



Comparison of machine learning and process-based SWAT model in simulating streamflow in the Upper Indus Basin

Khalil Ur Rahman¹ · Quoc Bao Pham² · Khan Zaib Jadoon³ · Muhammad Shahid⁴ · Daniel Prakash Kushwaha⁵ · Zheng Duan⁶ · Babak Mohammadi⁶ · Khaled Mohamed Khedher^{7,8} · Duong Tran Anh⁹

Received: 21 August 2021 / Accepted: 9 May 2022 / Published online: 14 June 2022
© The Author(s) 2022

Abstract

This study appraised and compared the performance of process-based hydrological SWAT (soil and water assessment tool) with a machine learning-based multi-layer perceptron (MLP) models for simulating streamflow in the Upper Indus Basin. The study period ranges from 1998 to 2013, where SWAT and MLP models were calibrated/trained and validated/tested for multiple sites during 1998–2005 and 2006–2013, respectively. The performance of both models was evaluated using nash–sutcliffe efficiency (NSE), coefficient of determination (R^2), Percent BIAS (PBIAS), and mean absolute percentage error (MAPE). Results illustrated the relatively poor performance of the SWAT model as compared with the MLP model. NSE, PBIAS, R^2 , and MAPE for SWAT (MLP) models during calibration ranged from the minimum of 0.81 (0.90), 3.49 (0.02), 0.80 (0.25) and 7.61 (0.01) to the maximum of 0.86 (0.99), 9.84 (0.12), 0.87 (0.99), and 15.71 (0.267), respectively. The poor performance of SWAT compared with MLP might be influenced by several factors, including the selection of sensitive parameters, selection of snow specific sensitive parameters that might not represent actual snow conditions, potential limitations of the SCS-CN method used to simulate streamflow, and lack of SWAT ability to capture the hydropeaking in Indus River sub-basins (at Shatial bridge and Bisham Qila). Based on the robust performance of the MLP model, the current study recommends to develop and assess machine learning models and merging the SWAT model with machine learning models.

Keywords Hydrological modeling · Glacier · SWAT · MLP · Upper Indus Basin

Introduction

Glaciers are considered an icon of climate change, and they can clearly represent the emergence of climate globally (IPCC, 2018). Several studies have reported that

mountain glaciers will significantly contribute to sea-level rise in the coming years, and it may change the hydrology of basins covered by permanent snow and glacier (Beniston et al. 2018; Hock et al. 2019). Snowmelt is the primary source of fresh water in many regions worldwide,

✉ Duong Tran Anh
ta.duong@hutech.edu.vn

¹ State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

² Institute of Applied Technology, Thu Dau Mot University, Binh Duong Province, Vietnam

³ Department of Civil Engineering, Islamic International University, Islamabad 44000, Pakistan

⁴ NICE, SCCE, National University of Sciences & Technology (NUST), Islamabad 44000, Pakistan

⁵ Department of Soil and Water Conservation Engineering, College of Technology, G.B. Pant University of Agriculture & Technology, Pantnagar 263145, India

⁶ Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, SE-223 62 Lund, Sweden

⁷ Department of Civil Engineering, College of Engineering, King Khalid University, Abha 61421, Saudi Arabia

⁸ Department of Civil Engineering, High Institute of Technological Studies, Mrezgua University Campus, 8000 Nabeul, Tunisia

⁹ HUTECH University, 475A, Dien Bien Phu, Ward 25, Binh Thanh District, Ho Chi Minh City, Vietnam

and it is extremely important for the community living in Hindukush-Karakorum-Himalayas (HKH). HKH is also named as the “third pole” and “roof of the world” due to substantial glacial coverage in the high elevated basins (Yao et al. 2012; You et al. 2016). Glacial and snow cover in the HKH region constitute from 70 to 80% of the mean annual available freshwater from the Upper Indus Basin (UIB) (Immerzeel et al. 2009).

