

On the value of glacier mass balances for hydrological model calibration

Markus Konz^{a,*}, Jan Seibert^{b,c}

^aETH Zurich, Institute of Environmental Engineering, Hydrology and Water Resources Management, Wolfgang-Pauli-Str. 15, 8093 Zurich, Switzerland

^bUniversity of Zurich, Department of Geography, Winterthurerstrasse 190, 8057 Zürich, Switzerland

^cDepartment of Physical Geography and Quaternary Geology, Stockholm University, 106 91 Stockholm, Sweden

ARTICLE INFO

Article history:

Received 7 December 2009

Received in revised form 12 February 2010

Accepted 19 February 2010

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Erwin Zehe, Associate Editor

Keywords:

Model calibration

Glacier mass balances

Monte Carlo simulations

HBV model

ABSTRACT

Hydrological modelling of glacierized catchments is challenging because internal inconsistencies might be hidden due to ice melt which represents an additional source of water. This is even more significant if there are no data available to evaluate model simulations, as is often the case in remote areas. On the other hand, these glacierized catchments are important source regions for water, and detailed knowledge of water availability is a prerequisite for good resource management strategies. An important question is how useful a limited amount of data might be for model applications. Therefore, in this study the predictive power of limited discharge measurements, mass balance observations and the combination of both was analyzed by means of Monte Carlo analyses with multi-criteria model performance evaluation. Ensembles of 100 parameter sets were selected by evaluating the simulations based on a limited number of discharge measurements, glacier mass balance, and the combination of discharge and mass balance observations. Then the ensemble simulation of runoff was evaluated for the entire runoff series. The result indicated that a single annual glacier mass balance observation contained useful information to constrain hydrological models. Combining mass balance observations with a few discharge data improved the internal consistency and significantly reduced the uncertainties compared to parameter set selections based on discharge measurements alone. To obtain good ensemble predictions, information on discharge was required for at least 3 days during the melting season. This demonstrated that the timing of runoff measurements is important for the information contained in these data.

© 2010 Elsevier B.V. All rights reserved.

Introduction

Glacierized headwaters are source regions for many important rivers in the world. These ‘water towers’ supply water to surrounding lowlands and are even more important in arid climates (Viviroli et al., 2007). However, due to the remoteness and the harsh climate, there is often a shortage of continuous discharge measurements, which hinders the assessment of water resources of high mountain catchments. Modelling ungauged catchments is always challenging, but even more so for glacierized catchments because here the catchment water balance is not constrained. The issue of ungauged catchments recently has received increased attention due to the PUB (prediction in ungauged basins) initiative (Sivapalan et al., 2003). While the predictive capabilities of runoff models have increased significantly, calibration of parameters is usually still important and requires adequate data. While, it is difficult to make predictions in totally ungauged basins, it has been demonstrated that a limited number of gaugings can help to achieve acceptable model simulations (Seibert and Beven, 2009). This raises the question of which and what amount of data are needed

to find reliable parameter sets for the application of conceptual models in high mountain areas.

Contributions in the literature which focus on the quantitative aspect of calibration data are limited. Sorooshian et al. (1983) and Yapo et al. (1996) suggested that discharge data series in the order of years are needed for reliable model calibration. This would mean that it is in practice not possible to collect the necessary calibration data for applying a model for a previously ungauged basin. On the other hand, while not perfect, a limited amount of gaugings could still contain enough information to select acceptable model parameter values in the case of ungauged basins. Seibert and Beven (2009) have shown for a set of Swedish catchments that even a limited amount of measured discharge data were sufficient to find good parameter sets for ensemble predictions of discharge.

This approach to allow predictions in ungauged sites by measuring runoff at a limited number of points in time or for shorter periods has been tested by others. Perrin et al. (2007) showed that using about 100–350 observation days spread randomly over a longer time period provided robust parameter estimates and that results hardly improved when more days for model calibration were included. Prior knowledge in the form of regionalized parameter estimates in combination with a few observation days can deliver reliable parameter estimates (Rojas-Serna et al., 2006).

* Corresponding author. Tel.: +41 44 633 21 30.

E-mail address: markus.konz@ifu.baug.ethz.ch (M. Konz).

Another option could be to also collect further meaningful variables such as soil moisture or groundwater levels in order to constrain models (e.g. Seibert, 2000; Binley and Beven, 2003). Juston et al. (2009) demonstrated for a catchment in Central Sweden that a limited number of stream flow and groundwater gaugings can provide basically the same information as continuous time series.

In alpine catchments, glacier mass balances can be used as an additional source of information. However, a direct comparison of measured and modelled mass balance can be difficult, because modelled values usually refer to an areal mean, while direct measurements are point values covering only part of the catchment and a comparison, thus, can be biased by local peculiarities dominating the accumulation and ablation processes. In particular, regions where the mass balance is influenced by small-scale topographic undulations or redistribution of snow (due to wind or avalanches) may exhibit broader deviations between modelled and observed patterns (Paul et al., 2009). However, the mean mass balance over an entire glacier or a sufficiently large section of a glacier can usually be reproduced accurately (Paul et al., 2009; Koboltschnig et al., 2008). Konz et al. (2006, 2007) have shown that glacier mass balances contain important information that improve the reliability of calibrated model parameters in poorly gauged Himalayan catchments. In fact, glacier mass balance measurements are beneficial for the evaluation of internal consistency of hydrological models. Since glacier melt is an additional source of water besides precipitation, the runoff models in glacierized catchments possibly underestimate basin precipitation and partly compensate this error by overestimating glacier melt, or vice versa. Schaeffli et al. (2005) presented a conceptual glacio-hydrological model with daily resolution for the glacierized Rhone catchment in Switzerland in which simulated glacier melt was validated using glacier mass balance observations. Braun and Aellen (1990) already presented a study where direct single point mass balance observations at different elevations of the Aletsch Glacier (Switzerland) were employed for an additional performance criterion for runoff modelling.

Compared to continuous discharge measurements, glacier mass balance observations do not require the construction of gauging stations. Thus, the observations are especially suited for data collection in ungauged remote areas, where it is complicated to install and maintain gauging stations.

