# Assessing the influence of water sampling strategy on the performance of tracer-aided hydrological modeling in a mountainous basin on the Tibetan Plateau

Yi Nan[1], Zhihua He[2], Fuqiang Tian[1], Zhongwang Wei[3], Lide Tian[4]

[1] Department of Hydraulic Engineering, State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing, China

[2] Center for Hydrology, University of Saskatchewan, Saskatchewan, Canada

[3] Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, School of Atmospheric Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China

[4] Institute of International Rivers and Eco-security, Yunnan University, Kunming, China

***Corresponding to*:** Fuqiang Tian

Email: tianfq@tsinghua.edu.cn

**Abstract**

Tracer-aided hydrological models integrating water isotope module into the simulation of runoff generation are useful tools to reduce uncertainty of hydrological modeling in cold basins that are featured by complex runoff processes and multiple runoff components. However, there is little guidance on the strategy of field water sampling for isotope analysis to run tracer-aided hydrological models, which is especially important for large mountainous basins on the Tibetan Plateau (TP) where field water sampling work is highly costly. This study conducted a set of numerical experiments based on the THREW-T (Tsinghua Representative Elementary Watershed - Tracer-aided version) model to evaluate the reliance of the tracer-aided modeling performance on the availability of site measurements of water isotope in the Yarlung Tsangpo River (YTR) basin on the TP. Data conditions considered in the numerical experiments included the availability of glacier meltwater isotope measurement, quantity of site measurements of precipitation isotope, and the variable collecting strategies for stream water sample. Our results suggested that: (1) In high-mountain basins where glacier meltwater samples for isotope analysis are not available, estimating glacier meltwater isotope by an offset parameter from the precipitation isotope is a feasible way to force the tracer-aided hydrological model. Using a set of glacier meltwater $\delta^{18}O$ that were 2‰~9‰ lower than the mean precipitation $\delta^{18}O$ resulted in only small changes in the model performance and the quantifications of contributions of runoff components (CRCs, smaller than 5%) to streamflow in the YTR basin; (2) strategy of field sampling for site precipitation to correct the global gridded isotope product of isoGSM for model forcing should be carefully designed. Collecting precipitation samples at sites falling in the same altitude tends to be worse at representing the ground pattern of precipitation $\delta^{18}O$ over the basin than collecting precipitation samples from sites in a range of altitudes; (3) Collecting weekly stream water samples at multiple sites in the wet and warm seasons is the optimal strategy for calibrating and evaluating a tracer-aided hydrological model in the YTR basin. It is highly recommended to increase the number of stream water sampling sites rather than spending resource on extensive sampling of stream water at a sole site for multiple years. These results provide important implications for collecting site measurements of water isotope for running tracer-aided hydrological models to improve quantifications of CRCs in the high-mountain basins.

## 1. Introduction

Catchments located in mountainous regions generally provide important water resources for downstream regions (Viviroli et al., 2003). As typical mountainous cryosphere, the Tibetan Plateau (TP) is the source region for several large rivers in Asia, and has been called as a ‘water tower’ because of its importance for downstream livelihoods and agricultural irrigations (Schaner et al., 2012). Dominant characteristic of mountainous catchments on TP is the multiphase of water sources that generate runoff and the consequently complex hydrological processes, highlighting the importance of accurately quantifying the contributions of runoff components (CRCs) to streamflow for better understandings the runoff dynamics under changing climate. This task is difficult due to the complex hydrological processes being insufficiently represented by typical hydrological models, leading to large uncertainty of hydrological simulations (He et al., 2018). Due to the strong inter-compensation of runoff processes induced by different water sources and runoff pathways (Duethmann et al., 2015), uncertainties of the modeled CRCs in mountainous basins on the TP are rather high. Utilizing more datasets to evaluate the model performance is a feasible way to constrain modeling uncertainty and improve quantifications of CRCs in cold regions (Chen et al., 2017).

Tracer-aided hydrological models integrating environmental tracer (e.g., stable oxygen isotope, $^{18}$O) modules into runoff generation processes have proved helpful for parameter calibration, model structure diagnosis and CRC quantification (Son and Sivapalan, 2007; Birkel et al., 2011), and are increasingly adopted in cold catchments (e.g., Ala-aho et al., 2017; He et al., 2019; Nan et al., 2021a). Recent studies indicated that estimates of precipitation $\delta^{18}$O from outputs of isotopic general circulation models (iGCMs) perform well on forcing tracer-aided models in large basins with a high cost of water sampling (Delavau et al., 2017; Nan et al. 2021b). Similarly to the tracer-based end-member mixing methods that utilize the different tracer signatures of water sources to separate the hydrograph and quantify CRCs (Klaus and McDonnell, 2013; He et al., 2020), the tracer-aided hydrological models used the differed isotopic compositions of runoff components to regulate the water apportionments in runoff generation. The isotopic compositions of runoff components strongly differ in high-mountain basins resulting from the following two reasons: One is the significantly more depleted $\delta^{18}$O of meltwater compared to that of rain, due to the altitude and temperature effects, and the fractionation effect during melting processes (Xi, 2014; Boral and Sen, 2020). Another is the damping and lagging isotopic variability of subsurface runoff pathway, compared to that of surface runoff, as a result of the catchment hydrological functions of storing, mixing and transporting water (Bowen et al., 2019; Birkel and Soulsby, 2015; McGuire and McDonnell, 2006). Consequently, water isotope signatures show potential to improve the representations of internal hydrological processes in hydrological models, if observations of water isotopes were involved in the model calibration and evaluation procedures (McGuire et al., 2007; He et al., 2019).

Although a plenty of isotope-based works have been conducted in mountainous catchments on the TP to improve understandings of local hydrological processes (e.g., Li et al., 2020; Kong et al., 2019; Tan et al., 2021), few of them provided guidance on data collection of water isotope for hydrological applications in large mountainous areas. Some water sampling works in large mountainous catchments were conducted in a single field campaign (e.g., Xia et al., 2019; Dong et al., 2018), which is, although helpful to understand the generations of short-term runoff events, not suitable for the calibration of tracer-aided models in a multi-year simulation period (Knapp et al., 2019; Zhang et al., 2019). An exception is Stevenson et al. (2021) who utilized a 7-year dataset of stream water $\delta^{18}O$ in a 3.2 $km^2$ catchment to analyze the effects of stream water sampling strategies on the calibration of a tracer-aided hydrological model. Challenges arise when transferring their findings to the application of tracer-aided hydrological models in large high-mountain basins: First, it is questionable that whether sampling stream water at one site can adequately represent the isotope signature of stream water over the whole large basin, considering the strong spatial variability of hydrological processes caused by the heterogeneity in meteorological factors and land surface conditions in mountains (Wang et al., 2021; Li et al., 2020). Second, the influences of data collection of precipitation isotope on the performance of tracer-aided hydrological models remain unclear. Results of He et al. (2019) indicated that monthly sampling of precipitation at two sites seems to be able to capture the isotope variations in a 233 $km^2$ catchment. However, the requirement of isotope data quantity to adequately capture the spatial pattern of precipitation isotope signature for forcing tracer-aided models in large basins (~$10^5$ $km^2$) is poorly explored (Nan et al., 2021b). Third, in glacierized mountainous catchments where streamflow was fed by additional water source of glacier melt, the requirement of glacier meltwater samples for the forcing and evaluation of tracer-aided hydrological models is also unclear. Consequently, better understandings of how water sampling strategies influence the value of water isotope data for aiding hydrological modeling, is highly helpful for guiding the establishment of monitoring systems of water isotope in large mountainous regions. Considering the high costs of human and financial resources of collecting water samples in TP area, it is important to take efficient strategies for water sampling that balance the trade-off between field work burden and data adequacy well (Sprenger et al., 2019).

Motivated by the mentioned backgrounds, we conducted detailed analysis on the tracer-aided model performance in a large mountainous basin on the TP under different assumed situations with respect to the collection strategy of site water isotope data, based on a numerical experiment method. We adopted the tracer-aided hydrological model THREW-T developed by Nan et al. (2021a), which was forced by the global gridded isotope outputs of iGCM being merged with measurements of precipitation $\delta^{18}O$, to achieve the research aim. Three specific questions were addressed: (1) how does the estimated isotopic composition of glacier meltwater influence the performance of tracer-aided hydrological modeling when no glacier meltwater samples were available, (2) how does the collection strategy of site precipitation samples for precipitation isotope data merging influence the model performance, and (3) how does the sampling strategy of stream water influence the model calibration and evaluation? This study focused on the sampling strategy of precipitation and stream water, while the influence of glacier/snow meltwater isotope data sampling was not within the scope of this study.

## 2. Materials and methodology

### 2.1 Study area

The Yarlung Tsangpo River (YTR) basin, located in the southern TP (Fig. 1), extends in the ranges of 27°N -32°N and 82°E -97°E, with an elevation extent of 2900-6900 m above sea level (a.s.l.), which is one of the largest basins on the TP. The mean annual precipitation in the YTR basin is around 470mm featured by a distinct wet season from June to September, due to the dominance of the South Asian monsoon. Drainage area above the Nuxia hydrological station at the basin outlet is approximately $2\times10^5$ km$^2$, and around 2% of which is covered by glacier.

The Karuxung River (KR) catchment is located in the upper regions of the YTR basin, and was chosen as a supplementary experiment catchment, because of the long term field work of water sampling in this catchment. The KR originates from the Lejin Jangsan peak of the Karola mountain (7206m a.s.l.), and flows into the Yamdrok Lake (4550m a.s.l.), draining an area of around 286 km$^2$. Streamflow in the KR catchment is strongly influenced by glaciers which cover an area of 58 km$^2$.

**[Figure 1]**

### 2.2 Hydro-meteorological and water isotope data

Elevation of the YTR basin was derived from a digital elevation model (DEM) with a spatial resolution of 30m from the Geospatial Data Cloud (https://www.gscloud.cn). Daily meteorological inputs including precipitation, temperature and potential evapotranspiration were collected from the 0.1°×0.1° China Meteorological Forcing Dataset (CMFD, Yang and He, 2019). The second glacier inventory data set of China (Liu, 2012) and the Tibetan Plateau Snow Cover Extent product (TPSCE, Chen et al., 2018) were used to denote the glacier and snow coverages. Vegetation coverages were extracted from the MODIS satellite products of eight-day leaf area index (LAI) dataset MOD15A2H (Myneni et al., 2015) and monthly normalized difference vegetation index (NDVI) dataset MOD13A3 (Didan et al., 2015). Soil types and properties in the tested basins were collected from the Harmonized World Soil Database (HWSD, He, 2019). Observations of daily streamflow during 2000-2015 at the Nuxia, and that during 2000-2010 at Yangcun and Nugesha stations were used for hydrological model evaluation.

In the KR catchment, daily temperature and precipitation during 2006-2012 were collected at the Langkazi meteorological station. Altitudinal distributions of temperature and precipitation across the KR catchment were estimated based on the lapse rates reported in Zhang et al. (2015). Daily streamflow during 2006-2012 was measured at the Wengguo hydrological station.

Outputs of the scripps global spectral model with water isotopes incorporated (isoGSM, Yoshimura et al., 2008) with the spatial and temporal resolutions of 1.875°×1.875° and 6h were extracted to represent the spatio-temporal pattern of the precipitation isotope in the YTR basin. According to a previous evaluation of the isoGSM product (Nan et al., 2021b), while it can effectively capture the seasonal variation of precipitation $\delta^{18}O$, it had two major flaws: it overestimated precipitation $\delta^{18}O$ in the YTR basin, and performed poorly on accurately capturing the isotope signature of specific precipitation events and time periods. Higher elevation stations typically had a stronger bias. To obtain measurement precipitation $\delta^{18}O$ data, grab samples of precipitation were collected in the wet season of 2005 at four stations along the main channel of YTR, i.e., Nuxia (3691 m a.s.l.), Yangcun (4541m a.s.l.), Nugesha (4715m a.s.l.) and Lazi (4889m a.s.l.). The precipitation water samples were collected as soon as possible after the precipitation event in order to avoid the effect of evaporation. Stream water samples were collected weekly during the same period from river at the four stations.

The isoGSM isotope products were merged with measurement precipitation isotope data according to Eqs. 1-3 to provide input data for model: First, the bias of isoGSM product was assumed to be linearly related to altitude. Relation between the mean bias of isoGSM products and altitude was estimated by a least square method using $\delta^{18}O$ measurements of precipitation samples and gridded isoGSM estimates at the four sampling sites (Eqs. 1-2); Second, in each REW, precipitation $\delta^{18}O$ was determined by Eq. 3, based on the average altitude and the availability of $\delta^{18}O$ measurements from precipitation site samples on the date.

$$B_i = \overline{\delta^{18}O_{i,\mathrm{M}}} - \overline{\delta^{18}O_{i,\mathrm{G}}} \tag{1}$$

$$B = a \cdot H + b \tag{2}$$

$$\delta^{18}O_{k,j,\mathrm{Merged}} = \begin{cases} \delta^{18}O_{k,j,\mathrm{G}} + B_k, & \text{for date } j \text{ with no data} \\ \frac{\sum_{i=1}^{4}\delta^{18}O_{i,j,\mathrm{M}}}{4} - \frac{\sum_{i=1}^{4}\overline{\delta^{18}O_{i,\mathrm{M}}}}{4} + \overline{\delta^{18}O_{k,\mathrm{G}}} + B_k, & \text{for date } j \text{ with data, but unit } k \text{ containing no sampling site} \\ \delta^{18}O_{k,j,\mathrm{M}}, & \text{for date } j \text{ with data, and unit } k \text{ containing sampling site} \end{cases} \tag{3}$$

where, $B_i$ is the bias of isoGSM at sites *i*. $\overline{\delta^{18}O_{i,\mathrm{M}}}$ and $\overline{\delta^{18}O_{i,\mathrm{G}}}$ are the weighted average of the site measurement and isoGSM estimate over the sampling period at sites *i*, respectively. H is the altitude of the sampling site. Parameters *a* and *b* are the linear regression coefficients, which were estimated as -0.0046 and 14.96 by the least square method in this study. $\delta^{18}O_{k,j,\mathrm{Merged}}$

is the precipitation isotope obtained by merging isoGSM and measurement data, and $\delta^{18}O_{k,j,\mathrm{G}}$ refers to the original isoGSM isotope estimate at the hydrological model unit $k$ on the date $j$.