UIB is the main source of freshwater resources and plays a pivotal role in the sustainable development of Pakistan (Yaseen et al. 2020). The Upper Indus River system supplies sustainable water to the large population downstream of UIB for agriculture, industrial, and domestic purposes (Immerzeel et al. 2010). The seasonal water from UIB accounts for approximately half of the mean annual surface water available in Pakistan, which is essential for producing 3500 MW hydropower potential at Tarbela Dam (Hasson et al. 2017). Further, UIB also contributes to Pakistan's agrarian economy by satisfying the extensive irrigation requirements to meet rising food demand. Most of the south and southeast Asian basins are dependent on the summer monsoon; however, UIB is dependent on melted water from its ample glacial and snow coverage (Hasson et al. 2014).

Forecasting the streamflow and hazard management in such glacial basins plays a critical role in the region's sustainable development. Population in the Indus Basin depend on the river flows from UIB; therefore, streamflow forecasting is crucial for people living downstream of the HKH. Water resources of the Indus River basin should be managed by using real-time early warning systems (Krajewski et al. 2017). However, these systems usually require huge investments, which is difficult for communities living in UIB and the Indus River basin. Therefore, inexpensive, accurate, and innovative forecasting and simulation techniques are strongly recommended across the entire Indus Basin.

Estimating snowmelt and glacier-melt streamflow is vital for effective planning and management of surface water in the UIB. The changes in glaciers under the influence of climate change will strongly impact river flow and hydrological regimes in the UIB (Huss and Hock, 2018). However, the precise estimation of streamflow in a basin characterized by mountains covered with permanent snow and glacier is considered an unsolved problem, which deserves the attention of hydrology community (Lettenmaier et al. 2015).

Hydrological models neglect the trivial information related to the structure of the model, comprehend hydrological processes and powerful tools for effective decision-making related to the sustainable management of water resources (Nguyen et al. 2019; Rahman et al. 2020a; Tuo et al. 2016). Hydrological models play a vital role in allowing users to

explain, estimate, and predict hydrological processes in basins characterized with limited, non-accessible, cost-efficient, and time-consuming *in-situ* observations (Baffaut et al. 2015). The structure of these hydrological models varies from a complex process-based distributed model to simple lumped models. Soil and Water Assessment Tool (SWAT) model belongs to the complex process-based distributed model (Arnold et al. 1998; Gassman et al. 2007; Nguyen et al. 2019).

SWAT model has been extensively used in modeling several types of hydrological processes in river basins across the globe (Abbaspour et al. 2015; Duan et al. 2019; Francesconi et al. 2016; Golmohammadi et al. 2017; Liu et al. 2016; Malagò et al. 2016; Nguyen et al. 2019; Rahman et al. 2020a; Tuo et al. 2016). SWAT is also extensively employed to assess the impact of snow on the water cycle across the mountainous basins (Debele et al. 2010; Rostamian et al. 2008; Grusson et al. 2015; Troin and Caya, 2014; Shahid et al. 2021). The temperature-index method has been widely used to model the snow processes in different basins using SWAT (Hock, 2003; Walter et al. 2005; Zhang et al. 2008), which is proved remarkably accurate in several studies (Debele et al. 2010; Luo et al. 2013).

The data-driven models, on the other hand, have been effectively used in several hydrological applications and such models provided high accuracy even without prior knowledge of underlying processes (LV et al. 2020; Senent-Aparicio et al. 2019; Yang et al. 2020). The approaches such as artificial intelligence (AI), soft computing (SC), data mining (DM), computational intelligence (CI), and machine learning analyze the system-related data and provide linkage between input and output variables without considering the explicit physical behavior of the objective system (Solomatine et al. 2009).

