

RESEARCH

Open Access

Parameter uncertainty and identifiability of a conceptual semi-distributed model to simulate hydrological processes in a small headwater catchment in Northwest China

Shuai Ouyang¹, Heike Puhmann², Shunli Wang³, Klaus von Wilpert⁴ and Osbert Jianxin Sun^{1*}

Abstract

Introduction: Conceptual hydrological models are useful tools to support catchment water management. However, the identifiability of parameters and structural uncertainties in conceptual rainfall-runoff modeling prove to be a difficult task. Here, we aim to evaluate the performance of a conceptual semi-distributed rainfall-runoff model, HBV-light, with emphasis on parameter identifiability, uncertainty, and model structural validity.

Results: The results of a regional sensitivity analysis (RSA) show that most of the model parameters are highly sensitive when runoff signatures or combinations of different objective functions are used. Results based on the generalized likelihood uncertainty estimation (GLUE) method further show that most of the model parameters are well constrained, showing higher parameter identifiability and lower model uncertainty when runoff signatures or combined objective functions are used. Finally, the dynamic identifiability analysis (DYNIA) shows different types of parameter behavior and reveals that model parameters have a higher identifiability in periods where they play a crucial role in representing the predicted runoff.

Conclusions: The HBV-light model is generally able to simulate the runoff in the Pailugou catchment with an acceptable accuracy. Model parameter sensitivity is largely dependent upon the objective function used for the model evaluation in the sensitivity analysis. More frequent runoff observations would substantially increase the knowledge on the rainfall-runoff transformation in the catchment and, specifically, improve the distinction of fast surface-near runoff and interflow components in their contribution to the total catchment runoff. Our results highlight the importance of identifying the periods when intensive monitoring is critical for deriving parameter values of reduced uncertainty.

Keywords: Dynamic identifiability analysis; HBV-light model; Hydrological modeling; Sensitivity analysis; Uncertainty analysis

Introduction

Hydrological models are important tools for water resource planning and management and in assessing the effects of climate and land use change on the hydrological cycles and runoff regimes (Pechlivanidis et al., 2011; Zhang et al., 2012). Conceptual hydrological models are widely used to simulate the land phase of hydrological

cycles since they can capture the dominant catchment dynamics whilst remaining parsimonious and computationally efficient whilst requiring input data that are usually readily available and relatively simple and easy to use (Thyer et al., 2009; Kavetski and Clark, 2010).

Parameters in conceptual hydrological models need to be estimated through model calibrations because they cannot be directly determined from the physical characteristics of the catchment (Madsen, 2000; Madsen et al., 2002). However, when parameter calibration is employed, different parameter sets may simulate the observed system behavior

* Correspondence: sunjianx@bjfu.edu.cn

¹Institute of Forestry and Climate Change Research, Beijing Forestry University, 35 Qinghua East Road, Haidian District, Beijing 100083, China
Full list of author information is available at the end of the article

equally well, which is termed “equifinality” (Beven and Freer, 2001). Commonly, the calibrated model is tested against some independent (validation) dataset to ensure the applicability of the model to situations/periods not used in the model calibration. Typically, split-sampling or differential split-sampling are used to divide the entire dataset into two parts (Xevi et al., 1997; Henriksen et al., 2003; Moriasi et al., 2007). Deteriorating model behavior for the validation dataset may hint at parameter identification problems. However, with regard to the relatively large number of free parameters in a rainfall-runoff model, a single measure of performance is a weak criterion to assess and declare (or refuse) modeling success against the background of omnipresent equifinality (Beven, 2001). It is difficult to characterize the different aspects of model performance for a particular rainfall-runoff model with only one or two statistical criteria (Shakti et al., 2010). There have been suggestions that the information from runoff data can be much better utilized and the information for model calibration is increased when using objective functions based on hydrological signatures rather than purely statistical measures (Shamir et al., 2005; Gupta et al., 2008; Wagener and Montanari, 2011). Hydrological signatures are defined as hydrologic response characteristics that provide insight into the hydrologic functional behavior of catchments (Sawicz et al., 2011). Such response characteristics are often indicative of a specific watershed and how its response differs from others; examples include common descriptors of the hydrograph shape such as the runoff duration curve and the time to peak flow (Shamir et al., 2005). Moreover, different objective functions judge the goodness of a certain parameter set by different aspects and, hence, a model’s success at simulating runoff may be better quantified by using several evaluation measures (Dawson et al., 2007) and the so-called Pareto optimality, which describes solutions in which an objective function cannot be improved without decreasing other objective functions.

It is therefore important for hydrologists to identify the dominant parameters controlling model behavior by using sensitivity analysis, which also helps to better understand the model structure, the main sources of model output uncertainty, and the identification issues (Ratto et al., 2007). Among a variety of global sensitivity analysis methods currently available, the regional sensitivity analysis (RSA; Hornberger and Spear, 1981), also known as the generalized sensitivity analysis, is very popular and widely used (Ratto et al., 2007; Saltelli et al., 2008).

Hydrological modeling involves multiple steps, each with uncertainties of different origins that render uncertainty in the final model predictions (Butts et al., 2004). Realistic assessment of various sources of uncertainty is not only important for science-based decision making

but also helps to improve model structure and to reduce model uncertainty. In recent years, quantification of uncertainties in hydrological modeling has received a surge of attention, and several methods have been developed to derive meaningful estimates of uncertainties bound on model predictions. Among these methods, the generalized likelihood uncertainty estimation (GLUE) method proposed by Beven and Binley (1992), and the Bayesian methods (Thiemann et al., 2001; Engeland et al., 2005) are widely used for simultaneous calibration and uncertainty assessment of different hydrological models (e.g., Freer et al., 1996; Kuczera et al., 2006; Blasone et al., 2008a; Vrugt et al., 2009; Dotto et al., 2012, 2014). Both methods have been discussed with respect to their philosophies and the mathematical rigor they rely on (Gupta et al., 2003; Kavetski et al., 2006; Kuczera et al., 2006; Blasone et al., 2008a; Vrugt et al., 2009; Jin et al., 2010; Dotto et al., 2012, 2014). The popularity of GLUE lies in its conceptual simplicity and relative ease of implementation, requiring no modifications to the existing source codes of simulation models (Vrugt et al., 2009). Moreover, GLUE makes no assumption regarding the distribution of the model residuals, and it allows a flexible definition of the model performance (likelihood function), making it capable of including several variables in model calibration and uncertainty assessment (Blasone et al., 2008b). The main critical point with GLUE is that the obtained confidence bounds are dependent on some subjective choices (e.g., the cut-off value between behavioral and non-behavioral simulations; see the methods section), and therefore represent the empirical rather than the true distribution of model uncertainty.

Based on the RSA and the GLUE, Wagener et al. (2003) developed the so-called dynamic identifiability analysis (DYNIA), which is an approach to locating periods of high identifiability (i.e., low uncertainty) for individual parameters and to detect failures of model structure in an objective manner. The main motivation behind the DYNIA is an attempt to avoid the loss of information through aggregation of the model residual in time (Wagener et al., 2003). This methodology can be applied to track the variation of parameter optima in time, to separate periods of information and noise, or to test whether model components (and therefore parameter values) represent those processes of intention (Wagener et al., 2003).

The Qilian Mountains in northwestern China are the origin of several key inland rivers, including the Heihe, Shiyang, and Shule Rivers (He et al., 2012), and are highly valued for their ecosystem services in conservation of water resources and biodiversity. Urban water supply and irrigation agriculture in the Heihe river basin depend largely on the steady water yield from the mostly non-perennial tributaries in the source regions in the

Qilian Mountains. However, a declining forest cover in recent decades has imposed a potential risk of increased water runoff following heavy rainfall events because of reduced water conservation by vegetation, contributing to highly fluctuating water outputs. The lower altitudinal limit of the forest line retreated from 1,900 m a.s.l. in 1949 to around 2,300 m a.s.l. during the 1990s mainly because of overgrazing damage by goats and cattle and timber harvesting, and, as a consequence, the forest cover decreased from 22.4% to only 12.4% in the Qilian Mountains over the same period (Wang and Cheng, 1999). This, together with the local impacts of global climate change, causes a great concern on declining water conservation capacity of the Qilian Mountains and thus the eco-safety of the region. As a result, great efforts are being directed at assessing the hydrological and ecological consequences of vegetation and climate change in the tributaries of the Qilian Mountains. Hydrological modeling is explored as an operational tool for effective assessment of changes in hydrological processes relating to modification of land cover and climate change.

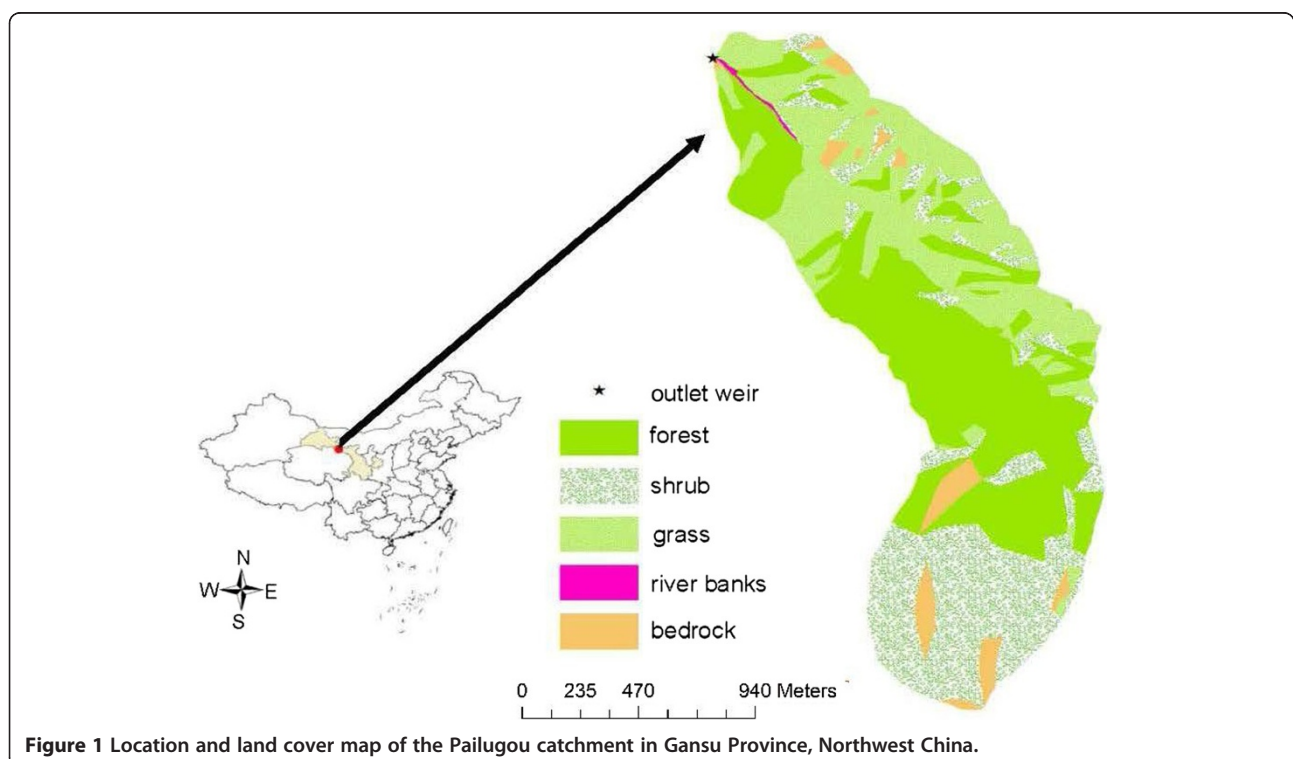
In this study, we investigated the applicability of the HBV-light model (Seibert, 2005) in simulating hydrological processes in the Pailugou catchment of Qilian Mountains, and determined sources and relative contributions of uncertainties in modeling procedures. The Pailugou catchment is a small headwater catchment in the Qilian Mountains, which drains into the Dayehkou