In glacierized catchments one can use glacier mass balance data in addition to discharge data to find appropriate parameter sets. In this study we explored the information contained in glacier mass balance data compared to that of a limited number of discharge measurements. Well-observed glacierized catchments in Austria and Switzerland were treated as if they were ungauged, and subsets of measured discharge data and glacier mass balances were taken for calibration of the HBV model. A Monte Carlo approach was used to generate 10,000 model runs, and ensemble mean runoff data using acceptable parameter sets were selected to predict runoff. These sets were identified based on model performance for randomly selected date combinations of discharge observations within a single year and glacier mass balance simulations. This approach is similar to that used by Seibert and Beven (2009), but with using glacier mass balances as an additional selection criterion.

Methods

Study catchments

Three catchments located in Austria and Switzerland were selected for this study with different glacier coverage and catchment areas (Table 1). The largest test catchment, the Venter Ache catchment, is located in the southern Oetztal Alps in Austria and has a catchment area of 165 km². The Vernagtferner catchment

Table 1
Characteristics of the three study sites.

	Area (km ²)	Glacier area (%)	Altitude ranges (m a.s.l.)
Vernagtferner	11.4	78	2640–3627
Venter Ache	165	38	1891–3768
Gletsch	39.9	52	1761–3600

(11.4 km²) is a subcatchment of the Venter Ache catchment. Discharge and meteorological data were available for the station in Vent (1891 m a.s.l., Zentralanstalt für Meteorologie und Geophysik, Vienna, for meteorological data and Federal Ministry of Agriculture, Forestry, Environment and Water Management, Vienna, for discharge data) and Vernagtbach (2640 m a.s.l.). Long-term meteorological, glaciological and hydrological observations of the Vernagtferner catchment were provided by the Commission of Glaciology, Bavarian Academy of Sciences. The meteorological variables (air temperature and precipitation) of the Vernagtferner catchment were measured near the gauging station Vernagtbach (2640 m a.s.l.), which is about 1 km downstream from the terminus of the glacier. For Venter Ache we used exclusively the data from the Vent station. Vernagtferner mostly faces to the east, south, or south-west, and topographic shading plays only a minor role during summer (Reinwarth and Escher-Vetter, 1999). Mass balance measurements started in 1965 on Vernagtferner (extent 2006: 8.172 km², 2765–3627 m a.s.l.) and both mass balances of the entire glacier as well as distributed observations for the fixed-date system (1 October–30 September) are available. In this study we used annual mass balances of the entire glacier for the period 1976–1987 for model analysis of Vernagtferner catchment. For the Venter Ache the mass balance of the south-facing Taschachjoch glacier section was used. This mass balance was compared to the simulation of the HBV model for the relevant elevation zones and exposition class south based on area-specific values. The Gletsch catchment is located in the Rhone basin and annual glacier mass balances are available for the Rhone glacier (16.45 km², 2197–3600 m a.s.l.) for the years 1979/1980, 1980/1981 and 1981/1982 (Laboratory of Hydraulics, Hydrology and Glaciology (VAW), ETH Zurich; Huss et al., 2008). The catchment area is 38.9 km² with glacier coverage of 52.2%. Rhonegletscher is south-exposed and a medium-sized valley glacier with a relatively dry local climate at the glacier terminus, and regional advection effects cause higher precipitation amounts in the accumulation area. Discharge and meteorological data are provided by Meteo Swiss for stations outside the catchment, which have been used to generate climatic input data of the Gletsch catchment. The gauging station is located at an altitude of 1761 m a.s.l.

HBV model

The HBV model is a widely used conceptual precipitation–runoff model which simulates catchment runoff based on precipitation, temperature and long-term potential evaporation at a daily time step. The original version of the HBV model was developed starting in the 1970s at SMHI (Swedish Meteorological and Hydrological Institute) (Bergström, 1992). The model was later developed further by various research groups. At ETH Zurich and at the Commission for Glaciology, Bavarian Academy of Sciences, in Munich, for instance, the HBV model was modified for use in Alpine catchments (Braun and Renner, 1992; Hottel et al., 1993; Konz, 2003). The HBV model describes hydrological processes in a simplified way and while the parameters have some physical meaning, in practice they cannot be measured but have to be determined by calibration. An important advantage of the modelling concept is that the runoff can be simulated using only daily temperature and precipitation data, both of which can be measured relatively easily.

In this study a modified version of the “HBV light” model (Seibert, 1997, 1999) was used, in which the snow melt routine was changed slightly and glacier ice melt and accumulation were represented. A spatial discretization in aspect classes South, North and East–West–Horizontal was introduced to represent the effect of different expositions on snow and ice melt. The usual degree-day equation, where snow melt, M_{Snow} [mm d^{-1}], is computed based on a degree-day factor, C_{DD} [$\text{mm d}^{-1} \text{ } ^\circ\text{C}^{-1}$], and the difference between mean daily temperature, T [$^\circ\text{C}$], and a threshold temperature, T_T [$^\circ\text{C}$], was modified. Following Hottelet et al. (1993) the degree-day factor for south-facing slopes was increased by a factor, F_{ASP} , whereas melt on north-facing slopes was reduced by dividing by F_{ASP} (Eq. (1)). Glacier ice melt, M_{Ice} [mm d^{-1}], was simulated using the degree-day method with the degree-day factor being increased by a factor (parameter F_{Ice}) for the melting of ice compared to snow due to the lower albedo (Eq. (2)). Furthermore

$$M_{Snow} = C_{DD} \cdot F_{ASP} \cdot (T - T_T) \quad (1)$$

$$M_{Ice} = C_{DD} \cdot F_{Ice} \cdot F_{ASP} \cdot (T - T_T) \quad (2)$$

For the glacierized part of each elevation zone, glacier melt did not start before all snow was melted. A certain percentage of the accumulated snow was transferred to the glacial storage in each time step. This parameter was fixed to 0.1% which means that after 1 year ~30% of the accumulated snow would be added to the glacial storage. Mass balances of the glacial storage were computed on an annual basis (year starting and ending on October 1). More detailed descriptions of the HBV model can be found elsewhere (Bergström, 1992, 1995; Harlin and Kung, 1992; Seibert, 1999; Konz, 2003; Konz et al., 2006).

The annual mean net specific glacier mass balances were calculated for each orientation class and altitude belt as the difference between specific mass gain (mass gain divided by glacier area of the altitude belt and orientation class) and specific net ablation in order to allow for comparison with observed mass balance measurements.

Model application

In order to test the value of a limited number of stream data measurements, glacier mass balances measurements or combined mass balance and discharge data, we generated 10,000 random parameter sets sampled from uniform distributions (Fig. 1). The limits for the uniform distributions were set based on previous modelling experiences (Schulz, 1999; Konz et al., 2006) (Table 2). The parameter sets were evaluated for different subsets of measured discharge values assuming limited data availability for the respective catchment.

