Hydrol. Earth Syst. Sci. Discuss., 7, 8661–8702, 2010 www.hydrol-earth-syst-sci-discuss.net/7/8661/2010/ doi:10.5194/hessd-7-8661-2010 © Author(s) 2010. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Simulation of high mountainous discharge: how much information do we need?

B. Schaefli^{1,2} and M. Huss^{3,4}

 ¹Laboratory of Ecohydrology, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
 ²Water Resources Section, Delft University of Technology, Delft, The Netherlands
 ³Department of Geosciences, University of Fribourg, Fribourg, Switzerland
 ⁴Laboratory of Hydraulics, Hydrology and Glaciology (VAW), ETH Zurich, Zurich, Switzerland
 Received: 11 October 2010 – Accepted: 17 October 2010 – Published: 1 November 2010
 Correspondence to: B. Schaefli (bettina.schaefli@epfl.ch)

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



Abstract

The hydrologic cycle of high mountainous catchments is frequently simulated with simple precipitation-discharge models representing the snow accumulation and ablation behavior of a very complex environment with a set of lumped equations accounting for altitudinal temperature and precipitation gradients. In this study, we present a methodology to include sparse snow depths measurements into the calibration process. Based on this methodology, we assess for a case study, the Rhonegletscher catchment (Switzerland), how much observed information we need to reliably calibrate the model, such that it reproduces the dominant system dynamics, discharge, as well as glacier mass balance. Here, we focus on the question whether observed discharge is sufficient as a calibration variable or whether we need annual or even seasonal glacier mass balance data. Introducing seasonally variable accumulation and ablation parameters is sufficient to enable the simple model to reproduce observed seasonal mass balances for the Rhonegletscher. Furthermore, our results suggest that calibrat-

- ing the hydrological model exclusively on discharge can lead to wrong representations of the intra-annual accumulation and ablation processes and to a strong bias in long term glacier mass balance simulations. Adding only a few annual mass balance observations considerably reduces this bias. Calibrating exclusively on annual balance data can, in turn, lead to wrong seasonal mass balance simulations. Even if these results are case study specific, our conclusions provide valuable new insights into the
- results are case study specific, our conclusions provide valuable new insights into the benefit of different types of observations for calibrating hydrological models in glacier catchments. The presented multi-signal calibration framework and the simple method to calibrate a semi-lumped model on point observations has potential for application in other modeling contexts.



1 Introduction

Continuous simulation of discharge has become a standard tool for water resources management at the catchment-scale, e.g. to estimate extreme events from long simulated time series (e.g., Haberlandt et al., 2008; Hingray et al., 2010) or to predict climate and land use change impacts on hydrologic regimes (e.g., Horton et al., 2006). For this type of application, conceptual models play an important role: they represent the dominant rainfall-runoff processes directly at the catchment-scale instead of resolving the small scale physical processes (see Sivapalan et al., 2003). In high mountain-

ous catchments, the focus of the present paper, such conceptual models represent the snow accumulation and ablation behavior of a very complex environment with a set of lumped equations accounting only for altitudinal temperature and precipitation gradients (e.g., Schaefli et al., 2005; Klok et al., 2001; Hock, 2003; Stahl et al., 2008).

These models are generally calibrated so that the model reproduces as closely as possible a series of observed discharge measurements. The question inevitably arises

¹⁵ whether hydrologic predictions of such a calibrated model are actually reliable. Knowing that a model performs well for historic situations does not necessarily imply that it will perform well for future catchment conditions, i.e. that it can be used for climate change or land use change impact prediction.

This well-known problem of calibrated models becomes particularly relevant for hydrologic systems that are undergoing such fast evolution as glacierized catchments (see, e.g. Moore et al., 2009). A first, although not sufficient step to answer this question is investigating whether with a calibrated model, we are "getting the right answers for the right reasons" (Kirchner, 2006). Adopting the perspective that catchments are systems that partition incoming water into different storages (soil, snow, groundwater

etc.), from which water is released at some later moment in time (see Wagener et al., 2007), this essentially translates into getting all these three functions right (partition-ing, storage and release). In glacierized catchments, this would for example mean that a precipitation-runoff model should not just mimic observed discharge but also



reproduce snow accumulation and melt dynamics or the glacier mass change (e.g., Schaefli et al., 2005; Koboltschnig et al., 2008). This is essential to be able to correctly simulate runoff dynamics under the strongly changing climatic conditions expected for the next decades.

One of the challenges is hereby the integration of very different types of data containing spatially integrated information such as discharge or glacier mass balance or point information such as snow height or ice ablation observations. We, therefore, developed a methodology that can make use of sparse point observations to calibrate and evaluate a lumped precipitation-discharge simulation model for high mountainous
 catchments.

This work was triggered by the analysis of a case study in the Swiss Alps, the Rhonegletscher catchment, for which Schaefli et al. (2005) have presented a conceptual hydrological model (GSM-SOCONT) and Huss et al. (2008) a more complex glaciological model with higher spatial resolution. Both models simulate monthly, seasonal and an-

- ¹⁵ nual glacier mass balances. However, the hydrological model GSM-SOCONT yields on average much less winter accumulation than the glaciological model and slightly more ablation during summer months. This striking difference, which would namely result in rather different predictions of glacier surface evolution, led us to the question whether a conceptual precipitation-runoff model calibrated (i) exclusively on runoff or
- (ii) on runoff and glacier-wide annual mass balance can reproduce the seasonal glacier mass balances or whether we need to incorporate some in-situ measurements of snow accumulation and ablation processes; and, if yes, how to incorporate such point observations.

Hereafter, we first present the case study (Sect. 2) before discussing in detail the method developed to merge different types of glacio-hydrological data into a precipitation-discharge model (Sect. 3). As illustrated and discussed in Sects. 4 and 5, this method provides valuable new insights into the use of multi-signal calibration for model development in general and into the question of how much information we need for discharge prediction in high Alpine environments.



2 Case study

2.1 Rhonegletscher catchment: study site and field data

The Rhonegletscher catchment at Gletsch has a size of 38.9 km² with a mean elevation of 2719 m a.s.l. The outlet is located at 1761 m a.s.l. Discharge measurements in daily resolution are available continuously since 1956.

Rhonegletscher is a medium-sized valley glacier and presently covers almost half of the catchment. Its size was 17.3 km² in 1980 and 15.9 km² in 2007 (Bauder et al., 2007). Several other small glacier and ice patches are present in the catchment (around 2.3 km²), which are included in the total ice-covered area for discharge simulations. There are no meteorological measurement stations located within the catchment. The nearest stations are Grimsel (altitude 1980 m a.s.l., latitude 46°34′18″, longitude 8°19′60″), Ulrichen (1346 m a.s.l., 8°18′29″, 46°30′17″) and Oberwald (1375 m a.s.l., 8°20′48″, 46°32′03″). The station at Oberwald only measures precipitation and stopped recording in 1999. Grimsel is the closest station (distance to nearest catchment point 1.5 km) but is located in a neighboring valley. Oberwald as well as Ulrichen lie downstream of the catchment outlet with a distance of 3.0 km and 7.4 km, respectively. The choice of the meteorological time series will be further discussed in Sect. 4.2.1.

The hypsometry of the catchment is derived from a digital elevation model with a resolution of 25 m (SwissTopo, 1995). For Rhonegletscher repeated topographical information documenting the changes in glacier surface elevation and area over the last decades are available from photogrammetry (Bauder et al., 2007). These digital terrain models were used to divide the glacier into five surface elevation bands with equal area.

A comprehensive data set of seasonal point mass balance observations is available from the direct glaciological measurements over five years, including a winter survey in April to May and a late summer field visit in September (Funk, 1985; Huss et al., 2008). Winter accumulation is measured using probings of the snow depth distributed all over



the glacier and annual mass balance is determined using stakes drilled into the ice. Field data are available for the years 1979/1980 to 1981/1982 and for 2006/2007 and 2007/2008. For the individual surveys between 10 and 50 ablation measurements and up to 800 snow soundings were performed.

In addition, the repeated digital elevation models allow the determination of changes in glacier ice volume in periods of some years to a few decades (Bauder et al., 2007). These geodetic glacier mass balances are highly accurate and provide the change in long-term glacial storage of water over the entire 20th century. This data set was used by Huss et al. (2008) to derive monthly mass balances for Rhonegletscher using a
 distributed model calibrated by all available field data (see details in Sect. 2.3). This data set is considered as the reference result for glacier mass changes over the last couple of decades and is used for final model validation (see Fig. 6d).