Glacier meltwater $\delta^{18}$O was assumed to be constantly lower than the weighted average of precipitation $\delta^{18}$O by an offset parameter ($\Delta_{\delta}$) during the study period (Eq. 4) because of the unavailability of glacier meltwater samples, which is generally within the range of 2-9‰ in the worldwide mountain regions (Rai et al., 2019; Wang et al., 2016; He et al., 2019; Ohlanders et al., 2013; Jeelani et al., 2017) and is adopted as 5‰ from Boral and Sen (2020) in the YTR basin.

$$\delta^{18}O_{k,\mathrm{GM}} = \overline{\delta^{18}O_{k,\mathrm{Corr}}} - \Delta_{\delta} \tag{4}$$

In the KR catchment, grab samples of precipitation and stream water were collected at the Wengguo station in 2006-2007 and 2010-2012 for isotope analysis. The spatial distribution of precipitation $\delta^{18}$O was estimated based on an altitudinal lapse of -0.34‰/100 as reported in Liu et al. (2007). Glacier meltwater $\delta^{18}$O was assumed to be constantly as -18.9‰ during the study period (as reported by Gao et al. 2009). Details of precipitation and stream water samples in the YTR and KR catchments were summarized in Table 1.

**[Table 1]**

### 2.3 Tracer-aided hydrological model

A distributed tracer-aided hydrological model, THREW-T (Tsinghua Representative Elementary Watershed - Tracer-aided version) model developed by Tian et al. (2006) and Nan et al. (2021a) was adopted for streamflow and isotope simulations. This model uses the representative elementary watershed (REW) method for spatial discretization of catchments (Reggiani et al., 1999). The study catchment is first divided into REWs based on DEM, and each REW is further divided into two vertical layers (surface and subsurface layers), including eight hydrological subzones based on the land cover and soil properties. In total, 63 and 41 REWs were extracted for the YTR basin and KR catchment, respectively (Tian et al., 2020; Nan et al., 2021a, 2021b). Areal averages of the gridded estimates of meteorological variables, vegetation cover and soil property were calculated in each REW to drive the model. A module representing glacier melting and snowpack evolution was incorporated into the model for application in cold regions (He et al., 2015; Xu et al., 2019; Tian et al., 2020). Accumulation and melting processes of snowpack were simulated according to temperature and precipitation, to update the snow water equivalent (SWE) of each REW. The snow cover area (SCA) was then calculated using the snow cover depletion curve (Fassnacht et al., 2016) and SWE threshold value (Parajka and Bloschl, 2008) for YTR basin and KR catchment, respectively, due to the different catchment scales. For simplification, the evolution of glacier was not simulated in the model. The temperature-index approach was used to calculate the amount of glacier melting, and it was assumed that the glacier melting water would directly contribute to streamflow through surface runoff pathway.

The tracer-aided module was developed by Nan et al. (2021a). The isotope was assumed to mix completely in each hydrological simulation unit within a simulation step. The Rayleigh fractionation method was adopted to simulate the isotope fractionation during water evaporation (similarly to He et al. 2019, Hindshaw et al. 2011, Wolfe et al. 2007). No parameters related to isotope modeling were introduced, since the isotope concentration was updated based on the water content of each unit and fluxes among them, which have been calculated in the runoff generation and flow concentration modules of the model. Forced by the inputs of precipitation and glacier meltwater isotopic compositions, the model simulates the isotope evolution in all the water storages in the watershed, including stream water, soil water and snowpack. The glacier evolution processes were not simulated in the hydrological model, therefore an assumed constant value was adopted to determine the isotope mass carried by glacier meltwater, instead of updating the isotope composition of glacier like other water storages. The iGCM isotope products properly corrected by $\delta^{18}O$ measurements of precipitation samples have proved feasible to force the THREW-T model in large catchments like YTR on the TP (Nan et al., 2021b). More details of hydrological model together with the snowpack evolution and tracer-aided module are given in Tian et al. (2006) and Nan et al. (2021a)

The THREW-T model quantified the contributions of runoff components (CRC) to streamflow based on two definitions of runoff components as reviewed in He et al. (2021). The first definition is based on the individual water sources in the total water input triggering runoff processes, including rainfall, snowmelt and glacier melt. The second definition is based on pathways of runoff-generation processes, resulting in surface and subsurface runoff (baseflow).

Physical basis and value ranges of the calibrated parameters in the THREW-T model were described in Table 2. The value of parameter was assumed to be universal for all the REWs. Two kinds of calibration approaches were conducted: (1) a bi-objective calibration using discharge and SCA, and (2) a tri-objective calibration using discharge, SCA and stream water $\delta^{18}O$. Metrics used to evaluate the model performance are listed in Eqs. 5-8. The Nash-Sutcliffe efficiency coefficient (NSE) was used to optimize the simulation of discharge and isotope, whereas the root-mean-square error (RMSE) was used for the evaluation of SCA simulation. The Logarithmic Nash-Sutcliffe efficiency coefficient (lnNSE) was used additionally for discharge calibration to assess the simulation of baseflow. The model parameters were calibrated by streamflow and SCA observations during 2001-2010 (at Nuxia station) and 2006-2012 in the YTR and KR basins, respectively. The model performance in YTR basin was validated by the Nuxia streamflow and SCA observations during 2011-2015, and the streamflow observations at Yangcun and Nugesha stations during 2001-2010.

$$\mathrm{NSE_{dis}} = 1 - \frac{\sum_{i=1}^{n}(Q_{\mathrm{o},i} - Q_{\mathrm{s},i})^2}{\sum_{i=1}^{n}(Q_{\mathrm{o},i} - \overline{Q_\mathrm{o}})^2} \quad (5)$$

$$\mathrm{NSE_{lndis}} = 1 - \frac{\sum_{i=1}^{n}(\ln Q_{\mathrm{o},i} - \ln Q_{\mathrm{s},i})^2}{\sum_{i=1}^{n}(\ln Q_{\mathrm{o},i} - \overline{\ln Q_\mathrm{o}})^2} \quad (6)$$

$$\mathrm{RMSE_{SCA}} = \sqrt{\frac{\sum_{i=1}^{n}(SCA_{\mathrm{o},i} - SCA_{\mathrm{s},i})^2}{n}} \quad (7)$$

$$\mathrm{NSE_{iso}} = 1 - \frac{\sum_{i=1}^{n}(\delta^{18}\mathrm{O}_{\mathrm{o},i} - \delta^{18}\mathrm{O}_{\mathrm{s},i})^2}{\sum_{i=1}^{n}(\delta^{18}\mathrm{O}_{\mathrm{o},i} - \overline{\delta^{18}\mathrm{O}_\mathrm{o}})^2} \quad (8)$$

where, *n* is the total number of observations. Subscripts of "o" and "s" refer to observed and simulated variables, respectively.

An automatic algorithm Python Surrogate Optimization Toolbox (pySOT) developed by Eriksson et al. (2017) was adopted for the multiple-objective optimization. The pySOT algorithm used a surrogate model to guide the search for improved solutions, with the advantage of needing few function evaluations to find a good solution. In each pySOT running, the optimization procedure was stopped if a maximum number of allowed function evaluations was reached, which was set as 3000 in this study. For the bi- and tri-objective calibrations, $0.5\cdot(\mathrm{NSE_{dis}}+\mathrm{NSE_{lndis}})-\mathrm{RMSE_{SCA}}$ and $0.5\cdot(\mathrm{NSE_{dis}}+\mathrm{NSE_{lndis}})-\mathrm{RMSE_{SCA}}+\mathrm{NSE_{iso}}$ were chosen as the combined optimization objectives. For each scenario, the pySOT algorithm was repeated 100 times, and behavioral parameter sets were selected among the 100 final results according to the performance metric thresholds, i.e., only the parameter sets producing metrics better than certain threshold values were regarded as behavioral parameter sets. The model uncertainty was evaluated based on the model performance driven by the behavioral parameter sets. The threshold values of evaluation metrics were used as $0.5\cdot(\mathrm{NSE_{dis}}+\mathrm{NSE_{lndis}})>0.8$, $\mathrm{RMSE_{SCA}}<0.08$ in the YTR basin; and $\mathrm{NSE_{dis}}>0.7$, $\mathrm{RMSE_{SCA}}<0.15$ in the KR catchment. Different values were adopted for the $\mathrm{NSE_{iso}}$ threshold among different scenarios, which would be introduced in the Result section.

**[Table 2]**

## 2.4 Numerical experiments

The influences of isotope data conditions on model performance were evaluated in three aspects as listed in Table 3: the assumed glacier meltwater isotope composition, the site measurement of precipitation isotope for data merging, and the stream water sampling strategy for model calibration.

**[Table 3]**

*Experiment 1: influence of assumed glacier meltwater isotope*

The first experiment was designed to test the reliance of model performance on the assumed glacier meltwater isotope, as glacier melt water samples are typically not available for isotope analysis in high mountain basins on the TP. In this experiment, variable glacier melt isotope signatures were adopted to calculate the isotopic contribution from glacier meltwater to streamflow, assuming the glacier meltwater $\delta^{18}O$ is 1‰, 3‰, 7‰ and 9‰ (i.e., $\Delta_\delta$ values in Table 3) lower than the long-term average $\delta^{18}O$ of precipitation. A benchmark model running by the literature based $\Delta_\delta$ value of 5‰ was used as a baseline reference to assess the influence of the assumed glacier meltwater isotope on the model performance.

*Experiment 2: influence of site measurement of precipitation isotope*

The second experiment was designed to test the reliance of the model performance on the availability of measured site precipitation isotope that was merged with the isoGSM product. The benchmark model running was forced by the merging precipitation isotope data based on measurements of precipitation isotope from all the four sampling sites (Figure 1). Three scenarios regarding the availability of measured precipitation isotope were designed as shown in Table 3. First, we assumed that only precipitation isotope measured at the two downstream sites of Nuxia and Yangcun are available for data merging (i.e., scenario P_2stationNY in Table 3). Second, we assumed that precipitation isotope measurement at the most upstream site Lazi is available. In addition to the measurement at the downstream site Nuxia (i.e., scenario P_2stationNL in Table 3). Third, we assumed that only precipitation isotope measurement at the most downstream site Nuxia is available for the data merging (i.e., scenario P_1station in Table 3).

*Experiment 3: influence of stream water sampling strategy*

The third experiment was conducted to analyze the influence of stream water sampling strategy on the model performance. Two types of stream water sampling strategies were considered, i.e., a time series sampling strategy based on regular and continuous sampling work at a certain point, and a spatially distributed sampling strategy based on one-time field campaigns of sampling work. For the time series sampling strategy, 7 scenarios ( “RT_YTR_” scenarios in Table 3) were designed to analyze the influences of the sampling frequency, the duration of the sampling period, and the number of sampling sites. For the spatially distributed sampling strategy, two scenarios (Figure 1b) were designed to represent typical field campaign activities: colleting samples along the mainstream of the basin (RS_YTR_Main, Table 3), and collecting water samples additionally from major tributaries (RS_YTR_Tributary, Table 3). Considering the limited availability of stream water $\delta^{18}O$ measurement in the YTR basin (only wet season in one year, Table 1), a supplementary experiment was designed to test the influence of sampling period duration on the model performance using the relatively long time-series isotope dataset in the small catchment KR (“RT_KR_” scenarios in Table 3).

To evaluate the influence of isotope data availability on the model performance, we carried out benchmark model simulations forced by full datasets of input isotope and stream water isotope data in the YTR and KR catchments (Table 3). The benchmark model runs were calibrated by a bi-objective calibration using SCA and streamflow observations, and a tri-objective calibration using additional stream water isotope, respectively. It is noted that, in the scenarios of experiment 3 in YTR basin (i.e., "RT_YTR_" and "RS_YTR_" scenarios in Table 3), the assumed data availability was beyond the actual measurement dataset. Consequently, the assumed stream water $\delta^{18}$O measurement data were obtained from a model simulation driven by a benchmark parameter set (rather than a subset of actual measurement stream water $\delta^{18}$O), which was selected from the behavioral parameters of the BM_YTR scenario calibrated by the tri-objective approach. The influence of the availability of stream water $\delta^{18}$O measurement on the tracer-aided model were evaluated by comparing the estimated CRCs and corresponding uncertainties with the assumed true values that were derived from the tri-objective calibrated benchmark running. Mean absolute error (MAE) and standard deviation (STD) were used to quantify the accuracy and uncertainty of CRC, which were calculated in Eqs. 9 and 10.

$$\mathrm{MAE}^{k} = \frac{\sum_{i=1}^{n}\left|\mathrm{CRC}_{\mathrm{s},i}^{k} - \mathrm{CRC}_{\mathrm{o}}^{k}\right|}{n} \tag{9}$$

$$\mathrm{STD}^{k} = \sqrt{\frac{\sum_{i=1}^{n}\left(\mathrm{CRC}_{\mathrm{s},i}^{k} - \overline{\mathrm{CRC}_{\mathrm{s}}^{k}}\right)^{2}}{n}} \tag{10}$$

where, $n$ is the number of behavioral parameter sets, and superscript $k$ indicates the runoff component (one of rainfall, snowmelt, glacier melt and baseflow). Subscript s and o indicate the simulated and observed value (observed value is the CRC produced by the tri-objective calibrated benchmark running). $\mathrm{CRC}_{\mathrm{s},i}^{k}$ is the contribution of runoff component $k$ simulated by the parameter set $i$. $\overline{\mathrm{CRC}_{\mathrm{s}}^{k}}$ is the average CRC simulated by all the behavioral parameter sets.

In the scenarios of experiments 1 and 2, the model was calibrated towards the complete stream water $\delta^{18}$O measurement dataset (Table 1), and the influence of isotope data availability on model performance were quantified by changes in model performance in the validation period and internal validate hydrological stations, as well as the uncertainty of CRC estimated by Eq. 10. In the scenarios of experiment 3 in the KR catchment (i.e., "RT_KR_" scenarios in Table 3), subsets of stream water $\delta^{18}$O measurement dataset (Table 1) with different length were picked out for model calibration.

