Reply to comments from Dr. Schaefli

General comment

1. **This paper is a re-submission of a paper previously discussed in HESSD. The authors made a considerable effort to revise the text and the model to meet the reviewers' concerns. The model now has a separated degree-day factor for snow and ice and the description of the model is clearer (but still not entirely clear).**

   Reply: Thanks.

2. **The proposed step-wise calibration method is tested to show how robust it is if applied to different periods and with different hydrograph separation criteria. The method is certainly transferable to other catchments and interesting for the readership of HESS and I recommend publication in HESS after minor revisions.**

   Reply: Thanks.

3. **Before giving some detailed comments hereafter, I would like to point out here that I do not agree with the authors' view that an observed time series can be manipulated such as to "expand the measurement dimension". Information can be extracted from data but the information content of data cannot be increased by any manipulation. Could you please comment on this?**

   Reply: Thank you for this suggestion. We have modified the related concepts in the paper. 'measurement dimension' has been modified as 'signature dimension' in the revised paper.

Detailed comments:

1. Abstract

1. 1. In the abstract, the hydrograph is partitioned according to water sources but then "the hydrological model parameters are grouped by the associated *runoff generation mechanism*"; please use coherent wording according to the very first review of the 1st submission to HESSD. Same holds for section 3, and for the conclusion.

Reply: We have modified the 'runoff generation mechanism' as 'runoff water sources' in the revised paper.

1.2. The abstract does not mention any results, conclusions or outlooks, simply summarizes the method.

Reply: We have expanded the abstract section by adding more details about the results *and conclusions:*

"Results show that the proposed calibration approach performed reasonably well. Cross validation and comparison to an automatic calibration method indicated its robustness."

2. Introduction:

2.1. Good literature summary.

Reply: Thanks.

2.2. I do not agree with wording "hydrograph partitioning is another possible way to expand RM". The measurement dimension cannot be expanded otherwise than by adding data; hydrograph partitioning might help to extract the meaningful information pieces and to match them with the corresponding parameter groups. This helps in parameter search since the parameters are not trying to match a piece of information which they are not supposed to simulated. But this does not "add measurements" and the measurement dimension is thus not expanded.

Reply: We have done the related modification in the revised paper by replacing the 'measurement dimension' with the 'signature dimension'. Here this sentence has been corrected as "However, glacier mass data and baseflow data are usually not available in some mountain basins. In these cases, hydrograph partitioning is another possible way to exploit information from available data." in the revised paper.

3. Case study

3.1. I re-iterate my comment: why is the case study qualified as "alpine"? For botany, "alpine" might be a general term referring to any high elevation mountain range, for hydrology, "alpine" refers to my understanding to a hydro-climatic regime with a winter season with snow accumulation and a summer season with melt occurring due to high temperatures; is this the case here? Or do we have a regime where accumulation and melt occur both during the summer as in the Himalaya? On web of science, I could find a single paper mentioning the words "alpine hydrology and Tianshan". Could you not just say why the area has alpine hydrology? Or simply replace alpine area by mountainous area? Namely also on p. 13398 and 13399 where the more general "mountainous area" should be used instead of alpine.

Reply: To the authors' understanding, the term 'alpine' is an alternative word (and short) *for high mountain area. It has no hydrological meaning in this manuscript as referred by the Referee. Thanks for your suggestion. To avoid misunderstanding, we have replaced all the word "alpine" with the word "mountain" or "mountainous" in the revised paper.*

4. Method

4.1. I still do not understand how you connect the accumulation and melt of snow with the modis image. The paper says that snow accumulation and potential melt are simulated per subcatchment, I conclude that SWE is also computed per subcatchment. How do you connect this to the area that experiences melt as obtained from the MODIS image? Do you multiply the potential melt (mm/day) with the area that experiences melt? But then, how do you update the SWE? What do you do if your computed SWE is non-zero but the MODIS image does not show any snow pixels? And what if SWE is zero but MODIS shows snow?

Reply: In response to this comment, we have added the below discussion in the revised paper:

"To be noted, snowfall in each subcatchment is calculated according to the daily precipitation and temperature. And snowmelt is simulated using the degree-day method. However, the snow water equivalent in the snow cover zone (non-glacier area) is not computed. The existing of snow cover in each subcatchment is only determined by MODIS snow image. When the MODIS image indicates the existing of snow cover and meanwhile the daily temperature is higher than 0 ℃, then snowmelt will occur, otherwise, snowmelt will not occur. The identification of snow cover by MODIS image is in accordance with the fact that the partitioning of snowmelt dominant hydrograph is based on MODIS snow products. If the existing of snow cover is determined by snow water equivalent, the temperature parameters to calculate snowfall can have significant effects on the estimation of the degree-day factor for snowmelt. To partly reduce this effect, we calibrate the degree-day factor for snowmelt on the basis of MODIS snow cover products. Although in this way, the water balance of snow cover is not taken into account in the snow cover zone, it should not impact the calibration of the degree-day factor for snowmelt. It's worth noting that snow water balance in the glacier zone is updated by calculation of snow water equivalent where snow cover level should be relatively low."

4.2. The use multi-letter parameter names is banned by HESS.

Reply: Thanks, we have modified all the multi-letter parameter names into subscripts. 'KKA' is corrected to 'K_A', 'KKD' is corrected to 'K_D' and 'WM' is corrected to 'W_M'.

5. Results

5.1. I recommend explicitly commenting on the fact that clearly, the automatic calibration cannot find the solution to the optimization problem, otherwise it *HAS* to find a solution that is better than the step-wise solution. If the automatic solution found by optimizing NSE has a lower NSE or higher RMSE than the manual calibration, this means that the algorithm could not find the optimum.

Reply: In response to this comment, we have added the discussion below in the revised paper:

"The automatic calibration algorithm has run for about 5 weeks (840 hour on a desktop equipped with an Intel Core i7 CPU with 2.8GHz) to obtain the current results. Its performance can increase if the algorithm keeps on running, and even get higher performance than the step-wise calibration method. The comparison here is intending to show that the step-wise calibration method based on hydrograph partition can achieve considerable performance more effectively. The automatic algorithm here treats all the parameters equally during the calibration period. Each parameter should be optimized when searching for the optimal parameter set. This searching algorithm hampers the efficiency of the calibration procedure without identifying the dominant sub-periods for different parameters. In the step-wise calibration method, only parameters that are responsible for the simulation of corresponding hydrograph partition are optimized in each step. And also the calibration of parameter by this method reflects the role of each parameter for the basin runoff generation."

5.2. Again, I do not agree with the wording "Benefitting from the partitioning curves, however, the stepwise calibration method increases the dimension of measurement information to four. The measurement dimension is now equal to the number of parameter groups," The information content of data cannot be expanded by data manipulation. It can only be extracted. Otherwise you would create information.

Reply: In the revised paper, we have revised this sentence as "Benefitting from the *partitioning curves, however, the stepwise calibration method increases the dimension of hydrological signature to four. The signature dimension is now equal to the number of parameter group."*

5.3. What means "to extracting index information"?

Reply: it have been corrected as "to extract hydrological *signatures".*

Reply to comments from Dr. Zappa

Remarks:

1. **This manuscript is a re-submission of I manuscript I already evaluated in March 2014. The original manuscript was already rather interesting concerning topic and concepts, but rather unripe in its realization, analysis and presentation. In this new version the problematic issues have been addressed.**

   Reply: Thanks.

2. **In its current form the paper is very well embedded in scientific literature on the topic. Also the description of the test area is well documented and referenced. As in the original manuscript I appreciate the use of field data for estimating the lapse rates (Sections 2.2.1 and 2.2.2). This is a nice example of confining uncertainty by adding additional information from observations.**

   Reply: Thanks.

3. **Concerning the improvements we have now in Table 5 a good overview including calibration and evaluation periods.**

   Reply: Thanks.

4. **In the original submission I was complaining because I found your model was not able to capture peaks due to storm rainfall and rapid reaction by the basin. In this version I found this issue is almost solved. Did you some adjustments in the process description? Or is this an improvement stemming from the changes in the snowmelt and icemelt components (Page 13402)?**

   Reply: The model has been slight modified in Section 3.2. We have improved the process for runoff generated from rainfall directly in glacier area in the model. Given the relative large glacier coverage and the steep terrain in the study basin, rainfall provides storm runoff and flows into the stream network directly, which flows into the bare soil zone and reaches the stream network slowly in the previous model. The simulation of peak flows have been improved significantly benefiting from these modifications.

Points to be addressed:

1. **I already mentioned in the original submission, that you should be careful in defining your partition a "dominant runoff mechanism". In this manuscript you confuse and**

mix this again. I remember we suggested to use "dominant source of water".

Reply: Thanks. We have modified all the 'runoff generation mechanism' as 'runoff water sources' in the revised paper.

2. **On page 13400 you present your rules to separate the hydrograph. In Figure 6 we see the temporal distribution of the 4 options presented in Eq. 6. I understand you want to keep the rules easy, but if I correctly interpret Figure 6 you have surely small rain events in April. The red and green categories are very marginal in your test area, as they should focus on temperature driven snow and icemelt short before and short after the rain season. How do these rain events with obvious generation of Qr affect your calibrated data sets?**

   Reply: Given the seasonality of precipitation in our test area (shown in Figure 3), we neglected the rain events in the period from October to April for the test of the proposed calibration method. We acknowledge that this is a rough assumption, and surely small rain events will occur during this period. To take the effects of these rain events on the calibration into account, an iteration calibration procedure is adopted in this study. The parameters for melt and rainfall runoff are firstly calibrated on their dominant hydrograph parts (red and green, blue in Figure 6) separately, then the melt parameters are re-calibrated on the basis of the calculation of rainfall runoff using the parameters already calibrated in the first step. This calibration procedure is repeated until the parameter values getting a stable level. In this way, the effects of rainfall events in April on the calibration can be partly taken into account. And also, we have done some work to evaluate the sensitivity of the calibration to the partition of the rainfall event dominant hydrograph in Section 4.5. Results in Table 6 and Figure 10 show the rainfall events can have an important role on the calibration on the rainfall runoff parameter (i.e. WM), while have relatively slighter effects on the calibration of melt and groundwater parameters. The accurate partition of the rainfall runoff dominant hydrograph should be improved based on the more accurate measurement of rainfall in the test area, which can be working for further study.

3. **13403: As table 6 demonstrate their sensitivity to your approach, can you give some more information on the meaning of KKA and KKD. You call both of them**

"coefficient used to calculate calibrated subsurface flow", which is for me no useful information. Are the two factors linkable to some physical property (infiltration, storage coefficient or so?)

Reply: We have added the below sentence in Section 3.2 in the revised manuscript: *"K_A and K_D are outflow coefficients of groundwater storage. Their sum determines the flow rate of groundwater baseflow and their ratio (K_D / K_A) dominate the proportion of free groundwater storage. Infiltration and storage should have effects on the calibration of the two parameters. "*

Minor issues:

1. **13390-15: Typo: "slope"**

   Reply: We have revised it.

2. **13400: The notation chosen in Equation 6 is rather odd (minus signs in the indices to describe the mathematical equivalence). It is surely how you implemented it in your algorithm, but it is not very elegant in a manuscript. Wouldn't be better to have maybe a table instead?**

   Reply: We have improved it in the form as follow in the revised manuscript:

$$Q=\begin{cases} Q_{SB} & \text{for } S_i=0,\ G_i=0, \text{ and } D_i=0 \\ Q_{SB}+Q_{SM} & \text{for } S_i=1,\ G_i=0, \text{ and } D_i=0 \\ Q_{SB}+Q_{SM}+Q_{GM} & \text{for } S_i=1,\ G_i=1, \text{ and } D_i=0 \\ Q_{SB}+Q_{SM}+Q_{GM}+Q_R & \text{for } D_i=1 \end{cases}$$

3. **Table 3: on which basis you decide to have identical hydraulic conductivity in the u-zone and s-zone?**

   Reply: The soil layer in the test area is very thin. Soil storage capacity is relative low. Subsurface flow is mainly generated from groundwater. To make the simulation of subsurface flow simple, we assumed the hydraulic conductivity of the u-zone is same to the s-zone.

Final considerations:

I thank the authors for having made the effort to invest some more time to improve this manuscript. I listen now only few point they should now address. If this is achieved then I can recommend the paper for acceptance.

Reply: Thanks. The related points have been addressed in the revised manuscript.

List of relevant changes.

Dear Editor,

This is a revised version of the hessd-11-13385-2014 paper. In making the new version of the paper, we have carefully addressed all the comments and suggestions provided by two Referees (i.e. Dr. Schaefli and Dr. Zappa). In response to the concern on the using of MODIS snow cover image in the model by Dr. Schaefli, we have added some new sentences in Section 3.2 to describe the connecting between accumulation and melt of snow and MODIS image in more detail. We have also corrected all of the "runoff generation mechanism" as "runoff water sources" in the revised manuscript, as pointed out by both the two Referees. In response to other minor comments by the two Referees, we have also added some sentences and corrected some words in this new manuscript. In particular:

1) We have added some details about the study results in the abstract section.
2) The related "measurement dimension" have been modified as "signature dimension" in the revised manuscript.
3) The word "alpine" has been replaced with "mountain" or "mountainous".
4) We have modified the multi-letter parameter names into subscripts. i.e., '*KKA*' is corrected to 'K_A', '*KKD*' is corrected to 'K_D' and '*WM*' is corrected to 'W_M'.
5) In response to the comments on the automatic calibration algorithm by Dr. Schaefli, we have added a discussion paragraph in Section 4.3.
6) We have added some sentences to describe the meaning of parameter K_A and K_D in Section 3.2 in response to the comments by Dr. Zappa.
7) The format of Equation 6 have been improved, as pointed out by Dr. Zappa.

Thank you very much for your attention and consideration. The revised new manuscript is presented as follows, and all the changes have been marked as red.

Sincerely yours,

Fuqiang TIAN

Department of Hydraulic Engineering, Tsinghua University, Beijing 100084

Diagnostic calibration of a hydrological model in a mountain area by hydrograph partitioning

Z. H. He[1], F. Q. Tian[1]*, H. V. Gupta[2], H. C. Hu[1], H. P.Hu[1]

1. State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China
2. Department of Hydrology and Water Resources, The University of Arizona, Tucson, Arizona, 85721, USA

*Corresponding author information:

Email: tianfq@tsinghua.edu.cn

Tele: +86 010 6277 3396

Fax: +86 010 6279 6971

Manuscript submitted to Hydrology and Earth System Sciences

2015.03.02

Abstract

Hydrological modeling can exploit informative signatures extracted from long time sequences of observed streamflow for parameter calibration and model diagnosis. In this study we explore the diagnostic potential of hydrograph partitioning for model calibration in mountain areas, where meltwater from snow and glaciers are important sources for river runoff (in addition to rainwater). We propose an index-based method to partition the hydrograph according to dominant runoff water sources, and a diagnostic approach to calibrate a mountain hydrological model. First, by accounting for the seasonal variability of precipitation and the altitudinal variability of temperature and snow/glacier coverage, we develop a set of indices to indicate the daily status of runoff generation from each type of water source (i.e., glacier meltwater, snow meltwater, rainwater, and groundwater). Second, these indices are used to partition a hydrograph into four parts associated with four different combinations of dominant water sources (i.e., groundwater, groundwater + snow meltwater, groundwater + snow meltwater+ glacier meltwater, groundwater + snow meltwater + glacier meltwater + rainwater). Third, the hydrological model parameters are grouped by the associated runoff sources, and each group is calibrated to match the corresponding hydrograph partition in a stepwise and iterative manner. Similar to use of the regime curve to diagnose seasonality of streamflow, the hydrograph partitioning curve based on a dominant runoff water source (more briefly called the partitioning curve, not necessarily continuous) can serve as a diagnostic signature that helps relate model performance to model components. The proposed methods are demonstrated via application of a semi-distributed hydrological model (THREW) to the Tailan River basin (1324 km^2) in the Tianshan Mountain of China. Results show that the proposed calibration approach performed reasonably well. Cross validation and comparison to an automatic calibration method indicated its robustness.