2.2 Hydrological model

The hydrological model GSM-SOCONT uses a semi-lumped simulation approach. It ¹⁵ computes snow accumulation and snow- and ice melt per elevation bands (here 5 elevation bands for the glacier, 5 for the non-glacier part of the catchment). For each band, the reference precipitation and temperature time series are interpolated according to the difference between its mean elevation and the reference station altitude. For temperature, this interpolation uses a temperature lapse rate ρ [°C m⁻¹]; for precipita-²⁰ tion, it uses a proportional increase with altitude, γ_p [% m⁻¹], of the amount observed at

the reference meteorological station. In the original formulation, the aggregation state of precipitation at the mean elevation of the band is determined based on a simple temperature threshold T_c [°C].

To reduce the model parameter dependency on the spatial resolution (i.e. the number of elevation bands), we adopt here a slightly modified snowfall estimation method. We compute the amount of snowfall W(t,z) and of rainfall V(t,z) at a given time step t and a altitude z and integrate over the hypsometric curve of each elevation band.



The average melt per elevation band, is computed based on a direct linear relation with positive air temperature (Hock, 2003) using a simple degree-day approach:

$$M(t) = \begin{cases} a_{\rm s}(T - T_{\rm m}) \text{ if } T > T_{\rm m}, \ H_{\rm s}(t) > 0\\ a_{\rm i}(T - T_{\rm m}) \text{ if } T > T_{\rm m}, \ H_{\rm s}(t) = 0\\ 0 \qquad \text{if } T < T_{\rm m} \end{cases}$$
(1)

where a_s and a_i are the degree-day factor for snow- respectively ice melt (mm d⁻¹°C⁻¹), ⁵ T_m is the threshold temperature for melting that is set to 0 °C and $H_s(t)$ is the depth of the snow cover at time step t. The transformation of firn (old snow that lasted more than a season) to ice is not accounted for.

For comparison purposes, we use the parameter set from Schaefli et al. (2005) (see Table 1) as a reference simulation.

Glaciological model 2.3 10

15

Huss et al. (2008) applied a distributed accumulation and melt model to the Rhonegletscher to obtain homogenous estimates of glacier mass balance in high temporal and spatial resolution. The model is driven by daily meteorological data and calculates the components of glacier surface mass balance for 25 m × 25 m grid cells. Thus, the spatial distribution of mass balance quantities is addressed in a more sophisticated way than in GSM-SOCONT.

Snow and ice melt is computed using a distributed temperature-index model (Hock, 1999). Degree-day factors are varied in space as a function of potential direct solar radiation to account for the effects of slope, aspect and shading. Instead of simply

- applying an altitude dependent precipitation gradient, Huss et al. (2008) derive a spatial 20 snow accumulation pattern from the detailed winter accumulation measurements and use this pattern to redistribute the precipitation input. Thus, effects of preferential snow deposition and wind-driven snow redistribution are taken into account. The amount of snowfall is estimated based on a linear transition from snowfall to rainfall between the thresholds of 0.5 °C and 2.5 °C. Measured precipitation is corrected using a multiplier
- 25

Discussion Paper HESSD 7,8661-8702,2010 Simulation of high mountainous discharge **Discussion** Paper B. Schaefli and M. Huss **Title Page** Introduction Abstract Conclusions References **Discussion** Paper Tables **Figures** 14 Back Close **Discussion** Paper Full Screen / Esc **Printer-friendly Version** Interactive Discussion

so that the observed winter accumulation in the five seasons with direct field data is matched (Huss et al., 2008).

The parameters of the glaciological model are calibrated on the observed ice volume changes available for subdecadal to multidecadal periods, and the spatial mass balance variability is constrained by the seasonal in-situ field measurements (see Huss et al., 2008). Total annual discharge observations are used to validate the estimates for catchment precipitation, but no other discharge data entered the calibration procedure.

2.4 Optimization algorithm

If nothing else is stated, parameter optimization is completed with the global optimization algorithm called Queueing Multi-Objective Optimiser (QMOO) developed by Leyland (2002). For an application of this optimizer to hydrology, see Schaefli et al. (2004, 2005). The algorithm has been designed to identify difficult-to-find optima and to solve far more complex problems than the ones presented in this paper, where we try to identify the optimal values for a few parameters (up to 14). We, therefore, assume that all identified parameter sets correspond to the best identifiable solution for a given calibration run.

3 Merging data and model

3.1 Theoretical background

Merging model and data is a crucial step in any model development and evaluation process. In hydrology, the main focus was, until recently, on how to integrate different data types into model parameter estimation (e.g. Seibert and McDonnell, 2002; Gupta et al., 1998). This focus is currently shifting to what we can learn about model structural deficiencies (e.g. Fenicia et al., 2008) or about the information content of different data sets (e.g. Winsemius et al., 2009; Reusser et al., 2009). In this paper, we combine



different methods traditionally used in model calibration to extract as much information as possible from the available data sets. Thus, rather than aiming at simply estimating the globally best parameter set, we aim at "informing the model" about potential data and model structural deficiencies while identifying the most likely ranges of parameter values.

We call these most likely parameter ranges posterior or updated ranges as in classical Bayesian parameter estimation (e.g. Kavetski et al., 2006), since they result from updating the modeller's a priori knowledge about the parameter values with all available information about the system behavior.

¹⁰ If this updating includes several different types of data (as in the present case study), there are two possible approaches: we can either identify the posterior parameter range by confronting the model *sequentially* with different types of data or by confronting it *simultaneously*. As we discuss hereafter, the two different approaches will yield different types of results and, thus, extract different information from the data.

3.1.1 Model improvement through multi-objective optimization

Simultaneous incorporation of different types of reference data results in a Multi-Objective Optimization (MOO) problem, where for every reference data type an appropriate objective function is formulated (e.g. Gupta et al., 1998; Yapo et al., 1998). MOO tries to identify parameter sets that correspond to points on the so-called Paretooptimal frontier (POF), the limit in the objective function space beyond which it is not

²⁰ optimal frontier (POF), the limit in the objective function space beyond which it is not possible to improve one objective function without worsening the others (see an illustration for the case of 2 objective functions to be minimized in Fig. 1).

A real-world model with imperfect model structure and imperfect input data can generally not exactly reproduce any of the observed reference data sets. The trade-off be-

tween the different objective functions contains information about how well the model structure can explain the observed data. (This information is partly lost in sequential calibration where the result of each calibration step represents the joint information extracted from all previous steps).



It is not straightforward to use this objective function space trade-off for model development or for statistical interpretation of the results (see Reichert and Mieleitner, 2009). The trade-off between the objective functions is related to the non-dominated volume of the POF (see Fig. 1), i.e. the volume of the objective function space between the POF and the origin. Since each objective function has its own scale, no absolute meaning can be attached to this non-dominated volume. However, if the system to be optimized is modified (e.g. changed model structure) while keeping the objective function definitions unchanged, then the POF with the lower non-dominated volume has a lower trade-off. This property of the POF can be used to improve the model structure

- ¹⁰ (for examples, see Schaefli et al., 2004; Fenicia et al., 2007, 2008) but also to select the most suitable model input data: a simulation resulting from model M_1 with input \mathbf{Y}_1 is judged to be better than a simulation from model M_1 with input \mathbf{Y}_2 if its POF has at least one better end-point or a smaller non-dominated volume.
- It is noteworthy that in MOO, there is always the risk that the used optimization algorithm was simply not able to find points on the true POF. The above assessment can, thus, be misleading for difficult to optimize problems (e.g. a very high number of parameters). A general guideline is that increasing the degree of freedom of a model should always result in a lower trade-off (see Fig. 1), since the model is more flexible. If this is not the case, the results should be interpreted with care.

20 3.1.2 Sequential merging

In hydrological modeling, sequential merging of the reference data with the model is either done in a formal Bayesian framework (e.g. Kuczera and Parent, 1998) or in a rejectionist framework (e.g. Freer et al., 2004).

Bayesian parameter estimation assumes that the data is distributed around some true value and that the model can predict this true value if the suitable parameter set is identified. If several different data sets are available, the parameters can be estimated in sequence. After having estimated a first parameter distribution with a first data set, every additional data set can be used to update the parameter distribution.