## 3. Results

### 3.1 Performance of the tracer-aided hydrological model

Figure 2 shows performance of the benchmark model running (i.e., BM_YTR scenario in Table 3) forced and calibrated by the full available isotope dataset. The $\mathrm{NSE}_{\mathrm{iso}}$ threshold by which behavioral parameter sets were selected in tri-objective calibration was set as 0.5. Seasonal variations in discharge and SCA were reproduced well by the bi-objective calibration (Figure 2a and 2b), indicated by the high values of $\mathrm{NSE}_{\mathrm{dis}}$ (>0.8) and $\mathrm{lnNSE}_{\mathrm{dis}}$ (>0.8), and a low

$RMSE_{SCA}$ (<0.08). The peak flows were less well reproduced by the model in comparison to the simulation of baseflow processes, partly due to the inaccurate precipitation input data at the high altitudes. The model showed extremely poor performance for the simulation of stream water isotope when looking at the large uncertainty range (Figure 2c) and low $NSE_{iso}$ (-0.72). The tri-objective calibration significantly improved the isotope simulation (Figure 2f), without bringing much sacrifice to the performance in simulating discharge and SCA (considering the minimum values of $NSE_{dis}$ and $lnNSE_{dis}$ are around 0.7 in Figure 2d and 2e). Moreover, the tri-objective calibration slightly reduced uncertainty for simulation of the rising hydrograph in the 2009 spring (Figure 2d). The seasonal variations in stream water $\delta^{18}O$ were captured well at all the four stations by simulations from the tri-objective calibration. The mean contributions of rainfall and snowmelt to annual streamflow estimated by the bi-objective calibration were 62.8% and 10.8%, which were around 1%-7% smaller than those estimated by the tri-objective calibration (Table 4). In contrast, the contribution of glacier melt estimated by the tri-objective calibration (17.1%) was lower than that estimated by the bi-objective calibration (26.4%). Surface runoff which was mainly fed by glacier melt in the YTR showed a larger proportion in the total streamflow simulated by a bi-objective calibration (52.1%) than that in the simulation of a tri-objective calibration (44.7%), while baseflow contribution quantified by the bi-objective calibration is smaller. Standard deviation values of the quantified CRCs indicated that the tri-objective calibration estimated smaller uncertainties for the quantifications of runoff components.

**[Figure 2]**

**[Table 4]**

The uncertainties of behavioral parameter set obtained by bi- and tri-objective calibration are shown in Figure 3. Apart from the hillslope roughness coefficient (*nt*), the uncertainties of all the parameters were reduced by tri-objective calibration, especially the parameters related to melting ($DDF_N$ and $T_o$) and flow concentration processes (*C1* and *C2*). The higher melting temperature threshold ($T_0$) obtained by tri-objective calibration was consistent with the lower contribution of melting water. The lower water storage capacity (*WM*) and higher shape coefficient (*B*) of tri-objective calibration should result in higher saturation area and consequently higher contribution of surface runoff, which was however not in agreement with the estimated CRC, indicating the important contribution of glacier melt in surface runoff. A benchmark parameter set that performed well on multiple objectives was selected among the behavioral parameters of BM_YTR calibrated by tri-objective method (as shown in Table 5), to produce stream water $\delta^{18}O$ for model calibration in experiment 3 in YTR basin. It is noted that this benchmark parameter set was only used to produce stream water $\delta^{18}O$ data for model calibration in experiment 3 in YTR basin, not an optimal parameter set representing the true hydrological processes.

**[Figure 3]**

**[Table 5]**

Figure 4 shows model performances in the KR catchment. The parameter sets producing positive $NSE_{iso}$ were selected as behavioral for tri-objective calibration. Variations of discharge and SCA were reproduced comparably well by the bi- and tri-objective calibrations indicated by the similar metric values. However, the bi-objective calibration produced extremely poor performance for the isotope simulation with low $NSE_{iso}$ and a large simulation error of ~5‰ (Figure 4c). The tri-objective calibration captured the seasonal variations in stream water $\delta^{18}O$ during the study period well. Similarly to YTR, the tri-objective calibration resulted in lower uncertainty in the simulated hydrograph (e.g., early 2010, 2006 and 2008), benefiting from involving isotope for the model calibration to reject parameter sets that produced good performance for discharge and SCA simulations but poor performance for isotope simulation. Regarding the CRCs to total streamflow, the bi-objective and tri-objective calibrations estimated similar results with differences up to 3%. The mean contributions of rainfall, snowmelt and glacier melt to annual streamflow in the KR catchment were around 45%, 22% and 33%, respectively. Contribution of surface runoff estimated by the bi-objective calibration, however, was 13% lower than that estimated by the tri-objective calibration. In contrast, baseflow is more important in the total streamflow simulated by the bi-objective calibration (accounting for 38%) in comparison to the simulation of the tri-objective calibration (accounting for 25%). Again in the KR catchment, uncertainties of CRCs quantified by the tri-objective calibration are much smaller than those estimated by the bi-objective calibration (Table 4).

**[Figure 4]**

### 3.2 Changes in model simulations forced by different assumed glacier meltwater isotopes

Behavioral parameter sets of experiment 1 were selected based on the same $NSE_{iso}$ threshold (0.5) with the benchmark running. Model simulations forced by assumed glacier meltwater $\delta^{18}O$ that are 5‰ (scenario BM_YTR, $\Delta_\delta$=5‰) and 7‰ (scenario G_Δ7, $\Delta_\delta$=7‰) lower than the long-term average precipitation $\delta^{18}O$ showed the best discharge simulations in the validation period (2011-2015) and stations (Yangcun and Nugesha), indicated by the high average metric values (Figure 5a-d). It is noted that simulations of all the glacier meltwater isotope input scenarios in experiment 1 except G_Δ1 performed better than the bi-objective calibration in which isotope data was not involved for parameter identification. Discharge simulation in the scenario of G_Δ1 estimated higher performance in the validation period than the bi-objective calibration (Figure 5a), but lower performance at internal stations (Figure 5b and 5c).

**[Figure 5]**

Figure 5 (e-h) shows the average CRCs and corresponding uncertainties estimated by the different glacier melt isotope inputs. Scenarios with larger $\Delta_{\delta}$ values (i.e., glacier meltwater isotope is much lower than precipitation isotope) tended to result in higher contributions of precipitation and lower contributions of glacier melt. This can be expected, as stream water $\delta^{18}O$ is a mixture mainly from $\delta^{18}O$ of precipitation and glacier meltwater in YTR basin and precipitation $\delta^{18}O$ is fixed in all the scenarios. Result of scenario G_Δ1, however, estimated a smaller contribution of glacier melt than the scenario G_Δ3. This was likely due to that the behavioral parameter sets were selected based on the performance of both discharge and isotope simulations. Parameter sets that estimated higher glacier melt contribution with good performance in isotope simulation but performed poorly on discharge simulation were excluded from the behavioral set in the G_Δ1 scenario.

### 3.3 Changes in model performance forced by isoGSM product merged with different site measurements of precipitation isotope

Figure 6 shows the relationship between REW-scale weighted averages of precipitation $\delta^{18}O$ and the longitude/elevation of corresponding REW for the scenarios in experiment 2. The precipitation $\delta^{18}O$ showed similar spatial pattern in the scenarios merging isoGSM with measurement data at more than one sites. In scenario P_1station that isoGSM was merged with measurement data only at the most downstream station Nuxia, however, spatial pattern was different, showing significantly higher precipitation $\delta^{18}O$ than other scenarios. The different precipitation $\delta^{18}O$ pattern was mainly a result of different altitudinal lapse rates of the isoGSM bias (i.e., parameter *a* in equation 2). Representing the bias characteristic in the whole basin solely by the data measured at the most downstream station resulted in significantly smaller isoGSM bias, and consequently overestimated precipitation $\delta^{18}O$.

Different precipitation $\delta^{18}O$ input data inevitably resulted in different simulations of stream water $\delta^{18}O$ as shown in Figure 7. The $NSE_{iso}$ threshold was set as 0.5 except for scenario P_1station, which produced extremely poor $\delta^{18}O$ simulation due to the high bias in precipitation $\delta^{18}O$ input data (Figure 7d). The other three scenarios all performed well in stream $\delta^{18}O$ simulation (Figure 7a-c), among which scenario P_2stationNL produced highest behavior, followed by P_4station and P_2stationNY.

**[Figure 6]**

**[Figure 7]**

Different precipitation isotope input data also led to different performance in hydrological modeling (Figure 8a-d). While different scenarios produced similar SCA simulations in the validation period (Figure 8d), the performance of discharge simulation significantly differed among the precipitation isotope input scenarios. In scenarios BM_YTR and P_2stationNL, the model performed better than the bi-objective calibration in the validation period (Figure 8a) and stations (Figure 8b and 8c), showing higher average values and smaller ranges of $NSE_{dis}$, which indicated that the model benefitted from involving isotope data for calibration. The model performance forced by scenario P_2stationNY was close to that of the bi-objective calibration, with poorer discharge simulation at internal stations (Figure 8b and 8c). Using precipitation isotope input from the scenario P_1station, however, the model performance was significantly worse than that of the bi-objective calibration. Reasons for the variable model performance forced by the precipitation isotope input scenarios could be: Site measurements of precipitation isotope used in scenarios BM_YTR (using data at four sampling stations) and P_2stationNL (using data at the most downstream sampling station and the most upstream sampling station) tended to provide more informative spatial distribution of precipitation $\delta^{18}O$ in the basin and were the most valuable data for the precipitation isotope data merging; in the scenario of P_1station, on the contrary, the bias of isoGSM product was inadequately corrected by site precipitation isotope measured only at the most downstream station Nuxia, resulting in much errors in the isoGSM product at high altitudes. Although precipitation isotope input data did not influence the simulation of hydrological processes, the calibration process that attempted to match simulated stream $\delta^{18}O$ with measurement influenced the parameter and consequently affected the internal hydrological processes.

**[Figure 8]**

Figure 8 (e-h) shows the average CRCs and corresponding uncertainties estimated by the different precipitation isotope input scenarios. All scenarios produced lower uncertainties than the bi-objective calibration, which can be expected as they were calibrated by a tri-objective approach. The variable precipitation input scenarios resulted in contribution differences of around 10% in runoff components of rainfall, glacier melt and baseflow. The sort of estimated contribution of rainfall (P_2stationNL > BM_YTR > P_2stationNY > P_1station) was opposite to that of average precipitation $\delta^{18}O$ shown in Figure 6, which was consistent with the estimation based on the end-member mixing method.

Among the evaluation metrics, discharge simulation at Nugesha station showed the largest sensitivity to precipitation isotope inputs. As shown in Figure 9, scenarios P_2stationNY and P_1station estimated higher contribution of meltwater, earlier discharge onset timing and higher peak flow. The discharge began to rise especially early (around February) in scenario P_1station, because of the low calibrated value for the melting temperature threshold $T_0$ (-4.5°C), resulted in extremely poor discharge simulation (average NSE is around 0, Figure 9d).

**[Figure 9]**

### 3.4 Model performance constrained by different stream water sampling strategies

Figure 10 shows the accuracy and uncertainty metrics of CRCs produced by experiment 3 in the YTR basin. The $NSE_{iso}$ threshold was set as 0.8, because the stream isotope data for model calibration was generated by a benchmark parameter set, towards which good simulation was rather easy to produce. In comparison to the baseline scenario of RT_TYR_BM, collecting stream isotope data in the dry season (i.e., from November to next February in scenario RT_YTR_WholeYear) brought little benefits to the estimation of water sources proportions, but significantly improved the quantifications of runoff generation pathways indicated by the lower MAE and STD in Figure 10b. The stream water in dry season was fed mainly by groundwater. Stream water isotope data collected in this period reflect the release of groundwater storage, thus helping to constrain the partition between surface and subsurface runoff pathway. On the other hand, reducing the frequency of stream isotope data from weekly to monthly (i.e., scenario RT_YTR_Monthly) led to significantly higher MAE and STD for both the partitions of water sources and runoff pathways, which indicated that stream water isotope data collected by a monthly sampling strategy could provide less constrains to model calibration. Extending the duration of stream isotope sampling period by one or two years (i.e., scenarios RT_YTR_2year and RT_YTR_3year) did not bring much benefits to the quantifications of CRCs regarding the similar metric values. Using stream water isotope data from a three years' sampling (RT_YTR_3year) even led to higher MAE and STD than that using stream water isotope data from a 2 years' sampling (RT_YTR_2year), which might be an occasional result obtained by the random calibration procedure (100 pySOT runs). In comparison to simulations constrained by stream water isotope data from multiple sampling years, results constrained by stream water isotope data from multiple sampling sits (i.e., scenarios of RT_YTR_2station and RT_YTR_4station) yielded lower MAE and STD for the quantified CRCs.

**[Figure 10]**

Model simulations calibrated by spatially distributed stream $\delta^{18}O$ data collected in a one-time field campaign reduced the CRC uncertainty compared to the bi-objective calibration (Figure 10). However, its MAE and STD for the quantifications of CRCs were higher than that estimated by the model when calibrated by weekly sampled time series of stream $\delta^{18}O$. Additionally using stream isotope data from four major tributaries (i.e., scenario RS_YTR_Tributary) brought little benefits to the model performance than using isotope data from the main stream solely (RS_YTR_Main), partly due to the signatures of stream water isotope from tributaries were already reflected by water samples collected at confluences on the main river channel.

In the KR catchment, stream isotope data was collected from five continues years, providing better data basis for the evaluation of the influence of sampling period duration. The $NSE_{iso}$ threshold was set as 0, same with the benchmark scenario in KR catchment. Figure 11 compares the CRC estimations and their uncertainty metric STD of variable scenarios. For the estimate of water sources, the model produced rather large uncertainty ranges of ~20% and ~40% for the contributions of rainfall and glacier melt when calibrating the model using discharge and SCA. Using one-year's stream water isotope data for model calibration, the uncertainty ranges were reduced by rejecting some outliers as shown in Figure 11a-c, but the STD was still large (Figure 11e). The STD can be reduced by increasing the number of calibration isotope data at a rate of ~1%/year. Using isotope data collected from five years, however, didn't result in further decrease in the CRC uncertainties compared to the result calibrated by isotope data collected in a four-year sampling period. The situation, however, was quite different for the estimates of runoff pathways. The bi-objective calibration produced a large uncertainty of ~40% and a STD of ~10% (Figure 11d) for the contribution of baseflow. Using one-year's data for model calibration, the uncertainty range was significantly reduced by about half of that modelled by the bi-objective calibration (from ~10% to ~5%). However, further increase in the duration of sampling period did not bring much improvements on constraining the uncertainties in quantifications of runoff pathways with STD fluctuating around only 4%. It is indicated that model calibration upon more stream isotope data was useful to better constrain the uncertainties of the model simulations and modeled CRCs, but benefit would disappear after a certain duration of stream water sampling period has been reached.