1 Introduction

1.1 Background

Parameter calibration has been singled out as one of the major issues in the application of hydrological models (Johnston and Pilgrim, 1976; Gupta and Sorooshian, 1983; Beven and Binley, 1992; Boyle *et al.*, 2000). Commonly, one or more objective functions are selected as criteria to evaluate the similarity between observed and simulated hydrographs (Nash and Sutcliffe, 1970; Brazil, 1989; Gupta *et al.*, 1998; van Griensven and Bauwens, 2003). As model complexity increases, parameter dimensionality also increases significantly, which makes it much more difficult to calibrate model parameters manually. For this reason, automatic calibration procedures have been developed to identify the optimal parameter set (Gupta and Sorooshian, 1985; Gan and Biftu, 1996; Vrugt *et al*, 2003a,b). However, due to limitations in process understanding and measurement technologies, one can find different parameter sets within a chosen space that may acceptably reproduce the observed aspects of the catchment system (Sorooshian and Gupta, 1983; Beven and Freer, 2001). This phenomenon, which has been called “equifinality”, causes uncertainty in simulation and prediction (Duan *et al.*, 1992; Beven, 1993, 1996), and highlights the need for methods that are powerful enough to ‘diagnostically’ evaluate and correct models, i.e., that are capable of indicating to what degree a realistic representation of the real world has been achieved and pointing towards how the model should be improved (Spear and Hornberger, 1980; Gupta *et al.*, 1998, 2008).

Traditional regression-based model evaluation strategies (e.g., based on the use of Mean Squared Error or Nash Sutcliffe Efficiency as performance criteria) are demonstrably poor in their ability to identify the roles of various model components or parameters in the model output (Van Straten and Keesman, 1991; Zhang *et al.*, 2008; Gupta *et al.*, 2008; Yilmaz *et al.*, 2008; Hingray *et al.*, 2010), which is due in part to the loss of meaningful information when projecting from the high dimension of the data set (like hydrograph) down to the low (often one) dimension of the measure (Yilmaz *et al.*, 2008; Gupta *et al.*, 2009). A diagnostic evaluation method should match the number of unknowns (parameters) with the number of pieces of information by making use of multiple measures of model performance (Gupta *et al.*, 1998, 2008, 2009; Yilmaz *et al.*, 2008). One way to exploit hydrological information is to analyze the spatiotemporal characteristics of hydrological variables that can be related to specific hydrological processes in the form of "signature indices"(Richter *et al.*, 1996; Sivapalan *et al.*, 2003; Gupta *et al* 2008, Yilmaz *et al.*, 2008). Ideally, a "signature" should represent some "invariant" property of the system, be readily identifiable from available data, directly reflect some system function, and be maximally related to some "structure" or "parameter" in the model.

Attention to hydrological signatures, therefore, constitutes the natural basis for model diagnosis (Gupta *et al.*, 2008). Placed in this context, the body of literature on the topic is indeed large. Jothityangkoon *et al.* (2001) proposed a downward approach to evaluate the model's performance against appropriate signatures at progressively refined time scale. Signatures that govern the evaluation of model complexity are the inter-annual variability, mean monthly variation in runoff (called regime curve), and the flow duration curve (FDC). Farmer *et al.* (2003) evaluated the climate, soil and vegetation controls on the variability of water balance through four signatures: gradient of the annual yield frequency graph, average yield over many years for each month, FDC, and magnitude and shape of the hydrograph. Shamir *et al.* (2005a) described a parameter estimation method based on hydrograph descriptors (total flow, range between the extreme values, monthly rising limb density of the hydrograph, monthly maximum flow and negative/positive change) that characterize dominant streamflow patterns at three timescales (monthly, yearly, and record extent). Detenbeck *et al.* (2005) calculated several hydrologic indices including daily flow indices (mean, median, coefficient of variation, and skewness), overall flood indices (flood frequency, magnitude, duration, and flood timing of various levels), low flow variables (mean annual daily minimum), and ranges of flow percentiles to study the relationship of the streamflow regime to watershed characteristics. Shamir *et al.* (2005b) presented two streamflow indices to describe the shape of the hydrograph (rising/declining limb density, i.e., RLD and DLD) for parameter estimation in 19 basins of United States. Yadav *et al.* (2007) used similarity indices and hydrological signatures (runoff ratio and slope of the FDC) to classify catchments. Westerberg *et al.* (2011) selected several evaluation points on the FDC to calibrate models, and compared two selection methods to evaluate their effects on parameter calibration.

Generally, the reported signatures have the following two characteristics: (1) they concentrate on the extraction of hydrologically meaningful information contained in hydrographs, and (2) they focus on either an entire study period or a special continuous section of the entire period. They have occasionally considered temporal variability of runoff components and dominance of different runoff sources during different periods (e.g., the seasonal switching of runoff sources discussed in Tian *et al.*, 2012). However, a hydrograph could be dominated by various components or water sources at different response times (Haberlandt *et al.*, 2001; Eder *et al.*, 2005). Within this in mind, a few studies have explored the use of hydrological information in time dimension for stepwise calibration. For example, Schaefli *et al.* (2005) presented a stepwise calibration method for 7 parameters in a high mountainous area: snow and ice melt degree-day factors were conditioned by mass balance, slow reservoir parameters were determined by base flow, reservoir coefficients were calibrated by summer runoff, and the direct runoff coefficient was used to control discharge during precipitation events. Another notable example is Hingray *et al*. (2010), in which the authors estimated the value of snowmelt degree-day factor in a mountain basin by progressively minimizing the differences between observed and simulated values of different magnitude hydrographs. There are also many other follow up studies.

In mountain areas, streamflow is composed of both snow/glacier meltwater and rainwater. The energy-based and temperature-index models are two principal approaches to simulate snow and glacier melt (Rango and Martinec, 1979; Howard, 1996; Kane *et al.*, 1997; Singh *et al.*, 2000; Fierz *et al.*, 2003). To describe significant heterogeneity of temperature, precipitation, snow, and glacier, distributed hydrological models are generally used for precipitation-runoff modeling in mountain regions (Daly *et al.,* 2000; Klok *et al.*, 2001 etc.). Also, the utilization of remotely sensing products of precipitation and snow cover data in the mountain runoff modeling has become more popular in recent years (Swamy and Brivio, 1997; Akyurek *et al.*, 2011; Liu *et al.*, 2012 etc.). Most of these studies report sound simulation results. However, the need to develop an appropriate calibration strategy for precipitation-runoff modeling in mountain areas remains a key issue for two reasons: first, the hydrological processes are usually more complex (with snow/glacier melt and possibly soil freezing/thawing) than those in warmer areas, which implies a larger dimension of parameter (R^P) in the corresponding hydrological model; second, measured data set useful for model identification is usually limited due to a sparse gauge network, which produces a small measurement dimension (R^M) far lower than R^P.

To address this problem, related studies are putting effort into two directions. One is to reduce the calibrated R^P by estimating some of the parameters based on basin characteristics *a priori*. For example, Gurtz *et al.* (1999) proposed a parameterization method based on elevation, slope and shading derived from basin terrain. Gomez-Landesa and Rango (2002) obtained model parameters of ungauged basins from gauged basins by basin size, proximity of location, and shape similarities. Eder *et al.* (2005) estimated most of the parameters *a priori* from basin physiography before an automatic calibration was applied. The parameterization method may involve some uncertainties but be useful for the determination of insensitive parameters.

The second direction is to exploit hydrological information from implicit measure data. For instance, Dunn and Colohan (1999) used baseflow data as additional criteria for model evaluation. Mendoza *et al.* (2003) exploited recession-flow data to estimate hydraulic parameters. Stahl *et al.* (2008) used glacier mass balance information combined with stream hydrographs to constrain melt factors. Huss *et al.* (2008) used annual ice volume change data for optimizing melt and radiation factors, and glacier equilibrium line altitude for precipitation correction factors. Schaefli and Huss (2011) integrated the seasonal information of point glacier mass balance for model calibration by modifying the GSM-SOCONT model. Jost *et al.* (2012) introduced glacier volume loss calculated by high-resolution digital elevation models to calibrate hydrologic model. Knowledge acquired from the aforementioned research indicates that the use of additional information (e.g., baseflow, recession flow, and glacier mass balance) can effectively help reduce parameter uncertainty by significantly expanding R^M.

However, glacier mass data and baseflow data are usually not available in some mountain basins. In these cases, hydrograph partitioning is another possible way to exploit information from available data. Information about dominant hydrological processes contained in a hydrograph can be extracted by hydrograph partitioning or separation; this has long been a topic of interest in hydrology. Several different kinds of methods have been proposed (Pinder and Jones, 1969; McCuen, 1989; Nathan, 1990; Arnold *et al.*, 1995, 1999; Vivoni *et al.*, 2007), which can generally be classified into graphical methods, analytical methods, empirical methods, geochemical methods and automated program techniques (Nejadhashemi *et al.*, 2009). Most of them primarily focus on the partitioning of baseflow and are not capable of identifying more than two components. With the advent of isotope methods, multi-component hydrograph separation models have been developed. However, these models need be run for an extended period of time (usually a minimum of one hydrologic year) for the assumption that the isotopes of components are conserved to hold (Hooper and Shoemaker, 1986) and call for volumes of field data that are seldom available in poorly gauged and difficult to access mountain basins.

1.2 Objectives and Scope

This paper explores the benefits of partitioning the hydrograph into several parts, each related to one combination of dominant water sources for runoff generation. The parameter group controlling each type of runoff sources is then calibrated using the corresponding partitioning hydrographic curves via a stepwise approach, and model deficiencies are diagnosed by evaluating the model simulations associated with each partitioning curve (as a diagnostic signature). We demonstrate the potential of this approach in a mountain area where streamflow is the result of complex runoff generation processes arising from combinations of storm events and snow/glacier melt. The influence of each type of water source (groundwater, snow meltwater, glacier meltwater, or rainwater) varies in time and can be determined by an analysis of the dynamic spatiotemporal information in the available data series.

The paper is organized as follows. Section 2 contains a description of the geographic and hydrological characteristics of the study basin, including the main data sources and data preprocessing. Section 3 details the proposed method of hydrograph partitioning and parameter calibration based on a semi-distributed model coupled with the temperature-index method. Section 4 presents the results and discusses the possible sources of uncertainty. Section 5 provides a summary of this study and discusses further applications of the partitioning strategy.

2 Study Area and Data

2.1 Overview of the Study Area

The study mountain area (Tailan River basin, TRB) is on the south slope of the Tianshan Mountain (one of the highest mountain areas in China) in the Xinjiang Uygur Autonomous Region of China and extends from 41° 35′N to 42° 05′ N and 80° 04′E to 80° 35′E, covering a drainage area of 1324 km^2. Elevation ranges from 1600 m to 7100 m a.s.l. with an average value as high as 4100 m a.s.l. Precipitation occurs mainly in summer and rarely in winter, and winter precipitation always comes in the form of snowfall. Snow coverage accumulates in winter and ablates from spring into late summer when it melts away completely; the snow coverage dynamics can be obtained from MODIS data (see Figure 4). The basin is highly glacierized with approximately 33% of the basin area covered by glacier ice (see Figure 1). The glacier coverage stretches from approximately 3000 m to 7100 m a.s.l. and exists mainly at an altitude range of 4000 m to 5000 m a.s.l. Glacier melt and snowmelt form runoff as long as the temperature rises above a certain threshold and provide primary sources for downstream discharge.

TRB is a heavily studied mountain watershed in northwestern China. The relevant literature (Kang and Zhu, 1980; Shen *et al.*, 2003; Xie *et al.*, 2004; Gao *et al.*, 2011; Sun *et al.*, 2012) are reviewed below, and the main conclusions about the hydrometeorological characteristics are summarized as follows:

(1) The climate presents strong altitudinal variability. The mean annual precipitation in higher mountain areas is approximately 1200 mm (Kang *et al.*, 1980), while it is approximately only 180 mm in the outlet plain area (Xie *et al.*, 2004). The mean annual temperature ranges from below 0℃ in mountain areas to approximately 9℃ at the basin outlet (Sun *et al.*, 2012).

(2) Meltwater is the principal source of streamflow. Snow and glacier meltwater account for approximately 63% of the annual runoff (Kang *et al.*, 1980). The contribution of rainwater is relatively lower and occurs mainly in the storm rain period (May to September) (Xie *et al.*, 2004). Groundwater baseflow is smaller but dominates the streamflow in the winter (January, February and December), during which either rainfall or melt rarely occur (Kang *et al.*, 1980).

(3) The TRB river network is a simple fan system. Given large topographic drop and moderate drainage area, the runoff concentration time is no longer than one day (Xie *et al.*, 2004). Melting and falling water can quickly flow into the main channel and reach the basin outlet.

2.2 Data & Preprocessing

The Tailan gauging station (THS, 1602 m a.s.l.) is located the outlet of the watershed, where runoff, precipitation and temperature have been measured since 1957. To collect temperature and precipitation data at higher elevation, two automatic weather stations (AWS, product type TRM-ZS2) were set up in June 2011 (i.e., XT AWS, at 2116 m a.s.l. and TG AWS, at 2381 m a.s.l.). This relatively short record (from July 1, 2011-December 31, 2012) was used to estimate the lapse rate of precipitation and temperature (see below). The Bingtan automatic weather station (BT AWS, at 3950 m a.s.l.) located in an adjacent catchment (Kumalak basin) was used to validate the estimated temperature lapse rates. A digital elevation model (DEM) with a spatial resolution of 30 m was provided by the International Scientific & Technical Data Mirror Site, Computer Network Information Center of the Chinese Academy of Sciences (http://www.gscloud.cn). Remotely sensed snow cover area (SCA) data were downloaded from the MODIS website; the MOD10A2 and MYD10A2 products were used, both of which have a spatial resolution of 500m and a temporal resolution of eight-days. Daily snow cover data was obtained by linear interpolation of the eight-day data. The China Glacier Inventory (CGI) (Shi, 2008) was used to derive glacier coverage in the TRB. In our experience, most of the snow melts away after the warm summer period and the lowest snow/ice coverage in the year should, therefore, be roughly equal to the glacier coverage. Based on an analysis of filtered MODIS SCA (see Sect. 2.2.3), the lowest values of snow/ice coverage in the study period (2003-2012) are almost the same, which indicates that TRB glacier coverage is relatively stable during the study period. The DEM, river system, gauging stations and glacier distribution are shown in Fig.1.

2.2.1 Temperature Lapse Rate

Altitudinal distribution of temperature can be estimated through the lapse rate (Rango and Martinec, 1979; Tabony, 1985). According to Aizen *et al.* (2000), rates of temperature decrease with increasing elevation are quite different in various months, and ignoring this difference may lead to significant errors in the simulation of snow accumulation and melt. The lapse rate was therefore estimated for each month. Temperature variations with altitude can be estimated by the following equation, i.e.:

$$T = T_o + T_p \cdot (H - h) \tag{1}$$

where, T_o is the temperature value at low altitude (THS in this study), and T_p is the temperature lapse rate (usually negative), H and h are the elevation values at high and low positions, i.e., the mean elevation of two AWS and the elevation of THS, respectively. The values of T_p in different months are obtained by minimizing the error function, i.e.:

$$\min : z = \sum (T_i - (T_{oi} + T_p \cdot (H - h)))^2 \tag{2}$$

where, i indicates the i^{th} day in the analyzed month, T_i is the observed temperature in AWS, which is the mean value of the TG AWS and XT AWS in this study.

