AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2025WR040043

# Winter Baseflow Calibration's Critical Role in Hydrological Modeling for the Pamir Region

**J. Huang**[1], **M. Barandun**[1], **J. Richard-Cerda**[2], **M. Hoelzle**[1], **and E. Pohl**[1]

[1]Department of Geosciences, University of Fribourg, Fribourg, Switzerland, [2]Institute of Applied Geosciences, Technische Universität Darmstadt, Darmstadt, Germany

**Key Points:**

- Incorporation of winter baseflow calibration improves model estimates of snow melt runoff and baseflow
- Dual-layer groundwater module better captures baseflow dynamics
- Calibration strategy and forcing quality deserve equal attention in hydrological modeling

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

J. Huang,
jingheng.huang@unifr.ch

**Citation:**

Huang, J., Barandun, M., Richard-Cerda, J., Hoelzle, M., & Pohl, E. (2026). Winter baseflow calibration's critical role in hydrological modeling for the Pamir region. *Water Resources Research*, *62*, e2025WR040043. https://doi.org/10.1029/2025WR040043

Received 22 JAN 2025
Accepted 15 DEC 2025

**Author Contributions:**

**Conceptualization:** J. Huang, E. Pohl
**Formal analysis:** J. Huang, M. Barandun, E. Pohl
**Methodology:** J. Huang, M. Barandun, J. Richard-Cerda, M. Hoelzle, E. Pohl
**Supervision:** E. Pohl
**Validation:** J. Huang, J. Richard-Cerda, E. Pohl
**Visualization:** J. Huang, E. Pohl
**Writing – original draft:** J. Huang, J. Richard-Cerda, E. Pohl
**Writing – review & editing:** J. Huang, M. Barandun, J. Richard-Cerda, M. Hoelzle, E. Pohl



## Abstract

The Pamir Mountains, a critical water source for Central Asia, require accurate quantification of runoff components for water resource management under climate change. Uncertainties in precipitation data are known to greatly affect hydrological model accuracy, leading to the widespread use of multi-data calibration methods to avoid internal error compensation effects between snow and glacier accumulation and melt processes. Traditional approaches incorporating runoff, snow cover fraction, and glacier mass balance are frequently employed in the region's hydrological model calibration; yet we find this calibration approach to still result in significant uncertainties in the quantification of baseflow, snowmelt, and glacier runoff. Here we show winter baseflow calibration to provide a previously overlooked yet powerful constraint on model parameters, not only constraining baseflow but also enhancing the estimation of snowmelt and glacier runoff through groundwater parameters' control on hydrograph characteristics. Even with low-quality forcing data, winter baseflow calibration guides parameters toward more realistic values of runoff estimates, improving model reliability. Using five different forcing data sets, we show that incorporating winter baseflow alongside traditional calibration variables (runoff, snow cover, and glacier mass balance) reduces uncertainty ranges from 34%–61% to 8%–21% for snowmelt, 5%–17% to 3%–11% for glacier runoff, and 33%–50% to 7%–21% for baseflow estimates. Though parameter equifinality remains a challenge, winter baseflow calibration consistently enhances model accuracy, emphasizing its vital role in refining hydrological predictions in alpine, data-scarce, and climate-sensitive regions.

## Plain Language Summary

The Pamir Mountains provide vital water resources for Central Asia and face increasing climate vulnerability. Accurate hydrological modeling is essential for understanding and predicting water availability, but model accuracy depends heavily on both meteorological input data quality and calibration methods. Our study shows that incorporating winter river flow measurements significantly enhances model accuracy in this region. By incorporating this winter baseflow data alongside traditional measurements like total river flow and snow cover, our calibration method reduces uncertainty in estimating how much water comes from snowmelt, glacier melt, and groundwater. The approach remains effective even with poor meteorological inputs, making it valuable for data-scarce mountain regions.

## 1. Introduction

The Pamir Mountains, where glaciers and seasonal snow dominate the hydrological cycle, represent a crucial water supply for Central Asia, supporting ~136 million people (Pritchard, 2019; Varis & Kummu, 2012). This cryosphere-dependent system faces mounting pressure from climate change (Immerzeel et al., 2020), making accurate hydrological modeling essential for water resource management (Siegfried et al., 2010). However, uncertainties in meteorological input data (particularly precipitation) significantly affect model accuracy, in particular when quantifying the contributions of glacier and snow melt and groundwater to total runoff, and can lead to unreliable model parameter estimation (Dembélé et al., 2020; van Tiel et al., 2020). Modeling uncertainty becomes more pronounced when models rely solely on streamflow for calibration, as this approach often leads to equifinality, where the same runoff can result from either increased glacier or snow melt contributions for conditions with reduced precipitation input and vice versa (Beven, 2006; van Tiel et al., 2020). A multi-data calibration approach is strongly recommended to reduce such internal error compensation effects and better constrain the model parameter calibration (van Tiel et al., 2020). However, reliable model calibration data are often not available in high-mountain regions, especially in the Pamir (Unger-Shayesteh et al., 2013).

Mountain hydrology research has traditionally focused on cryospheric features such as glacier and snowpack changes (e.g., Lutz et al., 2014; Pritchard, 2019). In high-mountain catchments, seasonal snow accumulation and glacier storage changes complicate solving the water balance ($Q = P - \mathrm{ET} \pm \Delta S$; where $P$ is the precipitation, $Q$ is the streamflow, ET is the evapotranspiration, and $\Delta S$ is the storage change). To constrain these cryospheric variables in hydrological models, three complementary variables are typically used for model calibration: total runoff (R), glacier mass balance (GMB), and snow cover fraction (SCF). Calibrating the model using only R poses the risk of equifinality as this variable aggregates rainfall, snowmelt, and glacier melt. Adding SCF constrains the spatial extent and timing of snow melt, while GMB constrains glacier mass changes and thus glacier melt (van Tiel et al., 2020). While the R + SCF + GMB calibration approach has become standard practice for quantifying snow and glacier melt contributions to total runoff in high mountain regions (e.g., Huang et al., 2022), this method faces fundamental challenges in data-scarce regions such as the Pamir. Region-wide GMB estimates show high uncertainties, with high-resolution time series available for only a few glaciers (Barandun et al., 2021). Hydrological models often produce uncertain estimates of glacier melt contribution given a simplified glacier-change representation calibrated with limited in situ GMB observations (Tarasova et al., 2016) or no GMB data at all (Pohl et al., 2017).

Despite the increasing recognition of groundwater discharge as a significant contributor to mountain streamflow (Andermann et al., 2012; Cochand et al., 2019), its representation and use for calibration in hydrological models are limited as its quantification remains highly uncertain (Chapman, 1999; Partington et al., 2012). Strong interactions between surface runoff and aquifers have been observed in high mountain regions (Carroll et al., 2024; Cochand et al., 2019; Tiel et al., 2024) and may be intensified with permafrost thaw under climate warming (Diak et al., 2023). However, in the absence of relevant observational data, studies often assume a minor role of groundwater reservoirs, frequently represented by models through shallow and poorly developed soils, resulting in small and short-lived groundwater reservoirs (e.g., Su et al., 2022). The modeled contribution of groundwater discharge to total streamflow, as well as how much this contributes during meteorologically extreme years in the Pamir, varies between studies (Beck et al., 2013; Pohl et al., 2015; Tarasova et al., 2016). In the absence of substantial aquifers, the watersheds in the Pamirs respond relatively fast to snow and glacier melt, which constitute the bulk of the total annual flow. Models thus can effectively capture the highly seasonal flow patterns, often achieving a high Nash–Sutcliffe efficiency (NSE, Nash & Sutcliffe, 1970) despite poorly representing winter baseflow sustained by groundwater discharge (e.g., Huang et al., 2022).

Groundwater modeling requires an understanding of the aquifer's physical and hydrologic properties. These properties result from geological and geomorphological basin characteristics and can be reflected in the baseflow recession curve (Brutsaert & Nieber, 1977; Cheng et al., 2016; Hall, 1968; Tallaksen, 1995; Zecharias & Brutsaert, 1988; Zhang & Schilling, 2006). In the Pamir region, a fast hydrograph recession is followed by a slow recession in winter (Pohl et al., 2015, 2017); a behavior that has elsewhere been attributed to the coexistence of shallow high hydraulic conductivity sedimentary aquifers and deep low hydraulic conductivity bedrock aquifers (Hayashi, 2020). Traditional single-layer groundwater models with simple power law relationships fail to capture this complexity (Hayashi, 2020; Somers & McKenzie, 2020). The oversimplification of the groundwater module can lead to significant baseflow underestimation (Floriancic et al., 2018), which subsequently affects the accuracy of the quantification of snowmelt, glacier, and rainfall runoff. To better represent the groundwater discharge in the Pamir region, Pohl et al. (2015) implemented a two-layer groundwater approach: a fast recession component representing quick release from permeable sediments and a slow recession component capturing gradual discharge from deeper groundwater sources. However, despite these advances in groundwater representation, the understanding of how baseflow calibration (i.e., giving a separate weight to the model's performance during the baseflow period) affects model performance remains limited due to the scarcity of baseflow data. The available field methods, such as seepage meters and isotopic tracers, do not currently allow an accurate determination of baseflow (Becker et al., 2004; Cook et al., 2008; Klaus & McDonnell, 2013). As a result, various indirect methods for estimating baseflow have been developed (e.g., Brutsaert & Nieber, 1977; Cheng et al., 2016; Lyne & Hollick, 1979; Sloto & Crouse, 1996), but their performance varies by catchment characteristics (e.g., soil and topographical features) and is difficult to evaluate due to the absence of ground truth data (Chapman, 1999; Eckhardt, 2008; Klaus & McDonnell, 2013; Nathan & McMahon, 1990; Partington et al., 2012; Smakhtin, 2001; Xie et al., 2020).

The Gunt River basin in the Pamirs provides a unique opportunity to investigate how calibrating baseflow in hydrological model calibration influences model performance. The basin's winter streamflow (November–

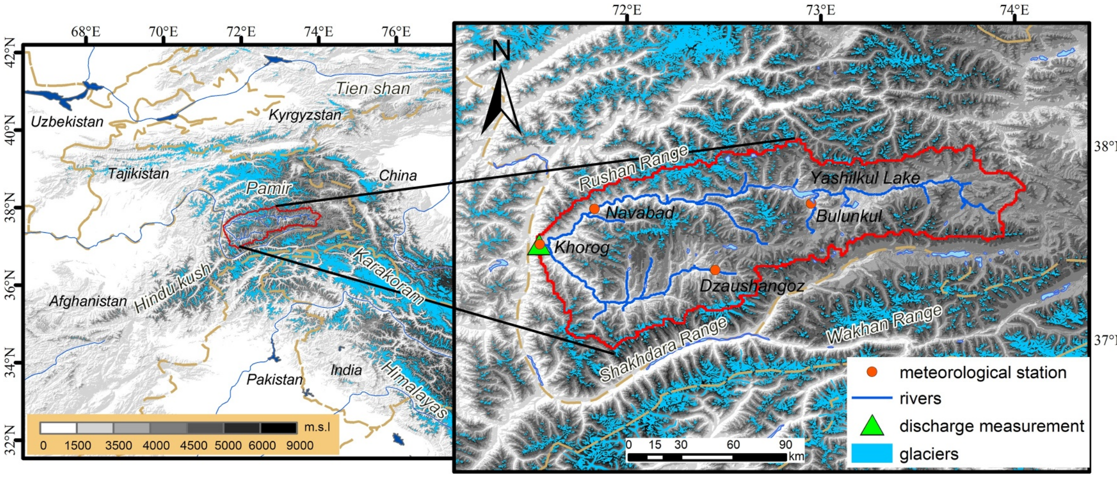


**Figure 1.** Elevation map of the Gunt basin showing the available meteorological stations (Navabad, Bulunkul, and Dzaushangoz) and the Khorog discharge measurement location. Glacier cover data obtained from Randolph Glacier Inventory 6.0.

February) derives almost exclusively from baseflow because temperatures consistently remain below freezing, preventing rainfall and snow melt (Pohl et al., 2015). This scenario provides a robust benchmark to evaluate baseflow simulation. The basin shares key characteristics with the broader Pamir region in terms of topographic relief, climate (e.g., negative west–east precipitation gradient), and nival–glacial hydrological regime (Pohl et al., 2017), making it a suitable case study for providing insights into hydrological model calibration in the complex Pamir mountain system. Here, we use the spatial processes in hydrology (SPHY) model (Terink et al., 2015), a widely used tool in high mountain hydrological simulations (e.g., Immerzeel et al., 2020; Lutz et al., 2016), to evaluate how baseflow calibration affects hydrological model performance in the Gunt River basin. We modify the model with an additional linear groundwater storage component to better represent varying delays in the baseflow recession as in Pohl et al. (2015). We aim to answer the research questions of how incorporating baseflow calibration alongside traditional variables (R, SCF, and GMB) can (a) improve model performance and (b) reduce parameter uncertainty. We use multiple reanalysis-based meteorological forcing data sets, the most feasible data source for the data-sparse mountain region (Zandler et al., 2019), to assess the robustness of this calibration approach under different meteorological forcing conditions. The present study thus provides an improved framework for hydrological modeling in the Pamir region, in particular for better quantification of runoff components.

## 2. Study Area

The Gunt River basin is located in the Pamir Mountains and covers an area of approximately 14,000 km$^2$ (Figure 1). The basin's elevation extends from 2,000 m to 6,700 m a.s.l., with an average of about 4,300 m a.s.l. Land cover is predominantly barren or sparsely vegetated, accounting for 67.6% of the area (Pohl et al., 2015). About 5% of the catchment area is covered by glaciers (Randolph Glacier Inventory 6.0).

The Westerlies primarily drive moisture supply in the Gunt basin from west to east (Baldwin & Vecchi, 2016; Schiemann et al., 2008). Over 85% of precipitation falls as snow in winter and spring (Haag et al., 2019). The Gunt basin is located in the transition zone between the westerly winds and the Indian summer monsoon (Fuchs et al., 2013; Palazzi et al., 2013), and can be characterized as a dry and continental basin. The western part receives higher annual precipitation (∼300 mm) than the eastern part (∼100 mm; Farinotti et al., 2020; Palazzi et al., 2013; Pohl et al., 2015; Wang et al., 2017). Although the eastern part is generally higher in elevation and receives less precipitation than the western part, no apparent correlation has been found between altitude and precipitation (Pohl et al., 2015). According to meteorological records, the temperature gradient in the Gunt basin follows topography, with temperatures staying below 0°C during the winter months (November–February), favoring snow accumulation (Haag et al., 2019; Pohl et al., 2015).

In the arid Pamir, sublimation can be significant due to strong solar radiation, low humidity, and high wind speeds. Studies from the neighboring Tien Shan mountains have reported sublimation to account for approximately 5%–15% of the total annual ablation at Urumqi Glacier No. 1 and Kyzylsu glacier (Fugger et al., 2024; Yao et al., 2012). Given the dry conditions in the study area, relative sublimation losses may be even higher.

## 3. Data

### 3.1. Climate Data

The SPHY model simulates the water balance components, where $Q = P - \mathrm{ET} \pm \Delta S$. The model requires daily precipitation sum and temperature (maximum, minimum, and average) as forcing inputs. Four meteorological stations in the Gunt basin record precipitation; however, their location in valleys below 3,800 m (Figure 1) renders the spatial interpolation and derivation of local precipitation gradients highly uncertain (Teegavarapu, 2014). For temperature, while its general decrease with elevation is well understood, local variations from katabatic winds and complex topography render interpolation from limited valley stations problematic, in particular over glaciers where near-surface cooling can significantly modify local temperatures (Salerno et al., 2023).

Therefore, we choose reanalysis products given their broad spatial coverage, physical basis, and accessibility. To reduce modeling uncertainty associated with using a single forcing data set (Barandun & Pohl, 2023), we incorporated several data sets. For precipitation, we use the land component of the European Center for Medium-Range Weather Forecasts (ECMWF) fifth-generation reanalysis (ERA5-Land) precipitation (Hersbach et al., 2020), derived from ERA5 through bilinear interpolation at a resolution of ∼31 to ∼9 km. We also use the climatologies at high resolution for the Earth's land surface areas (CHELSA) precipitation (Karger et al., 2021), which also builds on ERA5 but applies a more sophisticated downscaling approach to achieve a ∼1 km resolution by incorporating orographic predictors, including wind fields, valley exposition, and boundary layer height, combined with satellite-derived cloud frequencies. To enable direct comparison with previous work in the region (Pohl et al., 2015), we also include the High Asia Refined analysis version 1 (HAR v1; Maussion et al., 2014) precipitation data set, even though this is a previous-generation product not extended beyond 2013. HAR v1 provides dynamically downscaled daily precipitation at 10 km resolution through the Weather Research and Forecasting (WRF) model (Barker et al., 2012).