**[Figure 11]**

## 4. Discussions

### 4.1 Implications for water sampling for isotope analysis in high mountains of TP

This study tested the reliance of the benefits of using tracer-aided hydrological model on isotope data availability in two mountainous catchments YTR and KR on the TP. Our findings consistently showed that the model robustness, with respect to performance in the validation period and internal stations and the quantifications of CRCs, can be significantly improved by involving isotope data for parameter calibration, similarly to previous tracer-aided modeling studies (e.g., He et al., 2019; Ala-aho et al., 2017; Birkel et al., 2010). It can be expected that more data help to provide more constrains on identification of model parameters. Nonetheless, water sampling in high mountains on the TP is restricted by environment accessibility, financial and human costs (Stevenson et al., 2021, Li et al., 2020). It is therefore highly needed to find optimal strategies of collecting water samples that balance well between data adequacy for model running and affordable sampling cost (Sprenger et al., 2019).

As an important water source in mountainous catchment on the TP, sampling of glacier meltwater was expected to be favorable for the determination of glacier meltwater isotopic composition and its contribution to total streamflow (He et al., 2019). Field campaign for sampling of glacier melt water is strongly challenging in the YTR basin in this study, due to the harsh accessibility of very high altitudes where glaciers lie. We thus assumed that glacier meltwater $\delta^{18}O$ was lower than the average local precipitation $\delta^{18}O$ by an offset parameter ($\Delta_{\delta}$).

This simple assumption turned to work well on driving the tracer-aided hydrological model and produced better performance than the bi-objective calibration in both validation periods and internal stations. Experiments by using different $\Delta_\delta$ values indicated that the prior assumed isotopic compositions of glacier melt have small influence on the estimated glacier meltwater contribution in the YTR basin. It should be noted that this was different from the results of some hydrograph separation works (e.g., Pu et al., 2020; Lone et al., 2021), which indicated important influence of meltwater isotope composition in estimating the CRC. Those works were based on the end-member mixing approach, which was applied in a short time scale, and was more dependent on the isotope composition of each runoff component. However, this study applied the tracer-aided hydrological model in a longer time scale, where the temporal variability of isotope composition played a more important role than its absolute value, on the parameter calibration. Consequently, when the temporal variabilities of isotope composition of each water source were reproduced properly, the glacier melt $\delta^{18}O$ value in a reasonable range would have little influence on the model performance. The $\Delta_\delta$ values ranging from 2‰-9‰ led to only ~5% difference in the estimated contributions of glacier melt. Using a $\Delta_\delta$ to estimate glacier meltwater $\delta^{18}O$ could serve as an option to force the tracer-aided hydrological models in high-mountain catchments where collecting glacier meltwater samples is highly challenging.

Results of experiment 2 indicated that the original isoGSM precipitation $\delta^{18}O$ data showed large bias in the high mountain basins on TP, and must be corrected by or merged with measurement data before using to force the tracer-aided hydrological model. Our experiments showed that measurement of precipitation isotope at only two sampling sites (scenario P_2stationNL) in the large YTR basin of $2\times10^5$ km$^2$ can be highly valuable for isotope data merging. Forced by isoGSM data that was merged with precipitation $\delta^{18}O$ measurements from two sampling sites, the model performed better than the bi-objective calibration in simulating discharge in the validation period and internal stations, and performed comparably to the simulations of a benchmark running which used precipitation $\delta^{18}O$ measurements from four stations for the data merging. This benefitted from the large altitudinal range covered by the two sampling sites (a most downstream site Nuxia and a most upstream site Lazi) to represent the spatial pattern of isoGSM bias. Likewise using measurement data at two sites in the scenario P_2stationNY, model performance deteriorated visibly, as the sampling sites (Nuxia and Yangchun) were both located in the downstream regions, being worse at representing the spatial pattern of precipitation $\delta^{18}O$ over the basin. Consequently, the strategy of collecting precipitation samples for isotope data merging should be carefully designed; spending high cost on collecting precipitation samples within a small region might be not worth at improving the performance of the tracer-aided hydrological model.

Measurements of stream water $\delta^{18}O$ are essential for the calibration and evaluation of tracer-aided hydrological models. Three kinds of sampling strategies in YTR basin were evaluated in experiment 3: one-time campaign field sampling, continuous sampling at a fixed location for a long period, and continuous sampling at multiple fixed locations during a short period. It is indicated that continuously sampled stream water $\delta^{18}O$ at a fix location is more valuable for aiding hydrological model than that collected by one-time field sampling campaigns at distributed sites. Seasonality of stream water $\delta^{18}O$ referring to the processes of water storage, mixture and transport in the basin can be better captured by continuous time series measurements of $\delta^{18}O$ data (McGuire and McDonnell, 2006). Spatially sampled stream water $\delta^{18}O$ data by one-time field sampling campaigns possibly miss seasonal $\delta^{18}O$ signatures of stream water that were caused by seasonal runoff generation processes (Kendall and Coplen, 2001; Nan et al., 2019), and provide less constrains for the model calibration. Sampling of stream water during dry season (scenario RT_YTR_WholeYear) brought little improvements to the modeling of water source proportions, which is consistent with the findings in Stevenson et al. (2021). High frequent like weekly sampling of stream water in the dry season makes small senses on improving the stream $\delta^{18}O$ data quality, as stream $\delta^{18}O$ in this season has little variations due to small precipitation triggered runoff inputs. Monthly sampling of stream water (RT_YTR_Monthly) turned to be insufficient to capture the strong hydrological variations in the wet season (Birkel and Soulsby, 2015). For large basins like YTR, increasing the number of sampling site for stream water $\delta^{18}O$ is more useful than extending the years of sampling period at fixed sites, as seasonality of $\delta^{18}O$ signatures of water sources should be similar among years in a short study period. Consequently, continuous sampling at multiple locations in a short period like one or two years seems to be the optimal stream sampling strategy for running tracer-aided hydrological model in mountainous basins like YTR on the TP. The value of extending sampling period was more significant in a smaller catchment KR. The uncertainty of CRC estimation kept decreasing until the data series length reached four years and two years, for the aspects of water source and runoff pathway, respectively. This was consistent with the finding by Stevenson et al. (2021) that the benefits from isotope plateaued after a certain year number, which was five for that study.

### 4.2 Uncertainties and limitations

This study used simulated stream $\delta^{18}O$ of a benchmark model running to represent the fully available dataset of stream $\delta^{18}O$ for water sampling in the YTR basin, due to the limited stream water samples. This procedure likely caused the inherent correlation of the stream $\delta^{18}O$ dataset, which made the model easily reproduce the assumed measurements of stream $\delta^{18}O$ and may underestimate the value of stream $\delta^{18}O$ data collected in extended sampling years and sampling sites. Results in this study serve to provide preliminary understandings of the influences of stream water sampling strategy on the model performance. More solid evaluations, however, can be further benefited from using more real field measurements of stream $\delta^{18}O$ in the mountain basins.

Our study tried to look for optimal water sampling strategies to provide isotope input and calibration data for the tracer-aided hydrological model in the YTR basin and KR catchment on the TP. The transferability of our findings to other basins can be partly expected. For example, we can expect that in catchments where precipitation $\delta^{18}O$ and runoff processes show small spatial heterogeneity, collecting water samples at multiple stations would bring few additional benefits for the modeling work than collecting water samples at a sole station. The influence of assumed glacier meltwater would differ with the glacier covered area fraction in the basins. However, situations in catchments with different geographical and climatic characteristics were not evaluated in this study, which is restricted by the fact that high-quality water isotope data in a set of mountain basins on the TP were hardly available currently (Birkel and Soulsby, 2015). The authors suggest tracer-aided modeling researchers to publish their water isotope data to improve the evaluation of the reliance of tracer-aided modeling performance on water sampling strategy (similarly to He et al. 2021; Niinikoski et al., 2016; Yde et al., 2016).

The model performances were evaluated based on the behavioral parameter sets, which were selected by the threshold values of evaluation metrics. The threshold values were determined by looking at the graph comparing simulation and observation values, and artificially judging whether good fitness has been achieved. This process was rather subjective and had inevitable influence on the evaluation result. However, this was a widely used method (e.g., Birkel et al., 2011; Delavau et al., 2017; He et al., 2019), and the threshold values were set at levels achieved by the studies conducted in the same region (e.g., Zhang et al., 2015; Chen et al., 2017).

## 5. Conclusion

The value of water isotope data for aiding hydrological modeling in large mountainous catchments was tested by a set of numerical experiments in the YTR basin. Reliance of the tracer-aided model performance on the availability of input isotope data and evaluation stream water isotope data was extensively investigated in the numerical experiments. Results could provide important guidance for collecting water samples and establishing tracer-aided hydrological model in mountainous regions on the TP. Our main finds are as follows:

1. In high-mountain basins where glacier meltwater samples for isotope analysis are not available, estimating isotopic composition of glacier meltwater by an offset parameter from precipitation isotope is a feasible way to force the tracer-aided hydrological model. Our test indicated that using a set of glacier meltwater $\delta^{18}O$ that are 2‰~9‰ lower than the mean precipitation $\delta^{18}O$, resulted in small changes in the model performance and the quantifications of CRCs (smaller than 5%) in the YTR basin. This influence, however, is expected to change with the glacier area coverages in other mountain basins.

2. Strategy of field sampling for precipitation to collect measurement precipitation $\delta^{18}O$ merged with isoGSM product should be carefully designed. Collecting precipitation samples at sites from the same altitude tends to be worse at representing the spatial pattern of precipitation $\delta^{18}O$ over the basin than collecting precipitation samples from sites covering a range of altitudes. Measurements of precipitation isotope at only two sampling sites covering an elevation range of 2900-6900m in the large YTR basin of $2\times10^5$ $km^2$ can be highly valuable for precipitation isotope data merging.

3. Colleting weekly stream water samples at multiple sites in the wet and warm seasons is the optimal strategy to capture more hydrological process variability for calibrating and evaluating a tracer-aided hydrological model in the YTR basin. It is highly recommended to increase the number of stream water sampling sites in the high-mountain basins rather than extending the duration of sampling period at a sole site. Benefits from extending the duration of sampling period is more visible in a small catchment but smaller in large basins, and tend to disappear when a certain duration of sampling period has been reached.

## Code and data availability

Code and data availability. The isotope data and the code of THREW-T model used in this study are available from the corresponding author (tianfq@tsinghua.edu.cn). Other data sets and the calibration program pySOT are publicly available as follows: DEM (http://www.gscloud.cn/sources/details/310?pid=302, last access: 1 January 2019, Geospatial Data Cloud Site, 2019), CMFD (https://doi.org/10.11888/AtmosphericPhysics.tpe.249369.file, Yang and He, 2019), glacier data (https://doi.org/10.3972/glacier.001.2013.db, Liu et al., 2012), NDVI (https://doi.org/10.5067/MODIS/MOD13A3.006, Didan et al., 2015), LAI (https://doi.org/10.5067/MODIS/MOD15A2H.006, Myneni et al., 2015), HWSD (https://data.tpdc.ac.cn/zh-hans/data/3519536a-d1e7-4ba1-8481-6a0b56637baf/?q=HWSD, last access: 1 January 2019, He, 2019) and the pySOT program (https://doi.org/10.5281/zenodo.569554, Eriksson et al., 2017). These data sets and programs are also referred to in the main text (Yang et al., 2010; Chen et al., 2018).

## Author contribution

YN, ZH and FT conceived the idea; ZW provided the isoGSM data; LT provided the measurement isotope data; YN, ZH and FT conducted analysis; ZW and LT provided comments on the analysis; all the authors contributed to writing and revisions.

## Financial support

This study has been supported by the National Natural Science Foundation of China (grant no. 92047301 and 51879136) and the Shuimu Tsinghua Scholar Program.

## Competing interests

At least one of the (co-)authors is a member of the editorial board of Hydrology and Earth System Sciences.

## References


Ala-aho, P., Tetzlaff, D., McNamara, J. P., Laudon, H., and Soulsby, C.: Using isotopes to constrain water flux and age estimates in snow-influenced catchments using the STARR (Spatially distributed Tracer-Aided Rainfall–Runoff) model, Hydrology and Earth System Sciences, 21, 5089-5110, 10.5194/hess-21-5089-2017, 2017.

Birkel, C., Dunn, S. M., Tetzlaff, D., and Soulsby, C.: Assessing the value of high-resolution isotope tracer data in the stepwise development of a lumped conceptual rainfall-runoff model, Hydrological Processes, 24, 2335-2348, 10.1002/hyp.7763, 2010.

Birkel, C., Tetzlaff, D., Dunn, S. M., and Soulsby, C.: Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models, Water Resources Research, 47, 10.1029/2010wr009547, 2011.

Birkel, C., and Soulsby, C.: Advancing tracer-aided rainfall-runoff modelling: a review of progress, problems and22nrealizedd potential, Hydrological Processes, 29, 5227-5240, 10.1002/hyp.10594, 2015.

Bloeschl, G., and Montanari, A.: Climate change impact: throwing the dice?, Hydrological Processes, n/a-n/a, 10.1002/hyp.7574, 2009.

Boral, S., and Sen, I. S.: Tracing 'Third Pole' ice meltwater contribution to the Himalayan rivers using oxygen and hydrogen isotopes, Geochemical Perspectives Letters, 48-53, 10.7185/geochemlet.2013, 2020.

Bowen, G. J., Cai, Z., Fiorella, R. P., and Putman, A. L.: Isotopes in the Water Cycle: Regional- to Global-Scale Patterns and Applications, in: Annual Review Of Earth And Planetary Sciences, Vol 47, edited by: Jeanloz, R., and Freeman, K. H., Annual Review of Earth and Planetary Sciences, 453-+, 2019.

Capell, R., Tetzlaff, D., and Soulsby, C.: Can time domain and source area tracers reduce uncertainty in rainfall - runoff models in larger heterogeneous catchments?, Water Resources Research, 48, 10.1029/2011wr011543, 2012.

Chen, X., Long, D., Hong, Y., Zeng, C., and Yan, D.: Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the Upper Brahmaputra River basin?, Water Resources Research, 53, 2431-2466, 10.1002/2016wr019656, 2017.

Chen, X., Long, D., Liang, S., He, L., Zeng, C., Hao, X., and Hong, Y.: Developing a composite daily snow cover extent record over the Tibetan Plateau from 1981 to 2016 using multisource data, Remote Sen. Environ., 215, 284–299, https://doi.org/10.1016/j.rse.2018.06.021, 2018.