This approach is precluded for two cases: (i) if the different data sets are incompatible, i.e. if they contain inconsistent information (this could namely arise if the field data contains systematic measurement errors); (ii) if the available data is sparse and not sufficient to reasonably formulate a statistical model (i.e. not enough data to develop

and test the assumptions). Beven and Binley (1992) introduced the Generalized Like-lihood Uncertainty Estimation (GLUE), partially to address the above problems. Later on, GLUE built the basis of a rejectionist framework (Freer et al., 2004), where acceptable parameter sets are selected among all a priori possible parameter sets by rejecting those that have too low model objective function performance or lie beyond
 some limits of acceptability (see, e.g. Liu et al., 2009).

While the formal and informal Bayesian approaches have been shown to be useful for model calibration on observed reference flux data (in particular discharge, see, e.g. Kavetski et al., 2006; Freer et al., 1996), its use with observed state variables is often precluded because these observations are sparse in time and in space (here we have for example data for only 5 time steps).

Since the problem at hand is typical for the calibration of hydrological models on times series of observed fluxes and some sparse observations of state variables, we propose here a model calibration approach combining elements of both of the above approaches and classical statistical hypothesis testing.

20 3.2 Proposed model calibration framework

15

25

We propose a step-wise calibration framework using MOO as well as sequential updating of an a priori parameter range. In a first step, we use MOO with the observed fluxes and state variables to improve the model structure based on all information that we can extract from the objective function trade-off; the objective functions for this optimization step are discussed in Sect. 3.3. In a second and third step, we sequentially calibrate the resulting model.

This sequential calibration combines a statistical hypothesis testing approach to extract the available information from the state observations and proceeds with classical



model calibration minimizing the sum of squared residuals between an observed and a simulated times series.

As mentioned before, such a sequential parameter range updating approach only works if the different data sets do not contain inconsistent information. Here, this is ⁵ ensured by the preliminary procedure of input data set selection and model structure updating.

3.3 Model performance criteria

3.3.1 Performance with respect to point observations

To assess the model performance with respect to seasonal glacier mass balance, we need to compare the simulated seasonal balance (average per elevation band) to the observed point balances. In the case of large sample sizes and assuming that the samples have been taken at representative locations within the elevation band, the straightforward method would be to estimate the spatial mean from the data and then minimize the residuals between this reference value and the simulated mean value.

Glaciological measurements consist typically of very small samples (i.e. 5 to 10 points per elevation band of a few hundred meters). Without any further source of information, it is highly questionable whether such a small sample contains enough information to estimate the spatial average.

Assuming that the samples have been taken at points that reflect the range of sea-²⁰ sonal balances, we can, however, at least test whether the observed sample is *compatible* with the simulated average value for the elevation band. We use a two-sided sign test of the null hypothesis that the sample data comes from an arbitrary continuous distribution with a given median *m* (e.g Gibbons and Chakraborti, 2003). This is a simple non-parametric test where the test statistics is computed as follows: count the number of samples s_i that are lower than the median *m*, $s_i < m$ and the number

of sample values $s_i > m$. The test statistics equals to the smaller of the two numbers. This test statistics is known to follow a binomial distribution for *n* trials (where *n* is the



sample size) and a probability of success of each trial of 0.5. We can reject the null hypothesis at the 0.05 level that the sample comes from a continuous distribution with median m if the p-value (p) of the test statistics is smaller than 0.025. This test is robust for small sample sizes. We used the data from the winters 2006/2007 and 2007/2008, which have a high spatial resolution, to verify that the observed point balances have sufficiently symmetric distributions to not bias the statistical test when we apply the above statistical test to the mean balance (the hydrological model only yields mean

values) instead of the median.

The performance criterion for a band j and a year t_y becomes:

¹⁰
$$g(j,t_y,\theta) = \begin{cases} 0 \text{ if } p \gg 0.025 \\ 1 \text{ if } p < 0.025 \end{cases}$$

5

The performance criterion $G(\theta)$ considering all available years and elevation bands becomes

$$G(\boldsymbol{\theta}) = \frac{1}{J \cdot n} \sum_{j=1}^{J} \sum_{t=1}^{n} g(j,t,\boldsymbol{\theta})$$

It is important to point out that the above performance criterion *G* can be used to select acceptable parameter sets but not to find a single optimum parameter set. This simply follows from the binary nature of $g(j,t_y,\theta)$, which does not allow a further discriminate between acceptable sets.

3.3.2 Performance criteria for time series

For discharge, we use the classical Nash-Sutcliffe index (Nash and Sutcliffe, 1970) as ²⁰ a performance criterion

$$N_{Q}(\boldsymbol{\theta}) = 1 - \sum_{t=1}^{n} \left[\widetilde{Q}(t) - \widehat{Q}(t|\boldsymbol{\theta}) \right]^{2} \left[\widetilde{Q}(t) - \overline{Q}(.|\boldsymbol{\theta}) \right]^{-2}$$

$$8673$$
(4)

Discussion Paper HESSD 7,8661-8702,2010 Simulation of high mountainous discharge **Discussion** Paper B. Schaefli and M. Huss **Title Page** Introduction Abstract Conclusions References Discussion Paper **Tables Figures** 14 Back Close **Discussion** Paper Full Screen / Esc **Printer-friendly Version** Interactive Discussion

(2)

(3)

where $\tilde{Q}(t)$ [mm d⁻¹] is the observed discharge on day t, \overline{Q} the corresponding annual mean value, $\hat{Q}(t|\theta)$ is the simulated discharge given parameter set θ and n is the number of simulated time steps. (\hat{C} denotes a value obtained from a model simulation, whereas \hat{C} denotes an observed value).

Since for MOO we have to formalize the optimization problem as a minimization problem, we use the performance criterion $f_Q(\theta)$ for optimization:

 $f_{\mathsf{Q}}(\boldsymbol{\theta}) = 1 - N_{\mathsf{Q}}(\boldsymbol{\theta})$

20

If time series of the glacier-wide annual mass balance are available, a similar performance criterion can be defined:

10
$$f_{\rm B}(\boldsymbol{\theta}) = \sum_{t=1}^{n_y} \left[\widetilde{B}(t_y) - \widehat{B}(t_y|\boldsymbol{\theta}) \right]^2$$
 (6)

where $\widetilde{B}(t_y)$ [mm y⁻¹] is the observed glacier-wide mass balance for the year t_y , $\widehat{B}(t_y|\theta)$ is the corresponding simulated value and n_y is the number of available yearly values.

3.3.3 Sequential parameter range updating and final calibration

Once an optimal model structure has been identified, its parameters are sequentially ¹⁵ updated starting with some a priori knowledge about the possible range of the parameter values:

- 1. Select a state observation data set (e.g. point mass balance $\tilde{b}[x,y,t|(x,y) \in j]$) and the corresponding model simulation $\hat{b}[j,t|\theta]$.
- 2. Among all model parameters $\boldsymbol{\theta}$, select the subset $\boldsymbol{\theta}_{v}$ that influences the simulation of $\hat{b}[j,t|\boldsymbol{\theta}]$.
- 3. Generate *r* random realizations of $\boldsymbol{\theta}_{v}$ drawn in the prior or updated parameter ranges and compute the corresponding $\hat{b}[j,t|\boldsymbol{\theta}](t|\boldsymbol{\theta}_{v})$.



(5)

- 4. For each $\hat{b}[j,t|\theta](t|\theta_v)$, test the null hypothesis H_0 that the observed sample $\tilde{b}[x,y,t|(x,y) \in j]$ (point measurements) comes from a distribution having as median $\hat{b}[j,t|\theta](t|\theta_v)$ against the alternative hypothesis that it does not (see Sect. 3.3.1).
- 5. Retain the parameter sets for which H_0 cannot be rejected for all available data points v(t).
 - 6. Redefine the minimum and the maximum of the possible parameter ranges based on the minimum and maximum values of the parameters retained in the previous step.
- ¹⁰ 7. If available, repeat points 1 to 8 for additional state observation data sets.
 - 8. Use the updated parameter ranges as prior parameter ranges to calibrate the model on the available time series (discharge, glacier-wide mass balance) (see Sect. 3.3.2).