For temperature, we use ERA5-Land temperature, which applies a constant lapse rate of $-0.0065°\mathrm{C\ m^{-1}}$ in the downscaling process (∼31 to ∼9 km resolution). Additionally, we use the Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Land Surface Temperature/Emissivity Daily Global product (MOD11C1 version 6) night land surface temperature (∼5 km resolution), which has shown good agreement with station measurements in previous studies and thus its usability as a surrogate for air temperatures in the region (Pohl et al., 2015).

These data sets result in five main forcing combinations: ERA5 + ERA5, ERA5 + MODIS, CHELSA + ERA5, and CHELSA + MODIS, plus the HAR + MODIS combination following Pohl et al. (2015). We excluded CHELSA temperature data as they derive directly from ERA5, using an identical methodology to that of ERA5-Land. An evaluation of these products against station observations is presented in Section 4.4.

While numerous other precipitation and temperature products are available (Sun et al., 2018), we selected just a representative set of products, as a comprehensive assessment of forcing data quality lies beyond the scope of the present study.

### 3.2. Glacier and Snow Cover Data

Including GMB observations as one of the calibration data sets helps improve the accuracy of glacial runoff simulation. However, in situ GMB observations are scarce in the Pamirs (Barandun et al., 2020; Hoelzle et al., 2017). In the Gunt basin, only one glacier monitoring site, established in 2020, is available (glacier No. 457). Several geodetic-based GMB data sets (Hugonnet et al., 2021) are available. While these studies often cover extensive areas, their temporal resolution is typically limited to periods of 5 years or longer. As a result, they fail to capture interannual or even seasonal variations. Here, we use an annual GMB data set based on a distributed mass balance model calibrated with sub-seasonal snowline observations and geodetic mass balances (Barandun et al., 2021). Repeated snowline observations during the melt season enable a robust calibration, offering a significant advantage over modeling studies constrained by limited glaciological measurements or the coarse temporal resolution of geodetic GMB.

Initial glacier distributions in the model are derived from the Randolph Glacier Inventory (RGI) version 6 (Pfeffer et al., 2014). The RGI v6 represents the glacier outlines of the 2000s in the Pamir, and the glacier area in each model grid is updated yearly. Outlines of supraglacial debris-covered glaciers from RGI v6 are derived from Scherler et al. (2018) based on Landsat 8 and Sentinel-2 optical satellite imagery. The glaciers' initial ice thickness is estimated using the GlabTop2 model (Frey et al., 2014; Linsbauer et al., 2012). We use the MODIS MOD10A2 snow product (Hall & Riggs, 2007), which has been extensively used and validated (Arsenault et al., 2014), to estimate the monthly SCF in the Gunt basin. MOD10A2 offers a spatial resolution of 500 m and a temporal resolution of 8 days, with data available from 2000 to the present.

### 3.3. Discharge Data

We obtained daily discharge data (with gaps) from 1998 to 2013 at the basin outlet (Khorog) from the State Administration for Hydrometeorology of Tajikistan. The Gunt River hydrograph presents a nival–glacial regime showing low discharge in winter, which rapidly increases in spring, peaks with sharp daily fluctuations in summer, and sharply declines in early fall (Pohl et al., 2015). Summer hydrographs show strong interannual fluctuations, contrasting with the low and relatively constant winter hydrographs. Snow and glacier melt therefore play a significant role in the hydrological cycle of the Gunt basin, although the contribution of snow melt and glacier runoff to total runoff is debated (Pohl et al., 2015; Tarasova et al., 2016).

Groundwater discharge is generally recognized as the primary and most significant component of the baseflow, especially during periods of low flow (Bachman, 1998; Hewlett & Hibbert, 1967). The baseflow recession analysis (Figure S1 in Supporting Information S1), following Safeeq et al. (2013), also indicates that groundwater discharge dominates runoff in the Gunt basin during winter (when recession constant $<0.065$). In the present study, we assume winter runoff to correspond entirely to baseflow and to come entirely from groundwater discharge. However, this assumption is not without uncertainty (see Section 6).

Since 2006, a naturally dammed lake has been regulated to maintain constant winter discharge for a downstream hydropower plant near Khorog (Figure 1). The drawdown of the winter discharge is compensated by retaining water during the summer, but the regulation effect on summer flows is minor (less than 10% reduction), and the summer lake level is maintained within its natural fluctuation range to protect the ecosystem (World Bank, 2008). Therefore, this naturally dammed lake primarily affects winter low flows, while it does not significantly alter the annual or seasonal runoff patterns in the catchment. Consequently, total runoff data after 2006 are included in the statistics for model calibration and validation, while winter runoff data after 2006 are excluded. Other human impacts on discharge are negligible, with summer irrigation in the valleys accounting for only 0.4% of the mean annual discharge in Khorog (Pohl et al., 2015).

### 3.4. Other Data

For soil hydraulic properties, we use HiHydroSoil v2.0 high-resolution global maps, which provide a comprehensive global inventory of soil hydraulic variables in a gridded format with a spatial resolution of 250 m (Simons et al., 2020). We derive land use data from the global land cover map based on ENVISAT medium resolution imaging spectrometer (MERIS) Level-1B data, with a spatial resolution of approximately 300 m (Defourny et al., 2009). This map includes 22 land cover classes defined by the United Nations (UN) Land Cover Classification System (LCCS). The overall accuracy of these land cover classes is approximately 68% (Simons et al., 2020).

## 4. Methods

### 4.1. Hydrological Model

We use the SPHY model, a spatially distributed, physically based leaky-bucket type of model designed for cryospheric–hydrological studies in high mountain regions (Terink et al., 2015). We selected the SPHY model mainly for its computational efficiency and suitability for simulating cryosphere–hydrological processes in data-scarce mountain regions. The SPHY model integrates parameterizations of the dominant hydrological processes, including (a) rainfall runoff, (b) lake/reservoir outflow, (c) cryospheric processes (snow and glacier), (d) dynamic vegetation, (e) evapotranspiration, and (f) soil hydrological processes. The model has been widely used in hydrological modeling in mountain regions, especially in the Tibetan Plateau (e.g., Khanal et al., 2021; Lutz et al., 2016). In the present study, the model runs at a daily time step at 1 × 1 km spatial resolution. The calculated runoff components are summed up for each grid cell and then routed to yield the total runoff. The SPHY model uses the simple flow accumulation routing scheme, with a flow recession coefficient to account for flow delay. A detailed description of the model can be found in Terink et al. (2015).

SPHY model calculates reference evapotranspiration using the Hargreaves method, followed by the calculation of potential evapotranspiration using the crop coefficient approach (Hargreaves & Samani, 1982). Surface runoff is calculated based on saturation excess runoff, known as Hewlettian runoff (Hewlett, 1961). The baseflow calculation in SPHY is based on the steady-state response of groundwater flow to recharge (Hooghoudt, 1940), which is also used in the Soil and Water Assessment Tool (SWAT) model (Neitsch et al., 2009). The model uses an exponential decay weighting function (Venetis, 1969) to estimate the groundwater recharge. Water table fluctuations are modeled as a nonsteady response to periodic recharge (Rycroft & Smedema, 1983). Detailed procedures for calculating baseflow can be found in Terink et al. (2015).

A distributed temperature index model is used to estimate glacier and snow melt (Hock, 2003). Snowfall occurs when the air temperature falls below 0°C, calculated for each grid cell (1 × 1 km) based on precipitation input. Snow melt and rain can infiltrate and refreeze in the snowpack. The amount of freezing and meltwater retention in the snowpack is determined by the snow storage capacity, which can be calibrated. Melt is proportional to air temperature and is calculated at a daily resolution for all grid cells except for those covered by glaciers. The total catchment-wide snow melt reported in the present work results from the sum of snow melt on the ground and snow melt from glaciers (Terink et al., 2015).

To simulate glacier changes, we use a subgrid of 100 × 100 m for each RGI glacier outline to account for elevation-dependent glacier melt. A temperature lapse rate extrapolates the model forcing temperature for the corresponding elevation band (see Section 4.4) to each subgrid cell of the glacier. The downscaled forcing data define the precipitation distribution and are not redistributed to each subgrid cell. We use three different degree-day factors (DDFs): one for clean ice (DDF_ice), one for debris-covered ice (DDF_debris), and one for snow (DDF_snow). The DDFs are constant in space and time. We multiply these factors by the subgrid cell temperature to calculate melt when the air temperature is above the critical temperature of 0°C. Accumulation is described as all solid precipitation that occurs below the critical temperature. Melt can infiltrate the snowpack and be retained on the glacier until maximum retention capacity is reached. Any additional melt will run off and may infiltrate into the ground or become surface runoff after leaving the glacier surface.

The annual mass balance is the sum of accumulation and ablation at the end of the hydrological year and is used to calculate ice thickness and volume change. At the end of the hydrological year, the accumulated snow is directly converted into ice and redistributed over the ablation area to account for flow dynamics based on the ice volume (Morlighem et al., 2011). The goal of this redistribution is to account for the geometric adjustment of the glacier. In the present study, the modeled glacier area change is less than 3% of the total glacier area, and its impact on the calculated mass balance is thus minor.

### 4.2. Model Improvements

The SPHY model has been successfully applied to simulate runoff in glacierized high alpine catchments (e.g., Khanal et al., 2021; Lutz et al., 2016). However, the distinct hydrological characteristics of the Pamir region require two key model improvements: (a) incorporation of snow melt infiltration and (b) implementation of a dual groundwater system.

Compared with often rainfall-dominated systems with a majority of groundwater recharge resulting from rainfall infiltration in the Himalayas (Lone et al., 2021), snow infiltration is of paramount importance for groundwater recharge in nival-glacial hydrological systems (Meeks et al., 2017; Mohammed et al., 2019; Urrutia et al., 2019). This is also the case in Pamir, where precipitation occurs predominantly in solid form during the winter months. We modify the original SPHY model, lacking snow melt infiltration, to incorporate snow melt infiltration following saturation excess principles (Hewlett, 1961); in doing so, snow melt is able to contribute to groundwater recharge while generating surface runoff only when the upper soil layer reaches saturation.

In the original SPHY model, baseflow recession is modeled as an exponential function of time (Terink et al., 2015). Figure S1 in Supporting Information S1 shows the winter baseflow recession in the Gunt River basin from 1 November 2000 to 28 February 2001, reflecting the basin's typical winter recession pattern. We fitted the

**Table 1**
*Most Sensitive Parameters in Improved SPHY Model*

| Parameter | Description | Unit |
|---|---|---|
| PCF | Precipitation intensity correction factor | – |
| TC | Temperature correction factor | °C |
| RootDF | Field capacity correction factor in the first soil layer | mm/mm |
| SubDF | Field capacity correction factor in the second soil layer | mm/mm |
| dtGw1 | Recharge delay time for GW1 | day |
| dtGw2 | Recharge delay time for GW2 | day |
| alpGw1 | Control the recession speed for GW1 | – |
| alpGw2 | Control the recession speed for GW2 | – |
| SplitGW | Recharge proportion to ground layer2 | % |
| DDFI | Degree-day factor for clean ice | mm w.e. $day^{-1}$ $°C^{-1}$ |
| DDFDG | Degree-day factor for debris-covered ice | mm w.e. $day^{-1}$ $°C^{-1}$ |
| DDFS | Degree day factor for snow | mm w.e. $day^{-1}$ $°C^{-1}$ |
| SnowSC | Storage capacity of snowpack | mm/mm |
| GlacF | Percentage of glacier melt that becomes surface runoff | % |
| Kx | Recession coefficient of routing | – |

winter discharge based on exponential recession behavior (Safeeq et al., 2013) and constructed the recession curve using the adapted matching strip method (Posavec et al., 2006). The recession curve presents two stages: the first shows a rapid recession with a recession constant of 0.0067, while the second shows a slower recession rate with a recession constant of 0.0007 (Figure S1 in Supporting Information S1). The use of a single recession curve underestimates baseflow during the slow recession period (Figure S1 in Supporting Information S1). To represent this dual recession system, a double groundwater layer module is used instead: the first reservoir (GW1) simulates a fast recession through sedimentary deposits, and the second reservoir (GW2) represents a traditional deep groundwater layer with a slower recession rate. The "fast" layer receives recharge from percolation through the unsaturated soil profile, and a fraction of the recharge can be routed to the deep aquifer, which is controlled by the parameter SplitGW (Table 1). Those two layers can generate the baseflow in parallel, showing different recession (alphaGW1, alphaGW2) and recharge (deltaGW1, deltaGW2) rates. Details regarding the double aquifer can be found in Luo et al. (2012) and Neitsch et al. (2009). The groundwater module is parameterized through calibration using observed winter discharge data (1998–2006), which is considered baseflow and results from groundwater discharge.

### 4.3. Definition of Each Runoff Component

In the improved SPHY model, there are four types of runoff: (a) Rainfall runoff is defined as the water flow generated when rainfall exceeds the soil's infiltration capacity. (b) Snowmelt runoff comprises surface snowmelt runoff (including snowmelt on glaciers) and subsurface drainage generated when the infiltrated snowmelt exceeds the soil's field capacity. (c) Glacier runoff refers specifically to bare ice melt from glaciers after the overlying snow has melted. While some of the ice melt infiltrates through glacier-margin sediments to contribute to baseflow (Tiel et al., 2024), we define glacier runoff here as only the surface runoff component of ice melt. (d) Baseflow combines the groundwater contributions from all sources: snow melt infiltration, glacier melt infiltration, and rainfall percolation, regulated through the dual groundwater system described in Section 4.2.

### 4.4. Forcing Downscaling

All forcing datasets—except CHELSA precipitation—need to be downscaled from their original resolution to match our SPHY model's 1 × 1 km spatial resolution. We adopt an elevation-based downscaling approach, given the role of elevation in climate variability in mountain regions (Liston & Elder, 2006).

We divide the Gunt basin equally into 10 elevation bands with 10 equal percentiles to better characterize precipitation–elevation relationships, since precipitation gradients can vary significantly at different elevations in high-mountain regions (Immerzeel et al., 2015; Ruelland, 2020). For each elevation band, we obtain the gradients from CHELSA data, which is considered to be able to capture small-scale precipitation gradients (Karger et al., 2017). We use CHELSA data to obtain the probability density distribution of precipitation gradients in the Gunt River (Figure S2 in Supporting Information S1). We consider a grid cell in CHELSA together with its eight neighboring points as one group, and the precipitation gradient within this group is calculated from the regression of precipitation against elevation. This process is repeated for each grid cell to obtain the probability density function of the precipitation gradients. The precipitation gradient distribution in each elevation band shows very similar characteristics, with a median of 2%/100 m to 2.2%/100 m (Figure S2 in Supporting Information S1). The gradient from CHELSA is substantially lower than the median value of 9.8%/100 m found in the Upper Indus basin through inverse modeling (Immerzeel et al., 2015). The lower precipitation gradients in the present study can either reflect the more arid conditions of the Pamir region, which receives less moisture from westerly circulation systems compared with the Upper Indus basin (Pohl et al., 2015; Schiemann et al., 2008), or might simply be related to unresolved altitude-related gradients within the CHELSA data set. While more sophisticated downscaling approaches exist (e.g., Fiddes & Gruber, 2014), developing high-resolution precipitation fields is beyond the scope of this study. To simplify, we apply the gradient of 2%/100 m to all elevation bands for downscaling the ERA5-Land and HAR precipitation from 10 × 10 km to 1 × 1 km. To downscale temperature from ERA5-Land and MOD11C1, we use lapse rates calculated from the products themselves (0.55°C/100 m and 0.5°C/100 m, respectively). To evaluate the downscaled products, we compare them with independent station observations (2000–2005) from elevations between 2,080 and 2,751 m; however, these stations cover only the lower elevation range of the catchment (Figures S3 and S4 in Supporting Information S1). To enable direct comparison, gridded data are first adjusted to station elevations using the same elevation gradients applied in the downscaling process. These stations are not assimilated into the ERA5-Land reanalysis (Hersbach et al., 2020). Monthly precipitation correlations with observations vary among the data sets (Figure S3 in Supporting Information S1), with ERA5-Land showing higher correlations ($r$ = 0.5–0.7) compared with those of CHELSA ($r$ = 0.3–0.5) and HAR ($r$ = 0.4–0.5). ERA5-Land and HAR consistently overestimate precipitation compared with observations, with long-term monthly ratios ranging from 1.2 to 1.8 for ERA5-Land and from 1.7 to 3.8 for HAR (Figure S3 in Supporting Information S1). Considering these overestimations, correction factors of 0.56–0.83 for ERA5-Land and 0.26–0.59 for HAR should be applied to initial model parameter values. CHELSA shows no systematic bias but higher temporal variability.

We apply a similar elevation adjustment for temperature comparisons. Compared with the in situ observations, MODIS tends to exhibit a cold bias, while ERA5 shows no clear pattern. ERA5 temperature shows a root mean square error (RMSE) of 4.17°C–5.58°C across stations, whereas MODIS temperatures show greater variability, with an RMSE ranging from 7.65°C to 12.2°C.

