The Cryosphere, 7, 1227–1245, 2013
www.the-cryosphere.net/7/1227/2013/
doi:10.5194/tc-7-1227-2013

The Cryosphere
Open Access

Reanalysing glacier mass balance measurement series

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

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

Received: 1 February 2013 – Published in The Cryosphere Discuss.: 4 March 2013
Revised: 10 June 2013 – Accepted: 21 June 2013 – Published: 6 August 2013

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 recently, comprehensive uncertainty assessments have rarely been carried out and mass balance data have often been applied using rough error estimation or without consideration of errors. In this study, we propose a framework for reanalysing glacier mass balance series that includes 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. Reanalysing glacier mass balance series needs to become a standard procedure for every monitoring programme to improve data quality, including reliable uncertainty estimates.

1 Introduction

Changes in glacier mass are a key subject 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, most of these data series consist 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 on relatively small mountain and valley glaciers (Zemp et al., 2009). Combined with multi-annual geodetic surveys, these long-term glaciological mass balance series provide a unique opportunity for 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

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

studies have carried out extensive homogenization and uncertainty assessments for reanalysing 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 no guidelines available yet for standardization of the 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 (GAA, 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 and to find a consensus on best practices, mainly for homogenization, validation, and calibration of (long-term) observation series.

The present paper is a joint outcome of that workshop and aims at proposing best practices for reanalysing mass balance series. First, we provide a brief review of observation methods, related uncertainties, and reanalysing procedures for observation series. Second, we present results from a select number of glaciers with long-term mass balance programmes and discuss these in light of the proposed reanalysing scheme. Finally, we conclude with recommendations for data producers and summarize 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" of Cogley et al. (2011). "Homogenization" and "reanalysis" of observational data series are well established in climatology (cf. Kalnay et al., 1996; Aguilar et al., 2003; Begert et al., 2005). However, the corresponding methods, developed for treating an atmospheric continuum, cannot be directly applied to discrete glaciers. We, hence, use these terms in their general climatological meaning but specify the methodological implementation for glacier mass balance series in Sect. 3. In the same chapter, general definitions and glacier-specific explanations are given for "validation" and "calibration", which are used differently among communities. In terms of the uncertainty assessment, we differentiate between random (i.e. noise) and systematic (i.e. bias) errors (i.e. disagreements between measured and true values).

The mass balance of a glacier is defined as the sum of all components of accumulation (acc) and ablation (abl), and a distinction can be made between surface (sfc), internal (int), and basal (bas) balances (see Table 1 and Fig. 2 in Cogley et al., 2011). 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

$$\dot{m} = \dot{\mathrm{acc}}_{\mathrm{sfc}} + \dot{\mathrm{abl}}_{\mathrm{sfc}} + \dot{\mathrm{acc}}_{\mathrm{int}} + \dot{\mathrm{abl}}_{\mathrm{int}} + \dot{\mathrm{acc}}_{\mathrm{bas}} + \dot{\mathrm{abl}}_{\mathrm{bas}} + \frac{q_{\mathrm{in}} + q_{\mathrm{out}}}{\mathrm{d}s}, \quad (1)$$

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 b_a (or more generally over the span of time from t_0 to t_1) is linked to the mass balance rate by

$$b_\mathrm{a} = m\,(t_\mathrm{a}) - m\,(t_0) = \int_{t_0}^{t_\mathrm{a}} \dot{m}\,(t)\,\mathrm{d}t. \quad (2)$$

To obtain the glacier-wide mass balance, the point balances are integrated over the glacier mean area S over the same time span:

$$B_\mathrm{a} = \frac{1}{S} \cdot \int_S b_\mathrm{a} \mathrm{d}s. \quad (3)$$

Note that this study focuses on land-terminating glaciers; the balance components of lake and marine floating glacier tongues and ice shelves are not considered here because their mass balance is often dominated by frontal and basal terms not addressed by the glaciological method (cf. Kaser et al., 2003).

2.2 Glaciological observation method

The principal steps of the glaciological observation method are the measurement of ablation and/or accumulation at individual points as well as the interpolation between the measurement points and extrapolation to unmeasured regions of the glacier. Often, the interpolation and extrapolation process incorporates mass balance indicators, such as snowline observations and related expert knowledge. The glaciological 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 earliest measurements. However, the detailed implementation does vary between different glaciers and observers. The number and density of stake and snow pit observations varies from glacier to glacier and through time (e.g. Fountain and Vecchia, 1999; Miller and Pelto, 1999; Van Beusekom et al., 2010). Another typical variation 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 analysis or 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. On some glaciers the measurements are performed at monthly (on some inner-tropical glaciers) or even at daily resolution (at a few points during summer seasons). Observers at some cold or polythermal glaciers account for internal accumulation too (e.g. Josberg et al., 2007). The results are usually reported 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 metre water equivalent per year ($\mathrm{m\,w.e.\,a^{-1}}$). Equilibrium line altitude (ELA), accumulation area ratio (AAR), and mass balance gradients are usually calculated from mass balance distribution with elevation (ranges).

There are three main sources of random and systematic errors in the glaciological method: the field measurements at point locations, the spatial averaging of these results over the entire glacier, and the changes of glacier in area and elevation. The field measurements are subject to errors in (i) height determination (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 measurement errors and associated assumptions (with errors expected to be larger for snow and firn than for ice); (iii) superimposed ice, which is difficult to measure and of which the spatial variability is often not well captured by the stake network (e.g. Schytt, 1949; Wright et al., 2007); and (iv) flux divergence which is irrelevant to the glacier-wide balance (cf. Cuffey and Paterson, 2010) unless the sampling between divergence and convergence zones is unbalanced (Vallon, 1968). Error sources related to the spatial averaging of the point measurements are (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; e.g. Fountain and Vecchia, 1999; Pelto, 2000), (vi) the method (e.g. contour, profile, kriging) used for interpolation between the point observations and for extrapolation to unmeasured regions (e.g. Hock and Jensen, 1999; Escher-Vetter et al., 2009), and (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). Common to all mass balance series is (viii) The issue of the glacier elevation and area changing over time: the (changing) coordinates and elevation of observation points can directly be measured whereas 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 bands) for every year in order to provide "conventional" balances (cf. Elsberg et al., 2001; Huss et al., 2012). A simple analytical solution to this type of inhomogeneity is given in Sect. 3.2. The few studies which have attempted to quantify all these errors include Thibert et al. (2008), Huss et al. (2009), Fischer (2010), Zemp et al. (2010), Hynek et al. (2012), and references therein.

Observation principles were mainly developed on and for land-terminating, mid-latitude glaciers in the Northern Hemisphere, which mainly change by winter accumulation and summer ablation. In practice, these principles and the relative importance of the error sources listed above might be of limited applicability to the seasonal analysis of glaciers in other regions. However, cumulative annual balances eliminate these seasonal complexities which, hence, might not be relevant to the following comparison with multi-annual geodetic balances.