Ten thousand random date combinations of 4, 8, 16, 32, 64, 128 and 258 days out of 365, respectively, were assembled for each of the available hydrological years of the three investigated catchments. For the cases of 4, 8 and 16 days the 10,000 date combinations were selected in a constrained way so that the dates were selected from only the low flow period, the high flow period, the falling limb, the rising limbs and a combination of all periods (2000 date combinations each), in order to ensure a variety of different date combinations. The date combinations for 32 and more days were selected randomly out of the 365 days. Based on evaluation at these randomly selected dates and annual glacier mass-balance data the 100 best parameter sets out of the 10,000 were selected, where ‘best’ was defined in three different ways:

- (1) Runoff at the randomly selected dates using the sum of squared errors as objective function; these parameter sets will in the following be referred to as Q selected parameter sets.
- (2) Glacier mass balance for one particular year (different years were tested) evaluated as volume errors between measured and simulated mass balance (MB selected parameter sets). The measured glacier mass balances of the entire glacier were used for Gletsch and Vernagtferner, whereas specific balances of the south-facing part of Vernagtferner (Taschachjoch) were used to evaluate the simulated mass balances for the corresponding elevation-aspect-zones in the Venter Ache simulations.
- (3) Combination of runoff at randomly selected dates and glacier mass balances for 1 year (always the first year of the simulation period). For the combination the parameter sets were ranked according to the two different criteria, the ranks were summed and the 100 sets with the lowest rank sums were selected (MB + Q selected parameter sets).

Runoff was then simulated for the entire period for each of the 100 parameter sets and an ensemble mean was computed as arithmetic mean. Such ensemble predictions have previously shown to outperform predictions using single parameter sets (Seibert and Beven, 2009). Finally the model efficiency (Nash and Sutcliffe, 1970) was computed based on the runoff observations for the entire period, referred to as global efficiency in the following. This was done for each of the three different parameter set selections (Q, MB and MB + Q) and for each of the randomly selected date combinations.

In the previous analyses, discharge data subsets were selected from single years. Another approach would be to select repeated surveys over several years. To test this, it was assumed that during

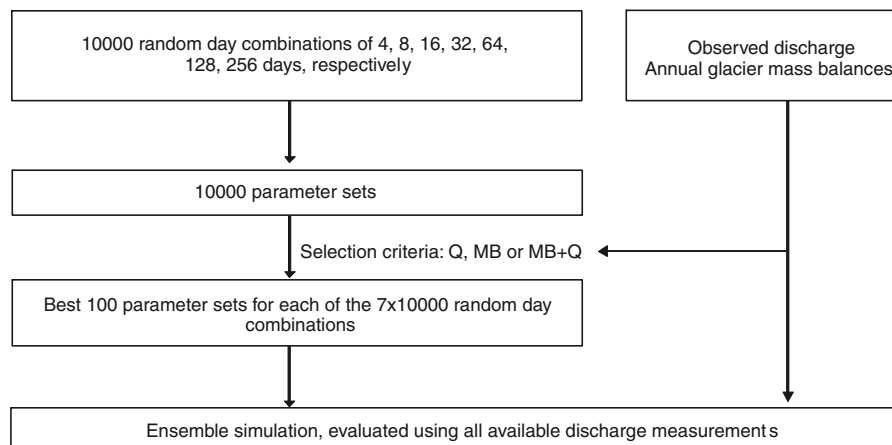


Fig. 1. Analysis scheme to select and evaluate parameter sets based on glacier mass balances (MB), limited discharge measurements (Q) and the combination of both (MB + Q).

Table 2

Parameter ranges for the Monte Carlo analysis.

	Lower limit	Upper limit	Description
TT	−2	2	Threshold temperature
DDF	1	10	Degree-day factor
SFCF	0.4	1.2	Snowfall correction factor
F_{ice}	1.5	2	Ice melt factor
F_{ASP}	1.2	2	Aspect factor
FC	50	200	Maximum soil moisture storage
LP	0.2	1	Soil moisture value above which ET_{act} reaches ET_{pot}
BETA	0.1	6	Parameter that determines the relative contribution to runoff from rain or snow melt
PERC	0	3	Maximum percolation to lower zone
K1	0.01	0.4	Recession coefficient upper storage
K2	0.00001	0.2	Recession coefficient lower storage
MAXBAS	1	7	Runoff transformation parameter

the 11-year period discharge had been measured on the same days each year.

As a benchmark, 100 of the 10,000 parameter sets were also selected randomly and the resulting performance of the ensemble simulation was computed. This was done 10,000 times and the median as well as the 5th and 95th percentiles of the model efficiencies were calculated.

In order to allow a comparison of model performance, two benchmark simulations were used: as a lower benchmark, parameter sets were selected arbitrarily out of the 10,000 without any selection criteria and as an upper benchmark, the best 100 parameter sets according to the entire runoff time series were selected. In both cases an ensemble mean of the simulated runoff series was computed, and the model efficiencies of these series was also computed. The lower benchmark of Vernagtferner exhibits significantly lower model efficiencies than the two other catchments (see Fig. 2). This reflects the high sensitivity of discharge simulations of this highly glaciated catchment (78%) for melt simulations. A wrong representation of melt input immediately causes low model efficiencies.

Results

Predictive quality of limited stream data

The ensemble discharge was calculated for each of the 10,000 random date combinations of 4, 8, 16, 32, 64, 128 and 258 days, using the best 100 parameter sets according to the lowest sum of squared error. The averaged values of medians of the 10,000 Nash–Sutcliffe coefficients (ensemble simulations) of the 11 years for each random date combination ranged from 0.6 when parameter sets were selected based on measurements on only four days to 0.85 when 256 observations were used (Fig. 2). The range of simulation performances due to the different selections of observation days was large, especially when only few days were used. Each of the 11 years was taken to select the parameter sets and the ensemble performance was evaluated against the entire discharge time series. The strongest increase in efficiency could be observed between 16 and 32 observed days. This could be caused partly by our date selection procedure. Four randomly selected days could result in relatively good predictions but the variability of prediction quality largely varies between different date combinations used for model calibration. This variability when using runoff at only a few dates for calibration is lower in the Venter Ache catchment compared to the two other catchments. The randomly selected parameter sets acting as a lower benchmark resulted in a median for the global efficiency of −1.2 for the Vernagtferner, 0.53 for Venter Ache and 0.53 for Gletsch. The range of resulting model performances was large with 5th and 95th percentiles of −2.01 to −0.58 in the case of Vernagtferner, 0.37–0.60 for Venter Ache and 0.37–0.67 for Gletsch. The upper benchmark, i.e., using