These biases provide reference values for initial model parameters in the subsequent calibration. However, it should be noted that the comparison is inherently limited by a scale mismatch between point measurements and gridded data (Tustison et al., 2001) and the relatively low elevation range of available stations.

### 4.5. Model Parameter Estimation and Optimization

The model calibration process is illustrated in Figure 2. The improved SPHY model used in present work presents a high number of parameters ($n = 60$; Table S1 in Supporting Information S1). While such high number inherently increases equifinality risks, it also allows us to test whether winter (November–February) baseflow (WBR) can effectively constrain even highly parameterized models. We identify the most sensitive parameters (Table 1) using the Morris global sensitivity analysis (Morris et al., 2014). We iteratively adjust the most sensitive parameters to obtain an overall optimal set of calibration parameters (Immerzeel et al., 2015). Note that ice typically shows a higher DDF than snow does; the ratio between DDF for ice and DDF for snow is therefore fixed at 1.9 based on the study by Barandun et al. (2021).

We conduct model simulations over the time period 1995–2015, constrained by the availability of discharge and GMB observations, with the first 3 years (1995–1997) designated as a warm-up period to initialize model states.

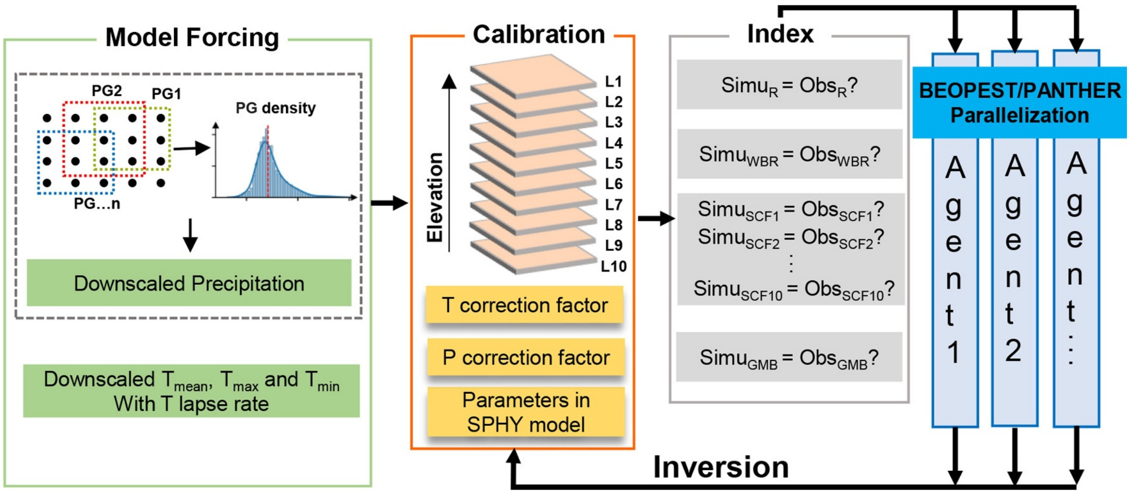


**Figure 2.** Flowchart of the model calibration process. Note PG, precipitation gradient; T, temperature; L, elevation band; Simu, simulation; Obs, observation; R, total runoff; WBR, winter (November–February) baseflow; SCF, snow cover fraction from MODIS MOD10A2; GMB, glacier mass balance based on snowline-aided method (Barandun et al., 2021).

All forcing data sets cover this entire period except for HAR precipitation, which is only available until 2013. For model calibration, we use the period 2001–2005, when all calibration variables (R, WBR, SCF, and GMB) are concurrently available. While we acknowledge that snowline-derived GMB data show higher uncertainty during the calibration period compared with that shown in post-2013 observations (Barandun et al., 2021), we accept this uncertainty as a necessary trade-off to capture natural hydrological conditions. For validation, we use two separate time periods to validate the model: 1998–1999 (precalibration) and 2008–2015 (postcalibration). WBR after 2006 is excluded from validation due to flow regulation from the dam. However, the regulation of the dam is assumed to have minimal impact (<10%) on annual runoff magnitude and seasonality (see Section 3.3), so total runoff is included in the statistics.

Given the complexity of mountain hydrology, we implement a simultaneous calibration strategy in which all parameters are optimized concurrently using the complete set of observational constraints. The simultaneous calibration, though computationally intensive, better captures parameter interdependencies than stepwise calibration, in particular when dealing with sparse observational data or potential systematic measurement errors (Schaefli & Huss, 2011). While stepwise calibration can provide explicit constraints on model internal consistency (van Tiel et al., 2020), it may also obscure the model's capacity to represent coupled cryosphere–hydrological processes in data-scarce mountain regions.

To automatically calibrate the SPHY model, we use the parameter estimation (PEST) suite with a global parameter search algorithm (Doherty, 2015). PEST is a nonlinear parameter estimation and optimization package offering model-independent optimization routines. PEST can undertake local or global optimization using differential evolution methods. Its role is to minimize the weighted sum of squared differences between simulations and observations, which is calculated as:

$$\phi = \sum_{i=1}^{m} (w_i r_i)^2 \tag{1}$$

where $\phi$ is the objective function, $m$ is the total number of time steps in the calibration period, $w_i$ is the observation weight, and $r_i$ is the difference between the simulations and the observations.

The fact that these discrepancies can be weighted renders some observations more important than others in determining the optimization outcome. Choosing the proper weight can significantly impact the calibration results and often represents a trial-and-error process. Ideally, the weight should be set at a level commensurate with the measurement noise level. Following Clausnitzer and Hopmans (1995), we calculate the weight as:

$$w_i = \left(\frac{1}{n_i \cdot \sigma_i^2}\right)^{1/2} \quad (2)$$

where $n_i$ is the number of observations and $\sigma_i^2$ is the variance of the observations.

The optimization process terminates when a stopping criterion (e.g., the convergence of the total objective function and the convergence of the parameter set) is satisfied. Many sets of parameter values can yield close to identical optimal objective function values (global minima or maxima), an issue known as the equifinality problem. During each iteration, PEST undertakes first-order second-moment (FOSM)-based predictive analysis to find a parameter set that results in the maximum or minimum model prediction while still calibrating the model.

The NSE is used as an assessment criterion of the runoff, while the relative and absolute biases serve as efficiency criteria for assessing model performance in simulating SCF and GMB. Following Moriasi et al. (2007), for runoff simulation, the model performance can be classified as very good (NSE > 0.75; bias < 10%), good (0.65 < NSE <0.75; 10% < bias < 15%), satisfactory (0.50 < NSE < 0.65; 15% < bias < 25%), or unsatisfactory (NSE < 0.50 or bias > 25%). For SCF, we applied a ±20% bias threshold, considering the documented MODIS accuracy of ~60%–80% in mountain regions, with lower accuracy at higher elevations (Frei et al., 2012; Liang et al., 2017; Şorman et al., 2007; Wang et al., 2008). For GMB, we used a ±0.25 m w.e. $yr^{-1}$ bias threshold based on the reported uncertainty for the Gunt basin (Barandun et al., 2021).

### 4.6. Winter Baseflow Separation

To assess the uncertainty of our reference winter baseflow, we employ four graphical separation methods, six digital filter methods, one recession analysis method, and one method that combines the recession analysis and the digital filter to quantify winter baseflow from the total discharge. We choose those methods because they require only streamflow data as input. The graphical methods include the smoothed minima method (UKIH, UK Institute of Hydrology, 1980) and three methods developed by the United States Geological Survey (Fixed, Sliding, and Local; Sloto & Crouse, 1996). The six digital filter methods are the LH method (Lyne & Hollick, 1979), the Chapman method (Chapman, 1991), the Boughton method (Boughton, 2004), the Eckhardt method (Eckhardt, 2005), the EWMA method (Tularam & Ilahee, 2008), and the HydRun method (Tang & Carey, 2017). The recession analysis method is from Cheng et al. (2016) and provides an objective methodology for automated selection of base flows based on the BN77 method (Brutsaert & Nieber, 1977). The last method that combines the recession analysis and the LH filter is from Duncan (2019).

## 5. Modeling Results

### 5.1. Model Calibration and Validation

The SPHY model is calibrated with R + WBR + SCF + GMB and driven by different forcing inputs. The optimal parameter set for each forcing is automatically estimated by PEST, representing the overall best model performance (Equation 1). Figure 3 and Table 2 show the model performance in simulating the daily R and WBR. For total runoff during calibration, ERA5 precipitation-forced models achieved very good performance (NSE: 0.84 to 0.92; bias: −2.3% to −3.4%), and HAR + MODIS showed similar quality (NSE: 0.90; bias: −5.9%). CHELSA-precipitation-forced models performed satisfactorily (NSE: 0.57 to 0.59; bias: −20.9% to −21%). The lower performance with CHELSA forcing is likely due to anomalously low precipitation estimates in certain years (Figure S5 in Supporting Information S1). For example, only 50 mm were estimated for 2001, far below the 2001–2005 average of 300 mm, and much lower than the other precipitation data sets used in this study (Figure S5 in Supporting Information S1). During validation periods, ERA5 precipitation-forced models performed best, maintaining satisfactory to good performance (NSE: 0.68 to 0.83; bias: 6.4%–17.0%). HAR + MODIS showed satisfactory performance for 2008–2013 (NSE: 0.78; bias: 12.6%), while CHELSA-precipitation-forced models showed satisfactory performance but with consistently larger bias (NSE: 0.73–0.77; bias: 18.8%–23.0%).

For WBR simulation during calibration, CHELSA + ERA5 showed the best performance (NSE: 0.83; bias: −0.4%), followed by ERA5-forced models (NSE: 0.73 to 0.76; bias: −1.5% to −1.8%). HAR + MODIS showed the poorest WBR performance (NSE: 0.62; bias: −0.42%) despite its good performance in total runoff simulation

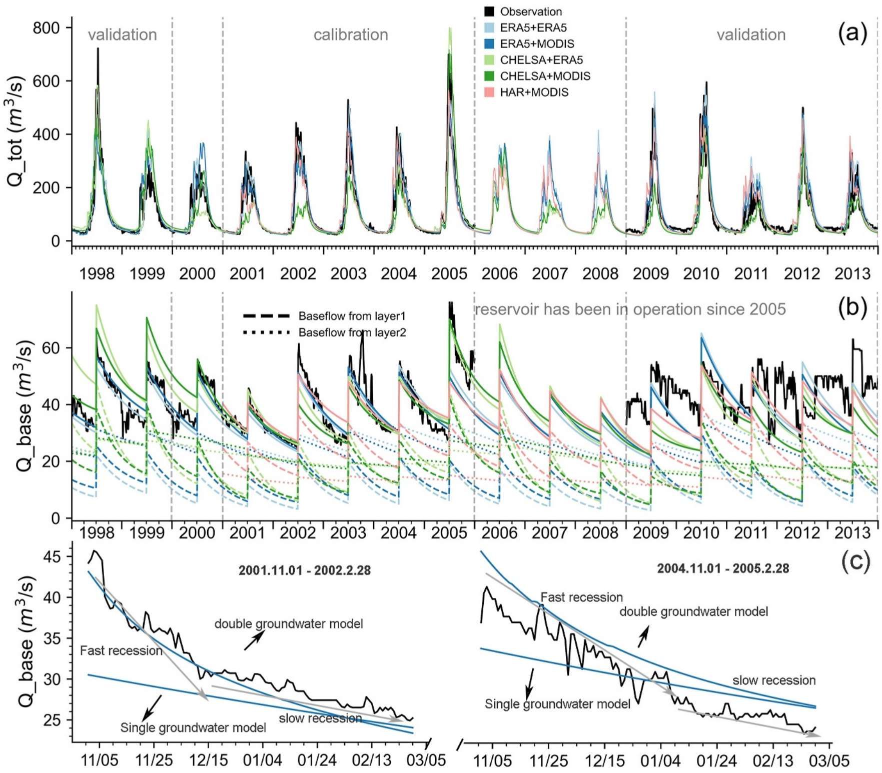


**Figure 3.** (a) Comparison between the best individual model simulation per input data combination and observed total discharge. The parameter set for each forcing is automatically estimated by PEST, representing the best overall model simulation performance. (b) Comparison between the simulated and observed WBR. Since 2006, Yashilkul Lake regulation has significantly affected WBR; thus, WBR has been excluded from statistics since 2006. The regulation is assumed to not significantly impact the overall annual runoff magnitude (<10%), nor runoff seasonality, and total runoff is therefore included in the statistics. (c) Simulated WBR comparing single- versus dual-layer groundwater representation (GW1 + GW2). For clarity purposes, only models driven by ERA5+ERA5 are shown for 2001–2002 and 2004–2005.

**Table 2**
*Nash–Sutcliffe Efficiency (NSE) and Relative Bias (%) in Calibration and Validation Periods*

| | | | ERA5+ ERA5 | ERA5+ MODIS | CHELSA+ ERA5 | CHELSA+ MODIS | HAR+ MODIS |
|---|---|---|---|---|---|---|---|
| R | Calibration (2001–2005) | NSE | 0.84 | 0.92 | 0.57 | 0.59 | 0.90 |
| | | Bias | −2.3 | −3.4 | −20.9 | −21 | −5.9 |
| | Validation (1998–1999) | NSE | 0.80 | 0.69 | 0.77 | 0.73 | * |
| | | Bias | 6.4 | 9.0 | 20.1 | 18.8 | * |
| | Validation (2008–2013) | NSE | 0.68 | 0.83 | 0.75 | 0.74 | 0.78 |
| | | Bias | 17.0 | 14.5 | −20.5 | −23.0 | 12.6 |
| WBR-double | Calibration (2001–2005) | NSE | 0.76 | 0.73 | 0.83 | 0.76 | 0.62 |
| | | Bias | −1.5 | −1.8 | −0.4 | −3.0 | −0.42 |
| | Validation (1998–1999) | NSE | 0.74 | 0.62 | −2.05 | −1.04 | 0.54 |
| | | Bias | 1.4 | 6.1 | 29.7 | 22.7 | 5.3 |
| WBR-single | Calibration (2001–2005) | NSE | 0.32 | 0.33 | 0.48 | 0.35 | 0.53 |
| | | Bias | −9.6 | −6.5 | −10.1 | −10.3 | 5.8 |
| | Validation (1998–1999) | NSE | 0.58 | −0.05 | −3.9 | 0.29 | 0.27 |
| | | Bias | −4.3 | −12.8 | 34.5 | 6.42 | 11.3 |

*Note.* R, total discharge; WBR, winter (November–February) baseflow; -single, single groundwater layer; -double, two groundwater layers. Forcing combinations denoted as precipitation + temperature source: ERA5+ERA5, ERA5+MODIS, CHELSA + ERA5, CHELSA + MODIS, HAR + MODIS. *HAR lacking data before 2001. The regulation is assumed to not significantly impact the overall annual runoff magnitude (<10%), nor runoff seasonality, and total runoff is therefore included in the statistics. The modes were calibrated with R + SCF + GMB + WBR.