Delavau, C. J., Stadnyk, T., and Holmes, T.: Examining the impacts of precipitation isotope input on distributed, tracer-aided hydrological modelling, Hydrology and Earth System Sciences, 21, 2595-2614, 10.5194/hess-21-2595-2017, 2017.

Didan, K.: MOD13A3 MODIS/Terra vegetation Indices Monthly L3 Global 1km SIN Grid V006, NASA EOSDIS Land Processes DAAC [data set], https://doi.org/10.5067/MODIS/MOD13A3.006, 2015.

Dong, G., Weng, B., Chen, J., Yan, D., and Wang, H.: Variation characteristics of stable isotopes in water along main stream of Naqu River in source area of Nujiang River, Water Resources and Hydropower Engineering, 49, 108-114, 2018.

Duethmann, D., Bolch, T., Farinotti, D., Kriegel, D., Vorogushyn, S., Merz, B., Pieczonka, T., Jiang, T., Su, B., and Guentner, A.: Attribution of streamflow trends in snow and glacier melt-dominated catchments of the Tarim River, Central Asia, Water Resources Research, 51, 4727-4750, 10.1002/2014wr016716, 2015.

Dunn, S. M., McDonnell, J. J., and Vaché, K. B.: Factors influencing the residence time of catchment waters: A virtual experiment approach, Water Resources Research, 43, 10.1029/2006wr005393, 2007.

Eriksson, D., Bindel, D., and Shoemaker, C.: Dme65/Pysot: V0.1.35, Zenodo [code], https://doi.org/10.5281/zenodo.569554, 2017.

Gao J., Tian L., and Liu, Y.: Oxygen isotope variation in the water cycle of the Yamzho Lake Basin in southern Tibetan Plateau, Chinese Sci. Bull., 54, 2758–2765, 2009.

Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, Hydrological Processes, 22, 3802-3813, 10.1002/hyp.6989, 2008.

He, Y.: Pan-TPE soil map based on Harmonized World Soil Database (V1.2), National Tibetan Plateau Data Center [data set], https://data.tpdc.ac.cn/zh-hans/data/3519536a-d1e7-4ba1-8481-6a0b56637baf/?q=HWSD, 2019

He, Z. H., Tian, F. Q., Gupta, H. V., Hu, H. C., and Hu, H. P.: Diagnostic calibration of a hydrological model in a mountain area by hydrograph partitioning, Hydrology and Earth System Sciences, 19, 1807-1826, 10.5194/hess-19-1807-2015, 2015.

He, Z., Unger-Shayesteh, K., Vorogushyn, S., Weise, S. M., Kalashnikova, O., Gafurov, A., Duethmann, D., Barandun, M., and Merz, B.: Constraining hydrological model parameters using water isotopic compositions in a glacierized basin, Central Asia, Journal of Hydrology, 571, 332-348, 10.1016/j.jhydrol.2019.01.048, 2019.

He, Z., Unger-Shayesteh, K., Vorogushyn, S., Weise, S. M., Duethmann, D., Kalashnikova, O., Gafurov, A., and Merz, B.: Comparing Bayesian and traditional end-member mixing approaches for hydrograph separation in a glacierized basin, Hydrology and Earth System Sciences, 24, 3289-3309, 10.5194/hess-24-3289-2020, 2020.

He, Z., Duethmann, D., and Tian, F.: A meta-analysis based review of quantifying the contributions of runoff components to streamflow in glacierized basins, Journal of Hydrology, 603, 126890, 10.1016/j.jhydrol.2021.126890, 2021.

Hindshaw, R. S., Tipper, E. T., Reynolds, B. C., Lemarchand, E., Wiederhold, J. G., Magnusson, J., Bernasconi, S. M., Kretzschmar, R., and Bourdon, B.: Hydrological control of stream water chemistry in a glacial catchment (Damma Glacier, Switzerland), Chemical Geology, 285, 215-230, 10.1016/j.chemgeo.2011.04.012, 2011.

Immerzeel, W. W., van Beek, L. P. H., and Bierkens, M. F. P.: Climate Change Will Affect the Asian Water Towers, Science, 328, 1382-1385, 10.1126/science.1183188, 2010.

Jeelani, G., Shah, R. A., Jacob, N., and Deshpande, R. D.: Estimation of snow and glacier melt contribution to Liddar stream in a mountainous catchment, western Himalaya: an isotopic approach, Isotopes in environmental and health studies, 53, 18-35, 10.1080/10256016.2016.1186671, 2017.

Kendall, C., and Coplen, T. B.: Distribution of oxygen-18 and deuterium in river waters across the United States, Hydrological Processes, 15, 1363-1393, 10.1002/hyp.217, 2001.

Klaus, J., and McDonnell, J. J.: Hydrograph separation using stable isotopes: Review and evaluation, Journal of Hydrology, 505, 47-64, 10.1016/j.jhydrol.2013.09.006, 2013.

Knapp, J. L. A., Neal, C., Schlumpf, A., Neal, M., and Kirchner, J. W.: New water fractions and transit time distributions at Plynlimon, Wales, estimated from stable water isotopes in precipitation and streamflow, Hydrology and Earth System Sciences, 23, 4367-4388, 10.5194/hess-23-4367-2019, 2019.

Kong, Y., Wang, K., Pu, T., and Shi, X.: Nonmonsoon Precipitation Dominates Groundwater Recharge Beneath a Monsoon-Affected Glacier in Tibetan Plateau, Journal of Geophysical Research: Atmospheres, 124, 10913-10930, 10.1029/2019jd030492, 2019.

Laudon, H., Taberman, I., Ågren, A., Futter, M., Ottosson-Löfvenius, M., and Bishop, K.: The Krycklan Catchment Study-A flagship infrastructure for hydrology, biogeochemistry, and climate research in the boreal landscape, Water Resources Research, 49, 7154-7158, 10.1002/wrcr.20520, 2013.

Li, Z.-J., Li, Z.-X., Song, L.-L., Gui, J., Xue, J., Zhang, B. J., and Gao, W. D.: Hydrological and runoff formation processes based on isotope tracing during ablation period in the source regions of Yangtze River, Hydrology and Earth System Sciences, 24, 4169-4187, 10.5194/hess-24-4169-2020, 2020.

Li, Z., Feng, Q., Li, Z., Yuan, R., Gui, J., and Lv, Y.: Climate background, fact and hydrological effect of multiphase water transformation in cold regions of the western china: a review, EARTH SCIENCE REVIEWS, 190, 33-57, https://doi.org/10.1016/j.earscirev.2018.12.004, 2019.

Liu, S.: The second glacier inventory dataset of China (version 1.0) (2006–2011), National Tibetan Plateau Data Center [data set], https://doi.org/10.3972/glacier.001.2013.db, 2012.

Liu, Z., Tian, L., Yao, T., Gong, T., Yin, C., and Yu, W.: Temporal and spatial variations of delta O-18 in precipitation of the Yarlung Zangbo River Basin, J. Geogr. Sci., 17, 317–326, https://doi.org/10.1007/s11442-007-0317-1, 2007.

Lutz, A. F., Immerzeel, W. W., Shrestha, A. B., and Bierkens, M. F. P.: Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation, Nature Climate Change, 4, 587-592, 10.1038/nclimate2237, 2014.

McGuire, K. J., and McDonnell, J. J.: A review and evaluation of catchment transit time modeling, Journal of Hydrology, 330, 543-563, 10.1016/j.jhydrol.2006.04.020, 2006.

McGuire, K. J., Weiler, M., and McDonnell, J. J.: Integrating tracer experiments with modeling to assess runoff processes and water transit times, Advances in Water Resources, 30, 824-837, 10.1016/j.advwatres.2006.07.004, 2007.

Myneni, R., Knyazikhin, Y., and Park, T.: MOD15A2H MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006, NASA EOSDIS Land Processes DAAC [data set], https://doi.org/10.5067/MODIS/MOD15A2H.006, 2015.

Nan, Y., Tian, F., Hu, H., Wang, L., and Zhao, S.: Stable Isotope Composition of River Waters across the World, Water, 11, 1760, 10.3390/w11091760, 2019.

Nan, Y., He, Z., Tian, F., Wei, Z., and Tian, L.: Can we use precipitation isotope outputs of isotopic general circulation models to improve hydrological modeling in large mountainous catchments on the Tibetan Plateau?, Hydrology and Earth System Sciences, 25, 6151-6172, 10.5194/hess-25-6151-2021, 2021b.

Nan, Y., Tian, L., He, Z., Tian, F., and Shao, L.: The value of water isotope data on improving process understanding in a glacierized catchment on the Tibetan Plateau, Hydrology and Earth System Sciences, 25, 3653-3673, 10.5194/hess-25-3653-2021, 2021a.

Niinikoski, P. I. A., Hendriksson, N. M., and Karhu, J. A.: Using stable isotopes to resolve transit times and travel routes of river water: a case study from southern Finland, Isotopes in environmental and health studies, 52, 380-392, 10.1080/10256016.2015.1107553, 2016.

Ohlanders, N., Rodriguez, M., and McPhee, J.: Stable water isotope variation in a Central Andean watershed dominated by glacier and snowmelt, Hydrology and Earth System Sciences, 17, 1035-1050, 10.5194/hess-17-1035-2013, 2013.

Pomeroy, J. W., Gray, D. M., Brown, T., Hedstrom, N. R., Quinton, W. L., Granger, R. J., and Carey, S. K.: The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence, Hydrological Processes, 21, 2650-2667, 10.1002/hyp.6787, 2007.

Rai, S. P., Singh, D., Jacob, N., Rawat, Y. S., Arora, M., and BhishmKumar: Identifying contribution of snowmelt and glacier melt to the Bhagirathi River (Upper Ganga) near snout of the Gangotri Glacier using environmental isotopes, Catena, 173, 339-351, 10.1016/j.catena.2018.10.031, 2019.

Reggiani, P., Hassanizadeh, S. M., Sivapalan, M., and Gray, W. G.: A unifying framework for watershed thermodynamics: constitutive relationships, Advances In Water Resources, 23, 15-39, 10.1016/s0309-1708(99)00005-6, 1999.

Son, K., and Sivapalan, M.: Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, Water Resources Research, 43, 10.1029/2006wr005032, 2007.

Sprenger, M., Stumpp, C., Weiler, M., Aeschbach, W., Allen, S. T., Benettin, P., Dubbert, M., Hartmann, A., Hrachowitz, M., Kirchner, J. W., McDonnell, J. J., Orlowski, N., Penna, D., Pfahl, S., Rinderer, M., Rodriguez, N., Schmidt, M., and Werner, C.: The Demographics of Water: A Review of Water Ages in the Critical Zone, Reviews Of Geophysics, 57, 800-834, 10.1029/2018rg000633, 2019.

Stevenson, J. L., Birkel, C., Neill, A. J., Tetzlaff, D., and Soulsby, C.: Effects of streamflow isotope sampling strategies on the calibration of a tracer‐aided rainfall‐runoff model, Hydrological Processes, 35, 10.1002/hyp.14223, 2021.

Tan, H., Chen, X., Shi, D., Rao, W., Liu, J., Liu, J., Eastoe, C. J., and Wang, J.: Base flow in the Yarlungzangbo River, Tibet, maintained by the isotopically-depleted precipitation and groundwater discharge, The Science of the total environment, 759, 143510, 10.1016/j.scitotenv.2020.143510, 2021.

Tetzlaff, D., Birkel, C., Dick, J., Geris, J., and Soulsby, C.: Storage dynamics in hydropedological units control hillslope connectivity, runoff generation, and the evolution of catchment transit time distributions, Water Resour Res, 50, 969-985, 10.1002/2013WR014147, 2014.

Tian, F., Hu, H., Lei, Z., and Sivapalan, M.: Extension of the Representative Elementary Watershed approach for cold regions via explicit treatment of energy related processes, Hydrology And Earth System Sciences, 10, 619-644, 10.5194/hess-10-619-2006, 2006.

Tian, F., Xu, R., Nan, Y., Li, K., and He, Z.: Quantification of runoff components in the Yarlung Tsangpo River using a distributed hydrological model, Advances in Water Science, 31, 324-336, 2020.

Tong, R., Parajka, J., Salentinig, A., Pfeil, I., Komma, J., Széles, B., Kubáň, M., Valent, P., Vreugdenhil, M., Wagner, W., and Blöschl, G.: The value of ASCAT soil moisture and MODIS snow cover data for calibrating a conceptual hydrologic model, Hydrology and Earth System Sciences, 25, 1389-1410, 10.5194/hess-25-1389-2021, 2021.

Viviroli, D., Weingartner, R., and Messerli, B.: Assessing the hydrological significance of the world's mountains, Mountain Research And Development, 23, 32-40, 10.1659/0276-

4741(2003)023[0032:athsot]2.0.co;2, 2003.

Wang, C., Dong, Z., Qin, X., Zhang, J., Du, W., and Wu, J.: Glacier meltwater runoff process analysis using δD and δ18O isotope and chemistry at the remote Laohugou glacier basin in western Qilian Mountains, China, Journal of Geographical Sciences, 26, 722-734, 10.1007/s11442-016-1295-y, 2016.

Wang, Y., Wang, L., Zhou, J., Yao, T., Yang, W., Zhong, X., Liu, R., Hu, Z., Luo, L., Ye, Q., Chen, N., and Ding, H.: Vanishing Glaciers at Southeast Tibetan Plateau Have Not Offset the Declining Runoff at Yarlung Zangbo, Geophysical Research Letters, 48, 10.1029/2021gl094651, 2021.

Wolfe, B. B., Karst-Riddoch, T. L., Hall, R. I., Edwards, T. W. D., English, M. C., Palmini, R., McGowan, S., Leavitt, P. R., and Vardy, S. R.: Classification of hydrological regimes of northern floodplain basins (Peace–Athabasca Delta, Canada) from analysis of stable isotopes (δ18O, δ2H) and water chemistry, Hydrological Processes, 21, 151-168, 10.1002/hyp.6229, 2007.

Xi, X.: A Review of Water Isotopes in Atmospheric General Circulation Models: Recent Advances and Future Prospects, International Journal of Atmospheric Sciences, 2014, 1-16, 10.1155/2014/250920, 2014.

Xia, X., Li, S., Wang, F., Zhang, S., Fang, Y., Li, J., Michalski, G., and Zhang, L.: Triple oxygen isotopic evidence for atmospheric nitrate and its application in source identification for river systems in the Qinghai-Tibetan Plateau, The Science of the total environment, 688, 270-280, 10.1016/j.scitotenv.2019.06.204, 2019.