The temperature series data from July 1, 2011 to December 31, 2012 at THS, TG AWS and XT AWS were used to estimate the temperature lapse rate. The results (Table1) indicate significant month-to-month variation ranging from -0.30℃100 m^{-1} in December to -0.86℃100 m^{-1} in August. To validate the temperature lapse rates, the estimated and observed temperature data at BT AWS were compared (Fig. 2). We also compared the estimated temperature by an annual constant lapse rate (-0.62℃100 m^{-1}, a similar value to previous studies, e.g., Tabony (1985) and Tahir *et al.*(2011)). This constant value is optimized by the same method in Eqn. (2) but using all daily temperature measurements. Figure 2 indicates that the monthly lapse rate method performs better than the annual constant rate method at the BT station for all months throughout the year. Further, the temperature curves estimated by monthly lapse rates for April to August match the observed ones rather well. Note that the estimated temperatures tend to underestimate observed ones for the rest of the months, which, however, will not affect the melt runoff significantly due to the general freezing condition during this period.

2.2.2 Precipitation Lapse Rate

Based on the precipitation series measured at THS, the monthly precipitation to annual precipitation ratio (Fig.3) for the study period (2003-2012) indicates that precipitation occurs mainly in May to September. The lapse rate of precipitation was also estimated monthly, and a similar procedure as temperature was applied. The different is that the precipitation analysis was conducted at a weekly rather than daily time step, and the maximum measured precipitation of the two installed AWS was used instead of the mean value. The analyzed period is limited to the storm rain period (May to September). Other months are not included due to the relatively small precipitation amount. The weekly precipitation lapse rates are listed in Table2. Daily precipitation differences between higher and lower altitudes can be estimated as the weekly precipitation lapse multiplied by the ratio of daily precipitation to the corresponding weekly amount in THS. The precipitation lapse rate was not validated against BT AWS because of significant differences in precipitation distribution between the two basins (i.e., Tailan and Kumalak).

2.2.3 Filtering of MODIS Snow Cover Area Data

Snow cover extent was obtained from MODIS products. The MOD10A2 and MYD10A2 products were downloaded from the website http://reverb.echo.nasa.gov. In total, we obtained 460 eight-day images (two tiles, h23v04 and h24v04) from 2003 to 2012 for each product. Given that the accuracy of the MODIS SCA product is affected by cloud coverage to a significant degree, the remotely sensed images should be filtered to avoid the noise from clouds before using it for hydrological modeling (Ackerman *et al*., 1998). The following three successive steps are adopted to filter the products based on previous reports (Gafurov and Bardossy, 2009; Wang et al., 2009; Lopez-Burgos et al., 2012):

(1) Satellite combination: The snow cover products of two satellites, Terra (MOD10A2) and Aqua (MYD10A2) were combined. As long as the value of a pixel is marked as snow in either satellite, the pixel value is marked as snow.

(2) Spatial combination: Inspecting the values of the nearest four pixels around one center pixel marked as cloud, if at least three of the four surrounding pixels are marked as snow, the center pixel is modified as snow.

(3) Temporal combination: If one pixel is marked as cloud, its values in the previous and following observations are investigated. If both of the two observed values are snow, then the present value of the same pixel is snow.

As an example, the filtered results from year 2004-2005 shown in Fig.4 demonstrate a significant reduction in fluctuation of the SCA products. We find that the lowest values of snow/ice coverage in all years (2003-2012) are relatively stable (from 2003 to 2012 are: 35%, 34%, 39%, 36%, 37%, 34%, 41%, 35%, 38%, 39%, showing no obvious trend), which is close to the glacier coverage area (33%) derived from the CGI data mentioned in Sect.2.2. As mentioned before, MODIS snow/ice covered area in later summer is mainly composed of glacier coverage when snow has been melt away completely. The filtered results indicate a relatively stable coverage of glacier in TRB.

2.2.4 Altitudinal Cumulative Melt Curve

The daily temperature of each cell in MODIS SCA images can be estimated by a temperature lapse rate based on its elevation and daily temperature measured at THS. As long as the temperature exceeds a specific threshold value for melt (assumed to be 0℃ in this study), a given cell was labeled as an active cell in terms of melt. The land cover type for each cell was classified into glacier, snow, and other land cover according to the CGI and MODIS SCA product. To obtain the area covered by snow only, we subtracted the glacier area in CGI from the SCA (a similar procedure can be found in Luo *et al.*, 2013). When a glacier or snow cover cell is active, it is labeled as a melt cell, and the melt area is computed as the number of active cells multiplied by the area of a cell.

Organizing the melt area by elevation from low to high and summing the melt area at each elevation, we can get the altitudinal cumulative melt curve, which can be used to describe the spatiotemporal distribution of melt area. The altitudinal cumulative melt curves calculated from 2003 to 2012 for all months (Fig.5) show that melt mainly occur from May to September, which coincides with the precipitation period. Snowmelt starts at an elevation of approximately 1650 m a.s.l., while glacier melt starts at an elevation of approximately 2950 m a.s.l, which has an important implication for hydrograph partitioning.

3 Methodology

Theoretically, every drop of water in the streamflow comes ultimately from precipitation. Practically, we can consider water sources for runoff generation in mountain areas as mainly consisting of meltwater from snow and glacier, rainwater, and groundwater. Groundwater at the basin scale is recharged by direct infiltration and run-on infiltration of meltwater or rainwater, and it is mainly discharged as baseflow via a subsurface flow path (especially in mountain areas where the large elevation gradient favors baseflow discharge). For the purpose of hydrograph partitioning, we can consider recharge to be a separate water source for streamflow, independent of meltwater and rainwater, which principally forms the baseflow part of a hydrograph. The remaining part of a hydrograph is principally formed by meltwater and rainwater via surface flow path (Blöschl *et al.*, 2013). We develop three indices to indicate the water sources for runoff generation at the daily time scale. The hydrograph is further partitioned into several sub-parts based on the indices values. Each sub-part is dominated by one or more water sources for runoff generation. With the partitioning hydrographic curves, the parameters of hydrological models are correspondingly grouped by runoff sources and calibrated in a stepwise fashion. We use the THREW model coupled with a temperature-index module as an exploratory tool. To better demonstrate usefulness of the proposed methods, only the runoff generation related parameters, which are also significantly sensitive parameters (see Sect.4.6), are calibrated. Other insensitive parameters are fixed at their initial values, specified *a priori* from the literature or by expert knowledge.

3.1 An Index-based Method for Hydrograph Partitioning

In mountain areas, the relative contribution of different runoff water sources to the total streamflow varies throughout the year (Martinec *et al.*, 1982; Dunn and Colohan, 1999; Yang *et al.*, 2007). For the rainwater source, Fig.3 shows that precipitation in TRB presents strong seasonality and primarily concentrates (more than 76%) in the storm rain period from May to September. During the relatively dry period from October to April, mean precipitation gauged at the THS is just 43 mm, while precipitation in the higher mountainous region is mainly snowfall. Therefore, surface runoff induced by rainwater can rarely occur during relative dry period. It is reasonable to assume that the rainwater source can only contribute to the surface runoff part of a hydrograph on the same day during the storm rain period (May to September) except for the baseflow occurring much later.

For the meltwater sources, the altitudinal cumulative melt curves (Fig.5) show that the areas experiencing glacier melt and snowmelt change significantly with elevation. Melt of glacier and snow begins at different elevations in different months, i.e., glacier melt can only occur in the areas higher than 2950 m (the lower elevation limit of glacier coverage) while snowmelt can occur in areas higher than 1650 m. It can be deduced that snowmelt generally occurs at lower elevations than glacier melt. Remember that temperature decreases with increase in altitude. There should exist a period of time during which temperature at 1650 m is higher than snowmelt threshold while temperature above 2950 m is lower than glacier threshold and thus snowmelt does occur but glacier melt not.

The groundwater source should be a dominant source for the baseflow part of a hydrograph and, of course, it dominates the recession limb of a hydrograph (part of a baseflow partition) when no rainfall or melting occurs.

Based on the above physical understanding, we can partition the hydrograph using the following three indices:

(1) Date index (D_i): D_i is used to distinguish the dates on which rainfall and thus possible rainwater directly runoff process occurs. For simplicity, in this study we use D_i to distinguish dry period and storm rain period and assume no rainfall runoff in the dry period, i.e.,

$$D_i = \begin{cases} 1, & \text{for days in storm rain period from May to September} \\ 0, & \text{for days in relative dry period from October to April} \end{cases} \quad (3)$$

(2) Snowmelt index (S_i): S_i indicates whether snowmelt possibly occurs on a given day:

$$S_i = \begin{cases} 1, & \text{for days when temperature at altitude 1650 m is higher than 0 ℃} \\ 0, & \text{for other days} \end{cases} \quad (4)$$

(3) Glacier melt index (G_i): G_i is used to identify days when glacier melt possibly occurs:

$$G_i = \begin{cases} 1, & \text{for days when temperature at altitude 2950 m is higher than 0 ℃} \\ 0, & \text{for other days} \end{cases} \quad (5)$$

The hydrograph is then partitioned according to the three indices by using the following rules:

$$Q = \begin{cases} Q_{SB} & \text{for } S_i = 0,\ G_i = 0, \text{ and } D_i = 0 \\ Q_{SB} + Q_{SM} & \text{for } S_i = 1,\ G_i = 0, \text{ and } D_i = 0 \\ Q_{SB} + Q_{SM} + Q_{GM} & \text{for } S_i = 1,\ G_i = 1, \text{ and } D_i = 0 \\ Q_{SB} + Q_{SM} + Q_{GM} + Q_R & \text{for } D_i = 1 \end{cases} \quad (6)$$

where, Q is the overall streamflow series, Q_{SB} stands for the baseflow generated by groundwater source, Q_{SM} for snow meltwater runoff, Q_{GM} for glacier meltwater runoff, and Q_R for rainwater directly runoff. The partitioning principles are described as follows:

(1) Groundwater is the dominant component (Q=Q_{SB}) when both melt and rainwater directly runoff do not occur. This condition is mathematically equivalent to S_i+G_i+D_i=0, which requires S_i=0, G_i=0, and D_i=0;

(2) Snow meltwater and groundwater are the dominant components(Q=Q_{SB}+Q_{SM}) when the temperature is higher than 0 ℃ at 1650 m a.s.l. and lower than 0 ℃ at 2950 m a.s.l. (requires S_i=1, G_i=0, and D_i=0);

(3) Snow meltwater and glacier meltwater coupled with groundwater dominate (Q=Q_{SB}+Q_{SM}+Q_{GM}) on days when the temperature at 2950 m a.s.l. exceeds 0℃in October to April. This means G_i=1, D_i=0, and S_i=1, noting that S_i must be equal to 1 when G_i=1 for the decreasing nature of temperature along altitude;

(4) Finally, all sources are mixed (Q=Q_{SB}+Q_{SM}+Q_{GM}+Q_R) for other days in the storm rain

period (May to September, D_i=1). Each category contains days that could be continuous or discontinuous in time and could lie within different weeks due to temporal variability of precipitation and temperature.

3.2 Tsinghua Representative Elementary Watershed Hydrological Model

The Tsinghua Representative Elementary Watershed model (THREW model) used for the hydrological simulation in this study, has been successfully applied in many watersheds in both China and the United States (see Tian *et al.*, 2008, 2012; Li *et al.*, 2012; Liu *et al.*, 2012 etc.), including an application to a high mountainous catchment of Urumqi River basin by Mou *et al.* (2008). The THREW model adopts the REW (Representative Elementary Watershed) approach to conceptualize a watershed, where REW is the sub-catchment unit for hydrological modeling. The study basin was divided into several units (REW) based on a digital elevation model. Sub-catchment units were further divided into a surface and sub-surface layer, each layer containing several sub-zones. The sub-surface layer is composed of two zones: saturated zone and unsaturated zone, and the surface layer consists of six zones: vegetated zone, bare soil zone, snow covered zone, glacier covered zone, sub-stream-network zone, and main channel reach; see Tian *et al*. (2006) for further details.

The main runoff generation processes simulated by the THREW model include rainfall surface runoff, groundwater baseflow, snowmelt and glacier melt. Rainfall surface runoff is simulated by a Xin'anjiang module, which adopts a water storage capacity curve to describe non-uniform distribution of water storage capacity of a sub-catchment (Zhao, 1992). The storage capacity curve is determined by two parameters (spatial averaged storage capacity W_M and shape coefficient B). Rainfall surface runoff forms on areas where storage is replete. Replete areas are calculated by the antecedent storage and current rainfall. The saturation excess runoff is computed based on water balance. The remainder of rainfall can infiltrate into soil and become additional contributions to groundwater. Groundwater forms baseflow that is separately calculated by two coefficients (K_A and K_D). K_A and K_D are outflow coefficients of groundwater storage. Their sum determines the flow rate of groundwater baseflow and their ratio (K_D / K_A) dominate the proportion of free groundwater storage. Infiltration and storage should have effects on the calibration of the two parameters. The Xin'anjiang module has been successfully applied to the Qiedeke, Kaidu, Manasi and Kahai basins in Tianshan Mountain by different authors (Jiang, 1987; Yang *et al.*, 1987;Mu and Jiang, 2009), which indicates its applicability in our study area.

For the simulation of melt processes in this study, the THREW model was modified to couple with the temperature-index method, given the easy accessibility of air temperature data and generally good model performance of the temperature-index model (Hock, 2003; Singh *et al.*, 2000). Snow and glacier melt are simulated using separate degree-day factors (snowmelt degree day factor D_s and glacier melt degree day factor D_g). Glacier melt only occurs in glacier area according to CGI, which remains stable during the study period (2003-2012, see discussion in Sect. 2.2.3). Precipitation in the snow and glacier zone is divided into rainfall and snowfall according to two threshold temperature values (0℃ and 2.5℃ are adopted in this study according to Wu and Li (2007)), i.e., when temperature is higher than 2.5℃, all precipitation is rainfall, when temperature is lower than 0℃, all precipitation is snowfall, and when temperature falls between the two thresholds, precipitation is divided into rainfall and snowfall half by half (a simple division scheme adopted here). Rainfall on glacier areas forms runoff and flows into the stream-network directly without infiltration into soil. Snow water equivalent (SWE) on glacier areas is updated by combining snowfall and snowmelt, and for simplicity, snow is assumed to cover all glacier areas when the corresponding SWE is not zero. Snowmelt in glacier areas is simulated using snow degree-day factor D_s until it melts away completely. Snow cover area in non-glacier area is updated using MODIS data. To be noted, snowfall in each subcatchment is calculated according to the daily precipitation and temperature. And snowmelt is simulated using the degree-day method. However, the snow water equivalent in the snow cover zone (non-glacier area) is not computed. The existing of snow cover in each subcatchment is only determined by MODIS snow image. When the MODIS image indicates the existing of snow cover and meanwhile the daily temperature is higher than 0℃, then snowmelt will occur, otherwise, snowmelt will not occur. The identification of snow cover by MODIS image is in accordance with the fact that the partitioning of snowmelt dominant hydrograph is based on MODIS snow products. If the existing of snow cover is determined by snow water equivalent, the temperature parameters to calculate snowfall can have significant effects on the estimation of the degree-day factor for snowmelt. To partly reduce this effect, we calibrate the degree-day factor for snowmelt on the basis of MODIS snow cover products.

Although in this way, the water balance of snow cover is not taken into account in the snow cover zone, it should not impact the calibration of the degree-day factor for snowmelt. Since MODIS SCA products (i.e., MYD10A2) are available from 2003, the model simulation period is from 2003 to 2012, of which 2003-2007 for calibration and 2008-2012 for evaluation. The time step for simulation is daily.

