance) and, hence, allows for an improved detection of a systematic difference.

For the same dataset (using the DEMs from Koblet et al., 2010), the entire PoR (1959–99) shows a large cumulated discrepancy of more than 3 m w.e. As a consequence, H is to be accepted at the 95 % but to be rejected at the 90 % confidence interval. After checking all assumptions of the uncertainty assessment, the reason is most probably to be found in the above-mentioned overestimation of the internal accumulation. This is also indicated by the fact that for all periods, the reduced discrepancies are positive, i.e. the geodetic results are more negative than the glaciological ones. The correction applied here for internal accumulation (i.e. 3–5 % of the annual accumulation) which is based on estimates by Schneider and Jansson (2004) of re-freezing of percolation water in cold snow and firn as well as of the freezing of water trapped by capillary actions in snow and firn by the winter cold based on data from 1997/98 and 1998/99. Reijmer and Hock (2008) find the internal accumulation to amount to as much as 20 % of the winter accumulation in 1998/99 based on a snow model coupled to a distributed energy- and mass balance model. Our comparison with the geodetic method indicates that these estimates might be valid for the investigated periods but – applied as a general correction to all year – seem to exaggerate the contribution of the internal accumulation to the annual balance.

Finally, the results of the new DEMs (from Koblet et al., 2010) compared with the glaciological balances excluding corrections for internal accumulation show the best fit

TCD
7, 789–839, 2013

with smallest cumulated and reduced discrepancies, and clear acceptances of H for all periods. As a consequence, no calibration of the glaciological balance is needed over the re-analyzed period (1959–99). However, In spite of the relatively small discrepancies between the two methods ($< 100\,\mathrm{mm\,w.e.\,a^{-1}}$) there still is a great risk of not detecting a remaining bias. Future research can address this by trying to reduce the errors such as by a co-registration of the existing elevation grids to a high-precision reference DEM of a new survey.

5.2 Calibration of glacier mass balance series: the example of Silvretta-gletscher

Comparison of glaciological and geodetic mass balance series of Silvrettagletscher for the periods 1994–2003, and 2003–2007 indicates a significant bias beyond the uncertainties. Huss et al. (2009) homogenized the measurement series by recalculating seasonal mass balances based on the raw data and calibrated the cumulative glaciological balance with the geodetically determined mass change. Here, an example for the calibration of the original mass balance series for SIL is provided according to the theoretical framework described in Sect. 3.6.

For the two PoRs, the differences between glaciological mass balance and the geodetic surveys are considerable (Fig. 2a). Whereas the cumulative glaciological balance 1994–2007 is 3.1 m w.e., the geodetic mass balance indicates a cumulative balance of $-7.9\,\mathrm{m\,w.e.}$ over the same period. According to the statistical method (Sect. 3.4) this difference is significant at the 95 % level and H is rejected. Since the related error source (s) could not be clearly identified and corrected, the series for the two PoRs 1994–2003 and 2003–07 thus need to be calibrated.

First, the centered glaciological balance β_t is calculated as the deviation from the period mean $\overline{B_{\mathrm{glac.corr.}a}}$ (see Eq. 28), and β_t is subsequently shifted with the mean annual geodetic mass balance $\overline{B_{\mathrm{geod.corr.}a}}$ (see Eq. 30). This results in a calibrated series that represents a conventional mass balance covering all components of the

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
TCD
7, 789–839, 2013

surface balance. The long-term (decadal) changes in glacier mass are provided by the geodetic surveys, and the year-to-year variability of the original series based on the direct glaciological method is preserved (Fig. 2a). The mean annual bias between the original glaciological and the geodetic balance is distributed equally over all balance years between two geodetic surveys.

A calibration of the annual mass balance series requires consequent changes to be applied to the seasonal balances, the altitudinal mass balance distribution, as well as ELA and AAR values in order to provide a consistent set of variables. As in the case of SIL the measurements of winter accumulation are independent of the annual surveys, and there is no indication that the winter balance is biased; the misfit is fully attributed to the summer balance (Fig. 2b). The mass balance elevation-distribution remains similar, but is shifted for each year according to the mean annual bias (see Eqs. 34 and 35). ELA and AAR for the calibrated series are determined from the corrected mass balance distribution (Fig. 2b).

5.3 Uncertainties of glacier mass balance series: comparison of a larger sample

The development of the abovedescribed conceptual framework for the re-analysis of mass balance series strongly builds on the experience from glaciers with detailed and long-term mass balance monitoring programmes. Here we analyze glaciological and geodetic balances with related uncertainties from roughly 50 PoRs with data available from the glaciers in Table 1. All reported values and statistics are given in Supplement Table C. For summary statistics, only independent PoRs are analysed, omitting the overall PoR for glaciers with more than one PoR. It is to be noted that for most of these periods and glaciers it was not possible to quantify systematic and random errors for all potential error sources. However, the available sample nevertheless reveals interesting insights into the uncertainty of glacier mass balance.

On average, the (corrected) glaciological balances of the investigated PoRs are negative with $-454\,\mathrm{mm\,w.e.\,a^{-1}}$ and a corresponding random error of $340\,\mathrm{mm\,w.e.\,a^{-1}}$ The related random uncertainty of the field measurements at point location is estimated to

140 mm w.e. a^{-1}. For all but two glaciers, this estimate refers to the point measurement itself; uncertainties for density measurement and superimposed ice are not specified or assumed to be zero. This value is within the range but at the lower end of corresponding estimates found in the literature, e.g. Meier et al. (1971): 100–340, Lliboutry (1974): 300, Cogley and Adams (1998): 200–400, Gerbeaux et al. (2005): 100 for ice ablation, 250–400 for firn ablation, Vallon and Leibva (1981) for drilling in accumulation area: 300 (all values in mm w.e. a^{-1}).

The spatial integration of the glaciological point measurement is attributed with random uncertainties of 278 mm w.e. a^{-1} which is attributed to the local representativeness and to the interpolation method with 325 and 140 mm w.e. a^{-1}, respectively. The extrapolation to unmeasured areas is usually not specified and/or considered to be covered by the uncertainty of the interpolation method. These estimates are larger than corresponding values found in the literature, e.g. by Vallon and Leiva (1981; 70 mm w.e. a^{-1}) or by Fountain and Vecchia (1999). Uncertainties due to reference area changing over time are attributed with an average absolute bias of 12 mm w.e. a^{-1} and a corresponding random uncertainty of 62 mm w.e. a^{-1}.

The average geodetic balance of the investigated PoRs (i.e. −574 mm w.e. a^{-1}) is slightly more negative than the glaciological result. Accounting for the time period between the geodetic surveys, the mean annual error is 146 mm w.e. a^{-1}. The remaining elevation bias (from PoRs without DEM co-registration) is estimated on the average to 3 mm w.e. a^{-1} with a random uncertainty of 114 mm w.e. a^{-1}. The spatial correlation of the elevation differences are not specified so far. The uncertainty related to the density conversion is attributed to 74 mm w.e. a^{-1}.

