Abstract: The impacts of climate change on water resources in snow- and glacier-dominated basins are of great importance for water resource management. The Snowmelt Runoff Model (SRM) was developed to simulate and predict daily streamflow for high mountain basins where snowmelt runoff is a major contributor. However, there are many sources of uncertainty when using an SRM for hydrological simulations, such as low-quality input data, imperfect model structure and model parameters, and uncertainty from climate scenarios. Among these, the identification of model parameters is considered to be one of the major sources of uncertainty. This study evaluates the parameter uncertainty for SRM simulation based on different calibration strategies, as well as its impact on future hydrological projections in a data-scarce deglaciating river basin. The generalized likelihood uncertainty estimation (GLUE) method implemented by Monte Carlo sampling was used to estimate the model uncertainty arising from parameters calibrated by means of different strategies. Future snowmelt runoff projections under climate change impacts in the middle of the century and their uncertainty were assessed using average annual hydrographs, annual discharge and flow duration curves as the evaluation criteria. The results show that: (1) the strategy with a division of one or two sub-period(s) in a hydrological year is more appropriate for SRM calibration, and is also more rational for hydrological climate change impact assessment; (2) the multi-year calibration strategy is also more stable; and (3) the future runoff projection contains a large amount of uncertainty, among which parameter uncertainty plays a significant role. The projections also indicate that the onset of snowmelt runoff is likely to shift earlier in the year, and the discharge over the snowmelt season is projected to increase. Overall, this study emphasizes the importance of considering the parameter uncertainty of time-varying hydrological processes in hydrological modelling and climate change impact assessment.

Keywords: Snowmelt Runoff Model; parameter uncertainty; data-scarce deglaciating river basin; climate change impacts; generalized likelihood uncertainty estimation

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) stated that the temperature will continue to increase in the 21st century [1]. Potential impacts of global warming on water availability, especially in...
snow- and glacier-dominated regions are of great concern [2]. Due to the sensitivity and vulnerability of snow and glaciers due to climate change, previous studies have paid considerable attention to how climate change impacts snow accumulation [3–5], as well as to the timing and amount of snowmelt [6–8]. Water resources in snow- and glacier-dominated regions will suffer from the probable risk of declining amounts of snow and size of glaciers. In order to assess the climate change impacts on the water resources in snow- and glacier-dominated watersheds, it is common to use data from climate simulations to drive snow/glacier hydrological models [9–11]. In these models, snow- and glacier-related processes are usually simulated by energy budget methods [12–14] and/or temperature index methods [15,16], as well as/or multi-layer snow models with snow depth data assimilation methods [17–19]. The physically-based energy-budget methods and snow models incorporate energy fluxes occurring within the snowpack and interacting with the surface [20], while the empirical temperature-index method assumes a linear relationship between air temperature and the rate of snowmelt [21]. Since the energy budget method requires more meteorological data as input, it is often impractical to conduct at high-altitude snow and glacier basins with sparse observations. Temperature-index methods are more practical for snowmelt simulation in data-scarce mountain regions with insufficient data, as the air temperature is relatively widely available compared to other components [21].

The Snowmelt Runoff Model (SRM) is one of the most commonly used temperature-index models. It is a conceptual model developed to simulate and predict daily streamflow based on the degree-day method in high mountain basins where snowmelt is a major contribution [16]. It has been used to assess the climate change impacts on river discharge in mountainous environments [11,22–26]. For instance, Elias et al., [11] investigated the climate change impacts on streamflow timing and runoff volume utilizing the SRM in the Upper Rio Grande basin, Colorado and New Mexico, USA, and found that the large decrease in runoff volume in summer would pose challenges to water management in this region. Tahir et al., [25] applied the SRM to assess the climate change impacts on snowmelt runoff in the Upper Indus Basin and showed that the SRM is effective in estimating snowmelt runoff in high mountain regions. Xie et al., [26] proposed and applied a progressive segmented optimization algorithm (PSOA) for calibrating time-variant parameters of the SRM on the Manas River basin of Xinjiang, China. Their study showed that the PSOA can effectively calibrate the time-variant model parameters while avoiding the high increase in computational time caused by a significant increase of parameter dimensionality.

However, there are multiple sources of substantial uncertainties when applying the SRM for snowmelt runoff modeling and projection. Snowmelt modeling generally involves three uncertainty sources related to forcing data (i.e., input data and data for calibration), imperfections in the model structure and model parameters [27]. In climate change impact studies, uncertainties arise from different climate scenarios, as well as from parameter instability due to the possible changes in a basins’ physical characteristics [28]. Among these uncertainties, the identification of model parameters is an essential step for any hydrological modeling, as there are parameters that cannot be derived directly from watershed characteristics but rather need to be inferred via a trial-and-error process [29,30].

Parameter uncertainty may be especially amplified in future climate change impact studies [31–33]. Therefore, quantifying uncertainties due to parameter instability, identifiability and non-uniqueness should be undertaken routinely, and robust methodologies should be implemented. Much attention has been given to uncertainty in model parameters in the hydrological literature [28,31,34–36]. Among the methods developed to deal with parameter uncertainty, the generalized likelihood uncertainty estimation (GLUE) method [37] has been widely used in recent decades [38–41]. The GLUE method is easy to implement and allows a flexible definition of the so-called likelihood function. The likelihood function, which is used to separate behavioural and non behavioural solutions is particularly valuable for assessment of integrated, distributed models that operate with multi-variable, multi-site and multi-response criteria [39]. Moreover, a numerical Monte Carlo sampling technique [42] which is used to sample from the prior parameter distribution is usually coupled with GLUE for quantifying
parameter uncertainties in hydrological modeling [40,43]. Conditioned on the available observational data and associated uncertainty bounds, this approach produces a well-distributed set of parameter distributions, the acceptability of which are similarly good at producing model simulations.