2.3 Geodetic observation method

The geodetic observation method determines volume change by repeated mapping and differencing of glacier surface elevations. Common methods are ground surveys using theodolites (e.g. Lang and Patzelt, 1971) or global navigation satellite systems (e.g. Hagen et al., 2005), airborne or spaceborne surveys with photogrammetry (e.g. Finsterwalder and Rentsch, 1981; Berthier et al., 2007), and SAR interferometry (e.g. Magnusson et al., 2005; Berthier et al., 2007) or various forms of laser-altimetry (e.g. Sapiano et al., 1998; Geist et al., 2005; Moholdt et al., 2010). For large ice fields and ice caps, accurate elevation data are often limited to a selection of survey profiles along glacier centre lines or a systematic pattern of ground-tracks from satellite altimetry. A comparison with the glaciological balance can be done along common centre lines or after extrapolation to the entire glacier. The uncertainty and potential bias related to this extrapolation (e.g. Arendt et al., 2002; Berthier et al., 2010) need to be accounted for in a similar manner as for the glaciological method. The methodological description below focuses on DEM (digital elevation model) differencing over 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:

$$\Delta V = r^2 \sum\nolimits_{k=1}^{K} \Delta h_k, \tag{4}$$

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 interannual 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 metre (Eq. 4). Commonly, the geodetic balance is obtained by making an assumption about the density of the volume gained or lost (see Eq. 5 in Sect. 2.4). If it is true that the change of bed elevation is 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 of the atmosphere. Plotting errors 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. Additional systematic errors in geodetic volume changes can originate from changing reference areas (e.g. due to frontal fluctuations or ice divide migrations) and from glacier regions not covered by the geodetic survey(s). It is therefore important to keep the glacier masks (and areas) consistent both within and between glaciological and geodetic analyses. Physical modelling of above 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 performing 3-D co-registration of the DEMs. An analytical relationship and simple solution for DEM misalignment is presented in Nuth and Kääb (2011), and the procedure is explained briefly in Supplement A, Eqs. (A1)–(A4).

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, defined as the standard deviation divided by the square root of the number of independent items of information in the sample, indicates an uncertainty when spatially averaging the data such as for estimating glacier-wide changes. However, for the calculation of the standard error the number of independent items cannot be assumed to be equal to the number of items in the sample (i.e. pixels) because spatial auto-correlation is commonly present in elevation data (e.g. Schiefer et al., 2007) and must be accounted for (Etzelmüller, 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.

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, glacier surfaces with low visible contrast may have larger random errors than high-contrast bedrock surfaces), the slope distribution of the surrounding topography versus glacier topography (Kääb et al., 2012), and/or whether the differenced elevation data are of varying resolutions (Paul, 2008).

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) and 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. Also, both methods must use common reference areas (with regard to ice divides and glacier boundary definitions) in order to ensure that results of the same glacier system are being compared. Accounting for the generic differences basically means to quantify (if possible) the following mass balance components and related uncertainties: internal ablation (including heat conversion from changes in gravitational potential energy), internal accumulation, basal ablation (including ice motion, geothermal heat, and basal melt due to basal water flow), and basal accumulation.

For a comparison with the glaciological balance, the geodetic volume change must be converted into a specific mass balance over a period of record (PoR) in the unit metre water equivalent (m w.e.):

$$B_{\mathrm{geod.PoR}} = \frac{\Delta V}{\overline{S}} \cdot \frac{\overline{\rho}}{\rho_{\mathrm{water}}}, \tag{5}$$

where $\overline{\rho}$ is the average density of ΔV, assuming no change in bulk glacier density over the balance period, and $\overline{S}$ is the average glacier area of the two surveys at time *t0* and *t1* assuming a linear change through time as

$$\overline{S} = \frac{S_{t0} + S_{t1}}{2}. \tag{6}$$

Glacier elevation changes are a combined result of changes in surface, internal, and basal balance, and the flux divergence at a point (Cuffey and Paterson, 2010). Below the

ELA, changes are either ice ablation or emergence, so that the appropriate density is 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 appropriate density depends on the relative contributions of surface and dynamical components to the elevation change and is commonly between 500 and 900 $\mathrm{kg\,m^{-3}}$. Special cases occur when a change in elevation results solely from firn compaction or 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 to use a glacier-wide average density together with a plausible uncertainty range, such as $\overline{\rho} = 850 \pm 60\,\mathrm{kg\,m^{-3}}$ (cf. Sapiano et al., 1998; Huss, 2013). If biases are suspected, then sensitivity tests can help to determine the potential magnitude of bias in these density assumptions (e.g. Moholdt et al., 2010; Kääb et al., 2012; Nuth et al., 2012; Huss, 2013).

3 Conceptual framework for reanalysing glaciological and geodetic mass balance series

Reanalysis is defined by Cogley et al. (2011) as the reexamination and possible modification of a series of measurements in the light of methods or data not available when the measurements were made. In order to avoid confusion with the climatological "reanalysis" product (cf. Kalnay et al., 1996), we use the terms "reanalyse" or "reanalysing" in this paper.

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 (Cox and March, 2004; Thibert et al., 2008). The geodetic balance is able to cover the entire glacier but requires a density conversion and is carried out at multi-annual intervals. Hence, the ideal way to reanalyse a mass balance series is to combine the glaciological method with multi-annual geodetic surveys (Hoinkes, 1970; Haeberli, 1998). In the following, we present a comprehensive scheme for the entire reanalysing process including six principal steps (Fig. 1). The observation step (Sect. 3.1) includes measurements and documentation of glacier mass balance which are subject to methodological and observer-related inhomogeneities. The aim of the homogenization (Sect. 3.2) is to reduce these inhomogeneities whereas the uncertainty assessment (Sect. 3.3) is concerned with the estimation of remaining systematic (ε) and random (σ) errors. Validation (Sect. 3.4) compares the glaciological with the geodetic balance. In the event of significant differences, the iteration step (Sect. 3.5) is designed to identify and quantify the corresponding error sources. Should a large difference of an unknown origin be revealed, the glaciological balance is calibrated to the geodetic balance (Sect. 3.6).

3.1 Observations

Observations are generally defined as the recording of measurements and related meta-data. 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 typically decadal intervals the individual surveys are usually carried out with different sensors and platforms, by different operators and analysts, and using different software packages and interpretation approaches. For later reanalysing the observation series, it is important that the related meta-data are stored and made available with the observational results.

3.2 Homogenization

Homogenization is defined as the procedure to correct measurement time series for artefacts and biases that are not natural variations of the signal itself but originate from changes in observational or analytical practice (Cogley et al., 2011). The aim of this step is to use available data and meta-data to detect and reduce inhomogeneities so that the observation series are internally consistent. Both the glaciological and the geodetic data series need to be homogenized independently.

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. The latter issue of changing glacier area (and elevation) over time is an inhomogeneity common to all mass balance series. The following approach provides conventional balances by adjusting the surface area and recalculating the specific balance for each elevation band of the glacier.

Assuming a linear area change over a period of record covering N years, the area S of an elevation band e is calculated for each year t as

$$S_{e.t} = S_{e.0} + \frac{t}{N} \cdot (S_{e.N} - S_{e.0}), \tag{7}$$

with elevation bin areas $S_{e.0}$ and $S_{e.N}$ from the first and the second geodetic survey, respectively. The time t is zero in the year of the first survey.

The conventional balance for the entire glacier is now regularly computed as the area-weighted sum of all (E)

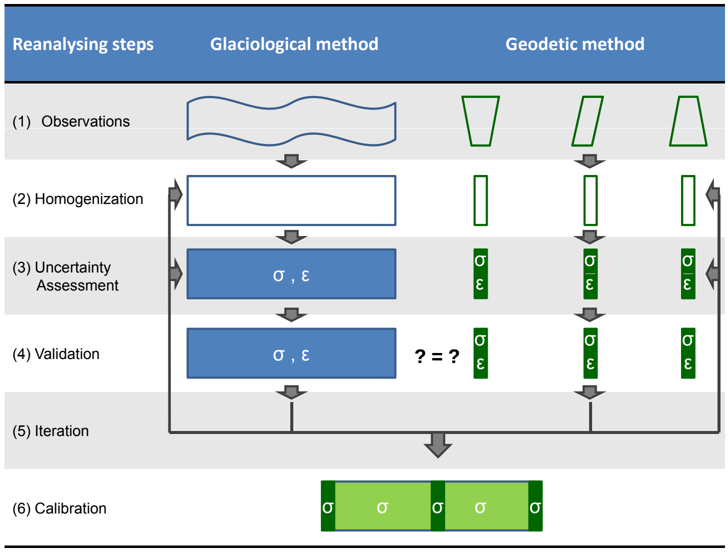

Fig. 1. Generic scheme for reanalysing glacier mass balance series in six steps, as described in Sects. 3.1–3.6. A series of N annual glaciological observations and three multi-annual geodetic surveys are independently homogenized and assessed for systematic (ε) and random (σ) errors. Resulting glaciological balances are validated and calibrated (if necessary) against geodetic balances in order to reduce unexplained differences identified as significant according to common confidence levels.