basin and finally feeds into the Heihe River. The vegetation and partial attributes of hydrological processes in the catchment have been intensively investigated by the Academy of Water Resource Conservation Forests of Qilian Mountains in Zhangye, Gansu Province (AWRCFQM). The on-site investigations include a long-term meteorological observation, runoff monitoring, assessment on forest growth and health, and characterization of site conditions. Based on data from the monitoring program of the AWRCFQM and simulations with the HBV-light model, we aim to determine how runoff signatures would help with improving the model calibration, and to identify the periods when intensive monitoring is critically required for deriving parameter values of reduced uncertainty.

Methods

Study catchment

The Pailugou catchment (latitude 38°24'N, longitude 100°17'E, and elevation 2,660–3,788 m a.s.l.) is located in the Qilian Mountains, near Zhangye City, in northwestern China's Gansu province, covering an area of 2.53 km² (Figure 1). Based on the climate record (1990–2010) from the meteorological station at the outlet of the catchment, mean annual temperature is 0.5°C and mean annual precipitation is 378.5 mm. Over 80% of the precipitation falls from June to September (Zheng et al., 2014). Mean annual temperatures decrease with elevation by



0.58°C/100 m and mean annual precipitation increases with elevation by 4.3%/100 m (Wang et al., 2001). The main parental materials in the catchment are calcareous rocks; from these, relatively shallow soils developed, which commonly have a coarse texture, an intermediate organic matter content, and pH values ranging from 7 to 8 (He et al., 2012). The soils are mainly classified as Capsic luvisol, Haplic cambisol, and Hapludoll using the FAO-UNESCO (1988) soil classification system (Yu et al., 2010). Permanently and seasonally frozen soils are widespread at middle and higher elevations. Vegetation comprises patches of forest stands, shrub communities, and pastures. Qinghai spruce (*Picea crassifolia* Kom.) is the only arbor tree species in the catchment and occurs primarily on shaded (north-facing) and semi-shaded (east- or west-facing) slopes at intermediate elevations between 2,600 and 3,300 m a.s.l. The sunny (south-facing) slopes in this altitudinal range are mostly occupied by the grassland plants *Carex lansuensis*, *Pedicularis muscicola* Maxim., and *Polygonum viviparum*. Shrubs, including *Dasiphora fruticosa*, *Caragana jubata* (Pall.) Poir., and *Salix gilashanica*, are mainly found at elevations above 3,300 m a.s.l. (Yu et al., 2010).

Data collection

HBV-light requires input forcing data consisting of daily precipitation and air temperature as well as monthly estimates of potential evapotranspiration. We obtained meteorological data for the full period 2000–2003 from a monitoring station near the catchment outlet at 2,570 m a.s.l. The meteorological data included air temperature, solar radiation, relative humidity, wind velocity, and precipitation. The daily mean air temperature was derived as the arithmetic average of temperatures recorded at 02:00, 08:00, 14:00, and 20:00 h Beijing Standard Time (BST). Monthly mean potential evapotranspiration was calculated from observed meteorological data using the FAO Penman-Monteith method described by Allen et al. (1998).

Runoff was measured manually at the catchment outlet with a V-notch weir, three times a day (i.e., at 08:00, 14:00, and 20:00 h BST) in summer (from May to September), and at a five-day intervals in winter (between October and April), from 1 January 2000 through 31 December 2003. Missing daily values of runoff between October and April were approximated by linear interpolation. Table 1 shows characteristics of the average annual rainfall, runoff, and potential evapotranspiration derived from the available data for the period 2000–2003.

Topographical data were derived from a Digital Elevation Model with a resolution of 1 m, which was produced by the AWRCFQM from laser scanner data. A land use classification for the Pailugou catchment

Table 1 Annual rainfall, runoff, and potential evapotranspiration for the years 2000 to 2003 in the Pailugou catchment

Year	Rainfall (mm a ⁻¹)	Runoff (mm a ⁻¹)	Potential evapotranspiration (mm a ⁻¹)
2000	353	71	842
2001	301	44	874
2002	411	105	764
2003	416	103	772

was obtained from the AWRCFQM (Figure 1); it distinguishes five land use types in the catchment: forest (40.4% of the catchment), grassland (29.5%), shrubland (25.2%), exposed bedrocks (4.7%), and river banks (0.2%). For the modeling, we disintegrated river banks and exposed bedrocks into forest, grassland, and shrubland. Table 2 gives an overview of the vegetation distribution of the three main vegetation types (forest, shrubs, and grassland) for different altitudinal ranges in the catchment.

Model description

The HBV-light model (Seibert, 2005) used in this study is a conceptual rainfall-runoff model modified from the original HBV model by Bergström (1976). There are two minor changes in the modified model corresponding in general to the original version described by Bergström (1992). The first is that, instead of starting the simulation with some user-defined initial state values, the HBV-light v3.0.0.1 uses a “warming-up” period during which state variables evolve from standard initial values to their correct values according to meteorological conditions and

Table 2 Distribution of altitudinal ranges in the Pailugou catchment and corresponding percentage of cover by forest, shrubland, and grassland

Altitude (m a.s.l.)	Percentage of land area (%)	Forest (%)	Shrub (%)	Grass (%)
2,660–2,750	5.04	44.84	8.33	46.83
2,750–2,850	19.19	32.4	10.9	56.7
2,850–2,950	26.99	43.57	8	48.43
2,950–3,050	12.38	72.54	3.31	24.15
3,050–3,150	7.94	67.13	26.2	6.68
3,150–3,250	5.16	84.3	11.43	4.26
3,250–3,350	4.8	46.05	51.87	2.08
3,350–3,450	5.13	6.82	87.33	5.85
3,450–3,550	5.49	92.71	7.29	0
3,550–3,650	5.14	0	100	0
3,650–3,788	2.74	0	100	0

parameter values. Secondly, the restriction that only integer values are allowed for the routing parameter, *MAXBAS*, has been removed to allow the use of all real (non-integer) values.

HBV-light simulates catchment runoff at a daily time step and requires daily values of precipitation and air temperature as well as data on potential evapotranspiration (based on either long-term daily or monthly averages) as forcing variables. It includes four main components: a distributed snow routine, a distributed soil moisture routine, a lumped response routine, and a routing routine. All incoming precipitation first enters the snow routine. Precipitation is simulated to be either snow or rain depending on whether the temperature is above or below a threshold temperature, *TT* (°C). All precipitation simulated to be snow, i.e., falling when the temperature is below the *TT*, is multiplied by a snowfall correction factor, *SFCF* (-). The amount of snow melt, *Melt* (mm d⁻¹), and the refreezing of melt water, *Refreezing* (mm d⁻¹), are calculated, respectively, by:

$$Melt(t) = CFMAX \cdot (T(t) - TT) \quad (1)$$

$$Refreezing(t) = CFR \cdot CFMAX \cdot (TT - T(t)) \quad (2)$$

where *T* (°C) is the mean daily air temperature, *CFMAX* (mm d⁻¹ °C⁻¹) is the degree-day factor, *CFR* (-) is the refreezing coefficient, and *t* is time.

The sum of rainfall and snowmelt from the snow routine enters the soil moisture routine, which calculates the changes in soil moisture storage as the difference between effective precipitation (rain or snowmelt), *P* (mm d⁻¹), and actual evapotranspiration, *ETA* (mm d⁻¹). *ETA* is calculated from potential evapotranspiration, *ETP* (mm d⁻¹), by a linear function of the soil moisture storage, *SM* (mm):

$$ETA(t) = ETP(t) \cdot \min\left(\frac{SM(t)}{FC \cdot LP}, 1\right) \quad (3)$$

where *FC* (mm) is the maximum possible soil moisture storage, and *LP* (-) indicates the relative filling of the soil moisture storage above which *ETA* reaches *ETP*.

The seepage from the soil moisture storage (i.e., the contribution of the effective precipitation to the groundwater module), ΔR (mm d⁻¹), is calculated as a non-linear function of the current filling of the soil moisture storage, *SM* (mm), by

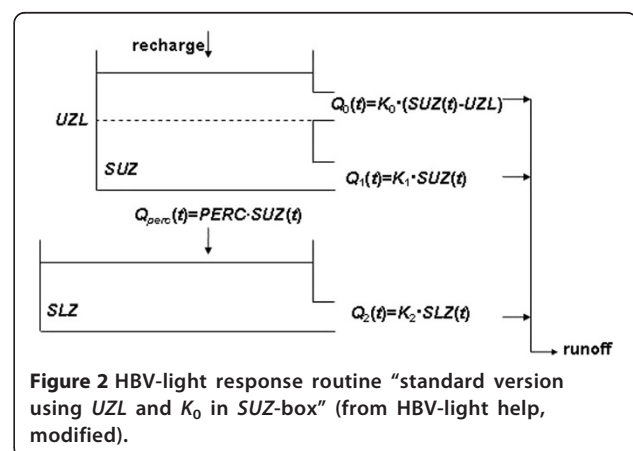
$$\Delta R(t) = P(t) \left(\frac{SM(t)}{FC}\right)^{BETA} \quad (4)$$

where *BETA* (-) is an empirical shape parameter.