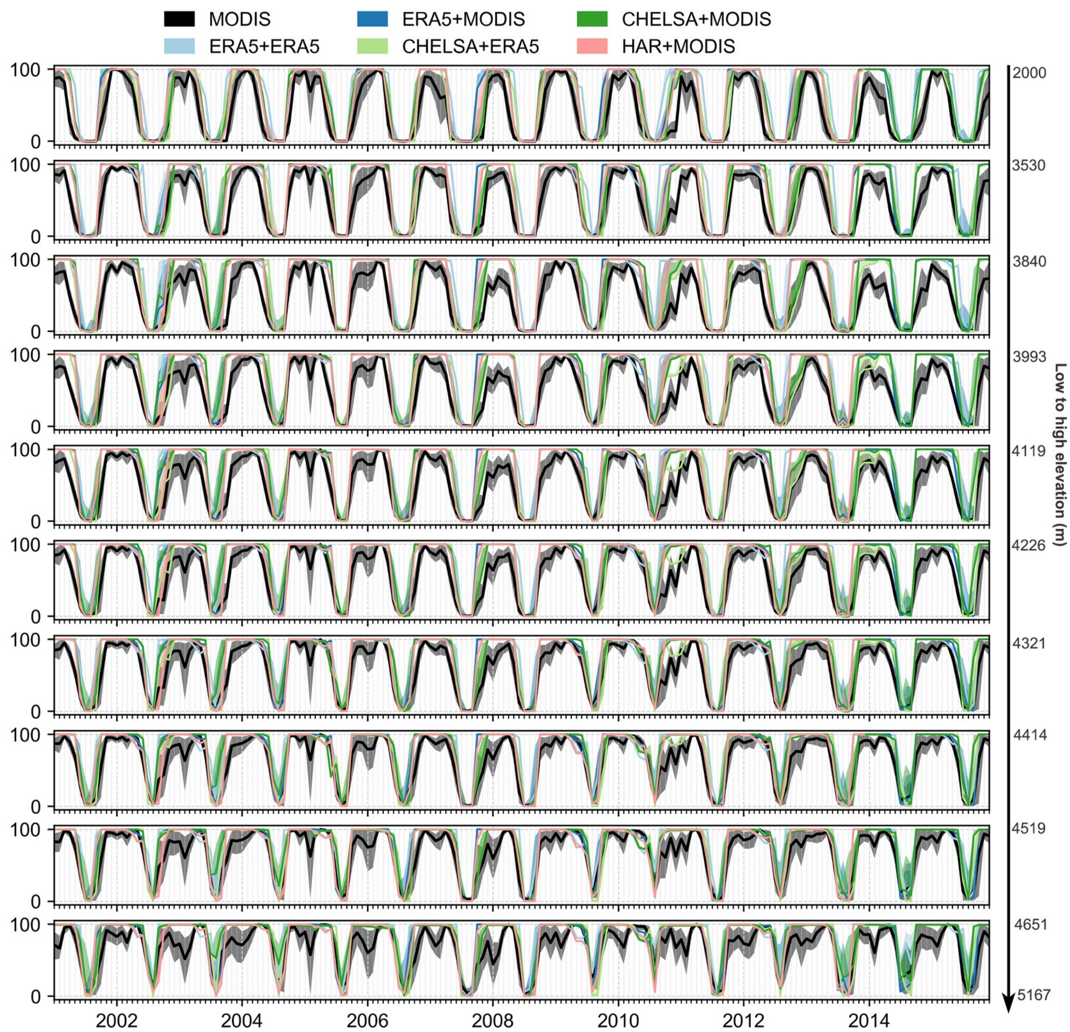


**Figure 4.** SCF comparison between model simulations and MODIS observations across 10 elevation bands from 2001 to 2016. Elevation bands range from 2,000 to 5,167 m, increasing from top to bottom (specific elevation zones are displayed on the right side of the figure). Analysis excludes glacier-covered grid cells due to MODIS limitations in distinguishing between snow and ice surfaces (Rondeau-Genesse et al., 2016). For model results, shaded areas show SCF ranges when using different snow depth thresholds (1, 5, and 10 mm) to define snow presence. For MODIS data (500-m resolution), where each pixel is originally classified as either snow-covered (1) or snow-free (0), the shaded areas represent SCF ranges when using different thresholds (25%, 50%, and 75% of snow-covered MODIS pixels) to aggregate to the model 1 km resolution.

(NSE: 0.90; bias: −5.9%). During validation, ERA5-precipitation-forced models maintained good performance (NSE: 0.62 to 0.74; bias: 1.4%–6.1%), while both HAR and CHELSA-precipitation-forced models showed unsatisfactory performance, with the latter showing even negative NSE values.

The dual-layer groundwater representation significantly improved WBR simulation capabilities (Table 2 and Figure 3c). The model with a single groundwater layer failed to capture the observed WBR pattern, showing much lower NSE (0.32–0.53) during calibration compared with that of the dual-layer version (NSE: 0.62–0.83). Figure 3c shows how only the dual-layer model captured the characteristic two-phase recession with a rapid initial decline from water retained in the coarse sediments followed by a slower recession representing gradual discharge from deeper groundwater sources. For all forcing combinations, baseflow is estimated to dominate total runoff from October to April (Figure S6 in Supporting Information S1). Annual baseflow contributions to total runoff, however, exhibit variability (Figure S7 in Supporting Information S1) driven by uncertainties in precipitation inputs (Figure S5 in Supporting Information S1).

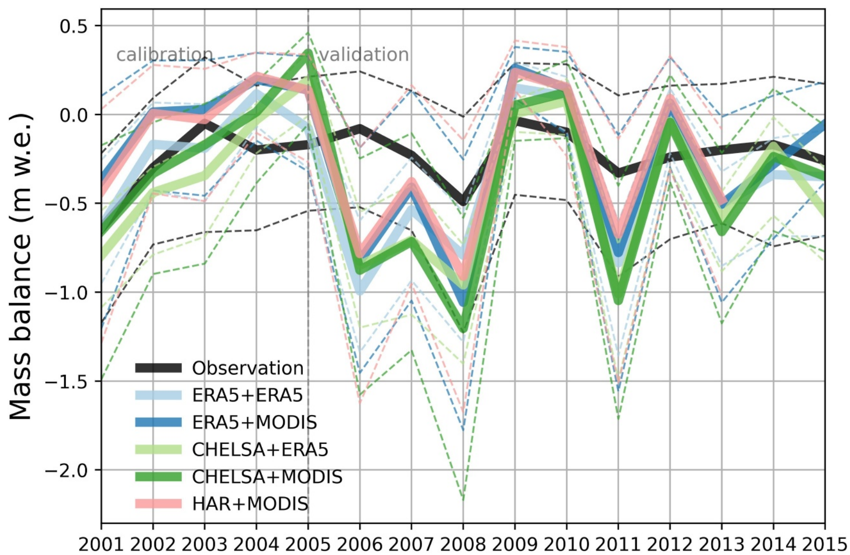


**Figure 5.** Comparison between modeled annual glacier mass balances and surface mass balance from Barandun et al. (2021). Solid lines show median values from different forcing combinations to minimize the influence of extreme values; dashed lines indicate their lower and upper quartiles. The black line represents the reference mass balance with uncertainty bounds (dashed black lines) from Barandun et al. (2021).

To verify the conclusion that winter runoff entirely originates from baseflow, we also quantify baseflow using the classical baseflow separation methods (see Section 4.6). Figure S8 in Supporting Information S1 shows the comparison between the reference winter baseflow in our study and the baseflow estimation from the different methods. The results show that, although the BN77 method indicates that not all points represent pure baseflow, the bias between winter baseflow estimates from various separation methods and the reference winter baseflow is within 5% (Table S2 in Supporting Information S1). The strong consensus among these methods justifies considering the winter runoff in the Gunt River basin as baseflow. One exception is the result from the Chapman method that assumes only 80% as baseflow (Table S2 in Supporting Information S1). Other studies have also noted that this method tends to underestimate baseflow during low-flow periods due to its fixed parameters (Helfer et al., 2025; Xie et al., 2020). We noticed that, although those methods generally agree well in terms of the overall bias, their capabilities in identifying runoff variability differ, with NSE ranging from 0.51 to 0.92 (Table S2 in Supporting Information S1). Such difference is not surprising, as different regions with varying hydro-climatic settings may require different baseflow separation approaches (e.g., Gonzales et al., 2009; Pelletier & Andréassian, 2020; Xie et al., 2020).

The model simulates the SCF well during both calibration (2001–2005) and validation (1998–1999, 2008–2013) periods, with similar correlation coefficients ranging from 0.71 to 0.90 (Figure 4). However, the model consistently overestimates SCF across elevation bands, with biases ranging from 5.2% to 19.7%. These overestimations are predominantly concentrated in the eastern regions and depend on the precipitation products (Figures S9–S13 in Supporting Information S1). The modeled catchment-scale GMB agrees well with the results in Barandun et al. (2021), with an average bias ranging from −0.28 to −0.08 m w.e. $yr^{-1}$ depending on the meteorological forcing combination (Figure 5). The best-performing forcing data sets were the ERA5 precipitation and temperature. The mass balances are in better agreement for the calibration period (mean bias: ∼ −0.14 m w.e. $yr^{-1}$) than for the validation period (mean bias: ∼ −0.20 m w.e. $yr^{-1}$) for all forcing combinations.

### 5.2. Performance of Different Model Calibration Strategies

The model was calibrated using different combinations of calibration data to explore their impact on model performance (Figure 6). The results show that the WBR calibration was essential for any model setup to accurately simulate the winter stream flow (Figure 6). The traditional R + SCF + GMB method yielded satisfactory results for R, SCF, and GMB (as shown in Table 3) but did not perform well for WBR. Consequently, it was

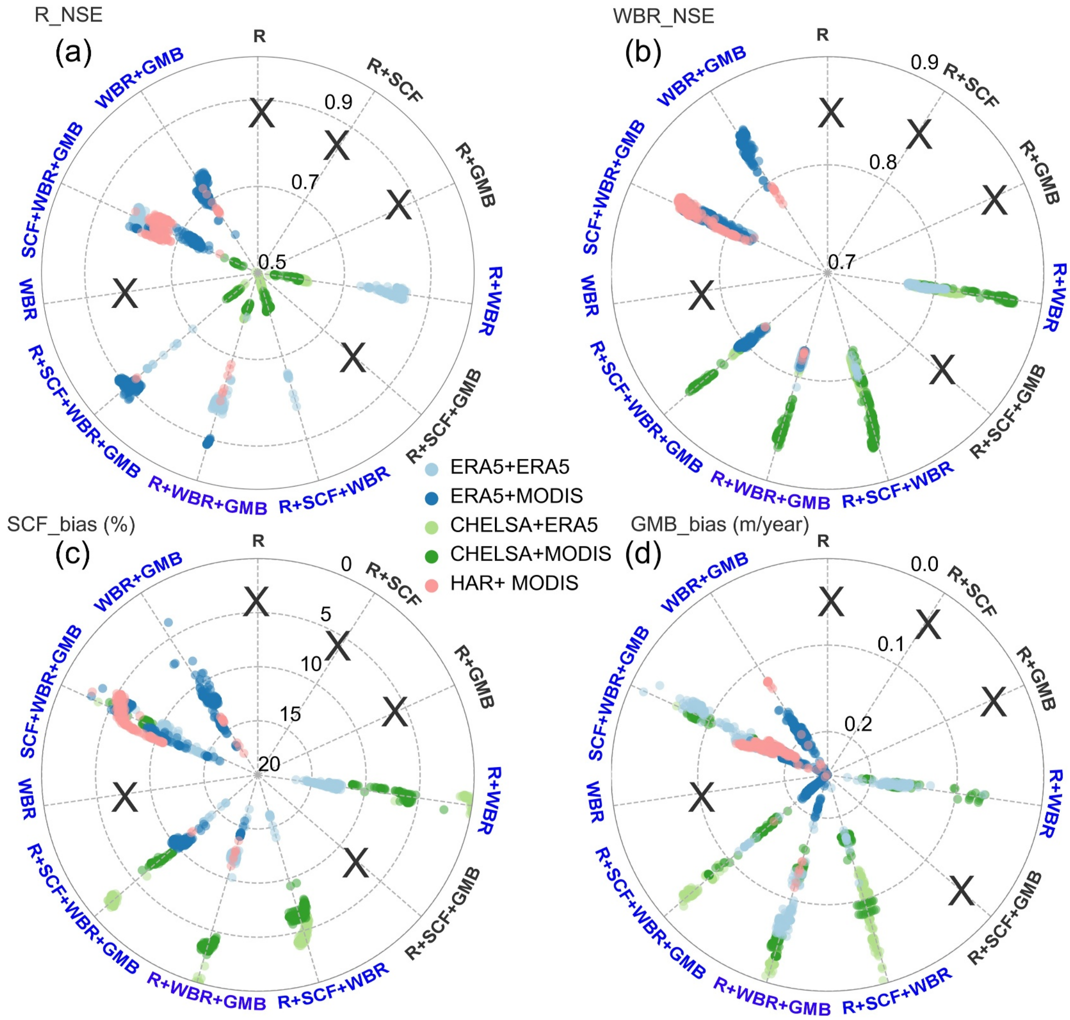


**Figure 6.** Model performances according to different objective functions for different data combinations in the calibration period. (a) R_NSE, NSE for total discharge; (b) WBR_NSE, NSE for winter baseflow; (c) SCF_bias, snow cover fraction bias; (d) GMB_bias, glacier mass balance bias. Calibration group names with WBR calibration are marked in blue; calibration strategies that failed in any of the other performance categories are indicated with a cross. Good minimum performances are presented as "standard" in Table 3.

categorized as a "failed" calibration strategy. Calibration strategies for optimal model performance depend on the forcing data set. For example, for ERA5 + ERA5, good model performance can be achieved by combining WBR with any other calibration variable. However, for ERA5 + MODIS, to achieve optimal model performance, GMB and WBR must be used simultaneously as calibration variables. These results highlight the fundamental role of WBR calibration as a constraint for model performance, though the choice of additional calibration data depends on the forcing data set used.

**Table 3**
*Standards of Overall Good Model Performance*

| R | | WBR | | SCF1…10 | GMB |
|---|---|---|---|---|---|
| NSE | RB (%) | NSE | RB (%) | AB (%) | AB (m/y) |
| >0.5 | <25 | >0.75 | <10 | <20 | <0.25 |

*Note.* RB, relative bias; AB, absolute bias. If a strategy does not meet the performance criteria in any single category, it is classified as a "failed" calibration strategy. The standard for R is lower than that for WBR because it was the highest achievable one with CHELSA data.

### 5.3. Impact of Different Calibration Strategies on Runoff Component Quantification

WBR calibration effectively constrained the uncertainty in runoff component quantification (Figure 7). The uncertainty range, defined as the spread between maximum and minimum values estimated by each forcing data set, is substantial without WBR calibration: 34.2%–60.9% for snowmelt, 5.0%–17.3% for glacier runoff, and 33.4%–50.4% for baseflow (Table 4). Incorporating WBR calibration, especially when combined with any other

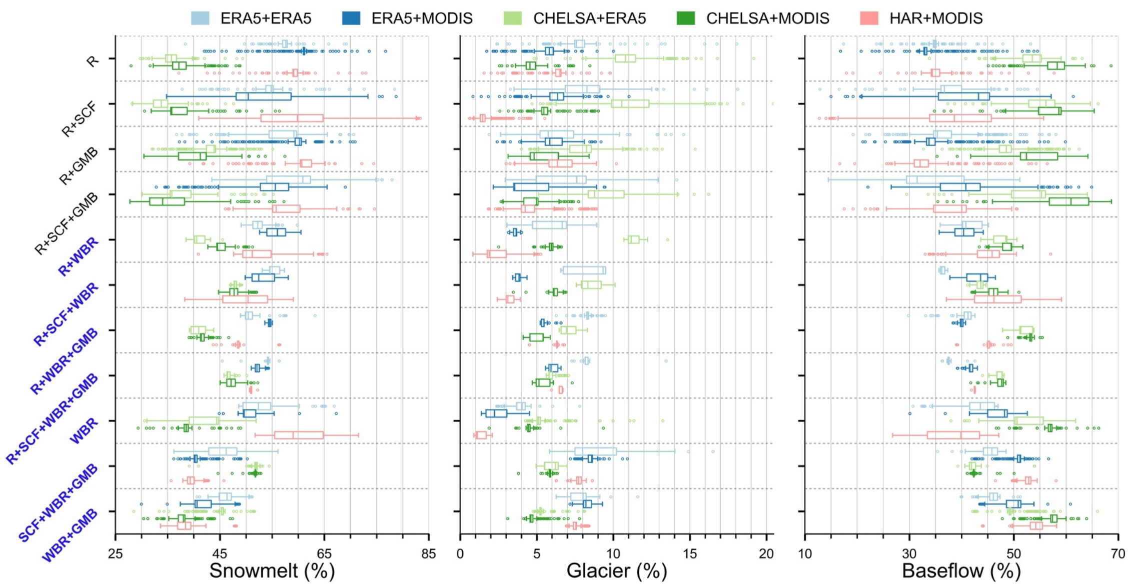


**Figure 7.** Box plots of major runoff component contributions with different calibration strategies during the time period 2001–2005. Groups with R + WBR calibration are highlighted in blue. Model runs are included when meeting their respective calibration objectives (Table 3). Rainfall runoff contribution was negligible (<3%), and we thus excluded this from further analysis.

calibration objective (R, SCF, or GMB), substantially narrowed these uncertainty ranges to 8.2%–21.2%, 2.6%–10.7%, and 6.8%–21.3%, respectively (Figure 7; Table 4).