elevation bands:

$$B_{\text{glac}.t} = \frac{\sum_{e=1}^{E} B_{\text{glac}.e.t}\, S_{e.t}}{S_t}. \tag{8}$$

For glaciers with strongly non-linear area changes, the (normalized) front variation series might be used to weight the interannual area changes. Complex balance gradients or large changes in surface elevation might need to be addressed by re-integrating the point observations, such as by using a distributed mass balance model (e.g. Huss et al., 2012). Note that the analysis in this paper is focussed on conventional balances. Obtaining reference-surface balances would require correcting both to the reference area and to the reference elevation. This can only be solved with a distributed mass balance model (e.g. Paul, 2010; Huss et al., 2012) and would introduce further elements of uncertainty.

For the geodetic method, the main task is to ensure that the DEMs from the different surveys are appropriately co-registered and that there is sufficient stable terrain surrounding the glacier, or other 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).

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 systematic and random errors in the homogenized glaciological and geodetic data series as well as in the generic differences between the two balances. Therefore, the uncertainties related to the lists of potential error sources above (Sects. 2.2, 2.3 and 2.4) need to be estimated and cumulated for time periods between geodetic surveys. The resulting variables can be summarized as follows.

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

$$\overline{B_{\text{glac.a}}} = \frac{1}{N} \sum_{t=1}^{N} B_{\text{glac.a}.t}. \tag{9}$$

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 Sects. 2.2 and 3.2.

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

$$\overline{\varepsilon_{\text{glac.total.a}}} = \frac{\varepsilon_{\text{glac.total.PoR}}}{N} = \frac{\sum_{t=1}^{N} (\varepsilon_{\text{glac.point}.t} + \varepsilon_{\text{glac.spatial}.t} + \varepsilon_{\text{glac.ref}.t})}{N}. \tag{10}$$

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

$$\overline{\sigma_{\text{glac.total.a}}} = \frac{\sigma_{\text{glac.total.PoR}}}{\sqrt{N}} = \frac{\sqrt{\sum_{t=1}^{N}(\sigma^2_{\text{glac.point}.t} + \sigma^2_{\text{glac.spatial}.t} + \sigma^2_{\text{glac.ref}.t})}}{\sqrt{N}}. \quad (11)$$

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

$$\overline{B_{\text{geod.a}}} = \frac{B_{\text{geod.PoR}}}{N} \quad (12)$$

together with estimates for systematic (ε) and random (σ) errors related to the combined DEM uncertainty; see Sect. 2.3. The mean annual systematic error is expressed as

$$\overline{\varepsilon_{\text{geod.total.a}}} = \frac{\varepsilon_{\text{geod.total.PoR}}}{N} = \frac{\varepsilon_{\text{geod.DEM.PoR}}}{N} \quad (13)$$

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

The corresponding mean annual random error is estimated as

$$\overline{\sigma_{\text{geod.total.a}}} = \frac{\sigma_{\text{geod.total.PoR}}}{N} = \frac{\sqrt{\sigma^2_{\text{geod.DEM.PoR}}}}{N} = \frac{\sqrt{\sigma^2_{\text{coreg}+}\sigma^2_{\text{autocorr}}}}{N} \quad (14)$$

and integrates uncertainties related to the remaining elevation error after co-registration (σ_{coreg}) and to the spatial autocorrelation in the elevation differences (σ_{autocorr}) as root sum of squares. Note that, for scaling random errors at the annual time step, the division is by the number of years (not by the square root of N) as a unit conversion because the uncertainty over the period of record originates from the two geodetic surveys and is independent from the number of years in between. In cases where the geodetic survey only partly covers the glacier, special measures need to be taken to determine the best extrapolation procedure and to quantify related additional uncertainties.

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 assigned to the geodetic balance. Deducting internal and basal balance estimates, if they are known, from the geodetic balance results in comparing surface balances (Sect. 3.4) and ensures maintaining surface balances in case of a later calibration (Sect. 3.5). The resulting corrected balances and their random errors are expressed as

$$\overline{B_{\text{glac.corr.a}}} = \overline{B_{\text{glac.a}}} + \overline{\varepsilon_{\text{glac.total.a}}} \quad (15)$$

with

$$\overline{\sigma_{\text{glac.corr.a}}} = \overline{\sigma_{\text{glac.total.a}}}, \quad (16)$$

and

$$\overline{B_{\text{geod.corr.a}}} = \overline{B_{\text{geod.a}}} + \overline{\varepsilon_{\text{geod.total.a}}} + \overline{\varepsilon_{\text{sd.a}}} - \overline{B_{\text{int.a}}} - \overline{B_{\text{bas.a}}} \quad (17)$$

with

$$\overline{\sigma_{\text{geod.corr.a}}} = \sqrt{\overline{\sigma_{\text{geod.total.a}}}^2 + \overline{\sigma_{\text{dc.a}}}^2 + \overline{\sigma_{\text{sd.a}}}^2 + \overline{\sigma_{\text{int.a}}}^2 + \overline{\sigma_{\text{bas.a}}}^2}. \quad (18)$$

Mean annual values as calculated above allow for direct comparison of balances and error estimates between different glaciers and time periods. For a comparison of the two methods as discussed below, cumulated values over common balance periods are more convenient.

3.4 Validation

Validation can be defined as the comparison of a data series with independent observations (cf. Rykiel, 1996). 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 geodetic balance (cf. Eq. 19). The first check is to discern whether the discrepancy between the two methods can be explained by their natural dispersion: if the random uncertainties of the two methods are large enough, 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 levels), the following statistical test supports decisions concerning whether to accept the null-hypothesis H_0: the cumulative glaciological balance is not statistically different from the geodetic balance. We define the discrepancy Δ over the period of record PoR as the difference between the cumulative glaciological and the geodetic balances, both corrected for identified systematic errors and generic differences:

$$\Delta_{\text{PoR}} = B_{\text{glac.corr.PoR}} - B_{\text{geod.corr.PoR}} \quad (19)$$

The common variance of the two methods is defined as the sum of both random uncertainties, cumulated over the balance period, following the law of error propagation assuming that they are uncorrelated:

$$\sigma_{\text{common.PoR}} = \sqrt{\sigma^2_{\text{glac.corr.PoR}} + \sigma^2_{\text{geod.corr.PoR}}}, \quad (20)$$

and represents the total dispersion of the data.

Finally, we can define the reduced discrepancy

$$\delta = \frac{\Delta_{\text{PoR}}}{\sigma_{\text{common.PoR}}}. \quad (21)$$

The more consistent the two methods, the closer δ is to zero. The common variance (σ in Eq. 20) is considered to be perfectly estimated (i.e. with an infinite degree of freedom) because most estimates of the measurement uncertainties result from physical approaches. The measurement difference (Δ_{PoR}) is therefore expected to follow a normal law with a variance $\sigma_{\mathrm{common.PoR}}$. Acceptance of H_0 is then tested whether the reduced discrepancy δ follows a centred Gaussion of unit variance (Fig. 4). Working with a 95 % confidence level (i.e. the so-called $1.96 \times \sigma$ confidence interval which corresponds to the often used 2×sigma error), we can accept the hypothesis H_0 (i.e. $\Delta_{\mathrm{PoR}} = 0$) if $-1.96 < \delta < 1.96$. Under this condition, there is a probability of $\alpha = 5\,\%$ of making a wrong decision and rejecting H_0 although the results of the two methods are actually equal (i.e. error of type I, false alarm). Alternatively using a 90 % confidence level, we can accept H_0 if $-1.64 < \delta < 1.64$ with a probability for an error of type I of $\alpha = 10\,\%$. This means that mistaken rejection of H_0 is twice as likely and more series qualify for calibration.