Excess water from the soil moisture zone replenishes the groundwater storage, which in our case is configured as the “standard version using *UZL* and *K₀* in *SUZ*-box” (Figure 2). The system consists of two conceptual groundwater boxes: an upper box with two outflows (fast runoff *Q₀* and delayed runoff *Q₁*) with different recession coefficients, and a lower box with one outflow (slow baseflow *Q₂*). Recharge from precipitation or snow melt firstly enters the upper groundwater box. *Q₀* becomes active only when the water level in the upper groundwater box, *SUZ* (mm), exceeds the threshold filling *UZL* (mm). The percolation from the upper to the lower groundwater box, *Q_{perc}* (mmd⁻¹), depends on the filling of the upper groundwater box, *SUZ* (mm). The maximum percolation rate from the upper to the lower groundwater box is defined by the parameter *PERC* (mmd⁻¹).

In the routing routine, the total runoff at the catchment outlet (the sum of the outflows from two or three linear reservoirs depending on whether the water level in the upper groundwater box, *SUZ*, is above *UZL*) is computed using an equilateral triangular weighting function with the base *MAXBAS*.

With the designated model structure, there are a total of 34 parameters involved. We simplified the model structure by fixing the generally less sensitive parameter *CWH* at a value of 0.2, based on the suggestion by Uhlenbrook et al. (1999). The three vegetation zones were not differentiated for the other snow routine parameters (*TT*, *CFMAX*, *SFCF*, *CFR*), hence *TT_{forest}* = *TT_{shrub}* = *TT_{grass}* = *TT*, etc. With this, the final model structure comprises 21 free parameters. We further constrained possible parameter values by defining the following bounds: *FC_{forest}* > *FC_{grass}* > *FC_{shrub}* (taking into account the measurements by Wang et al., 2005), *BETA_{forest}* > *BETA_{shrub}* > *BETA_{grass}*, and *LP_{forest}* < *LP_{shrub}* < *LP_{grass}*.



Objective function definition

Assessing performance of a hydrological model requires estimates of the “closeness” of the simulated behavior of the model to the observations. In this study, a number of efficiency criteria (or objective functions) were used to evaluate the model performance, each emphasizing on a specific type of simulated and observed behavior. We used the coefficient of determination (R^2) and Nash-Sutcliffe efficiencies (R_{eff} , $R_{eff,log}$) to describe the model fit with respect to the entire hydrograph:

$$R^2 = \frac{\left(\sum (Q_{obs}(t) - \overline{Q_{obs}}) \cdot (Q_{sim}(t) - \overline{Q_{sim}})\right)^2}{\sum (Q_{obs}(t) - \overline{Q_{obs}})^2 \cdot \sum (Q_{sim}(t) - \overline{Q_{sim}})^2} \quad (5)$$

$$R_{eff} = 1 - \frac{\sum (Q_{obs}(t) - Q_{sim}(t))^2}{\sum (Q_{obs}(t) - \overline{Q_{obs}})^2} \quad (6)$$

$$R_{eff,log} = 1 - \frac{\sum (\ln Q_{obs}(t) - \ln Q_{sim}(t))^2}{\sum (\ln(Q_{obs}(t) - \ln \overline{Q_{obs}}))^2} \quad (7)$$

where $Q_{obs}(t)$ and $Q_{sim}(t)$ are the observed and predicted runoff at time step t , respectively, and $\overline{Q_{obs}}$ and $\overline{Q_{sim}}$ are the mean values of observed and simulated runoff, respectively. Values of R^2 vary between 0 and 1, values of R_{eff} and $R_{eff,log}$ between $-\infty$ and 1. R_{eff} emphasizes runoff peaks, while $R_{eff,log}$ is more sensitive to the model performance during low flow.

In addition to the statistical efficiency measures, runoff signatures including the volumetric efficiency (S_{VE}), flow duration curve (S_{FDC}), the peak flow (S_{PQ}), and the time to peak (S_{PT}) were used to assess model performance. The volumetric efficiency S_{VE} represents the fraction of water delivered at the correct time and ranges from 0 to 1 (perfect fit):

$$S_{VE} = 1 - \frac{\sum |Q_{obs}(t) - Q_{sim}(t)|}{\sum Q_{obs}(t)} \quad (8)$$

The flow duration curve was used as a second runoff signature. The flow duration curve represents the relationship between the magnitude and the frequency of runoff, providing an estimate of the percentage of time the runoff was equaled or exceeded over a given time period. The objective function is defined as:

$$S_{FDC} = 1 - \frac{\sum_{i=1}^n (Q^o - Q^s)^2}{\sum_{i=1}^n (Q^o - \overline{Q^o})^2} \quad (9)$$

where Q^o and Q^s are the observed and simulated runoff corresponding to a given percentage of exceedance, i , in the flow duration curve ($i = [0, 1, 2, \dots, 100]$).

$\overline{Q^o}$ is the mean of the observed runoff of all exceedance percentages. Values of S_{FDC} vary between $-\infty$ and 1.

Another criterion was used to evaluate the model performance with respect to the simulation of peak runoff:

$$S_{PQ} = 1 - \frac{\sum |Q_{peak,obs}(t) - Q_{peak,sim}(t)|}{\sum Q_{peak,obs}(t)} \quad (10)$$

where $Q_{peak,sim}(t)$ and $Q_{peak,obs}(t)$ are simulated peak runoff and observed peak runoff at time t , respectively. Peaks in the runoff time series were defined as days for which the preceding day and the following day both had smaller runoff values than the present day. Values of S_{PQ} vary between 0 and 1.

The correct timing of the simulated runoff peaks was assessed by:

$$S_{PT} = 1 - \frac{\sum |T(Q_{peak,obs}) - T(Q_{peak,sim})|}{\max(\sum |T(Q_{peak,obs}) - T(Q_{peak,sim})|)} \quad (11)$$

where $T(Q_{peak,sim})$ is the day of the simulated peak runoff, and $T(Q_{peak,obs})$ is the day of the observed peak runoff. Only peaks with $T(Q_{peak,sim}) - T(Q_{peak,obs}) < 4$ were considered in Eqs. 10 and 11. Values of S_{PT} vary between 0 and 1.

The different objective functions given above judge the goodness of a certain parameter set focusing on different aspects in the runoff characteristics. One parameter set can, for example, give a good model performance according to R^2 but only a poor performance in terms of R_{eff} and vice versa. In this study we combined the objective functions of statistical measures (R^2 , R_{eff} , $R_{eff,log}$) with those for the runoff signatures (S_{VE} , S_{FDC} , S_{PQ} , S_{PT}) in order to obtain a best compromise of the parameter fit, satisfying as best as possible most of the objective functions under consideration. The combined objective function, C_{OF} , is a weighted sum of the different objective functions:

$$C_{OF} = w_1 \cdot R^2 + w_2 \cdot R_{eff} + w_3 \cdot R_{eff,log} + w_4 \cdot S_{VE} + w_5 \cdot S_{FDC} + w_6 \cdot S_{PQ} + w_7 \cdot S_{PT} \quad (12)$$

The weights $w_1, w_2, w_3, \dots, w_7$ in Eq. 12 were chosen based on the parameter sensitivity with respect to the corresponding objective functions. The weight for

each objective function was derived from the standard weight:

$$w_{standard}(p_i) = \begin{cases} 1/n_s(p_i) & n_s(p_i) > 0 \\ 1/n & n_s(p_i) = 0 \end{cases} \quad (13)$$

where n is the number of objective functions included in Eq. 12 (in our case, $n = 7$) and n_s is the total number of at least slightly sensitive objective functions (according to our classification, see next section) with respect to parameter p_i . The standard weight was then multiplied by a “sensitivity factor” $c_{OF}(p_i)$ accounting for the different sensitivities of the different objective functions (OF) with respect to the same parameter (p_i):

$$w_{OF}(p_i) = c_{OF}(p_i) \cdot w_{standard}(p_i) \quad (14)$$

The “sensitivity factor” is 7 if the objective function OF (i.e., either R^2 , R_{eff} , $R_{eff,log}$, S_{VE} , S_{FDC} , S_{PQ} , S_{PT}) is highly sensitive with respect to parameter p_i . Similarly, $c_{OF}(p_i) = 3$ for moderately sensitive objective functions, $c_{OF}(p_i) = 1$ for slightly sensitive objective functions, and $c_{OF}(p_i) = 0$ for insensitive objective functions.