3.3 Stepwise Calibration of Grouped Parameters Upon Partitioning Curves

Model parameters are grouped *a priori* according to their connection with causal physical mechanisms (see Table 3). According to Xie *et al.* (2004) and Kang *et al.* (1980), parameters that control groundwater baseflow, snowmelt, glacier melt, and rainwater surface runoff should be the most sensitive parameters for the runoff simulation (also see our sensitivity analysis in Sect. 4.6). These parameters are subjected to calibration in this study. They are related to the corresponding hydrograph parts and then calibrated in a stepwise manner: first, groundwater baseflow parameters (K_A and K_D) are estimated based on the Q_{SB} part of the hydrograph. Second, snowmelt degree day factor (D_s) is calibrated upon the $Q_{SB}+Q_{SM}$ part. Third, glacier melt degree-day factor (D_g) is determined according to the $Q_{SB}+Q_{SM}+Q_{GM}$ part. Finally, rainfall surface runoff parameters (B, W_M) are calibrated on days when D_i equals to1, i.e., the $Q_{SB}+Q_{SM}+Q_{GM}+Q_R$ part of hydrograph.

In each step, only the specific parameter group is subjected to calibration. The parameters determined in the previous steps are kept constant, and all other parameters that will be calibrated in the next steps adopt their initial values. As the simulation in each step can, to some degree, be affected by the initial conditions produced in the preceding step, an iterative procedure is implemented to progressively minimize this influence. The parameter groups are first calibrated based on the corresponding hydrograph parts, and then the stepwise sequence is repeated until the calibrated parameters converge, i.e., the difference in parameter values between two contiguous iterations is less than 10%. In each calibration step, we use *RMSEln* (Eqn. (7), emphasizing low flow) or *RMSE* (Eqn. (8), emphasizing high flow) as objective function for parameter optimization. The remaining, insensitive, parameters are determined *a priori* according to previous modeling experience (mainly from Sun *et al.* (2012)) and listed in Table 3. The initial values of the calibrated parameters are also determined *a priori* according to Sun *et al.* (2012) and Tian *et al.* (2012).

The overall streamflow can be simulated with all calibrated parameters, which is evaluated with *NSE* and *NSEln* (logarithm Nash Criterion) values. Given that it is relatively easier to obtain high evaluation merit values in snowmelt driven basins due to strong seasonality of streamflow, we further adopt a simple benchmark model (the inter-annual mean value for every calendar day) to evaluate performance of the proposed method by subtracting streamflow seasonality. This benchmark model is proposed by Schaefli and Gupta (2007) for basins having a relatively constant seasonality. The improvement of a model comparing to the benchmark model is quantified by the *BE*, see Eqn. (9) for detail.

$$RMSE\ln = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log Q_{obs}(i) - \log Q_{sim}(i))^2} \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_{obs}(i) - Q_{sim}(i))^2} \tag{8}$$

$$BE = 1 - \frac{\sum_{i=1}^{n}(Q_{obs}(i) - Q_{sim}(i))^2}{\sum_{i=1}^{n}(Q_{obs}(i) - Q_{ben}(i))^2} \tag{9}$$

4 Results and Discussion

4.1 Partitioning Hydrographic Curves

The hydrograph from 2003 to 2012 was partitioned based on Eqn. (6). In total, we obtained four kinds of partitioning curves, i.e. Q_{SB} part, $Q_{SB}+Q_{SM}$ part, $Q_{SB}+Q_{SM}+Q_{GM}$ part and $Q_{SB}+Q_{SM}+Q_{GM}+Q_R$ part. As an example, the partitioning curves in 2003 are shown in Fig. 6, in which the melting period ranges from late February to late November (labeled as red and green dots). Snowmelt (red dots) starts in February and ends in November, while glacier melt (green dots) starts later (March) and stops earlier (October). This melt situation agrees well with the previous studies of Kang *et al.* (1980) and Sun *et al.* (2012). Hydrograph parts dominated by groundwater source mainly fall into December, January and February and are denoted by black dots. The rainwater surface runoff occurs in the storm rain period only (May to September, denoted by blue dots). The total number of days of $Q_{SB}+Q_{SM}$ part from 2003 to 2007 is 365, and that of $Q_{SB}+Q_{SM}+Q_{GM}$ part is 249, while the $Q_{SB}+Q_{SM}+Q_{GM}+Q_R$ part occupies 765 days. The numbers of non-melt days (i.e. the Q_{SB} part, due to glacier melt generally occurs in the $Q_{SB}+Q_{SM}+Q_{GM}+Q_R$ part) in the five years are 114, 80, 89, 96, and 68, respectively.

Correspondingly, the mean temperatures in those years gauged at the THS are 8.9, 10.1, 9.9, 10.4, and 11.3℃, respectively. A lower mean annual temperature causes a longer non-melt period in that year and vice versa. Note that the partitioning curves can be discontinuous in time due to the spatial-temporal variability of temperature.

4.2 Model Calibration by the Stepwise Method

The six key parameters (K_A, K_D, D_s, D_g, W_M, and B) were firstly calibrated by the proposed stepwise and iterative method. To focus on baseflow generated by the groundwater source during the Q_{SB} period, the *RMSEln* metric that emphasizes low flow is chosen as the evaluation criterion for the calibration of parameters K_A and K_D,. Conversely, high flow is our focus for the remaining periods ($Q_{SB}+Q_{SM}$, $Q_{SB}+Q_{SM}+Q_{GM}$, $Q_{SB}+Q_{SM}+Q_{GM}+Q_R$) and the *RMSE* metric is chosen as the evaluation criterion for calibration of parameters D_s, D_g, and W_M and B. To deal with interaction between steps, an iterative calibration approach was adopted. A total of five iterations was implemented until the parameter estimates became stable; the simulation of each kind of partitioning curve in each step of the last iteration is presented in Fig. 7.The calibrated parameters are shown in Table 4 and the evaluation merits are listed in Table 5.

Figure 7a shows that the magnitude of baseflow in Q_{SB} part was captured well at most of the times. The *RMSEln* merit is 0.302 m^3/s, and the parameters K_A and K_D were determined as 1.1 and 0.002 respectively. Streamflow in the $Q_{SB}+Q_{SM}$ part is dominated by both snow meltwater and groundwater. The Fig.7b shows that melt peak flow events have also been captured well by a calibrated D_s as 2.5 mm ℃$^{-1}$ day^{-1} after the determination of K_A and K_D in the first step. For the $Q_{SB}+Q_{SM}+Q_{GM}$ part, glacier meltwater began to control the streamflow in combination with snow meltwater and groundwater. Snowmelt and baseflow were determined *a priori* by previously calibrated parameters. The remaining residual between the simulated and observed discharge can be attributed to glacier melt alone, which was thus used for the calibration of glacier melt factor D_g. The *RMSE* value for this hydrograph partition was optimized as 4.784 m^3/s and we obtained a sound simulation by a calibrated D_g as 7.2 mm ℃$^{-1}$ day^{-1} as shown in Fig.7c.During the storm rain periods ($Q_{SB}+Q_{SM}+Q_{GM}$ $+Q_R$ part), rainwater directly runoff is an additional important component of river runoff. Similarly, parameters W_M and B can be calibrated separately after *priori* determination of melt runoff and groundwater baseflow. The simulated *RMSE* value in this period is 12.650 m^3/s, with calibrated W_M=10.50cm and B=0.80. The overall daily streamflow simulation is obtained by combining the four partitions together (see Figure 8a). The corresponding *NSE* index is 0.881 and *NSEln* is 0.929. Generally the results suggest a sound simulation compared to the observation.

To be noted, the calibrated values of melt degree day factors D_s (2.5mm ℃$^{-1}$ day^{-1}) and D_g (7.2mm ℃$^{-1}$ day^{-1}) are similar to the values obtained in other studies in Tainshan area, e.g., D_s is calibrated as 2.5 mm ℃$^{-1}$ day^{-1} by Liu *et al.* (2012), and D_s and D_g are estimated as 3.1 mm ℃$^{-1}$ day^{-1} and 7.3 mm ℃$^{-1}$ day^{-1} respectively based on observed mass balance data by Liu *et al.* (1999), which indicates the robustness of our calibration method.

4.3 Comparison to Automatic Calibration Method

For comparison, we also carry out an automatic calibration with the help of the ε-NSGAII algorithm, an optimization method developed by Deb *et al.* (2002) and Kollat and Reed (2006). The six parameters were calibrated together and evaluated by *NSE* value of the overall hydrograph. The run time of the automatic algorithm is about 5 weeks (840 hour on a desktop equipped with an Intel Core i7 CPU with 2.8GHz). The *NSE* value for the final optimized parameters is 0.868, and the *NSEln* value is 0.846 (Fig. 8b), both of which are lower than the values obtained by the proposed stepwise method. The parameters calibrated by ε-NSGAII are listed in Table 4, and are different from those calibrated by the stepwise method. Specifically, the snowmelt degree-day factor (D_s) and groundwater baseflow parameters (K_A and K_D) obtained by ε-NSGAII are 2.03mm ℃$^{-1}$ day^{-1} and 5.6 and 99.1 respectively. The evaluation merits of *RMSE* and *RMSEln* for each partitioning curve are also shown in Table 5. In general, the simulation by the automatic algorithm is not as good as that by the stepwise method, especially for the low and middle flow partitions (Q_{SB}+Q_{SM} and Q_{SB}+Q_{SM}+Q_{GM}). This may be due to the tendency of NSE-based automatic calibration to emphasize high flows.

To make a further evaluation, a benchmark model suggested by Schaefli and Gupta (2007) is used for the comparison, which simply simulates daily runoff as the inter-annual daily mean value. Simulation results by the benchmark model are shown in the Figure 8c, which shows *NSE* value as 0.815 and *NSEln* value as 0.923. The high *NSE* and *NSEln* values can be attributed to the strong seasonality of stream discharge in the study basin (Schaefli and Gupta, 2007).The *BE* index (Eqn. (9), see Table 5) is used to measure the improvement of simulations by the calibration methods compared to the benchmark model. A positive value for *BE* means that the evaluated method outperforms the benchmark model. Figure 8 shows the simulations of daily streamflow by the three methods (Fig.8a by stepwise calibration method, Fig.8b by automatic calibration method and Fig.8c by benchmark model), which shows better simulation by the two calibration runs with THREW model than the benchmark model (*BE* values are both positive). The stepwise calibration run obtained a *BE* value of 0.355, while *BE* of the automatic calibration run is 0.271. The benchmark model describes the mean value of daily discharge on each calendar day. The higher the *BE* value is, the better the seasonal variability of the hydrograph is captured by the evaluation method. The higher *BE* value in the stepwise calibration method can be attributed to the better simulation of middle and low flows which are dominated by groundwater and melt water (Fig.8a). However, *BE* values simulated by two calibrated parameter sets are both relatively low, which is attributed to the poor mimic of the (rapidly rising and falling) peaks.

Note that the automatic calibration method based on *NSE* value of the overall hydrograph adopts 1D measurement information to optimize four parameter groups. Benefitting from the partitioning curves, however, the stepwise calibration method increases the dimension of hydrological signature to four. The signature dimension is now equal to the number of parameter groups, and the grouped parameters can be optimized according to their corresponding runoff sources separately. A sound simulation of the overall hydrograph is obtained by the reasonable reproduction of the separate partitioning curves. Therefore, parameters calibrated by the stepwise method are inclined to have more explicit physical basis.

In regards to computation efficiency, the stepwise calibration required 385 runs of the model to complete, with each model run taking about 1.5 minutes and the total computation time being about 10 hrs. In contrast, the state-of-the-art automatic calibration algorithm required about 5 weeks of CPU time consumption on a desktop equipped with an Intel Core i7 CPU and 2.8GHz. The comparison indicates that the stepwise calibration method is both more physically based as well as more computationally efficient.

It is worth noting, the performance of the automatic calibration algorithm can increase if the algorithm keeps on running, and even be higher than that of the step-wise calibration method. The comparison here is intending to show that the step-wise calibration method based on hydrograph partition can achieve considerable performance more effectively. The automatic algorithm here treats all the parameters equally during the calibration period. Each parameter should be optimized when searching for the optimal parameter set. This searching algorithm hampers the efficiency of the calibration procedure without identifying the dominant sub-periods for different parameters. In the step-wise calibration method, only parameters that are responsible for the simulation of corresponding hydrograph partition are optimized in each step. And also the calibration of parameter by this method reflects the role of each parameter for the basin runoff generation.

4.4 Evaluation for the Stepwise Calibration Method

The parameter set calibrated by the stepwise method is applied to the evaluation period (2008~2012), and the daily discharge simulation is shown in Fig.9a.The evaluation merits are listed in Table 5. The *NSE*, *NSEln* and *RMSE* values for the whole period indicate sound evaluation results but general lower performance compared to calibration period. However, the evaluation results by the stepwise method are still significant better than the benchmark model, which obtained a *NSE* value as low as 0.577 (Fig. 9b and Table 5). The *BE* value in evaluation period by the stepwise calibration method is 0.413. Furthermore, from the partition perspective, the *RMSEln* and *RMSE* values for four partitions in Table 5 show that the low flow simulations (Q_{SB}, Q_{SB}+Q_{SM}, and Q_{SB}+Q_{SM}+Q_{GM} parts) are pretty good and even outperform the calibration simulations. The high flow simulation (Q_{SB}+Q_{SM}+Q_{GM}+Q_R part) is, however, insufficient, with *RMSE* 16.727m^3/s (compared to 12.65 m^3/s in calibration period). The lower performance of overall evaluation should be attributed to the insufficiency in storm rain days, especially for some extreme storm events in the summer of 2010 (see Fig. 9a). The underestimation of these events is likely due to inadequate observations of rainfall, which are principally due to the strong spatial variability of rainfall in mountainous areas. It is widely acknowledged that the extreme runoff events are difficult to capture in mountain area, where gauged station is scarce, on the daily scale (Aizen *et al.*, 2000; Jasper *et al.*, 2002). However, the accuracy of our results is similar to Li and Williams (2008) (used SRM model) and Liu *et al.*(2012) (who used the MIKE-SHE model) who performed similar work in a basin that is close to TRB in Tianshan Mountains. Their Nash values for daily discharge varied from 0.51 to 0.78, and also failed to simulate the peak flows in summer. They also attributed the low efficiency to the heavy precipitation.

To further evaluate the robustness of the stepwise calibration method based on partitioning curves, cross validation was implemented. The hydrograph in the evaluation period was partitioned based on dominant runoff sources, as was done in the calibration years 2003-2007. We calibrated the model to 2008-2012 and evaluated it for 2003-2007. The new calibrated parameter values are K_A=0.9, K_D=0.003, D_s=2.2 mm ℃$^{-1}$ day^{-1}, D_g=7.4 mm ℃$^{-1}$ day^{-1}, W_M=10.2cm and B=0.77, which are similar to the values calibrated in 2003-2007 listed in Table 4. The *NSE*, *NSEln* and *RMSE* values for calibration period 2008-2012 and evaluation period 2003-2007 are 0.757, 0.900, 10.892m^3/s and 0.883, 0.910, 8.589m^3/s, respectively, using this new calibrated parameter set. The simulations of the two periods by cross validation are presented in Fig.9c-d, which shows similar performance by two calibrated parameter sets and further demonstrates the robustness of the proposed stepwise calibration method.