Differences in time system and reference areas are quantified with a bias of −3 mm w.e. a^{-1} and a random uncertainty of 18 mm w.e. a^{-1}. Differences due to internal and basal balances are usually not specified or assumed to be zero. The few estimates account about 10 mm w.e. a^{-1} for internal ablation (STO and SAR), between 4 and 100 mm w.e. a^{-1} for internal accumulation (STO, SAR, VER) and between 1 and 9 mm w.e. a^{-1} for basal ablation (STO, GRS, SIL). Note that the sample of estimates

TCD
7, 789–839, 2013

for the generic differences is rather small and biased to temperate glaciers. Nevertheless, it becomes obvious that more research is needed to show whether estimates of internal and basal components, typically derived from short measurement periods, can be applied to long-term mass balance series. As such, estimates for internal and basal ablation are similar to the ones found at Gulkana Glacier by March and Trabant (1997; 5 mm w.e. a^{-1} due to both geothermal heat flow and ice motion as well as 60 mm w.e. a^{-1} due to potential energy loss by water flow). However, Alexander et al. (2011) find a much higher contribution, i.e. > 10 % of basal melt to the total ablation of Franz Josef Glacier which they explain by the strongly maritime environment of the glacier. Proposed corrections for internal accumulation seem to be very large; cf. Trabant and Mayo (1985) for Alaskan glaciers: 7–64 %, Trabant and Benson (1986) for McCall: 40 %, Schneider and Jansson (2004) for STO: 3–5 %, Reijmer and Hock (2008) for STO: 20 % (all relative to annual accumulation).

On average, overall uncertainties are 340 and 146 mm w.e. a^{-1} for (corrected) glaciological and geodetic balances, respectively. Individual PoRs with values greater than 500 m w.e. a^{-1} are found for HEF (glac), KWF (glac) and GRS (glac). Absolute values for bias corrections are mostly below 50 mm w.e. a^{-1}. Note that all analyzed series were at least partly homogenized which aims at reducing these systematic errors Regression analyses show no correlations between the balances and the discrepancy or the common variance. This means that neither the difference between the two methods nor the random uncertainties depends on the value or the sign of the balances.

A comparison of glaciological and geodetic balance results, corrected for biases and generic differences, is shown in Fig. 3. In the case of a perfect fit, all points would align on the line of equal glaciological and geodetic balances. There is a slight tendency for the points to be located below this line which is an indication of more negative geodetic balances. The mean value for the annual discrepancy (Δ_a) of our sample is +120 mm w.e. a^{-1} with a root mean square of mm w.e. a^{-1} 226. This tendency of more negative geodetic balances can stem from a positive or negative undetected bias in the glaciological or in the geodetic balance, respectively, or in an underestimation of the generic differences or density assumptions required to fit the glaciological to the geodetic balance. However, for the majority of the points, the deviation from this line is within the random uncertainties. The few exceptions are two points of HEF (1964–69, 1979–91), one point of KWF (1997–06), two points of SIL (1994–03, 2003–07), and all three PoRs of ENG. For SIL the bias is probably related to the reduction in the number of stakes in the mid-1980s. Measurement errors in the accumulation zone could have contributed to the differences between glaciological and geodetic method (Huss et al., 2009). Cogley (2009) compared direct and geodetic based on 105 common PoRs from 29 glaciers but without an individual glacier uncertainty assessment. He found a (statistically also not significant) negative mean annual discrepancy of −74 mm w.e. a^{-1} and a root mean square of +378 mm w.e. a^{-1}

The decision as to whether the corrected glaciological balance is significantly different from the corrected geodetic balance (i.e. rejection of H) is based on the reduced discrepancy which considers both the cumulated discrepancy between the results of the two methods and the common variance. The reduced discrepancies of all PoRs are plotted in Fig. 4. Working with a 95 % confidence interval, H is accepted for 35 out of all 46 PoRs and 11 (2× HEF, 1× KWF, 3× SIL, GLP, 1× VER, 3× ENG) are candidates for a calibration. Setting the confidence interval to 90 % increases the number of candidates for a calibration to 14 (adding 2× SAR, WLP). This is because a lowering of the confidence interval increases the detectability of the lowest systematic difference between the two methods. The location of points in the middle of the Gaussian curve is to be seen as a preliminary indication for the agreement of the results from the two methods. Checking for large common variances helps identify PoRs where the ability to detect a systematic difference between the methods is low. In the case of STO, the PoR (1990–99) with a reduced variance close to 1.0 is probably more reliable then the three other decadal PoRs with reduced variances between −0.2 and +0.4: their common variances, and hence their lowest detectable differences, are twice or more the value of the period 1990–99 (with a lowest detectable bias of about 200 mm w.e. a^{-1}). Other PoRs with accepted H but low ability for detecting systematic differences (i.e. lowest detectable bias > 500 mm w.e. a^{-1}) are found for GRS (all decadal periods), HEF (all but one decadal periods), and KWF (1997–2006).

Working with a 90 % confidence interval, the mean of all lowest detectable differences decreases from 476 to 387 mm w.e. a^{-1}. At the same time, lowering the confidence interval increases the probability α for an error of type I (i.e. false alarm leading to a calibration exercise of series without a bias) and reduces the probability β of an error of type II (i.e. not detecting erroneous series) In general, this β risk of not calibrating an incorrect series is quite high (for series in Supplement Table C around 7 % in average for $\alpha = 5$ %), except if very large discrepancies are observed between the glaciological and the geodetic results. This risk decreases just slightly if we choose to calibrate more systematically (62 %), running a higher risk α of 10 % in recalibrating (uselessly) a correct series. For example, for GRS (1967–79) where glaciological and geodetic results match well ($\delta = 0.02$), β only decreases from 95 to 90 % if α is set from 5 to 10 %. β decreases to a greater extent for series showing strong discrepancies between geodetic and glaciological methods, as Silvretta (2003–07) with β decreasing from 50 to 37 % when the stress on α risk set loose. For co-registered DEMs from high-quality geodetic surveys with very small random uncertainties such as from airborne laser scanning (cf. Joerg et al., 2012), one could, hence, consider setting the confidence interval to zero and calibrating the glaciological balances in any case. However, as long as the geodetic uncertainties are in the same order of magnitude as those of the glaciological method and as long as the differences between the two methods (i.e. density conversion, internal and basal balances) cannot be well quantified, the 90 % confidence interval can be deemed a good selection criterion.

From the above statistical exercise, it becomes evident that the ability to detect a systematic difference between the glaciological and the geodetic method depends primarily on the size of the random uncertainty. The level of confidence sets the threshold for calibration. In other words, setting the estimates for the random uncertainties in an extremely conservative manner (i.e. assuming worst case values for every potential source of error) reduces the ability to detect a systematic difference between the methods. Figure 5 shows that the detectable annual bias decreases with the length of the PoR. This is explained by error propagation since the glaciological balance has to be cumulated for comparison to the geodetic balance: if a systematic error occurs annually in the glaciological balance, it grows linearly from year to year. Along the PoR, the random errors of the cumulated glaciological balance also accumulate but only in proportion to the square root of the number of years. For all data series gathered in the present study, approximately one decade is required for systematic error accumulation to surpass random error sum enough to become detectable with a significant confidence level (90 % in our calculations). Thus long periods are required to detect biases among the natural scattering of the observations and this statistical comparison test is weaker over PoRs of less than 10 yr. Consequently, we recommend testing to discover whether calibration is worthwhile when a long period (> 10 yr) of control is available from the geodetic balance.

