Hydrological Sciences Journal
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t911751996

Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments

B. Hingray[a]; B. Schaefli[b]; A. Mezghani[a]; Y. Hamdi[c]
[a] Laboratoire d'Etude des Transferts en Hydrologie et Environnement, Grenoble, Cedex, France [b] Water Resources Section, Delft University of Technology, GA, Delft, The Netherlands [c] Ecole Nationale d'Ingénieurs de Gabés, Gabés, Tunisia

Online publication date: 20 August 2010

To cite this Article Hingray, B. , Schaefli, B. , Mezghani, A. and Hamdi, Y.(2010) 'Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments', Hydrological Sciences Journal, 55: 6, 1002 — 1016
To link to this Article: DOI: 10.1080/02626667.2010.505572
URL: http://dx.doi.org/10.1080/02626667.2010.505572

Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments

B. Hingray[1], B. Schaefli[2], A. Mezghani[1] & Y. Hamdi[3]

[1]*Laboratoire d'Etude des Transferts en Hydrologie et Environnement, CNRS, BP 53, F-38041 Grenoble Cedex 9, France*
benoit.hingray@gmail.com

[2]*Water Resources Section, Delft University of Technology, TU Delft, NL-2600 GA Delft, The Netherlands*

[3]*Ecole Nationale d'Ingénieurs de Gabés, Gabés 6011, Tunisia*

Received 15 October 2009; accepted 28 May 2010; open for discussion until 1 March 2011

Citation Hingray, B., Schaefli, B., Mezghani, A. & Hamdi, Y. (2010) Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments. *Hydrol. Sci. J.* **55**(6), 1002–1016.

Abstract This paper presents a calibration framework for a precipitation–runoff model for flood prediction in a mesoscale Alpine basin with discharges strongly influenced by hydraulic works. The developed methodology addresses two classical hydrological calibration challenges: computational limitations to run optimization algorithms for distributed hourly models and the absence of concomitant meteorological and natural discharge time series. The presented processes-oriented, multi-signal approach is based on hydrological data from a variety of sources and for different periods, corresponding to various spatial scales. The model parameters are calibrated by sequentially minimizing differences between observed and simulated values for different hydrological signals and signatures such as: (a) the phase of precipitations, (b) the time evolution of point-scale snow heights, (c) the mean inter-annual cycle of daily discharges, and (d) timing of snowmelt-induced spring runoff. We compare the model performance to a benchmark model obtained by simply using the globally optimal parameter values from the nearest gauged and non perturbed catchment. For prediction of flow seasonality and also extreme events, the calibration methodology outperforms the benchmark.

Key words distributed model; hydrological signatures; step-wise calibration; multi-signal calibration; process-oriented calibration; Alpine hydrology; ungauged basin; anthropogenic influence

Calage d'un modèle basé sur des signatures pour la prédiction hydrologique dans des bassins alpins de méso-échelle

Résumé Nous présentons une méthodologie pour calibrer les paramètres d'un modèle hydrologique distribué appliqué sur un bassin alpin aux débits fortement influencés par divers ouvrages hydrauliques. Orientée processus et multi-signal, elle repose sur des données de qualités et d'origines variées, provenant de périodes différentes et correspondant à des échelles spatiales diverses. Les paramètres sont calibrés de façon séquentielle en minimisant les différences entres valeurs observées et simulées pour différents signaux et signatures hydrologiques dont: (a) la nature des précipitations, (b) l'évolution temporelle des hauteurs de neige, (c) le cycle interannuel moyen des débits journaliers, et (d) la synchronisation des débits de hautes eaux. Cette approche offre une performance modèle meilleure que celle obtenue avec une approche globale mono-objectif. Pour les bassins perturbés, la performance est supérieure à celle obtenue avec des paramètres estimés sur le bassin non perturbé le plus proche.

Mots clefs modèle distribué; signatures hydrologiques; calibration séquentielle; calibration multi-signal; calibration orientée processus; bassin alpin; bassin non jaugé; aménagements hydrauliques

1 INTRODUCTION

Hydrological prediction in any real-world catchment faces classical challenges, such as poor data availability and apparently intractable parameter identification problems with computationally intense distributed models. In this paper, we present a real-world, mesoscale Alpine case study, where these challenges are exacerbated by the strong perturbation of natural discharges through the hydraulic infrastructure, but also by known meteorological phenomena, such as sudden temperature decrease during precipitation events, that are not captured by the measurement stations. Thus, we had to deal with the problem of how to extract information about the natural system behaviour from

ISSN 0262-6667 print/ISSN 2150-3435 online

doi: 10.1080/02626667.2010.505572
http://www.informaworld.com

the discharge records and how to make use of it to calibrate an hourly discharge model for which one model run takes 3 h on a personal computer.

Hydrological models are generally calibrated on observed discharge, which requires meteorological time series (precipitation, temperature) concomitant with the observed discharge (for a review, see Gupta *et al*., 2005). A fundamental problem arises in ungauged basins (Sivapalan *et al*., 2003a); however, many catchments are not completely ungauged: there is some information about the dominant hydrological processes but this information has first to be extracted somehow from the available data and then used for model calibration (e.g. Winsemius *et al*., 2009).

Therefore, we developed a signature-based model calibration approach (see, e.g. Sivapalan *et al*., 2003b; Winsemius *et al*., 2009): the core of such an approach is to not calibrate the model directly on observed discharge (which in our case would be too strongly perturbed), but to identify in all available data hydrologically meaningful patterns about the system behaviour (see Yilmaz *et al*., 2008) and to calibrate the model on these signatures. Signatures have a long tradition in hydrology (a well-known signature is the flow-duration curve), but their value for model calibration was perhaps underestimated during the recent phase of intense development of automatic model calibration procedures (e.g. Vrugt *et al*., 2010). The current resurgence of interest (Gupta *et al*., 2008; Yilmaz *et al*., 2008; Winsemius *et al*., 2009) is related to the strong need for new tools to understand how hydrological models work, in particular in view of the International Association of Hydrological Sciences Decade on Prediction in Ungauged Basins (2003–2012) (Sivapalan *et al*., 2003a).

A more classical solution for setting up hydrological models in ungauged or poorly gauged catchments is parameter regionalization (Bárdossy, 2007; Viviroli *et al*., 2009), in which calibrated model parameters are transposed to ungauged catchments, assuming that spatial proximity or landscape and climatic similarity lead to a similar hydrological response, and that this response can be simulated with similar hydrological parameters.

The calibration method that we present here for the hourly discharge simulation in the Upper Rhône River (URR) basin, Switzerland, is a mixture of parameter regionalization and signature-based calibration. Where possible, we calibrate the model parameters on hydrological time series (e.g. discharges and snow height series), and signatures (e.g. mean annual discharge regime or patterns in the snowmelt dynamics). If unavoidable (complete absence of data), we regionalize the parameters. A key feature here is that we do not regionalize individual parameter values, but parameter groups that encode a dominant hydrological process. This is a pre-requisite to obtain hydrologically meaningful parameter sets (see Bárdossy, 2007).

The paper is organized as follows: Section 2 describes the context of the simulation study, flood estimation in a mesoscale Alpine catchment that is largely ungauged and has strongly perturbed sub-basin discharge; Section 3 details the different components of the hydrological model and the methodology developed for the calibration of its parameters; the results and model performance are presented in Section 4; followed by a discussion and conclusions in Section 5.

2 CASE STUDY – CONTEXT AND DATA

2.1 Study area

The URR basin covers the southwestern part of the Swiss Alps. The most important basin characteristics, as well as the relevant sub-basins for design flood estimation, are given in Table 1. The mean annual precipitation is around 1050 mm, the mean annual temperature at mean altitude is 0.8°C.

Some of the URR sub-basins (namely those in Table 1) have streamflow gauging stations (see http://www.bafu.admin.ch/hydrologie), but almost all the URR tributaries have large accumulation reservoirs for hydropower production, or water intakes from

Table 1 Characteristics of the main URR sub-basins. Each basin has itself a number of sub-basins; some of which are nested.