4 Results

20

15 4.1 Reference data sets for parameter calibration

We use the following reference data sets for parameter calibration:

- Daily discharge, the years 1974–1982 for calibration, 1983–1999 for validation (performance criterion $N_{\rm Q}$);
- Observed winter point accumulation (direct measurements) for winters 1979/1980 to 1981/1982; for each year, we only use the 3 highest bands (to ensure that the ice melt parameter does not influence the model result). One sample (elevation band 4, winter 1981/1982) has less than 10 points and is excluded from the analysis. In total, we have 3 × 3 1 = 8 reference samples.

iscussion Pa	HES 7, 8661–8	HESSD 7, 8661–8702, 2010					
aper Discussion	Simulatio mount disch B. Schaeflia	Simulation of high mountainous discharge B. Schaefli and M. Huss					
Paper [Title Abstract Conclusions	Page Introduction References					
Discussion Pap	Tables	Figures					
er Discus	Back Full Scre	Close een / Esc					
ssion Paper	Printer-friendly Version Interactive Discussion						

- Observed annual point balance (direct measurements) for the corresponding hydrological years. Only the two lowest elelvation bands have more than 10 observations, resulting in 6 reference samples.
- Glacier-wide annual balance (performance criterion $f_{\rm B}$): for the years 1979/1980
- to 1981/1982, we have reference values estimated from direct measurements. For the validation of the hydrological model, we complete these reference values with the glacier-wide annual balances simulated with the glaciological model of Huss et al. (2008)

4.2 MOO for modeling system improvement

5

10 4.2.1 Selection of meteorological input series

There is the typical problem that the available meteorological time series have been observed at points located outside the catchment. Assessing the information content of the available series for the modeling purpose is, thus, an essential modeling step. Here, the overall goal is to simulate discharge and glacier mass balance. As discussed

- in Sect. 3.1.1, analyzing the trade-off between the two corresponding objective functions for the different input series, is a valuable tool to identify the most suitable input time series. Figure 3a shows the trade-offs between the model performances for discharge simulation and for glacier-wide annual mass balance simulation using the different available input time series of precipitation and temperature. The 3 Pareto-optimal
- frontiers (POFs) show clearly that using the precipitation time series observed at Oberwald leads to much better discharge simulations and to a smaller trade-off between obtaining good results for discharge and for mass balance. Accordingly, we prefer this measurement station for this study (even if it stopped recording in 1999). Assessing in detail the differences in information content of the available meteorological series is beyond the scope of this paper.



4.2.2 Model structure modification

The GSM-SOCONT model (Schaefli et al., 2005) has 6 parameters that influence the glacier mass balance simulations: the ice and snow melt factors, a_i and a_s , the precipitation correction factor γ_p , the temperature lapse rate ρ and the threshold temperature

- ⁵ for rainfall/snowfall separation T_c . In the original version, only the first 3 parameters are calibrated, the remaining were fixed, $\rho = 0.0065 \,^{\circ}\text{Cm}^{-1}$ and $T_c = 0 \,^{\circ}\text{C}$. The original model, calibrated according to Schaefli et al. (2005) on discharge and glacier-wide annual balance for 3 hydrological years (1979/1980 to 1981/1982), cannot reproduce the seasonal mass balances. To illustrate this result, Fig. 2a shows the observed ac-
- ¹⁰ cumulation quantiles and the simulated winter accumulation per elevation band for one winter; the simulated mean value does not always lie between the observed 10% and 90% quantiles. This suggests that the inclusion of the winter balance observations adds additional information in the sense that it constrains the model differently than without using it.

4.2.3 New model structure for point balance simulation

An attempt of re-calibration of all of the above model parameters on the seasonal point balances, according to the performance criteria described in Sect. 3.3, showed that GSM-SOCONT cannot reproduce both winter and annual point balances: for the winter balance, the smallest achievable value of $G_w(\theta)$ is 2 and for the annual balance $G_a(\theta) = 3$. This means that for the winter balance, there are least two bands for which we can reject the hypothesis that the observed sample has the median value simulated

by the model.

20

25

In a first step, we tried to use a linear transition between rainfall and snowfall (e.g. Rohrer et al., 1994) and to calibrate the threshold parameters between which there is a mixture of snowfall and rainfall. Even if this makes the model more flexible, it is still

not possible to achieve $G_w(\theta) = 0$ and $G_a(\theta) = 0$ simultaneously.



We next tested the use of seasonally varying model parameters. We successively increased the number of seasonally varying parameters, starting with the temperature lapse rate. It is well known that temperature lapse rates show strong seasonal fluctuations, under the combined effect of dry and saturated adiabatic lapse rates and their dependence on temperature. The temperature decrease with altitude is on average lower during the cold season than during the warm season. This is also confirmed by an analysis of seasonal temperature variations at all meteorological stations within some tens of kilometers around Rhonegletscher, which showed that temperature lapse

5

¹⁰ We, thus, introduced the seasonal temperature lapse rates ρ_w and ρ_s (winter is defined as from 1 November to 30 April, which corresponds more or less to the dates of winter accumulation observation.) With this additional degree-of-freedom, the model can achieve $G_w(\theta) = 0$ and $G_a(\theta) = 0$, i.e. there are parameter sets that reproduce both the reference point data for winter accumulation and annual balance. This is a priori not surprising; we increased the degree-of-freedom and provided the model with more flexibility to reproduce observed data. It can now, for example, simulate snow melt

rates are strongest in June and reach a minimum from November to January.

during the winter season. This modification also reduces the performance trade-off between glacier-wide annual balance and discharge simulation (Fig. 3b). It is noteworthy that the performance ²⁰ with the Grimsel input station is still much worse, confirming our initial choice not to consider this station. The new model version has, however, still an important shortcoming: considering all good solutions (i.e. parameter sets) on the POF of Fig. 3b, the seasonal lapse rates show a distinct dependence on each other for good discharge parameter sets versus good mass balance parameter sets (see scatterplot enclosed in

²⁵ Fig. 3b). This implies that this model version needs different pairs of temperature lapse rates for mass balance simulation than for discharge simulation.

The seasonal lapse rates have, for evident reasons, a strong interaction with the precipitation input correction factor. A solution to overcome the above structural short-coming is the introduction of seasonal input correction factors. As tests showed, this



8679

leads to complex parameter interactions with the rainfall/snowfall separation thresholds, which can be solved by modifying the input pre-processing: we replace the altitudinal precipitation correction factor with a solid input (snow) correction factor, γ_s [% m⁻¹] and an constant rainfall correction factor, γ_r [% m⁻¹]).

The new interpolation of precipitation as a function of elevation reads as

5

10

15

$$P(z) = W(z) + V(z) = \left[1 + \gamma_{s}(z - z_{ref})\right]W(z) + \left[1 + \gamma_{r}(z - z_{ref})\right]V(z)$$
(7)

where W(z) and V(z) are the snowfall and the rainfall at elevation *z*. This is also more meaningful from a physical perspective: rainfall and snowfall occur under different meteorological situations. Their average scaling with elevation can, therefore, be assumed to be quite different.

The new model structure has the nice property that we can fix the rainfall/snowfall separation thresholds (to 0 °C and 2 °C) and still obtain parameter sets that reproduce the reference winter accumulation and annual point balance. This result holds independent of the chosen meteorological input series (but, of course, for each series, we obtain different parameter sets).

We would like to add here that, in this area of the Alps, it is difficult to derive precipitation increase factors directly from observed data (e.g. Sevruk, 1997). Accordingly, it is tempting to think that the above input correction parameters are just "fetch" factors that allow a correction of the observed input for model calibration. The use of a

²⁰ separate snow accumulation correction factor is, however, in line with the finding of Sevruk (1997) that, in the Valais area, significant altitude gradients of precipitation can be found above 1500 m (but not below). This is also about the lower limit of the seasonal snowcover. For the Rhonegletscher, this observation is confirmed by the snow accumulation measurements that show a strong increase with elevation only for high elevations (Huss et al., 2008).

In summary, after confronting the model with the available data sets of winter accumulation and annual point balance, we retain the following model structure:

- Fixed rainfall/snowfall separation thresholds (0 °C and 2 °C).



- Seasonal temperature lapse rates ($\rho_{\rm w}, \rho_{\rm s}$).
- Constant ice melt factor a_i, snow melt factor a_s and snow accumulation correction factor γ_s.