5.4 Recommendations for principal investigators and implications for data users

From the presented exercise in assessing uncertainties and re-analysis of mass balance measurement, we make twelve general statements as guidelines for the benefit of data producers and users.

Recommendations for investigators of glacier mass balance: (I) glaciological mass balance programmes are ideally complemented from the very beginning with decadal geodetic surveys. (II) Such geodetic surveys should use sensors optimized for snow and ice surveys, be carried out towards the end of the ablation season (i.e. with minimal snow cover), and cover the entire glacier system as well as surrounding stable terrain (for uncertainty assessments). (III) As a rule of thumb, absolute differences between the glaciological and geodetic balances (a) smaller than 100 mm a^{-1} indicate that the two methods are probably consistent, (b) between 100 and 200 mm a^{-1} indicate that a re-analysis might be appropriate, and (c) larger than 200 mm a^{-1} indicate that a re-analysis is urgently needed. (IV) Mass balance series longer than 20 yr should be

TCD
7, 789–839, 2013

re-analyzed in any case. (V) Every mass balance series should be clearly flagged in publications and databases with its re-analysis status. (VI) More research is needed to better understand and quantify the potential error sources and related systematic and random errors (cf. Sects. 2.2–2.4). Important issues are the influence of the interpolation method on the glaciological balance, the density conversion of the geodetic balance, and the quantification of the internal balance components especially for polythermal and cold glaciers.

Implications for users of glacier mass balance data: (VII) the glaciological method measures the components of the surface balance. (VIII) The geodetic method measures all components of the surface, internal, and basal balances. (IX) The results of the two methods provide conventional balances which incorporate both climate forcing and changes in glacier hypsometry and represent the glacier contribution to runoff and sea level rise; for climate-glacier investigations, the reference-surface balance might be a more relevant quantity (cf. Elsberg et al., 2001; Huss et al., 2012). (X) The results of the two methods can be compared as long as temporal and spatial differences in the survey as well as the internal and basal balances are accounted for or can be assumed to be negligible. (XI) Both the glaciological and the geodetic balances are subject to systematic and random errors related to various sources. Overall uncertainties are typically a few hundred but sometimes a few thousands mm w.e. per year. (XII) Re-analysis of (especially of long-term) mass balance series based on both methods allows the quantification of the related uncertainties and of remaining unexplained biases.

Finally, calibration of a glaciological mass balance series (as explained in Sect. 3.6) implies a large biases of unknown origin and efforts should focus on determining the source of the biases or at least suggesting potential error sources for future research and re-analysis.

TCD
7, 789–839, 2013

6 Conclusions

Based on the experience from long-term monitoring and a series of three workshops, this paper briefly summarizes the glaciological and geodetic methods and provides a comprehensive list of potential error sources for both methods. We propose a conceptual framework for re-analysing glacier mass balance series including a statistical toolset for assessing random and systematic errors as well as for calibrating the glaciological with the geodetic balances.

Our analysis of 50 periods of record from a dozen of European glaciers shows that both the glaciological and the geodetic balances are subject to errors from various sources. Thereby, systematic errors have typical values below one hundred mm w.e. per year but cumulate with the length of the observation record for cumulative glaciological series. The cumulative bias in the derived glacier mass balance presents a challenge to the efforts to detect climatic trends. Random uncertainties have typical values of a few hundred mm w.e. per year, but cumulate due to the law of error propagation. Biases between glaciological and geodetic balances are therefore easier to detect for longer time spans of a decade or more.

The proposed re-analysis scheme allows the direct comparison of the results from glaciological and geodetic methods and helps identify significant differences between the two data series. If the detected difference cannot be explained and removed by a re-evaluation of the homogenization and of the uncertainty assessment, we recommend calibrating the glaciological to the geodetic balances. The balance series calibrated in this way maintains the inter-annual variation of the original glaciological data at the same time as cumulative values become consistent with the decadal geodetic balances and provides an optimal estimate of the surface balance until the source(s) of bias can be identified and quantified.

The re-analysis of glacier mass balance series should become a standard procedure for every mass balance monitoring programme with increasing importance for long time series. Users of re-analyzed datasets profit from improved data quality and uncertainty

TCD
7, 789–839, 2013

estimates which will lead, it is hoped, to a more thorough interpretation of glacier mass balance results in the future.

Supplementary material related to this article is available online at: http://www.the-cryosphere-discuss.net/7/789/2013/tcd-7-789-2013-supplement.zip.

Acknowledgements. We are grateful to Wilfried Haeberli for his experienced input into both the concept and the manuscript and to Susan Braun-Clarke for carefully editing the English of this paper. The present work is an outcome of a "Workshop on Measurement and Uncertainty Assessment of Glacier Mass Balance" held July 9–11, 2012, in Tarfala, Sweden and organized by the World Glacier Monitoring Service in collaboration with Stockholm University. The workshop was a "Marcus Wallenberg Symposium" sponsored by the Marcus Wallenberg Foundation for International Cooperation in Science and co-sponsored by the International Union of Geology and Geophysics (IUGG) through the International Association of Cryospheric Sciences (IACS) as well as by the International Glaciological Society (IGS).

References

Ageta, Y. and Fujita, K.: Characteristics of mass balance of summer-accumulation type glaciers in the Himalayas and Tibetan Plateau, Z. Gletscherkd. Glazialgeol., 32, 61–65, 1996.

Albrecht, O., Jansson, P., and Blatter, H.: Modelling glacier response to measured mass-balance forcing, Ann. Glaciol., 31, 91–96, 2000.

Alexander, D., Shulmeister, J., and Davies, T.: High basal melting rates within high-precipitation temperate glaciers, J. Glaciol., 57, 789–795, 2011.

Andreassen, L. M., Kjøllmoen, B., Rasmussen, A., Melvold, K., and Nordli, Ø.: Langfjordjøkelen, a rapidly shrinking glacier in northern Norway, J. Glaciol., 58, 581–593, 2012.

Arendt, A. A., Echelmeyer, K. A., Harrison, W. D., Lingle, C. S., and Valentine, V. B.: Rapid wastage of Alaska glaciers and their contribution to rising sea level, Science, 297, 382–386, 2002.