In search of potential systematic errors in the observations, the substantial power of the statistical test is given by the ability to reject H_0 when it is actually false and a significant difference ε really exists. If the test outcome is to accept H_0 in that case, an error of type II 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

$$\beta = F\left(u_\alpha - \frac{\varepsilon}{\sigma_{\mathrm{common.PoR}}}\right) - F\left(-u_\alpha - \frac{\varepsilon}{\sigma_{\mathrm{common.PoR}}}\right), \qquad (22)$$

where F denotes the cumulative distribution function of the standard (zero-mean, unit-variance) normal distribution, and u_α is such that $F(u_\alpha) = \alpha$. For type-I risks α of 5 % and 10 %, u_α equals 1.96 and 1.64, respectively. Under higher type-I risk α (more series being flagged for calibration), the risk β of maintaining an incorrect glaciological series (not to 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 $\varepsilon_{\mathrm{limit}}$ which is detectable. This detection limit can be calculated as

$$\varepsilon_{\mathrm{limit.PoR}} = \left(u_{1-\alpha/2} + u_{1-\beta}\right)\sqrt{\sigma^2_{\mathrm{glac.PoR}} + \sigma^2_{\mathrm{geod.PoR}}}, \qquad (23)$$

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 2.9, so that the detectable error is a little less than 3 times the common variance.

Adapting Eq. (23) for annual values of random errors for the glaciological balances (cf. Eq. 11) indicates how the threshold of difference detection is lowered for longer time series. Hence, Eq. (23) becomes

$$\varepsilon_{\mathrm{limit.PoR}} = \left(u_{1-\alpha/2} + u_{1-\beta}\right)\sqrt{N\sigma^2_{\mathrm{glac.a}} + \sigma^2_{\mathrm{geod.PoR}}} \qquad (24)$$

so that the detectable annual difference $\varepsilon_{\mathrm{limit.a}}$ is given by

$$\varepsilon_{\mathrm{limit.a}} = \frac{\varepsilon_{\mathrm{limit.PoR}}}{N} = \left(u_{1-\alpha/2} + u_{1-\beta}\right)\sqrt{\frac{\sigma^2_{\mathrm{glac.a}}}{N} + \frac{\sigma^2_{\mathrm{geod.PoR}}}{N^2}}. \qquad (25)$$

Since the uncertainty over the period of record in the geodetic balance and the annual uncertainty in the glaciological balance do not depend on N, the detectable systematic error is lowered as the period of record increases, and it decreases as $1/\sqrt{N}$ for long time series (see Fig. 5) as $1/N \gg 1/N^2$ when N tends to large values under the square root of Eq. (25).

Calculation examples for Eqs. (19)–(25) are given in Supplement C.

3.5 Iteration

Once a systematic difference between the two methods is detected with high confidence, a first step is to locate the corresponding error source by going back to the homogenization process and/or the uncertainty assessment. The statistical exercise above thus helps to identify the survey period with the greatest discrepancies. Re-evaluating the available meta-data 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. 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 discrepancies require calibration and further research.

3.6 Calibration

Calibration can be defined as the adjustment of a data series to independent observations (cf. Rykiel, 1996). If a significant difference 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 because of the annual observation intervals and the spatial integration issue – with the geodetic results. The aim of the calibration is to maintain the relative seasonal/annual variability of the glaciological method while adjusting to the absolute (multi-annual) 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 invoking 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 difference, the divergence from the geodetic balance is corrected in a first step by calibrating the annual glaciological balances as follows.

Over a balance period 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. 15).

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

$$\beta_t = B_{\text{glac.corr.a.}t} - \overline{B_{\text{glac.corr.a}}}. \tag{26}$$

Over the entire balance period it results that

$$\sum\nolimits_{t=1}^{N} \beta_t = 0. \tag{27}$$

Over the same balance period, the mean annual geodetic balance $\overline{B_{\text{geod.corr.a}}}$ is calculated (see Eq. 17).

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

$$B_{\text{cal.}t} = \beta_t + \overline{B_{\text{geod.corr.a}}}, \tag{28}$$

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

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

$$B_{\text{cal.}n} = \sum\nolimits_{t=1}^{n} B_{\text{cal.}t} = n \cdot \overline{B_{\text{geod.corr.a}}} + \sum\nolimits_{t=1}^{n} \beta_t. \tag{29}$$

For the last year of the balance period ($n = N$), the cumulative calibrated balance equals the product of N times the corrected annual geodetic balance because the last term of Eq. (29) sums to 0 due to Eq. (27).

In a second step, the seasonal balances are calibrated. Unless there is a clear hint to a bias in the spring observations, the winter balance B_{w} remains untouched as it is usually independent from the annual survey

$$B_{\text{cal.w}} = B_{\text{glac.w}}, \tag{30}$$

and the difference in the annual balance B_{a} is fully assigned to the summer balance B_{s} as

$$B_{\text{cal.s}} = B_{\text{cal.a}} - B_{\text{cal.w}}. \tag{31}$$

Note that this does not imply that the summer balance is more prone to systematic errors than the winter balance. The proposed approach attributes the difference by default to the annual observations and leaves the winter balance untouched. The summer balance, in most cases, is not directly measured but calculated from annual and spring observations.

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

$$\beta_{\text{e.}t} = B_{\text{glac.}e.t} - B_{\text{glac.corr.a.}t}. \tag{32}$$

Then, the calibrated elevation band balance is defined as

$$B_{\text{cal.}e.t} = \beta_{e.t} + B_{\text{cal.}t}. \tag{33}$$

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 bands, i.e. by fitting a curve to the calibrated surface balance data as a function of altitude (cf. Cogley et al., 2011). Note that this approach does not require changing directly observed (end-of-summer) snowlines, which are often used as (annual) equilibrium line for mass balance calculations at glaciers where all mass exchange is expected to occur at the glacier surface and with no superimposed ice. In fact, deviations of the calibrated ELA (and corresponding AAR) from the spatially averaged altitude of the observed snowline (and the topographic AAR) might help identifying remaining error sources in the glaciological method.

Reanalysed mass balance series and derived parameters need to be flagged accordingly in any databases in which they are stored. This can be done by linking both glaciological and geodetic mass balance series through a lookup table including information on the reanalysing status (e.g. not reanalysed, homogenized only, validated but no calibration needed, validated and calibrated) and providing reference to related publications.

The calibration of glaciological mass balance series implies a difference 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 and snowline observations but fits the glacier-wide results to the geodetic balance (see Sect. 5.2). This allows for reproducibility, later reanalysing exercises when new information about potential error sources or a new geodetic DEM becomes available, and/or application of statistical treatments (e.g. Lliboutry's variance analysis model).