While runoff component quantification inherently depended on forcing inputs (Figure 7), WBR calibration substantially mitigated these input-dependent variations, in particular for snowmelt and baseflow estimates. For example, before WBR calibration, median estimates across different forcing data sets varied widely, with snowmelt contributions ranging from 35.0% to 60.9% and baseflow contributions from 33.4% to 58.3% (Table 4);

**Table 4**
*Runoff Contribution to Total Runoff Estimated With Different Calibration Strategies*

| | | | R | R + SCF | R + GMB | R + SCF + GMB | R + WBR | R + SCF + WBR | R + WBR + GMB | R + SCF + WBR + GMB | WBR | SCF + WBR + GMB | WBR + GMB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Snowmelt | Range | 38.3 | 43.9 | 34.2 | 45.2 | 11.5 | 8.2 | 14.3 | 10.9 | 22.4 | 20 | 21.2 |
| E | (%) | Median | 60.9 | 54.8 | 59.6 | 55.9 | 53.6 | 52.7 | 50.5 | 52.1 | 50.7 | 42.6 | 40.7 |
| R | Glacier | Range | 16.3 | 17.3 | 12.7 | 12 | 5.9 | 6.1 | 4.3 | 7.9 | 6.4 | 10.7 | 5.3 |
| A | (%) | Median | 6.1 | 6.6 | 5.8 | 5.0 | 3.7 | 3.9 | 8.3 | 5.9 | 3.1 | 8.6 | 8.2 |
| 5 | Baseflow | Range | 37.7 | 48.5 | 40.6 | 50.4 | 21.3 | 10.8 | 12.7 | 6.8 | 22.5 | 21.3 | 21 |
| | (%) | Median | 33.4 | 36.8 | 34.9 | 40.3 | 40.8 | 43.4 | 41.2 | 41.7 | 46 | 47.1 | 50.6 |
| C | Snowmelt | Range | 21.4 | 28.2 | 27 | 29.4 | 12.7 | 7.3 | 7.6 | 7.2 | 22.6 | 15.3 | 23.1 |
| H | (%) | Median | 36.3 | 35 | 43.8 | 35.3 | 42.7 | 47.8 | 41.4 | 46.7 | 39.9 | 51.8 | 38.1 |
| E | Glacier | Range | 16.6 | 17.9 | 12.2 | 13.6 | 11 | 6.6 | 4.2 | 2.6 | 11.3 | 3.2 | 5 |
| L | (%) | Median | 10 | 9.9 | 8.1 | 5 | 10.8 | 6.5 | 5.4 | 5.8 | 5.0 | 5.9 | 4.7 |
| S | Baseflow | Range | 34.8 | 31.3 | 31.5 | 39.2 | 8.4 | 9.8 | 7.6 | 6.7 | 28.3 | 14.7 | 23.9 |
| A | (%) | Median | 55.3 | 56.4 | 49 | 58.3 | 48.6 | 44.5 | 53.1 | 47.7 | 53.9 | 42.3 | 57.3 |
| | Snowmelt | Range | 48.7 | 55.2 | 44 | 50.2 | 27 | 20.8 | 24.1 | 11.3 | 42.2 | 20.5 | 23.1 |
| H | (%) | Median | 57.7 | 52.5 | 55 | 53.8 | 51.4 | 48.3 | 48.5 | 51.6 | 49.6 | 41.5 | 38.7 |
| A | Glacier | Range | 17.6 | 19.9 | 13.4 | 14.4 | 12.7 | 7.7 | 5.4 | 8.7 | 12.3 | 12.7 | 8.5 |
| R | (%) | Median | 6.3 | 5.9 | 6.8 | 5.0 | 3.7 | 6.0 | 6.7 | 5.9 | 3.6 | 7.8 | 5.4 |
| | Baseflow | Range | 51.8 | 53.4 | 47.5 | 54.2 | 24 | 23.5 | 25.6 | 12.2 | 39.5 | 22.7 | 26.1 |
| | (%) | Median | 35.1 | 45.2 | 38.1 | 41.7 | 44.5 | 44.5 | 45.2 | 42.4 | 46.5 | 50.5 | 55.2 |

*Note.* The first column represents different precipitation inputs. Median and range (max - min) from Figure 7 show model results under different forcing inputs and calibration data.

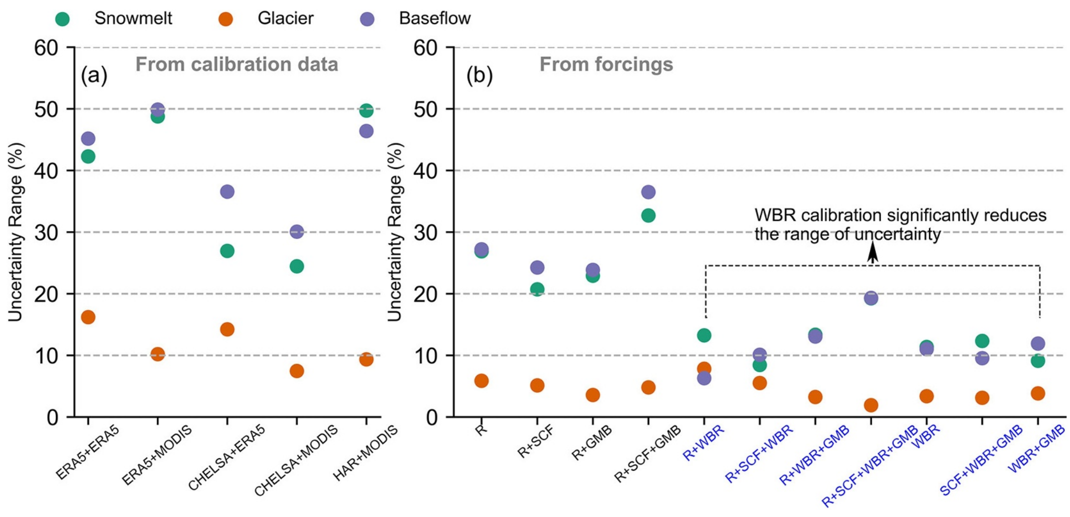


**Figure 8.** Runoff component uncertainty range induced by (a) calibration strategy and (b) forcing inputs. Uncertainty range refers to the difference between the estimated maximum and minimum values of runoff components.

but, after incorporating WBR with other calibration data, these ranges converged to 38.1%–53.6% for snowmelt and 40.8%–57.3% for baseflow, reducing the uncertainty by 40.2% and 33.7%, respectively (Table 4).

To further quantify the uncertainties resulting from either calibration strategy or forcing data, Figure 8 shows the relative differences for each category. The specific methods for quantifying these uncertainties are described in Text S1 of Supporting Information S1. The results suggest that the choice of a particular calibration strategy can introduce larger uncertainties in runoff component estimation than those introduced by differences in the forcing data. Traditional calibration approaches (R + SCF + GMB) constrain these uncertainties less effectively than approaches incorporating WBR calibration. WBR calibration consistently reduces uncertainty ranges for snowmelt runoff (21%–33% to 9%–13%) and baseflow (24%–38% to 7%–19%). Differences in glacier melt uncertainty are mainly improved by incorporating GMB calibration. These findings indicate that, while the quality of forcing data is crucial in hydrological modeling, the calibration strategy is at least equally important for reducing uncertainty in the runoff component estimates.

## 6. Discussion

### 6.1. WBR Calibration

Incorporation of WBR as an additional calibration variable has been proven significantly beneficial for constraining model parameter ranges across different forcing data sets. As outlined in Terink et al. (2015), SPHY balances the surface runoff and groundwater recharge to achieve a desired baseflow (Text S4 in Supporting Information S1). By constraining the amount of baseflow in the calibration process, the model requires to adjust percolation water either through liquid precipitation or snow melt, leading to, for example, different SCF representations between different forcing data sets. The WBR calibration substantially narrows the uncertainty ranges for certain GW parameters and those affecting meltwater infiltration rates, notably reducing the uncertainty of DeltaGW1 by over 90% and SplitGW by 30%–90% (Figures S14–S18 in Supporting Information S1). However, some other parameter ranges were not significantly improved (Figures S14–S18 in Supporting Information S1). The overall uncertainty reduction (Figures 7 and 8) might indicate either a low sensitivity of such parameters (Table S1 in Supporting Information S1) or equifinality related to snow accumulation and melt processes, which are driven by the precipitation and temperature correction factors (PCF, TCF) that indirectly control the groundwater recharge from melt. The high uncertainty of those parameters can also be related to the limited calibration data, the simplified model, and the choice of the objective function (Cui, Li, & Tian, 2021; Cui, Sreekanth, et al., 2021; Her & Chaubey, 2015; Smith et al., 2022; Tague et al., 2004; Wagener & Gupta, 2005). Despite some of the parameters not being narrowed down, the calibration strategy attests additional value over classical calibration strategies for better constrained runoff component estimates (Figures 7 and 8), although with limited value for deriving a robust and transferable model. For the latter, the dependence on forcing data still appears to significantly influence the derivation of the model parameters.

Note that WBR alone is insufficient for good model performance, as its calibration alone shows limited effectiveness in constraining parameter uncertainty (Figures S14–S18 in Supporting Information S1), and thus must be combined with other data types to calibrate the model. Combining WBR with total runoff is the most effective way to reduce parameter uncertainty. Calibrating on total runoff can lead to equifinality, as matching the total flow is sufficient without constraining the processes involved. By adding WBR as an individual runoff component in the calibration process, we force the model structure to account for both meltwater infiltration for groundwater storage recharge and a dual-layer groundwater discharge that produces the characteristic fast and slow recession baseflow. As groundwater parameters directly influence how precipitation translates into baseflow over months to years (Pohl et al., 2017), the WBR calibration affects multiple calibration parameters simultaneously, which in combination lead to a specific baseflow behavior. This also explains, to some degree, why different forcing data produce different parameter sets. However, these parameter sets seem to be consistent with their respective forcing data. It is striking that the WBR calibration is efficient for the parameter estimation even over such a short calibration period as the 5-year one used in this study (2001–2005), as well as for the forcing data sets with evident data issues.

As Pohl et al. (2015) demonstrated for the Gunt catchment, even inaccurate precipitation data can reproduce total discharge if basic seasonality is maintained; the only requirement is to adjust the precipitation intensity to satisfy the water balance. Her et al. (2019) further suggest that the uncertainties from model parameters can be larger than the uncertainties from the forcing inputs for the quantification of slow components such as soil moisture and groundwater. Therefore, even with the anomalously low precipitation in 2001 (50 vs. 300 mm average) for CHELSA, the CHELSA-forced models still can achieve satisfactory WBR performance during the calibration period. However, validation results show significant degradation in WBR simulation performance. In addition to the well-known forcing data uncertainty problem, the reduced performance during the validation period can also be attributed to several other factors, including overfitting during calibration (e.g., Her & Chaubey, 2015; Zhao et al., 2019), limitations in calibration data quality (e.g., Montanari & Di Baldassarre, 2013), and the choice of the objective function (e.g., Her & Seong, 2018). The impact of each aspect on this study is discussed in Text S2 of Supporting Information S1.

With long-lasting snow covers (Immerzeel et al., 2009) and increasing dominance of snow- and glacier-fed streams toward the western end of the Himalayas, hydrological study results can differ significantly regarding the importance of runoff components (Azam et al., 2021). The WBR calibration strategy appears promising for exploiting these nival–glacial discharge patterns and deriving much better constrained estimates of runoff components. Even in the absence of continuous discharge estimates, we have shown that WBR + GMB can already yield more constrained results than classical calibration strategies. As the baseflow recession can be determined (at least for the slow recession component) at a monthly scale, the WBR calibration should allow for using historic data records for Central Asian rivers available at the monthly temporal resolution (e.g., GRDC; Elmi et al., 2024). In combination with an increasing availability of satellite-derived geodetic GMB (e.g., Barandun et al., 2021; Hugonnet et al., 2021), snowline altitude (Barandun et al., 2021; Loibl et al., 2025), and gravimetric terrestrial water storage data (e.g., from the Gravity Recovery and Climate Experiment (GRACE)), additional boundary conditions and constraints can be implemented. The complementary data on snow accumulation and melt should help to further increase model robustness due to snowmelt's paramount importance on groundwater recharge in cold regions (Jeelani et al., 2021; Mohammed et al., 2019; Tiel et al., 2024), and the terrestrial water storage can further constrain the water balance at the mountain range scale (Bai et al., 2018; Lo et al., 2010; Sun et al., 2010).

### 6.2. Runoff Component Quantification

Only two similar studies conducted in the study area (Pohl et al., 2015; Tarasova et al., 2016) are available for direct comparison with the present work. Taking the median of the results from the R + SCF + WBR + GMB calibration as the baseline, the estimated contributions of snowmelt runoff, glacier runoff, and baseflow to total runoff from 2001 to 2005 are 46.7%–52.7%, 5.8%–5.9%, and 41.7%–47.7%, respectively (Table 4). Using HAR + MODIS forcing data, Pohl et al. (2015) estimated lower snowmelt (28%) and baseflow (38%) contributions but much higher glacier runoff (34%). Pohl et al. (2015) did not incorporate GMB data to constrain the modeled GMB. On the other hand, the higher glacier contribution in Pohl et al. (2015) probably results somewhat from snow melt on the glacier surface being included in the glacier meltwater contribution, whereas we separate the snow melt on ice from the glacier meltwater contribution and add this to the snow runoff component instead. Armstrong et al. (2018) calculated the contribution of ice and snow melt on the glacier separately and showed that, on an annual scale, snow melt on ice often exceeds the contribution of ice melt (up to seven times). However, in their study, the authors report that exposed ice melt contributes 8% to the annual runoff of the Amu Darya basin, while snow on ice contributes only 4%. Furthermore, without constraining GMB, Pohl et al. (2015) also modeled significantly more negative GMB for some of their forcing combinations. For comparison, Tarasova et al. (2016) estimated a contribution of glacier runoff (without snow melt on ice) to total runoff of 5.8%–14.3% using R + GMB calibration in a multi-objective function; these results are much closer to those presented here. The GMB in the present work shows a dominant negative bias over the whole study period for most forcing combinations when compared with the GMB reported in Barandun et al. (2021). Therefore, the glacier meltwater could be overestimated in our study.

All forcing data sets tend to overestimate the SCF in the eastern part of the Gunt basin during May and June (Figures S9–S13 in Supporting Information S1), suggesting the potential bias in the modeled seasonal distribution of the snow melt runoff. The SCF overestimation likely stems from two key factors: first, the Pamir region exhibits a strong negative precipitation gradient, with annual total precipitation decreasing from west to east due to westerly dominated moisture transport being progressively depleted as air masses traverse the mountain ranges (Baldwin & Vecchi, 2016). While ERA5 and HAR generally capture this pattern, CHELSA shows an opposite pattern with higher precipitation in the east (Figure S19 in Supporting Information S1). Second, in these eastern dry areas, thin and discontinuous snow cover may not be well detected by MODIS (Hall & Riggs, 2007). MODIS snow detection is limited in the winter months over the high mountain regions, for example, due to low solar illumination angles and extensive cloud cover (Dietz et al., 2012; Rittger et al., 2013). Adjusting the depth threshold of the modeled snow cover changes the model's performance. For example, the bias is reduced by 2%–7% when increasing the snow depth threshold from 5 to 20 mm for identifying snow presence in the Gunt basin (Figure S20 in Supporting Information S1).

We estimate that groundwater discharge contributes 41.7%–47.7% to annual total streamflow. The significant groundwater contribution could provide resilience against climate change impact because groundwater acts as a stable, slow-release reservoir (Cochand et al., 2019; Hayashi, 2020; Somers & McKenzie, 2020). The loss of glacier cover jeopardizes their buffering capacity on streamflow during extreme meteorological conditions (Pohl et al., 2017). However, as shown in the Canadian Rocky Mountains, groundwater reservoirs might increase their storage capacity in regions that experience glacier loss (Castellazzi et al., 2019). More accurate groundwater representation is thus crucial, since geological properties strongly control streamflow response to climate warming—higher permeability systems can better maintain summer flows, but they risk faster groundwater depletion during droughts (Tague et al., 2004). While permafrost thaw may increase groundwater contributions in some regions (Walvoord & Kurylyk, 2016), changing precipitation patterns and earlier snow melt could ultimately reduce recharge (Cuthbert et al., 2019; Meixner et al., 2016; Taylor et al., 2013). Future work comparing WBR-constrained versus unconstrained projections could help quantify uncertainty in predictions of the groundwater's stabilizing role for mountain water resources, which is increasingly critical as climate change continues to alter the timing and magnitude of different water sources (van Tiel et al., 2020).

### 6.3. Model Equifinality

The reduction in variability of individual runoff components using WBR calibration (Sections 5.1 and 6.2) contrasts with a significant variability in parameter value ranges across models with different forcing data (Figures S13–S17 in Supporting Information S1). We attribute such disparity primarily to the outlined inherent dependence of groundwater recharge on snowmelt (Section 6.1) that prioritizes calibration of parameters leading to good groundwater recharge and subsequent discharge representation (e.g., CHELSA models compared with ERA5 models prioritize lower infiltration with higher field capacity; Figure S16 in Supporting Information S1). As this parameterization affects snowmelt infiltration into the ground in years with biased precipitation estimates, both over- and underestimation, the groundwater parameterization reflects a compromise for the calibration period that might fail in the validation period, for which precipitation estimates might differ significantly. Therefore, the model may perform poorly during the validation phase compared with the calibration phase. We did not test how the choice of calibration and validation periods would affect the model performance because the available period in which we were able to compare all the different forcing setups is overall short, and also considering the constraints of river flow regulations in place after 2005.