Bader, H.: Sorge's Law of densification of snow on high polar glaciers, J. Glaciol., 2, 319–323, 1954.
Berthier, E., Arnaud, Y., Kumar, R., Ahmad, S., Wagnon, P., and Chevallier, P.: Remote sensing estimates of glacier mass balances in the Himachal Pradesh (Western Himalaya, India), Remote Sens. Environ., 108, 327–338, 2007.
Berthier, E., Schiefer, E., Clarke, G. K. C., Menounos, B., and Rémy, F.: Contribution of Alaskan glaciers to sea-level rise derived from satellite imagery, Nat. Geosci., 3, 92–95, 2010.
Berthier, E., Scambos, T. A., and Shuman, C. A.: Mass loss of Larsen B tributary glaciers (Antarctic Peninsula) unabated since 2002, Geophys. Res. Lett., 39, L13501, doi:10.1029/2012GL051755, 2012.
Braithwaite, R. J.: Assessment of mass-balance variations within a sparse stake network, Qamanârssûp Sermia, West Greenland, J. Glaciol., 32, 50–53, 1986.
Braithwaite, R. J., Konzelmann, T., Marty, C., and Olesen, O. B.: Errors in daily ablation measurements in northern Greenland, 1993–94, and their implications for glacier climate studies, J. Glaciol., 44, 583–588, 1998.
Chinn, T.: Structure and equilibrium of the Dry Valley glaciers, New Zealand Antarctica Records, 6, 73–88, 1985.
Cogley, J.: Geodetic and direct mass-balance measurements: comparison and joint analysis, Ann. Glaciol., 50, 96–100, 2009.
Cogley, J. G. and Adams W. P.: Mass balance of glaciers other than the ice sheets, J. Glaciol., 44, 315–325, 1998.
Cogley, J. G., Hock, R., Rasmussen, L. A., Arendt, A. A., Bauder, A., Braithwaite, R. J., Jansson, P., Kaser, G., Möller, M., Nicholson, L., and Zemp, M.: Glossary of Glacier Mass Balance and Related Terms. IHP-VII Technical Documents in Hydrology No. 86, IACS Contribution No. 2, UNESCO-IHP, Paris, 114 pp., 2011.
Cox, L. H. and March, R. S.: Comparison of geodetic and glaciological mass-balance techniques, Gulkana Glacier, Alaska, USA, J. Glaciol., 50, 363–370, 2004.
Cuffey, K. and Paterson, W. S. B.: The Physics of Glaciers, 5th edn., Massachusetts Academic Press, Burlington, 704 pp., 2010
Eckert, N., Thibert, E., and Vincent, C.: Extracting the temporal signal from a winter and summer mass-balance series: application to a six-decade record at Glacier de Sarennes, French Alps, J. Glaciol., 57, 134–150, 2011

Elsberg, D. H., Harrison, W. D., Echelmeyer, K. A., and Krimmel, R. M.: Quantifying the effects of climate and surface change on glacier mass balance, J. Glaciol., 47, 649–658, 2001.
Elvehøy, H., Jackson, M., and Andreassen, L. M.: The influence of drainage boundaries on specific mass-balance results: a case study of Engabreen, Norway, Ann. Glaciol., 50, 135–140, 2009.
Escher-Vetter, H., Kuhn, M., and Weber, M.: Four decades of winter mass balance of Vernagtferner and Hintereisferner, Austria: methodology and results, Ann. Glaciol., 50, 87–95, 2009
Etzelmüller, B.: On the quantification of surface changes using grid-based Digital Elevation Models (DEMs), Transactions in GIS, 4, 129–143, 2000.
Finsterwalder, R. and Rentsch, H.: Zur Höhenänderung von Ostalpengletschern im Zeitraum 1969–1979, Z. Gletscherkd. Glazialgeol., 16, 111–115, 1981.
Fischer, A.: Glaciers and climate change: interpretation of 50 years of direct mass balance of Hintereisferner, Global Planet. Change, 71, 13–26, 2010.
Fischer, A.: Comparison of direct and geodetic mass balances on a multi-annual time scale, The Cryosphere, 5, 107–124, doi:10.5194/tc-5-107-2011, 2011.
Fischer, A. and Markl, G.: Mass balance measurements on Hintereisferner, Kesselwandferner and Jamtalferner 2003 to 2006: database and results, Z. Gletscherkd. Glazialgeol., 42, 47–83, 2009.
Fountain, A. G. and Vecchia, A.: How many stakes are required to measure the mass balance of a glacier?, Geogr. Ann. A, 81, 563–573, 1999.
Funk, M., Morelli, R., and Stahel, W.: Mass balance of Griesgletscher 1961–1994: different methods of determination, Z. Gletscherkd. Glazialgeol., 33, 41–56, 1997.
Geografiska Annaler: Special issue on Methods of Mass Balance Measurements and Modelling, Geogr. Ann. A, 81, 461–796, 1999.
Geist, T., Elvehøy, H., Jackson M., and Stötter J.: Investigations on intra-annual elevation changes using multitemporal airborne laser scanning data – case study Engabreen, Norway, Ann. Glaciol., 42, 195–201, 2005.
Gerbaux, M., Genthon, C., Etchevers, P., Vincent, C., and Dedieu, J. P.: Surface mass balance of glaciers in the French Alps: distributed modeling and sensitivity to climate change, J. Glaciol., 51, 561–572, 2005.
Haeberli, W.: Historical evolution and operational aspects of worldwide glacier monitoring, in: Into the Second Century of Worldwide Glacier Monitoring: Prospects and Strategies, edited by: Haeberli, W., Hoelzle, M., and Suter S., UNESCO-IHP, Paris, France, 35–51, 1998.

Hagen, J. O., Eiken, T., Kohler, J., and Melvold, K.: Geometry changes on Svalbard glaciers: mass-balance or dynamic response?, Ann. Glaciol., 42, 255–261, 2005.
Hagg, W., Braun, L., Uvarov, V., and Makarevich, K.: A comparison of three methods of mass-balance determination in the Tuyuksu glacier region, Tien Shan, Central Asia, J. Glaciol., 50, 505–510, 2004.
Haug, T. Rolstad, C., Elvehøy, H., Jackson, M., and Maalen-Johansen, I.: Geodetic mass balance of the western Svartisen ice cap, Norway, in the periods 1968–1985 and 1985–2002, Ann. Glaciol., 50, 119–125, 2009.
Hoinkes, H. C.: Methoden und Möglichkeiten von Massenhaushaltsstudien auf Gletschern: Ergebnisse der Messreihe Hintereisferner (Ötztaler Alpen) 1953–1968, Z. Gletscherkd. Glazialgeol., 6, 37–90, 1970.
Holmlund, P.: Maps of Storglaciären and their use in glacier monitoring studies, Geogr. Ann. A, 78, 193–196, 1996.
Holmlund, P., Jansson, P., and Pettersson, R.: A re-analysis of the 58 year mass-balance record of Storglaciären, Sweden, Ann. Glaciol., 42, 389–394, 2005.
Hock, R. and Jensen, H.: Application of kriging interpolation for glacier mass balance computations, Geogr. Ann. A, 81, 611–619, 1999.
Huss, M.: Density assumptions for converting geodetic glacier volume change to mass change, The Cryosphere Discuss., 7, 219–244, doi:10.5194/tcd-7-219-2013, 2013.
Huss, M., Bauder, A., and Funk, M.: Homogenization of long-term mass-balance time series, Ann. Glaciol., 50, 198–206, 2009.
Huss, M., Hock, R., Bauder, A., and Funk, M.: Conventional versus reference-surface mass balance, J. Glaciol., 58, 278–286, 2012.
Hynek, B., Unger, R., and Schöner, W.: Comparison of geodetic and glaciological mass balances for 5 small Alpine glaciers, in preparation, 2013.
IGS: Thematic issue on mass balance measurements and modelling, Ann. Glaciol., 50, 220 pp., 2009.
Jansson, P.: Effect of uncertainties in measured variables on the calculated mass balance of Storglaciären, Geogr. Ann. A, 81, 633–642, 1999.
Joerg, P. C., Morsdorf, F., and Zemp, M.: Uncertainty assessment of multi-temporal airborne laser scanning data: a case study on an Alpine glacier, Remote Sens. Environ., 127, 118–129, 2012.