Regional sensitivity analysis

A RSA (Hornberger and Spear, 1981) was performed to distinguish between the sensitive model parameters, which have a large impact on the model output, and the non-sensitive model parameters using a Monte Carlo procedure. For the Monte Carlo simulations, 10,000 parameter sets were generated by sampling from a uniform distribution within the given range for each parameter (Table 3). The Monte Carlo sets were split into two groups yielding either “behavioral model runs” or “non-behavioral model runs”. Distinction between the behavioral runs and the non-behavioral runs was made according to the model’s performance. We assigned the runs yielding the 500 (5% of all runs) highest objective function values to the class of behavioral runs; all other runs were classed into the non-behavioral runs. The Kolmogorov-Smirnov two-sample test was used to determine whether the cumulative distribution of the parameter values in the group of behavioral model runs was significantly different from the group of non-behavioral model runs. The Kolmogorov-Smirnov test calculates a test statistic from the maximum distance D between two cumulative distribution functions, $F(p_b)$ and $G(p_n)$, by:

$$D = \max|F(p_b) - G(p_n)| \quad (15)$$

Table 3 Model parameters and their value ranges (lower and upper limits) used in the Monte Carlo runs

Index	Parameter	Lower limit	Upper limit	Units
Catchment				
1	$PCALT$	4.0	14.0	%/100 m
2	$TCALT$	0	0.5	°C/100 m
Snow routine				
3	TT	-2.5	2.5	°C
4	$CFMAX$	1.0	9.0	$\text{mmd}^{-1} \text{°C}^{-1}$
5	$SFCF$	0.2	0.65	-
6	CFR	0	0.8	-
Soil routine				
7	FC_{forest}	200	580	mm
8	FC_{shrub}	25	300	mm
9	FC_{grass}	30	568	mm
10	LP_{forest}	0	0.60	-
11	LP_{shrub}	0.50	0.89	-
12	LP_{grass}	0.70	0.90	-
13	$BETA_{forest}$	3.0	6.0	-
14	$BETA_{shrub}$	2.0	5.0	-
15	$BETA_{grass}$	1.0	3.0	-
Response routine				
16	$PERC$	0.01	3.0	mm d^{-1}
17	UZL	15	70	mm
18	K_0	0.4	1.0	d^{-1}
19	K_1	0.035	0.20	d^{-1}
20	K_2	0.020	0.035	d^{-1}
Routing routine				
21	$MAXBAS$	6.0	11.0	d

where $F(p_b)$ is the cumulative distribution function for the behavioral model runs, $G(p_n)$ is the cumulative distribution function for the corresponding non-behavioral model runs, p_b are the behavioral parameter sets, and p_n are the non-behavioral parameter sets. We grouped the parameter sensitivity into four categories based on the test statistic D and the corresponding P value: highly sensitive ($D > 0.2$, $P \leq 0.05$), moderately sensitive ($0.1 \leq D \leq 0.2$, $P \leq 0.05$), slightly sensitive ($D < 0.1$, $P \leq 0.05$), and insensitive ($P > 0.05$).

Uncertainty analysis

The uncertainty in the simulated runoff is assessed using the GLUE method (Beven and Binley, 1992; Beven and Freer, 2001), which is based on the concepts of RSA. Performance of the GLUE analysis includes the following steps, with steps i to iii being identical to the RSA procedure: i) a large number of model runs with randomly

chosen parameter sets selected from a chosen probability distribution; ii) definition of the likelihood function (Eqs. 1 to 8) and calculation of likelihood values corresponding to the parameter sets; iii) selection of a cutoff threshold value or a fixed percentage of the number of sample parameter sets for the likelihood function to distinguish between the behavioral parameter sets and the non-behavioral parameter sets (the runs yielding the 500 highest objective function values [i.e., 5% of the total runs] were classed as behavioral runs, similar to the cut-off used in the RSA analyses); iv) rescaling of the cumulative likelihood values of all behavioral models to unity; and v) calculation of the percentiles of the cumulative distribution of the likelihood measure. GLUE integrates the outputs of all behavioral models in an ensemble prediction. For each time step of the simulation, the output prediction is obtained as the median of the distribution of all ensemble members, and its uncertainty bounds are estimated as 2.5% and 97.5% percentiles of the distribution.

Dynamic identifiability analysis (DYNIA)

DYNIA was developed by Wagener et al. (2003) and is based on elements from both RSA and GLUE. Similar to RSA, DYNIA calculates the probability distribution of parameter values in behavioral parameter sets, but doing so for each individual model time step. It estimates the parameter sensitivity and derives from this the amount of information available for identifying a specific parameter at a given time. Periods of high parameter sensitivity contain a large amount of information for identifying a given parameter; following Wagener et al. (2003), we term them periods of high “parameter identifiability”. The development of parameter identifiability over time can also be used to detect failure of model structures, as was shown by Wagener et al. (2003). The DYNIA procedure begins with the same Monte Carlo simulations as performed for the RSA and GLUE analyses. However, rather than calculating an error criterion which integrates over the entire simulation period (as in Eqs. 5 to 12), DYNIA estimates an error for each individual model time step. The model error at a given time is taken as the mean squared error for a moving window of $2n + 1$ time steps around the current time step:

$$MSE(t) = \frac{1}{(2n + 1)} \left(\begin{aligned} &Q_{obs}(Q_{obs}(t-n) - Q_{sim}(t-n))^2 + (Q_{obs}(t-n-1) - Q_{sim}(t-n-1))^2 + \dots \\ &+ (Q_{obs}(t) - Q_{sim}(t))^2 + \dots \\ &+ (Q_{obs}(t+n-1) - Q_{sim}(t+n-1))^2 + (Q_{obs}(t+n) - Q_{sim}(t+n))^2 \end{aligned} \right) \quad (16)$$

Taking into account considerations of Wagener et al. (2003) and based on previous experiences in other applications, we used a window size of 5 days (i.e., $n = 2$) for all parameters.

For each individual model time step, the parameter sets are ranked according to the value of the model error, and the top 5% performing parameter sets are taken as the behavioral sets. As in the RSA analysis, the identifiability of each parameter is quantified from the shape of the cumulative likelihood distribution of the parameter values. The parameter ranges are split into m bins (in our study $m = 40$) of equal width, and the gradient of the cumulative likelihood distribution in each bin is calculated from the difference of the cumulative likelihood distribution between adjacent bins. This gradient is an indicator of the identifiability of the parameter: a larger gradient indicates that the parameter value is more likely to be contained in that bin, i.e., the parameter is more constrained in this value range. Hence, the distribution of the parameter values in the bins can be understood as the information content of the runoff data (objective function) for constraining a certain parameter. The information content (IC) of the observation data at a given time step t with respect to the identifiability of a parameter p_i is calculated by:

$$IC_i(t) = 1 - \frac{p_{i,u}(t) - p_{i,l}(t)}{p_{i,max} - p_{i,min}} \quad (17)$$

where $p_{i,u}$ and $p_{i,l}$ are the parameter values at the upper and lower confidence limits at time step t , and $p_{i,max}$ and $p_{i,min}$ are the upper and lower value bounds used in the Monte Carlo sampling (Table 3). IC values range between 0 and 1, with high values indicating a high identifiability.

Results and discussion

Parameter sensitivity

The RSA analyses confirm that model parameter sensitivity is largely dependent upon the objective function used (Table 4). Among the objective functions based on statistical measures, R^2 shows the highest parameter sensitivity. It exhibits a high sensitivity with respect to 7 model parameters, a moderate sensitivity to 3 parameters, and a slight sensitivity to 5 parameters; it is insensitive to 6 parameters. R_{eff} and $R_{eff,log}$ are generally less sensitive. Among the objective functions based on runoff signatures, S_{PT} has the highest

Table 4 RSA parameter sensitivities for various objective functions

Parameters	Statistical measures			Runoff signatures				Combined objective functions	
	R^2 (1)	R_{eff} (2)	$R_{eff, log}$ (3)	S_{VE} (4)	S_{FDC} (5)	S_{PQ} (6)	S_{PT} (7)	C_{OF} (8–28)	
Catchment									
<i>PCALT</i>	***	***	***	**	***	***	***	(8)	***
<i>TCALT</i>	***	*	–	*	*	–	–	(9)	**
Snow routine									
<i>TT</i>	***	***	*	**	**	*	*	(10)	**
<i>CFMAX</i>	*	**	*	*	–	–	–	(11)	**
<i>SFCF</i>	***	***	**	***	–	**	**	(12)	***
<i>CFR</i>	–	–	–	–	–	*	–	(13)	*
Soil routine									
FC_{forest}	*	*	*	–	*	–	–	(14)	–
FC_{shrub}	***	**	**	*	***	***	**	(15)	***
FC_{grass}	*	–	**	–	**	**	*	(16)	*
LP_{forest}	–	–	*	*	–	–	–	(17)	*
LP_{shrub}	**	*	*	–	–	**	**	(18)	–
LP_{grass}	–	–	–	–	–	–	**	(19)	**
$BETA_{forest}$	–	–	–	–	–	–	*	(20)	*
$BETA_{shrub}$	*	*	*	–	–	**	**	(21)	**
$BETA_{grass}$	**	–	*	*	*	***	**	(22)	***
Response routine									
<i>PERC</i>	***	**	**	***	***	***	***	(23)	*
<i>UZL</i>	*	**	–	–	***	–	–	(24)	***
K_0	–	–	–	–	–	–	–	(25)	–
K_1	***	***	*	**	**	*	***	(26)	**
K_2	**	**	**	**	–	*	*	(27)	**
Routing routine									
<i>MAXBAS</i>	–	**	–	**	***	**	***	(28)	***

***, high sensitivity; **, moderate sensitivity; *, slight sensitivity; –, insensitive parameters. The numbers in the parentheses are the indexes for all objective functions: including 3 statistical measures (1–3), 4 runoff signatures (4–7), and 21 combined objective functions (8–28).

parameter sensitivity. S_{PT} exhibits a high sensitivity with respect to 4 parameters, a moderate sensitivity to 6 parameters, and a slight sensitivity to 4 parameters; it is insensitive to 7 parameters. S_{FDC} , S_{PQ} , and S_{VE} are generally less sensitive. Using a weighted combination of all objective functions (last column in Table 4) leads to parameter sensitivities which in most cases lie between the most sensitive objective function and the least sensitive objective function. When applying the combined objective function, 13 out of 21 parameters have at least a moderate sensitivity. However, the sensitivity with respect to *PERC*, FC_{forest} , and LP_{shrub} decreases remarkably. The combined objective function also fails – as do all other objective functions – in identifying the parameter K_0 .

Among the different model routine parameters, the catchment parameter *PCALT* is the most sensitive, with moderate or high sensitivities for all objective functions.

PCALT describes the linear gradient of precipitation with altitude. Since the climate data for the modeling are derived from a monitoring station just below the catchment outlet, the linear extrapolation of the precipitation to the catchment area by means of *PCALT* highly influences the assumed areal precipitation input and, hence, the potential recharge water and the catchment runoff. The second catchment parameter, the temperature gradient *TCALT*, is less sensitive (high sensitivity with respect to R^2 , and moderate sensitivity for C_{OF}). All snow routine parameters are sensitive with respect to most of the objective functions, with the only exception being *CFR*, which is sensitive only with respect to S_{PQ} . The Pailugou catchment experiences long periods of snow cover; runoff almost ceases in winter, and is highly influenced by snow melt and refreezing processes during late spring and summer as well as the start of snow

accumulation in fall. Therefore, the dominant role of the snow routine parameters for the model performance is not astonishing.