4.5 Sensitivity Analysis on Index-based Partitioning Method

The stepwise calibration method relies heavily on the hydrograph partition for different runoff sources. The indices defined in Sect. 3.1 are keys to identify the dominant days for melt water and rainwater. The definitions for elevation bands for the 0 ℃ Isotherm and for storm rain days in the year producing rainwater runoff should have significant influence on the parameter calibration. In this study, the elevation band of 0 ℃ Isotherm for snowmelt is fixed and defined as 1650m.This value should have minimal effect on the snowmelt simulation, as the occurrence of snowmelt is actually determined by the MODIS snow cover data. Glacier cover area is assumed as constant, which is very rough for we have only one CGI data. In this section, we define different elevation bands of 0 ℃ Isotherm for glacier to analyze the effect of glacier area variation on the model calibration. We also select different seasons as the storm rain period to analyze its sensitive effect.

According to the CGI data, the glacier area extends from the altitude of 2950m in 2002. Considering the possible variability, we define four different lowest elevation bands for the glacier area (LEG), i.e., -500m (2450m), -200m (2750m), +200m (3150m) and +500m (3450m). As an example, various hydrograph partition patterns in year 2003 are shown in Fig. 10. For the storm rain period (SRP), new seasons are defined as April to October, April to September, May to October, and June to August compared to the benchmark period May to September. A new hydrograph partition pattern in year 2003 is also shown in Fig. 10.The left column in Fig.

10 shows that the $Q_{SB}+Q_{SM}+Q_{GM}$ partition becomes longer while the $Q_{SB}+Q_{SM}$ partition becomes shorter when the LEG is lower. Therefore, glacier melt starts earlier and ends later in the years with lower LEG. In the right column, the $Q_{SB}+Q_{SM}+Q_{GM}$ partition becomes longer with the shorter SRP, while the variation of the $Q_{SB}+Q_{SM}$ partition can be negligible. Parameters were re-calibrated according to the new partition curves, and the results are shown in Table 6, indicating the increase of degree-day factor for glacier melt (D_g) with the increase of the LEG. The value of D_g is also found to become higher when the SRP falls in the warmer months. The variation of LEG imposes significant impacts on the calibration of D_g, with a result ranging from 5.8 to 8.0mm ℃$^{-1}$ day^{-1}, while the variation of SRP principally impacts the calibration of parameter W_M, with a result ranging from 8.2 to 10.5cm. However, the *NSE* values (see Table 6) for different settings show minimal differences. This can be attributed to the fact that parameters are optimized on separate partitioning curves in the stepwise calibration method. Each hydrograph partition can be well simulated by adjusting the parameter values. The partition patterns can influence the value of parameters significantly but only slightly influence the discharge simulation. Among various LEGs, the setting of 2950m leads to the highest *NSE* value. Glacier melt degree day factor (D_g) calibrated with this LEG is 7.2 mm ℃$^{-1}$ day^{-1}, which is very close to the value estimated as 7.3 mm ℃$^{-1}$ day^{-1} by Liu *et al.*(1999), in which the D_g is estimated according to the observed glacier mass balance data in Tianshan area. This can further demonstrate the reasonability of the assumption in Sect. 3.2 that the glacier area is stable and its lowest elevation is fixed at 2950m during the study period. For the various storm rain periods (SRP), when the May to October period is adopted, the discharge simulation is slightly better than the benchmark setting of SRP, i.e. May to September. This phenomenon seems to indicate the importance of precipitation measurement as discussed in Sect. 4.4. With the help of more advanced precipitation measurement, the storm rain period can be determined more precisely to improve the model simulation.

To evaluate the relative dominance of multiple runoff sources on the total runoff, we compute their contributions to total runoff by various LEG and SRP in Fig.11. The mean contributions of every runoff source are as follows: groundwater contributes 17%, snow meltwater contributes 16.5%, glacier meltwater contributes 40% and rainwater directly runoff contributes 26.5%. Total melt water (snowmelt and glacier melt) occupies approximately 56.5%

and is close to the ratio 63% suggested by Kang *et al.* (1980).

4.6 Sensitivity Analysis on Parameters

The number of parameters to be calibrated is determined by the parameter sensitivity and *a priori* analysis. To evaluate the effect of different parameters on the simulation of different hydrograph partitions, we implemented a simple parameter sensitivity procedure that is carried out by a “one-at-a-time” approach. Parameters from different groups in Table 3 are selected for sensitivity analysis, including saturated hydraulic conductivity for u-zone K_s^u, saturated hydraulic conductivity for s-zone K_s^s, subsurface flow coefficient K_A and K_D, manning roughness coefficient for hillslope n^t, spatial heterogeneous coefficient for infiltration capacity α^{IFL}, ground surface depression storage capacity $Fmax^b$, shape coefficient to calculate the saturation excess runoff area from the Xin’anjiang model B, spatial averaged tension water storage capacity in the Xin’anjiang model W_M, glacier degree day factor D_g and snowmelt degree per day factor D_s. Parameter are varied from -50% to +50% of the calibrated values using the stepwise method in Table 4. The relative change (R_{MS}) of simulated measure merits (*RMSEln or RMSE*) for different hydrograph partitions are used to evaluate the sensitivity (Eqn. (10)), where MS is the value of measure merits by the calibrated parameter, MS_+ is the merits value obtained by the parameter +50% of the calibrated one, and MS_- is the merits value obtained by the parameter -50% of the calibrated one. The sensitivity simulation results are shown in Table 7, which demonstrates the dominant control of parameter K_A, K_D, W_M, B, D_s and D_g. Some parameters have significant effects on simulation of multi hydrograph partitions. For example, parameters controlling the $Q_{SB}+Q_{SM}+Q_{GM}+Q_R$ period can also have significant effect on the other periods. To minimize this interaction, iterative calibration was implemented in the calibration procedure. The number of calibrated parameters is determined as six, which control the main runoff sources (i.e. groundwater baseflow, snowmelt, glacier melt and rainwater directly runoff). Note that the low dimension of parameter calibration should not account for the low efficiency of peak flow simulation, referring to the similar study in Tianshan mountain areas by Li and Williams (2008), and Liu *et al.*(2012), in which the models have a higher parameter dimension (higher than six), and the peak flow simulations are still inadequate.

$$R_{MS} = \left| \frac{MS_+ - MS_-}{MS} \right| \times 100\% \qquad (10)$$

5 Summary and Conclusion

This study proposes diagnostic calibration approach to extracting hydrological signatures from available data series in a mountain area, which can be further used to partition the hydrograph into dominant runoff sources. The parameters of a hydrological model were grouped according to runoff sources and then related to the corresponding hydrologic partitioning curve. Each parameter group was calibrated to improve the simulation of the corresponding partitioning curve in a stepwise way. In this way, the dimension of hydrological signature is expanded to equal the number of parameter groups. The parameter uncertainty due to interaction of parameters is reduced via an iterative calibration procedure. Application to a mountain watershed in the Tianshan Mountain in northwestern China showed that the approach performed reasonably well. Cross validation and comparison to an automatic calibration method indicated its applicability.

Note that a semi-distributed hydrological model was utilized to illustrate the proposed diagnostic calibration approach in the high mountainous Tailan River Basin. Glacier mass balance is not simulated in the model and the glacier coverage was kept fixed during the study period, which can be subject to significant change in the context of global warming. According to existing studies (Stahl *et al.*, 2008; Schaefli and Huss, 2011; Jost *et al.*, 2012), glacier mass balance data is useful to constrain the parameter uncertainty for hydrological modeling in a glaciered basin. While arguing that our assumption of unchanged glacier coverage will not weaken the importance of the proposed approach, we acknowledge that an improved model coupled with glacier mass balance equations will improve the accuracy of hydrological simulation aided by glacier mass balance observations. This is left for future research.

A prerequisite for the proposed approach is hydrograph partitioning based on dominant runoff sources. The key to the partition procedure is to identify the functional domain of each runoff source from signature information extracted from easily available data. A partition can be achieved in which the relative roles of different runoff sources in the basin runoff vary significantly with time. The mountain watershed is an area in which the runoff sources can be separated by the combination of topography, ground-gauged temperature and precipitation, and remotely sensed snow and glacier coverage. Other areas with strong temporal variability of catchment wetness along with precipitation (e.g., monsoon zones) could also be suitable for the proposed approach. The Dunne runoff is prone to dominate the hydrograph when the catchment is wet and it could switch to Hortonian runoff rapidly under the combination of high evaporative demand and less precipitation, as shown by Tian *et al*. (2012) in the Blue River basin of Oklahoma. This is, however, also left for future research.

Acknowledgments. We wish to thank Mr. Wang Xinhui for his assistance in collecting hydrometeorology data in the Tailan River basin, and thank Charlie Luce and Viviana Lopez-Burgos who provided great help in MODIS snow coverage product filtering. The authors would also like to thank sincerely two Referees (B. Schaefli and M. Zappa) and Editor Markus Weiler for his careful comments, which improve the quality of manuscript significantly. This study was supported by the National Science Foundation of China (NSFC 51190092, U1202232, 51222901) and the foundation of the State Key Laboratory of Hydroscience and Engineering of Tsinghua University (2012-KY-03, 2014-KY-01). Their support is greatly appreciated.

References

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C. and Gumley, L. E.: Discriminating clear sky from clouds with MODIS, J. Geophys. Res.,103, 32141-32157, 1998.

Aizen, V., Aizen, E., Glazirin, G. and Loaiciga, H. A.: Simulation of daily runoff in Central Asian alpine watersheds, J. Hydrol., 238, 15-34, 2000.

Akyurek, Z., Surer, S. and Beser, O.: Investigation of the snow-cover dynamics in the Upper Euphrates Basin of Turkey using remotely sensed snow-cover products and hydrometeorological data, Hydrol. Process.,25 (23), 3637-3648, 2011.

Arnold, J. G. and Allen, P. M.: Automated methods for estimating baseflow and ground water recharge from streamflow records, Journal of the American Water Resources Association, 35, 411-424, 1999.

Arnold, J. G., Allen, P. M., Muttiah, R. and Bernhardt, G.: Automated base-flow separation and recession analysis techniques, Ground Water, 33, 1010-1018, 1995.

Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, Adv. Water Resour., 16, 41-51, 1993.

Beven, K.: Equifinality and uncertainty in geomorphological modelling, The Scientific Nature of Geomorphology: Proceedings of the 27th Binghamton Symposium in Geomorphology, 289-313, 1996.

Beven, K. and Binley, A.: The future of distributed models-model calibration and uncertainty prediction, Hydrol. Process., 6, 279-298, 1992.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol.,249, 11-29, 2001.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A. and Savenije, H.(Eds.): Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales, Cambridge Univ. Press, New York,2013.

Boyle, D. P., Gupta, H. V. and Sorooshian, S.: Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, Water Resour. Res., 36, 3663-3674, 2000.

Brazil, L.: Multilevel calibration strategy for complex hydrologic simulation models, NOAA Technical Report, NWS 42, Fort Collins, 217 pp, 1989.

Bulygina, N., McIntyre, N. and Wheater, H.: Conditioning rainfall-runoff model parameters for ungauged catchments and land management impacts analysis, Hydrol. Earth Syst. Sci., 13 (6), 893-904, 2009.

Daly, S. F., Davis, R., Ochs, E. and Pangburn, T.: An approach to spatially distributed snow modelling of the Sacramento and San Joaquin basins, California, Hydrol. Process, 14 (18SI), 3257-3271, 2000.

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Transactions on evolutionary computation, 6, 182-197, 2002.

Detenbeck, N. E., Brady, V. J., Taylor, D. L., Snarski, V. M. and Batterman, S. L.: Relationship of stream flow regime in the western Lake Superior basin to watershed type characteristics, J. Hydrol., 309, 258-276, 2005.

Duan, Q., Sorooshian, S. and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resour. Res., 28, 1015-1031, 1992.

Dunn, S. M. and Colohan, R. J. E.: Developing the snow component of a distributed hydrological model: a step-wise approach based on multi-objective analysis, J. Hydrol., 223, 1-16, 1999.

Eder, G., Fuchs, M., Nachtnebel, H. and Loibl, W.: Semi-distributed modelling of the monthly water balance in an alpine catchment, Hydrol. Process., 19, 2339-2360, 2005.

Farmer, D., Sivapalan, M. and Jothityangkoon, C.: Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis, Water Resour. Res., 39, 1035, 2003.

Fierz, C., Riber, P., Adams, E., Curran, A., Fohn, P., Lehning, M. and Pluss, C.: Evaluation of snow-surface energy balance models in alpine terrain, J. Hydrol.,282 (1-4), 76-94, 2003.

Gafurov, A. and Bardossy, A.: Cloud removal methodology from MODIS snow cover product, Hydrol. Earth Syst. Sci., 13, 1361-1373, 2009.

Gan, T. Y. and Biftu, G. F.: Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure, Water Resour. Res., 32, 3513-3524, 1996.

Gao, W., Li, Z. and Zhang, M.: Study on Particle-size Properties of Suspended Load in Glacier Runoff from the Tomor Peak, Arid Zone Research, 28, 449-454, 2011(in Chinese).

Gomez-Landesa, E. and Rango, A.: Operational snowmelt runoff forecasting in the Spanish Pyrenees using the snowmelt runoff model, Hydrol. Process.,16, 1583-1591,2002.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80-91, 2009.

Gupta, H. V., Sorooshian, S. and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, Water Resour. Res., 34, 751-763, 1998.

Gupta, V. K. and Sorooshian, S.: Uniqueness and observability of conceptual rainfall-runoff model parameters: The percolation process examined, Water Resour. Res., 19, 269-276, 1983.

Gupta, V. K. and Sorooshian, S.: The Automatic Calibration of Conceptual Catchment Models Using Derivative-Based Optimization Algorithms, Water Resour. Res., 21, 437-485, 1985.

Gupta, H. V., Wagener, T. and Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, Hydrol. Process., 22, 3802-3813, 2008.

Gurtz, J., Baltensweiler, A. and Lang, H.: Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountainous basins, Hydrol. Process.,13,2751-2768, 1999.

Haberlandt, U., Klocking, B., Krysanova, V. and Becker, A.: Regionalisation of the base flow index from dynamically simulated flow components - a case study in the Elbe River Basin, J. Hydrol., 248, 35-53, 2001.

Hingray, B., Schaefli, B., Mezghani, A. and Hamdi, Y.: Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments, Hydrolog. Sci. J., 55 (6), 1002-1016, 2010.

Hock, R.: Temperature index melt modelling in mountain areas, J. Hydrol., 282, 104-115, 2003.

Hooper, R. P. and Shoemaker, C. A.: A Comparison of Chemical and Isotopic Hydrograph Separation, Water Resour. Res., 22, 1444-1454, 1986.

Howard, C.: Revisiting the degree-day method for snowmelt computations – Discussion, Water Resources Bulletin, 32 (2), 411-413, 1996.

Huss, M., Farinotti, D., Bauder, A. and Funk, M.: Modelling runoff from highly glacierized alpine drainage basins in a changing climate, Hydrol. Process., 22 (19SI), 3888-3902, 2008.

Jasper, K., Gurtz, J. and Herbert, L.: Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model, J. Hydrol.,267 (1-2),40-52, 2002.

Jiang, H. F.: Snow ablation modeling and its application to Qiedeke basin, Journal of Xinjiang Agricultural University, 1, 67-75, 1987 (in Chinese).

Johnston, P. R. and Pilgrim, D. H.: Parameter optimization for watershed models, Water Resour. Res., 12, 477-486, 1976.

Jost, G.,Moore, R. D., Menounos, B. and Wheate, R.: Quantifying the contribution of glacier runoff to streamflow in the upper Columbia River Basin, Canada, Hydrol. Earth Syst. Sci.,16,849-860,2012.

Jothityangkoon, C., Sivapalan, M. and Farmer, D. L.: Process controls of water balance variability in a large semi-arid catchment: downward approach to hydrological model development, J. Hydrol., 254, 174-198, 2001.