4 Selected glaciers with long-term observation programmes

The following analysis in Sect. 5 is based on selected glaciers with long-term measurements including both glaciological and geodetic surveys and with available information for estimating related uncertainties. In general, the reanalysing 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 balance periods used is provided in Table 1. The analysed dataset consists of a total of 46 balance periods from 12 glaciers, including

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

PU	Name	LAT	LON	glac. surveys (first/last/# obs. years)	geodetic surveys (year $_{\mathrm{method}}$)	literature
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}, $\mathbf{-98}_{\mathbf{aP}}$, $\mathbf{2009}_{\mathbf{aL}}$	Hynek et al. (2012)
AT	Hintereisferner (HEF)	46.80 N	10.77 E	1953/2012/60	1893_{tP}, $\mathbf{1953}_{\mathbf{tP}}$, $\mathbf{-64}_{\mathbf{tP}}$, $\mathbf{-67}_{\mathbf{tP}}$, $\mathbf{-69}_{\mathbf{aP}}$, $\mathbf{-79}_{\mathbf{aP}}$, $\mathbf{-91}_{\mathbf{aP}}$, $\mathbf{-97}_{\mathbf{aP}}$, $\mathbf{2006}_{\mathbf{aL}}$	Fischer and Markl (2009), Fischer (2010)
AT	Jamtalferner (JAM)	46.87 N	10.17 E	1989/2012/24	1969_{aP}, $\mathbf{-96}_{\mathbf{aP}}$, -2002_{aP}, $\mathbf{-06}_{\mathbf{aL}}$	Fischer and Markl (2009), Fischer (2010)
AT	Kesselwandferner (KWF)	46.84 N	10.79 E	1953/2012/60	1969_{aP}, $\mathbf{-71}_{\mathbf{aP}}$, $\mathbf{-97}_{\mathbf{aP}}$, $\mathbf{2006}_{\mathbf{aL}}$	Fischer and Markl (2009), Fischer (2010)
AT	Kleinfleisskees (FLK)	47.05 N	12.95 E	1999/2012/14	1931_{tP}, -53_{aP}, -69_{aP}, -79_{aP}, -92_{aP}, $\mathbf{-98}_{\mathbf{aP}}$, $\mathbf{2009}_{\mathbf{aL}}$	Hynek et al. (2012)
AT	Wurtenkees lower part (WLP) upper part (WUP)	47.04 N	13.00 E	1983/2012/30	1969_{aP}, -79_{aP}, -91_{tP}, $\mathbf{-98}_{\mathbf{aP}}$, $\mathbf{2006}_{\mathbf{aL}}$	Hynek et al. (2012)
AT	Vernagtferner (VER)	46.87 N	10.82 E	1965/2012/48	1889_{tP}, 1912_{tP}, -38_{tP}, $\mathbf{-69}_{\mathbf{aP}}$, $\mathbf{-79}_{\mathbf{aP}}$, $\mathbf{-90}_{\mathbf{aP}}$, $\mathbf{-99}_{\mathbf{aP}}$, $\mathbf{2006}_{\mathbf{aL}}$, $\mathbf{-09}_{\mathbf{aL}}$	Moser et al. (1986), Reinwarth and Rentsch (1994), Reinwarth and Escher-Vetter (1999)
CH	Griesgletscher (GRS)	46.44 N	8.33 E	1962/2012/51	$\mathbf{1961}_{\mathbf{aP}}$, $\mathbf{-67}_{\mathbf{aP}}$, $\mathbf{-79}_{\mathbf{aP}}$, $\mathbf{-86}_{\mathbf{aP}}$, $\mathbf{-91}_{\mathbf{aP}}$, $\mathbf{-98}_{\mathbf{aP}}$, $\mathbf{2003}_{\mathbf{aP}}$, $\mathbf{-07}_{\mathbf{aP}}$	Huss et al. (2009)
CH	Silvrettagletscher (SIL)	46.85 N	10.08 E	1960/2012/53	$\mathbf{1959}_{\mathbf{aP}}$, $\mathbf{-73}_{\mathbf{aP}}$, $\mathbf{-86}_{\mathbf{aP}}$, $\mathbf{-94}_{\mathbf{aP}}$, $\mathbf{2003}_{\mathbf{aP}}$, $\mathbf{-07}_{\mathbf{aP}}$	Huss et al. (2009)
FR	Sarennes (SAR)	45.07 N	6.07 E	1949/2012/64	1908_{tP}, $\mathbf{-52}_{\mathbf{aP}}$, $\mathbf{-81}_{\mathbf{aP}}$, $\mathbf{2003}_{\mathbf{aP}}$	Eckert et al. (2011), Thibert et al. (2008)
NO	Engabreen (ENG)	66.40 N	13.50 E	1970*/2012/43	$\mathbf{1968}_{\mathbf{aP}}$, $\mathbf{2001}_{\mathbf{aL}}$, $\mathbf{-08}_{\mathbf{aL}}$	Geist et al. (2005), Elvehøy et al. (2009), Haug et al. (2009), Kjøllmoen et al. (2011, and earlier issues)
SE	Storglaciären (STO)	67.90 N	18.57 E	1946/2012/67	1910_{tP}, -49_{aP}, $\mathbf{-59}_{\mathbf{aP}}$, $\mathbf{-69}_{\mathbf{aP}}$, $\mathbf{-80}_{\mathbf{aP}}$, $\mathbf{-90}_{\mathbf{aP}}$, $\mathbf{-99}_{\mathbf{aP}}$, 2008_{aP}	Holmlund (1996), Albrecht et al. (2000), Holmlund et al. (2005), Koblet et al. (2010), Zemp et al. (2010)

* 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.

38 multi-annual periods with an average time span of 11 yr (ranging from 4 to 32 yr), 8 overall periods for glaciers with more than one balance period, and additional 9 balance periods with alternative calculations for the example glacier Storglaciären (cf. Sect. 5.1). Lower and upper parts of Goldbergkees and Wurtenkees are analysed separately due to the disintegration of the glacier before the analysed balance periods (1998–2009 and 1998–2006). Details about glaciological and geodetic surveys and related uncertainty assessments are found within the publications listed in Table 1. 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 Reanalysing glacier mass balance series: the example of Storglaciären

Glaciological mass balance measurements on Storglaciären have been carried out without interruption since 1945/1946 together with aerial surveys at approximately decadal intervals (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 has been challenged by inaccuracies in the maps and methodologies, which 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) reanalysed 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 reanalysing 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_0) are shown in Table 2, as explained in Sect. 3.4. For each balance period, the cumulative glaciological balance is corrected for systematic errors as well as for generic differences from 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 that allows statistical quantification of whether 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 1969–1980 and 1980–1990 both have reduced discrepancies far beyond the 90 % and 95 % confidence levels and, hence, show that the glaciological are significantly different from the geodetic balances (i.e. H_0 to be rejected). Interestingly, there is no such discrepancy for the overall balance period (1959–1990). This is because the two strongly erroneous decades have cumulated discrepancies of opposite signs most probably caused by errors in the map of the 1980 survey. This nicely demonstrates the importance of testing both the entire balance period and the individual (multi-annual) intervals. In such a case, there are two options: identify the error source in another iteration of the reanalysing process or calibrate the glaciological balance over the two balance periods with significant differences to the geodetic balances.

Koblet et al. (2010) chose the first option and homogenized all DEMs. Comparing these new geodetic results with the glaciological findings shows the improvements in the periods 1969–1980 and 1980–1990 with much smaller cumulated discrepancies. H_0 is now clearly accepted for both periods. However, the additional balance period (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 period of record (1959–1999) shows a large cumulated discrepancy of more than 3 m w.e. As a consequence, H_0 is to be accepted at the 95 % but to be rejected at the 90 % confidence level. After checking all assumptions of the uncertainty assessment, the reason is most probably to be found in the above-mentioned over-estimation 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) 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 1997/1998 and 1998/1999. 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 years – seem to exaggerate the contribution of the internal accumulation to the annual balance.

Table 2. Summary statistics for the comparison of glaciological and geodetic balances of Storglaciären, Sweden. For different dataset combinations, the table shows analysed 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. (19)–(21). The acceptance (rejection) of the hypothesis (H_0: the two balances are equal) is evaluated on the 95 % and 90 % confidence level (i.e. a type-I risk), which corresponds to reduced discrepancies inside (outside) the ±1.96 and ±1.64 range, respectively. For the same confidence levels, the type-II risk β (cf. Eq. 22) of not detecting an erroneous series is also given.