The soil moisture routine parameters are generally more sensitive with respect to the runoff signatures, especially with those focusing on runoff peaks (S_{PT} , S_{PQ}). The storage capacity of the soil moisture reservoir, FC , has a larger impact on the model performance than $BETA$, which influences the amount of percolation from the soil moisture storage to the groundwater in times when the soil is not saturated. LP , which describes the reduction of the potential evapotranspiration in drier soils, has the least sensitivity. The sensitivities of $BETA$, LP , and FC vary between different vegetation classes. Although forests cover almost half of the catchment, the parameters of this vegetation class influence the model performance less than the parameters of the two other classes do.

The response routine parameters are generally sensitive with respect to most of the objective functions, except K_0 . The parameter $PERC$, which represents the maximum percolation rate from the upper groundwater box to the lower groundwater box, is the most sensitive parameter of the response routine. The large influence of $PERC$ on the model fit indicates the importance of slow groundwater runoff in the catchment. K_0 is the least sensitive model parameter, and no objective function under consideration identified K_0 . K_0 controls the fast runoff when the filling in the upper groundwater box exceeds the threshold UZL (Figure 2). Precipitation in the Pailugou catchment is generally very low and it is realistic to assume that fast surface or near-surface runoff is a rare event, occurring only after exceptionally

high rain storms. However, the main reason for the low sensitivity of K_0 is most likely the low time resolution of the outflow data and the model, which is larger than the reaction time of fast runoff in this small and very reactive headwater catchment. Contrary to the fast outflow coefficient, K_0 , the second outflow coefficient of the upper groundwater box, K_1 , proves to be highly sensitive, which indicates the importance of fast interflow for the runoff generation in the catchment. K_2 , the recession coefficient of the lower groundwater box, is moderately sensitive to most objective functions and plays a lesser role for the objective functions focusing on runoff peaks (S_{PQ} , S_{PT}). The threshold level, UZL , above which fast runoff from the upper groundwater box occurs, is generally not very sensitive. However, its sensitivity is much increased when considering the flow duration curve as efficiency criteria.

The routing parameter $MAXBAS$ shows a higher sensitivity with respect to the objective functions based on runoff signatures and the combined objective functions than to the more statistical measures.

Uncertainty analysis

Figure 3 shows the value distribution for each analyzed model parameter in the behavioral model runs with respect to the original value range used for the Monte Carlo runs (Table 4). To compare different model parameters, the original value range was scaled to [0, 1]. It is obvious from Figure 3, that for all the objective functions, values in the behavioral runs spread across the entire value range considered. However, when looking at the interquartile ranges of the boxplots, some parameters appear to be more constrained than others. Noticeable differences in

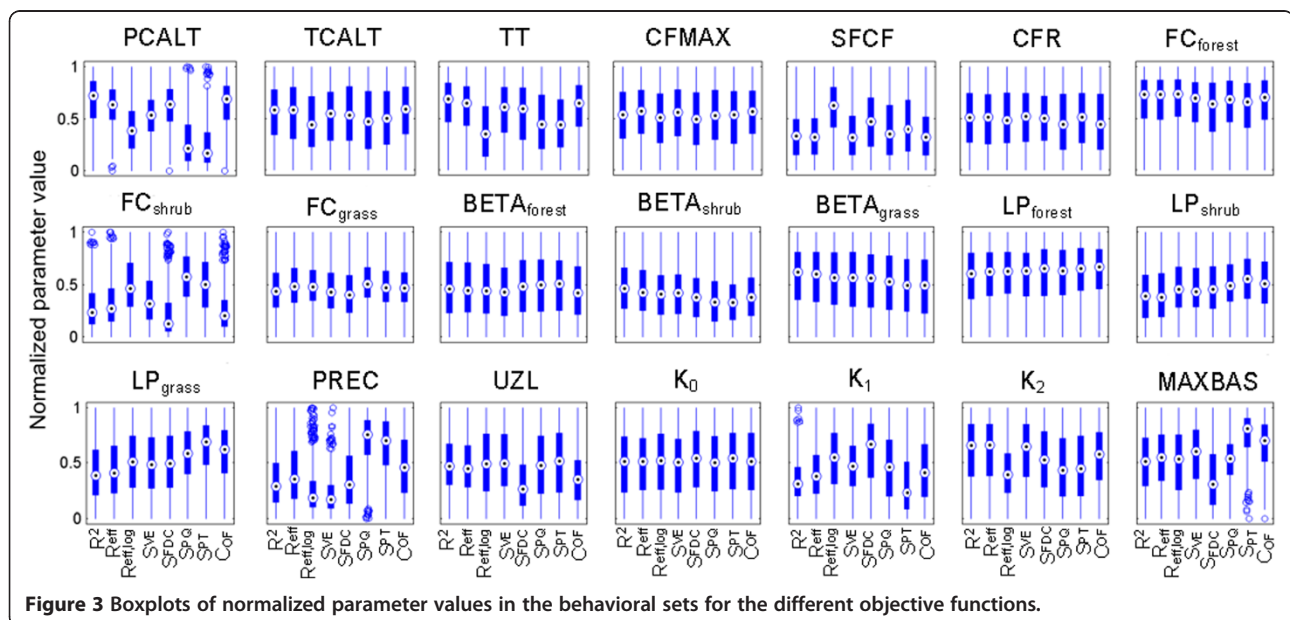


Figure 3 Boxplots of normalized parameter values in the behavioral sets for the different objective functions.

the constraining of the parameters are found between different objective functions. The routing parameter *MAXBAS* is constrained in the interquartile range to around 25% of the original value range when using the runoff-peak oriented objective functions S_{PT} or S_{PQ} , or the combined objective function. It is much less constrained when using any of the other objective functions. Similarly, the groundwater recharge parameter *PERC* is much more constrained by the volume efficiency (S_{VE}) or $R_{eff,log}$, which both put a larger weight on low flow conditions. Using combined objective functions is beneficial especially for the identification of the soil routine parameters. Five out of nine soil routine parameters are best identified when using the combined objective functions, and for the others, the objective functions rank 2nd or 3rd with respect to constraining of the interquartile range.

As clearly illustrated in Figure 3, some parameters are constrained in different value ranges depending on which objective function is used to assess the model behavior. For example, values of *PCALT* in the behavioral runs are relatively large when considering R^2 , R_{eff} or C_{OF} as the objective function, and are significantly lower when using S_{PQ} or S_{PT} . The snow correction factor, *SFCF*, attains higher values in the behavioral runs when considering $R_{eff,log}$ as the objective function. Values of the response routine *PERC* are particularly low when based on S_{PQ} and S_{PT} . The values of the routing routine *MAXBAS* are especially low when based on S_{FDC} , but much higher values are attained when based on other objective functions.

Figure 4 displays the uncertainty bands (e.g., lower and upper bounds of the 95% confidence intervals) of the GLUE estimates, their median, and the observed runoff for some of the objective functions (R^2 , $R_{eff,log}$, S_{FDC} , C_{OF2}). The GLUE results indicate that HBV-light is generally capable of simulating the runoff in the Pailugou catchment, and most of the time yields a good agreement with the observed runoff for all objective functions under consideration, although most models generally underestimate peak runoffs. The GLUE simulations based on the various objective functions diverge greatly, in particular following pronounced snow melt events (Figure 5). Using the objective functions R^2 , R_{eff} , S_{FDC} , C_{OF2} , C_{OF7} , C_{OF8} , C_{OF9} , C_{OF17} , and C_{OF21} leads to a general over-prediction of the runoff; the other objective functions systematically under-predict the runoff (Figure 5). C_{OF2} has the lowest absolute cumulative difference with a cumulated value at the end of the simulation period of 3.3 mm above the observations; C_{OF12} has the largest absolute cumulative difference with a value of 161.2 mm of the objective functions above the observations (Figure 5). An obvious systematic error in the simulations may originate from the linear interpolation of the meteorological input data from one meteorological station (below the catchment outlet), by using linear altitudinal gradients for precipitation and temperature (*PCALT*, *TCALT*). Another source of error could be the daily time steps of the HBV-light simulations, which may be too coarse to adequately describe the rainfall-runoff transformation during high-intensity rainfall or snow melt events. Moreover, the daily discharges

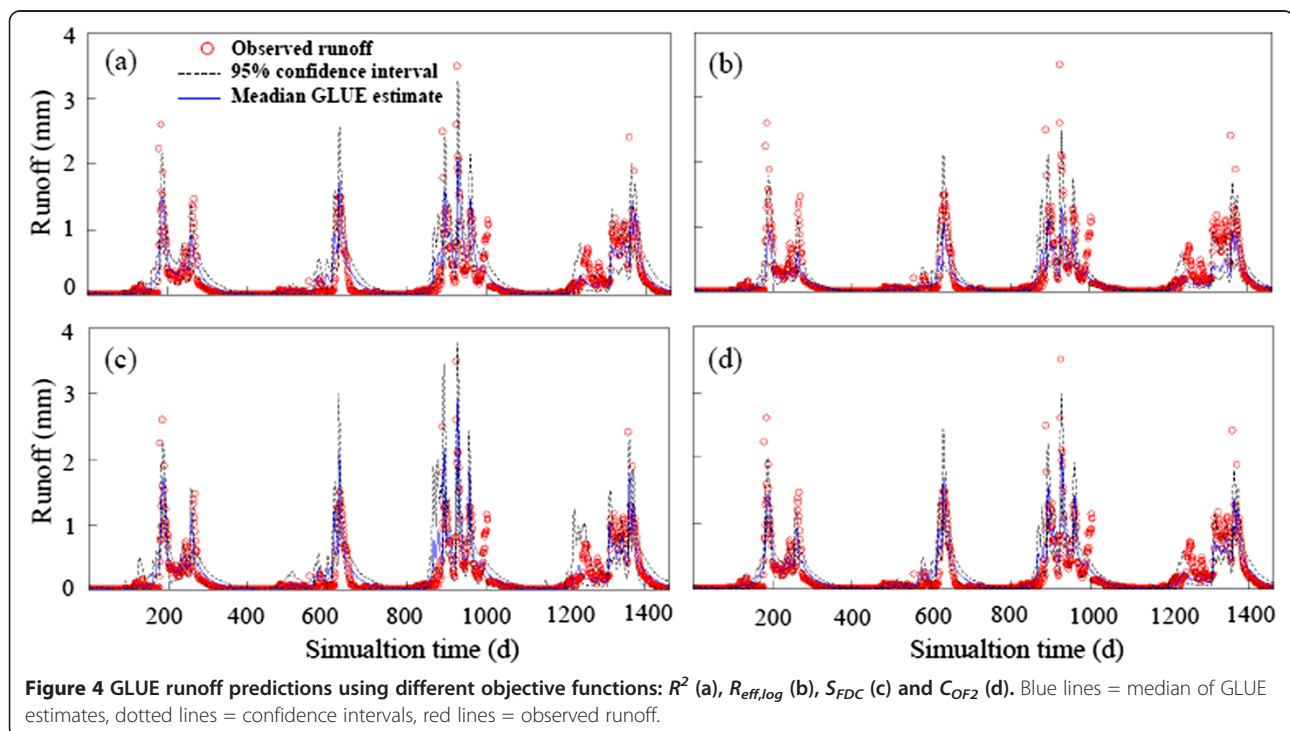


Figure 4 GLUE runoff predictions using different objective functions: R^2 (a), $R_{eff,log}$ (b), S_{FDC} (c) and C_{OF2} (d). Blue lines = median of GLUE estimates, dotted lines = confidence intervals, red lines = observed runoff.