Juston, J., Seibert, J. and Johansson, P.: Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment, Hydrol. Process.,23 (21), 3093-3109, 2009.

Kane, D. L., Gieck, R. E., and Hinzman, L. D.: Snow Modeling at Small Alaskan Arctic Watershed, Journal of Hudrologic Engineering, 2 (4), 204-210, 1997.

Kang, E., Zhu, S. and Huang, M.: Some Results of the Research on Glacial Hydrology in the Region of MT. Tuomuer, Journal of Glaciology and Geocryology, 2, 18-21, 1980(in Chinese).

Klok, E. J., Jasper, K., Roelofsma, K. P., Gurtz, J. and Badoux, A.: Distributed hydrological modelling of a heavily glaciated Alpine river basin, Hydrolog. Sci. J.,46 (4), 553-570, 2001.

Kollat, J. B. and Reed, P. M.: Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design, Adv. Water Resour., 29, 792-807, 2006.

Li, H. Y., Sivapalan, M. and Tian, F. Q.: Comparative diagnostic analysis of runoff generation processes in Oklahoma DMIP2 basins: The Blue River and the Illinois River, J. Hydrol., 418, 90-109, 2012.

Li, X. G. and Williams M. W.: Snowmelt runoff modelling in an arid mountain watershed, Tarim Basin, China, Hydrol. Process.,22 (19SI), 3931-3940, 2008.

Liu, D. F., Tian, F. Q., Hu, H. C. and Hu, H. P.: The role of run-on for overland flow and the characteristics of runoff generation in the Loess Plateau, China, Hydrolog. Sci. J., 57, 1107-1117, 2012.

Liu, S. Y., Xie, Z. C., Wang, N. L. and Ye, B. S.: Mass balance sensitivity to climate change: a case study of glacier no. 1 at Urumqi riverhead, Tianshan mountains, China, Chin. Geogr. Sci., 9, 134-140, 1999.

Liu, T., Willems, P., Feng, X. W., Li, Q., Huang, Y., Bao, A. M., Chen, X., Veroustraete, F. and Dong, Q. H.: On the usefulness of remote sensing input data for spatially distributed hydrological modelling: case of the Tarim River basin in China, Hydrol. Process.,26 (3), 335-344, 2012.

Lopez-Burgos, V., Gupta, H. V. and Clark, M.: A probability of snow approach to removing cloud cover from MODIS Snow Cover Area products, Hydrol. Earth Syst. Sci. Discuss, 9, 13693-13728, 2012.

Luo, Y., Arnold, J., Liu, S., Wang, X. and Chen, X.: Inclusion of glacier processes for distributed hydrological modeling at basin scale with application to a watershed in Tianshan Mountains, northwest China, J. Hydrol., 477, 72-85, 2013.

Martinec, J., Oeschger, H., Schotterer, U. and Siegenthaler, U.: Snowmelt and groundwater storage in alpine basin, In Hydrological Aspects of Alpine and High Mountain Areas, Wallingford, United Kingdom: IAHS Press, 169–175, 1982.

McCuen, R. H.: Hydrologic analysis and design, Prentice Hall, New Jersey pp.355-360, 1989.

Mendoza, G. F., Steenhuis, T. S., Walter, M. T. and Parlange, J. Y.: Estimating basin-wide hydraulic parameters of a semi-arid mountainous watershed by recession-flow analysis, J. Hydrol.,279,57-69,2003.

Mou, L., Tian, F., Hu, H. and Sivapalan, M.: Extension of the Representative Elementary Watershed approach for cold regions: constitutive relationships and an application, Hydrol. Earth Syst. Sci., 12, 565-585, 2008.

Mu, Z. X. and Jiang, H. F.: Establishment of snowmelt type Xin'anjiang watershed model based on digital elevation model, Journal of Xinjiang Agricultural University, 5 (32), 75-80, 2009 (in Chinese).

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, J. Hydrol., 10, 282-290, 1970.

Nathan, R. J., McMahon, T. A.: Evaluation of automated techniques for base flow and recession analyses, Water Resour. Res., 26, 1465-1473, 1990.

Nejadhashemi, A. P., Shirmohammadi, A., Sheridan, J. M., Montas, H. J. and Mankin, K. R.: Case Study: Evaluation of Streamflow Partitioning Methods, J. Irrig. Drain. Eng., 135, 791-801, 2009.

Pellicciotti, F., Brock, B., Strasser, U., Burlando, P., Funk, M. and Corripio, J.: An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland, Journal of Glaciology, 51 (175), 573-587, 2005.

Pinder, G. F. and Jones, J. F.: Determination of the ground-water component of peak Determination of the ground-water component of peak discharge from the chemistry of total runoff, Water Resour. Res., 5, 438-445, 1969.

Rango, A. and Martinec, J.: Application of a Snowmelt-runoff Model Using Landsat Data, Nord. Hydrol., 10, 225-238, 1979.

Richter, B. D., Baumgartner, J. V., Powell, J. and Braun, D. P.: A method for assessing hydrologic alteration within ecosystems, Conservation Biology, 10, 1163-1174, 1996.

Schaefli, B. and Gupta, H. V.: Do Nash values have value, Hydro. Process., 21 (15), 2075-2080, 2007.

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

Schaefli, B. and Huss, M.: Integrating point glacier mass balance observations into hydrologic model identification, Hydrol. Earth Syst. Sci.,15,1227-1241,2011.

Shamir, E., Imam, B., Gupta, H. V. and Sorooshian, S.: Application of temporal streamflow descriptors in hydrologic model parameter estimation, Water Resour. Res., 41, W06021, doi:10.1029/2004WR003409, 2005a.

Shamir, E., Imam, B., Morin, E., Gupta, H. V. and Sorooshian, S.: The role of hydrograph indices in parameter estimation of rainfall-runoff models, Hydrol. Process.,19, 2187-2207, 2005b.

Shen, Y., Liu, S., Ding, Y. and Wang, S.: Glacier Mass Balance Change in Tailanhe River Watersheds on the South Slope of the Tianshan Mountains and its impact on water resources, Journal of Glaciology and Geocryology, 25, 124-129, 2003(in Chinese).

Shi, Y.: Concise Glacier Inventory of China, Shanghai Popular Science Press., Shanghai, China, 2008(in Chinese).

Singh, P., Kumar, N. and Arora, M.: Degree-day factors for snow and ice for Dokriani Glacier, Garhwal Himalayas, J. Hydrol, 235, 1-11, 2000.

Sivapalan, M., Bloschl, G., Zhang, L. and Vertessy, R.: Downward approach to hydrological prediction, Hydrol. Process.,17, 2101-2111, 2003.

Sorooshian, S. and Gupta, V. K.: Automatic calibration of conceptual rainfall-runoff models-the question of parameter observability and uniqueness, Water Resour. Res., 19, 260-268, 1983.

Spear, R. C. and Hornberger, G. M.: Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis, Water Research,14,43-49,1980.

Stahl, K., Moore, R. D., Shea, J. M., Hutchinson, D. and Cannon, A. J.: Coupled modelling of glacier and streamflow response to future climate scenarios, Water Resour. Res.,44,2008.

Sun, M., Yao, X., Li, Z. and Li, J.: Estimation of Tailan River Discharge in the Tianshan Mountains in the 21st Century, Advances on Climate Change Research, 8, 342-349, 2012(in Chinese).

Swamy, A. N. and Brivio, P. A.: Modelling runoff using optical satellite remote sensing data in a high mountainous alpine catchment of Italy, Hydrol. Process., 11 (11), 1475-1491, 1997.

Tabony, R. C.: The variation of surface temperature with altitude, Meteorolgical Magazine, 114, 37-48, 1985.

Tahir, A. A., Chevallier, P., Arnaud, Y., Neppel, L. and Ahmad, B.: Modeling snowmelt-runoff under climate scenarios in the Hunza River basin, Karakoram Range, Northern Pakistan, J. Hydrol, 409, 104-117, 2011.

Tian, F. Q., Hu, H. P. and Lei, Z. D.: Thermodynamic watershed hydrological model: Constitutive relationship, Science in China, Ser. E-Technological Sciences, 51, 1353-1369, 2008.

Tian, F., Hu, H., Lei, Z. and Sivapalan, M.: Extension of the Representative Elementary Watershed approach for cold regions, Hydrol. Earth Syst. Sci., 10, 619-644, 2006.

Tian, F. Q., Li, H. Y. and Sivapalan, M.: Model diagnostic analysis of seasonal switching of runoff generation mechanisms in the Blue River basin, Oklahoma, J. Hydrol, 418, 136-149, 2012.

Van Griensven, A. and Bauwens, W.: Multiobjective autocalibration for semidistributed water quality models, Water Resour. Res., 39, 1348, 2003.

Van Straten, G. T. and Keesman, K. J.: Uncertainty propagation and speculation in projective forecasts of environmental change: A lake-eutrophication example, J. Forecast., 10, 163-190, 1991.

Vivoni, E. R., Entekhabi, D., Bras, R. L. and Ivanov, V. Y.: Controls on runoff generation and scale-dependence in a distributed hydrologic model, Hydrol. Earth Syst. Sci., 11, 1683-1701, 2007.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W. and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resour. Res., 39, 1214, 2003a.

Vrugt, J. A., Gupta, H. V., Bouten, W. and Sorooshian, S.: A Shuffled Complex Evolution Metropolis Algorithm for Optimization and Uncertainty Assessment of Hydrological Model Parameters, Water Resour. Res., 39, 1201, doi:10.1029/2002WR001642, 8., 2003b.

Wang, X. W., Xie, H. J., Liang, T. G. and Huang, X. D.: Comparison and validation of MODIS standard and new combination of Terra and Aqua snow cover products in northern Xinjiang, China, Hydrol. Process.,23, 419-429, 2009.

Westerberg, I. K., Guerrero, J. L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E. and Xu, C. Y.: Calibration of hydrological models using flow-duration curves, Hydrol. Earth Syst. Sci.,15 (7), 2205-2227, 2011.

Wu, J., L. L.:A rain-on-snow mixed flood forecast model and its application, Engineering Journal of Wuhan University,40,20-23,2007(in Chinese).

Xie, C., Ding, Y., Liu, S. and Han, H.: Analysis on the Glacial Hydrological Features of the Glaciers on the South Slope of Mt. Tuomuer and the Effects on Runoff, Arid Land Geography, 27, 570-575, 2004(in Chinese).

Yadav, M., Wagener, T. and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Adv. Water Resour, 30, 1756-1774, 2007.

Yang, D. Q., Zhao, Y. Y., Armstrong, R., Robinson, D. and Brodzik, M. J.: Streamflow response to seasonal snow cover mass changes over large Siberian watersheds, J. Geophys. Res., 112, F02S22F2, 2007.

Yang, X. S., Jiang, H. F., Huang, C. R. Zheng, Z., and Yong, G.: An applied study on the snowmelt type of Xin'anjiang watershed model at the Kaidu river basin, Journal of Xinjiang Agricultural University,4, 82-90, 1987 (in Chinese).

Yilmaz, K. K., Gupta, H. V. and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, Water Resour. Res., 44,W09417, doi:10.1029/2007WR006716, 2008.

Zhang, Z. X., Wagener, T., Reed, P. and Bhushan, R.: Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization, Water Resour. Res., 44 (W00B04), doi:10.1029/2008WR006833, 2008.

Zhao, R. J.: The Xin'anjiang model applied in China, J. Hydrol,135,371-381,1992.

Table1. Estimated monthly temperature lapse rate in the TRB

Month	Temperature lapse rate (℃/day/100 m)
January	-0.38
February	-0.38
March	-0.66
April	-0.76
May	-0.80
June	-0.78
July	-0.82
August	-0.86
September	-0.66
October	-0.60
November	-0.54
December	-0.30
Annual	-0.62

Table 2. Estimated week-precipitation lapse rate in storm rain months

Month	Precipitation lapse rate (mm/week/ 100 m)
May	1.63
June	1.69
July	3.14
August	2.40
September	2.28

Table 3. Grouped parameters in the THREW model. Parameters subjected to calibration are highlighted in red.

Category	Symbol	Unit	Description	Value
Subsurface	K_s^u	m s^{-1}	Saturated hydraulic conductivity for u-zone	1.25E-05
	K_s^s	m s^{-1}	Saturated hydraulic conductivity for s-zone	1.25E-05
	K_A	-	Coefficient used to calculate subsurface flow	Calibrated
	K_D	-	Coefficient used to calculate subsurface flow	Calibrated
Routing	n^t	-	Manning roughness coefficient for hillslope, obtained from the literature according to land use and vegetation type	1.50E-01
	n^r	-	Similar to n^t, roughness coefficient for channel	3.00E-01
Infiltration	α^{EFL}	-	Spatial heterogeneous coefficient for exfiltration capacity	1.00E+00
	α^{IFL}	-	Spatial heterogeneous coefficient for infiltration capacity	1.50E+00
Interception	$F\max^b$	m	Ground surface depression storage capacity	0.00E+00
	α^{vb}	m	Maximum rainfall depth a single leaf can intercept and hold	1.00E-05
Rainfall runoff	B	-	Shape coefficient to calculate the saturation excess runoff area from the Xin'anjiang model	Calibrated
	W_M	cm	Spatial averaged tension water storage capacity in the Xin'anjiang model	Calibrated
Melt	D_g	mm℃$^{-1}$ day^{-1}	Glacier melt degree day factor	Calibrated
	D_s	mm℃$^{-1}$ day^{-1}	Snowmelt degree day factor	Calibrated

Table 4. Calibrated parameters by the stepwise and automatic methods

Parameter	Stepwise Calibrated	Automatic Calibrated
K_A	1.1	5.6
K_D	0.002	99.1
D_s(mm ℃$^{-1}$ day^{-1})	2.5	2.03
D_g(mm ℃$^{-1}$ day^{-1})	7.2	7.52
W_M(cm)	10.5	11.9
B	0.80	0.62

Table 5. Evaluation merits for the stepwise and automatic calibration methods

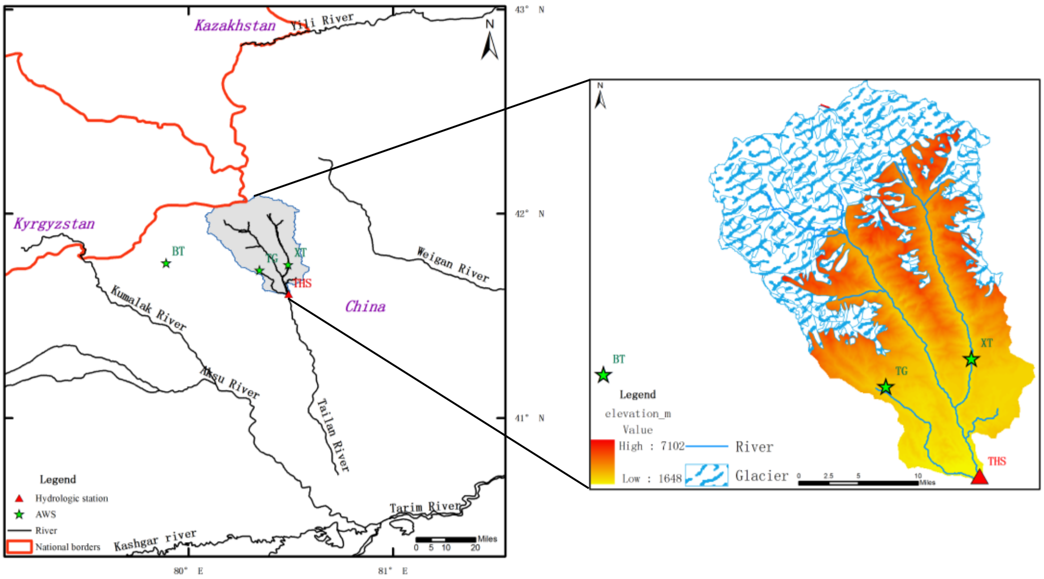
Merits	Calibration period Automatic method	Calibration period Stepwise method	Calibration period Benchmark model	Evaluation period Stepwise method	Evaluation period Benchmark model
RMSEln(Q_{SB}, m^3/s)	0.352	0.302	-	0.213	-
RMSE($Q_{SB}+Q_{SM}$, m^3/s)	2.807	1.811	-	1.762	-
RMSE($Q_{SB}+Q_{SM}+Q_{GM}$, m^3/s)	6.079	4.784	-	4.558	-
RMSE($Q_{SB}+Q_{SM}+Q_{GM}+Q_R$, m^3/s)	13.245	12.650	-	16.727	-
NSE	0.867	0.881	0.815	0.752	0.577
NSEln	0.841	0.929	0.923	0.894	0.844
RMSE (m^3/s)	8.990	8.459	10.534	11.021	14.381
BE	0.271	0.355	-	0.413	-

Table 6. Sensitive analysis of the calibrated parameters on lowest elevation band for glacier area (LEG) and storm rain period (SRP). *NSE* is the Nash Sutcliffe Efficiency value for the calibration period.