Dataset combination	PoR	B.glac.corr $\pm\sigma$	B.geod.corr $\pm\sigma^*$	Δ_{PoR}	$\sigma_{\mathrm{common.PoR}}$	δ	H_0: 95/90	β: 95/90
	Year	m w.e.	m w.e.	m w.e.	m w.e.	no unit	no unit	%
B.glac versus B.geod.old, incl. intACC	1959–1969	-3.06 ± 1.18	-2.13 ± 0.45	-0.93	1.26	-0.74	yes/yes	89/81
	1969–1980	-2.41 ± 1.09	-6.22 ± 0.90	3.81	1.42	2.69	no/no	23/15
	1980–1990	0.99 ± 0.59	3.72 ± 0.87	-2.73	1.05	-2.60	no/no	26/17
	1959–1990	-4.46 ± 1.72	-4.62 ± 1.33	0.16	2.84	0.05	yes/yes	95/90
B.glac versus B.geod.new, incl. intACC	1959–1969	-3.06 ± 1.18	-4.04 ± 0.45	0.98	1.26	0.78	yes/yes	88/80
	1969–1980	-2.41 ± 1.09	-2.79 ± 0.90	0.39	1.41	0.27	yes/yes	94/89
	1980–1990	0.99 ± 0.59	0.11 ± 0.87	0.88	1.05	0.84	yes/yes	87/78
	1990–1999	0.72 ± 0.43	-0.35 ± 0.24	1.07	0.49	2.19	no/no	41/29
	1959–1999	-3.76 ± 1.71	-7.00 ± 0.49	3.24	1.79	1.81	yes/no	56/43
B.glac versus B.geod.new, excl. intACC	1959–1969	-3.06 ± 1.19	-3.56 ± 0.45	0.50	1.26	0.40	yes/yes	93/87
	1969–1980	-2.41 ± 1.09	-2.16 ± 0.90	-0.25	1.42	-0.18	yes/yes	95/89
	1980–1990	0.99 ± 0.59	0.76 ± 0.87	0.23	1.05	0.22	yes/yes	94/89
	1990–1999	0.72 ± 0.43	0.23 ± 0.23	0.49	0.49	0.99	yes/yes	83/74
	1959–1999	-3.76 ± 1.71	-4.64 ± 0.52	0.88	1.79	0.49	yes/yes	92/86

*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.

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 with smallest cumulated and reduced discrepancies, and clear acceptances of H_0 for all periods. As a consequence, no calibration of the glaciological balance is needed over the reanalysed period (1959–1999). However, in spite of the relatively small discrepancies between the two methods (< 0.10 m w.e. a^{-1}) there still is a great risk of not detecting a remaining difference. 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 Silvrettagletscher

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

For the two balance periods, the differences between glaciological mass balance and the geodetic surveys are considerable (Fig. 2a). Whereas the cumulative glaciological balance 1994–2007 is -3.09 m w.e., the geodetic mass balance indicates a cumulative balance of -7.94 m w.e. over the same period. According to the statistical test (Sect. 3.4) this difference is significant at the 95 % level and H_0 is rejected. Since the related error source could not be clearly identified and corrected, the series for the two balance periods 1994–2003 and 2003–2007 thus need to be calibrated.

First, the centred glaciological balance β_t is calculated as the deviation from the period mean $\overline{B_{\mathrm{glac.corr.a}}}$ (see Eq. 26), and β_t is subsequently shifted to agree with the mean annual geodetic mass balance $\overline{B_{\mathrm{geod.corr.a}}}$ (see Eq. 28). This results in a calibrated series that represents a conventional mass balance covering all components of the surface balance. The long-term 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 difference between the original glaciological and the geodetic balance is distributed equally over all balance years between two geodetic surveys.

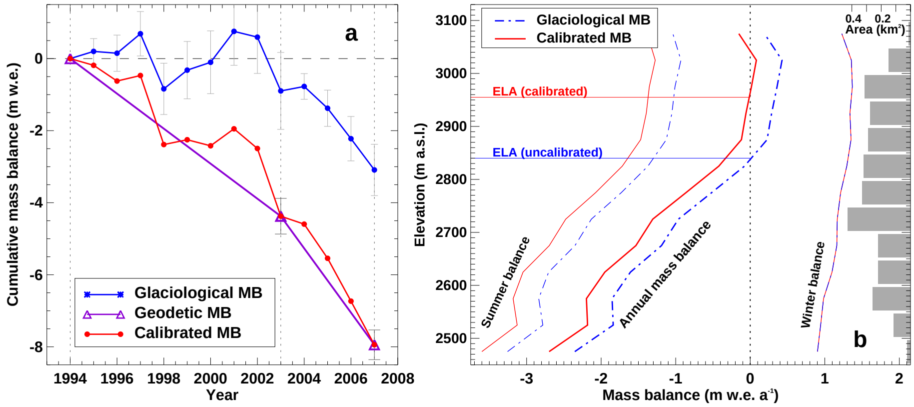

Fig. 2. Calibration of glaciological mass balance series for the periods 1994–2003 and 2003–2007 with the geodetic surveys for Silvrettagletscher (cf. Huss et al. 2009). **(a)** Cumulative mass balance (original glaciological mass balance and calibrated with the geodetic mass change). Uncertainties according to the uncertainty analysis are given. **(b)** Mass balance elevation distribution (original and calibrated glaciological series) as a mean over the period 2003–2007. Both seasonal and annual mass balances are shown, the original and calibrated ELA is indicated, and glacier hypsometry is given.

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. 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 assigned to the summer balance (Fig. 2b). The mass balance elevation distribution remains similar, but is shifted for each year according to the mean annual difference (see Eqs. 32 and 33). 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 conceptual framework described above for reanalysing mass balance series strongly builds on the experience from glaciers with detailed and long-term mass balance monitoring programmes. Here we analyse glaciological and geodetic balances with related uncertainties from roughly 50 balance periods with data available from the 12 glaciers (for name abbreviations see Table 1). All reported values and statistics are given in Supplement C. For summary statistics, only independent balance periods are analysed, omitting the overall balance period for glaciers with more than one period. 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 periods of records are negative with a mean of $-0.45\,\mathrm{m\,w.e.\,a^{-1}}$ and a corresponding random error of $0.34\,\mathrm{m\,w.e.\,a^{-1}}$. The related uncertainty of the field measurements at point locations is estimated to be $0.14\,\mathrm{m\,w.e.\,a^{-1}}$. 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): 0.10–0.34, Lliboutry (1974): 0.30, Cogley and Adams (1998): 0.20, Gerbeaux et al. (2005): 0.10 for ice ablation, 0.25–0.40 for firn ablation, Vallon and Leiva (1981) for drilling in accumulation area: 0.30 (all values in $\mathrm{m\,w.e.\,a^{-1}}$). The spatial integration of the glaciological point measurements has an estimated uncertainty of $0.28\,\mathrm{m\,w.e.\,a^{-1}}$ which are attributed to the local representativeness, the interpolation method, and the extrapolation to unmeasured areas. Estimates for these three uncertainties range between 0.10 and $0.50\,\mathrm{m\,w.e.\,a^{-1}}$. Thereby, 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; $0.07\,\mathrm{m\,w.e.\,a^{-1}}$) or by Fountain and Vecchia (1999). Uncertainties due to reference area changing over time have a root mean square of $0.01\,\mathrm{m\,w.e.\,a^{-1}}$ and a corresponding uncertainty of $0.10\,\mathrm{m\,w.e.\,a^{-1}}$.