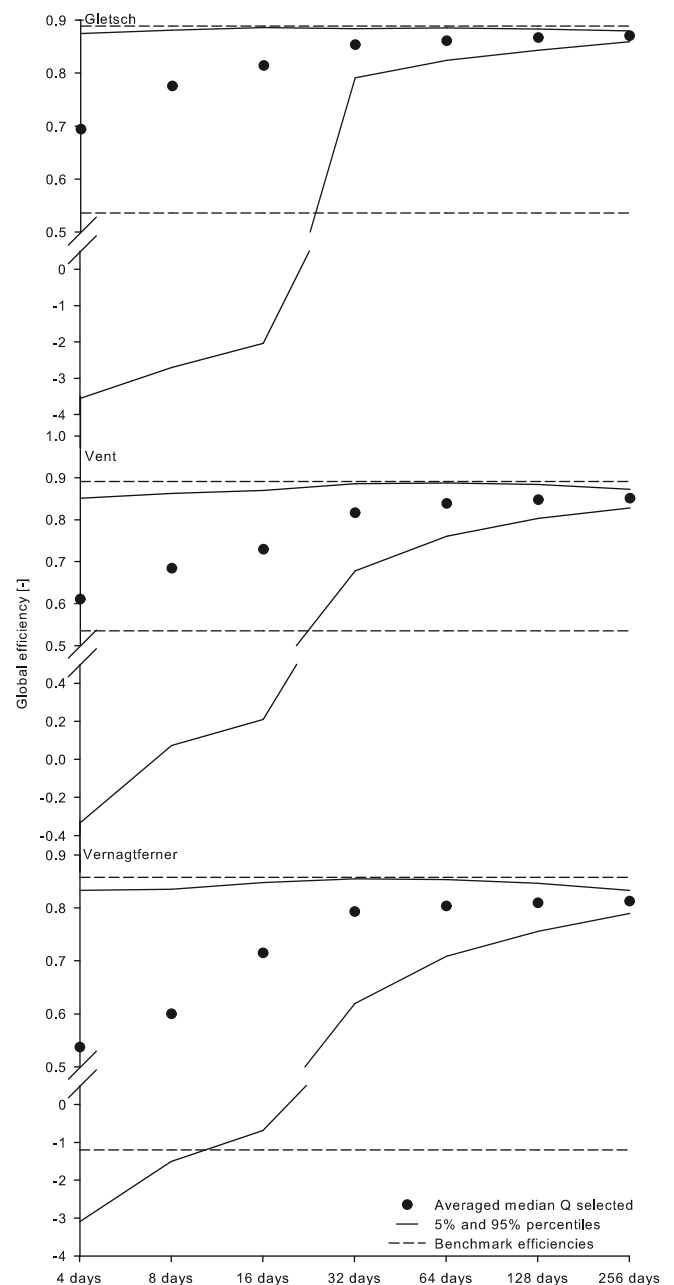


Fig. 2. Analysis of predictive capabilities of limited discharge measurements. The observations were taken from 1 year only, but the analysis was repeated for each of the 11 years and the results were averaged in the figure. Uncertainty ranges (lines) and the averaged medians of discharge global efficiencies (dots) are presented separately in order to improve the readability of the figure.

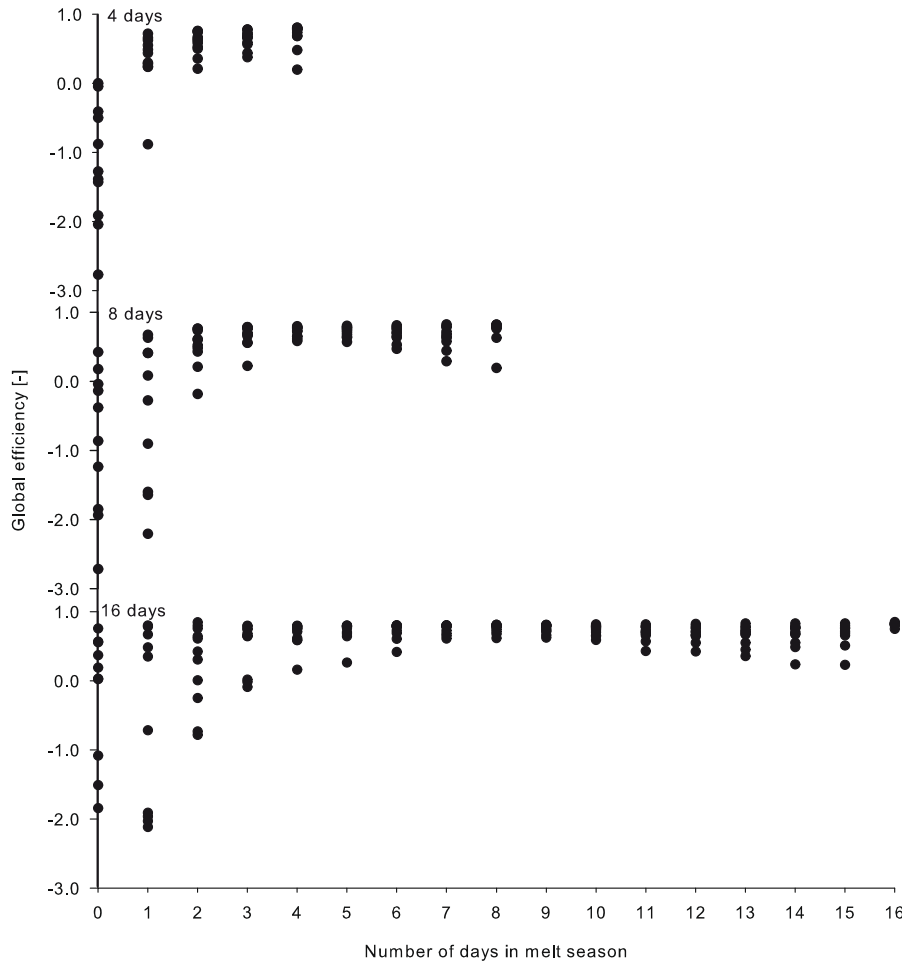


Fig. 3. Model performance according to the season from which the days of the random combinations have been selected. The melt season is defined by a threshold discharge of 0.6 mm/day. The analysis was conducted for the Vernagtferner catchment. The observations were taken from 1 year only, but the analysis was repeated for each of the 11 years. The dots in the figure stand for the different years which were used for parameter set selection.

the 100 best parameter sets with respect to the entire runoff series, resulted in model efficiencies of 0.86 (Vernagtferner), 0.89 (Venter Ache) and 0.89 (Gletsch).