Xu, R., Hu, H., Tian, F., Li, C., and Khan, M. Y. A.: Projected climate change impacts on future streamflow of the Yarlung Tsangpo-Brahmaputra River, Global and Planetary Change, 175, 144-159, 10.1016/j.gloplacha.2019.01.012, 2019.

Yang, K. and He, J.: China meteorological forcing dataset (1979–2018), National Tibetan Plateau Data Center [data set], https://doi.org/10.11888/AtmosphericPhysics.tpe.249369.file, 2019.

Yao, T., Masson-Delmotte, V., Gao, J., Yu, W., Yang, X., Risi, C., Sturm, C., Werner, M., Zhao, H., He, Y., Ren, W., Tian, L., Shi, C., and Hou, S.: A review of climatic controls on δ18O in precipitation over the Tibetan Plateau: Observations and simulations, Reviews of Geophysics, 51, 525-548, 10.1002/rog.20023, 2013.

Yde, J. C., Knudsen, N. T., Steffensen, J. P., Carrivick, J. L., Hasholt, B., Ingeman-Nielsen, T., Kronborg, C., Larsen, N. K., Mernild, S. H., Oerter, H., Roberts, D. H., and Russell, A. J.: Stable oxygen isotope variability in two contrasting glacier river catchments in Greenland, Hydrology and Earth System Sciences, 20, 1197-1210, 10.5194/hess-20-1197-2016, 2016.

Yong, B., Wang, C.-Y., Chen, J., Chen, J., Barry, D. A., Wang, T., and Li, L.: Missing water from the Qiangtang Basin on the Tibetan Plateau, Geology, 49, 867-872, 10.1130/g48561.1, 2021.

Yoshimura, K., Kanamitsu, M., Noone, D., and Oki, T.: Historical isotope simulation using Reanalysis atmospheric data, Journal of Geophysical Research, 113, 10.1029/2008jd010074, 2008.

Zhang, F., Zhang, H. B., Hagen, S. C., Ye, M., Wang, D. B., Gui, D. W., Zeng, C., Tian, L. D., and Liu, J. S.: Snow cover and runoff modelling in a high mountain catchment with scarce data: effects of temperature and precipitation parameters, Hydrol. Process., 29, 52–65, https://doi.org/10.1002/hyp.10125, 2015.

Zhang, Z., Chen, X., Cheng, Q., and Soulsby, C.: Storage dynamics, hydrological connectivity and flux ages in a karst catchment: conceptual modelling using stable isotopes, Hydrology and Earth System Sciences, 23, 51-71, 10.5194/hess-23-51-2019, 2019.

## List of Tables

**Table 1.** Summary of precipitation and stream water samples in the YTR and KR catchments.

| Catchment (Station) | Year | Sampling period | Precipitation | | | Stream | | |
|---|---|---|---|---|---|---|---|---|
| | | | Sample number | $\overline{\delta^{18}O}$ (‰) | Std (‰) | Sample number | $\overline{\delta^{18}O}$ (‰) | Std (‰) |
| YTR (Nuxia) | 2005 | 14/Mar to 23/Oct | 86 | -10.33 | 7.18 | 34 | -15.74 | 1.60 |
| YTR (Yangcun) | | 17/Mar to 05/ Oct | 59 | -13.17 | 7.10 | 30 | -16.57 | 1.69 |
| YTR (Nugesha) | | 14/Mar to 22/ Oct | 45 | -14.29 | 7.99 | 25 | -17.84 | 0.99 |
| YTR (Lazi) | | 06/ Jun to 22/Sep | 42 | -17.41 | 5.75 | 22 | -16.52 | 1.43 |
| KR (Wengguo) | 2006 | 06/Apr to 11/Nov | 24 | -15.22 | 3.83 | 31 | -17.35 | 1.68 |
| | 2007 | 23/Apr to 09/ Oct | 39 | -16.99 | 5.93 | 25 | -17.30 | 1.01 |
| | 2010 | 05/May to 18/ Oct | 63 | -19.25 | 5.03 | 23 | -17.44 | 1.29 |
| | 2011 | 28/Mar to 06/Nov | 69 | -13.99 | 5.90 | 32 | -17.11 | 1.30 |
| | 2012 | 16/ Jun to 22/ Sep | 42 | -13.88 | 6.21 | 14 | -17.01 | 0.60 |

**Table 2.** Calibrated parameters of the THREW-T model

| Symbol | Unit | Physical descriptions | Value range |
|---|---|---|---|
| *nt* | - | Manning roughness coefficient for hillslope | 0-0.2 |
| *WM* | cm | Tension water storage capacity, used in Xinanjiang model to calculate saturation area | 0-10 |
| *B* | - | Shape coefficient used in Xinanjiang model to calculate saturation area | 0-1 |
| *KKA* | - | Coefficient to calculate subsurface runoff in $Rg=KKD \cdot S \cdot K^S{}_S \cdot (y_S/Z)^{KKA}$, where $S$ is the topographic slope, $K^S{}_S$ is the saturated hydraulic conductivity, $y_s$ is the depth of saturated groundwater, $Z$ is the total soil depth | 0-6 |
| *KKD* | - | See description for *KKA* | 0-0.5 |
| $T_0$ | °C | Temperature threshold above which snow and glacier melt | -5-5 |
| $DDF_N$ | mm/°C/day | Degree day factor for snowmelt | 0-10 |
| $DDF_G$ | mm/°C/day | Degree day factor for glacier melt | 0-10 |
| *C1* | - | Coefficient to calculate the runoff concentration process using Muskingum method: $O_2=C_1 \cdot I_1+C_2 \cdot I_2+C_3 \cdot O_1+C_4 \cdot Q_{lat}$, where $I_1$ and $O_1$ is the inflow and outflow at prior step, $I_2$ and $O_2$ is the inflow and outflow at current step, $Q_{lat}$ is lateral flow of the river channel, $C_3=1-C_1-C_2$, $C_4=C_1+C_2$ | 0-1 |
| *C2* | - | See description for *C1* | 0-1 |

**Table 3.** Descriptions of water sampling scenarios in the three numerical experiments. $\delta^{18}O_{GM}$ is the assumed glacier meltwater isotope signature and $\overline{\delta^{18}O_{PR}}$ refers to the long term mean isotope signature of precipitation.

| Experiment | Scenarios | Isotope data conditions |
|---|---|---|
| Benchmark model running in the YTR basin | BM_YTR | Using assumed glacier meltwater isotope as: $\delta^{18}O_{GM}=\overline{\delta^{18}O_{PR}}$ - 5‰<br>Using IsoGSM outputs that were merged with sample measurements of precipitation isotope from four sampling sites<br>Using all available stream water samples in the study period to calibrate the model |
| Benchmark model running in the KR catchment | BM_KR | Using all available stream water samples in the study period to calibrate the model |
| Experiment 1: Estimate of glacier meltwater isotope | G_Δ1 | Assuming glacier meltwater isotope as: $\delta^{18}O_{GM}=\overline{\delta^{18}O_{PR}}$ - 1‰ |
| | G_Δ3 | Assuming glacier meltwater isotope as: $\delta^{18}O_{GM}=\overline{\delta^{18}O_{PR}}$ - 3‰ |
| | G_Δ7 | Assuming glacier meltwater isotope as: $\delta^{18}O_{GM}=\overline{\delta^{18}O_{PR}}$ - 7‰ |
| | G_Δ9 | Assuming glacier meltwater isotope as: $\delta^{18}O_{GM}=\overline{\delta^{18}O_{PR}}$ - 9‰ |
| Experiment 2: Site sampling data of precipitation isotope | P_1station | Using IsoGSM outputs merged with measurements of precipitation isotope collected at one station (Nuxia) in YTR |
| | P_2stationNY | Using IsoGSM outputs merged with measurements of precipitation isotope collected at two stations (Nuxia and Yangcun) in YTR |
| | P_2stationNL | Using IsoGSM outputs merged with measurements of precipitation isotope collected at two stations (Nuxia and Lazi) in YTR |
| Experiment 3: Stream water sampling strategy for model evaluation | RT_YTR_BM | Sampling strategy: time series sampling; Sampling timing: wet season; Sampling frequency: weekly; Duration of sampling period: 1 year (2005): Number of sampling site: 1 station (Nuxia) |
| | RT_YTR_WholeYear | Same to RT_YTR_BM, but with the sampling timing as the whole study years |
| | RT_YTR_Monthly | Same to RT_YTR_BM, but with the sampling frequency as monthly |
| | RT_YTR_2year | Same to RT_YTR_BM, but with the duration of sampling period as only 2 years (2005 and 2006) |
| | RT_YTR_3year | Same to RT_YTR_BM, but with the duration of sampling period as only 3 years (2005-2007) |
| | RT_YTR_2station | Same to RT_YTR_BM, but with the number of sampling site as 2 stations (Nuxia and Yangcun) |
| | RT_YTR_4station | Same to RT_YTR_BM, but with the number of sampling site as 4 stations (Nuxia, Yangcun, Nugesha and Lazi) |
| | RS_YTR_Main | Sampling strategy: spatially distributed sampling in a single field campaign; Location of sampling site: along the main stream |
| | RS_YTR_Tributary | Same to RS_YTR_Main, but using stream water samples from additional sites along the tributaries |
| | RT_KR_1year | Sampling strategy: time series sampling; Duration of sampling period: 1 year (2006) |
| | RT_KR_2year | Same to RT_KR_1year, but with the duration of sampling period as 2 years (2006 and 2007) |
| | RT_KR_3year | Same to RT_KR_1year, but with the duration of sampling period as 3 years (2006-2007, 2010) |
| | RT_KR_4year | Same to RT_KR_1year, but with the duration of sampling period as 4 years (2006-2007, 2010-2011) |
| | RT_KR_5year | Same to RT_KR_1year, but with the duration of sampling period as 5 years (2006-2007, 2010-2012) |

**Table 4.** Contributions (%) of runoff components in the YTR basin and KR catchment estimated by different calibration variants in the benchmark scenario.

| Runoff Component | YTR basin | | KR catchment | |
|---|---|---|---|---|
| | Bi-objective calibration[*] | Tri-objective calibration | Bi-objective calibration | Tri-objective calibration |
| Rainfall | 62.8 (±6.5) | 70.7 (±2.5) | 46.4 (±5.0) | 43.9 (±1.4) |
| Snowmelt | 10.8 (±1.1) | 12.2 (±0.4) | 22.6 (±2.4) | 21.4 (±0.7) |
| Glacier melt | 26.4 (±7.5) | 17.1 (±2.9) | 31.0 (±7.4) | 34.6 (±2.0) |
| Surface runoff | 52.1 (±10.5) | 44.7 (±6.7) | 62.0 (±10.9) | 75.1 (±3.3) |
| Subsurface runoff | 47.9 (±10.5) | 55.3 (±6.7) | 38.0 (±10.5) | 24.9 (±3.3) |

*: Values in brackets refer to the standard deviation of the contribution of runoff component produced by the behavioral parameter sets.

**Table 5**. Benchmark parameter set and corresponding model behavior that are used to produce stream water $\delta^{18}O$ data for model calibration in experiment 3 in YTR basin.

| Parameter value | | Model behavior | |
|---|---|---|---|
| *nt* | 0.09 | $NSE_{dis}$ (Nuxia,calibration) | 0.87 |
| *WM* | 0.92 | $NSE_{dis}$ (Nuxia,validation) | 0.80 |
| *B* | 0.62 | $RMSE_{SCA}$ (calibration) | 0.08 |
| *KKA* | 3.22 | $RMSE_{SCA}$ (validation) | 0.12 |
| *KKD* | 0.14 | $NSE_{iso}$ | 0.58 |
| $T_0$ | 1.59 | $NSE_{dis}$ (Yangcun) | 0.85 |
| $DDF_N$ | 8.04 | $NSE_{dis}$ (Nugesha) | 0.76 |
| $DDF_G$ | 8.28 | Contribution of rainfall | 70% |
| *C1* | 0.0004 | Contribution of snowmelt | 12% |
| *C2* | 0.075 | Contribution of glacier melt | 18% |
| | | Contribution of baseflow | 56% |

## List of Figures

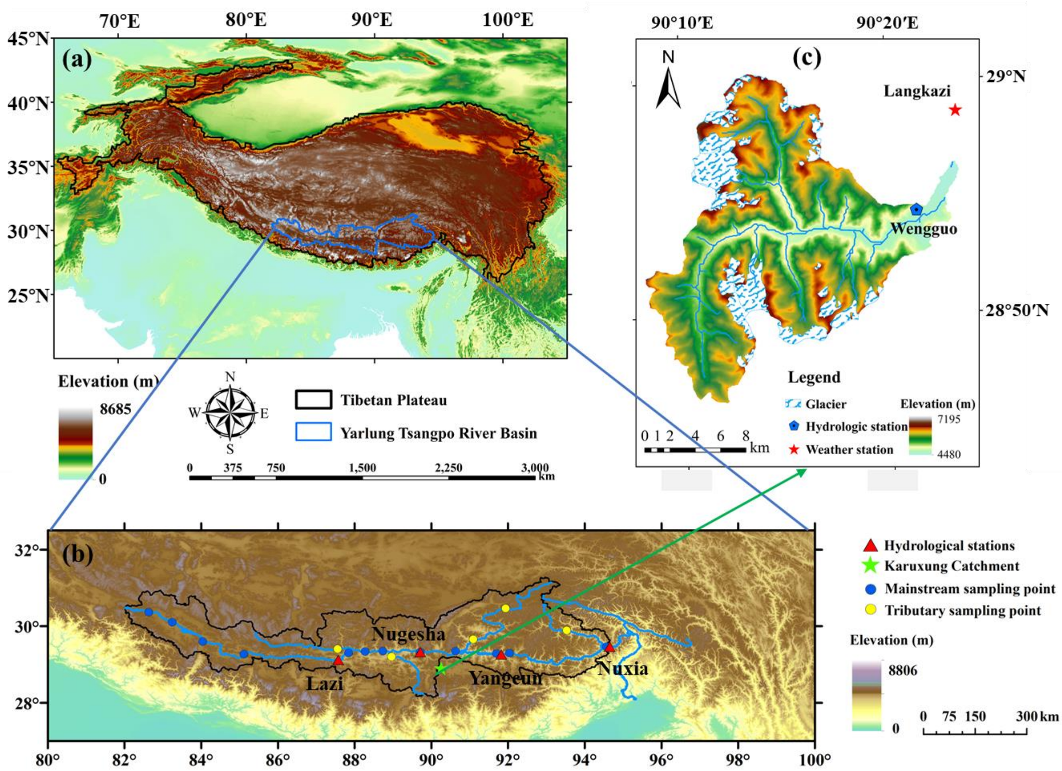


**Figure 1.** Locations and topography of the (a) Tibetan Plateau, (b) Yarlung Tsangpo river basin and (c) Karuxung catchment. Triangles in figure b refer to hydrometric stations and sampling sites for precipitation and stream water isotope. Dots in figure b refer to assumed stream water sampling locations in RD_YTR scenarios.