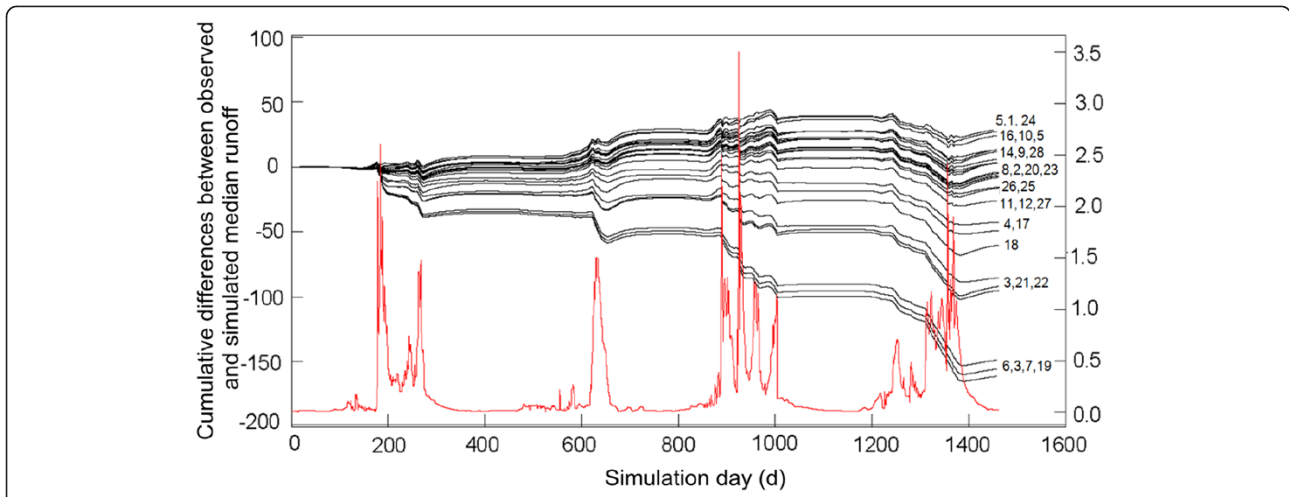


Figure 5 Cumulative differences between observed and simulated median runoff for the various objective functions. Black lines = cumulative difference, red line = observed runoff. The numbers are the indexes for all objective functions including the 3 statistical measures, 4 runoff signatures, and 21 combined objective functions in Table 4.

derived from the water level measurements at three times daily (at 08:00, 14:00, and 20:00 h BST) may have smoothed the very rapid flow characteristics at the site, failing to capture the dynamic nature of the rainfall-runoff transformation.

A desirable model fit would go along with a high precision (i.e., narrow confidence bands of the GLUE simulations) and a high accuracy (i.e., a large percentage of observations being enclosed by the confidence bounds). The precision and the accuracy of the GLUE runs based on the various objective functions are displayed in Figure 6. Both, the precision and the accuracy vary between the different objective functions. As an example, simulations based on $R_{eff,log}$ (index 3 in Figure 6) as objective function yield very narrow confidence bounds (median width of 0.07 mm), but they contain only about 60% of the observations. Conversely, using S_{FDC} (index 5

in Figure 6) implies a high uncertainty in the modeled runoff, but the wide confidence bounds (median width = 0.17 mm) include more than 80% of the observations. Figure 6 suggests the combined objective function with respect to parameter $TCALT$ (index 9 in Figure 6) to be a favorable objective function for the model conditioning. Using this objective function leads to confidence bounds which are in the intermediate range of all objective functions; at the same time, those relatively narrow confidence bounds contain already almost 80% of the observations.

It should be noted that the estimated model uncertainties are sensitive to the choice of threshold values which distinguish behavioral and non-behavioral model runs, which has been often considered as one of the main drawbacks of the GLUE technique (e.g., Montanari, 2005; Blasone et al., 2008b). However, in this study, we did not

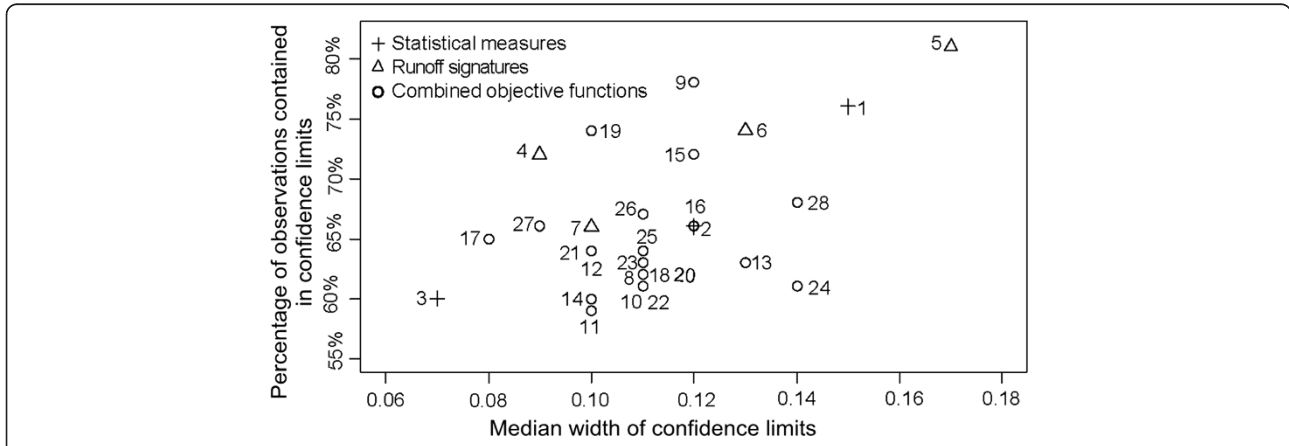


Figure 6 Relationships between percentage of observations contained in confidence limits and median width of confidence limits.

investigate how sensitive the model simulation results are to the cut off threshold values, therefore, further studies need to investigate how the threshold value should be chosen in order to provide stabilization (may be difficult) in the application of the GLUE method.

Temporal changes of parameter sensitivity

DYNIA was applied to analyze the temporal changes in the parameter identifiability over the period 2000–2003 for each of the 21 model parameters. Parameters are usually found to have specific periods where they play a more pronounced role for the simulated runoff and are therefore more sensitive – hence better identifiable – than in other periods (Figure 7).

Figure 7a displays the development of the *IC* (Eq. 13) over time for the catchment parameters (*PCALT*, *TCALT*) and the parameters of the snow routine (*TT*, *CFMAX*,

SFCF, *CFR*). The *IC* of *PCALT* varies largely over the simulated time period; it is highest during the rainy season, when the soil moisture storage and the groundwater storages are filled and additional precipitation produces runoff, and it decreases continuously during recession periods. The *IC* for *TCALT* is significantly higher in periods with active vegetation, when the soil moisture is high. The parameters of the snow routine – with exception of *CFR* which has a low *IC* over the entire simulation time – are generally better identifiable during the snow melt period, especially during the first main melt events in spring.

Figure 7b displays the development of the *IC* over time for the parameters of the soil moisture routine. *FC_{shrub}* shows a pronounced temporal dependency of the *IC*; it is more identifiable during the early vegetation period when the soil moisture storage is replenished by melt