All date combinations resulted in median efficiencies better than the lower benchmark efficiency, which shows that even a limited number of measurements used to rank model parameter sets improved the performance compared to random parameter selections. However, none of the date combinations fully reached the upper benchmark efficiency.

Generally, days selected from the melting season were better suited for parameter set selection than days from the low flow season (Fig. 3). The model performance for the Vernagtferner catchment in general increased when more days were selected from the melt season. This was especially apparent up to using three measurements during the melting season which significantly increased the predictive power regardless of how many measurements were used in total. Interestingly, the performance slightly decreases when all measurements were selected from the melting season. For 8 and 16 days the best results could be achieved if an equal distribution of melting and low flow season days are selected.

Using data from different years helped to further increase model performance (Fig. 4). Using discharge on four days in each of the 11 years allowed parameter sets to be selected which provided an improved ensemble simulation. The variability caused by the selection of the days ranged from -0.3 to 0.85 for 4 days compared to -3.1 to 0.79 for the single year selection. Obviously, the number

of observation days used for the parameter set selection increased by a factor of 11 in this case.

Predictive quality of glacier mass balance data

Parameter sets were ranked based on the difference between simulated and observed glacier mass balances (sum of squared

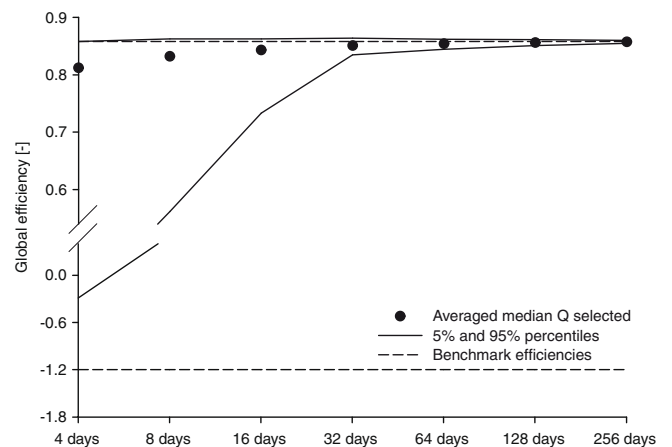


Fig. 4. Analysis of survey discharge measurements using annually repeated measurements of Vernagtferner catchment.

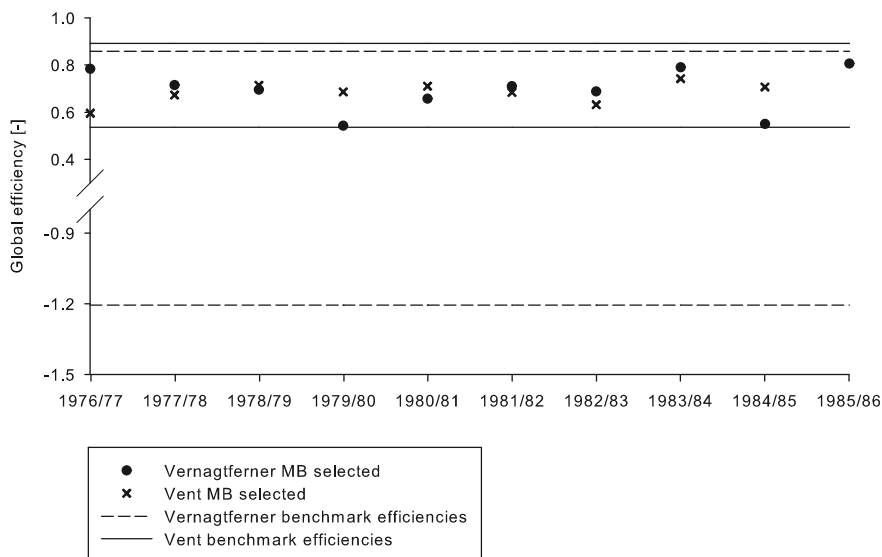


Fig. 5. Ensemble global efficiencies based on MB selected parameter sets and arbitrarily chosen parameter sets without selection criterion.

errors). Only one mass balance was taken for the identification of the 100 parameter sets. The 100 parameter sets were used to compute ensemble mean discharge time series that were evaluated against the entire discharge time series. Fig. 5 depicts the global efficiencies of the MB selected parameter sets taking mass balance data of different years into account for the simulation. These efficiencies were again compared with the benchmark simulations shown as horizontal lines. A single annual mass balance observation allowed parameter sets to be selected which resulted in ensemble predictions with efficiencies between 0.54 and 0.80. This is comparable to the performance of discharge simulations when calibration was based on discharge measured at 256 days during 1 year. The MB selected parameter sets deliver comparable global efficiencies for both the Vernagtferner and the Venter Ache catchments although the percentage glacier coverage of the Venter Ache catchments is considerably smaller. The parameter sets selected by mass balance data only resulted in model efficiencies higher than 0.54 for the Vernagtferner and 0.60 for the Venter Ache catchment.

The additional information from the mass balance data of Vernagtferner significantly improved the global model performance compared to the lower benchmark. For the larger Venter Ache catchment the differences between simulation results of MB selected parameter sets and randomly selected parameter sets (lower benchmark) were less important. However, even here the MB selected parameter sets deliver significantly better discharge simulations than the randomly chosen parameter sets. The best global efficiencies that were achieved with the 10,000 parameter sets were 0.87 for Vernagtferner and 0.88 for Venter Ache. The global efficiencies of the MB selected parameter sets do not attain these values.

Predictive quality of combined mass balance and discharge measurements

Combining discharge and mass balance measurements helped to narrow the uncertainty ranges significantly, especially for a small number of observed discharge data (Fig. 6). The probability of achieving good discharge simulations increases compared to Q selected parameter sets. That is true for all three catchments with different glacier coverage. For this analysis we used the first year's mass balance measurements for the different periods used for the respective catchment. Survey measurements with annually re-

peated discharge measurements show the same behaviour when adding a single mass balance observation as additional selection criterion.