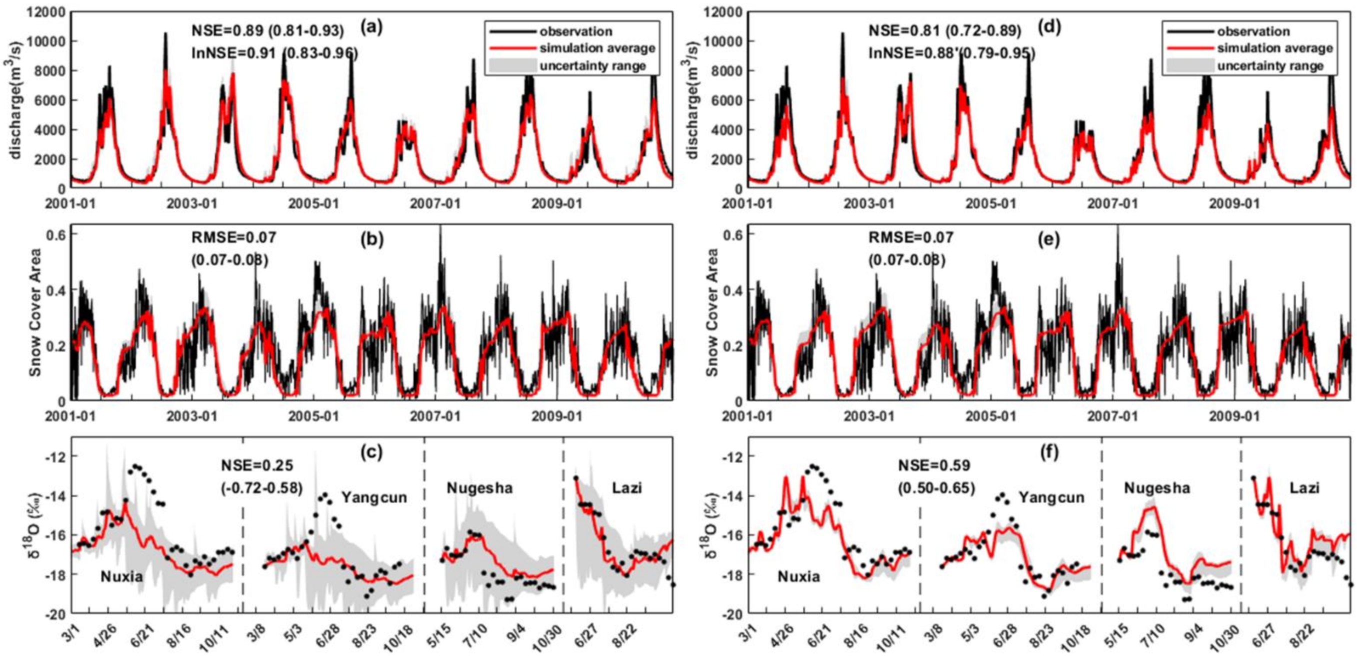


**Figure 2.** Uncertainty ranges and metrics values of the simulated discharge (Nuxia station), SCA, and stream δ$^{18}$O (at four stations during 2005) in the YTR basin, that were produced by the behavioral parameter sets of a bi-objective calibration (a-c) and a tri-objective (d-f) calibration in the benchmark model running.

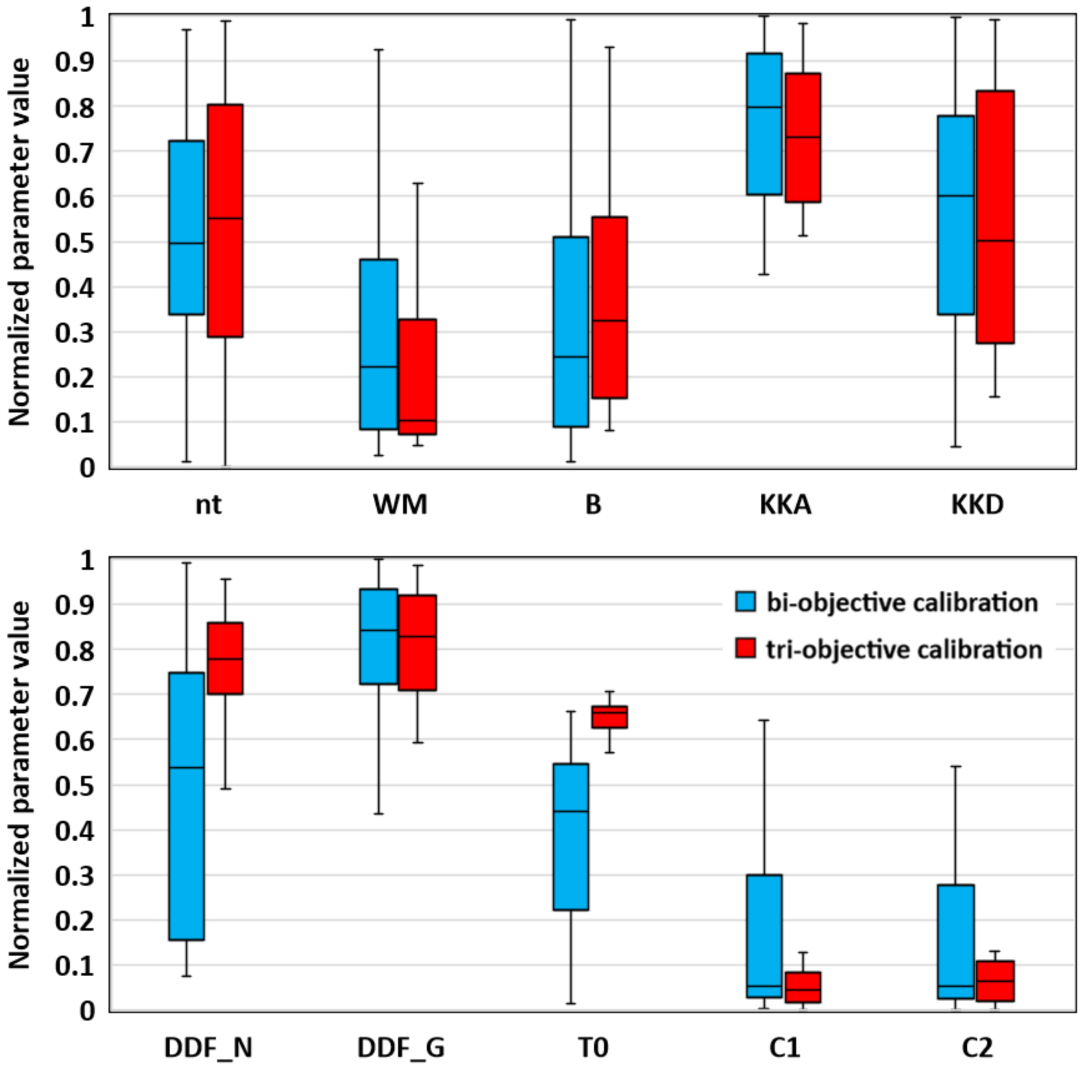


**Figure 3.** Uncertainties of behavioral parameter set obtained by bi- and tri-objective calibration methods for BM_YTR scenario in YTR basin.

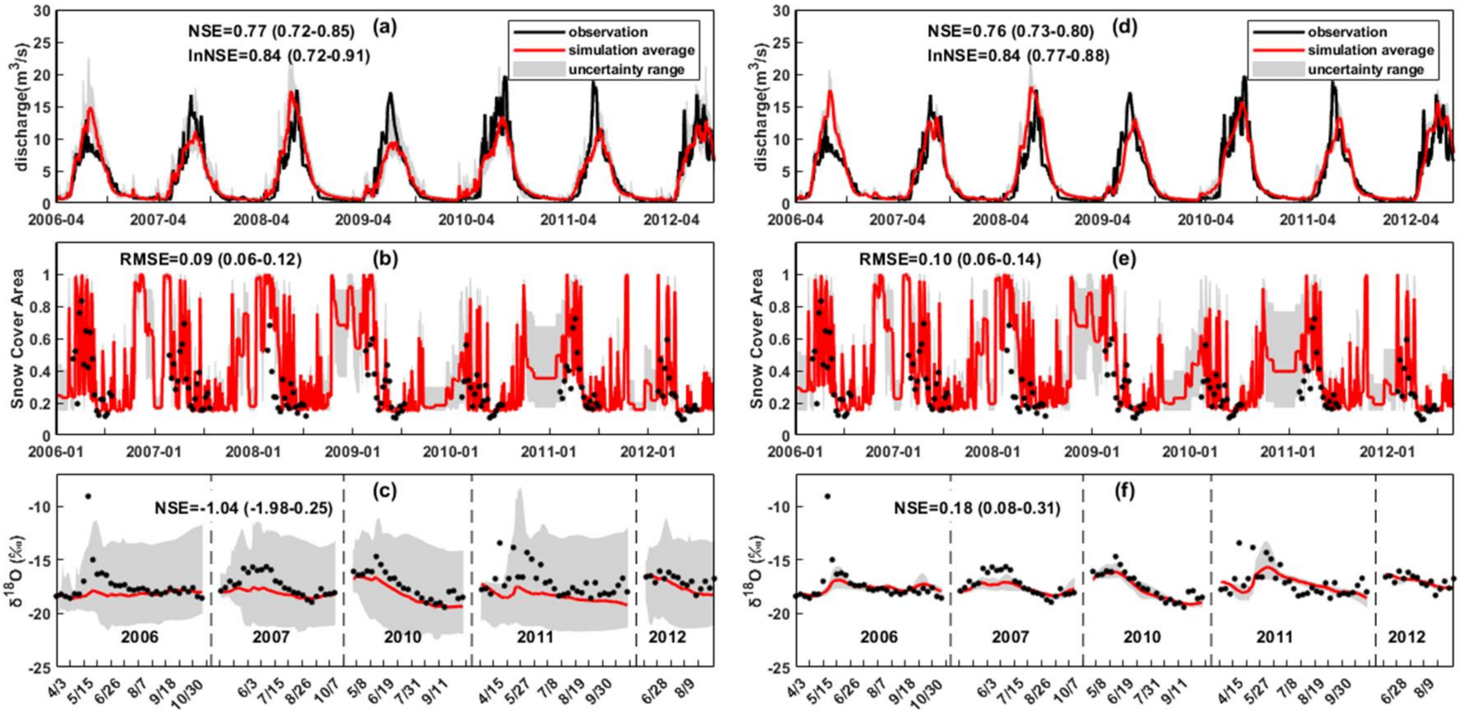


**Figure 4.** Uncertainty ranges and metrics values of the simulated discharge, SCA, and stream $\delta^{18}O$ in the KR catchment produced by the behavioral parameter sets of a bi-objective calibration (a-c) and a tri-objective (d-f) calibration in the benchmark model running.

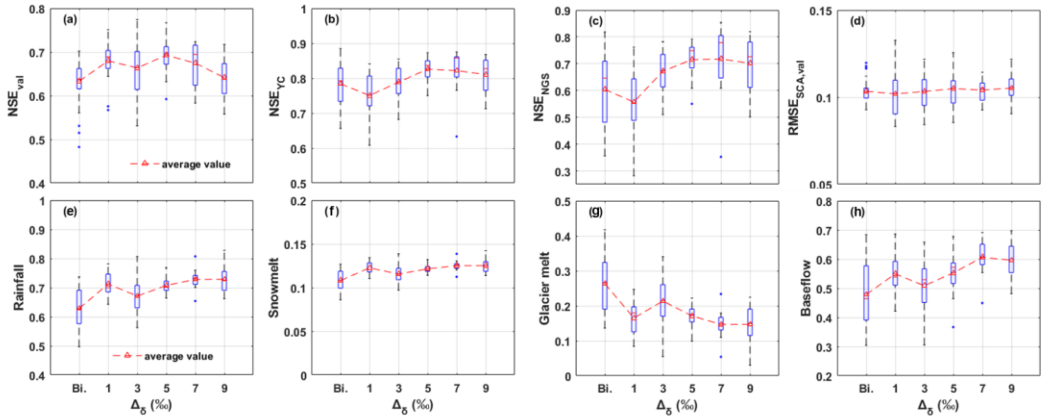


**Figure 5.** Model performances (a-d) and runoff component contributions (e-h) in the YTR basin in different scenarios using different glacier isotope input (experiment 1). Subplot (a) and (d) are the performances for Nuxia streamflow and SCA simulation in validation period, respectively. Subplot (b) and (c) are the performances for streamflow simulation at internal stations Yangcun and Nugesha, respectively. Subplot (e), (f) and (g) are the contribution of runoff components based on water source definition. Subplot (h) is the contribution of baseflow based on the runoff pathway definition.

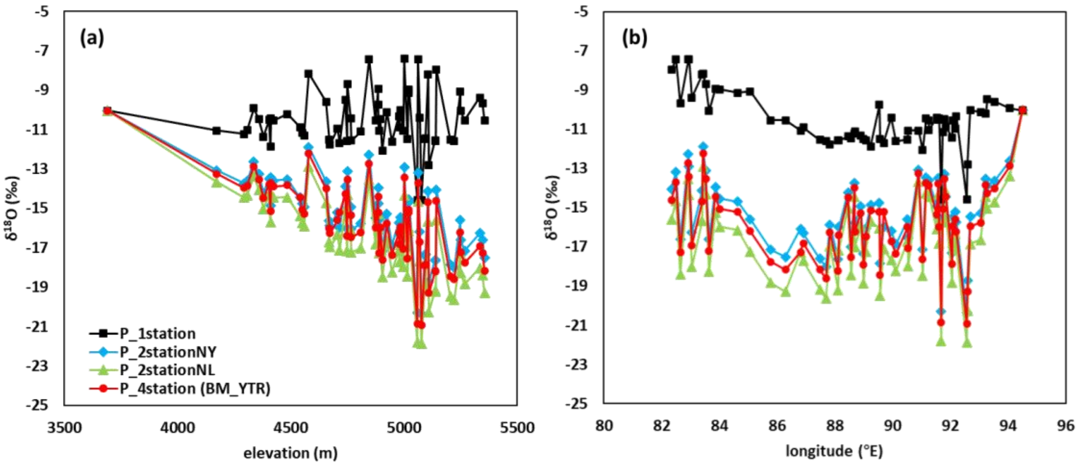


**Figure 6.** Comparisons of weighted averages of precipitation $\delta^{18}O$ on 63 REWs in the YTR by elevation (a) and longitude (b) in each scenario of experiment 2.