Kääb, A., Berthier, E., Nuth, C., Gardelle, J., and Arnaud, Y.: Contrasting patterns of early twenty-first century glacier mass change in the Himalayas, Nature, 488, 495–498, 2012.
Kaser, G. and Osmaston, H.: Tropical Glaciers, UNESCO International Hydrological Series, Cambridge University Press, Cambridge, UK, 207 pp., 2002.
Kaser, G., Fountain, A. G., and Jansson, P.: A Manual for Monitoring the Mass Balance of Mountain Glaciers with Particular Attention to Low Latitude Characteristics, a Contribution to the UNESCO HKH-Friend Programme, Paris, France, 107 pp., 2003.
Kaser, G., Cogley, J., Dyurgerov, M., Meier, M., and Ohmura, A.: Mass balance of glaciers and ice caps: consensus estimates for 1961–2004, Geophys. Res. Lett., 33, L19501, doi:10.1029/2006GL027511, 2006.
Kjøllmoen, B., Andreassen, L. M., Elvehøy, H., Jackson, M., and Giesen, R. H.: Glaciological investigations in Norway in 2010, NVE Report 2011-3, 89 pp., 2011.
Koblet, T., Gärtner-Roer, I., Zemp, M., Jansson, P., Thee, P., Haeberli, W., and Holmlund, P.: Reanalysis of multi-temporal aerial images of Storglaciären, Sweden (1959–99) – Part 1: Determination of length, area, and volume changes, The Cryosphere, 4, 333–343, doi:10.5194/tc-4-333-2010, 2010.
Lang, H. and Patzelt, G.: Die Volumenänderungen des Hintereisferners (Ötztaler Alpen) im Vergleich zur Massenänderung im Zeitraum 1953–64, Z. Gletscherkd. Glazialgeol., 7, 229–238, 1971.
Lliboutry, L.: Multivariate statistical analysis of glacier annual balances, J. Glaciol., 13, 371–392, 1974.
March, R. S. and Trabant, D. C.: Mass balance, meteorological, ice motion, surface altitude, and runoff data at Gulkana Glacier, Alaska, 1993 balance year, USGS Water-Resources Investigations Report 964299, 30 pp., 1997.
Meier, M. F., Tangborn, W. V., Mayo, L. R., and Post, A.: Combined ice and water balances of Gulkana and Wolverine Glaciers, Alaska, and South Cascade Glacier, Washington, 1965 and 1966 hydrologic years, USGS Prof. Pap. 715-A, 23 pp., 1971.
Miller, M. and Pelto, M.: Mass balance measurements on the Lemon Creek Glacier, Juneau Icefield, Alaska 1953–1998, Geogr Ann. A, 81, 671–681, 1999.
Moholdt, G., Nuth, C., Hagen, J., and Kohler, J.: Recent elevation changes of Svalbard glaciers derived from ICESat laser altimetry, Remote Sens. Environ., 114, 2756–2767, 2010.

TCD
7, 789–839, 2013

Uncertainties and re-analysis of glacier mass balance measurements

M. Zemp et al.

Moser, H., Escher-Vetter, H., Oerter, H., Reinwarth, O., and Zunke, D.: Abfluß in und von Gletschern, GSF-Bericht 41/86, Teil I und II, GSF Gesellschaft für Strahlen- und Umweltforschung, München, 408 pp. and 147 pp., 1986.

Nussbaumer, S. U., Zemp, M., and Jansson, P.: Summary report on the "Workshop on Measurement and Uncertainty Assessment of Glacier Mass Balance" 9–11 July 2012, Tarfala, Sweden, World Glacier Monitoring Service, Zurich, Switzerland, 11 pp., 2012.

Nuth, C. and Kääb, A.: Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change, The Cryosphere, 5, 271–290, doi:10.5194/tc-5-271-2011, 2011.

Nuth, C., Schuler, T. V., Kohler, J., Altena, B., and Hagen, J. O.: Estimating the long-term calving flux of Kronebreen, Svalbard, from geodetic elevation changes and mass-balance modelling, J. Glaciol., 58, 119–133, 2012.

Østrem, G. and Brugman, M.: Glacier Mass-Balance Measurements: A Manual for Field and Office Work, NHRI Science Report, Saskatoon, Canada, 224 pp., 1991.

Østrem, G. and Haakensen, N.: Map comparison or traditional mass-balance measurements: which method is better?, Geogr. Ann. A, 81, 703–711, 1999.

Paul, F.: Calculation of glacier elevation changes with SRTM: is there an elevation dependent bias?, J. Glaciol, 55, 945–946, 2008.

Radić, V. and Hock, R.: Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise, Nat Geosci, 4, 91–94, 2011.

Reijmer, C. and Hock, R: Internal accumulation on Storglaciären, Sweden, in a multilayer snow model coupled to a distributed energy- and mass-balance model, J. Glaciol., 54, 61–72, 2008.

Reinwarth, O. and Escher-Vetter, H.: Mass balance of Vernagtferner, Austria, from 1964/65 to 1996/97: results for three sections and the entire glacier, Geogr Ann. A, 81, 743–751, 1999.

Reinwarth, O. and Rentsch, H.: Volume and mass balance of Vernagtferner/Oetztal Alps, Z. Gletscherkd. Glazialgeol., 30, 99–107, 1994.

Rolstad, C., Haug, T., and Denby, B.: Spatially integrated geodetic glacier mass balance and its uncertainty based on geostatistical analysis: application to the western Svartisen ice cap, Norway., J. Glaciol., 55, 666–680, 2009.

Sapiano, J., Harrison, W., and Echelmeyer, K.: Elevation, volume and terminus changes of nine glaciers in North America, J. Glaciol., 44, 119–135, 1998.