Discussion

Discharge data alone were more meaningful in the larger basin with lower percentage glacier coverage (Venter Ache). Here, the reliability range was narrower for four measurements as in the two other catchments. With the increasing influence of glacier melt on runoff generation the predictive quality of limited (4–16) measurements diminished. The season in which the measurements were conducted plays a crucial role in the information content of the data. Days in the melting season were generally more useful for model calibration than days in the low flow season. The days of the melt season contain information about the melting conditions that are important to select the parameters of the snow and glacier routine. However, a limited number of low flow days helped to improve the calibration. This could be due to the information losses of low flow conditions if melting conditions alone are used for calibration. The information content of the low flow season is particularly important for adjusting the storage coefficients of the model's runoff generation routine. This effect would be even more pronounced if objective functions other than the Nash–Sutcliffe criterion that focus also more on low flows (e.g. Nash–Sutcliffe criterion computed on logarithmic runoff values) were used to define the model performance.

It is encouraging that MB selected parameter sets enabled acceptable discharge predictions to be made even in catchments with relatively small glacier extent. The mass balance information is the integral value over all accumulation and ablation processes of 1 year including snow cover processes. In high alpine catchments snow plays a crucial role in the runoff generation. In the HBV model ice melt is simulated similarly as snow melt, with the only difference of the parameter F_{ice} . For calibration of such a conceptual approach as used in HBV that is applied on ice and on snow with only one parameter difference, MB data are informative not only for the glacierized part of the catchment but also for the snow accumulation and melt in the entire catchment. This explains the good performance of the MB selected parameter sets for Venter Ache catchment although the glacier coverage is considerably lower than in the two other catchments. Future work could explore

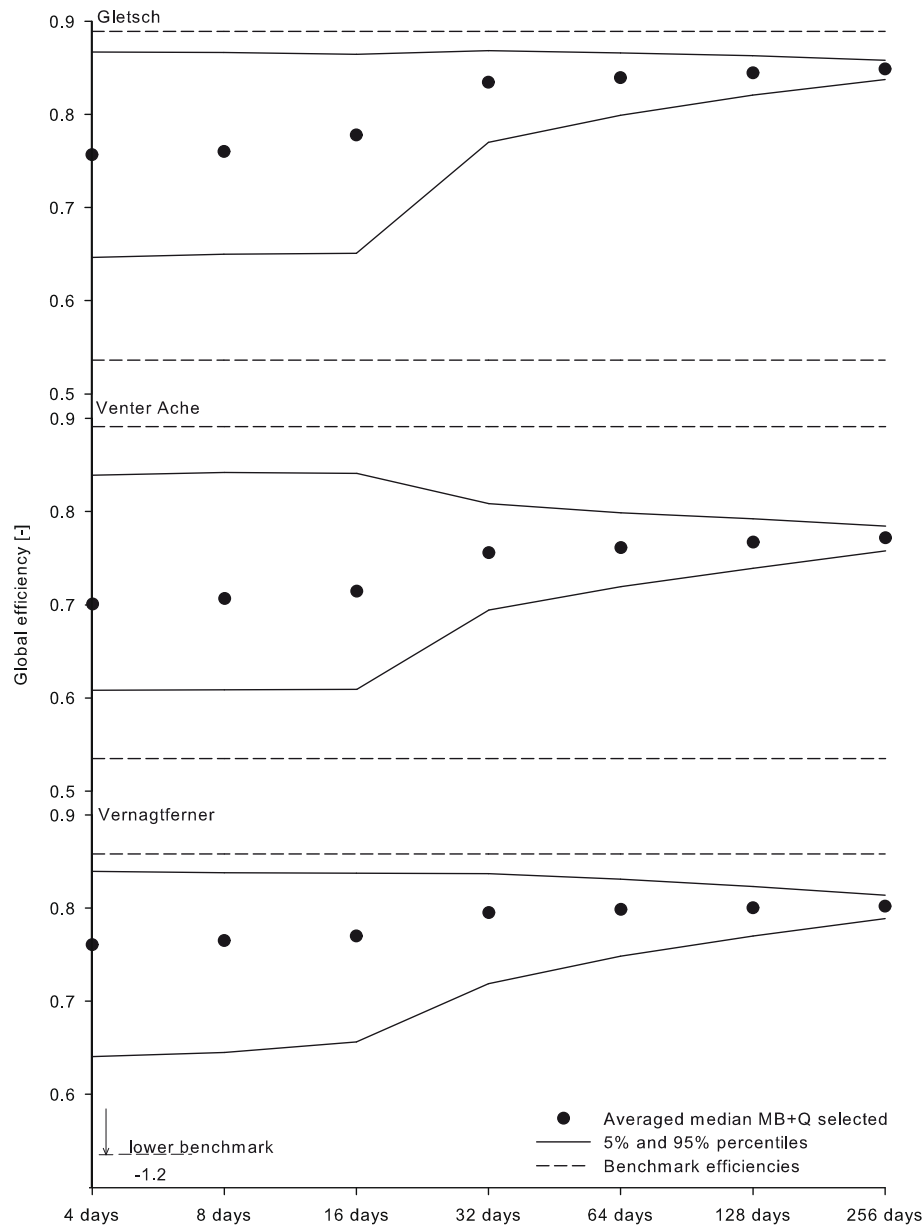


Fig. 6. Analysis of MB + Q selected parameter sets. The observations were taken only from 1 year only, but the analysis was repeated for each of the 11 years and the results were averaged in the figure. Only the first year's mass balance observation was taken for the analyses.