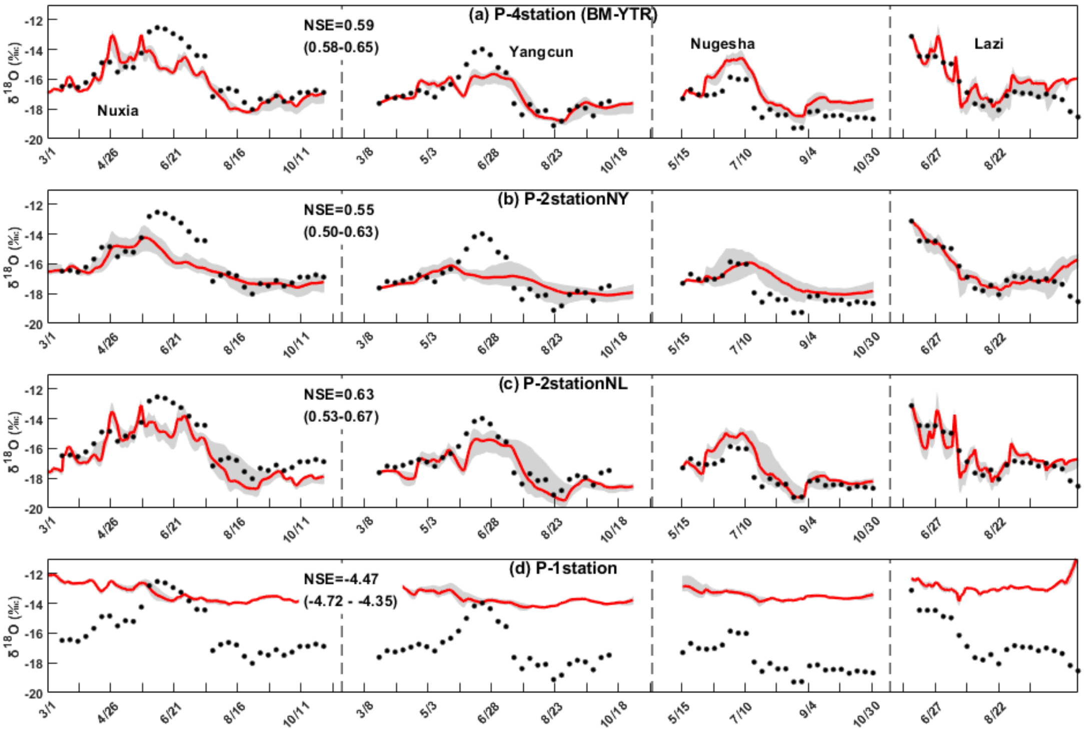


**Figure 7.** Uncertainty ranges of stream water $\delta^{18}O$ simulations at four stations in 2005 produced by the behavioral parameter sets of each scenario in experiment 2.

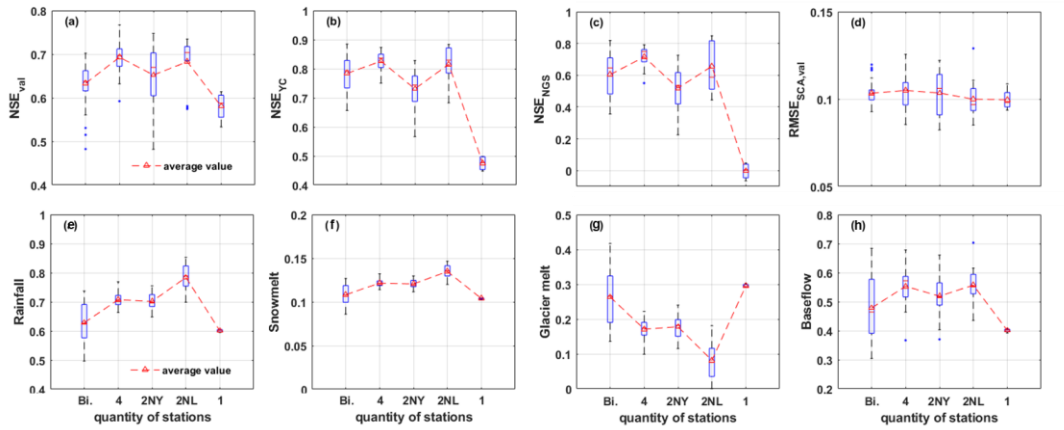


**Figure 8.** Model performances (a-d) and runoff component contributions (e-h) in the YTR basin in different scenarios using precipitation isotope measurements from different sampling sites (experiment 2). Subplot (a) and (d) are the performances for Nuxia streamflow and SCA simulation in validation period, respectively. Subplot (b) and (c) are the performances for streamflow simulation at internal stations Yangcun and Nugesha, respectively. Subplot (e), (f) and (g) are the contribution of runoff components based on water source definition. Subplot (h) is the contribution of baseflow based on the runoff pathway definition.

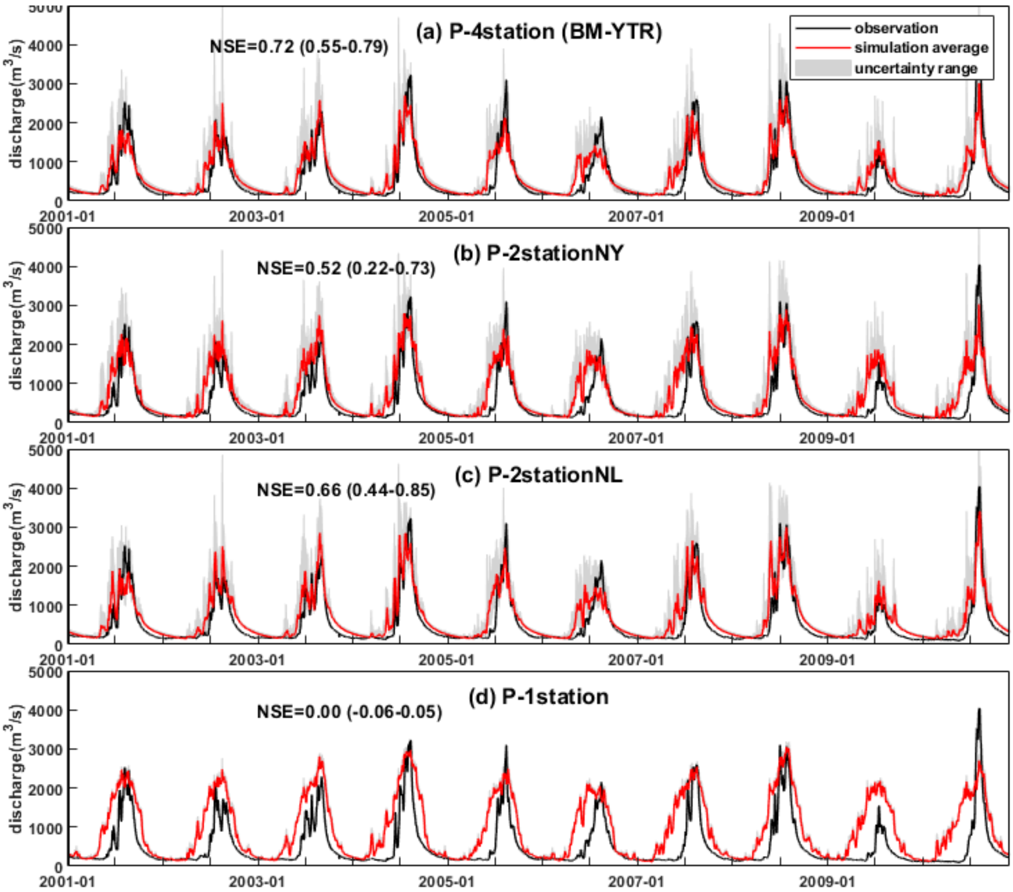


**Figure 9.** Uncertainty range and metrics values of simulated discharge at Nugesha station produced by the behavioral parameter sets of each scenario in experiment 2.

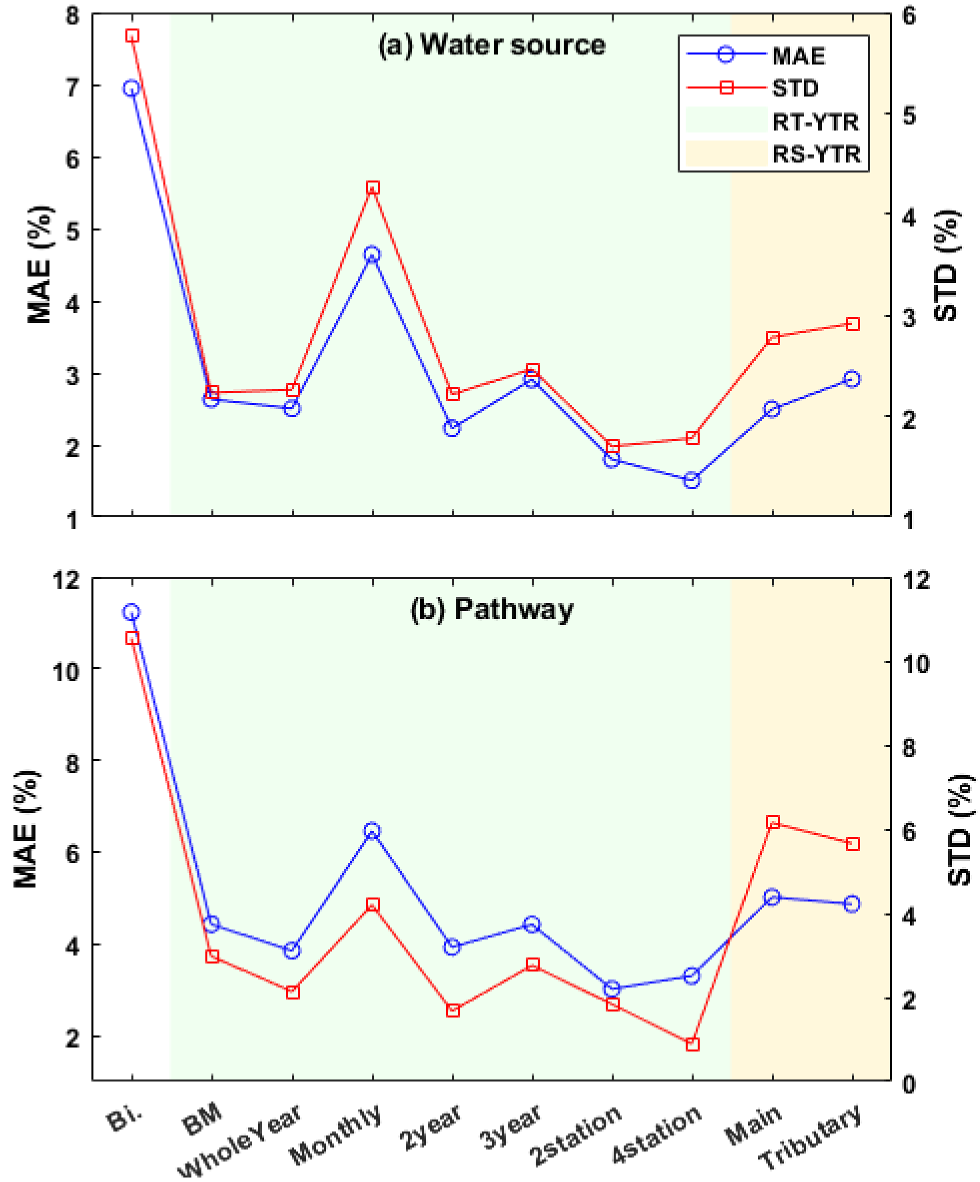


**Figure 10.** Accuracy and uncertainty metrics of estimated CRCs in the YTR basin derived from the different stream water sampling strategies (experiment 3). (a) for CRCs quantified under the definition of water source and (b) for CRCs quantified under the definition of runoff pathway.

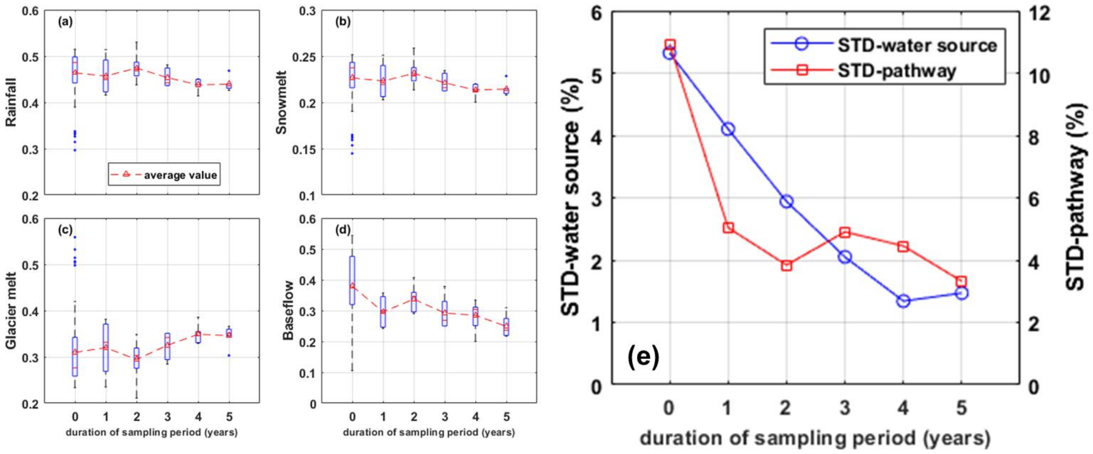


**Figure 11.** Uncertainties of the contributions of (a) rainfall, (b) snowmelt, (c) glacier melt and (d) baseflow in the KR catchment, estimated by scenarios with different durations of sampling period (experiment 3). The uncertainties of CRCs based on two different definitions are summarized in subplot (e).