4.2.4 New model structure for glacier-wide balance and discharge

- ⁵ If we optimize this model structure on annual glacier-wide mass balance and discharge, there remains an important trade-off between achieving good mass balance and good discharge simulations (see Fig. 4a, model 11p). Introducing a separate summer snow melt factor a_{ss} partly reduces this trade-off but good mass balance simulations lead to far too much winter discharge. Adding a separate summer snow accumulation correc-
- ¹⁰ tion factor considerably reduces the trade-off (see Fig. 4a, model 13p). However, there is still too much discharge in early spring for all good mass balance simulations (see Fig. 4b). This problem can be addressed by introducing a water retention capacity η_s for the snow layer. In the original GSM-SOCONT, snow melt $M_s(t)$ is directly transferred to the water-runoff transfer module (two linear reservoirs for the glacier part, a linear and a poplinear reservoir for the pop-glacier part). In the new version, only if the
- ⁵ linear and a nonlinear reservoir for the non-glacier part). In the new version, only if the ratio between melt water and snow layer depth is higher than η_s , snow layer outflow occurs.

The *final improved model structure*, used in the rest of this paper, has 14 parameters to be calibrated:

- Fixed rainfall/snowfall separation thresholds (0°C and 2°C) and linear transition between them.
 - Winter and summer temperature lapse rates ($\rho_{\rm w}, \rho_{\rm s}$).
 - Ice melt factor a_i , constant throughout the year.
 - Winter and summer snow melt factors (a_{sw}, a_{ss}) ,
- Winter and summer snow accumulation correction factors (γ_w , γ_s).

	HESSD 7, 8661–8702, 2010					
apor I Discussio	Simulation of high mountainous discharge B. Schaefli and M. Huss					
	Title Page					
	Abstract	Introduction				
_	Conclusions	References				
	Tables	Figures				
<u>с</u>	14	►I.				
		•				
-	Back	Close				
	Full Screen / Esc					
<u>.</u>	Printer-friendly Version					
	Interactive Discussion					
D						

- Summer rainfall correction factor γ_r .
- 6 water-runoff transfer parameters (A, log(k), β , k_{i} , k_{s} , η_{s}).

4.3 Sequential calibration

We successively constrain the model parameter ranges from physically possible (a priori) ranges (see Table 2) to ones that are plausible given the point reference data (winter and annual point balances) to the final best parameter set for discharge and glacier-wide annual balance simulation. During this sequential updating, we fix all parameters to an initial guess (see Table 1) and vary only the ones corresponding to an updating step.

10 Winter point balance

We start with the parameters that influence winter accumulation, i.e. the winter accumulation correction factor γ_w , the winter snow melt factor a_{sw} and the temperature lapse rate $\rho_{\rm w}$. During this period, the other parameters have no influence on the model output. Figure 5b, d and f show the distribution of the parameter sets for which we cannot reject the null hypothesis that the observed samples have the simulated median 15 value for all retained elevation bands and winters (Sect. 4.1, a total of 100 000 parameter sets was randomly drawn in the initial priors of Table 2). The temperature lapse rates are very small and they show a certain correlation with the snow melt factors (linear correlation of -0.6, see Fig. 5i). This result is to be expected. Melt factors are known to be lapse-rate specific (e.g. Shea et al., 2009) since, to melt a given quantity 20 of snow, the melt factor has to be higher if the temperature is lower (i.e. of the temperature lapse rate is more negative). For the ice melt factor, this relationship is obscured by the fact that ice melt only occurs if the glacier is snow-free. The other parameters do not show any pair-wise dependence structure.



Annual point balance

We reduce the prior for a_{sw} , γ_w and ρ_w to the above identified ranges and draw another 100 000 parameter sets including the summer melt factor, a_{ss} , the summer temperature lapse rate, ρ_s , the summer accumulation correction factor γ_s and the ice melt factor a_i .

- ⁵ The distributions of the parameter sets for which the null hypothesis cannot be rejected for all elevation bands and years are shown in Fig. 5 and summarized in Table 2. The summer snow melt factor a_{ss} and the ice melt factor a_i show a very strong dependence on each other (see Fig. 5h). This dependence shows the typical shape reported by previous studies (see, e.g. Schaefli et al., 2005, Fig. 4) or (Hock, 1999, Fig. 4).
- Figure 6d shows the simulated annual glacier-wide balance corresponding to one of the acceptable parameter sets. This simulation is strikingly different from the reference simulation of Huss et al. (2008). This illustrates that the point mass balances contain some information but not enough to constrain the model. This is partly due to the fact that for annual point balance, only the lowest elevation bands have sufficiently large samples.

Discharge and glacier-wide balance

We complete a final parameter calibration minimizing $f_Q = 1 - N_Q$ and f_B , limiting the search space to the parameter ranges that are acceptable given the point balance observations. Given that neither discharge nor glacier-wide annual balance contain ²⁰ direct information on the temperature lapse rates, we fix their value to a previously identified possible parameter couple (see Table 1). Even under the final updated model structure, there remains a trade-off between f_Q and f_B (Fig. 6a), corresponding to a considerable range of possible system representations. This is illustrated in Fig. 6c and d that show the discharge and the glacier-wide mass balance simulations for all parameter sets on the Pareto-optimal frontier. The best performing parameter sets under f_Q and f_B are given in Table 1. The performance of the best f_Q simulation in terms of mimicking the observed discharge is comparable to the calibration presented



in Schaefli et al. (2005) (Nash value of 0.92 and bias -0.01 for the calibration period, 0.90 and bias of 0.03 for the entire period 1976–1999). The rainfall/melt water-runoff transformation parameters of the non-glacier part have considerably different values then the original GSM-SOCONT, which is due to the modification of the transfer module with the introduction of η_s .

5

10

Figure 6d shows the simulated glacier-wide annual mass balance for all parameter sets on the POF. We compare this simulation to the one corresponding to the original model structure calibrated on f_Q and on f_B . Just as the original model, the new model strongly overestimates the negative cumulative mass balance at the end of the simulation period (further discussed in Sect. 5.2) leading to a strong bias, where the bias between a simulated variable *x* and a reference variable *y* is defined as $100(x - y)y^{-1}$.

The original model shows a mean bias compared to the simulation of Huss et al. (2008) of 294%. In the new model, this is reduced to 101% for the best simulation under N_Q and equals –33% for the best simulation under f_B .

¹⁵ The variability of the monthly glacier-wide balance of the reference simulation is also better reproduced by the new simulation. This can be measured by the coefficient of variation, which corresponds to the standard deviation of a variable *x* divided by the mean of *x*. The coefficient of variation of the reference simulation is –14.9, the new model gives –7.1 under $N_{\rm Q}$ and –13.3 under $f_{\rm B}$ whereas the original model had a far too low value (–3.9).

Finally, Fig. 6b illustrates an important detail finding; in previous work, we used a warm-up period of 2 years for model calibration, assuming that this is sufficient to build-up the initial snowpack. Given the new model formulation, including the snowpack retention capacity, a much longer warm-up period is required since during the 3rd and

²⁵ 4th year, the built-up firn (snow that lasts more than one season) is still too small to retain the meltwater in early spring.



5 Discussion

5.1 Multi-objective calibration for model improvement

The presented case study revealed the importance of considering flux and state variables for model development, i.e. of using several reference signals. First of all, for
 the choice of appropriate input series: considering discharge performance or mass balance performance individually, none of the available measurement stations would appear to contain more valuable information. Considering the trade-off between the two objective functions reveals that there is a considerable difference in information content. We believe that this type of analysis has great potential for hydrological modeling where assessing the (dis-)information content of input time series is still one of the big open questions.

During the model modification process, it is again the trade-off between the reproduction of these two model variables that provides clear indications about the value of adding a parameter or modifying a submodule. The multi-signal approach allows the detection of a subtle model improvement through the introduction of the snow layer retention capacity η_s : adding this parameter does not lead to a better discharge simulation, i.e. we cannot find a parameter set under the new model structure (including η_s) that outperforms the best discharge simulation under the old model structure. But the best mass balance parameter sets do also a fairly good job for discharge simulation and vice-versa; the trade-off (see Fig. 1) between discharge performance and mass

²⁰ and vice-versa; the trade-off (see Fig. 1) between discharge performance and mass balance performance is reduced.

Our results also underline the value of analyzing the parameter sets associated with the simulation trade-offs, i.e. of analyzing the parameter sets corresponding to points on the Pareto-optimal frontier (Fig. 3): the strikingly different relationship between win-

ter and summer lapse rate parameters for good discharge simulation versus good mass balance simulation was a very strong hint for a model structural deficiency. Again, this type of model diagnostics could be very valuable in other hydrological contexts.