Basin	Mean elevation (m a.s.l.)	Area (km^2)	Glaciation (%)	Number of sub-basins	Number of RHHUs
Viège at Viège	2660	778	29.5	4	41
Rhône at Brig*	2370	913	24.2	7	60
Rhône at Sion	2310	3349	18.4	22	200
Rhône at Branson	2250	3728	16.8	24	213
Rhône at Porte-du-Scex	2130	5220	14.3	35	299

* the basin is considered as non perturbed over the 1982–2001 period.
Note: RHHU: relatively homogenous hydrological units (see Section 3).

nearby reservoirs (see http://www.swissdams.ch). With a total storage volume of 1200 × 10^6 m^3 (roughly 20% of the annual precipitation of the URR basin), the accumulation reservoirs have a significant impact on the downstream hydrological regimes (Fig. 1). Details of the day-to-day reservoir management are confidential (for financial reasons) and, thus, cannot be used to reconstruct either the inflow into the accumulation reservoirs or the hypothetical natural flows at the downstream gauging stations. Major flood events usually occur in the autumn, at a time when the accumulation reservoirs are potentially all full. During the recent observed flood events, some of the reservoirs had space available and were used to store water, but

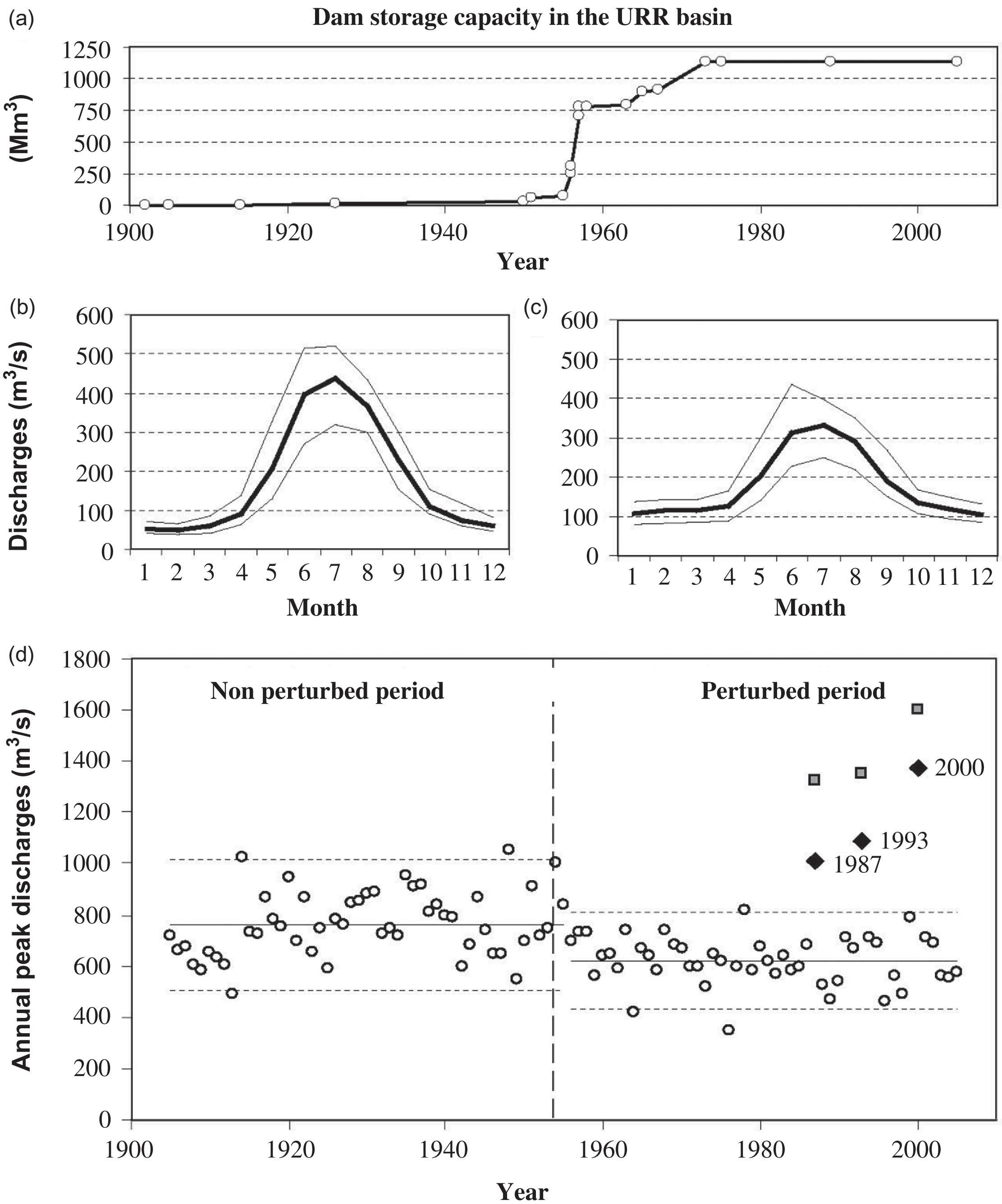

Fig. 1 Effects of water works on the URR basin hydrology: (a) dam storage capacity evolution over the 1905–2005 period; (b) and (c) 10, 50 and 90% percentiles of the mean monthly discharge at Porte du Scex for 1905–1955 and 1955–2005, respectively; and (d) annual maximum peak discharge. For 1987, 1993 and 2000 floods: ◆: observed discharge; □: reconstructed discharge (from Hingray *et al.*, 2009).

such storage volume is not guaranteed by the hydropower companies. Therefore, in agreement with the water management authorities, it was decided to consider the worst-case scenario in which the reservoirs have no delaying effect on floods. This explains why the model is set up here without modelling hydraulic works (for the operational flood forecasting model version including them, see Hernández *et al.*, 2009). Further combined with a stochastic weather generator developed for the URR basin (Mezghani & Hingray, 2009), the model was used to generate a number of design flood scenarios under pseudo-natural conditions (Hingray *et al.*, 2006b).

2.2 Data

In mountainous regions, the main meteorological variables that govern the intensity of floods are precipitation and temperature. The latter determines the amount of snow- or glacier melt and how much precipitation falls as snow and is temporarily stored on the hillslopes. Depending on the configuration, the amount of rainfall-induced runoff can be increased or reduced significantly.

In the URR basin, the temporal variability of precipitation and temperature is high and the response times of basins are often shorter than 12 hours. The development of a flood prediction simulation tool thus requires a 3-hour (or shorter) modelling time step (Obled *et al.*, 2009). Hourly meteorological variables are available from 1982 to 2001 for 11 automatic weather stations. To benefit from the higher density of daily precipitation stations, data from 37 daily raingauges were used in addition. Note that no meteorological stations exist above 2500 m a.s.l.

Mean daily discharge observations are available since the 1900s at Porte du Scex and since the 1940s for the other main sub-basins listed in Table 1. Hourly discharge series are available since the 1980s. Three major flooding events have occurred in the last 20 years (August 1987, September 1993 and October 2000). They resulted from large amounts of regional precipitation induced by warm and humid air masses arriving from the Mediterranean Sea across the Alps. Based on observed flows and data provided by the hydropower production companies, the Swiss Federal Office for the Environment has reconstructed peak discharges and/or hydrographs (Fig 1(d)) under "hypothetical" natural conditions for several gauging stations of the basin (OFEG, 2002).

For most basins for which discharge observations exist dated before the construction of all major dams in the 1950s, there are either no concomitant observed meteorological time series, or ones at too coarse a temporal resolution (daily instead of hourly). In such cases, the 1941–1956 period, for which non-perturbed discharge data are available, will be used to identify the relevant hydrological signatures for model calibration.

3 METHOD

3.1 Precipitation–discharge transformation model

The discharge simulation is completed with a distributed, hourly version of GSM-SOCONT, a lumped reservoir-based model for discharge simulation in Alpine catchments (Schaefli *et al.*, 2005). The URR basin is modelled with 35 sub-basins, defined according to the locations where discharge estimations are required and to the guidelines proposed by Obled *et al.* (2009). The ice-covered and the ice-free parts of each sub-basin are divided into elevation bands, further referred to as relatively homogeneous hydrological units (RHHUs), covering around 500 m each (Table 1).

For each RHHU, a potential evapotranspiration (PET) series is calculated based on the Penman-Monteith version of Burman & Pochop (1994). Hourly mean areal precipitation (MAP) and temperature (MAT) are estimated as weighted averages of the values observed at raingauges or climatic stations. The MAT is estimated for the mean elevation of each RHHU accounting for the regional elevation–temperature relationships obtained for each time step from locally measured temperatures, or from climatology if necessary (Hingray *et al.*, 2006b). The MAP is supposed to be solid if MAT is smaller than a critical temperature, Tc_1, liquid if MAT is higher than a critical temperature, Tc_2, and mixed otherwise.

The temporal evolution of the snowpack is computed based on these meteorological time series. At each time step and for each RHHU, solid and/or liquid precipitation are added to the snowpack and to its liquid water content respectively. If temperatures are positive, the liquid water content is increased by snowmelt, if they are negative, the snowpack is fed by refreezing water from its liquid water content. The hourly amount of snowmelt (or of refreezing water) is estimated with a temperature-index model (TIM) from the positive (resp. negative) temperature degrees for the considered hour. This TIM approach is similar to more classical degree-day approaches. To account for latent heat transfer in cases of rain-on-snow events, the temperature-index factor (TIF) for snowmelt is increased proportionally to the amount of rainfall. If the liquid water content of the snowpack reaches its retention capacity, the snowpack releases an "equivalent rainfall". The retention capacity is assumed to be proportional to the water equivalent of snow and a constant saturation ratio (see a similar approach in Kuchment & Gelfan, 1996). For glaciated RHHUs, ice melt is also estimated from a temperature-index approach. Ice melt only occurs when the glacier surface of the considered RHHU is free of snow.