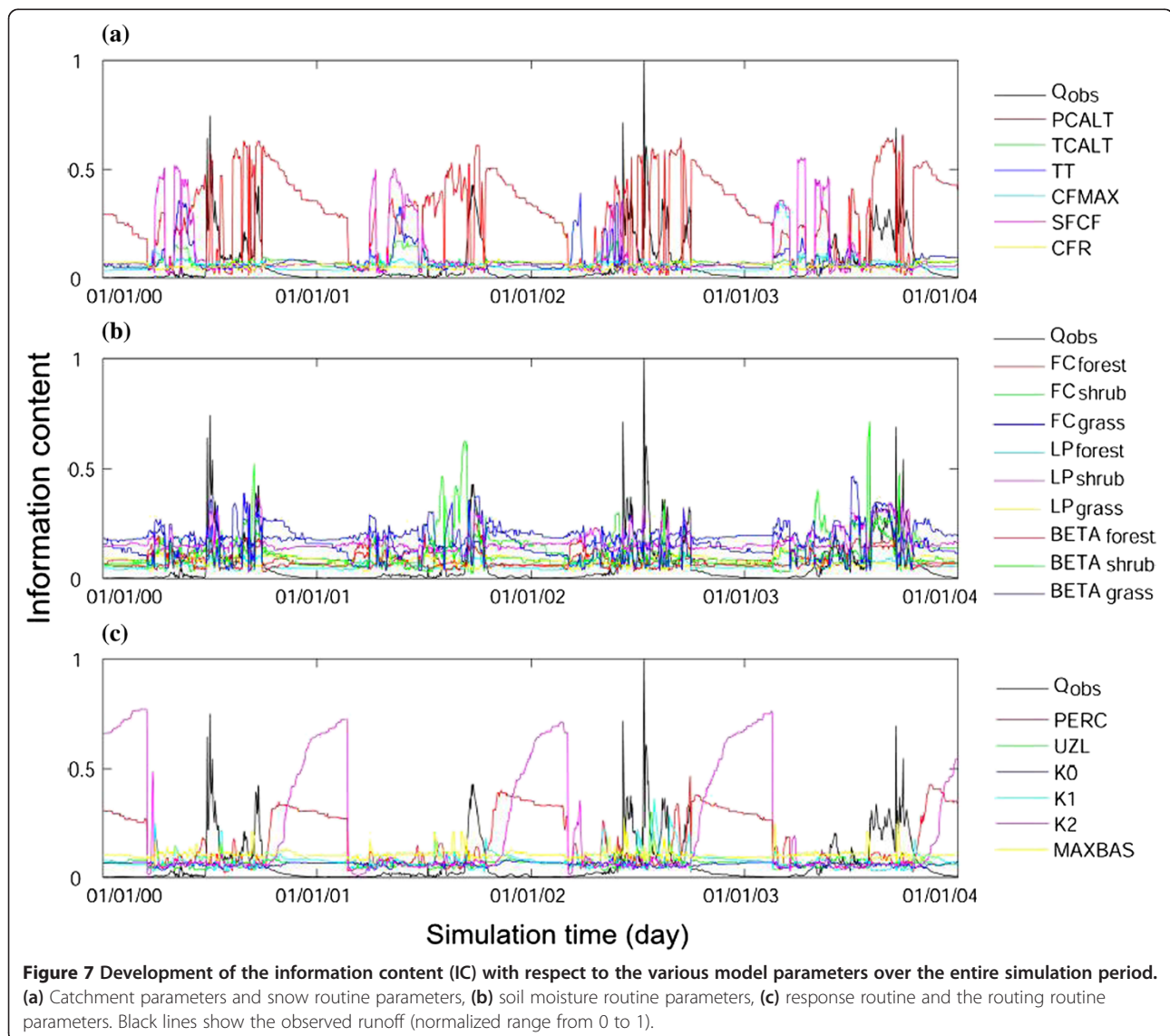


Figure 7 Development of the information content (*IC*) with respect to the various model parameters over the entire simulation period. **(a)** Catchment parameters and snow routine parameters, **(b)** soil moisture routine parameters, **(c)** response routine and the routing routine parameters. Black lines show the observed runoff (normalized range from 0 to 1).

water. The *ICs* for FC_{forest} and FC_{grass} show similar patterns, both exhibiting a somewhat higher identifiability during the periods when the catchment is already relatively wet and further precipitation increases the proportion of faster runoff components in the total catchment runoff. The *ICs* of *LP* and *BETA* of the same vegetation class show similar patterns; *LP* and *BETA* are generally better identifiable prior to runoff peaks, in times when the catchment's storages are filling.

The *IC* for the parameters of the response routine and the routing routine (Figure 7c) is directly linked to the dynamics of the runoff peaks. The *IC* for *PERC* is higher in early winter, in the course of declining percolation from the upper groundwater box to the lower groundwater box. The three recession constants, K_0 , K_1 and K_2 , show very distinct patterns of *IC*. While the *IC* of the fast runoff component (K_0) is always very low, that of K_1 is markedly increased in the falling limbs of runoff peaks and decreases with very high runoff events. The dominant role of the parameter K_1 for controlling recession after peak runoffs underlines the importance of fast subsurface flow in the catchment. The *IC* of K_2 , which influences the dynamics of the slow groundwater runoff, continuously increases during low flow periods, when recharge from the soil zone ceases and the base-flow from the lower groundwater box becomes the main runoff source. *UZL* becomes somewhat better identifiable in late summer when the catchment is already relatively wet from the summer rains, and additional precipitation causes the upper groundwater box to exceed the threshold filling *UZL* and initiates fast runoff. Not surprisingly, the routing parameter *MAXBAS* is more identifiable during pronounced runoff peaks following snow melt and in the wet season, especially in the rising limbs of the runoff peaks.

Temporal changes of optimal parameter values

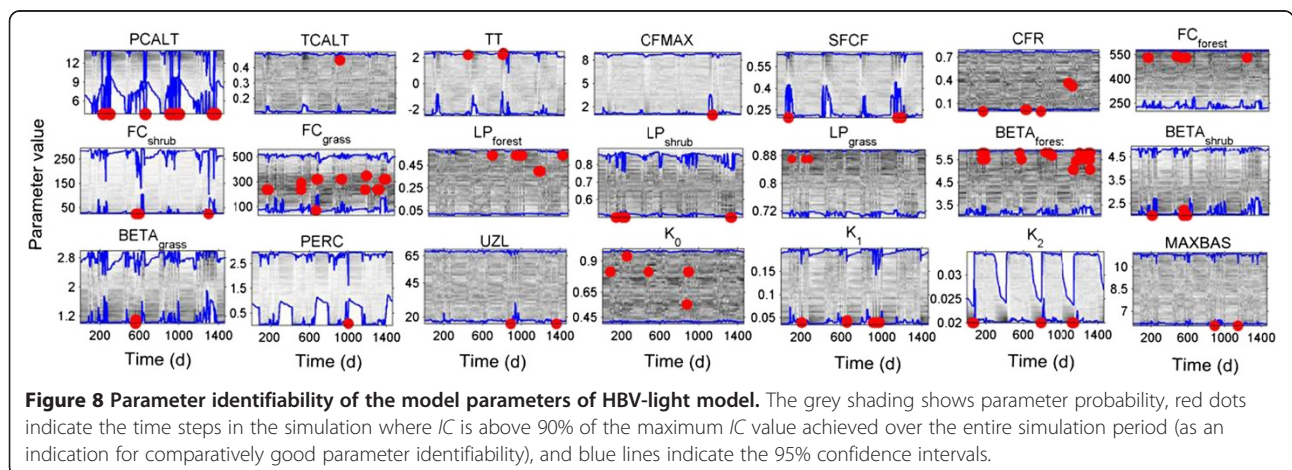
Figure 8 shows the *IC* of the error criterion (Eq. 12) with respect to a model parameter as well as the location of

the parameter values with the highest probability of occurrence in the behavioral runs. In Figure 8, darker grey indicates a higher probability density of the parameter value in the behavioral runs, indicating time periods when the parameter values are better identifiable, while lighter grey indicates time periods when the parameter cannot be identified. The blue lines show the 95% confidence limits for the parameter estimate at a given time step. The red dots in Figure 8 indicate the time steps in the simulation where *IC* is above 90% of the maximum *IC* value achieved over the entire simulation period, using this as an indication for comparatively good parameter identifiability.

The DYNIA analysis reveals different types of parameter behavior. The optimum values (i.e., red dots in Figure 8) of 11 model parameters (i.e., *PCALT*, *TCALT*, *CFMAX*, *SFCF*, FC_{shrub} , LP_{shrub} , LP_{grass} , *PERC*, *UZL*, K_2 , and *MAXBAS*) are constant over time (Figure 8). For these parameters, the same value would be identified, regardless of the time period used for the model conditioning/calibration. For five more parameters (i.e., *TT*, FC_{forest} , $BETA_{shrub}$, $BETA_{grass}$, K_1), the variation of the optimum parameter values (i.e., red dots) is less than 10% of the original parameter range, also indicating the possibility of a relatively stable and time-invariant parameter identification. For the other parameters (*CFR*, LP_{forest} , $BETA_{forest}$, K_0 , FC_{grass}), the optimum values of these parameters shift over the time domain (Figure 8). This can be attributed to the very low sensitivities (Figure 7) of these parameters or inadequacies within the model structure. Our results indicate the importance of identifying the periods when intensive monitoring is critical for deriving parameter values of reduced uncertainty.

Conclusions

In this study, with the objective of evaluating the model performance of a conceptual semi-distributed rainfall-



runoff model, the well-known HBV-light model is applied in a small headwater catchment in Qilian Mountains in Northwest China using RSA, GLUE, and DYNIA frameworks. Several main conclusions can be drawn from this study:

1. The results of RSA show that model parameter sensitivity is largely dependent upon the objective function used for the model evaluation in the sensitivity analysis. Most of the model parameters are sensitive when the runoff signatures and combined objective functions are used. The time resolution of the runoff observations and the HBV-light simulations is too coarse to satisfactorily describe the fast runoff processes in the catchment. More frequent runoff observations would substantially increase the knowledge on the rainfall-runoff transformation in the catchment and, specifically, improve the distinction of fast surface-near runoff and interflow components in their contribution to the total catchment runoff.
2. The results of GLUE show that the HBV-light model is generally able to simulate the runoff in the Pailugou catchment with an acceptable accuracy. However, a distinct pattern of mismatch is found in some high-intensity rainfall/snow melt events at a daily step. Most parameters are well constrained, showing higher parameter identifiability and lower model uncertainty when runoff signatures or the combined objective functions are used. The combined objective function focusing on the catchment parameter *TCALT* performed best in terms of model uncertainty and model precision.
3. The DYNIA analysis shows different types of parameter behavior. The optimum values of 11 model parameters are constant over time regardless of the time period used for the model conditioning/calibration. For 5 parameters, the variation of the optimum parameter values is less than 10% of the original parameter range, also indicating the possibility of relatively stable and time-invariant parameter identification. For the other 5 parameters optima change over the time domain. All of these indicate that model parameters have specific periods where they are more sensitive, more identifiable, and where they play a clearer role than during other periods. The hydrological process of snow routine could be better described if monitoring is intensified during snow melt. Our results also highlight the importance of identifying the periods when intensive monitoring is critical to derive parameter values of reduced uncertainty.

Changes in climate and/or land cover have significant implications to rainfall-runoff dynamics at watershed or catchment scales, which in turn affect regional ecosystem processes. Hydrological models are important tools for evaluating the potential impacts of climate change and land cover change on the hydrological cycles and runoff regimes. However, uncertainty in model parameters due to a lack of identifiability may greatly limit the use of models for purposes such as parameter regionalization or the investigation of land use or climate. Sensitivity analysis and identification of parameters with significant implications to changes in landscape features are a critical step in studying regional ecosystem processes in response to natural or anthropogenic perturbations. A higher identifiable parameter can reduce model uncertainty and is critical for evaluating the effects of climate change and land use disturbance on the hydrological cycles.

It should be noted here that the generality of results and conclusions of this study need to be verified through the application of HBV-light in other regions.