5.2 Reproduction of glacier mass balance and water balance

As our final calibrated model demonstrates, a simple semi-lumped hydrological model can indeed reproduce glacier mass balance dynamics if there is enough reference data to identify the model structure and calibrate the model parameters. We can, however,

5 not expect the hydrological model to reproduce the exact interannual glacier-wide mass balance dynamics of the more complex glaciological model, partly because we do not use the same meteorological input data.

The original GSM-SOCONT model systematically overestimates summer melt and underestimates winter accumulation, with respect to data (Fig. 2), as well as with respect to the reference simulation, even if the annual glacier-wide mass balance is fairly close to the reference simulation. This clearly highlights the need to include seasonal mass balance observations in the model development process.

The most important difference between the original GSM-SOCONT and the new model version becomes visible if we consider the individual mass balance terms (Ta-

¹⁵ ble 3). The total area-averaged precipitation is much higher for the new model than for the original model, evaporation is lower. These differences in the mass balance terms can result in big differences in the relative contribution of glacier storage change to total annual runoff (Fig. 6e) in individual years. This relative contribution is not only a proxy for climate sensitivity but also a key variable for water management and in particular
 ²⁰ hydropower production (see Schaefli et al., 2007).

The question which of the two water balance representations is closer to the "truth" is extremely difficult to answer. There are estimated reference values on a 2 km by 2 km grid for evaporation and precipitation (see Table 3) but given the resolution and the complexity of the topography, these are just rough estimates for such a small catchment. A

rather strong hint is, however, the following: for the years 1974–1990 (to which Table 3 refers to), the reference simulation of Huss et al. (2008) corresponds to a mean glacier storage change of –260 mm w.e. (water equivalent) per year relative to the ice-covered area (around 17 km²). Relative to the entire catchment area (38.9 km²), this equals



-110 mm w.e. a^{-1} , a value that is very close to the storage change simulated with the hydrological model for the same period (-100 mm w.e. a^{-1}), whereas the original model gives a value of -250 mm w.e. a^{-1} . It is noteworthy that the hydrological model slightly underestimates the overall catchment water storage change since it does not simulate any evaporative losses on the glacier-covered catchment part.

Neither the original nor the new model can reproduce the annual glacier-wide mass balance simulations at the end of the simulation period. As discussed by Huss et al. (2009), this period corresponds to a period of intense melting where traditional temperature-index models might become unreliable. Huss et al. (2009) found in fact consistently negative degree-day factor anomalies for this period. The question how to

account for such anomalies in temperature-index types hydrological models, especially for projections into the future, remains open to date.

5.3 Information content of glacio-hydrological data

5.3.1 Calibration on point balance observations

5

¹⁵ The proposed model calibration framework makes use of small samples of observed point mass balances (on average, around 20 samples per elevation band) to identify plausible parameter sets for which we cannot reject the hypothesis that the observed samples have the simulated values as median.

Our results suggest that even small samples of point mass balance observations contain valuable information to calibrate 4 additional hydro-meteorological model parameters with respect to the original GSM-SOCONT. The obtained trade-off for the reproduction of winter balance versus annual balance with the original model provides evidence that both data sets contain complementary information. The fact that this trade-off can be removed through simple model modifications suggests that the point data sets are not inconsistent.

Such observations are often deemed to be too sparse to be useful to calibrate a simple semi-lumped hydrological model. Our results, however, demonstrate that the data



is indeed useful to identify a model structure that can reproduce the seasonal mass balance dynamics and furthermore, it constrains the plausible ranges of the parameters. This conclusion is at least partly unexpected: in fact, it is often assumed that winter accumulation data does not contain any information on temperature lapse rates argu-

⁵ ing that in glacier catchments, the low winter temperatures lead to permanent snow fall independent of the temperature lapse rate.

The updated plausible parameter ranges contain other hints that support the value of the proposed method to extract information from the mass balance samples: (i) the temperature lapse rates are less strong in winter than in summer as expected for theoretical reasons (see Sect. 4.2.2) and (ii) the parameters do not show unexpected pair-wise dependencies, the dependence of a_{ss} and a_i shows the typical shape ob-

tained in classical calibration on time series (Fig. 5).

5.3.2 Is glacier-wide balance data sufficient?

15

10

Our case study shows that using only discharge and annual glacier-wide mass balance data can lead to wrong intra-annual mass balance dynamics. And using only point mass balance observations can lead to wrong glacier-wide mass balance. Both results can be easily understood if we consider the general water balance equation:

$$dS/dt = \sum_{j=1}^{J} F_{j}^{+}(t) - \sum_{j=1}^{J} F_{j}^{-}(t)$$
(8)

where $F_i^+(t)$ is the influx of a spatial unit, $F_t^-(t)$ the outflux and dS/dt the yearly storage

²⁰ change. In many catchments, it can be assumed that dS/dt is very small or even negligible. This does not hold, for evident reasons, in catchments with a significant glacier cover. The same storage change can be obtained with different compensations between the flux terms. For a reliable model calibration, observed reference data (or any other information about the behavior of the system) has to provide constraints (i) on dS/dt and either $F_i^+(t)$ or $F_t^-(t)$ or (ii) on $F_t^+(t)$ and $F_t^-(t)$.

Glaciological data provides direct information on the different mass balance terms: winter accumulation observations on $F_j^+(t)$, annual balance observations on $F_j^-(t)$ (often observed only for low elevation bands) and annual glacier-wide mass balance estimations on dS/dt. Discharge data also contains information on these terms but obscured by the hydrological processes in the ice-free part; it namely does not contain any direct information on dS/dt given that there is an essentially unknown outflux: evaporation.

Since in a glacier catchment, any lack of snow accumulation during winter can be compensated by ice melt, discharge effectively contains little direct information on snow- or rainfall (i.e. on $F_j^+(t)$). This shows that for a reliable model calibration, we have to pick winter accumulation and one or two other elements of the above list. If we have a good model structure, winter accumulation data and discharge data could be sufficient to calibrate the model. In practice, however, annual balance data will be needed to compensate for imperfect model structure. The same holds for the use of winter accumulation with annual glacier-wide mass balance data. In any case, using only discharge and annual glacier-wide data is highly likely to lead to at least partly erroneous estimated fluxes. Model calibration only on discharge data, common practice in hydrology, can probably be assimilated to mere guessing in the context of prediction of high Alpine hydrology.

20 6 Conclusions

Getting the water balance right is a basic requirement for any hydrologic model – yet, in many environments, this is very difficult to achieve. In the case of glacier catchments, it is often assumed that some glacier-wide annual mass balance estimations, in addition to observed discharge, is good enough to obtain reliable estimates of the different water balance terms and to develop prediction models. In this paper, we present evidence

²⁵ balance terms and to develop prediction models. In this paper, we present evidence that information on the seasonal mass balance is a pre-requisite to reliably calibrate a



hydrological model and we demonstrate the value of a simple, but promising method to use small samples of point observations for this calibration.

The observed seasonal mass balance data could only be reproduced through a model structure modification. The identified new model structure shows some inter-

- esting features; it has seasonal lapse rates and a separate snow accumulation and rainfall correction factor. These last two parameters are required for the model to be sufficiently flexible to compensate differently for missing processes during summer and winter (e.g. raingauge undercatch, redistribution through wind etc.). In summary, we can say that a simple conceptual model, as the one presented here, can be a valu-
- ¹⁰ able tool to project the behavior of a glacier catchment but only if we have enough (seasonal) information to constrain the parameters that directly affect the mass balance. Since in most catchments around the world, virtually no observed ground-based data is available, future research could focus on the question of how to extract such information from remotely-sensed data. It would also be interesting to assess how the ¹⁵ information content of seasonal mass balance data varies across hydro-climatological
- regions.

Glacier hydrology deals with systems where the main water storage term, the glaciers, can be directly observed - an invaluable advantage over other ecosystems. We showed in this paper possible ways of taking advantage of this direct data on storage and flux to select appropriate input time series and to improve the hydrologic model structure. We believe that such catchments offer ideal test cases for further developments in this direction. Since, in addition, these catchments are undergoing rapid change, they can also be used to assess the ability of hydrologic models to predict change.