According to previous studies [44–46], the determination of parameters is essential to utilizing the SRM. Among the SRM parameters, the runoff coefficients ($C_s$ and $C_r$) expressing the losses as ratios ($C_s/C_r$ corresponds to the ratio of runoff contributed by snowmelt/rainfall to precipitation) vary in time, and so are affected by different climatic conditions and land cover properties in a hydrological year. They are usually calibrated if the agreement between observed and simulated runoffs is not acceptable [47]. The calibration strategy most commonly used for time varying parameters is to calibrate them in several sub-periods during the whole calibration period, and then assume that these parameters are constant in each sub-period. For example, Senzeba et al., [48] used the SRM to simulate the snowmelt runoff in a seasonally snow-covered eastern Himalayan catchment in 2006, 2007 and 2009. Single values of $C_s$ and $C_r$ were applied for each calibration year and the mean value from the three calibration years was used for a validation year (hereafter referred to as 1-period calibration). Zhang et al., [49] divided the hydrological year into two sub-periods (i.e., May–June and July–April) to calibrate the runoff coefficients of the SRM (hereafter referred to as 2-period calibration). In the study of Fuladipanah & Jorabloo [50], more than two sub-periods were divided in a hydrological year based on the variations of climatic conditions (hereafter referred to as multi-period calibration).

In the last few decades, the SRM has been applied successfully to mountainous watersheds all over the world, and it was recommended by Abudu et al., [46] that the SRM has a potential to forecast streamflow and to evaluate the effects of climate change on runoff in the mountainous watersheds of northwestern China, especially in data-scarce watersheds in high-elevation regions. However, the parameter uncertainty of the SRM restricts its application in this respect, as its parameter stability is relatively limited due to the highly dynamic behavior of snowpack and snowmelt.

Accordingly, this study aims to evaluate the parameter uncertainty for SRM simulation by using different calibration strategies, and to assess its impact on future hydrological projections. More specific objectives are to: (1) estimate the uncertainty of SRM simulation based on parameter uncertainty (i.e., the overfitting problem) and different calibration strategies (i.e., yearly calibration and multi-year calibration); and (2) assess the uncertainty of future SRM discharge projections determined by utilizing different calibrated parameter sets in a data-scarce deglaciating river basin.

In order to have a wide review of calibration strategies, we apply yearly calibration (which obtains optimal parameter sets from year to year) and multi-year calibration (which gets one optimal parameter value set for many years) with different sub-periods. The GLUE method is used to assess the flexibility and sensitivity of the parameters and to model the uncertainty arising from calibrated parameters.

2. Study Area and Data

2.1. Study Area

The data-scarce deglaciating Yurungkash watershed was chosen to implement the proposed study (Figure 1). It is one of the tributaries of the snow- and glacier-fed Tarim River Basin, the biggest inland river basin in China. The Yurungkash watershed is located approximately between 80°–82° N latitude and 74°–75° E longitude in the Xinjiang Uygur Autonomous Region and covers a surface area of about 14,600 km². It originates from the Akshay Glacier located on the northern slope of Kunlun Mountain, and its elevation ranges from about 1280 m above sea level (asl) up to 6780 m asl. The mean altitude of the watershed is about 4680 m asl and nearly 60% of the area lies above 5000 m asl.
The studied watershed is typically snow- and glacier-characterized, with about 40% of the watershed covered by permanent snow and ice. There is no meteorological station installed within the watershed; Hotan station is the closest one (Figure 1), about 600 km far away from the outlet of the river basin, located at 37.13° N, 79.93° E and 1375 m asl. The mean annual temperature of the study region is about –8 °C, and it receives a mean total annual precipitation of 260 mm with 65% falling between June and September. The Tunguz Locke station is located at the basin outlet (36.81° N, 79.92° E) and records a mean annual flow of 80 m³/s. The annual runoff presents a large intra-annual variability with about 90% occurring in the snowmelt season (April–September).

2.2. Data

The Shuttle Radar Topography Mission (SRTM), comprised of a modified radar system and offering high-resolution DEM data [51], was used to derive the topography of the study area and delineate the elevation zones. The Yurungkash watershed was divided into six elevation zones and the properties of these zone areas are listed in Table 1. The mean elevation of each elevation zone was calculated by using hypsometric curves, which were also used to derive the temperature of different elevation zones by extrapolation of base station temperature. Daily air temperature observed at the Hotan station and daily precipitation from the China Ground Rainfall Daily Value 0.5° × 0.5° Lattice Dataset (CGRD) [52] for the 2003–2012 period were used for model simulation. Two different datasets were used for temperature and precipitation because precipitation at the Hotan station cannot well represent the whole study watershed, due to the large spatial variability of precipitation in mountain basins. The CGRD was generated by interpolating observed precipitation from more than 2000 meteorological stations in China, utilizing the thin plate spline approach which was developed to interpolate and smooth scattered data in geosciences [53]. The reliability of this dataset has been proven by Zhao et al., [54]. Daily runoff data at the Tunguz Locke station for the 2003–2012 period was used for SRM calibration and validation.
The daily snow-covered area (SCA) is another model input, derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) [55] snow cover product. The MODIS 8-day composite snow cover data product (MOD10A2) is generated by compositing observations from the MODIS Snow Cover Daily L3 Global 500 m Grid (MOD10A1) data set. For every eight-day period, the grid is mapped as snow if snow is observed on any single day, and cloud cover is reported only if the grid was cloud-obscured for all eight days. The MOD10A2 can better eliminate the influence of clouds and improves the precision of snow identification over that of the MOD10A1 [56]. The daily SCA derived from the MOD10A2 by using linear interpolating was used for model simulation in this study. The mean monthly watershed-averaged temperature (T), precipitation (P) and snow-covered area (SCA) for the 2003–2012 period is presented in Figure 2. This figure indicates that the T is above 0 °C during May–September. The watershed receives much more precipitation in July and August than in other months. The SCA decreases from April and increases after August, and then decreases again after October. The minimum SCA of the watershed is greater than 40%, which is the area of its permanent snow/ice.