For glaciated RHHUs, the rainfall and meltwater–runoff transformation is completed through two linear reservoirs: one for ice melt and one for the equivalent rainfall, which is only rainfall in absence of a snowpack. For non-glaciated RHHUs, the equivalent rainfall is separated into infiltration and effective rainfall. They are respectively transformed into runoff components through a nonlinear soil reservoir that produces the slow component of discharge and a linear reservoir for direct runoff. Effective rainfall and actual evapotranspiration are estimated from equivalent rainfall and potential evapotranspiration as a function of the filling rate of the slow reservoir.

The discharge simulated at the outlet of each sub-basin is the sum of the different discharge components produced by elevation bands from glaciated and non-glaciated areas. Discharges produced by the 35 sub-basins are routed in the river network based on the Muskingum method (Cunge, 1969).

Thirteen parameters had to be estimated for each RHHU (10 for RHHUs without glaciated areas) (see Table 2). The travel time and weighting factor of the Muskingum method were estimated for the 17 channel segments of the channel routing model.

3.2 Parameter estimation

The calibration method uses hydrological process knowledge to extract useful information from the very heterogeneous data set available in the region, coming from various catchments and time periods and corresponding to different hydrometeorological phenomena. It is thus multi-signal and sequentially minimizes differences between reference and simulated hydrological signals or signatures where the reference comes from the 1982–2001 period for non-perturbed basins, and from the unperturbed 1941–1956 historical period otherwise. The calibration starts with point-scale data from climatic stations, continues with data from gauged sub-basins located within or in the neighbourhood of the URR basin, then larger sub-basin areas, and finally the whole URR basin. Specific groups of parameters are assigned to each objective. Parameters calibrated in the corresponding calibration step may be readjusted in succeeding steps (Fig. 2). The calibration procedure starts with a reasonable initial parameter set available from previous studies on other gauged and non-perturbed basins in the Swiss Alps (Guex *et al*., 2002; Schaefli *et al*., 2005; Horton *et al*., 2006). For regionalization purposes, parameter values are assumed to be space-invariant for all RHHUs of a given sub-basin, but also over selected sub-areas and if possible over the whole domain. The main steps of the calibration strategy are detailed in the following.

3.2.1 Phase of precipitation The threshold air temperatures for snow/rainfall separation (Tc_1, Tc_2) have been determined based on the empirical frequencies of precipitation events falling as snow, rainfall or mixed phase as a function of air temperature. They

Table 2 Selected algorithms and parameters.

Module	Algorithm	Calibration parameters
Interpolation of meteorology	Weighted mean of nearest met. stations with hourly lapse rate	Threshold temperatures for snowfall computation (Tc_1, Tc_2)
Snowpack evolution	Modified TI approach	TIF for snowmelt (a_n) / refreezing (a_r), threshold temperature for melts / freezing (Tc_0), snowpack saturation ratio (θ_r), a_n correction coeff. if liquid precipitation (b_r)
Glacier melt	TI approach	TIF for ice melt (a_g)
Rainfall and meltwater – discharge transfer	Non glaciated areas	Maximum storage capacity of slow reservoir (A), recession coefficients for slow / rapid linear reservoirs (K_l, K_r)
	Glaciated areas	Recession coeffs for snow- and ice melt reservoirs (K_n, K_g)
Channel routing	Muskingum approach	Travel time (K_m) and weighting factor (X_m)

TI: temperature-index; TIF: temperature-index factor.

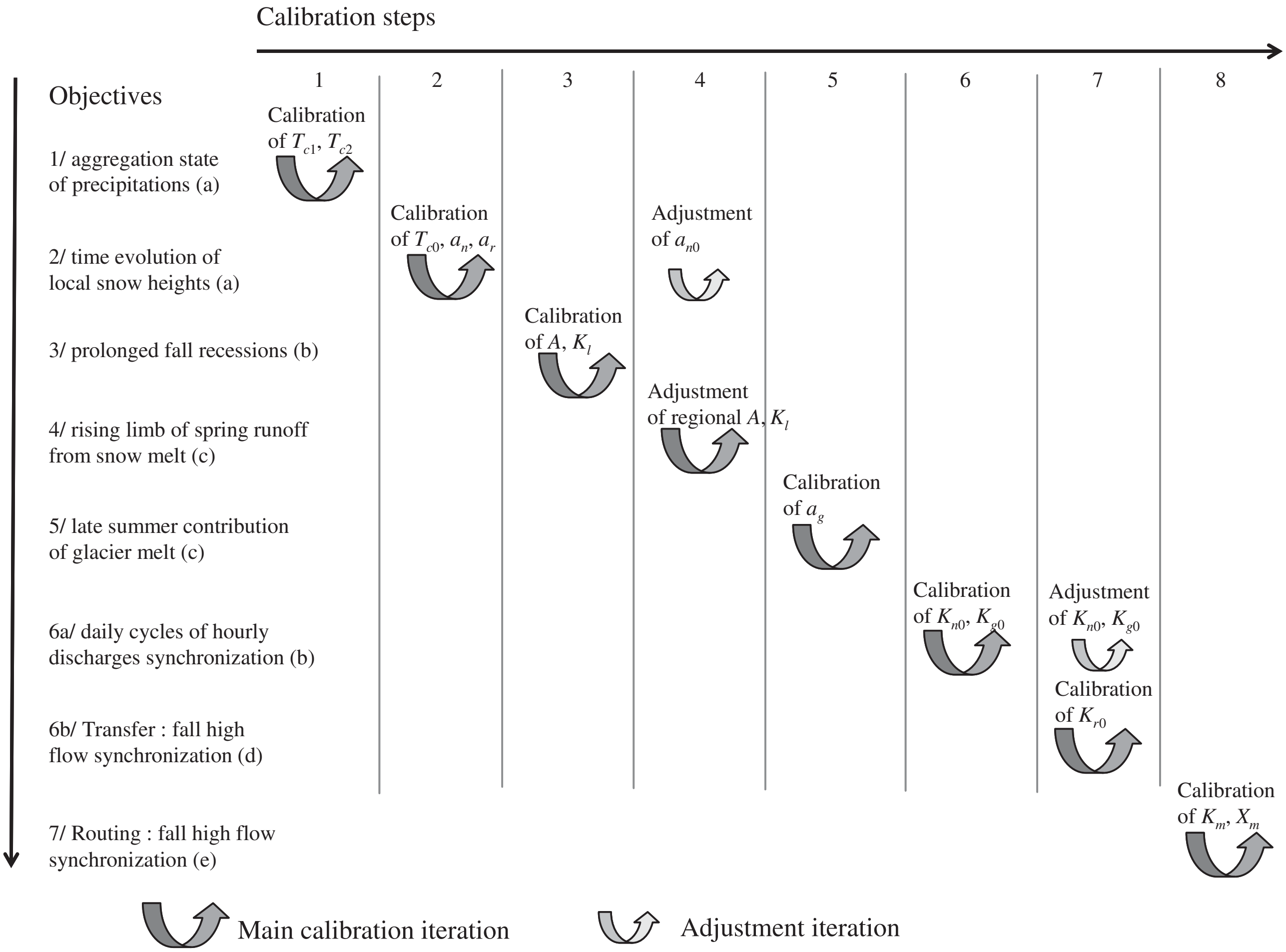

Fig. 2 Calibration steps and corresponding objectives. Calibration steps are based on: (a) point-scale data, (b) selected time periods in the case of a non-perturbed sub-basin, (c) the mean inter-annual cycle of daily discharges from a recent time period in the case of a non-perturbed sub-basin, or from historical data otherwise, and (d), (e) observed or reconstructed high-flow events, respectively, for 1987, 1993 or 2000.

are obtained from the observations of precipitation state available at 17 Swiss climatic stations located above 1000 m a.s.l. Threshold temperatures are similar for all stations and we thus retained space invariant values: $Tc_1 = 0°C$ and $Tc_2 = 2°C$.