The average geodetic balance of the investigated periods of records (i.e. $-0.58\,\mathrm{m\,w.e.\,a^{-1}}$) is slightly more negative than the glaciological result. The remaining elevation bias (from balance periods without DEM co-registration) is estimated on average to be $< 0.01\,\mathrm{m\,w.e.\,a^{-1}}$ with an uncertainty of $0.01\,\mathrm{m\,w.e.\,a^{-1}}$. The spatial correlations of the elevation differences are not specified. The uncertainty related to the density conversion is estimated to be $0.03\,\mathrm{m\,w.e.\,a^{-1}}$. Differences in time system and reference areas are quantified with a root mean square bias of 0.03 m w.e. a^{-1} and a corresponding uncertainty of 0.02 m w.e. a^{-1}.

Differences due to internal and basal balances are usually not specified or assumed to be zero. The few estimates include about 0.01 m w.e. a^{-1} for internal ablation (STO and SAR), between 0.01 and 0.10 m w.e. a^{-1} for internal accumulation (STO, SAR, VER) and < 0.01 m w.e. a^{-1} for basal ablation (STO, GRS, SIL). Note that the sample of estimates 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. Estimates for internal and basal ablation are similar to the ones found at Gulkana Glacier by March and Trabant (1997; < 0.01 m w.e. a^{-1} due to both geothermal heat flow and ice motion as well as 0.06 m 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 % contribution from basal melt to the total ablation of Franz Josef Glacier, which they explain by the strongly maritime environment of the glacier. Some of the proposed corrections for internal accumulation seem to be very large: Trabant and Mayo (1985) for Alaskan glaciers report 7–64 %, Trabant and Benson (1986) for McCall report 40 %, Schneider and Jansson (2004) for STO report 3–5 %, and Reijmer and Hock (2008) for STO report 20 % (all relative to annual accumulation).

On average, overall uncertainties are 0.34 and 0.07 m w.e. a^{-1} for (corrected) glaciological and geodetic balances, respectively. Individual balance periods with values greater than 0.50 m w.e. a^{-1} are found for HEF (glac), KWF (glac) and GRS (glac). Absolute values for bias corrections are mostly below 0.05 m w.e. a^{-1}. Note that all analysed series were at least partly homogenized, with the aim of 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 $+0.12$ m w.e. a^{-1} with a root mean square of 0.23 m w.e. a^{-1}. This tendency for 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 uncertain-

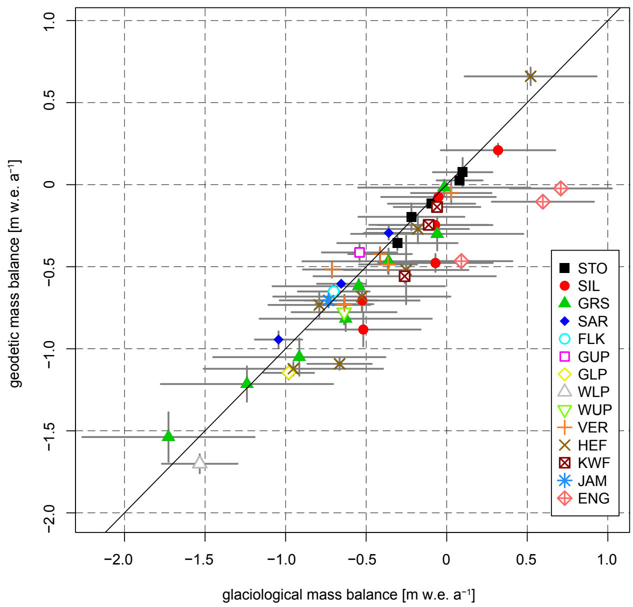

Fig. 3. Glaciological versus geodetic balances. Both series are corrected for biases and generic differences and plotted with random uncertainties. The black diagonal line marks equal balances from both methods.

ties. The few exceptions are two balance periods of HEF (1964–1969, 1979–1991), one period of KWF (1997–2006), two periods of SIL (1994–2003, 2003–2007), and all three periods 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 balance periods from 29 glaciers but without an individual glacier uncertainty assessment. He found a (statistically also not significant) negative mean annual discrepancy of -0.07 m w.e. a^{-1} and a root mean square of 0.38 m 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_0) 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 balance periods are plotted in Fig. 4. Working with a 95 % confidence level, H_0 is accepted for 37 out of all 46 balance periods and 9 ($1 \times$ HEF, $1 \times$ KWF, $2 \times$ SIL, GLP, $1 \times$ VER, $3 \times$ ENG) are candidates for a calibration. Setting the confidence level to 90 % increases the number of calibration candidates to 14 (adding $1 \times$ SIL, $2 \times$ SAR, WLP, $1 \times$ HEF). 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 of the agreement of the

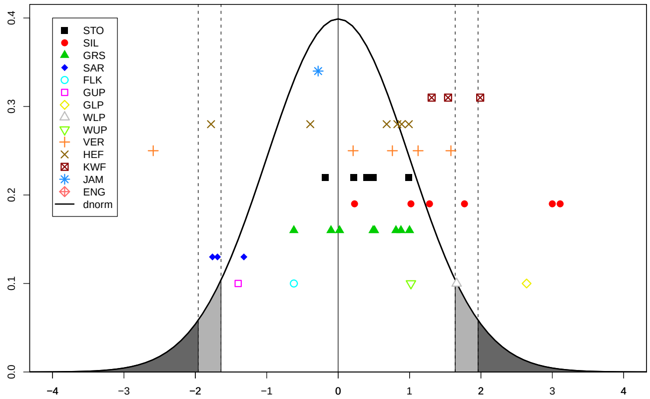

Fig. 4. Reduced discrepancies between all of the analysed periods of record. The reduced discrepancy (horizontal axis) has no unit and the values for the different glaciers are arbitrarily distributed along the vertical axis for a better overview. The curve labelled 'dnorm' denotes the probability density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

results from the two methods. Checking for large common variances helps identify balance periods where the ability to detect a systematic difference between the methods is weak. In the case of STO, the balance period (1990–1999) with a reduced variance close to 1.0 is probably more reliable than the three other approximately decadal balance periods with reduced variances between -0.2 and $+0.4$; their common variances, and hence their lowest detectable differences, are two or more times the value of the period 1990–1999 (with a lowest detectable difference of about 0.16 m w.e. a^{-1}). Other balance periods with accepted H_0 but weak ability for detecting systematic differences (i.e. lowest detectable difference > 0.50 m w.e. a^{-1}) are found for SIL (2003–2007), GRS (all decadal periods), HEF (1953–1964, 1964–1969, 1997–2006), and KWF (1997–2006). In the case of KWF, the difference is significant only for the entire period of record but not for the two individual balance periods.

Working with a 90 % confidence level, the mean of all lowest detectable differences decreases from 0.50 to 0.40 m w.e. a^{-1}. At the same time, lowering the confidence level 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 C, around 70 % on 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 (64 %), running a higher risk α of 10 % in recalibrating (uselessly) a correct series. For example, for GRS (1967–1979) 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 in the case of Silvretta (2003–2007), with β decreasing from 58 to 37 %. 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 level 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 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 difference decreases with the length of the period of record. 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. Over the course of the balance period, 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 the random-error sum by enough to become detectable with a useful confidence level (90 % in our calculations). Thus long periods are required to detect differences among the natural scatter of the observations. 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 reanalysing mass balance measurements, we make thirteen general statements as guidelines for the benefit of data producers and users.