For runoff projection, General Circulation Model (GCM) simulations from phase 5 of the Coupled Model Intercomparison Project [57] were used to drive the SRM. Although it is desirable to use as many GCMs as possible, the selection of a representative subset is usually necessary due to the high computational costs of model simulations.

The k-means clustering method [58], which has been used in several previous studies [59–61], was applied to select subsets from 26 GCMs under representative concentration pathway 4.5 (RCP 4.5) and RCP 8.5, respectively. The k-means clustering method divided the ensemble of 26 GCMs into a number of clusters to minimize the within-cluster sum of squared error (SSE). The SSE was characterized by the Euclidean distance based on the standard changes in temperature and precipitation.

<table>
<thead>
<tr>
<th>Zone</th>
<th>Elevation Range (m)</th>
<th>Mean Elevation (m)</th>
<th>Area (km²)</th>
<th>Area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1280–2700</td>
<td>2230</td>
<td>1324</td>
<td>9.07</td>
</tr>
<tr>
<td>B</td>
<td>2701–3700</td>
<td>3190</td>
<td>2248</td>
<td>15.40</td>
</tr>
<tr>
<td>C</td>
<td>3701–4700</td>
<td>4250</td>
<td>2170</td>
<td>14.86</td>
</tr>
<tr>
<td>D</td>
<td>4701–5200</td>
<td>4975</td>
<td>2485</td>
<td>17.02</td>
</tr>
<tr>
<td>E</td>
<td>5201–5700</td>
<td>5440</td>
<td>3238</td>
<td>22.18</td>
</tr>
<tr>
<td>F</td>
<td>5701–6780</td>
<td>6020</td>
<td>3135</td>
<td>21.47</td>
</tr>
<tr>
<td>Total</td>
<td>1280–6780</td>
<td>4470</td>
<td>14600</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of elevation zones.

Figure 2. The watershed-averaged temperature (T), precipitation (P) and snow-covered area (SCA) (Period 2003–2012). The T, P and SCA were derived from the Hotan station, China Ground Rainfall Daily Value 0.5° × 0.5° Lattice Dataset and MODIS product, respectively.
simulated by GCMs between the 2003–2012 and the 2041–2050 periods. A value of four clusters was determined to offer good group separation as well as a manageable number of simulations. Figure 3 presents the selected subsets of GCMs under RCP 4.5 and RCP 8.5. In each cluster, the colour marker which is closest to the centroid is the selected one. Table 2 presents the detailed information of the seven GCMs used in this study.

Table 2. Seven General Circulation Model simulations’ details of CMIP5.

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Institute/Country</th>
<th>Horizontal Resolution (lon×lat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPI-ESM-MR</td>
<td>Max Planck Institute for Meteorology, Germany</td>
<td>192 × 96</td>
</tr>
<tr>
<td>MPI-ESM-LR</td>
<td>Max Planck Institute for Meteorology, Germany</td>
<td>192 × 96</td>
</tr>
<tr>
<td>IPSL-CM5A-LR</td>
<td>Institute Pierre-Simon Laplace, France</td>
<td>96 × 96</td>
</tr>
<tr>
<td>GFDL-CM3</td>
<td>USA</td>
<td>144 × 90</td>
</tr>
<tr>
<td>NorESM1-M</td>
<td>Norwegian Climate Centre, Norway</td>
<td>144 × 96</td>
</tr>
<tr>
<td>MIROC5</td>
<td>MIROC, Japan</td>
<td>256 × 128</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>MIROC, Japan</td>
<td>128 × 64</td>
</tr>
</tbody>
</table>

3. Methods

3.1. SRM Modeling

The SRM is a conceptual, simple degree-day-based model that was designed to simulate and forecast daily streamflow resulting from snowmelt and rainfall in mountain basins. Since it was developed by Martinec et al., [16], the SRM has been successfully applied to many mountain basins of various sizes ranging from 0.76 to 917,444 km², and elevations ranging from 0 to 8848 m asl [62]. In recent years, the SRM has also been applied to investigate the impacts of climate change on runoff [7,11,25].

In this study, the SRM is applied to simulate and predict daily runoff for the Yurungkash watershed. It requires three daily input variables (temperature, precipitation and snow-covered area) and has seven parameters (runoff coefficients Cs and Cr, the degree-day factor, critical temperature, rainfall contribution area, recession coefficient and the time lag). A detailed description of the SRM model can be found in the Appendix A.
3.2. SRM Calibration Methods

Among the SRM parameters, the runoff coefficients (Cs and Cr) vary with time, and so they are the parameters that are most sensitive to different climatic conditions and land cover properties over a year-long time frame [47]. Therefore, they are usually the parameters chosen for calibration in SRM modelling. In order to have a broad perspective for calibrating runoff coefficients, both the yearly and multi-year calibration strategies with 1, 2, 4, 5, 6, and 7 sub-periods were used for comparison. X-period calibration means that a hydrological year is divided into × sub-periods based on similar characteristics in climatic data for runoff calibration. In total, our approach used 12 calibration strategies. Additional details, including the beginning and ending dates of each sub-period of the different calibration strategies, are presented in Figure 4. For the 2-period calibration, the ending date of the last period is 31 May in the following year. For the other calibration strategies, the ending date of the last period is 31 March in the following year. Apart from Cs and Cr, other model parameters are presented in Table 3, which were determined as described in the appendix. Calibration was carried out in the odd-numbered years and validation was performed in the even-numbered years for the 2003–2012 period.

![Figure 4](image-url)

Figure 4. The beginning and ending dates of each period in a hydrological year for different calibration strategies. Each color in each calibration strategy represents a sub-period. For example, there are 2 sub-periods in the 2-period calibration, the 1st period is June 1 August 31 and the 2 period is September 1 May 31. a represents June 20, b represents July 15 and c represents August 5.
Table 3. Parameters determined by hydrological judgment as described in the appendix.