catchments with even less glacier coverage. The importance of mass balance data for hydrological model calibration can also be seen with respect to internal model consistency (Braun and Aellen, 1990; Koboltschnig et al., 2008). The MB selected parameter sets seem to be less constrained and might provide good results for the 1 year used for the selection despite of internal inconsistencies. In other words, good mass balance simulations might be obtained for the wrong reasons. Ablation and accumulation could be over- or underestimated but volume errors are still low. These parameterizations then can lead to significant errors when simulating other years. Parameter sets which have been selected exclusively from the first year were evaluated against the mass balance observations of the entire period (Fig. 7). The combination of MB and Q delivers volume errors comparable to MB selected parameter sets or better. That is an interesting result because one would expect that the MB selected parameter sets should provide the best mass balance simulations. However, the relatively large mean absolute volume error of the MB selected parameter sets shows that the

information content of a single mass balance is limited and the variability of mass balance simulations is large over the years using the parameter sets selected from the first year. As the parameter sets have been selected from the first year's mass balance observation, the parameters are suited for the climatic conditions of this year and it is not surprising that all 100 parameter sets deliver almost perfect mass balance simulations with volume errors between 0.6 and 21 mm. These parameter sets are also suited to reproduce the second year's mass balance, however, from the third year onwards the performance is much poorer for at least some of the 100 parameter sets (Fig. 8). If additional discharge data are taken for parameter set selection (MB + Q selected) the internal consistency can be improved and more robust parameter sets are selected, which are better suited to handle the different climatic conditions of the 10 years and not only the first year's conditions (Fig. 7). Despite this finding simulated discharge using parameter sets selected based on 1 year's MB provides still a better estimate than a random selection of parameter values (Fig. 5). In other

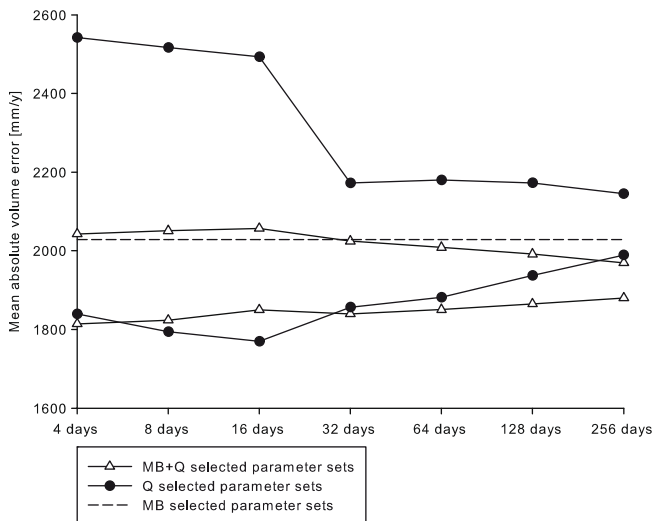


Fig. 7. Mean absolute volume errors of parameter sets selected by discharge data (Q), the combination of discharge data and mass balances (MB + Q) or by mass-balance data (MB). For volume errors simulated by Q selected and MB + Q selected the 5% and 95% percentiles of the different day selections are shown.

words, there is an added value of the mass balance data even if the efficiencies with values between 0.5 and 0.81 cannot be judged as good (Rango, 1992). Mass balance data provide information to reject the worst parameter sets but good mass balance simulations could be achieved for the wrong reasons which can lead to poorer model performances during years different from the one used for parameter selection. Adding even few discharge data improves the internal consistency and with it the robustness of the selected parameter sets as shown in good performance measures over the entire observation period. The mass balance data can provide good quantitative information on the input to the system in terms of melt and precipitation, whereas a limited number of discharge data can give relevant information on the hydrological response of the catchment. In that sense low flow data can be used to adjust the storage capacity and recession parameters of the catchment.

Glacier mass balance field measurements require sound methodological expertise and experienced fieldworkers. Remotely sensed glacier and snow information can be promising alternatives to ground data. Long-term glacier mass balances derived from dif-

ferent digital elevation models (DEMs) can also be used for model calibration. These long-term observations may contain even more information than annual mass balance data because they represent the integral of input and melt over several seasons including different climatic conditions. Additional to quantitative glacier mass balance data the applicability of qualitative (snow: yes–no) MODIS snow cover data for calibration should be taken into account. MODIS data are freely available all over the world and no additional DEMs are required. Parajka and Blöschl (2008) have recently demonstrated the applicability of MODIS data for model calibration.

Conclusions

In this study we systematically analyzed the predictive capabilities of a model where parameter sets were selected based on a limited number of discharge measurements, glacier mass balance observations and the combination of both. A single glacier mass balance observation for 1 year allowed the selection of good parameter sets that provided acceptable discharge simulations. The number of stream flow gaugings needed to obtain acceptable model simulations was also rather low, but, as expected, it was proved important that these gaugings were taken during the melting season. In fact, as few as three observed days of the melting season significantly improved the discharge simulations based on the parameter sets selected by the limited number of observations. Combining mass balance observation and a few discharge measurements significantly reduced the uncertainties of discharge simulations and improved the internal consistency of the modeling results. Annually repeated discharge measurements of few days (survey) also delivered reliable results. This is especially important if the glaciers are hard to access and mass balance observations are not available. As remotely sensed mass balance estimations become more accurate and more widely available, the importance of glacier mass balance data for parameter selections for runoff simulations is encouraging. Future work should evaluate the value of remotely sensed glacier mass balance data to constrain hydrological models.

Acknowledgements

We are grateful to the Commission of Glaciology of the Bavarian Academy of Sciences for providing the data of the Venter Ache and

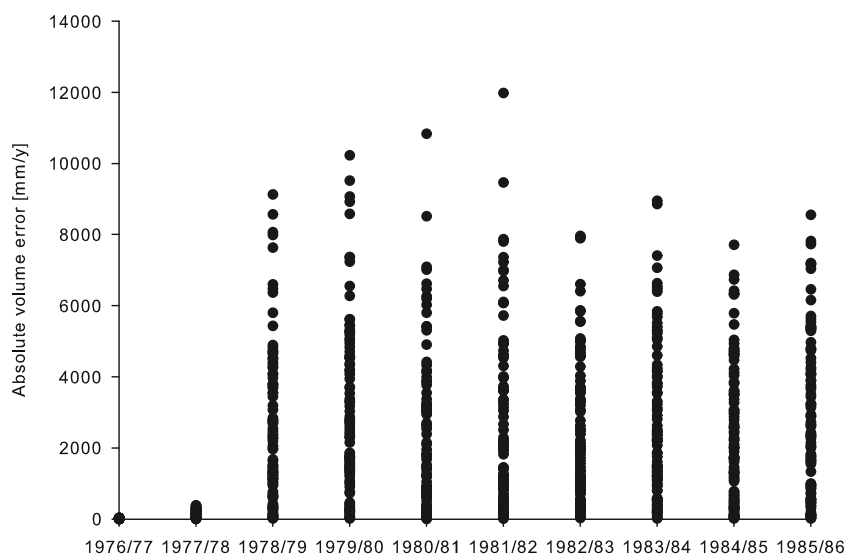


Fig. 8. Variability of the absolute volume errors of the 100 best MB selected parameter sets for the 10 mass balance years. Each dot stands for a parameter set.