3.2.2 Time evolution of local snow heights The saturation ratio of the snowpack (θ_r) and the correction coefficient (b_r) of the TIF have been specified based on values from the Canadian model HSAMI (Fortin, 2000). Series of daily snowpack heights observed at the 17 climatic stations were used for a first guess of the other snowpack module parameters. Hereby, the threshold temperatures for the temperature-index models were estimated such as to maximize synchronization between the spring starting dates of observed and simulated snowmelt at each climatic station. The TIFs for snowmelt and water refreezing, assumed to be proportional to each other, were next estimated such as to maximize the number of days with concomitant simulated and observed snowcover. For each of the 17 climatic stations, all parameter sets leading to reasonable simulation performances were identified (see Fig. 3 for details). Despite a considerable spread, the TIF for snowmelt was found to be an increasing function of the elevation. This relationship was modelled with a sigmoid function having a scale parameter a_{n0}. In a later calibration step, different values of this scale factor were assigned to different sub-regions.

3.2.3 Prolonged autumn recessions For non-perturbed low-elevation basins, the simulation of late autumn low-flow recessions is fully controlled by the outflow of the slow reservoir. Corresponding observed discharges can be used for estimating the recession

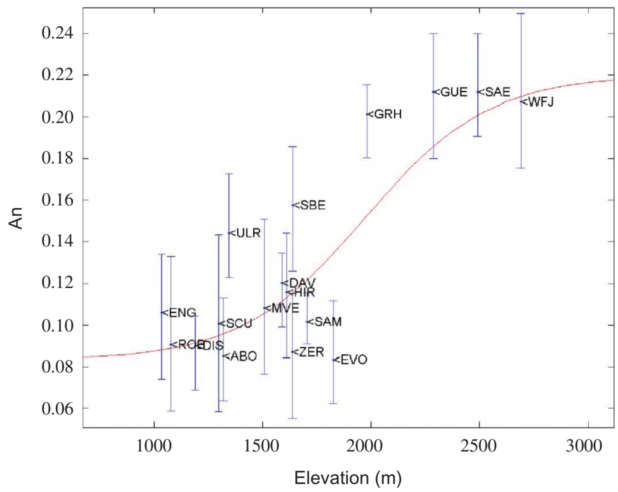

Fig. 3 TIFs for snowmelt (mm/°C/h) *versus* elevation (m). For each station, **<**: TIF leading to the largest number of days, N_C, with concomitant simulated and observed snowcover; bounds of ranges: TIFs leading to a 10% decrease in N_C.

constant (K_1) and the maximum storage capacity of the reservoir (A) following the methodology described in (Hingray *et al.*, 2006a). The estimation maximizes a Nash-Sutcliffe (NS) efficiency criterion (Nash & Sutcliffe, 1970) between simulated and observed discharges signals for late autumn recession periods. In the studied region, if not significantly modified by waterworks, late autumn recessions are often influenced by snowmelt from early snowfall events. The recession signals do therefore not necessarily only include release of underground water. Thus, a first guess was obtained for both parameters from their representative values extracted from a regionalization study on neighbouring non-perturbed and low-elevation basins (Guex *et al.*, 2002).

3.2.4 Mean inter-annual flood from spring snowmelt In Alpine basins, the high discharge seasonality is due to the seasonal accumulation of snow on hillslopes and to the temporary subsoil storage of infiltrated rainfall or meltwater. In the model, the seasonality of simulated discharges is governed by parameters conditioning accumulation and melt and by those associated to infiltration and low flow release. These parameters were calibrated on two signatures. The timing of the rising limb of the spring snowmelt flood is used to adjust the maximum storage capacity, A, of the slow reservoir (the reservoir can actually not produce significant flow if not partly filled: the greater its storage capacity, the later the beginning of simulated snowmelt flood). The slope of the rising limb is, in addition, an increasing function of the TIF for snowmelt; this signature was used to refine the scale factor, a_{n0}, of the temperature-index–elevation relationship at the basin scale and to further refine A and K_1. The reference values of these signatures were obtained from the 1982–2001 period for non-perturbed basins, and from the historical period 1941–1956 otherwise. Parameters are estimated so as to graphically minimize the difference between the reference and the simulated signature obtained for the 1982–2001 simulation period (Fig. 4(a)).

3.2.5 Mean inter-annual flow from glacier melt For partly glaciated basins, the late summer contribution of glacier melt to river discharge determines the summer water balance. In the model, this ice melt contribution is an increasing function of a_g and on the TIF for snow, which determines the length of the snow-free period (for further details, see Schaefli *et al.*, 2005). We, thus, used the mean falling limb of the meltwater induced flood to calibrate a_g (Fig. 4(b)) and to adjust if necessary the scaling factor, a_{n0}. We minimized the difference between reference and simulated mean inter-annual late summer discharge (from 15 July to 30 September). The reference value of this signature is obtained from the mean inter-annual cycle of daily discharge values, from the 1982–2001 period for non-perturbed basins, and from the historical 1941–1956 period otherwise.

3.2.6 Hourly discharge synchronization – sub-daily cycles and autumn high flows The downstream transfer of available water to the outlet of each glaciated sub-basin depends on three recession coefficients. From the beginning to the end of the melting period, they in turn modify daily cycles of melt water induced by daily variations of temperatures (K_r in spring for melt water from non-glaciated areas, K_r and K_n in early summer when the glaciers are snow-covered, K_n and K_g in late summer, when the non-glaciated areas are snow-free). In autumn, precipitation usually falls as snow over glaciated areas of the basin. High flows produced by heavy rainfall over the non-glaciated parts of the basin are determined by K_r. For unperturbed catchments, Schaefli *et al.* (2005) used this process knowledge to calibrate the transfer parameters successively: for each parameter, the estimation maximizes the temporal correlation between the corresponding observed and the simulated signals. In the absence of unperturbed data, a similar

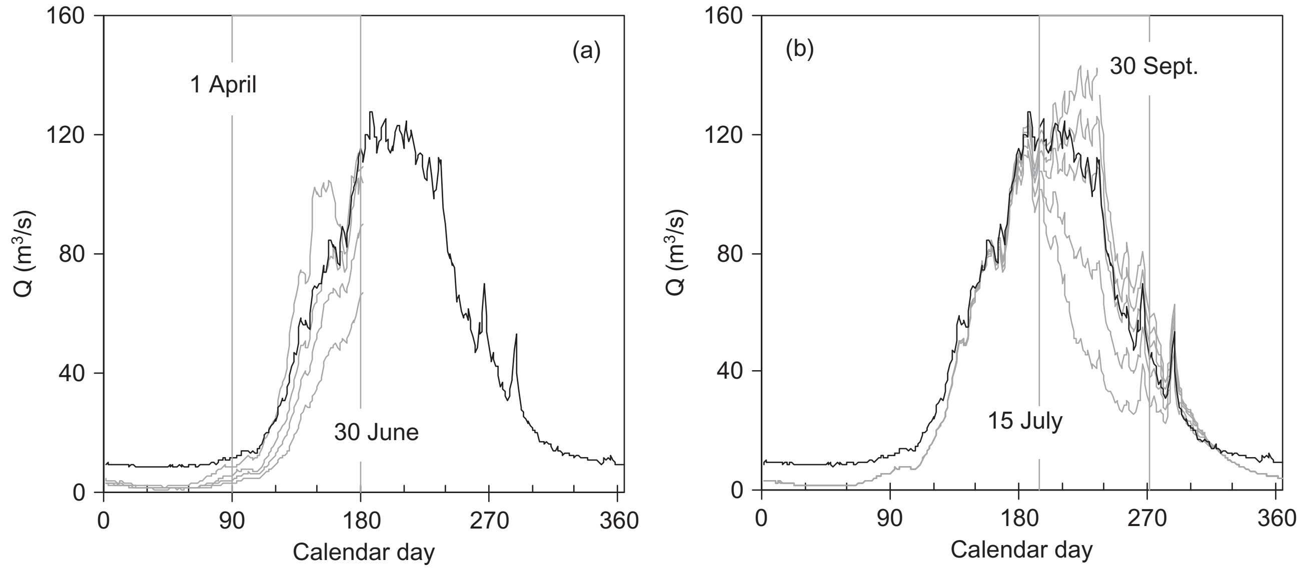

Fig. 4 Calibration of the TIFs: (a) for snowmelt – signature: rising limb of snowmelt flood (1 April–30 June); (b) for glacier melt – signature: late summer flow (15 July–30 September). Grey curves: simulated mean inter-annual discharge for different TIFs for either snow- or glacier-melt; black curve: observed values. Illustration for the Brig sub-basin (reference and simulated period: 1982–2001).

approach was not possible (we only had three reconstructed autumn flood events). Thus, we considered by default that the recession coefficients K_n and K_g were equal to the coefficient K_r, for which a regional scaling relationship had been obtained in earlier work (Guex *et al.*, 2002): K_r was assumed to depend on the basin surface area (S) and on its mean slope (p) via the following equation:

$$K_r = K_{r0} \left(\frac{S}{S_0}\right)^{0.5} \left(\frac{p}{p_0}\right)^{-0.5} \tag{1}$$

where K_{r0} is the scale factor of the relationship, i.e. the recession coefficient for a reference basin with surface area and mean slope being S_0 and p_0, respectively. For the URR basin, this relationship can account for a significant space variability of the K_r parameters (from 8 to 30 h according to the sub-basin characteristics). The estimation of the three recession coefficients for all sub-basins located upstream to a given estimation point thus reduces to the estimation, K_{r0}. Its value was estimated such as to reproduce the three historical floods (peak discharge or entire hydrograph if available).