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree Day Factor (cm/°C/day)</td>
<td>0.3</td>
</tr>
<tr>
<td>Lapse Rate (°C/100 m)</td>
<td>0.65</td>
</tr>
<tr>
<td>Threshold Temperature, $T_{\text{crit}}$ 1 (June–August); 3 (September–May)</td>
<td></td>
</tr>
<tr>
<td>Rainfall Contributing Area, RCA</td>
<td>1 (May–September); 0 (October–April)</td>
</tr>
<tr>
<td>Recession Coefficient, $K$, which is Determined by: $K_{n+1} = \frac{Q_{n+1}}{Q_n} = x \times Q_n^{y}$</td>
<td>$X_c = 0.9895$; $X_y = 0.026$</td>
</tr>
</tbody>
</table>

The Nash-Sutcliffe coefficient (NSE) [63] and the relative volume error (VE), two classical efficiency criteria in hydrological modelling, were used to evaluate the performance of the SRM. The NSE and the VE are computed as follows:

\[
\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(Q_i - Q'_i)^2}{\sum_{i=1}^{n}(Q_i - \overline{Q})^2} 
\]

\[
\text{VE}[\%] = \frac{V_R - V'_R}{V_R} \times 100
\]

where $Q_i$ is the measured daily discharge, $Q'_i$ is the simulated daily discharge, $\overline{Q}$ is the average measured discharge of the given time period or snowmelt season, $V_R$ is the measured yearly or seasonal runoff volume, $V'_R$ is the simulated yearly or seasonal runoff volume, and $n$ is the number of daily discharge values.

3.3. Bias Correction of GCM Outputs

GCM outputs are too coarse and biased to directly drive an SRM in climate change impact studies; therefore, a downscaling or bias correction process is needed [64]. In this study, the watershed-averaged value of the GCM–simulated precipitation was first calculated by the Thiessen Polygon method [65]. The GCM–simulated temperature was interpolated to the Hotan station using the Inverse Distance Weighted method [66] in order to be extrapolated later to the higher elevation zones. The precipitation and temperature were then corrected, respectively based on the watershed-averaged precipitation from CGRD and temperature at the Hotan station, by a distribution-based bias correction method, called the Daily Bias Correction (DBC) method in Chen et al., [67]. The DBC is a hybrid method combining the Local intensity scaling (LOCI) method [68] to correct the precipitation occurrence and the Daily translation (DT) method [69] to correct the frequency distributions of precipitation amounts and temperatures. Here are the two steps of this DBC method:

1. The LOCI method was used to correct the precipitation occurrence, which ensures that the frequency of the precipitation occurrence simulated by GCMs at the reference period equals that of the observed data for a specific month. A threshold for precipitation occurrence determined at the historical period was then applied to the future period.
2. The DT method was used to correct the empirical distribution of GCM-simulated precipitation and temperature magnitudes in terms of 100 quantiles from 0.01 to 1 with an interval of 0.01.

3.4. Future Snow Covered Area Projection

A changed climate would cause a change in the SCA in mountain snow basins. The method for future SRM simulation developed by Rango & Martinec [70] was used to evaluate the impact of future temperature and precipitation on SCA in this study. This method estimates the time shift of the SCA in the present climate over the snowmelt period to produce the SCA in a future climate. The characteristics of the watershed (i.e., the number of elevation zones, the area and the mean elevation of each zone) and the SCA parameters (i.e., the degree-day factor, the temperature lapse rate and the critical temperature)
determined in the present climate, as well as the bias corrected temperature and precipitation in the future climate are required for future SCA projection. More detailed procedures for calculating SCA can be found in Rango and Martinec [70].

3.5. GLUE Method for Uncertainty Estimates

The GLUE method [37,71] coupled with Monte Carlo sampling [42] was used in this study to investigate the parameter uncertainty of the SRM. In this method, a large number of model runs were made by using many different parameter sets randomly selected from a priori probability distribution. The acceptability of each run was evaluated against observed values and, the posterior parameter distribution was extracted based on a certain subjective threshold to maintain a good population of behavioral solutions. The parameter set with the corresponding model run considered to be non-behavioral was removed from further analysis. In this study, the GLUE scheme associated with the SRM includes the following steps:

1. 100,000 Monte Carlo sampling points of $C_s$ and $C_r$ were implemented from a feasible parameter space (0–1) with uniform distribution;
2. The likelihood values were calculated for all 100,000 model runs;
3. The likelihood functions (NSE and VE) and the threshold values (0.55 and ±10%, respectively) were specified as behavioral parameter sets; and
4. Posterior parameter sets were extracted depending on the threshold of the likelihood functions.

The sample sizes of the posterior parameter distributions were all over 1000. These posterior parameter sets were then used in validation and projection. It should be mentioned that for the yearly calibration, the posterior parameter sets of each year were used individually in the validation and projection.

The following three indices were used to evaluate the resolution, reliability and efficiency of the uncertainty interval estimates:

1. The Average Relative Interval Length (ARIL) [34] measures the resolution of the predictive distributions, which is defined over the entire simulation time period as:

$$ARIL = \frac{1}{n} \sum \frac{Q_{Upper,t} - Q_{Lower,t}}{Q_{obs,t}}$$  (3)

where $n$ is the number of days in the observed record, $Q_{Upper,t}$ and $Q_{Lower,t}$ are the upper and lower simulated discharges of the 95% confidence interval, respectively, and $Q_{obs,t}$ is the observed discharge.