Schiefer, E., Menounos, B., and Wheate, R.: Recent volume loss of British Columbian glaciers, Canada, Geophys. Res. Lett., 34, L16503, doi:10.1029/2007GL030780, 2007.
Schneider, T. and Jansson P.: Internal accumulation within firn and its significance for the mass balance of Storglaciären, Sweden J. Glaciol., 50, 25–34, 2004.
Schytt, V.: Re-freezing melt-water on the surface glacier ice Geogr. Ann., 31, 222–227, 1949.
Soruco, A., Vincent, C., Francou, B., Ribstein, P., Berger, T., Sicart, J., Wagnon, P., Arnaud, Y., Favier, V., and Lejeune, Y.: Mass balance of Glaciar Zongo, Bolivia, between 1956 and 2006, using glaciological, hydrological and geodetic methods, Ann. Glaciol., 50, 1–8, 2009.
Thibert, E. and Vincent, C.: Best possible estimation of mass balance combining glaciological and geodetic methods, Ann. Glaciol., 50, 112–118, 2009.
Thibert, E., Blanc, R., Vincent, C., and Eckert, N.: Glaciological and volumetric mass-balance measurements: error analysis over 51 years for Glacier de Sarennes, French Alps, J. Glaciol., 54, 522–532, 2008.
Trabant, D. C. and Benson, C. S.: Influence of internal accumulation and superimposed ice formation on mass balance of McCall Glacier in Alaska Data of Glaciological Studies, 58, 40–49, 1986.
Trabant, D C. and Mayo, L. R.: Estimation and effects of internal accumulation on five glaciers in Alaska, Ann. Glaciol., 6, 113–117, 1985.
Vallon, M.: Errors in the determination of ablation using stakes, J. Glaciol., 7, 132–133, 1968.
Vallon, M. and Leiva, J. C.: Bilans de masse et fluctuations récentes du Glacier de Saint-Sorlin (Alpes Francaises), Z. Gletscherkd. Glazialgeol., 17, 143–167, 1981.
Weber, M., Braun, L., Mauser, W., and Prasch, M.: Contribution of Rain, Snow- and Icemelt in the upper Danube Discharge today and in the Future, Bulletino del Comitato Glaciologico Italiano, Ser. 3 Geogr. Fis. Din. Quat., 33, 221–230, 2010.
WGMS: Fluctuations of Glaciers 2005–2010, vol. X, edited by: Zemp, M., Frey, H., Gärtner-Roer, I., Nussbaumer, S. U., Hoelzle, M., Paul, F., and Haeberli, W., ICSU(WDS)/IUGG(IACS)/UNEP/ UNESCO/WMO, World Glacier Monitoring Service, Zurich, Switzerland, publication based on database version: doi:10.5904/wgms-fog-2012-11, 336 pp., 2012.
Wright, A. P., Wadham, J. L., Siegert, M. J., Luckman, A., Kohler, J., and Nuttall, A. M.: Modeling the refreezing of meltwater as superimposed ice on a high Arctic glacier: a comparison of approaches, J. Geophys. Res., 112, 2007.

TCD
7, 789–839, 2013

Zemp, M., Hoelzle, M., and Haeberli, W.: Six decades of glacier mass-balance observations: a review of the worldwide monitoring network, Ann. Glaciol., 50, 101–111, 2009

Zemp, M., Jansson, P., Holmlund, P., Gärtner-Roer, I., Koblet, T., Thee, P., and Haeberli, W.: Reanalysis of multi-temporal aerial images of Storglaciären, Sweden (1959–99) – Part 2: Comparison of glaciological and volumetric mass balances, The Cryosphere, 4, 345–357, doi:10.5194/tc-4-345-2010, 2010.

TCD

7, 789–839, 2013

Table 1. Overview of glaciers used in this study with information about glaciological and geodetic surveys. Analyzed periods of record are indicated by included geodetic survey years highlighted in bold. Methods are abbreviated as follows: t = terrestrial, a = airborne, T = tachymetry, P = photogrammetry, L = laserscanning.

PU	Name	LAT	LON	glac. surveys (first/last/ # obs. years)	geodetic surveys (year$_{method}$)
AT	Goldbergkees lower part (GLP) upper part (GUP)	47.05° N	12.96° E	1989/2012/24	1909$_{tP}$, –31$_{tP}$, –53$_{aP}$, –69$_{aP}$, –79$_{aP}$, –92$_{aP}$, **–98**$_{aP}$, **2009**$_{aL}$
AT	Hintereisferner (HEF)	46.80° N	10.77° E	1953/2012/60	1893$_{tP}$, **1953**$_{tP}$, **–64**$_{tP}$, **–67**$_{tP}$, **–69**$_{aP}$, **–79**$_{aP}$, **–91**$_{aP}$, **–97**$_{aP}$, **2006**$_{aL}$
AT	Jamtalferner (JAM)	46.87° N	10.17° E	1989/2012/24	1969$_{aP}$, **–96**$_{aP}$, –2002$_{aP}$, **–06**$_{aL}$
AT	Kesselwandferner (KWF)	46.84° N	10.79° E	1953/2012/60	1969$_{aP}$, **–71**$_{aP}$, –97$_{aP}$, **2006**$_{aL}$
AT	Kleinfleisskees (FLK)	47.05° N	12.95° E	1999/2012/14	1931$_{tP}$, –53$_{aP}$, –69$_{aP}$, –79$_{aP}$, –92$_{aP}$, **–98**$_{aP}$, **2009**$_{aL}$
AT	Wurtenkees lower part (WLP) upper part (WUP)	47.04° N	13.00° E	1983/2012/30	1969$_{aP}$, –79$_{aP}$, –91$_{tP}$, **–98**$_{aP}$, **2006**$_{aL}$
AT	Vernagtferner (VER)	46.87° N	10.82° E	1965/2012/48	1889$_{tP}$, 1912$_{tP}$, –38$_{tP}$, **–69**$_{aP}$, **–79**$_{aP}$, **–90**$_{aP}$, **–99**$_{aP}$, **2006**$_{aL}$, **–09**$_{aL}$
CH	Griesgletscher (GRS)	46.44° N	8.33° E	1962/2012/51	**1961**$_{aP}$, **–67**$_{aP}$, **–79**$_{aP}$, **–86**$_{aP}$, **–91**$_{aP}$, **–98**$_{aP}$, **2003**$_{aP}$, **–07**$_{aP}$
CH	Silvrettagletscher (SIL)	46.85° N	10.08° E	1960/2012/53	**1959**$_{aP}$, **–73**$_{aP}$, **–86**$_{aP}$, **–94**$_{aP}$, **2003**$_{aP}$, **–07**$_{aP}$
FR	Sarennes (SAR)	45.07° N	6.07° E	1949/2012/64	1908$_{tP}$, **–52**$_{aP}$, **–81**$_{aP}$, **2003**$_{aP}$
NO	Engabreen (ENG)	66.40° N	13.50° E	1970[a]/2012/43	**1968**$_{aP}$, **2001**$_{aL}$, **–08**$_{aL}$
SE	Storglaciären (STO)	67.90° N	18.57° E	1946/2012/67	1910$_{tP}$, –49$_{aP}$, **–59**$_{aP}$, **–69**$_{aP}$, **–80**$_{aP}$, **–90**$_{aP}$, **–99**$_{aP}$, 2008$_{aP}$

[a] At ENG, the glaciological observations started one year after the first geodetic survey. The corresponding difference in time system (cf. Sect. 2.4) was accounted for using a positive degree-day model.

Table 2. Summary statistics for the comparison of glaciological and geodetic balances of Storglaciären, Sweden. For different dataset combinations, the table shows analyzed periods of record (PoR) with bias-corrected balances (B.corr) and related random uncertainties ($\pm\sigma$) for both the glaciological (glac) and the geodetic (geod) methods together with cumulated discrepancies (Δ_{PoR}), common variance ($\sigma_{\mathrm{common.PoR}}$), and reduced discrepancy (δ) calculated according to Eqs. (20)–(23). The acceptance or rejection of the hypothesis (H: the two balances are equal) is evaluated on the 95 % and 90 % confidence level (i.e. first type risk) which corresponds to reduced discrepancies outside the ± 1.96 and ± 1.64 range, respectively. For the same confidence intervals, the second type risk β (cf. Eq. 24) of not detecting an erroneous series is given too.