3.2.7 Hourly discharges synchronization – discharge routing The downstream routing of discharges through the river segments is fully determined by the two parameters of the Muskingum approach. These parameters were estimated maximizing the time correlation coefficient between simulated and observed hydrographs for the 1987 and 1993 historical floods. Note that we used here the "pseudo-simulated" hydrographs resulting from routing the reference (observed or reconstructed) hydrographs available at the upstream locations. The values of the Muskingum parameters are thus independent of all other estimated parameter values.

4 RESULTS

We give only the most important results in this section. Detailed results can be found in Hingray *et al.* (2006b). The ability of the model to reproduce the mean inter-annual cycle of daily discharge values – obtained for the calibration period or for the historical data – is quite good (Fig. 5). The Nash-Sutcliffe criterion estimated between the reference value and the simulated value of this signature varies between 0.95 and 0.98 depending on the station. The corresponding mean bias varies between –5% and +2%.

The model performance for the simulation of the three reconstructed flood events is rather modest and varies between the basins (Fig. 6). It is satisfactory for the URR at Brig, but particularly poor for the Viège River at Viège, with a large overestimation for 2000 and large underestimation for 1987 and 1993 (not shown). As simulation errors propagate downstream, it is also relatively poor for the downstream

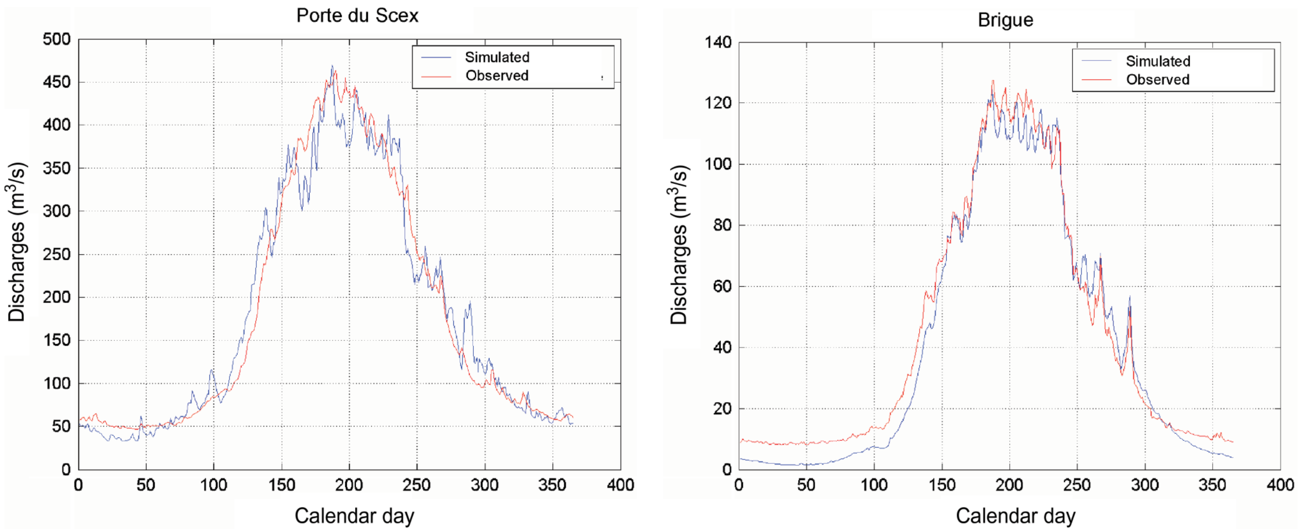

Fig. 5 Simulated and reference mean inter-annual cycle of daily discharges for the URR basin (Porte du Scex) and for the URR sub-basin at Brig (reference period: 1941–1956 for Porte du Scex; 1982–2001 for Brig; simulated period: 1982–2001).

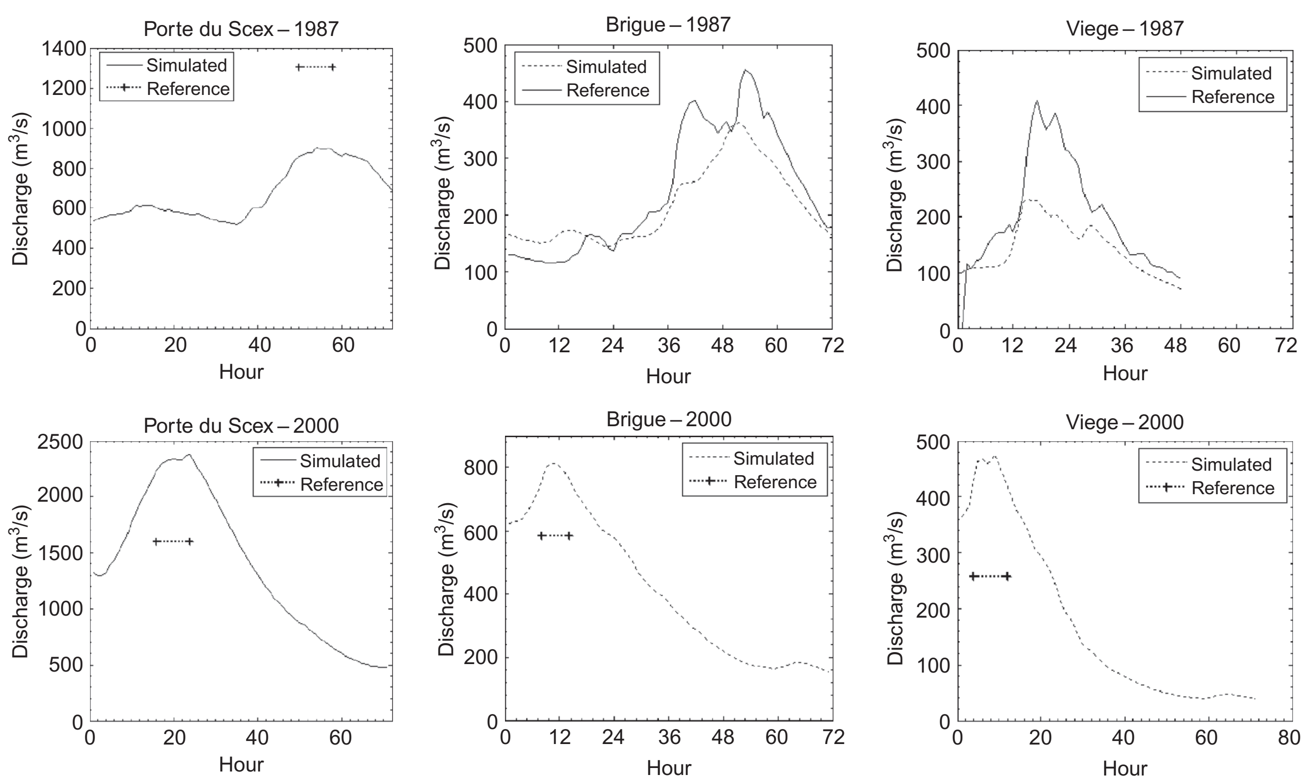

Fig. 6 Reproduction of the reconstructed 1987 and 2000 flood hydrographs or peak discharges for the URR basin (Porte du Scex) and two main sub-basins (URR at Brig, Viège at Viège). Straight lines correspond to peak discharge indirectly estimated from water line observations; in these cases hydrographs are not available and the exact timing of peak discharge is unknown.

stations. The quality of the model for each intermediate sub-basin is however rather good. When simulating the historical floods at these downstream stations using the reconstructed ("pseudo-observed") hydrograph for Viège instead of the simulated hydrograph, the resulting hydrographs are very close to the historical floods for all downstream stations (not shown).