Acknowledgements. The authors would like to acknowledge M. Funk for his efforts in collecting seasonal mass balance data of Rhonegletscher, and providing them for this study. The meteorological time series have been provided by the Swiss national weather service, MeteoSwiss, and the discharge data by the Swiss Federal Office for the Environment (Hydrological Section).



References

- Bauder, A., Funk, M., and Huss, M.: Ice-volume changes of selected glaciers in the Swiss Alps since the end of the 19th century, Ann. Glaciol., 46, 145–149, doi:10.3189/ 172756407782871701, 2007. 8665, 8666
- ⁵ Beven, K. J. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, doi:10.1002/hyp.3360060305, 1992. 8671
 - De Walle, D. and Rango, A.: Principles of snow hydrology, Cambridge University Press, Cambridge, 2008. 8695
- Fenicia, F., Solomatine, D. P., Savenije, H. H. G., and Matgen, P.: Soft combination of local models in a multi-objective framework, Hydrol. Earth Syst. Sci., 11, 1797–1809,
 - doi:10.5194/hess-11-1797-2007, 2007. 8670
 - Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Understanding catchment behavior through step-wise model concept improvement, Water Resour. Res., 44, W01402, doi:10. 1029/2006WR005563, 2008. 8668, 8670
- ¹⁵ Freer, J., Beven, K., and Ambroise, B.: Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, Water Resour. Res., 32, 2161– 2173, doi:10.1029/95WR03723, 1996. 8671
 - Freer, J. E., McMillan, H., McDonnell, J. J., and Beven, K. J.: Constraining dynamic TOP-MODEL responses for imprecise water table information using fuzzy rule based performance
- ²⁰ measures, J. Hydrol., 291, 254–277, doi:10.1016/j.jhydrol.2003.12.037, 2004. 8670, 8671 Funk, M.: Räumliche Verteilung der Massenbilanz auf dem Rhonegletscher und ihre Beziehung zu Klimaelementen, Ph.D. thesis, Eidgenössische Technische Hochschule Zürich, nebis, 1985. 8665, 8702

Gibbons, J. and Chakraborti, S.: Nonparametric Statistical Inference, Marcel Dekker, New York,

- 25 2003. 8672
 - Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, Water Resour. Res., 34, 751– 763, doi:10.1029/97WR03495, 1998. 8668, 8669

Haberlandt, U., Ebner von Eschenbach, A.-D., and Buchwald, I.: A space-time hybrid hourly

³⁰ rainfall model for derived flood frequency analysis, Hydrol. Earth Syst. Sci., 12, 1353–1367, doi:10.5194/hess-12-1353-2008, 2008. 8663



Hingray, B., Schaefli, B., Mezghani, A., and Hamdi, Y.: Signature-based model calibration for hydrologic prediction in mesoscale Alpine catchments, Hydrolog. Sci. J., 55, 1002–1016, doi:10.1080/02626667.2010.505572, 2010. 8663

Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, J. Glaciol., 45, 101–111, 1999. 8667, 8682

- solar radiation, J. Glaciol., 45, 101–111, 1999. 8667, 8682
 Hock, R.: Temperature index melt modelling in mountain areas, J. Hydrol., 282, 104–115, doi: 10.1016/S0022-1694(03)00257-9, 2003. 8663, 8667
 - Horton, P., Schaefli, B., Hingray, B., Mezghani, A., and Musy, A.: Assessment of climate change impacts on Alpine discharge regimes with climate model uncertainty, Hydrol. Process., 20, 2091–2109, doi:10.1002/hyp.6197, 2006. 8663
- Huss, M., Bauder, A., Funk, M., and Hock, R.: Determination of the seasonal mass balance of four Alpine glaciers since 1865, J. Geophys. Res., 113, F01015, doi:10.1029/2007JF000803, 2008. 8664, 8665, 8666, 8667, 8668, 8676, 8679, 8682, 8683, 8685, 8702

10

30

Huss, M., Funk, M., and Ohmura, A.: new paper showing strange behavior for 2007, Geophys. Res. Lett., 36, L23501, doi:10.1029/2009GL040789, 2009, 8686

Res. Lett., 36, L23501, doi:10.1029/2009GL040789, 2009. 8686 Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, Water Resour. Res., 42, W03407, doi:10.1029/2005WR004368, 2006. 8669, 8671

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, anal-

- yses, and models to advance the science of hydrology, Water Resour. Res., 42, W03S04, doi:10.1029/2005WR004362, 2006. 8663
 - Klok, E., Jasper, K., Roelofsma, K., Gurtz, J., and Badoux, A.: Distributed hydrological modelling of a heavily glaciated Alpine river basin, Hydrolog. Sci. J., 46, 553–570, doi: 10.1080/02626660109492850, 2001. 8663
- Koboltschnig, G. R., Schoner, W., Zappa, M., Kroisleitner, C., and Holzmann, H.: Runoff modelling of the glacierized Alpine Upper Salzach basin (Austria): multi-criteria result validation, Hydrol. Process., 22, 3950–3964, doi:10.1002/hyp.7112, 2008. 8664
 - Kuczera, G. and Parent, E.: Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm, J. Hydrol., 211, 69–85, doi:10.1016/ S0022-1694(98)00198-X, 1998. 8670
 - Leyland, G. B.: Multi-Objective Optimisation Applied to Industrial Energy Problems, Ph.D. thesis, available at http://library.epfl.ch/theses, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2002. 8668



- European Alps, 1971–1990, in: Hydrological Atlas of Switzerland, Plate 2.6, Service Hydrologique et Géologique National, Berne, 2001. 8696
- Alps: guantification of potential impacts and related modelling uncertainties, Hydrol. Earth Svst. Sci., 11, 1191–1205, doi:10.5194/hess-11-1191-2007, 2007, 8685 30 Schwarb, M., Daly, C., Frei, C., and Schär, C.: Mean annual precipitation throughout the
- Schaefli, B., Hingray, B., Niggli, M., and Musy, A.: A conceptual glacio-hydrological model for high mountainous catchments, Hydrol. Earth Syst. Sci., 9, 95-109, doi:10.5194/hess-9-95-2005, 2005. 8663, 8664, 8667, 8668, 8677, 8682, 8683, 8694, 8699, 8702 Schaefli, B., Hingray, B., and Musy, A.: Climate change and hydropower production in the Swiss
- Schaefli, B., Hingray, B., and Musy, A.: Improved calibration of hydrological models: use of a multi-objective evolutionary algorithm for parameter and model structure uncertainty estimation, in: Hydrology: Science and Practice for the 21st Century, edited by: Webb, B., British Hydrological Society, London, 362-371, 2004. 8668, 8670
- doi:10.5194/hess-13-999-2009, 2009. 8668 Rohrer, M. B., Braun, L. N., and Lang, H.: Long-term records of snow cover water equivalent in the Swiss Alps, 1. Analysis, Nord. Hydrol., 25, 53-64, 1994. 8677
- models with stochastic, time-dependent parameters, Water Resour, Res., 45, W10402, doi: 10.1029/2009WR007814, 2009. 8670 Reusser, D. E., Blume, T., Schaefli, B., and Zehe, E.: Analysing the temporal dynamics

of model performance for hydrological models, Hydrol. Earth Syst. Sci., 13, 999-1018,

- discussion of principles, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970. 8673 Reichert, P. and Mieleitner, J.: Analyzing input and structural uncertainty of nonlinear dynamic
- Jakob, M.: Glacier change in western North America: influences on hydrology, geomorphic hazards and water quality, Hydrol. Process., 23, 42-61, doi:10.1002/hyp.7162, 2009. 8663 Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models, Part I, a 10
- Berne, 1999. 8696 Moore, R. D., Fleming, S. W., Menounos, B., Wheate, R., Fountain, A., Stahl, K., Holm, K., and

Menzel, L., Lang, H., and Rohmann, M.: Mean annual actual evaporation 1973-1992, in: Hydrological Atlas of Switzerland, Plate 4.1, Service Hydrologique et Géologique National,

Liu, Y., Freer, J., Beven, K., and Matgen, P.: Towards a limits of acceptability approach to

doi:10.1016/j.jhydrol.2009.01.016, 2009. 8671

5

15

20

the calibration of hydrological models: Extending observation error, J. Hydrol., 367, 93–103,