Dataset combination middle	PoR Year	B.glac.corr $\pm\sigma$ [mm w.e.]	B.geod.corr $\pm\sigma^{a}$ [mm w.e.]	Δ_{PoR} [mm w.e.]	$\sigma_{\mathrm{common.PoR}}$ [mm w.e.]	δ no unit	H_0: 95/90 no unit	β : 95/90 %
B.glac versus B.geod.old, incl. intACC	1959–69	−3060 ± 1183	−2130 ± 449	−930	1265	−0.74	yes/yes	89/81
	1969–80	−2409 ± 1091	−6215 ± 899	3806	1414	2.69	no/no	23/15
	1980–90	990 ± 585	3720 ± 873	−2730	1051	−2.60	no/no	26/17
	1959–90	−4464 ± 1715	−4619 ± 1325	155	2167	0.07	yes/yes	95/90
B.glac versus B.geod.new, incl. intACC	1959–69	−3060 ± 1183	−4040 ± 449	980	1265	0.77	yes/yes	88/80
	1969–80	−2409 ± 1091	−2794 ± 899	385	1414	0.27	yes/yes	94/89
	1980–90	990 ± 585	110 ± 873	880	1051	0.84	yes/yes	87/78
	1990–99	720 ± 429	−351 ± 240	1071	492	2.18	no/no	41/29
	1959–99	−3760 ± 1714	−7000 ± 487	3240	1782	1.82	yes/no	56/43
B.glac versus B.geod.new, excl.intACC	1959–69	−3060 ± 1183	−3560 ± 449	500	1265	0.40	yes/yes	93/87
	1969–80	−2409 ± 1091	−2156 ± 899	−253	1414	−0.18	yes/yes	95/89
	1980–90	990 ± 585	760 ± 873	230	1051	0.22	yes/yes	94/89
	1990–99	720 ± 429	234 ± 240	486	492	0.99	yes/yes	83/74
	1959–99	−3760 ± 1714	−4640 ± 487	880	1782	0.49	yes/yes	92/86

[a] All random uncertainties are based on the new DEMs by Koblet et al. (2010) because the old ones by Albrecht et al. (2000) did not include terrain outside the glacier.

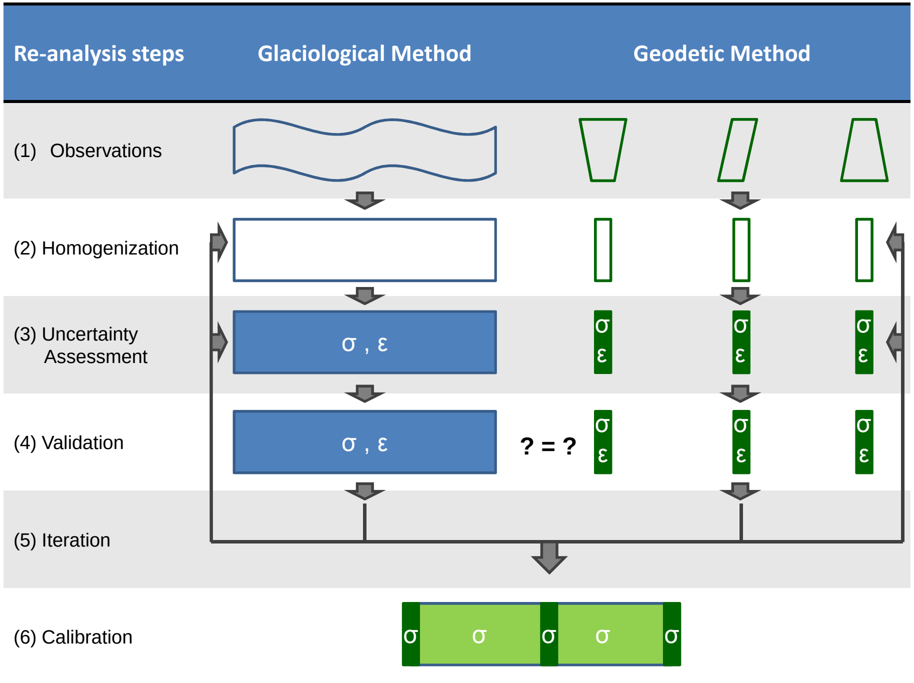

Fig. 1. Generic scheme for the re-analysis of glacier mass balance series in six steps as described in Sects. 3.1–3.6. A series of n annual glaciological observations and three decadal geodetic surveys are independently homogenized and assessed for systematic (ε) and random (σ) errors. Resulting glaciological balances are validated and calibrated (if necessary) against/with geodetic balances in order to reduce unexplained biases.

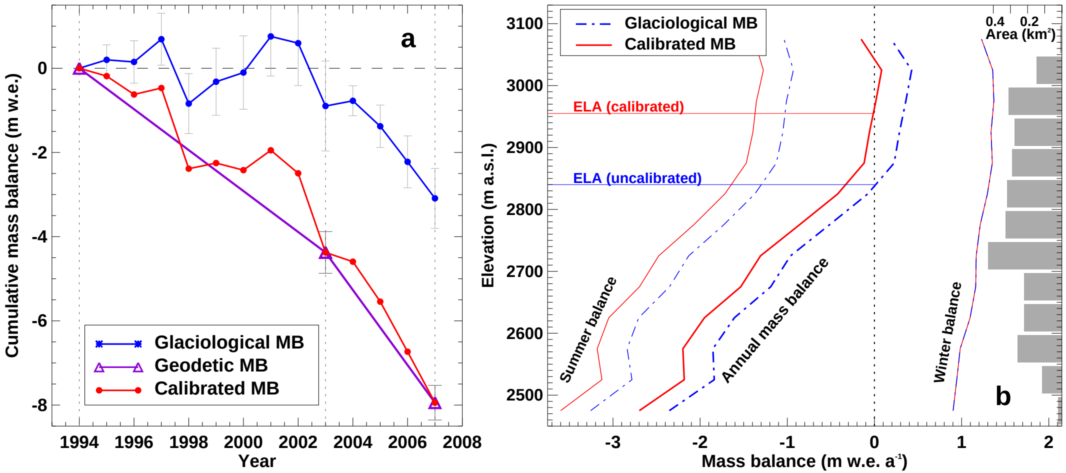

Fig. 2. Calibration of glaciological mass balance series for the periods 1994–2003, and 2003–07 with the geodetic surveys for Silvrettagletscher (cf., Huss et al., 2009). **(a)** Cumulative mass balance (original glaciological mass balance and calibrated with the geodetic mass change). Uncertainties according to the uncertainty analysis are given. **(b)** Mass balance elevationdistribution (original and calibrated glaciological series) as a mean over the period 2003–07. Both seasonal and annual mass balances are shown, the original and calibrated ELA is indicated, and glacier hypsometry is given.