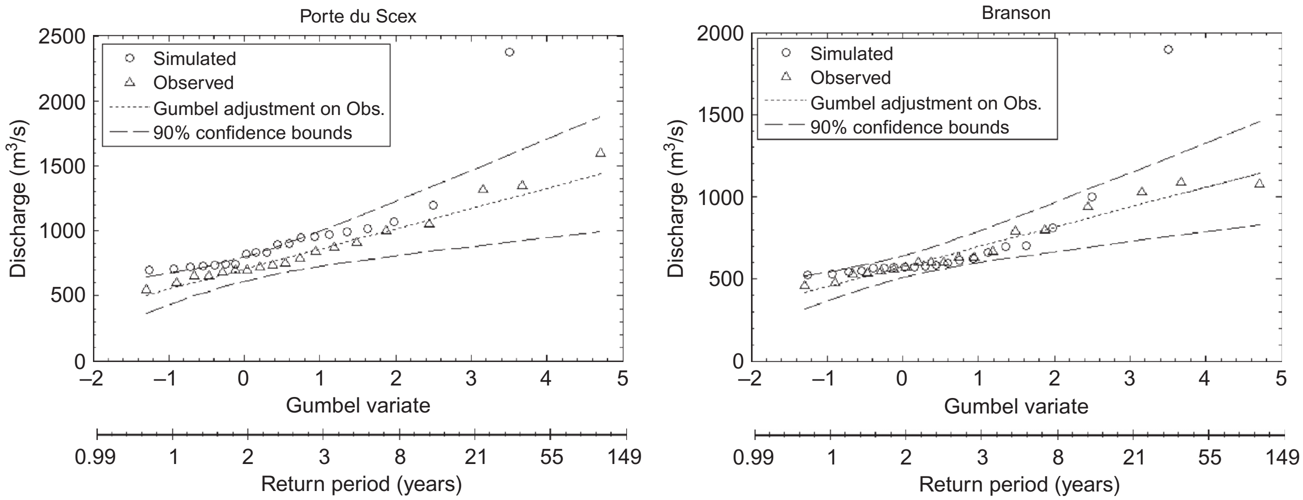

Fig. 7 Gumbel plots for hourly annual discharge maxima obtained for the reference period (1941–1956) and from simulation (1982–2001). The x-axis is the reduced variate for the given return period T, i.e. $u = -\ln(-\ln(1 - 1/T))$. Dashed lines correspond to 90% confidence bounds of the Gumbel distribution estimated based on the reference data.

As Monte Carlo simulations and ongoing research showed, the large under- or overestimation for the Viège basin cannot be explained by model parameter uncertainty. The main reason for the modest flood reproduction performance is the large uncertainty in area-averaged meteorological inputs (precipitation, temperature), which determine the amount of water available for flood generation. For October 2000, precipitation fell as rain over almost the entire URR basin, except in the Viège basin where a sudden decrease of temperature led to a temporary storage of precipitation as snow. It is generally agreed that without this meteorological phenomenon, the 2000 flood event would have been much larger in this basin (OFEG, 2002). However, this local cooling is not reflected in the observed time series; it was only observed by inhabitants of the valley. The simulated discharge values are thus significantly higher than the reconstructed hydrograph. A similar but inverse situation partially explains the large underestimation of the 1987 flood. The use of additional meteorological data from hydropower production companies would considerably improve the reproduction of these events, as recently demonstrated by Hernández *et al.* (2009). However, they were not available for this research.

For the perturbed URR basin and sub-basins, a classical calibration/validation procedure, in which parameters are estimated and validated on separate time periods, was not possible. Nevertheless, we completed a soft validation, comparing the statistical distribution of maximum annual discharges from the reference period to that obtained by simulation. However, as reference and simulated hourly annual discharge maxima do not refer to the same time period, we do not necessarily expect that they follow exactly the same distribution, contrary to what is expected for non-perturbed sub-basins (see results for Brig in Fig. 8). As illustrated in Fig. 7 for the URR at Branson, the simulated discharge were well within the confidence bounds obtained assuming a Gumbel distribution, except for the October 2000 flood, which, for the reasons discussed above, lies far outside the confidence bounds. For the URR at Porte du Scex, discharge maxima are all overestimated, even if still on the limit of the confidence bounds. However, this suggests that an improved estimation of parameters for the sub-basins located downstream of Branson could be found (for example, for the transfer parameters that seem to be too small).

5 DISCUSSION

Even if not fully satisfactory, the model simulates reasonably well the main hydrological signatures of the URR basin and sub-basins. However, considering individual high-flow events, the model shows a modest performance for some sub-basins. In the following, we discuss some input-related modelling uncertainties and potential improvements of the model calibration methodology.

For the first point, we completed extensive parameter sensitivity analyses, which showed that, for the historical flood events, the parameters influencing the effectively available liquid water for runoff generation are the most sensitive ones. The estimation of area-averaged liquid water input depends on the model

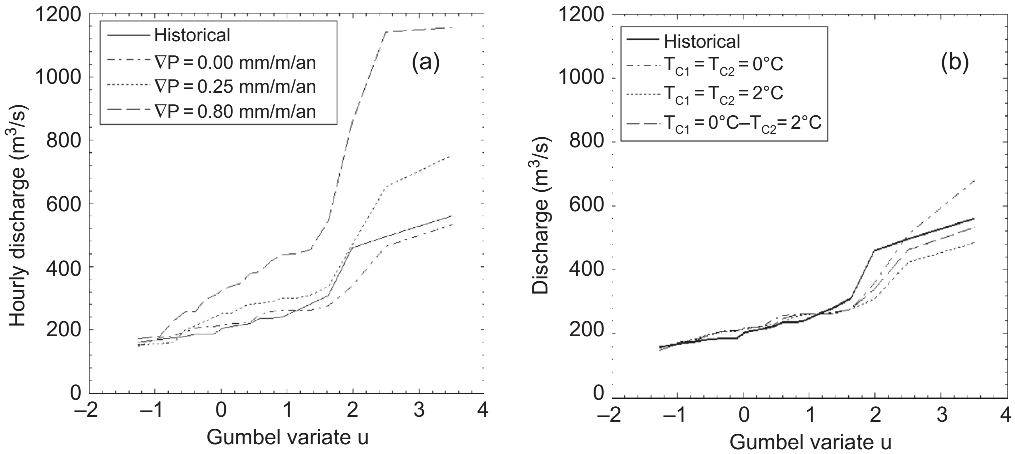

Fig. 8 Gumbel plots for observed (solid line) and simulated hourly annual discharge maxima (Brig sub-basin): (a) with different precipitation gradients ∇P: $\nabla P = 0$; $\nabla P = 0.25$ mm m^{-1} year^{-1} – estimated for the URR basin; $\nabla P = 0.8$ mm m^{-1} year^{-1} – estimated for the entire Swiss Alps by Kirchhofer & Sevruk (1991); and (b) with different thresholds for snow–rainfall separation: $Tc_1 = 0$°C and $Tc_2 = 2$°C; $Tc_1 = Tc_2 = 0$°C; $Tc_1 = Tc_2 = 2$°C.

altitudinal resolution, the temperature thresholds for rainfall/snow separation and the meteorological time series interpolation. We briefly discuss each of these sources of uncertainty below.

5.1 Areal precipitation estimation and altitudinal precipitation gradient

For estimating area-averaged hourly precipitation, we did not consider altitudinal gradients of precipitation as is frequently done in models of Alpine basins. The enhancement of precipitation with elevation primarily applies to annual precipitation, but does not hold for individual precipitation events. Higher precipitation amounts in higher elevation zones are largely due to more frequent precipitation events. Omitting an altitudinal gradient is thus a good option to simulate flood events, especially in cases of large precipitation events. Flood volumes and peak discharges are otherwise largely overestimated.

Statistical distributions of maximum annual discharges obtained for the Brig URR sub-basin with a regional or mesoscale precipitation altitudinal gradient are compared for illustration with those obtained without a gradient in Fig. 8(a). However, omitting an altitudinal gradient results in a mis-estimation of the annual water balance (Schädler & Bigler, 2002). This could potentially lead to bias in actual evapotranspiration and glacier-melt parameters that tend to compensate for water balance errors (Schaefli *et al.*, 2005). We assumed that these errors have little influence on flood simulations. An improved spatial interpolation of precipitation based on weather types (Gottardi *et al.*, 2008) could partly circumvent this problem.

5.2 Rain/snow separation, threshold temperatures and lapse rate

A poor estimation of the elevation–temperature relationship, partly due to data scarcity in high elevation zones, potentially leads to a mis-estimation of the fraction of liquid precipitation available for flood generation. The classical approach usually retained in models for Alpine basins is to use the average environmental lapse rate (–0.55°C/100m for the URR basin). The solution retained here is to use a time variable lapse rate estimated from the meteorological station network. This significantly improved the model performance (Hingray *et al.*, 2006b). However, the elevation–temperature relationship is not always linear and has to be better estimated. Almost no studies are available that investigate how to best treat this estimation problem in precipitation–runoff models. Further research is required to improve our method.