In our case, the inaccurate forcing data for the validation period shows higher deviation compared with the assumed more accurate forcing estimates (Table 2). Interestingly, the different precipitation forcing data (ERA5 and HAR) do show similar parameter estimates and a clear overall improvement in equifinality compared with the calibration without WBR (Figures S15 and S18 in Supporting Information S1). These results suggest that equifinality reduction with WBR calibration relies on some threshold quality of the forcing data, as said reduction was achieved with the different estimates for HAR and ERA5 but not with CHELSA. Besides forcing data, equifinality could also respond to the calibration's dependence on winter baseflow at the basin outlet, which overlooks spatial baseflow distribution (Cui, Li, & Tian, 2021; Cui, Sreekanth, et al., 2021; Smith et al., 2022; Tague et al., 2004; Wagener & Gupta, 2005). The simplified groundwater module may also produce parameters (e.g., DeltaGW1 and DeltaGW2) that are related to offset structural and input uncertainties (Cui, Li, & Tian, 2021; Cui, Sreekanth, et al., 2021).

Although the WBR calibration has overall reduced the model's equifinality issues, the effectiveness of WBR calibration can be influenced by the parameter space, as previous studies have shown that the number of calibration parameters significantly impacts model equifinality (e.g., Arsenault & Brissette, 2014; Casado-Rodríguez & del Jesus, 2022; Her & Chaubey, 2015). Furthermore, the uncertainties from model parameters can be larger than the uncertainties from the forcing inputs for the quantification of slow components such as soil moisture and groundwater (Her et al., 2019). To explore how model equifinality varies with parameter space in groups with and without WBR calibration, we conducted sensitivity experiments following the framework proposed by Her and Chaubey (2015).

The results suggest that the model equifinality generally increases with the number of parameters, though this trend becomes less pronounced when calibrating more than 40 parameters (Figure S21 and Table S3 in Supporting Information S1). The conclusion is consistent with the results from Her and Chaubey (2015), as the additional parameters with decreasing sensitivity have limited influence on the model behavior. Notably, the WBR calibration group exhibited 30%–46% lower equifinality compared with that of the conventional calibration group (R + SCF + GMB). However, in groups without WBR calibration, WBR simulations remain poor, regardless of the number of parameters. Even when all other calibration data were included in the calibration, the best WBR NSE achieved was only 0.56, below the threshold for "good performance" (Table 3). This is primarily due to WBR's small magnitude relative to total runoff, which limits its influence on overall model bias when not weighted (Equation 1). Consequently, groups without WBR calibration fail to improve WBR performance and thus do not provide effective constraints on the model.

### 6.4. Uncertainties

In addition to the known uncertainties in hydrological modeling, such as model structure (e.g., constant hydraulic conductivity in our model), forcing, calibration data, and parametrization (see Text S3 in Supporting Information S1 for details; Cui, Li, & Tian, 2021; Cui, Sreekanth, et al., 2021; Gupta & Govindaraju, 2019; Her et al., 2019; Her & Chaubey, 2015; Shin et al., 2023), one of the primary sources of uncertainty in the present study is the accuracy of the reference winter baseflow. As we use a definition of baseflow (see Section 4) that resembles the structure of our model, that is, baseflow results only from groundwater discharge, we do not account for other possible sources of baseflow. Likewise, the observational stream flow data might include localized melting of snow or ice (e.g., due to solar radiation, geothermal heat, or friction in streams) that can contribute to streamflow (Kapil et al., 2010; Liston & Winther, 2005), and the Yashilkul Lake in the Gunt basin can also contribute surface flow if the lake is not completely frozen. In addition, the quality of the observed data itself may also introduce uncertainty. For example, the winter baseflow recession curves for January–February in 1998 and 1999 are nearly identical, particularly in the rapid recession phase (Figure 3b), suggesting potential data quality issues. However, these errors are difficult to quantify.

For the interpretation of our results and the comparison with other studies (Section 6.2), the definition of baseflow used in this study is well-suited as all comparisons are made with other modeling studies that follow the same definition. By assuring a close representation of winter discharge with the WBR calibration, we can derive GW parameters that provide more accurate and reliable estimates at any other time of the year. Thus, the WBR calibration approach should improve the estimation of overall contribution of groundwater discharge to total runoff compared with that obtained with classical calibration without WBR.

Furthermore, considering the small bias between the baseflow obtained by different baseflow separation methods and the observed winter runoff (Section 5.1), using winter runoff for the WBR calibration approach seems justified in our case. In situations where liquid precipitation or other contributions to winter runoff perturb the stream flow signal, using a baseflow separation method, as described above, could be highly beneficial and allow the WBR calibration approach to work also under such conditions.

## 7. Conclusions

Nival–glacial-dominated hydrological systems in remote mountain regions offer significant challenges for hydrological modeling. Lack of data for model forcing, calibration, and validation, in particular, can lead to model equifinality. However, in the case of the Gunt River basin in the Pamir, we show that winter discharge can be a proxy of winter baseflow, which can be integrated into model calibration to improve model robustness. By representing the groundwater system as a dual-layer component with a fast and a slow recession and by incorporating winter baseflow as an independent weight in the objective function, not only runoff component contributions to total runoff and several GW model parameter estimates can be well constrained but also the variability of various other model parameters can be reduced. These results are not independent of the quality of the forcing data sets, but even different forcing datasets—if they provide somewhat reasonable estimates—seem to yield similar parameter ranges and reduce equifinality significantly.

The results obtained in the present work lead to the following conclusions:

1. The proposed calibration method improves overall NSE from 0.32–0.53 to 0.62–0.83 during the calibration period. The modified model effectively captures both rapid flow through coarse sediments (GW1) and slow discharge from deep aquifers (GW2), providing a more realistic representation of groundwater processes in mountain environments. This improvement is particularly evident in the simulation of the characteristic two-phase recession curve observed in the winter baseflow.
2. When winter baseflow is combined with traditional calibration variables (runoff, snow cover fraction, and glacier mass balance), this approach substantially narrows down uncertainty ranges for snowmelt (34.2%–60.9% to 8.2%–21.2%), glacier melt (5.0%–17.3% to 2.6%–10.7%), and baseflow (33.4%–50.4% to 6.8%–21.3%).
3. While forcing data quality is essential, the choice of the calibration strategy plays an even more critical role in reducing uncertainties in runoff component quantification. Winter baseflow calibration consistently reduces forcing-induced uncertainty in runoff component estimation, indicating that careful selection of calibration data combinations is paramount for achieving robust hydrological simulations.
4. The winter baseflow calibration approach relies on baseflow observations, which, in high alpine nival–glacial dominated hydrological regimes, can often be directly derived from discharge time series in winter. When this signal is obscured, baseflow separation methods can help to constrain these estimates but should be carefully selected as differences are apparent. These considerations might allow the application of the method also in catchments with less dominant nival–glacial hydrological regimes.

This study innovatively uses winter baseflow, often overlooked in high-mountain hydrological modeling, as a calibration constraint. Winter baseflow calibration substantially reduces uncertainties in runoff component estimates, providing a novel framework for model calibration in high-mountain regions worldwide. Improved accuracy in runoff component estimates could enhance future runoff predictions, supporting adaptive water management under climate change. It should be noted that winter baseflow estimation is subject to high uncertainties. Future research should focus on enhancing the precision of baseflow separation and developing rigorous objective criteria for selecting region-specific separation methods.

## Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

## Data Availability Statement

The data and code that support the findings of this study are available from the corresponding author upon request. The meteorological data for HAR can be obtained from the Technical University of Berlin via https://data.klima.tu-berlin.de/HAR/v1/ (Maussion et al., 2014); CHELSA data can be obtained via https://doi.org/10.16904/envidat.228.v2.1 (Karger et al., 2021); and ERA5-Land data (Hersbach et al., 2020) can be obtained via the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/). The mass balance data are published in Barandun et al. (2021). Debris cover information (Scherler et al., 2018) was obtained via https://doi.org/10.5880/GFZ.3.3.2018.005. Glacier outlines have been taken from the Randolph Glacier Inventory accessible at https://www.glims.org/RGI/. MOD10A2 and MOD11C1 data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (lpdaac.usgs.gov). We thank the State Administration for Hydrometeorology of Tajikistan for their cooperation and for providing the discharge data. Soil data from HiHydroSoil v2.0 (Simons et al., 2020) can be obtained via https://www.futurewater.eu/projects/hihydrosoil/, and land use data (Defourny et al., 2009) can be obtained from https://due.esrin.esa.int/page_globcover.php. Discharge data were provided by the State Administration for Hydrometeorology of Tajikistan and are available upon request.

**Acknowledgments**

This study was supported by Horizon Europe—Research and Innovation—European Union through the project "Water Efficient Allocation in a Central Asian Transboundary River Basin" (WE-ACT; contract no. 28000074) and by the Swiss Agency for Development and Cooperation (SDC) through the project "Cryospheric Observation and Modelling for Improved Adaptation in Central Asia" (CROMO-ADAPT; contract no. 81072443). We thank the project "Strengthening the resilience of Central Asian countries by enabling regional cooperation to assess glacio-nival systems to develop integrated methods for sustainable development and adaptation to climate change" funded Global Environment Facility / United Nations Development Programme / United Nations Educational, Scientific and cultural Organization (GEF/UNDP/UNESCO, contract no. 4500484501). Discharge data were obtained from the Administration for Hydrometeorology of Tajikistan http://www.meteo.tj/ as part of the BMBF (Federal Ministry of Education and Research) research programme PAMIR (FKZ 03G0815) for the Gunt River. We thank Ana R. Crespo for her suggestions on manuscript structure and editing. Open access publishing facilitated by Universite de Fribourg, as part of the Wiley - Universite de Fribourg agreement via the Consortium Of Swiss Academic Libraries.

## References

Andermann, C., Longuevergne, L., Bonnet, S., Crave, A., Davy, P., & Gloaguen, R. (2012). Impact of transient groundwater storage on the discharge of Himalayan rivers. *Nature Geoscience*, *5*(2), 127–132. https://doi.org/10.1038/ngeo1356

Armstrong, R. L., Rittger, K., Brodzik, M. J., Racoviteanu, A., Barrett, A. P., Khalsa, S.-J. S., et al. (2018). Runoff from glacier ice and seasonal snow in high Asia: Separating melt water sources in river flow. *Regional Environmental Change*, *19*(5), 1249–1261. https://doi.org/10.1007/s10113-018-1429-0

Arsenault, K. R., Houser, P. R., & De Lannoy, G. J. M. (2014). Evaluation of the MODIS snow cover fraction product. *Hydrological Processes*, *28*(3), 980–998. https://doi.org/10.1002/hyp.9636

Arsenault, R., & Brissette, F. P. (2014). Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches. *Water Resources Research*, *50*(7), 6135–6153. https://doi.org/10.1002/2013WR014898

Azam, M. F., Kargel, J. S., Shea, J. M., Nepal, S., Haritashya, U. K., Srivastava, S., et al. (2021). Glaciohydrology of the Himalaya-Karakoram. *Science*, *373*(6557), eabf3668. https://doi.org/10.1126/science.abf3668

Bachman, L. J. (1998). *Ground-water discharge and base-flow nitrate loads of nontidal streams, and their relation to a hydrogeomorphic classification of the Chesapeake Bay Watershed, middle Atlantic coast*. US Department of the Interior, US Geological Survey. https://doi.org/10.3133/wri984059

Bai, P., Liu, X., & Liu, C. (2018). Improving hydrological simulations by incorporating GRACE data for model calibration. *Journal of Hydrology*, *557*, 291–304. https://doi.org/10.1016/j.jhydrol.2017.12.025

Baldwin, J., & Vecchi, G. (2016). Influence of the Tian Shan on arid extratropical Asia. *Journal of Climate*, *29*(16), 5741–5762. https://doi.org/10.1175/JCLI-D-15-0490.1

Barandun, M., Fiddes, J., Scherler, M., Mathys, T., Saks, T., Petrakov, D., & Hoelzle, M. (2020). The state and future of the cryosphere in Central Asia. *Water Security*, *11*, 100072. https://doi.org/10.1016/j.wasec.2020.100072

Barandun, M., & Pohl, E. (2023). Central Asia's spatiotemporal glacier response ambiguity due to data inconsistencies and regional simplifications. *The Cryosphere*, *17*(3), 1343–1371. https://doi.org/10.5194/tc-17-1343-2023

Barandun, M., Pohl, E., Naegeli, K., McNabb, R., Huss, M., Berthier, E., et al. (2021). Hot spots of glacier mass balance variability in Central Asia [Dataset]. *Geophysical Research Letters*, *48*(11), e2020GL092084. https://doi.org/10.1029/2020GL092084

Barker, D., Huang, X.-Y., Liu, Z., Auligné, T., Zhang, X., Rugg, S., et al. (2012). The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA. *Bulletin of the American Meteorological Society*, *93*(6), 831–843. https://doi.org/10.1175/BAMS-D-11-00167.1

Beck, H. E., van Dijk, A. I. J. M., Miralles, D. G., de Jeu, R. A. M., Bruijnzeel, L. A., McVicar, T. R., & Schellekens, J. (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research*, *49*(12), 7843–7863. https://doi.org/10.1002/2013WR013918

Becker, M. W., Georgian, T., Ambrose, H., Siniscalchi, J., & Fredrick, K. (2004). Estimating flow and flux of ground water discharge using water temperature and velocity. *Journal of Hydrology*, *296*(1), 221–233. https://doi.org/10.1016/j.jhydrol.2004.03.025

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Boughton, W. (2004). The Australian water balance model. *Environmental Modelling & Software*, *19*(10), 943–956. https://doi.org/10.1016/j.envsoft.2003.10.007

Brutsaert, W., & Nieber, J. L. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resources Research*, *13*(3), 637–643. https://doi.org/10.1029/WR013i003p00637

Carroll, R. W. H., Niswonger, R. G., Ulrich, C., Varadharajan, C., Siirila-Woodburn, E. R., & Williams, K. H. (2024). Declining groundwater storage expected to amplify mountain streamflow reductions in a warmer world. *Nature Water*, *2*(5), 419–433. https://doi.org/10.1038/s44221-024-00239-0

Casado-Rodríguez, J., & del Jesus, M. (2022). Hydrograph separation for tackling equifinality in conceptual hydrological models. *Journal of Hydrology*, *610*, 127816. https://doi.org/10.1016/j.jhydrol.2022.127816

Castellazzi, P., Burgess, D., Rivera, A., Huang, J., Longuevergne, L., & Demuth, M. N. (2019). Glacial melt and potential impacts on water resources in the Canadian Rocky Mountains. *Water Resources Research*, *55*(12), 10191–10217. https://doi.org/10.1029/2018WR024295

Chapman, T. G. (1999). A comparison of algorithms for stream flow recession and baseflow separation. *Hydrological Processes*, *13*(5), 701–714. https://doi.org/10.1002/(sici)1099-1085(19990415)13:5<701::aid-hyp774>3.0.co;2-2

Chapman, T. G. (1991). Comment on "Evaluation of automated techniques for base flow and recession analyses". In R. J. Nathan & T. A. McMahon (Eds.), *Water Resources Research* (Vol. 27(7), pp. 1783–1784). https://doi.org/10.1029/91WR01007

Cheng, L., Zhang, L., & Brutsaert, W. (2016). Automated selection of pure base flows from regular daily streamflow data: Objective algorithm. *Journal of Hydrologic Engineering*, *21*(11), 06016008. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001427

Clausnitzer, V., & Hopmans, J. (1995). General-purpose optimization code based on the Levenberg-Marquardt algorithm. Land, air. *Water Resources Paper*.