	LEG(a.s.l. m)	D_s(mm/d/℃)	D_g(mm/d/℃)	W_M(cm)	B	K_A	K_D	*NSE*
	3450	2.2	8.0	10.1	0.70	0.7	0.002	0.870
	3150	2.5	7.9	10.1	0.75	0.7	0.002	0.871
SRP:	2950	2.5	7.2	10.5	0.80	1.1	0.002	0.881
May. To Sep.	2750	3.0	6.8	10.2	0.75	1.0	0.002	0.880
	2450	2.8	5.8	10.0	0.78	0.8	0.002	0.876
	SRP	D_s(mm/d/℃)	D_g(mm/d/℃)	W_M(cm)	B	K_A	K_D	*NSE*
	Jun. to Aug.	2.9	7.5	8.2	0.75	0.9	0.002	0.871
	May. to Oct.	2.8	6.9	9.4	0.76	0.8	0.002	0.882
LEG=2950m	May. to Sep.	2.5	7.2	10.5	0.80	1.1	0.002	0.881
	Apr. to Sep.	2.2	7.1	8.3	0.75	0.9	0.002	0.878
	Apr. to Oct.	2.6	6.9	9.4	0.77	1.1	0.002	0.881

Table 7. R_{MS} (%) for parameter sensitivity (R_{MS} values indicating the most sensitive parameters are labeled in bold and red)

	Subsurface				Routing	Infiltration	Interception	Rainfall Runoff		Melt	
Merits	K_s^u	K_s^s	K_A	K_D	n^t	α^{IFL}	$F\max^b$	W_M	B	D_s	D_g
RMSEln *(Q_{SB})*	9.70	11.14	**38.44**	**44.39**	15.70	0.12	0.08	1.07	18.51	7.53	2.88
RMSE *($Q_{SB}+Q_{SM}$)*	0.32	0.40	11.91	0.06	9.35	0.47	0.14	8.27	25.14	**51.22**	0.69
RMSE *($Q_{SB}+Q_{SM}+Q_{GM}$)*	0.22	0.21	0.62	0.64	10.00	0.17	0.25	7.92	0.29	26.28	**40.79**
RMSE *($Q_{SB}+Q_{SM}+Q_{GM}+Q_R$)*	0.17	0.85	0.57	0.97	1.84	0.08	0.06	**19.35**	**22.48**	10.78	11.57



(a) (b)

Figure 1. Location of the Tailan River basin in Xinjiang Uygur Autonomous Region, China. Two automatic weather stations (TG at 2381 m a.s.l. and XT at 2116 m a.s.l.) were set up in upstream mountain area in July, 2011. Additionally, the BT weather station (3950 m a.s.l.) located in the adjacent Kumalak River basin was used to validate the estimated temperature lapse rates. The Tailan Hydrologic Station (THS) has gauged streamflow data at the catchment outlet since 1957(a).Glacier occupies approximately 33% of the total basin area (b).

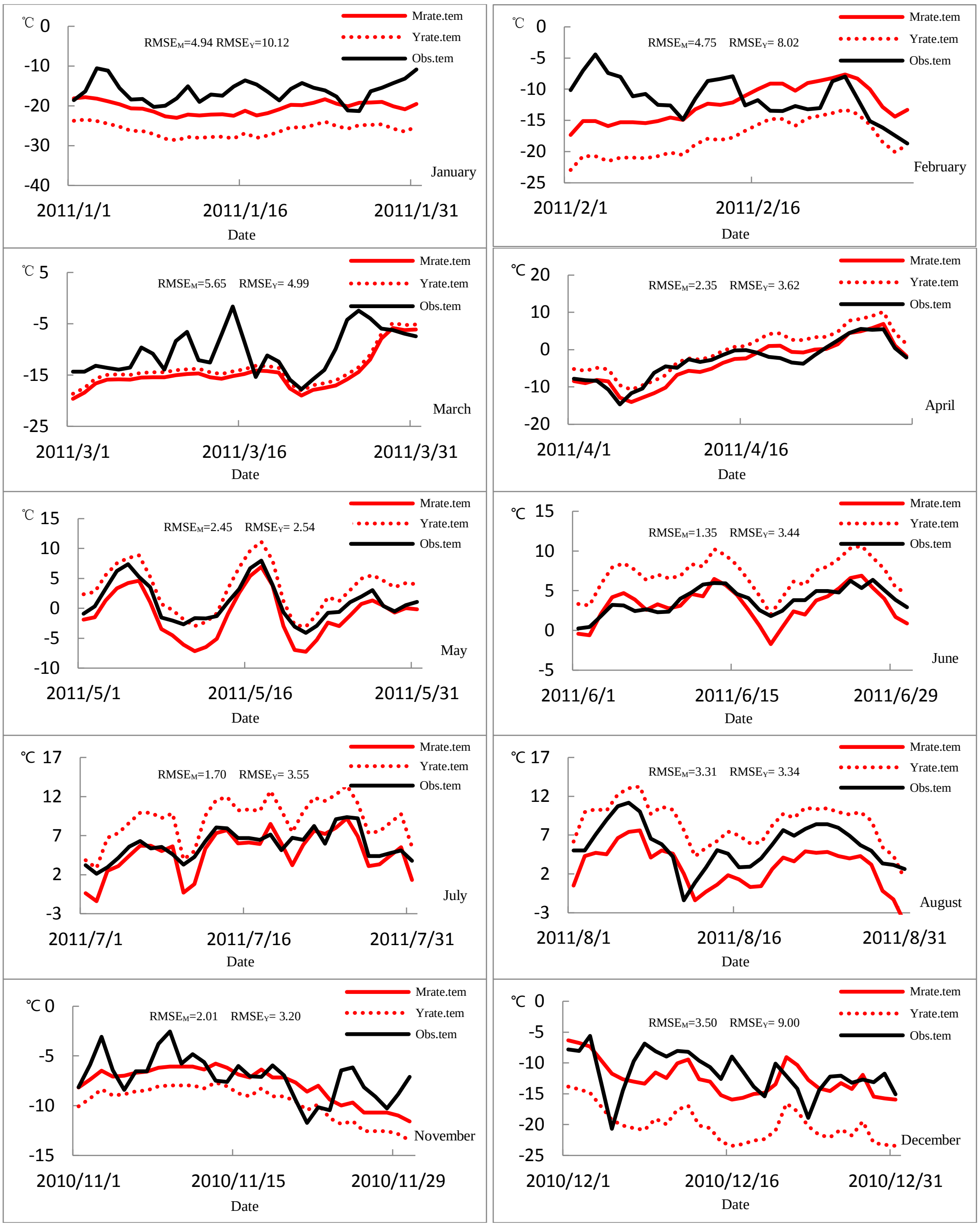


Figure 2. Evaluation of the estimated temperature lapse rate at the BT station. The black solid line is the observed temperature series at BT (Obs.tem); the red solid line is the estimated temperature by monthly lapse rate (Mrate.tem).The red dotted line indicates the estimated temperature based on annual constant rate (Yrate.tem). The goodness of fit between the observed and estimated temperature is measured by $RMSE_M$ for monthly lapse rate and $RMSE_Y$ for annual constant rate, respectively. The temperature series in September and October are absent at BT.

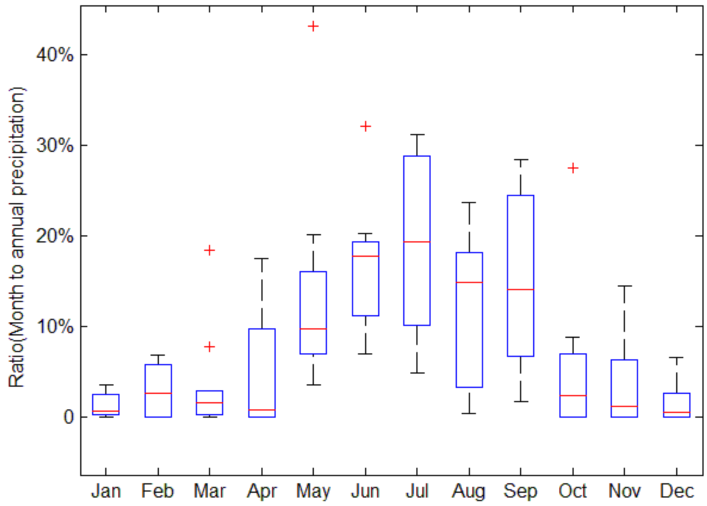


Figure 3. Proportion of monthly precipitation to annual amount (2003~2012). The red line in each box represents the median value for each month from 2003 to 2012. Red crosses indicate abnormal values that exceed 1.5 times the inter quartile range.

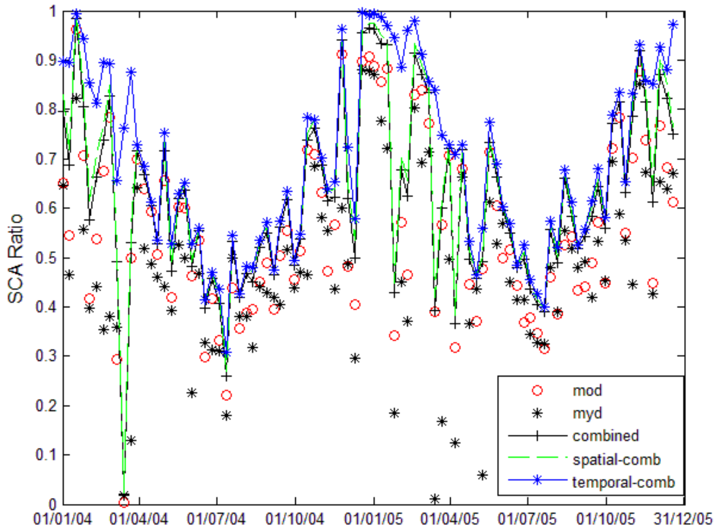


Figure 4. Filtered MODIS eight-day snow-cover products (2004-2005). The term 'mod' is the snow cover area from MOD10A2 products, 'myd' is MYD10A2 products, 'combined' is the combined result from step1, 'spatial-comb' from step2 and 'temporal-comb' from step3. See Sect. 2.2.3 for details.

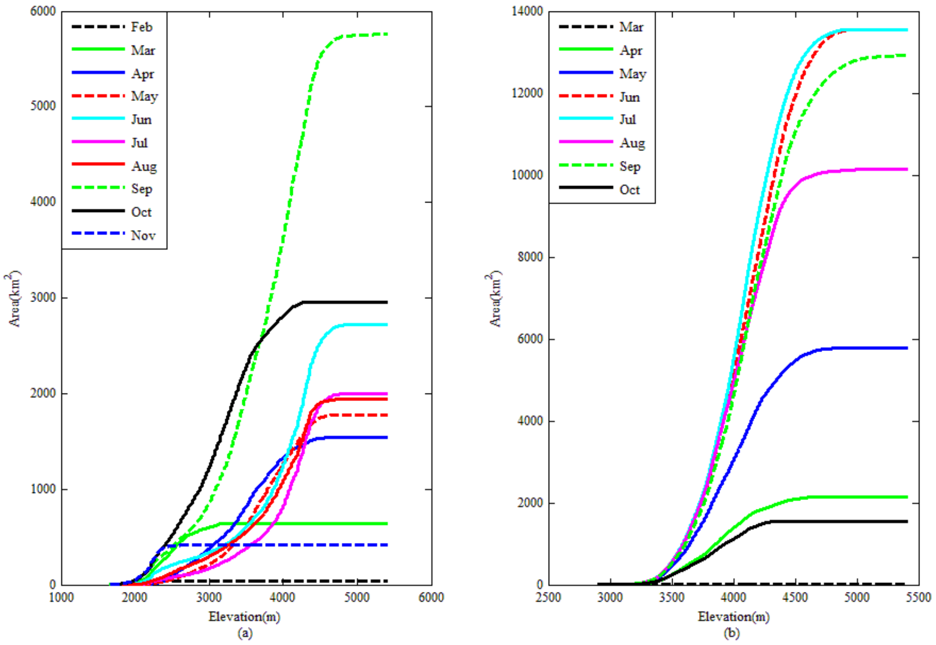


Figure 5. Altitudinal Cumulative Melt Curve. (a) Cumulative monthly snowmelt area distribution by elevation (2003~2012). (b) Cumulative monthly glacier melt area distribution by elevation (2003~2012). The snowmelt areas in December and January and the glacier melt areas in November, December, January and February are zero and are not shown in this figure.

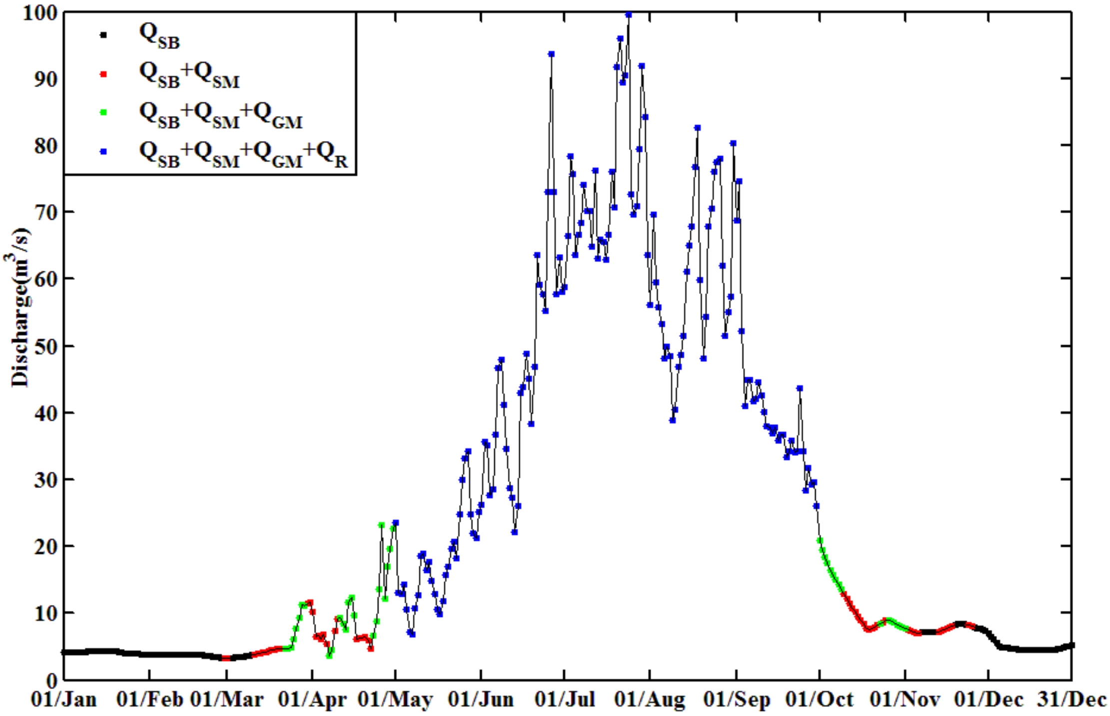


Figure 6. Hydrograph partition in 2003. Q_{SB} stands for subsurface baseflow generated by groundwater, Q_{SM} and Q_{GM} for snow meltwater and glacier meltwater respectively, and Q_R for rainwater directly runoff.