Discussion Paper 7,8661-8702,2010 Simulation of high

Discussion Paper

Discussion Paper

Discussion Paper

mountainous discharge

HESSD

B. Schaefli and M. Huss





- Seibert, J. and McDonnell, J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resour. Res., 38, 1241, doi:10.1029/2001WR000978, 2002. 8668
- Sevruk, B.: Regional dependency of precipitation-altitude relationship in the Swiss Alps, Climatic Change, 36, 255–369, doi:10.1023/A:1005302626066, 1997. 8679
- Shea, J. M., Moore, R. D., and Stahl, K.: Derivation of melt factors from glacier mass-balance records in western Canada, J. Glaciol., 55, 123–130, doi:10.3189/002214309788608886, 2009. 8681

Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrological prediction, Hydrol. Process., 17, 2101–2111, doi:10.1002/hyp.1425, 2003. 8663

Stahl, K., Moore, R., Shea, J., Hutchinson, D., and Cannon, A.: Coupled modelling of glacier and streamflow response to future climate scenarios, Water Resour. Res., 44, W02422, doi: 10.1029/2007WR005956, 2008. 8663

SwissTopo: Digital height model of Switzerland - DHM25, 1995. 8665

5

10

- ¹⁵ Wagener, T., Sivapalan, M., Troch, P. A., and Woods, R.: Catchment classification and hydrologic similarity, Geogr. Compass, 1, 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007. 8663
 - Winsemius, H., Schaefli, B., Montanari, A., and Savenije, H.: On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological informa-
- tion, Water Resour. Res., 45, W12422, doi:10.1029/2009WR007706, 2009. 8668
 Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-objective global optimization for hydrologic models, J. Hydrol., 204, 83–97, doi:10.1016/S0022-1694(97)00107-8, 1998. 8669



Table 1. Calibrated and default GSM-SOCONT parameter values; the default values are from (Schaefli et al., 2005). The calibrated values are for the final improved model structure with 14 parameters; the last column indicates in which model version the parameter is used, 8p stands for the original version with 8 parameters, > 8p stands for all model versions with more than 8 parameters.

Name	Unit	Default value	Model version	Best $N_{\rm Q}$	Best f _B
a _i	$mm d^{-1} °C^{-1}$	11.5	> 8p	6.3	3.5
a _{sw}	mm d ⁻¹ °C ⁻¹	6.6	12p	5.8	5.5
a _{ss}	mm d ⁻¹ °C ⁻¹	6.6	12p	3.1	2.6
a _s	mm d ⁻¹ °C ⁻¹	6.6	8p,11p	_	-
k _i	$mm d^{-1}$	4.7	> 8p	3.5	-
k _s	$mm d^{-1}$	5.2	> 8p	5.9	-
A	mm	2147	> 8p	1038	-
log(k)	$\log(s^{-1})$	-9.9	> 8p	-7.8	-
β	m ^{4/3} s ⁻¹	301	> 8p	20232	_
$\eta_{\sf s}$	-	-	14p	0.38	-
γ_{p}	% m ⁻¹	0.031	8p	-	-
ρ	°C m ⁻¹	-0.0065	8p	_	_
$ ho_{w}$	°C m ⁻¹	-0.0065	>11p	-0.0013	-0.0013
$ ho_{ m s}$	°C m ⁻¹	-0.0065	>11p	-0.0041	-0.0041
γ _r	% (100 m) $^{-1}$	3.1	>11p	-2.5	_
γ_{w}	% (100 m) $^{-1}$	3.1	>11p	7.5	6.7
γ_{s}	% $(100 \mathrm{m})^{-1}$	3.1	> 13p	6.4	0.02



	-
	_
	r
Table 2. Prior and updated parameter ranges for the generation of random parameter sets;	100
if there is no updated range, this implies that the point mass balance data does not contain	000
information for this parameter; the lower limit for ρ_w and ρ_s is the dry adiabatic lapse rate (e.g.	5
De Walle and Rango, 2008).	-

Name	Unit	Min. prior	Max. prior	Min. updated	Max. updated	Ref. data
a _i	$mm d^{-1} °C^{-1}$	2.0	12.0	3.5	6.5	Annual point bal.
a _{sw}	mm d ⁻¹ °C ⁻¹	0.5	8.0	1.5	8.0	Winter point bal.
a _{ss}	mm d ⁻¹ °C ⁻¹	0.5	8.0	1.5	8.0	Annual point bal.
$ ho_{w}$	°C m ⁻¹	-0.0065	0	-0.0016	0	Winter point bal.
$ ho_{ m s}$	°C m ⁻¹	-0.0065	0	-0.0042	0	Annual point bal.
$\gamma_{\rm w}$	% (100 m) $^{-1}$	0.0	20.0	4.5	18.5	Winter point bal.
γ_{s}	$\% (100 \mathrm{m})^{-1}$	0.0	20.0	0	20.0	Annual point bal.



Table 3. Water balance terms for the original GSM-SOCONT (called "Schaefli05") and the new final calibrated model (1974–1990) and the values extracted from the hydrological atlas of Switzerland, which provides estimated average values of precipitation on a 2 km by 2 km grid (1971–1990, Schwarb et al., 2001)) and of actual evaporation (1973–1992, Menzel et al., 1999) on a 1km by 1km grid; $\Delta S/\Delta t = P - Q - E$ is the storage variation; for the last column, this value is put in brackets since it is computed from data sources referring to different periods; snowfall is indicated relative to total precipitation.

Name	Schaefli05	Best N _Q	Hydro Atlas
Evaporation entire basin	180 mm	150 mm	280 mm
Precipitation (P) total	1940 mm	2670 mm	1950 mm
<i>P</i> glacier	2050 mm	3040 mm	
P non-glacier	1830 mm	2430 mm	
P (glacier)/P (non-glacier)	89%	80%	
$\Delta S / \Delta t$	$-250 \mathrm{mm}\mathrm{a}^{-1}$	$-140 \mathrm{mm}\mathrm{a}^{-1}$	$(-470 \mathrm{mm}\mathrm{a}^{-1})$
Discharge <i>Q</i>	2010 mm	2660 mm	2250 mm
Snowfall glacier	85%	82%	
Snowfall non-glacier	70%	71%	











Fig. 2. Winter accumulation and net annual balance per elevation band simulated for the winter 1979/1980, **(a)** simulated with the non-modified GSM-SOCONT model (Oberwald/Ulrichen input series) with default parameter values; **(b)** improved model with a behavioral parameter set for winter accumulation; grey bars are the 10% and 90% quantiles of the point observations, the red bar is the simulated value.





Fig. 3. Pareto-optimal frontiers (POF) illustrating the trade-off between **(a)** model performance for the simulation of discharge and of annual glacier-wide balance (the mass balance criterion $f_{\rm B}(\theta)$ is normalized by the observed value); **(b)** as (a) but for the model version with seasonal lapse rates, the enclosed figure shows the points on the POF in the parameter space rather than in the objective function space (note: the calibration presented in Schaefli et al., 2005, has $f_{\rm B} = 0.57$ and $f_{\rm Q} = 0.06$.)





Fig. 4. (a) Trade-off between discharge performance criterion and mass balance performance criterion for different versions of the modified GSM-SOCONT; 11p stands for 11 parameters (for parameters see Table 1; note: the figure does not show the tail end of the POF for the 11p model, the best performance for f_Q is 0.06 as for the other model versions. **(b)** discharge simulations corresponding to all parameter sets of (a).





Fig. 5. (a)–(g) Distribution of behavioral parameter sets for the seasonal point mass balances, **(h)** and **(i)** scatter plots of the parameters showing the strongest linear correlations.





Fig. 6. Model simulations: **(a)** Pareto-optimal frontier (POF) of model optimisation considering $N_{\rm Q}$ and $f_{\rm B}$ (computed for the 3 years corresponding to the data of (Funk, 1985)); **(b)** observed discharge and simulations corresponding to all parameter sets on the POF for the beginning of the year 1974 (for parameter values see Table 1); **(c)** as (b) for a later period; **(d)** annual glacierwide simulations of Huss et al. (2008), Schaefli et al. (2005) and of all parameters on the Pareto-curve; shown is also another simulation ($a_i = 3.9$, $a_{sw} = 7.6$, $a_{ss} = 7.3$, $\gamma_w = 0.12$, $\gamma_s = 0.08$ $\rho_w = -0.09$, $\rho_s = -0.33$, units see Table 2) that is acceptable considering only the point mass balance observations; **(e)** ratio between glacier storage change and annual discharge (the simulation of Huss et al. (2008) is divided by the observed annual discharge).