The percentage of observations bracketed by the Confidence Interval (PCI) [72] measures the amount of the observed data within a range of simulated intervals, defined as:

$$PCI = \frac{Q_{obs,in}}{n}$$  (4)

where $Q_{obs,in}$ is the number of observed discharges that are contained within the 95% confidence interval. A smaller ARIL means a narrower uncertainty interval and a larger PCI means that the interval has a greater reliability; and

2. The percentage of observations bracketed by the Unit Confidence Interval (PUCI) [35] based on ARIL and PCI, which is defined as:

$$PUCI = \frac{(1 - \text{Abs}(PCI - 0.95))}{ARIL}$$  (5)

Larger PUCI values represent smaller uncertainty of 95% confidence interval of discharge in general.

For the climate change impacts on runoff projection for the 2041–2050 period, the sub-periods’ division in a hydrological year may be different due to the changes in climatic conditions and land
cover properties, which may bring further uncertainty to hydrological projections. This study assumed there is no change of sub-period division in the future climate. To better represent the changed climate, the value of $C_s$, which reflects the decline of the snow coverage and the stage of vegetation growth, was shifted earlier by 30 days in the climate run, as recommended by previous studies [11,16,73]. The climate change impacts on hydrology were assessed at the 95% confidence intervals of average annual hydrographs, annual discharge, and flow duration curves.

4. Results

4.1. Parameter Uncertainty

The SRM was calibrated based on the yearly and multi-year calibration strategies, with different sub-periods (i.e., 1-, 2-, 4-, 5-, 6- and 7-period strategies) in the odd-numbered years, and validation was performed in the even-numbered years during the 2003–2012 period. Calibration performances with $\text{NSE} > 0.55$ and $-10\% < \text{VE} < 10\%$, as well as the corresponding validation performances are shown in Figure 5 for different calibration strategies. This figure shows that: (1) for both yearly and multi-year calibration strategies, the maximum NSE values at the calibration period increase from 1- to 5-period approaches, while there is no obvious improvement for the 6- and 7-period strategy; (2) the maximum NSE values are larger at the calibration period for yearly calibration, but smaller at the validation period than those for multi-year calibration; and (3) there are much more samples with NSE values below 0.55 and absolute values of VE above 10% when using more than two sub-periods at the validation period, especially for yearly calibration.

Figure 5. Scatter plots of the Nash-Sutcliffe coefficient (NSE) (upper) and the relative volume error VE (lower) of calibration versus validation for different calibration strategies. The left subplots present the yearly calibration and the right subplots present the multi-Year calibration.

4.2. Confidence Interval of Discharge

Three uncertainty indices for evaluating 95% confidence intervals of discharge derived from different calibration strategies are given in Table 4. The following features can be observed: (1) both
ARIL and PCI increase with the increase of sub-periods for both yearly and multi-year calibration strategies, making it difficult to identify which calibration strategy is better; (2) considering the PUCI, the 1- and 2-period calibrations have a larger value of PUCI than the other sub-periods, which indicates their results are more reliable and less uncertain; and (3) the PUCI values of the multi-year calibrations are generally larger than those of the yearly calibrations, which means the former approach is better.

Table 4. The Average Relative Interval Length (ARIL), the percentage of observations bracketed by the Confidence Interval (PCI) and the percentage of observations bracketed by the Unit Confidence Interval (PUCI) values of the different calibration strategies for the 95% confidence interval of discharge by different calibrations.

<table>
<thead>
<tr>
<th>Sub-Period(s)</th>
<th>Yearly Calibration</th>
<th>Multi-Year Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIL</td>
<td>PCI</td>
</tr>
<tr>
<td>1</td>
<td>0.454</td>
<td>0.204</td>
</tr>
<tr>
<td>2</td>
<td>1.009</td>
<td>0.430</td>
</tr>
<tr>
<td>4</td>
<td>1.287</td>
<td>0.500</td>
</tr>
<tr>
<td>5</td>
<td>1.531</td>
<td>0.586</td>
</tr>
<tr>
<td>6</td>
<td>1.546</td>
<td>0.596</td>
</tr>
<tr>
<td>7</td>
<td>1.548</td>
<td>0.606</td>
</tr>
</tbody>
</table>

The 95% confidence intervals of the simulated daily discharge are presented in Figure 6. The results from the 5-, 6- and 7-period calibrations are not shown in this figure because they are very similar to those from the 4-period calibration. Figure 6 shows that the confidence intervals of discharge derived by yearly calibration are slightly larger than those derived by multi-year calibration. The confidence intervals also become larger with the increase in sub-periods, which is consistent with the higher PCI values in Table 4. Furthermore, discharges can be observed lying outside the confidence intervals, which indicates that there are some other uncertainty sources related to forcing data and imperfections in the model structure that impact SRM discharge simulation in addition to the parameter uncertainty.

Figure 6. The 95% confidence intervals of daily discharge simulated by posterior parameter sets (color envelope), compared with observed runoff (black lines).
4.3. Projected Future Temperature, Precipitation, and SCA

By the degree-day method, daily temperature in a future climate modifies the SCA in the present climate to the approximate future SCA in the snowmelt season (April–September). Figure 7 depicts the daily watershed-averaged temperature, precipitation and SCA in snowmelt season from four selected GCMs under RCP 4.5 and RCP 8.5 for the 2041–2050 period (hereafter ‘seasonal’ means the average value from the snowmelt season). Four GCMs suggest an increase in temperature over the catchment between 1.2 to 2.6 °C under RCP4.5, and an increase of between 1.2 to 3.3 °C under RCP8.5. The precipitation over the catchment is also projected to increase by 12.0% to 39.8% under RCP4.5, and by 12.2% to 61.8% under RCP8.5. The SCA is projected to decrease in response to the high temperatures in late June. The figure shows that the predicted SCA decreases rapidly from a (June 21st, 58.0%) to b (July 31st, 40.1%) under RCP4.5 and from c (June 11st, 60.3%) to d (July 25th, 42.8%) under RCP8.5. The relative decrease in seasonal SCA ranges from 5.7% to 17.7% under RCP4.5, and from 1.2% to 14.8% under RCP8.5.