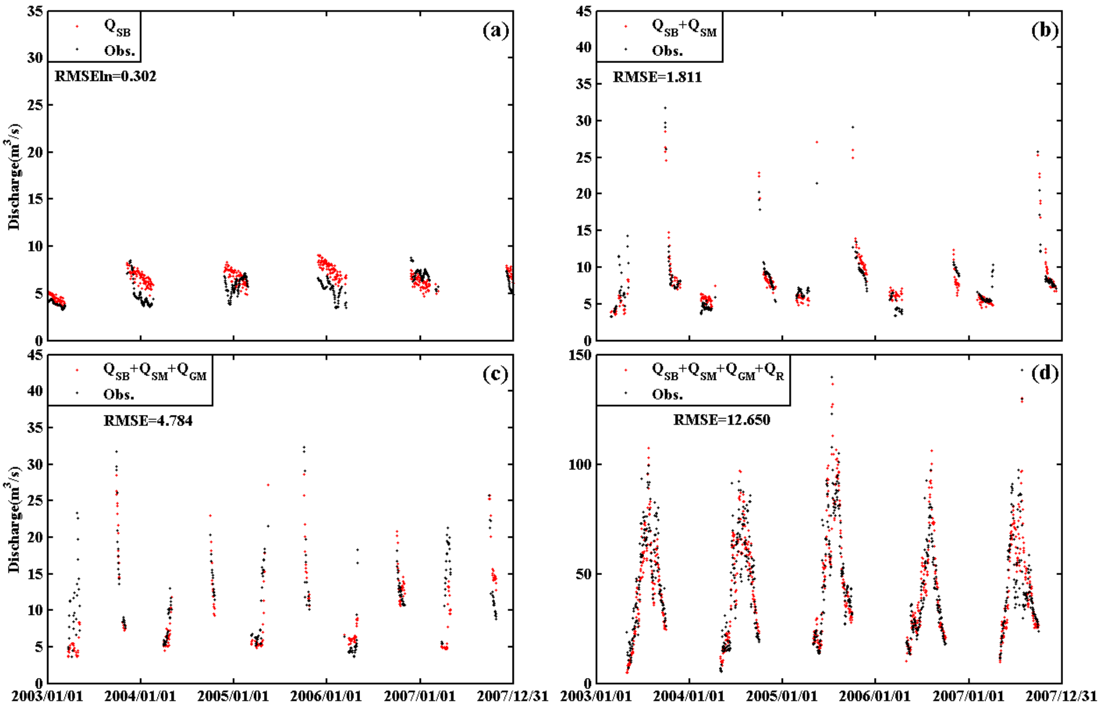


Figure 7. Stepwise calibration of grouped parameters upon partitioning curves. (a) Partitioning curves after calibrating K_A and K_D upon Q_{SB}. (b) Partitioning curves after calibrating D_s upon $Q_{SB}+Q_{SM}$. (c) Partitioning curves after calibrating D_g upon $Q_{SB}+Q_{SM}+Q_{GM}$. (d) Partitioning curves after calibrating W_M and B upon $Q_{SB}+Q_{SM}+Q_{GM}+Q_R$. The goodness of fit between observed and simulated discharge is measured by *RMSEln* (for Q_{SB} part) or *RMSE* (for other parts).

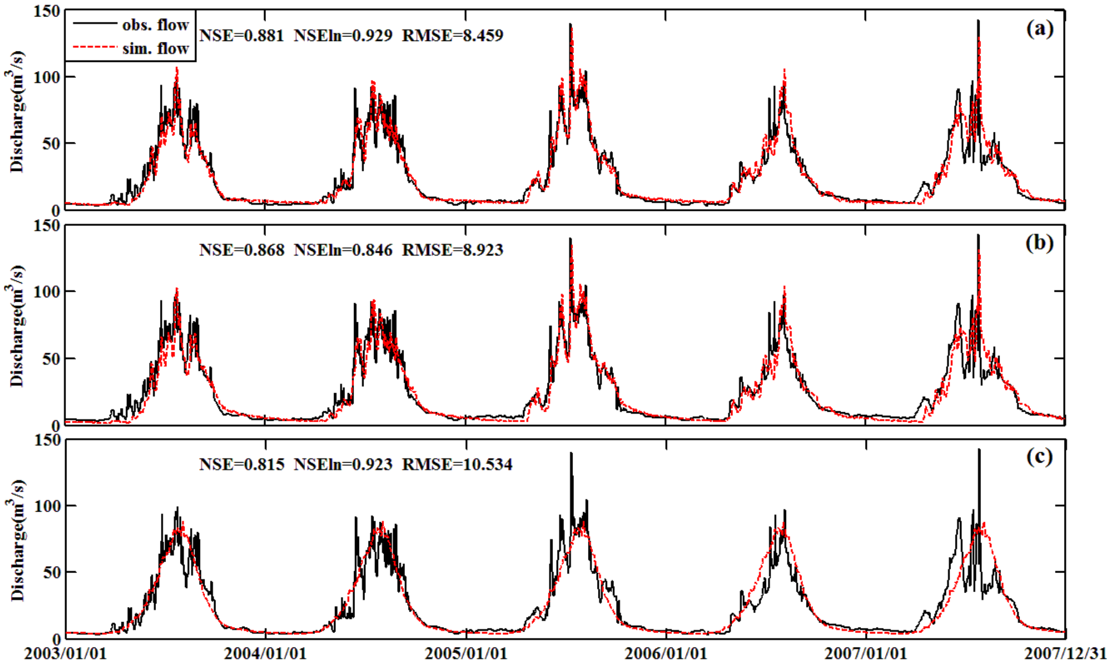


Figure 8. Simulation of daily streamflow by different methods from 2003 to 2007. (a) by the proposed stepwise method, (b) by the automatic calibration method, and (c) by the benchmark model. The performance of the simulations is measured in *NSE*, *NSEln* and *RMSE*.

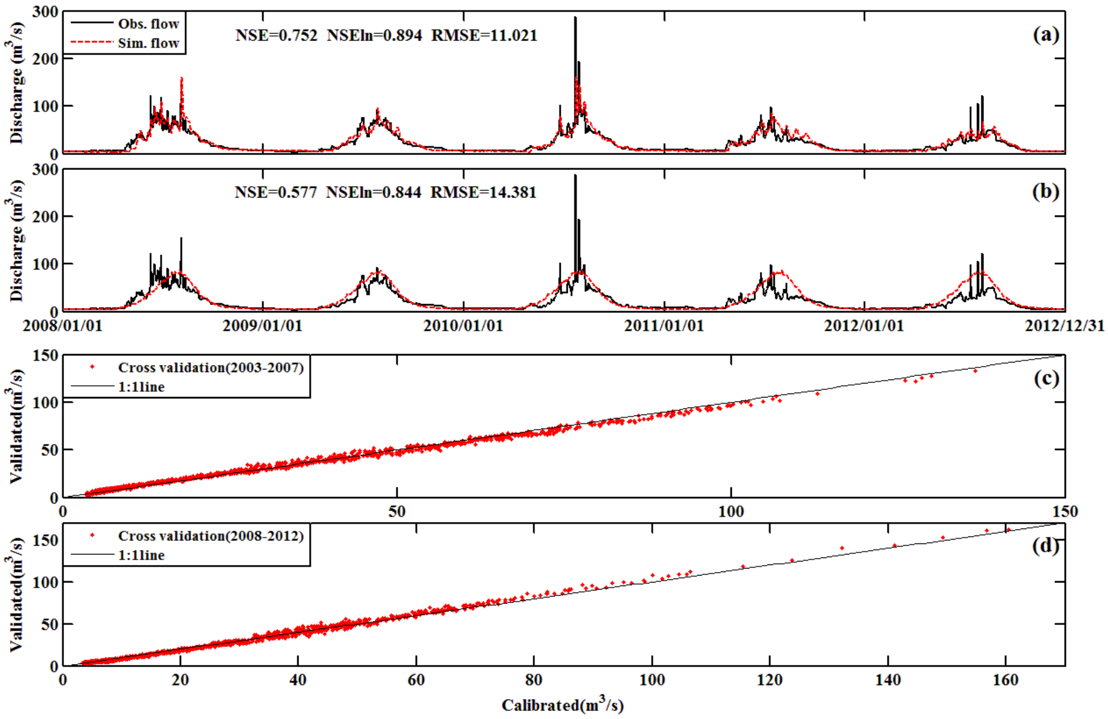


Figure 9. Evaluation of the stepwise calibration method. (a) discharge simulation in evaluation period 2008 to 2012 using the stepwise calibrated parameters in calibration period 2003 to 2007. (b) discharge simulation in evaluation period 2008 to 2012 by the benchmark model. (c) Cross validation simulation of daily discharge in 2003-2007. x-coordinate presents the simulated daily discharges by parameters calibrated in period 2003-2007. y-coordinate presents the simulated daily discharges by parameters calibrated in period 2008-2012. (d) Cross validation simulation of daily discharge in 2008-2012. x-coordinate presents the simulated daily discharges by parameters calibrated in period 2008-2012. y-coordinate presents the simulated daily discharges by parameters calibrated in period 2003-2007.

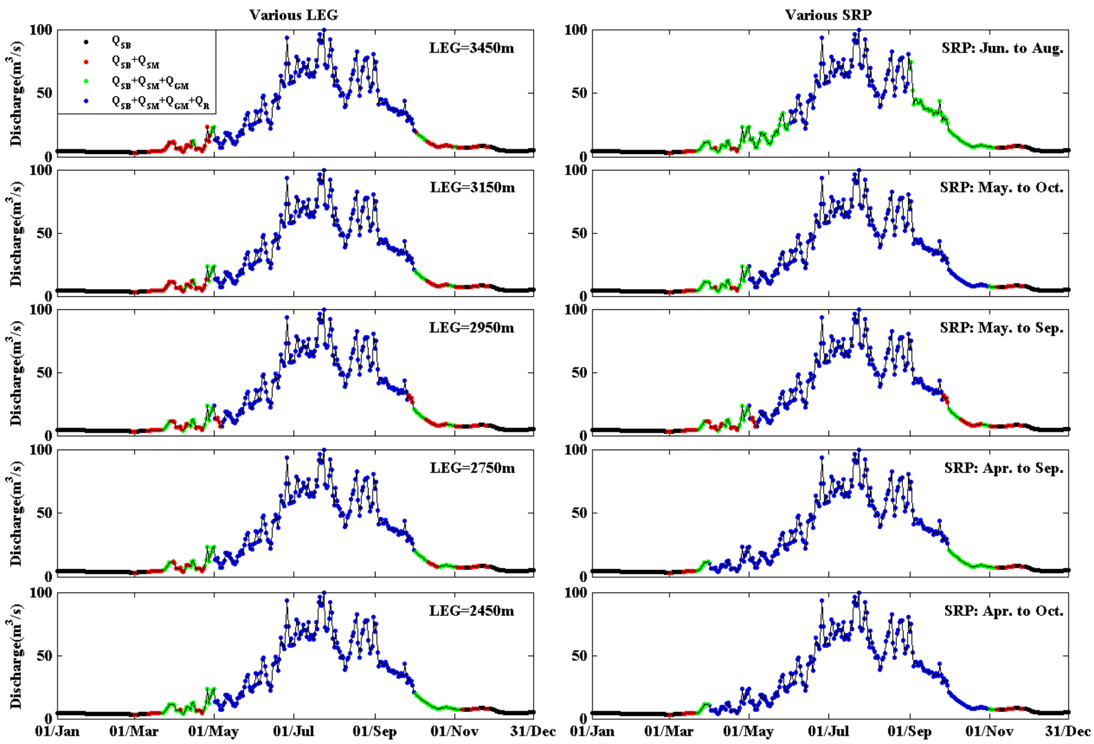


Figure 10. Sensitivity analysis for hydrograph partition. The first column is the hydrograph partition pattern using different lowest elevation band of the glacier area (LEG). The second column is the hydrograph partition pattern using different storm rain period (SRP).

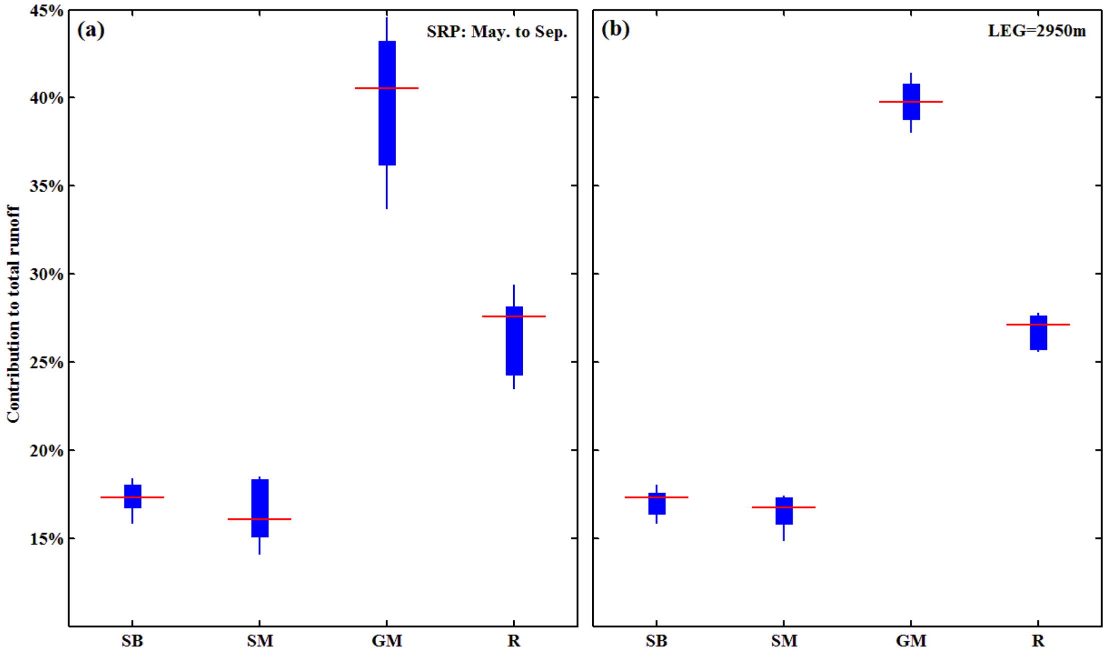


Figure 11. Sensitivity analysis on the contributions of different runoff sources to total runoff. (a) is the contribution pattern under different lowest elevation band of glacier area (LEG), where the storm rain period (SRP) is fixed as May to September. (b) is the contribution pattern under different SRPs, where the LEG is fixed as 2950m. The red line stands for the mean contribution for each runoff source, and the top/bottom end of each plot presents the highest/lowest contribution ratio. SB is groundwater baseflow, SM is snowmelt, GM is glacier melt and R is rainwater directly runoff.