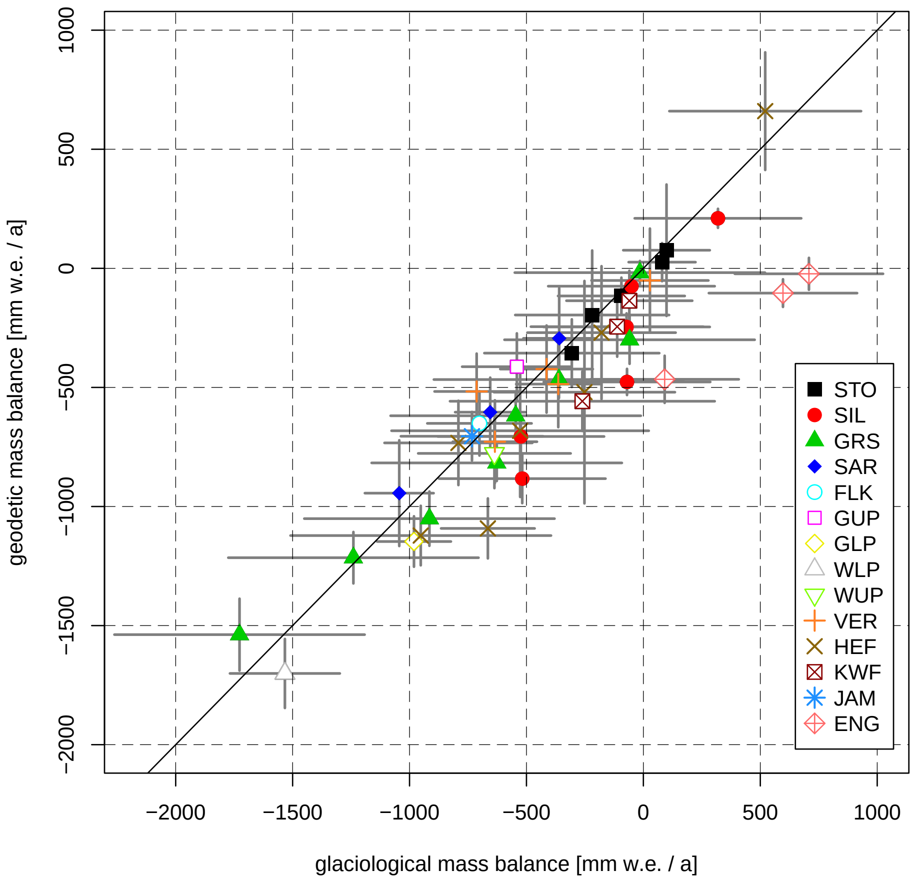

Fig. 3. Glaciological versus geodetic balances. Both series are corrected for biases and generic differences and plotted with random uncertainties. The black line marks equal balances from both methods.

TCD
7, 789–839, 2013

Uncertainties and re-analysis of glacier mass balance measurements

M. Zemp et al.

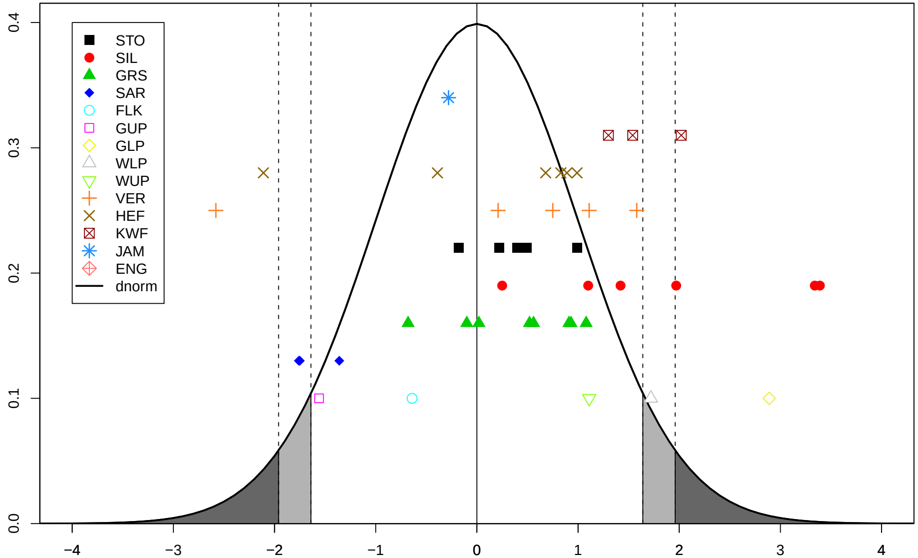

Fig. 4. Reduced discrepancies between all of the analyzed periods of record. The reduced discrepancy (x-axis) has no unit and the values for the different glaciers are arbitrary distributed along the y-axis for a better overview. The curve labelled "dnorm" denotes the probability density function for the standard (zero-mean, unit-variance) normal distribution. Shaded areas in dark and light grey indicate 95 % and 90 % confidence intervals, respectively. The values for ENG are all > 4 and, hence, not plotted.

TCD
7, 789–839, 2013

Uncertainties and re-analysis of glacier mass balance measurements

M. Zemp et al.

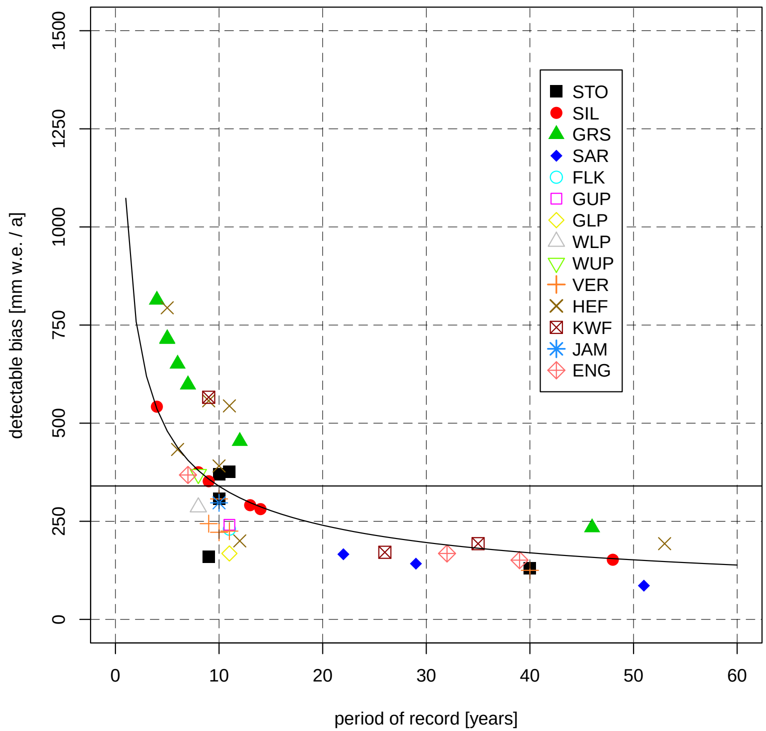

Fig. 5. Detectable bias as a function of the period of record. The horizontal line marks the mean random error of all glaciological balance series ($\overline{\sigma_{\mathrm{glac}.a}}$ = 340 mm w.e. a^{-1}). The curved line marks the minimum detectable annual bias (at 10 % error risk) as a function of the number of years of the series as given by Eq. (27) using average values of random errors (i.e. glaciological and geodetic measurements) for all data series. On average, ten years of data are required for the detectable bias to get lower than the annual random "noise" of the glaciological balance represented by $\overline{\sigma_{\mathrm{glac}.a}}$.

TCD
7, 789–839, 2013

Uncertainties and re-analysis of glacier mass balance measurements

M. Zemp et al.