Vernagtferner catchments. Dr. David Finger, ETHZ, supported this work with discussions and provided data of the Gletsch catchment. Comments of Bettina Schaeffli and Ludwig Braun helped to improve the paper.

References

- Bergström, S., 1992. The HBV Model: Its Structure and Applications. Swedish Meteorological and Hydrological Institute (SMHI), Hydrology, Norrköping. 35 pp.
- Bergström, S., 1995. The HBV model. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, Colorado, USA, pp. 443–476 (Chapter 13).
- Binley, A., Beven, K., 2003. Vadose zone flow model uncertainty as conditioned on geophysical data. *Ground Water* 41, 119–127.
- Braun, L.N., Aellen, M., 1990. Modelling discharge of glacierized basins assisted by direct measurements of glacier mass balance. In: Lang, H., Musy, A. (Eds.), *Hydrology in Mountainous Regions I. Hydrological Measurements, the Water Cycle* (Proceedings of the Lausanne Symposia, 1990). IAHS Publication No. 193. IAHS Press, Wallingford, pp. 99–106.
- Braun, L.N., Renner, C.B., 1992. Application of a conceptual runoff model in different physiographic regions of Switzerland. *Hydrol. Sci.-J.* 37, 3, 217–231.
- Härlin, J., Kung, C.S., 1992. Parameter uncertainty and simulation of design floods in Sweden. *J. Hydrol. (Amsterdam)* 137, 209–230.
- Hottelet, Ch., Braun, L.N., Leibundgut, Ch., Rieg, A., 1993. Simulation of Snowpack and Discharge in an Alpine Karst Basin. IAHS Publication No. 218, pp. 249–260.
- Huss, M., Bauder, A., Funk, M., Hock, R., 2008. Determination of the seasonal mass balance of four Alpine glaciers since 1865. *J. Geophys. Res.* 113, F01015. doi:10.1029/2007JF000803.
- Juston, J., Seibert, J., Johansson, P.O., 2009. Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment. *Hydrol. Process* 23 (21), 3039–3109.
- Koboltschnig, G.R., Schöner, W., Zappa, M., Kroisleitner, C., Holzmann, H., 2008. Runoff modelling of the glacierized alpine upper Salzach basin (Austria): multi-criteria result validation. *Hydrol. Process.* 22, 3950–3964.
- Konz, M., 2003. HBV3-ETH9 User's Manual. Internal Report Bavarian Academy of Sciences, Commission of Glaciology, Munich.
- Konz, M., Braun, L., Grabs, W., Shrestha, A., Uhlenbrook, S., 2006. Runoff from Nepalese Headwater Catchments – Measurements and Modelling. IHP/HWRP-Berichte, 4, Koblenz, 160 pp.
- Konz, M., Uhlenbrook, S., Braun, L., Shrestha, A., Demuth, S., 2007. Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater. *Hydrol. Earth Syst. Sci.* 11, 1323–1339.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, Part I – a discussion of principles. *J. Hydrol.* 10, 282–290.
- Parajka, J., Blöschl, G., 2008. The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models. *J. Hydrol.* 358, 240–258.
- Paul, F., Escher-Vetter, H., Machguth, H., 2009. Comparison of mass balances for Vernagtferner, Oetzal Alps, as obtained from direct measurements and distributed modelling. *Ann. Glaciol.* 50 (50).
- Perrin, C., Oudin, L., Andreassian, V., Rojas-Serna, C., Michel, C., Mathevet, T., 2007. Impact of limited streamflow data on the efficiency and the parameters of rainfall–runoff models. *Hydrol. Sci. J.* 52, 131–151.
- Rango, A., 1992. Worldwide testing of the snowmelt runoff model with application for predicting the effects of climate change. *Nord. Hydrol.* 23, 155–172.
- Reinwarth, O., Escher-Vetter, H., 1999. Mass balance of Vernagtferner, Austria, from 1964/65 to 1996/97: results for three sections and the entire glacier. *Geogr. Ann.* 81A (4), 743–751.
- Rojas-Serna, C., Michel, C., Perrin, C., Andreassian, V., 2006. Ungauged Catchments: How to Make the Most of a Few Streamflow Measurements? Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment – MOPEX, vol. 307. IAHS publication, pp. 230–236.
- Schaeffli, B., Hingray, B., Niggli, M., Musy, A., 2005. A conceptual glaciohydrological model for high mountainous catchments. *Hydrol. Earth Syst. Sci.* 9, 95–109.
- Schulz, M., 1999. Bestimmung der Wasserhaushaltsgrößen ausgewählter hochalpiner Einzugsgebiete mittels Messung und Simulation. Diploma thesis at the Institute of Geography, Ludwig Maximilians Universität Munich, Germany (unpublished).
- Seibert, J., 1997. Estimation of parameter uncertainty in the HBV model. *Nord. Hydrol.* 28, 247–262.
- Seibert, J., 1999. Regionalisation of parameters for a conceptual rainfall–runoff model. *Agr. Forest Meteorol.* 98, 279–293.
- Seibert, J., 2000. Multi-criteria calibration of a conceptual rainfall–runoff model using a genetic algorithm. *Hydrol. Earth Syst. Sci.* 4 (2), 215–224.
- Seibert, J., Beven, K.J., 2009. Gauging the ungauged basin: how many discharge measurements are needed? *Hydrol. Earth Syst. Sci.* 13, 883–892.
- Sivapalan, M., Schaake, J., Sapporo, J., 2003. PUB Science and Implementation Plan. V5. <<http://pub.iwmi.org/UI/Images/PUBSciencePlanV5.pdf>>. IAHS Decade on Predictions in Ungauged Basins (PUB).
- Sorooshian, S., Gupta, V.K., Fulton, J.L., 1983. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall–runoff models: Influence of calibration data variability and length on model credibility. *Water Resour. Res.* 19, 251–259.
- Viviroli, D., Dürr, H.H., Messerli, B., Meybeck, M., Weingartner, R., 2007. Mountains of the world, water towers for humanity: typology, mapping, and global significance. *Water Resour. Res.* 43, W07447. doi:10.1029/2006WR005653.
- Yapo, P.O., Gupta, H.V., Sorooshian, S., 1996. Automatic calibration of conceptual rainfall–runoff models: sensitivity to calibration data. *J. Hydrol.* 181, 23–48.