Cochand, M., Christe, P., Ornstein, P., & Hunkeler, D. (2019). Groundwater storage in high alpine catchments and its contribution to streamflow. *Water Resources Research*, *55*(4), 2613–2630. https://doi.org/10.1029/2018WR022989

Cook, P. G., Wood, C., White, T., Simmons, C. T., Fass, T., & Brunner, P. (2008). Groundwater inflow to a shallow, poorly-mixed wetland estimated from a mass balance of radon. *Journal of Hydrology*, *354*(1), 213–226. https://doi.org/10.1016/j.jhydrol.2008.03.016

Cui, T., Li, C., & Tian, F. (2021). Evaluation of temperature and precipitation simulations in CMIP6 models over the Tibetan Plateau. *Earth and Space Science*, *8*(7), e2020EA001620. https://doi.org/10.1029/2020EA001620

Cui, T., Sreekanth, J., Pickett, T., Rassam, D., Gilfedder, M., & Barrett, D. (2021). Impact of model parameterization on predictive uncertainty of regional groundwater models in the context of environmental impact assessment. *Environmental Impact Assessment Review*, *90*, 106620. https://doi.org/10.1016/j.eiar.2021.106620

Cuthbert, M. O., Gleeson, T., Moosdorf, N., Befus, K. M., Schneider, A., Hartmann, J., & Lehner, B. (2019). Global patterns and dynamics of climate–groundwater interactions. *Nature Climate Change*, *9*(2), 137–141. https://doi.org/10.1038/s41558-018-0386-4

Defourny, P., Schouten, L., Bartalev, S., Bontemps, S., Cacetta, P., De Wit, A. J. W., et al. (2009). Accuracy assessment of a 300 m global land cover map: The GlobCover experience [Dataset]. https://library.wur.nl/WebQuery/wurpubs/fulltext/14725

Dembélé, M., Hrachowitz, M., Savenije, H. H. G., Mariéthoz, G., & Schaefli, B. (2020). Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite data sets. *Water Resources Research*, *56*(1), e2019WR026085. https://doi.org/10.1029/2019WR026085

Diak, M., Böttcher, M. E., Ehlert von Ahn, C. M., Hong, W.-L., Kędra, M., Kotwicki, L., et al. (2023). Permafrost and groundwater interaction: Current state and future perspective. *Frontiers in Earth Science*, *11*, 1254309. https://doi.org/10.3389/feart.2023.1254309

Dietz, A. J., Kuenzer, C., Gessner, U., & Dech, S. (2012). Remote sensing of snow – A review of available methods. *International Journal of Remote Sensing*, *33*(13), 4094–4134. https://doi.org/10.1080/01431161.2011.640964

Doherty, J. (2015). *Calibration and uncertainty analysis for complex environmental models* (2nd ed.). Watermark Numerical Computing. Retrieved from http://docs.pesthomepage.org.s3.amazonaws.com/documents/pest_book_toc.pdf

Duncan, H. P. (2019). Baseflow separation – A practical approach. *Journal of Hydrology*, *575*, 308–313. https://doi.org/10.1016/j.jhydrol.2019.05.040

Eckhardt, K. (2005). How to construct recursive digital filters for baseflow separation. *Hydrological Processes*, *19*(2), 507–515. https://doi.org/10.1002/hyp.5675

Eckhardt, K. (2008). A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of Hydrology*, *352*(1), 168–173. https://doi.org/10.1016/j.jhydrol.2008.01.005

Elmi, O., Tourian, M. J., Saemian, P., & Sneeuw, N. (2024). Remote sensing-based extension of GRDC discharge time series - A monthly product with uncertainty estimates. *Scientific Data*, *11*(1), 240. https://doi.org/10.1038/s41597-024-03078-6

Farinotti, D., Immerzeel, W. W., de Kok, R. J., Quincey, D. J., & Dehecq, A. (2020). Manifestations and mechanisms of the Karakoram glacier Anomaly. *Nature Geoscience*, *13*(1), 8–16. https://doi.org/10.1038/s41561-019-0513-5

Fiddes, J., & Gruber, S. (2014). TopoSCALE v.1.0: Downscaling gridded climate data in complex terrain. *Geoscientific Model Development*, *7*(1), 387–405. https://doi.org/10.5194/gmd-7-387-2014

Floriancic, M. G., van Meerveld, I., Smoorenburg, M., Margreth, M., Naef, F., Kirchner, J. W., & Molnar, P. (2018). Spatio-temporal variability in contributions to low flows in the high Alpine Poschiavino catchment. *Hydrological Processes*, *32*(26), 3938–3953. https://doi.org/10.1002/hyp.13302

Frei, A., Tedesco, M., Lee, S., Foster, J., Hall, D. K., Kelly, R., & Robinson, D. A. (2012). A review of global satellite-derived snow products. *Advances in Space Research*, *50*(8), 1007–1029. https://doi.org/10.1016/j.asr.2011.12.021

Frey, H., Machguth, H., Huss, M., Huggel, C., Bajracharya, S., Bolch, T., et al. (2014). Estimating the volume of glaciers in the Himalaya-Karakoram region using different methods. *The Cryosphere*, *8*(6), 2313–2333. https://doi.org/10.5194/tc-8-2313-2014

Fuchs, M. C., Gloaguen, R., & Pohl, E. (2013). Tectonic and climatic forcing on the Panj river system during the Quaternary. *International Journal of Earth Sciences*, *102*(7), 1985–2003. https://doi.org/10.1007/s00531-013-0916-2

Fugger, S., Shaw, T. E., Jouberton, A., Miles, E. S., Buri, P., McCarthy, M., et al. (2024). Hydrological regimes and evaporative flux partitioning at the climatic ends of high mountain Asia. *Environmental Research Letters*, *19*(4), 44057. https://doi.org/10.1088/1748-9326/ad25a0

Gonzales, A. L., Nonner, J., Heijkers, J., & Uhlenbrook, S. (2009). Comparison of different base flow separation methods in a lowland catchment. *Hydrology and Earth System Sciences*, *13*(11), 2055–2068. https://doi.org/10.5194/hess-13-2055-2009

Gupta, A., & Govindaraju, R. S. (2019). Propagation of structural uncertainty in watershed hydrologic models. *Journal of Hydrology*, *575*, 66–81. https://doi.org/10.1016/j.jhydrol.2019.05.026

Haag, I., Jones, P. D., & Samimi, C. (2019). Central Asia's changing climate: How temperature and precipitation have changed across time, space, and altitude. *Climate*, *7*(10), 123. https://doi.org/10.3390/cli7100123

Hall, D. K., & Riggs, G. A. (2007). Accuracy assessment of the MODIS snow products. *Hydrological Processes*, *21*(12), 1534–1547. https://doi.org/10.1002/hyp.6715

Hall, F. R. (1968). Base-flow recessions—A review. *Water Resources Research*, *4*(5), 973–983. https://doi.org/10.1029/WR004i005p00973

Hargreaves, G. H., & Samani, Z. A. (1982). Estimating potential evapotranspiration. *Journal of the Irrigation and Drainage Division*, *108*(3), 225–230. https://doi.org/10.1061/JRCEA4.0001390

Hayashi, M. (2020). Alpine hydrogeology: The critical role of groundwater in sourcing the headwaters of the world. *Groundwater Series*, *58*(4), 498–510. https://doi.org/10.1111/gwat.12965

Helfer, F., Bernardi, F., de Barros, C. A. P., Allasia, D. G., Minella, J. P. G., Tassi, R., & Scariot, N. (2025). Enhanced baseflow separation in rural catchments: Event-specific calibration of recursive digital filters with tracer-derived data. *EGUsphere*, 1–36. https://doi.org/10.5194/egusphere-2025-244

Her, Y., & Chaubey, I. (2015). Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty. *Hydrological Processes*, *29*(19), 4220–4237. https://doi.org/10.1002/hyp.10487

Her, Y., & Seong, C. (2018). Responses of hydrological model equifinality, uncertainty, and performance to multi-objective parameter calibration. *Journal of Hydroinformatics*, *20*(4), 864–885. https://doi.org/10.2166/hydro.2018.108

Her, Y., Yoo, S.-H., Cho, J., Hwang, S., Jeong, J., & Seong, C. (2019). Uncertainty in hydrological analysis of climate change: Multi-parameter vs. multi-GCM ensemble predictions. *Scientific Reports*, *9*(1), 4974. https://doi.org/10.1038/s41598-019-41334-7

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis [Dataset]. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hewlett, J. D. (1961). Soil moisture as a source of base flow from steep mountain watersheds. *Southeastern Forest Experiment Station*, *132*. Retrieved from https://www.fs.usda.gov/research/treesearch/64674

Hewlett, J. D., & Hibbert, A. R. (1967). Factors affecting the response of small watersheds to precipitation in humid areas. *Forest Hydrology*, *1*, 275–290.

Hock, R. (2003). Temperature index melt modelling in mountain areas. *Journal of Hydrology*, *282*(1), 104–115. https://doi.org/10.1016/S0022-1694(03)00257-9

Hoelzle, M., Azisov, E., Barandun, M., Huss, M., Farinotti, D., Gafurov, A., et al. (2017). Re-establishing glacier monitoring in Kyrgyzstan and Uzbekistan, Central Asia. *Geoscientific Instrumentation, Methods and Data Systems*, *6*(2), 397–418. https://doi.org/10.5194/gi-6-397-2017

Hooghoudt, S. B. (1940). Bijdragen tot de kennis van eenige natuurkundige grootheden van den grond: Algemeene beschouwing van het probleem van de detailontwatering en de infiltratie door middel van parallel loopende drains, greppels, slooten en kanalen [in Dutch]. *Algemeene Landsdrukkerij*. Retrieved from https://library.wur.nl/WebQuery/wurpubs/fulltext/250838

Huang, J., Su, F., Yao, T., & Sun, H. (2022). Runoff regime, change, and attribution in the upper Syr Darya and Amu Darya, Central Asia. *Journal of Hydrometeorology*, *23*(10), 1563–1585. https://doi.org/10.1175/JHM-D-22-0036.1

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., et al. (2021). Accelerated global glacier mass loss in the early twenty-first century. *Nature*, *592*(7856), 726–731. https://doi.org/10.1038/s41586-021-03436-z

Immerzeel, W. W., Droogers, P., De Jong, S. M., & Bierkens, M. F. P. (2009). Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing. *Remote Sensing of Environment*, *113*(1), 40–49. https://doi.org/10.1016/j.rse.2008.08.010

Immerzeel, W. W., Lutz, A. F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., et al. (2020). Importance and vulnerability of the world's water towers. *Nature*, *577*(7790), 364–369. https://doi.org/10.1038/s41586-019-1822-y

Immerzeel, W. W., Wanders, N., Lutz, A. F., Shea, J. M., & Bierkens, M. F. P. (2015). Reconciling high-altitude precipitation in the upper Indus basin with glacier mass balances and runoff. *Hydrology and Earth System Sciences*, *19*(11), 4673–4687. https://doi.org/10.5194/hess-19-4673-2015

Jeelani, G., Lone, S. A., Nisa, A. U., Deshpande, R. D., & Padhya, V. (2021). Use of stable water isotopes to identify and estimate the sources of groundwater recharge in an alluvial aquifer of Upper Jhelum Basin (UJB), western Himalayas. *Hydrological Sciences Journal*, *66*(16), 2330–2339. https://doi.org/10.1080/02626667.2021.1985126

Kapil, J. C., Prasher, C., Datt, P., & Satyawali, P. K. (2010). Growth of melt–freeze clusters and formation of impeding layers to water flow in snow irradiated by a sun simulator under controlled laboratory conditions. *Annals of Glaciology*, *51*(54), 19–26. https://doi.org/10.3189/172756410791386599

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., et al. (2017). Climatologies at high resolution for the Earth's land surface areas [Dataset]. *Scientific Data*, *4*(1), 170122. https://doi.org/10.1038/sdata.2017.122

Karger, D. N., Wilson, A. M., Mahony, C., Zimmermann, N. E., & Jetz, W. (2021). Global daily 1 km land surface precipitation based on cloud cover-informed downscaling [Dataset]. *Scientific Data*, *8*(1), 307. https://doi.org/10.1038/s41597-021-01084-6

Khanal, S., Lutz, A. F., Kraaijenbrink, P. D. A., van den Hurk, B., Yao, T. D., & Immerzeel, W. W. (2021). Variable 21st century climate change response for Rivers in high Mountain Asia at seasonal to decadal time scales. *Water Resources Research*, *57*(5), e2020WR029266. https://doi.org/10.1029/2020WR029266

Klaus, J., & McDonnell, J. J. (2013). Hydrograph separation using stable isotopes: Review and evaluation. *Journal of Hydrology*, *505*, 47–64. https://doi.org/10.1016/j.jhydrol.2013.09.006

Liang, H., Huang, X., Sun, Y., Wang, Y., & Liang, T. (2017). Fractional snow-cover mapping based on MODIS and UAV data over the Tibetan Plateau. *Remote Sensing*, *9*(12), 1332. https://doi.org/10.3390/rs9121332

Linsbauer, A., Paul, F., & Haeberli, W. (2012). Modeling glacier thickness distribution and bed topography over entire mountain ranges with GlabTop: Application of a fast and robust approach. *Journal of Geophysical Research*, *117*(F3). https://doi.org/10.1029/2011JF002313

Liston, G. E., & Elder, K. (2006). A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). *Journal of Hydrometeorology*, *7*(2), 217–234. https://doi.org/10.1175/JHM486

Liston, G. E., & Winther, J.-G. (2005). Antarctic surface and subsurface snow and ice melt fluxes. *Journal of Climate*, *18*(10), 1469–1481. https://doi.org/10.1175/JCLI3344.1

Lo, M.-H., Famiglietti, J. S., Yeh, P. J.-F., & Syed, T. H. (2010). Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data. *Water Resources Research*, *46*(5). https://doi.org/10.1029/2009WR007855

Loibl, D., Richter, N., & Grünberg, I. (2025). Remote sensing-derived time series of transient glacier snowline altitudes for High Mountain Asia, 1985–2021. *Scientific Data*, *12*(1), 103. https://doi.org/10.1038/s41597-024-04309-6

Lone, S. A., Jeelani, G., Deshpande, R. D., Mukherjee, A., Jasechko, S., & Lone, A. (2021). Meltwaters dominate groundwater recharge in cold arid desert of Upper Indus River Basin (UIRB), western Himalayas. *Science of the Total Environment*, *786*, 147514. https://doi.org/10.1016/j.scitotenv.2021.147514

Luo, Y., Arnold, J., Allen, P., & Chen, X. (2012). Baseflow simulation using SWAT model in an inland river basin in Tianshan Mountains, Northwest China. *Hydrology and Earth System Sciences*, *16*(4), 1259–1267. https://doi.org/10.5194/hess-16-1259-2012

Lutz, A. F., Immerzeel, W. W., Kraaijenbrink, P. D. A., Shrestha, A. B., & Bierkens, M. F. P. (2016). Climate change impacts on the upper indus hydrology: Sources, shifts and extremes. *PLoS One*, *11*(11), e0165630. https://doi.org/10.1371/journal.pone.0165630

Lutz, A. F., Immerzeel, W. W., Shrestha, A. B., & Bierkens, M. F. P. (2014). Consistent increase in high Asia's runoff due to increasing glacier melt and precipitation. *Nature Climate Change*, *4*(7), 587–592. https://doi.org/10.1038/nclimate2237

Lyne, V., & Hollick, M. (1979). Stochastic time-variable rainfall-runoff modelling. In *Institute of engineers Australia national conference* (Vol. 79, pp. 89–93). Institute of Engineers Australia Barton. Retrieved from https://www.academia.edu/download/39814987/Stochastic_Time-Variable_Rainfall-Runoff20151108-28652-1m3nhta.pdf

Maussion, F., Scherer, D., Mölg, T., Collier, E., Curio, J., & Finkelnburg, R. (2014). Precipitation seasonality and variability over the Tibetan Plateau as resolved by the high Asia reanalysis [Dataset]. *Journal of Climate*, *27*(5), 1910–1927. https://doi.org/10.1175/JCLI-D-13-00282.1

Meeks, J., Moeck, C., Brunner, P., & Hunkeler, D. (2017). Infiltration under snow cover: Modeling approaches and predictive uncertainty. *Journal of Hydrology*, *546*, 16–27. https://doi.org/10.1016/j.jhydrol.2016.12.042

Meixner, T., Manning, A. H., Stonestrom, D. A., Allen, D. M., Ajami, H., Blasch, K. W., et al. (2016). Implications of projected climate change for groundwater recharge in the western United States. *Journal of Hydrology*, *534*, 124–138. https://doi.org/10.1016/j.jhydrol.2015.12.027

Mohammed, A. A., Pavlovskii, I., Cey, E. E., & Hayashi, M. (2019). Effects of preferential flow on snowmelt partitioning and groundwater recharge in frozen soils. *Hydrology and Earth System Sciences*, *23*(12), 5017–5031. https://doi.org/10.5194/hess-23-5017-2019

Montanari, A., & Di Baldassarre, G. (2013). Data errors and hydrological modelling: The role of model structure to propagate observation uncertainty. *Advances in Water Resources*, *51*, 498–504. https://doi.org/10.1016/j.advwatres.2012.09.007

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900. https://doi.org/10.13031/2013.23153

Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., & Aubry, D. (2011). A mass conservation approach for mapping glacier ice thickness. *Geophysical Research Letters*, *38*(19). https://doi.org/10.1029/2011GL048659

Morris, D. J., Speirs, D. C., Cameron, A. I., & Heath, M. R. (2014). Global sensitivity analysis of an end-to-end marine ecosystem model of the North Sea: Factors affecting the biomass of fish and benthos. *Ecological Modelling*, *273*, 251–263. https://doi.org/10.1016/j.ecolmodel.2013.11.019

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research*, *26*(7), 1465–1473. https://doi.org/10.1029/WR026i007p01465

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., & Williams, J. R. (2009). 1.1 overview of soil and water assessment tool (SWAT) model. *Tier B*, *8*, 3–23.