![Figure 7](image-url)

**Figure 7.** Projected mean monthly temperature (T), precipitation (P) and snow coverage area (SCA) (Period 2041–2050) compared with observations (Period 2003–2012). The pink shading represents the four chosen GCMs, whose mean value is represented by the red line. The same representation is used for the precipitation, indicated in blue. a represents (June 21st, 58.0%), b represents (July 31st, 40.1%), c represents (June 11st, 60.3%), d represents (July 25th, 42.8%).

4.4. Uncertainty in the Discharge Projections

The daily discharges are simulated by the SRM using bias corrected temperature and precipitation from the chosen GCMs under RCP4.5 and RCP8.5, as well as the projected SCA, utilizing the posterior parameter sets derived by the yearly and multi-year calibration strategies with 1-, 2- and 4-period segments. Figure 8 presents the 95% confidence interval of the mean daily discharge for the snowmelt season of the 2041–2050 period. The observed mean daily discharge of the 2003–2012 period is also presented for comparison. The confidence intervals represent uncertainties from four different GCMs as well as from the calibrated parameters. The following results can be observed: (1) the median values...
of the future discharge simulated by the yearly and multi-year calibration strategies are higher than those of the historical discharge, with the onset of snowmelt runoff shifted earlier. (2) The parameter uncertainty of the yearly calibration is larger than that of the multi-year calibration, as indicated by the wider pink intervals. (3) The discharge confidence interval is wider when there are more sub-periods, in particular for the 4-period strategy, which is much wider than the others.

**Figure 8.** The 95% uncertainty intervals for the GCM-srm predicted daily discharge (Period 2041–2050). The bottom and top of the colored intervals represent the 2.5th and 97.5th percentile values, respectively. In each subplot, the pink and blue shadings represent the yearly and the multi-year calibration strategies, respectively. Median values of the pink and blue envelopes are shown as the red and blue curves, respectively.

Figure 9 shows the 95% confidence intervals of the mean discharge in the snowmelt season from four GCMs under RCP4.5 and RCP8.5 for the 2041–2050 period. The observed mean discharge during the snowmelt season for the 2003–2012 period is also shown. Generally, the projections suggest increases in the mean discharge for the future period, especially under RCP8.5. The confidence intervals of mean discharge from the 1- and 2-period calibrations are smaller than those from the 4-period calibration. In order to compare different sources for uncertainty, these discharge confidence intervals were grouped into five sources (RCPs, GCMs, multi-year/yearly calibration strategies, division of sub-periods, and calibrated parameters). For example, to investigate the uncertainty related to the RCP, discharges were grouped by two RCPs. Each group includes discharges from four GCMs under RCP4.5 and RCP8.5 for the 2041–2050 period. The observed mean discharge in particular for the 4-period strategy, which is much wider than the others.

**Figure 9.** The 95% confidence intervals of the mean discharge in the snowmelt season from four GCMs under RCP4.5 and RCP8.5 for the 2041–2050 period. The observed mean discharge during the snowmelt season for the 2003–2012 period is also shown. Generally, the projections suggest increases in the mean discharge for the future period, especially under RCP8.5. The confidence intervals of mean discharge from the 1- and 2-period calibrations are smaller than those from the 4-period calibration. In order to compare different sources for uncertainty, these discharge confidence intervals were grouped into five sources (RCPs, GCMs, multi-year/yearly calibration strategies, division of sub-periods, and calibrated parameters). For example, to investigate the uncertainty related to the RCP, discharges were grouped by two RCPs. Each group includes discharges from four GCM, and yearly and multi-year calibration strategies with different sub-periods. The results show that uncertainty ranges related to the choice of RCP, GCM, multi-year/yearly calibration, sub-period and calibrated parameter are [283.11, 308.25], [260.92, 347.93], [267.30, 318.49], [201.86, 229.22] and [165.18, 307.91] with the unit of m$^3$/s, respectively. In terms of the uncertainty range, the largest uncertainty is found to be related to the SRM parameters.
Figure 9. The 95% confidence intervals for the GCM–SRM predicted average discharge of the snowmelt season (Period 2041–2050). The bottom and top of the colored bars represent the 2.5th and 97.5th percentile values, respectively. The black bar represents the average snowmelt season discharge during 2003–2012.

The 95% confidence intervals of the flow duration curve (FDC) for four GCMs under RCP8.5 for the 2041–2050 period are shown in Figure 10. Only the high-representative concentration pathway of RCP8.5 is presented to illustrate the uncertainties of the FDC and discuss the differences between GCMs and calibration strategies. Figure 10 reveals that the low, mid, and high flows are all predicted to increase in the Yurungkash watershed for the 2041–2050 period. The highest flow, which occurs less than 10% of the time, is predicted to increase or decrease for different GCMs. This suggests that the number of days with flows exceedingly above 10% will increase in the future, but there will be no significant change in the number of days with the highest flows.
Figure 10. The 95% confidence intervals for the GCM–SRM predicted flow duration curves under RCP8.5 (Period 2041–2050) compared with the observations (Period 2003–2012) shown as black line. The dotted and solid lines represent the 2.5th and 97.5th percentile values, respectively.

5. Discussion

It is widely acknowledged that calibration plays an important role in any hydrological modelling due to parameters that are spatially and temporally heterogeneous or that cannot be measured. Hence, the problem of parameter optimization and the associated uncertainty must be addressed [74]. Our results show that the simulation performance of the SRM is sensitive to calibration strategies with different sub-periods for calibrating the time-varying parameters. In terms of the NSE and the VE, the model’s calibration performance becomes better with more sub-periods. However, there are more NSE values below 0.55 when the sub-period is more than 4, indicating higher uncertainty. According to previous studies [32,75], increasing the number of calibrated parameters could yield better calibration performance, while also bringing larger uncertainties at the validation period. An overfitting problem may arise from over-parameterization when the amount of information contained in a hydrograph used for calibration is not enough to estimate a larger number of parameters [76]. An over-fitted model results in a smaller error for calibration but a larger error for validation. Hence, the 1– and 2-period strategies appear to be the most rational choice for calibration when considering the balance between simulation performances and the overfitting problem. In addition, the multi-year calibration would be more stable than the yearly calibration for SRM modeling, as the former performs better at the validation period.