Abbreviations

AWRCFQM: Academy of Water Resource Conservation Forests of Qilian Mountains in Zhangye; *BETA*: An empirical shape parameter; *BST*: Beijing Standard Time; *CFMAX*: The degree-day factor; *CFR*: The refreezing coefficient; *DYNIA*: Dynamic Identifiability Analysis; *FC*: Maximum soil moisture storage; *GLUE*: Generalized Likelihood Uncertainty Estimation; *IC*: Information content; K_0 : Recession coefficients of fast runoff; K_1 : Recession coefficients of delayed runoff; K_2 : Recession coefficients of low base-flow runoff; *LP*: Soil moisture value above which actual evapotranspiration above potential evapotranspiration; *MAXBAS*: A variable of length of triangular weighting function; *PCALT*: Change of precipitation with elevation; *PERC*: Maximum percolation rate from upper to lower groundwater box; R^2 : Coefficient of determination; R_{eff} : Nash-Sutcliffe efficiencies; *RSA*: Regional Sensitivity Analysis; S_{VE} : Volumetric efficiency; S_{FDC} : Flow duration curve; S_{PQ} : Peak flow; S_{PT} : Time to peak; *TCALT*: Change of temperature with elevation; *TT*: Temperature threshold for rain or snow; *UZL*: Threshold for the fast runoff.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

KW and OJS conceived the overall project framework. SO, HP, and SW performed the model simulations and data analysis. SO, HP, and OJS wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This research was jointly funded by Robert Bosch Foundation and Beijing Municipal Commission of Education (Key Laboratory for Silviculture and Conservation). Authors are grateful to the Academy of Water Resource Conservation Forest of Qilian Mountains (AWRCFQM), Zhangye, Gansu Province, China, for organizing the international joint project work and providing field data.

Author details

¹Institute of Forestry and Climate Change Research, Beijing Forestry University, 35 Qinghua East Road, Haidian District, Beijing 100083, China. ²Chair of Hydrology, University of Freiburg, Fahrenbergplatz, 79098 Freiburg i. Br., Germany. ³Academy of Water Resource Conservation Forests in Qilian Mountains, Zhangye 734000, Gansu, Province, China. ⁴Forest Research Institute of Baden-Württemberg, Department of Soils and Environment, Wonnhaldestr. 4, 79100 Freiburg i. Br., Germany.

Received: 21 March 2014 Accepted: 15 May 2014

Published online: 20 June 2014

References

- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration—guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper No. 56. Food and Agriculture Organization of the United Nations, Rome
- Bergström S (1976) Development and application of a conceptual runoff model for Scandinavian catchments. Swedish Meteorological and Hydrological Institute, Report No. RHO 7, Norrköping, Sweden
- Bergström S (1992) The HBV model—its structure and applications. Swedish Meteorological and Hydrological Institute, Report No. RHO 4, Norrköping, Sweden
- Beven KJ (2001) Rainfall-runoff modeling. Wiley, Chichester, UK
- Beven KJ, Binley AM (1992) The future of distributed hydrological models: model calibration and uncertainty prediction. *Hydrol Process* 6(3):279–298
- Beven KJ, Freer J (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. *J Hydrol* 105(1):157–172
- Blasone RS, Madsen H, Rosbjerg D (2008a) Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling. *J Hydrol* 353(1):18–32
- Blasone RS, Vrugt JA, Madsen H, Rosbjerg D, Zvyolosi GA, Robinson BA (2008b) Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Adv Water Resour* 31(4):630–648
- Butts BM, Payne TJ, Kristensen M, Madsen H (2004) An evaluation of the impact of model structure on hydrological modeling uncertainty for stream flow simulation. *J Hydrol* 298(1):242–266
- Dawson CW, Abrahart RJ, See LM (2007) HydroTest: a web-based toolbox of evaluation metric for the standardized assessment of hydrological forecasts. *Environ Modell Softw* 22(7):1034–1052
- Dotto CBS, Mannina G, Kleidorfer M, Vezzaro L, Henrichs M, McCarthy DT, Freni G, Rauch W, Deletic A (2012) Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling. *Water Res* 46(8):2545–2558
- Dotto CBS, Kleidorfer M, Deletic A, Rauch W, McCarthy DT (2014) Impacts of measured data uncertainty on urban storm water models. *J Hydrol* 508:28–42
- Engeland K, Xu CY, Gottschalk L (2005) Assessing uncertainties in a conceptual water balance model using Bayesian methodology. *Hydrol Sci J* 50(1):45–63
- FAO-UNESCO (1998) Soil map of the world, revised legend. FAO World Soil Resources report number 60, Food and Agricultural Organization of the United Nations. UNESCO, Rome
- Freer J, Beven KJ, Ambrose B (1996) Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resour Res* 32:2161–2173
- Gupta HV, Thieman M, Trosset M, Sorooshian S (2003) Reply to comment by Beven K and Young P on “Bayesian recursive parameter estimation for hydrologic models”. *Water Resour Res* 39(5):1117
- Gupta HV, Wagener T, Liu Y (2008) Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol Process* 22(18):3802–3813
- He ZB, Zhao WZ, Liu H, Tang ZX (2012) Effect of forest on annual water yield in the mountains of an arid inland river basin: a case study in the Pailugou catchment on northwestern China's Qilian Mountains. *Hydrol Process* 26(4):613–621
- Henriksen HJ, Troldborg L, Nyegaard P, Sonnenborg TO, Refsgaard JC, Madsen B (2003) Methodology for construction, calibration and validation of a national hydrological model for Denmark. *J Hydrol* 280(1):52–71
- Hornberger GM, Spear RC (1981) An approach to the preliminary analysis of environmental systems. *J Environ Manage* 12(1):7–18
- Jin X, Xu CY, Zhang Q, Singh VP (2010) Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *J Hydrol* 383(3):147–155
- Kavetski D, Clark MP (2010) Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resour Res* 46(10), W10511
- Kavetski D, Kuczera G, Franks SW (2006) Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resour Res* 42(3), W03407
- Kuczera G, Kavetski D, Franks S, Thyer M (2006) Towards a Bayesian total error analysis of conceptual rainfall-runoff models: characterising model error using storm-dependent parameters. *J Hydrol* 331(1):161–177
- Madsen H (2000) Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *J Hydrol* 235(3):276–288
- Madsen H, Wilson G, Ammentorp HC (2002) Comparison of different automated strategies for calibration of rainfall-runoff models. *J Hydrol* 261(1):48–59
- Montanari A (2005) Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water Resour Res* 41(8), W08406
- Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 50(3):885–900
- Pechlivanidis IG, Jackson BM, McIntyre NR, Wheeler HS (2011) Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global Nest J* 13(3):193–214
- Ratto M, Young PC, Romanowicz R, Pappenberge F, Saltelli A, Pagano A (2007) Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology. *Hydrol Earth Syst Sci* 11(4):1249–1266
- Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) Global sensitivity analysis: the primer. Wiley, Chichester
- Sawicz K, Wagener T, Sivapalan M, Troch PA, Carrillo G (2011) Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrol Earth Syst Sci* 15:2895–2911
- Seibert J (2005) HBV light version 2, user's manual. Uppsala University, Uppsala, In, Department of Earth Sciences
- Shakti PC, Shrestha NK, Gurung P (2010) Step wise multi-criteria performance evaluation of rainfall-runoff models using WETSPRO. *J Hydrol Meteor* 7(1):18–29
- Shamir E, Imam B, Gupta HV, Sorooshian S (2005) Application of temporal streamflow descriptors in hydrologic model parameter estimation. *Water Resour Res* 41(6), W06021
- Thyer M, Renard B, Kavetski D, Kuczera G, Franks SW, Srikanthan S (2009) Critical evaluation of parameter consistency and predictive uncertainty in hydrological modelling: a case study using Bayesian total error analysis. *Water Resour Res* 45(12):W00B14
- Thieman M, Trosser M, Gupta H, Sorooshian S (2001) Bayesian recursive parameter estimation for hydrologic models. *Water Resour Res* 37(10):2521–2535
- Uhlenbrook S, Seibert J, Leibundgut C, Rodhe A (1999) Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrol Sc J* 44(5):779–797
- Vrugt JA, ter Braak CJF, Gupta HV, Robinson BA (2009) Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stoch Environ Res Risk A* 23(7):1011–1026
- Wagener T, McIntyre N, Lees MJ, Wheeler HS, Gupta HV (2003) Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. *Hydrol Process* 17(2):455–476
- Wagener T, Montanari A (2011) Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resour Res* 47(6), W06301
- Wang G, Cheng G (1999) Water resource development and its influence on the environment in arid areas of China - the case of the Heihe River Basin. *J Arid Environ* 43(4):121–131
- Wang JY, Chang XX, Ge SL, Miao YX, Chang ZQ, Zhang H (2001) Vertical distribution of the vegetation and water and heat conditions of Qilian Mountains (northern slope). *J Northwest For Univ* 16:1–3
- Wang JY, Tian DL, Wang YH, Wang SL, Zhang XL, Geng SL (2005) Soil hydrological effect of forest and grass complex watershed in Qilian Mountains. *J Soil Water Conserv* 19:144–147
- Xevi E, Christiaens K, Espino A, Sewnandan W, Mallants D, Sørensen H, Feyen J (1997) Calibration, validation and sensitivity analysis of the MIKE-SHE model using the Neuenkirchen catchment as case study. *Water Resour Manage* 11(3):219–242
- Yu PT, Wang YH, Wu XD, Dong XH, Xiong W, Bu GW, Wang SL, Wang JY, Liu XD, Xu LH (2010) Water yield reduction due to forestation in arid mountainous regions, northwest China. *Int J Sediment Res* 25(4):423–430

Zhang A, Zhang C, Fu G, Wang B, Bao Z, Zheng H (2012) Assessments of impacts of climate change and human activities on runoff with SWAT for the Huifa River Basin, Northeast China. *Water Resour Manag* 26(8):2199–2217

Zheng XL, Zhao CY, Peng SZ, Jian SQ, Liang B, Wang XP, Yang SF, Wang C, Peng HH, Wang Y (2014) Soil CO₂ efflux along an elevation gradient in Qinghai spruce forests in the upper reaches of the Heihe River, northwest China. *Environ Earth Sci* 71:2065–2076

doi:10.1186/s13717-014-0014-9

Cite this article as: Ouyang *et al.*: Parameter uncertainty and identifiability of a conceptual semi-distributed model to simulate hydrological processes in a small headwater catchment in Northwest China. *Ecological Processes* 2014 **3**:14.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