Recommendations for investigators of glacier mass balance are as follows: (i) Glaciological mass balance programmes, based on a minimal network of long-term ablation and accumulation point measurements, should increase the observational network about every decade in order to reassess the spatial pattern of mass balance. (ii) Glaciological observations are ideally complemented from the very

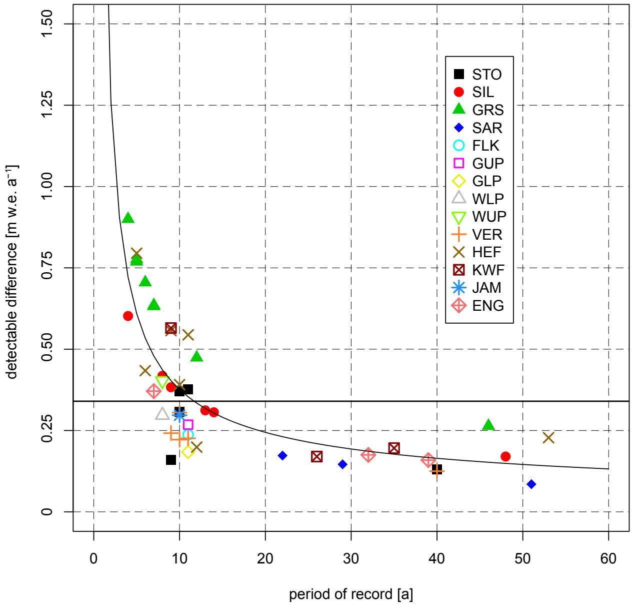

Fig. 5. Detectable difference as a function of the period of record. The horizontal line marks the mean random error of all glaciological balance series ($\overline{\sigma_{\text{glac.a}}} = 0.34$ m w.e. a^{-1}). The curved line marks the minimum detectable annual difference (at 10 % error risk) as a function of the number of years of the series as given by Eq. (25) 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 difference to become lower than the annual random "noise" of the glaciological balance represented by $\overline{\sigma_{\text{glac.a}}}$.

beginning with geodetic surveys at about decadal intervals. (iii) 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). (iv) As a rule of thumb, an absolute difference between the glaciological and geodetic balances that exceeds the annual random error estimated for the glaciological balance (e.g. > 0.30 m w.e. a^{-1} for our sample) indicates that reanalysing is urgently needed. (v) Mass balance series longer than 20 yr should be reanalysed in any case. (vi) Every mass balance series should be clearly flagged in publications and databases with its reanalysing status. (vii) 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.

There are implications for users of glacier mass balance data: (viii) the glaciological method measures the components of the surface balance. (ix) The geodetic method measures all components of the surface, internal, and basal balances. (x) 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). (xi) 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. (xii) 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 more than 0.50 m w.e. a^{-1}. (xiii) Reanalysing mass balance series, especially of long series, based on both methods allows the quantification of the related uncertainties and of remaining unexplained differences.

Finally, identification of a need to calibrate a glaciological mass balance series (as explained in Sect. 3.6) implies 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 reanalysing exercises.

6 Conclusions

Based on the experience from long-term monitoring programmes and a series of workshops, this paper briefly summarizes the glaciological and geodetic method and proposes a conceptual framework for reanalysing glacier mass balance series.

The glaciological method measures the surface mass balance components and is subject to error classes related to field measurements at point location, spatial integration over the entire glacier, and changes in glacier area and elevation over time. 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. The comparison of glaciological and geodetic balances requires accounting for survey differences in time system and reference areas, for internal and basal balance components, and for errors related to the density conversion of the geodetic balance.

Reanalysing glacier mass balance series includes six principal steps: (1) Observation and extrapolation provide a first estimate of glaciological and geodetic balances over common time periods. (2) Homogenization aims at identifying and removing artefacts and biases in order to achieve internally consistent observation series. (3) Uncertainty assessment provides estimates for both systematic and random errors for the glaciological and the geodetic balances as well as for the generic differences between the two methods. (4) Validation statistically tests whether the unexplained difference between the glaciological and the geodetic method is significant and provides estimates for the detectable bias. (5) Iteration aims at locating the corresponding error source by going back to the previous steps. (6) Calibration allows, in cases of unexplained discrepancies, for adjusting the (annual) glaciological to the (multi-annual) geodetic balances.

Our analysis of a dozen European glaciers with a total of 50 periods of both geodetic and glaciological balances shows that both methods are subject to errors from various sources. Systematic errors have typical values below one hundred mm w.e. per year but cumulate with the length of the observation record for glaciological series. The cumulative bias in the derived glacier mass balance presents a challenge to the efforts to detect climatic trends. Random errors have typical values of a few hundred mm w.e. per year and cumulate according to the law of error propagation. Differences between glaciological and geodetic balances are therefore easier to detect for longer time spans of a decade or more. The proposed calibration of the balance series maintains the interannual variation of the original glaciological data at the same time as cumulative values become consistent with the multi-annual geodetic balances and provides an optimal estimate of the surface balance until the source(s) of bias can be identified and quantified.

Reanalysing glacier mass balance series should become a standard procedure for every mass balance monitoring programme with increasing importance for long time series. Users of reanalysed datasets profit from improved data quality and uncertainty estimates, which will hopefully lead 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.net/7/1227/2013/tc-7-1227-2013-supplement.zip.

Acknowledgements. We are grateful to Wilfried Haeberli for his experienced input into both the concept and the manuscript, to Graham Cogley, Mauri Pelto, and Regine Hock for their thorough reviews and constructive comments, 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 9–11 July 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 Wenner-Gren Foundations, the International Union of Geology and Geophysics (IUGG) through the International Association of Cryospheric Sciences (IACS), and by the International Glaciological Society (IGS).

Edited by: V. Radić

References

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C., and Wieringa, J.: Guidelines on climate metadata and homogenization, World Climate Programme Data and Monitoring, 53, 50 pp., 2003.

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, 382–6, 2002.

Begert, M., Schlegel, T., and Kirchhofer, W.: Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000, Int. J. Climatol., 25, 65–80, 2005.

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.

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, Paris, UNESCO-IHP, 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.
GAA: 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.
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, 7, 877–887, doi:10.5194/tc-7-877-2013, 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., Schöner, W., Binder, D., and Weyss, G.: Comparison of direct and geodetic mass balances of three small glaciers in the Eastern Alps, Geophys. Res. Abstr., 14, EGU2012-9025, 2012.
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.
Josberg, E. G., Bidlake, W. R., March, R. S., and Kennedy, B. W.: Glacier mass-balance fluctuations in the Pacific Northwest and Alaska, USA, Ann. Glaciol., 46, 291–296, 2007.
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.
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Dennis, J.: The NCEP/NCAR 40-year Reanalysis Project, B. Am. Meteorol. Soc., 77, 437–471, 1996.
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.
Magnússon, E., Björnsson, H., Dall, J., and Pálsson, F.: Volume changes of Vatnajökull Ice Cap, Iceland, due to surface mass balance, ice flow, and subglacial melting at geothermal areas, Geophys. Res. Lett., 32, L05504, doi:10.1029/2004GL021615, 2005.
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, 96–4299, 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.

Moser, H., Escher-Vetter, H., Oerter, H., Reinwarth, and Zunke, O. 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.

Paul, F.: The influence of changes in glacier extent and surface elevation on modeled mass balance, The Cryosphere, 4, 569–581, doi:10.5194/tc-4-569-2010, 2010.

Pelto, M. S.: The impact of sampling density on glacier mass balance determination, Hydrol. Proc., 14, 3215–3225, 2000.

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.

Rykiel Jr., E. J.: Testing ecological models: the meaning of validation, Ecol. Model., 90, 229–244, 1996.

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. A, 31, 222–227, 1949.

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 Françaises), Z. Gletscherkd. Glazialgeol., 17, 143–167, 1981.

Van Beusekom, A. E., O'Neel, S. R., March, R. S., Sass, L. C., and Cox, L. H.: Re-analysis of Alaskan benchmark glacier mass-balance data using the index method, US Geological Survey Scientific Investigations Report, 5247, 16 pp., 2010.

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, Volume 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, 336 pp., publication based on database version: doi:10.5904/wgms-fog-2012-11, 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, F04016, doi:10.1029/2007JF000818, 2007.

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.