Directly related to this problem is the two-thresholds approach used to estimate the amount of liquid precipitation from total precipitation. Our hydrological model assumes a deterministic relationship between air temperature and rainfall. Since other atmospheric variables (e.g. relative air humidity) influence the phase of precipitation, this may lead to mis-estimation of liquid input for individual events. Precipitation can be fully liquid for 0°C air temperature or, conversely, fully solid for an air temperature of 2°C. Large errors in flood simulations can therefore be obtained (see also the previous discussion for the Viège basin). For illustration, statistical distributions of maximal annual discharges obtained using a one-threshold snow/rainfall separation approach (either 0°C or 2°C instead of the two-threshold temperatures approach) are presented in Fig. 8(b). The two-threshold approach performs the best. However, the important sensitivity of the distribution to the separation method highlights the need to further improve the conceptualization of the snow/rainfall separation process.

5.3 Point-scale process modelling and elevation bands

The computation of liquid water input depends, additionally, on the altitudinal resolution of the model. The key question here is how the extension of a point-scale rainfall/snow separation process to elevation bands affects the model performance. Since temperature varies with altitude, the scale-transition method is crucial. Here, we compute the elevation band snow/rainfall separation as a point process for the MAT. The induced error depends on the range of elevations covered by the band and the lapse rate (see Fig. 9). For the average lapse rate in the URR basin, when the band covers a 1000 m (resp. 2000 m) elevation range, the fraction of liquid precipitation (FLP) obtained for the band mean elevation is 0% for MAT = 0°C, while the real area-averaged FLP is around 33% (resp. 40%).

Area-averaged FLP is obtained from local FLPs estimated for all band locations according to the following hypotheses:

1. local FLP is a function of local air temperature; the local FLP–temperature relationship obtained from climatic data (Section 3.2) is space-invariant and thus independent of elevation; it is represented by the bold line in Fig. 9;
2. the elevation–temperature relationship is linear and the lapse rate is –0.55°C/100 m; and
3. the hypsometric curve of each band is a linear function of elevation.

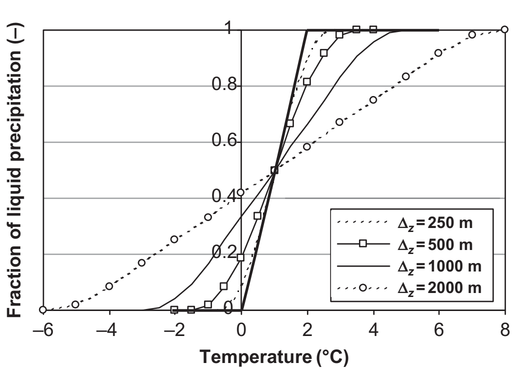

Fig. 9 Area-averaged fraction of liquid precipitation (FLP) *versus* MAT for four elevation bands covering different elevation ranges (Δz = 250, 500, 1000 and 2000 m).

Our sensitivity analysis showed that an altitudinal range of 500 m per band is a good compromise for the URR basin. The identification of the best altitudinal resolution also has to account for the altitudinal variability of the other snow related variables and processes. However, these results are in line with empirical results obtained by Schaefli *et al*. (2005) from discharge simulation results.

5.4 Sensitivity of model performance to calibration approach

The model performance could potentially also be enhanced through an improved calibration methodology. To assess the relevance of our multi-signal stepwise methodology, we additionally completed a classical global optimization using a single objective function. We tried to calibrate the model for the natural catchments of the URR basin using the so-called DREAM algorithm (Vrugt *et al*., 2010) a Markov Chain Monte Carlo sampler. Even with this very efficient sampler, we still need to run the model at least 50 000 times to calibrate the six model parameters (TIFs and the meltwater runoff transfer parameters). Computational limitations therefore restricted us to the use of only a couple of years for calibration (a complete model run for 20 years for all sub-basins takes around 3 hours on a personal computer). The results obtained show that such a global optimization using only observed discharge as a reference signal (instead of all available signatures) is not viable; the values of the key model parameters depend strongly on the chosen reference time period: the calibration either yields far too high ice-melt TIFs or too high snowmelt TIFs.

5.5 Comparison to a benchmark

The question remains how well our mixed parameter calibration/regionalization methodology performs. Therefore, we constructed a minimum benchmark simulation transposing the parameters obtained through multi-signal calibration for the best gauged catchment (Brig) to all ungauged sub-basins without including any further regionalization (i.e. using the same parameters for all sub-basins leading to a

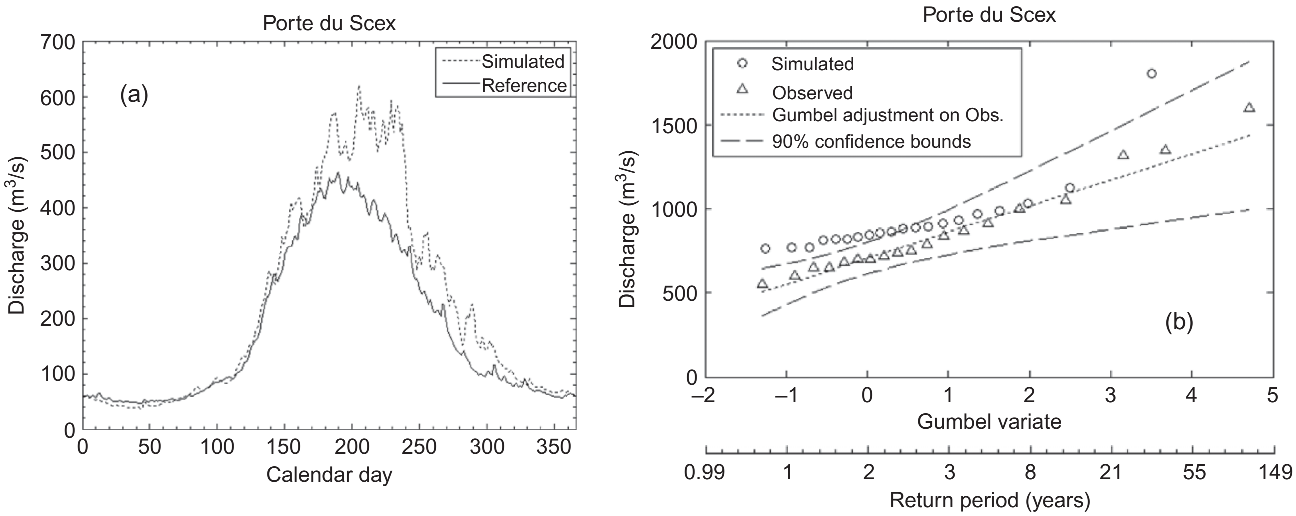

Fig. 10 Performance of a minimum benchmark model (URR at Porte du Scex): (a) same as Fig. 5; and (b) same as Fig. 7.

semi-lumped model). For extreme event prediction, but also for the prediction of the yearly cycle of mean inter-annual daily discharges, our calibration methodology outperforms this minimum benchmark (Fig. 10). This highlights the importance of using, whenever possible, all available hydrological information for model calibration (e.g. mean inter-annual signatures extracted from historical non perturbed periods in our case). The signatures presented here are, of course, case study specific and the available information could be very different for other catchments. The step-wise multi-signal approach presented here is flexible enough for these adaptations and our case study gives interesting hints about what kind of signatures are potentially useful for snow- and ice-melt influenced catchments. A further improvement could be obtained, for example, with a better use of the information contained in observed autumn and winter low flows for estimating the parameters of the soil reservoir, as proposed by Lafaysse *et al*. (2010).

5.6 Validity of the stationarity hypothesis

Finally, we briefly discuss the stationarity assumption and the question of how we could estimate the uncertainty inherent in the proposed signature-based model calibration methodology. The essential property of a signature is that it reflects some average catchment behaviour. Estimating this average behaviour based on historical data poses the question of how good the underlying (necessary!) stationarity assumption is. For the URR basin, land use as well as meteorological forcing are likely to have evolved significantly over the last century (Schmidli & Frei, 2005). We thus used only the most recent unperturbed historical discharge data (1942–1956), which to our view, is a good compromise between taking advantage of existing information and not mis-informing the model during parameter estimation. Winsemius *et al.* (2009) have recently proposed a methodology to quantify how uncertain historical signatures are: they compute the signatures for individual years and construct credibility limits around the mean inter-annual signature. Such an approach could further improve the presented results.

6 CONCLUSION

We have presented a calibration framework for the distributed discharge modelling of the Upper Rhône River in the Swiss Alps. Even if this mesoscale catchment shows an important number of observed hydrometerological time series, calibrating a distributed hourly model is challenging, first of all due to computational limitations, but more importantly because in such an Alpine environment most rivers are used for hydropower generation and the observed meteorological time series only cover the lower parts of the basins. We thus had to deal with similar problems as in poorly gauged catchments, where observed data are available but cannot be used directly for a classical model calibration and validation with long concomitant hydrological and meteorological time series.