Palazzi, E., von Hardenberg, J., & Provenzale, A. (2013). Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. *Journal of Geophysical Research: Atmospheres*, *118*(1), 85–100. https://doi.org/10.1029/2012JD018697

Partington, D., Brunner, P., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R., & Dandy, G. C. (2012). Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model. *Journal of Hydrology*, *458–459*, 28–39. https://doi.org/10.1016/j.jhydrol.2012.06.029

Pelletier, A., & Andréassian, V. (2020). Hydrograph separation: An impartial parametrisation for an imperfect method. *Hydrology and Earth System Sciences*, *24*(3), 1171–1187. https://doi.org/10.5194/hess-24-1171-2020

Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., et al. (2014). The Randolph Glacier Inventory: A globally complete inventory of glaciers. *Journal of Glaciology*, *60*(221), 537–552. https://doi.org/10.3189/2014JoG13J176

Pohl, E., Gloaguen, R., Andermann, C., & Knoche, M. (2017). Glacier melt buffers river runoff in the Pamir Mountains. *Water Resources Research*, *53*(3), 2467–2489. https://doi.org/10.1002/2016wr019431

Pohl, E., Knoche, M., Gloaguen, R., Andermann, C., & Krause, P. (2015). Sensitivity analysis and implications for surface processes from a hydrological modelling approach in the Gunt catchment, high Pamir Mountains. *Earth Surface Dynamics*, *3*(3), 333–362. https://doi.org/10.5194/esurf-3-333-2015

Posavec, K., Bačani, A., & Nakić, Z. (2006). A visual basic spreadsheet macro for recession curve analysis. *Groundwater Series*, *44*(5), 764–767. https://doi.org/10.1111/j.1745-6584.2006.00226.x

Pritchard, H. D. (2019). Asia's shrinking glaciers protect large populations from drought stress. *Nature*, *569*(7758), 649–654. https://doi.org/10.1038/s41586-019-1240-1

Rittger, K., Painter, T. H., & Dozier, J. (2013). Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, *51*, 367–380. https://doi.org/10.1016/j.advwatres.2012.03.002

Rondeau-Genesse, G., Trudel, M., & Leconte, R. (2016). Monitoring snow wetness in an Alpine Basin using combined C-band SAR and MODIS data. *Remote Sensing of Environment*, *183*, 304–317. https://doi.org/10.1016/j.rse.2016.06.003

Ruelland, D. (2020). Should altitudinal gradients of temperature and precipitation inputs be inferred from key parameters in snow-hydrological models? *Hydrology and Earth System Sciences*, *24*(5), 2609–2632. https://doi.org/10.5194/hess-24-2609-2020

Rycroft, D. W., & Smedema, L. K. (1983). *Land drainage: Planning and design of agricultural drainage systems*. Batsford Academic & Educational.

Safeeq, M., Grant, G. E., Lewis, S. L., & Tague, C. L. (2013). Coupling snowpack and groundwater dynamics to interpret historical streamflow trends in the western United States. *Hydrological Processes*, *27*(5), 655–668. https://doi.org/10.1002/hyp.9628

Salerno, F., Guyennon, N., Yang, K., Shaw, T. E., Lin, C., Colombo, N., et al. (2023). Local cooling and drying induced by Himalayan glaciers under global warming. *Nature Geoscience*, *16*(12), 1120–1127. https://doi.org/10.1038/s41561-023-01331-y

Schaefli, B., & Huss, M. (2011). Integrating point glacier mass balance observations into hydrologic model identification. *Hydrology and Earth System Sciences*, *15*(4), 1227–1241. https://doi.org/10.5194/hess-15-1227-2011

Scherler, D., Wulf, H., & Gorelick, N. (2018). Global assessment of supraglacial debris-cover extents [Dataset]. *Geophysical Research Letters*, *45*(21), 11798–11805. https://doi.org/10.1029/2018GL080158

Schiemann, R., Lüthi, D., Vidale, P. L., & Schär, C. (2008). The precipitation climate of Central Asia—Intercomparison of observational and numerical data sources in a remote semiarid region. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *28*(3), 295–314.

Shin, S., Her, Y., Muñoz-Carpena, R., & Khare, Y. P. (2023). Multi-parameter approaches for improved ensemble prediction accuracy in hydrology and water quality modeling. *Journal of Hydrology*, *622*, 129458. https://doi.org/10.1016/j.jhydrol.2023.129458

Siegfried, T., Bernauer, T., Guiennet, R., Sellars, S., Robertson, A. W., Mankin, J., & Bauer-Gottwein, P. (2010). Coping with international water conflict in central Asia: Implications of climate change and melting ice in the Syr Darya catchment. *International Studies*, 1–36.

Simons, G., Koster, R., & Droogers, P. (2020). Hihydrosoil v2. 0-high resolution soil maps of global hydraulic properties [Dataset]. *Future Works*. Retrieved from https://www.futurewater.nl/wp-content/uploads/2020/10/HiHydroSoil-v2.0-High-Resolution-Soil-Maps-of-Global-Hydraulic-Properties.pdf

Sloto, R. A., & Crouse, M. Y. (1996). *HYSEP: A computer program for streamflow hydrograph separation and analysis (No. 96–4040)*. Water-Resources Investigations Report. U.S. Geological Survey. https://doi.org/10.3133/wri964040

Smakhtin, V. U. (2001). Low flow hydrology: A review. *Journal of Hydrology*, *240*(3), 147–186. https://doi.org/10.1016/S0022-1694(00)00340-1

Smith, J. D., Lin, L., Quinn, J. D., & Band, L. E. (2022). Guidance on evaluating parametric model uncertainty at decision-relevant scales. *Hydrology and Earth System Sciences*, *26*(9), 2519–2539. https://doi.org/10.5194/hess-26-2519-2022

Somers, L. D., & McKenzie, J. M. (2020). A review of groundwater in high mountain environments. *WIREs Water*, *7*(6), e1475. https://doi.org/10.1002/wat2.1475

Şorman, A. Ü., Akyürek, Z., Şensoy, A., Şorman, A. A., & Tekeli, A. E. (2007). Commentary on comparison of MODIS snow cover and albedo products with ground observations over the mountainous terrain of Turkey. *Hydrology and Earth System Sciences*, *11*(4), 1353–1360. https://doi.org/10.5194/hess-11-1353-2007

Su, F., Pritchard, H. D., Yao, T., Huang, J., Ou, T., Meng, F., et al. (2022). Contrasting fate of Western third pole's water resources under 21st century climate change. *Earth's Future*, *10*(9), e2022EF002776. https://doi.org/10.1029/2022EF002776

Sun, A. Y., Green, R., Rodell, M., & Swenson, S. (2010). Inferring aquifer storage parameters using satellite and in situ measurements: Estimation under uncertainty. *Geophysical Research Letters*, *37*(10). https://doi.org/10.1029/2010GL043231

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107. https://doi.org/10.1002/2017RG000574

Tague, C., McMichael, C., Hope, A., Choate, J., & Clark, R. (2004). Application of the RHESSys model to a California semiarid Shrubland Watershed. *JAWRA Journal of the American Water Resources Association*, *40*(3), 575–589. https://doi.org/10.1111/j.1752-1688.2004.tb04444.x

Tallaksen, L. M. (1995). A review of baseflow recession analysis. *Journal of Hydrology*, *165*(1), 349–370. https://doi.org/10.1016/0022-1694(94)02540-R

Tang, W., & Carey, S. K. (2017). HydRun: A MATLAB toolbox for rainfall–runoff analysis. *Hydrological Processes*, *31*(15), 2670–2682. https://doi.org/10.1002/hyp.11185

Tarasova, L., Knoche, M., Dietrich, J., & Merz, R. (2016). Effects of input discretization, model complexity, and calibration strategy on model performance in a data-scarce glacierized catchment in Central Asia. *Water Resources Research*, *52*(6), 4674–4699. https://doi.org/10.1002/2015WR018551

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., et al. (2013). Ground water and climate change. *Nature Climate Change*, *3*(4), 322–329. https://doi.org/10.1038/nclimate1744

Teegavarapu, R. S. V. (2014). Statistical corrections of spatially interpolated missing precipitation data estimates. *Hydrological Processes*, *28*(11), 3789–3808. https://doi.org/10.1002/hyp.9906

Terink, W., Lutz, A. F., Simons, G. W. H., Immerzeel, W. W., & Droogers, P. (2015). SPHY v2.0: Spatial processes in HYdrology. *Geoscientific Model Development*, *8*(7), 2009–2034. https://doi.org/10.5194/gmd-8-2009-2015

van Tiel, M., Aubry-Wake, C., Somers, L., Andermann, C., Avanzi, F., Baraer, M., et al. (2024). Cryosphere-groundwater connectivity is a missing link in the mountain water cycle. *Nature Water*, *2*(7), 624–637. https://doi.org/10.1038/s44221-024-00277-8

Tularam, G. A., & Ilahee, M. (2008). Exponential smoothing method of base flow separation and its impact on continuous loss estimates. *American Journal of Environmental Sciences*, *4*(2), 136–144. https://doi.org/10.3844/ajessp.2008.136.144

Tustison, B., Harris, D., & Foufoula-Georgiou, E. (2001). Scale issues in verification of precipitation forecasts. *Journal of Geophysical Research*, *106*(D11), 11775–11784. https://doi.org/10.1029/2001JD900066

Unger-Shayesteh, K., Vorogushyn, S., Farinotti, D., Gafurov, A., Duethmann, D., Mandychev, A., & Merz, B. (2013). What do we know about past changes in the water cycle of Central Asian headwaters? A review. *Global and Planetary Change*, *110*, 4–25. https://doi.org/10.1016/j.gloplacha.2013.02.004

Urrutia, J., Herrera, C., Custodio, E., Jódar, J., & Medina, A. (2019). Groundwater recharge and hydrodynamics of complex volcanic aquifers with a shallow saline lake: Laguna Tuyajto, Andean Cordillera of northern Chile. *Science of the Total Environment*, *697*, 134116. https://doi.org/10.1016/j.scitotenv.2019.134116

van Tiel, M., Stahl, K., Freudiger, D., & Seibert, J. (2020). Glacio-hydrological model calibration and evaluation. *WIREs Water*, *7*(6), e1483. https://doi.org/10.1002/wat2.1483

Varis, O., & Kummu, M. (2012). The Major Central Asian River basins: An assessment of vulnerability. *International Journal of Water Resources Development*, *28*(3), 433–452. https://doi.org/10.1080/07900627.2012.684309

Venetis, C. (1969). A study on the recession of unconfined acquifers. *International Association of Scientific Hydrology. Bulletin*, *14*(4), 119–125. https://doi.org/10.1080/02626666909493759

Wagener, T., & Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*, *19*(6), 378–387. https://doi.org/10.1007/s00477-005-0006-5

Walvoord, M. A., & Kurylyk, B. L. (2016). Hydrologic impacts of thawing permafrost—A review. *Vadose Zone Journal*, *15*(6), vzj2016.01.0010. https://doi.org/10.2136/vzj2016.01.0010

Wang, Q., Yi, S., & Sun, W. (2017). Precipitation-driven glacier changes in the Pamir and Hindu Kush mountains. *Geophysical Research Letters*, *44*(6), 2817–2824. https://doi.org/10.1002/2017GL072646

Wang, X., Xie, H., & Liang, T. (2008). Evaluation of MODIS snow cover and cloud mask and its application in Northern Xinjiang, China. *Remote Sensing of Environment*, *112*(4), 1497–1513. https://doi.org/10.1016/j.rse.2007.05.016

World Bank. (2008). Environmental and Social Assessment for the Pamir I Hydropower Project: Regulation of Lake Yashilkul. World Bank.

Xie, J., Liu, X., Wang, K., Yang, T., Liang, K., & Liu, C. (2020). Evaluation of typical methods for baseflow separation in the contiguous United States. *Journal of Hydrology*, *583*, 124628. https://doi.org/10.1016/j.jhydrol.2020.124628

Yao, T. D., Thompson, L., Yang, W., Yu, W., Gao, Y., Guo, X. J., et al. (2012). Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nature Climate Change*, *2*(9), 663–667. https://doi.org/10.1038/nclimate1580

Zandler, H., Haag, I., & Samimi, C. (2019). Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions. *Scientific Reports*, *9*(1), 15118. https://doi.org/10.1038/s41598-019-51666-z

Zecharias, Y. B., & Brutsaert, W. (1988). Recession characteristics of groundwater outflow and base flow from mountainous watersheds. *Water Resources Research*, *24*(10), 1651–1658. https://doi.org/10.1029/WR024i010p01651

Zhang, Y.-K., & Schilling, K. E. (2006). Increasing streamflow and baseflow in Mississippi River since the 1940 s: Effect of land use change. *Journal of Hydrology*, *324*(1), 412–422. https://doi.org/10.1016/j.jhydrol.2005.09.033

Zhao, B., Mao, J., Dai, Q., Han, D., Dai, H., & Rong, G. (2019). Exploration on hydrological model calibration by considering the hydro-meteorological variability. *Hydrology Research*, *51*(1), 30–46. https://doi.org/10.2166/nh.2019.047

## References From the Supporting Information

Arenson, L. U., Harrington, J. S., Koenig, C. E. M., & Wainstein, P. A. (2022). Mountain permafrost hydrology—A practical review following studies from the Andes. *Geosciences*, *12*(2), 48. https://doi.org/10.3390/geosciences12020048

Bastola, S., Murphy, C., & Sweeney, J. (2011). Evaluation of the transferability of hydrological model parameters for simulations under changed climatic conditions. *Hydrology and Earth System Sciences Discussions*, *8*, 5891–5915.

Efstratiadis, A., & Koutsoyiannis, D. (2010). One decade of multi-objective calibration approaches in hydrological modelling: A review. *Hydrological Sciences Journal*, *55*(1), 58–78. https://doi.org/10.1080/02626660903526292

Evans, S. G., Ge, S., & Liang, S. (2015). Analysis of groundwater flow in mountainous, headwater catchments with permafrost. *Water Resources Research*, *51*(12), 9564–9576. https://doi.org/10.1002/2015WR017732

de Ferreira, P. M. L., da Paz, A. R., & Bravo, J. M. (2020). Objective functions used as performance metrics for hydrological models: State-of-the-art and critical analysis. *RBRH*, *25*, e42. https://doi.org/10.1590/2318-0331.252020190155

Gharari, S., Hrachowitz, M., Fenicia, F., & Savenije, H. H. G. (2013). An approach to identify time consistent model parameters: Sub-period calibration. *Hydrology and Earth System Sciences*, *17*(1), 149–161. https://doi.org/10.5194/hess-17-149-2013

Guse, B., Pfannerstill, M., Gafurov, A., Fohrer, N., & Gupta, H. (2016). Demasking the integrated information of discharge: Advancing sensitivity analysis to consider different hydrological components and their rates of change. *Water Resources Research*, *52*(11), 8724–8743. https://doi.org/10.1002/2016WR018894

Herrera, P. A., Marazuela, M. A., & Hofmann, T. (2022). Parameter estimation and uncertainty analysis in hydrological modeling. *WIREs Water*, *9*(1), e1569. https://doi.org/10.1002/wat2.1569

Horton, P., Schaefli, B., & Kauzlaric, M. C. (2022). Why do we have so many different hydrological models? A review based on the case of Switzerland. *WIREs Water*, *9*(1), e1574. https://doi.org/10.1002/wat2.1574

Merz, R., Parajka, J., & Blöschl, G. (2011). Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resources Research*, *47*(2). https://doi.org/10.1029/2010WR009505

Myers, D. T., Ficklin, D. L., Robeson, S. M., Neupane, R. P., Botero-Acosta, A., & Avellaneda, P. M. (2021). Choosing an arbitrary calibration period for hydrologic models: How much does it influence water balance simulations? *Hydrological Processes*, *35*(2), e14045. https://doi.org/10.1002/hyp.14045

Pohl, E., uulu Esenaman, M., Halimov, A., Amschwand, D., Saks, T., & Huang, J. (2024). A Central Asian pro-glacial discharge database for improving hydrological models, 16055. Presented at the EGU General Assembly Conference Abstracts. https://doi.org/10.5194/egusphere-egu24-16055

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, *46*(5). https://doi.org/10.1029/2009WR008328

Santos, L., Andréassian, V., Sonnenborg, T. O., Lindström, G., de Lavenne, A., Perrin, C., et al. (2025). Lack of robustness of hydrological models: A large-sample diagnosis and an attempt to identify hydrological and climatic drivers. *Hydrology and Earth System Sciences*, *29*(3), 683–700. https://doi.org/10.5194/hess-29-683-2025

Thirel, G., Santos, L., Delaigue, O., & Perrin, C. (2024). On the use of streamflow transformations for hydrological model calibration. *Hydrology and Earth System Sciences*, *28*(21), 4837–4860. https://doi.org/10.5194/hess-28-4837-2024

Westerberg, I. K., Sikorska-Senoner, A. E., Viviroli, D., Vis, M., & Seibert, J. (2022). Hydrological model calibration with uncertain discharge data. *Hydrological Sciences Journal*, *67*(16), 2441–2456. https://doi.org/10.1080/02626667.2020.1735638

Zhang, X., & Liu, P. (2021). A time-varying parameter estimation approach using split-sample calibration based on dynamic programming. *Hydrology and Earth System Sciences*, *25*(2), 711–733. https://doi.org/10.5194/hess-25-711-2021