In addition, much attention has been paid to parameter uncertainty assessment in hydroclimatic modeling recently [33,77]. In this study, both the yearly and multi-year calibration strategies with 1-, 2- and 4-periods were used for hydrological climate change impact assessment. Although the runoff projection contains large uncertainty, it shows that the onset of snowmelt runoff shifts earlier and the mean discharge of the snowmelt season increases. This result is similar to that of Wang et al., [78],
in which the response of snowmelt runoff to climate change was investigated in a basin located in northwest China.

Particularly, it is noticeable that the predicted hydrographs derived from the 4-period calibration do not shift earlier for peak river runoff. This is not consistent with the common view that the effects of a warming climate will lead to a peak flow shifting from summer and/or autumn to winter and/or early spring [2]. Elias et al., [11] applied a forward shift of 30 days for $C_s$ based on a 2-period calibration in the climate change impact assessment and predicted higher streamflow in March and April, which is similar to the results derived from the 1- and 2-period calibrations in our study. The result of the 4-period calibration is different from our expectations because it shifts a large value of $C_s$ is shifted to August, which leads to the peak flow occurring in August under climate change. This result indicates that the shift in runoff timing predicted by more sub-periods contains greater uncertainty. In addition, the predicted hydrograph and mean discharge show much wider intervals with the 4-period calibration than with the 1- and 2-period calibrations, which also indicates larger uncertainties. Hence, the 1- and 2-periods are recommended for SRM hydrological climate change impact assessment. This work also suggests that a multi-year calibration strategy would be more stable than a yearly calibration strategy in climate change impact studies.

It is necessary to point out some of the limitations of this study. This study adopted the GLUE method to estimate parameter uncertainty, one drawback of which is that the derived parameter distributions and uncertainty intervals are sensitive to some subjective factors, e.g., the threshold values of the estimation criteria [40,43]. Lower threshold values could result in a wider uncertainty interval of the posterior distribution [34] which may explain why the parameter uncertainty of SRM has the largest impact on the future hydrological projection compared to other uncertainties in this study. This result is similar to that of Tian et al. [79], in which parameter was found to be the major source of uncertainty for simulating future high flows over a southern river basin in China. However, this result is different from that of previous studies [64,80], which showed that hydrological model parameter uncertainty was relatively small compared to other uncertainty sources (RCP and GCM). In previous studies, hydrological model parameters were calibrated by a certain optimization algorithm, which is not recommended for the SRM [47]. However, the use of thresholds in this study cannot guarantee that all parameters are optimized. Nevertheless, the results indicate that the uncertainty associated with a parameter should be highlighted in SRM modeling. It is acknowledged that the parameter distribution and uncertainty intervals derived by GLUE have no clear statistical meaning [81]. Comparison of the results by different threshold values is needed to test the sensitivity of the GLUE, as recommended by Jin et al., [34].

Another limitation of this study is that only different calibration strategies were taken into account in the parameter estimation. It would be more comprehensive to calibrate parameters over different periods, such as wet and dry climate periods [82], as well as by a multi-objective calibration approach [83]. In addition, Seiller et al., [77] showed that the diagnosis of the impacts of climate change on water resources is very much affected by the objective functions used for hydrological model calibration. Investigation of the parameters’ nonstationarity in a changed climate could be another avenue for future studies [84].

Also, it is necessary to mention that although temperature-index models are more practical, they are often less accurate than full energy balance models [85]. In addition, the skill of temperature-index models is likely to decrease for future climate change impact studies due to the changes in the energy balance which is not a direct result of temperature (e.g., changes in albedo). The degree day factor has been shown to change over time scales of a decade or even over a single melt season [85–87].

6. Conclusions

This study evaluated the parameter uncertainty for SRM simulation by different calibration strategies and its impact on future hydrological projections in a data-scarce deglaciating river basin.
Different sub-periods were compared and the uncertainty of future discharge derived by the calibrated parameter sets was assessed. The following conclusions can be drawn:

1. The strategy with a division of 1 or 2 sub-period(s) in a hydrological year is appropriate for SRM modeling when considering the balance between simulation performance and the overfitting problem/uncertainty. In addition, the multi-year calibration approach is more stable than the yearly calibration for SRM hydrological simulation and projection, as the latter presents a lower validation performance combined with higher future projection uncertainty.

2. The future runoff projection contains large uncertainties, among which parameter uncertainty plays a significant role. The projection results indicate that the onset of snowmelt runoff is likely to shift earlier, and the discharge of the snowmelt season is projected to increase for the 2041–2050 period.

Overall, this research highlights the parameter uncertainty of SRM and its impact on hydrological climate change assessment in a data-scarce deglaciating river basin. The results imply that for hydrological models, like the SRM, time-variant parameters could result in over-parameterization issues and bring uncertainties in the hydrological simulation. Both of these have a significant effect on climate change impact assessment, and therefore should be routinely considered when using these models.

**Author Contributions:** Data curation, Y.X.; Funding acquisition, L.L. and J.C.; Investigation, Y.X.; Methodology, Y.X. and L.L.; Writing – original draft, Y.X.; Writing – review & editing, L.L., J.C., C.-Y.X., J.X., H.C. and J.L.

**Funding:** This research was funded by [the State Key Laboratory of Water Resources and Hydropower Engineering Science funding] grant number [2017SWG02], [the National Natural Science Foundation of China] grant number [51779176, 51525902, 51539009], and [the Thousand Youth Talents Plan from the Organization Department of CCP Central Committee (Wuhan University, China)]. The APC was funded by [the State Key Laboratory of Water Resources and Hydropower Engineering Science funding] grant number [2017SWG02].