The developed calibration method is based on hydrological signatures and combines hydrological process knowledge and insights into how the model works to extract relevant hydrological information from all available data sets. The core of the approach is the successive calibration of parameter groups in order to reproduce reference hydrological signatures that have been estimated based on observed data. Given that the type of hydrometeorological data we used is generally available in similar Alpine environments, the methodology can easily be transposed to other conceptual models for mountainous catchments. Furthermore, given the step-wise nature of the calibration procedure, new observed data that may become available in the future could be used to update the estimated parameters. Such a new data set could, for example, consist of glacier mass balance data or the currently confidential discharge data from hydropower companies (Schaefli *et al*., 2005).

The main objective of this paper was the discussion of the model calibration framework. At a later stage, this should be completed by a parameter uncertainty analysis. In particular, we are thinking of assessing the prediction uncertainty in a limits-of-acceptability approach (Winsemius *et al*., 2009). This method defines acceptable limits for each source of information and retains all parameter sets that yield simulations falling within these limits. As shown by Winsemius *et al*. (2009), this method is particularly promising for assessing prediction uncertainty in the presence of soft and hard information.

The partly poor simulation results certainly emphasize the limitations of the signature-based calibration approach for extreme event prediction for the presented case study. But they also highlight the inherent prediction uncertainties in similar Alpine environments due to uncertain meteorological model input. In this context, further research is required to obtain unbiased effective rainfall–runoff parameters in the presence of highly uncertain liquid water input, which is typical for high mountainous environments. Further research is also required to determine whether the hydrological model could be improved through additional observed hydrological data and process understanding, or whether prediction in this environment simply needs a denser meteorological network to better characterize the inputs.

Acknowledgements This work was part of the CONSECRU project (CONcept de SEcurité contre les CRUes) funded by the Wallis Canton and by the Swiss Federal Office for the Environment. We wish to thank the Swiss Federal Office for the Environment for the discharge data, the national weather service, MeteoSwiss, for the time series of meteorological and snow heights variables, and the Swiss Federal Institute of Topography for the topographic and land-use data. Special thanks are due to Frédéric Guex, Gabriel Faivre and Markus Niggli for their contribution to this work. We would also like to thank the two anonymous reviewers and the Guest Editors for their detailed comments that helped to improve our manuscript.

REFERENCES

Bárdossy, A. (2007) Calibration of hydrological model parameters for ungauged catchments. *Hydrol. Earth System Sci.* **11**, 703–710.

Burman, R. & Pochop, L. O. (1994) *Evaporation, Evapotranspiration and Climatic Data*. Amsterdam: Elsevier.

Cunge, J. A. (1969) Au sujet d'une méthode de propagation de crue. *J. Hydraul. Res.* **7**, 205–230.

Fortin, V. (2000) Le modèle météo-apport HSAMI: historique, théorie et application, Rapport de recherche, Institut de recherche d'Hydro-Québec, Varennes, Québec.

Gottardi, F., Obled, C., Paquet, E. & Gailhard, J. (2008) Régionalisation des précipitations sur les massifs montagneux français à l'aide de regressions locales et par type de temps. *Climatologie* 5.

Guex, D., Guex, F., Pugin, S., Hingray, B. & Musy, A. (2002) Regionalisation of hydrological processes in view of improving model transposability. WP3 Final Report of the Pan-European FRHYMAP Project "Flood Risk scenarios and Hydrological MAPping". Lausanne: EPFL, Contract no. CE: 3/NL/1/164/99 15 183 01.

Gupta, H. V., Beven, K. J. & Wagener, T. (2005) Model calibration and uncertainty estimation. In: *Encyclopedia of Hydrological Sciences* (M. G. Anderson, ed.), 2015–2032. Chichester: Wiley.

Gupta, H. V., Wagener, T. & Liu, Y. Q. (2008) Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol. Processes* **22**(18), 3802–3813.

Hernández, J. G., Boillat, J. L., Jordan, F. & Hingray, B. (2009) La prévision hydrométéorologique sur le bassin versant du Rhône en amont du Léman. *La Houille Blanche* **5**, 61–70.

Hingray, B., Guex, F., Guex, D., Pugin, S., Musy, A., Pfister, L., Idrissi, A., Iffly, J. F. & Hoffmann, L. (2006a) Simulation of streamflow by a regionalized lumped rainfall–runoff model over Luxembourg. In: *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment – MOPEX* (V. Andréassian, A. Hall, N. Chahinian & J. Schaake, eds.), 256–263. Wallingford: IAHS Press, IAHS Publ. 307.

Hingray, B., Mezghani, A., Schaefli, B., Niggli, M., Faivre, G., Guex, F., Hamdi, Y. & Musy, A. (2006b) Estimation des débits de crue du Rhône à Porte du Scex et autres points amont caractéristiques. Lausanne: EPFL, HYDRAM, Projet CONSECRU 2, Rapport final.

Hingray, B., Picouet, C. & Musy, A. (2009) *Hydrologie. 2. Une science pour l'ingénieur*. Lausanne: Presses Polytechniques Universitaires Romandes.

Horton, P., Schaefli, B., Mezghani, A., Hingray, B. & Musy, A. (2006) Assessment of climate-change impacts on Alpine discharge regimes with climate model uncertainty. *Hydrol. Processes* **20**(10), 2091–2109.

Kirchhofer, W. & Sevruk, B. (1991) *Mean annual corrected precipitation depths 1951–1980. Atlas hydrologique de la Suisse*. Bern: Service Hydrologique et Géologique National.

Kuchment, L. S. & Gelfan, A. N. (1996) The determination of the snowmelt rate and the meltwater outflow from a snowpack for modeling river runoff generation. *J. Hydrol*. **179**(1-14), 23–36.

Lafaysse, M., Hingray, B., Etchevers, P. & Martin, E. (2010) Improvement, calibration and validation of a physical-based hydrological model for Alpine catchments. Submitted to *J. Hydrol*.

Mezghani, A. & Hingray, B. (2009) A combined downscaling-disaggregation weather generator for stochastic generation of multisite hourly weather variables in complex terrain. Development and multi-scale validation for the Upper Rhone River Basin. *J. Hydrol.* **377** (3-4), 245–260.

Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models. Part I – A discussion of principles. *J. Hydrol.*, **10**(3), 282–290.

Obled, C., Zin, I. & Hingray, B. (2009) Choix des pas de temps et d'espace pour des modélisations parcimonieuses en hydrologie des crues. *La Houille Blanche* **5**, 81–87.

OFEG (2002) *Les crues 2000: Analyse des événements/Cas exemplaires. Protection contre les crues dans le canton du Valais*. Bern: Office Fédéral de l'Eau et de la Géologie.

Schädler, B. & Bigler, R. (2002) *Components of the natural water balance 1961–1990. Atlas hydrologique de la Suisse*. Bern: Service Hydrologique et Géologique National.

Schaefli, B., Hingray, B., Niggli, M. & Musy, A. (2005) A conceptual glacio-hydrological model for high mountainous catchments. *Hydrol. Earth System Sci.* **9**(1-2), 95–109.

Schmidli, J. & Frei, C. (2005) Trends of heavy precipitation and wet and dry spells in Swizterland during the 20th century. *Int. J. Climatol*. **25**, 753–771.

Sivapalan, M., Takeuchi, K., Franks, S., Schertzer, D., O'Connell, P. E., Gupta, V. K., McDonnell, J. J., Pomeroy, J. W., Uhlenbrook, S., Zehe, E. & Lakschmi, V. (2003a) IAHS Decade on Prediction in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. *Hydrol. Sci. J.* **48**(6), 857–880.

Sivapalan, M., Blöschl, G., Zhang, L. & Vertessy, R. (2003b) Downward approach to hydrological prediction. *Hydrol. Processes* **17**(11), 2101–2111.

Viviroli, D., Mittelbach, H., Gurtz, J. & Weingartner, R. (2009) Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland, Part II: Parameter regionalisation and flood estimation results. *J. Hydrol*. **377**(1-2), 208–225.

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Higdon, D., Robinson, B. A. & Hyman, J. M. (2010) Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int. J. Nonlinear Sci. Numerical Simulation* **10**(3), 273–290.

Winsemius, H., Schaefli, B., Montanari, A. & Savenije, H. H. G. (2009) On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information. *Water Resour. Res*. **45**, W12422, doi:10.1029/2009WR007706.

Yilmaz, K. K., Gupta, H. V. & Wagener, T. (2008) A process-based diagnostic approach to model evaluation: application to the NWS distributed hydrologic model. *Water Resour. Res*. **44**(9), W09417, doi:10.1029/2007wr006716.