**Acknowledgments:** This work was partially supported by the State Key Laboratory of Water Resources and Hydropower Engineering Science funding (No. 2017SWG02), the National Natural Science Foundation of China (Grant No. 51779176, 51525902, 51539009) and the Thousand Youth Talents Plan from the Organization Department of CCP Central Committee (Wuhan University, China). The authors would like to acknowledge the contribution of the World Climate Research Program Working Group on Coupled Modeling, and to thank climate modeling groups for making available their respective climate model outputs. The authors wish to thank the China Meteorological Data Sharing Service System and the Xinjiang Tarim River Management Bureau for providing the dataset for the Yurungkash river basin.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**Appendix A**

**Appendix A.1 SRM Modelling**

The following equation shows the daily discharge estimated by superimposing snowmelt and rainfall on the calculated recession flow:

\[ Q_{n+1} = \left[ C_{sn} \times a_n(T_n + \Delta T_n)S_n + C_{rn} \times P_n \right] \times A \times a \times (1 - K_{n+1}) + Q_n K_{n+1} \]  

(A1)

where:

- \( Q \) = the average daily discharge (m³/s)
- \( C_s/C_r \) = the runoff coefficient expressing the losses as a ratio of runoff to precipitation, with \( C_s \) referring to snowmelt and \( C_r \) referring to rainfall
- \( a \) = the degree-day factor (cm/°C/day)
- \( T \) = the number of degree-days (°C day)
- \( \Delta T \) = the adjustment by temperature lapse rate when extrapolating the temperature from the station to the mean elevation of the zone (°C day)
\( S = \) the ratio of the snow covered area to the total area (%)
\( P = \) the precipitation contributing to runoff (cm), which is determined by a preselected threshold temperature, \( T_c \), to be rainfall or snowfall. The contribution of rainfall is immediate while snowfall will be kept on storage until melting conditions occur.
\( A = \) the area elevation zone (km\(^2\))
\( \alpha = 10,000/86,400 \), the coefficient converting data from a runoff depth (cm\( \times \)km\(^2\)/day) to discharge (m\(^3\)/s)
\( K_{n+1} = \) the recession coefficient, \( K_{n+1} = \frac{Q_{n+1}}{Q_n} \)
\( n = \) the day number

It is recommended to subdivide the basin into several elevation zones if the elevation range exceeds 500 m. The simulated discharge of each zone is summed to estimate the runoff of the basin. The study area was divided into 6 elevation zones, characters of which are present in Table 1. Model structure are shown above, the determination of input variables and parameters are declared below:

Appendix A.1.1 Input Variables

\( T, P, S \) are three basic inputs of SRM.

The daily snowmelt depth is calculated by the number of degree-days, which is determined by the daily temperature. In order to run SRM for each elevation zone, daily temperature from Hotan station was extrapolated to the mean elevation of each zone based on the global lapse rate value of 6.5 °C/km.

The daily watershed-averaged precipitation was derived from the China Ground Rainfall Daily Value 0.5° × 0.5° Lattice Dataset by using the Thiessen polygon (TP) method, which was used for each elevation zone.

The daily SCA for each elevation zone was generated from the SCA of 8-day composite MOD10A2 satellite images by using linear interpolating.

Appendix A.1.2 Parameters

\( C_s, C_r, a, T_c, RCA, K, \) and time lag are the seven parameters to set up SRM, whose value are presented in Table 3.

The runoff coefficients \( C_s \) and \( C_r \) account for the losses between the runoff contributed by snowmelt and rainfall to precipitation. \( C_s \) and \( C_r \) are time-varying parameters, affected by lots of controlling factors such as climatic conditions and land cover properties. They can vary over seasonal, monthly, or even daily scale, which are the primary candidates for tuning if a runoff simulation is not immediately successful [47].

The degree-day factor \( a \) (cm\(^3\)/°C/day), converts the degree-days \( T \) (°C day) into the daily snowmelt depth \( M \) (cm) by Equation (A2). \( a \) is related to snow properties. According to the research for degree-day factors in western China conducted by Zhang et al. [88], 0.3 is adopted in this study.

\[ M = a \times T \]  

(A2)

The critical temperature \( T_c \) is used to determine whether the precipitation is rainfall (\( T > T_c \)) or snowfall (\( T < T_c \)). The \( T_c \) values for June–August and September–May are set to 3 °C and 0 °C, respectively, As recommended by Martinec [47].

The rainfall contributing area RCA represents the way that how to treat the rainfall determined by \( T_c \) and it needs to be determined according to the basin characteristics beforehand. When RCA is set to be 0, it is assumed that the rain falling on the snowpack is retained by the snow which is dry and deep. In this situation, rainfall depth is reduced by the ratio snow-free area/zone area. RCA is set to be 1 when the snowpack is ripe, which means that rainfall from the entire area is added to snowmelt. In this study, the snowpack is assumed to be ripe during May–September when the temperature is above 0 °C.

The recession coefficient \( K \) indicates the decline of discharge in a period without snowmelt or rainfall, and \((1 \times K)\) presents the proportion of the meltwater production immediately appearing in the
runoff. $K$ is determined by Equation (A3) [47], in which parameters $x$ and $y$ can be obtained through linear regression analysis using the historical discharge data.

$$K_{n+1} = \frac{Q_{n+1}}{Q_n} = x Q_n^{-y}$$

(A3)

There is a time lag ($L$) between the temperature cycle and the resulting discharge cycle in Equation (A1). In this study, a lag time of 18 h is adapted depending on the relation between $L$ and basin size [16]. In the case of 18 h, temperature measured on the $n$th day corresponds to the discharge on the $(n + 1)$ day.

References


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).