Recent studies have used several machine learning models to address different aspects of water resources management and hydrological modeling. For example, artificial neural network (ANN) is utilized to simulate rainfall-runoff, predict runoff, model river sediment process, predict storage inflow, and evaluate the water-powered energy (Choong et al. 2020; Pradhan et al. 2020). Fan et al. (2020) compared the short-term long memory (LSTM) model with SWAT and ANN models to simulate streamflow across Poyang Lake Basin. Results demonstrated the superior performance of LSTM model compared with SWAT and ANN models in simulating streamflow at a daily scale. Pradhan et al. (2020) investigated the performance of three ANN models and the SWAT model in predicting the streamflow and illustrated that ANN models have more accurate estimates as compared with SWAT. Similarly, Kumar et al. (2019) compared Emotional Neural Network (ENN) and ANN to simulate streamflow, where ENN was reported to have better performance. Koycegiz and Buyukyildiz (2019) compared SWAT with

support vector machine (SVM) and ANN in the headwater of Carsamba River, situated in Konya Closed Basin, Turkey. Results demonstrated that data-driven models (ANN and SVM) have better performance in streamflow simulation as compared with SWAT. Several other models, including Adaptive Neuro-Fuzzy Inference System (ANFIS) and SVM, are also used for hydrological modeling (He et al. 2014; Moradkhani et al. 2004).

The comparative studies among data-driven and physically-based models ensured the successful application, selection of robust models, interpretation of the model outputs and reliable results. The evaluation and comparison of data driven model across catchments like UIB is extremely critical. UIB, being the source of freshwater for the entire Indus Basin, has several problems such as extreme climate variability owing to the complex topography, streamflow is extremely seasonal, subject to severe climate and land use changes, and most importantly a data scarce region. Minimal studies are available that developed and evaluated data-driven models and performed a comparison with physical-based models in glacial regions like UIB. Therefore, the current study adds to the available literature; (i) compare the performance of SWAT and MLP models in streamflow simulation for the first time across UIB, (ii) simulate streamflow that is less influenced by precipitation and more dependent on topography, air temperature, relative humidity, and solar radiation, and (iii) propose model that is capable to capture the seasonality in streamflow. The findings from current study will educate researchers and policy makers about robust alternatives to data-intensive hydrological models for streamflow simulation in data scarce regions characterized with complex topography and diverse climate.

Study area

UIB is situated in the extreme north of Pakistan, mostly covered by permanent snow and glacier cover, and located between 33.67° – 37.20° N and 70.50° – 77.50° E. The elevation of UIB ranges from 8500 above mean sea level (a.m.s.l) to 200 m a.m.s.l with the mean elevation of 3750 m a.m.s.l. UIB shares its boundary with China (north), Afghanistan (west), and India (east), as shown in Fig. 1. UIB depicts significant topographic and climatic variations, which include a complex terrain of HKH mountain ranges. HKH has discrete topographical landscapes with conflicting climate change signals (Archer, 2003; Cheema and Bastiaanssen, 2012). HKH and Tibetan Plateau are the greatest glacial regions of the world with approximately 22,000 km² of glacial surface area and jointly host 11,000 glaciers (Ul Hasson et al. 2016). UIB is the originating source of freshwater for the Indus River, contributes to approximately half of the available surface water of

Pakistan (Yaseen et al. 2020), and plays an important role in Pakistan's sustainable economic development.

Figure 1 represents detailed information about the study area, including elevation, selected basins for hydrological modeling, the spatial distribution of rain gauges (RGs), climate and streamflow stations, and land use map. The study region covers important basins of the UIB, including Shyok, Shigar, Hunza, Astore, Gilgit, Indus River basin at Bisham Qila, and Indus River basin at Shatial bridge. Shyok and Shigar basins are situated in the eastern and central parts of Karakoram. Approximately 24% of Shyok and 33.33% of Shigar basins are covered by snow (Bhambri et al. 2013; Yaseen et al. 2020). Westerly disturbances and monsoon are the main seasons/sources for precipitation in Shyok and Shigar basins (Hasson et al. 2017; Hasson et al. 2014; Latif et al. 2020). Hunza basin is situated in the western Karakoram ranges, with 28% of the area covered with glaciers, which is 21% of the total UIB glacial coverage (Bhambri et al. 2013; Hasson et al. 2014). Three high-altitude climate stations, i.e., Khunjrab, Naltar, and Ziarat, are situated in the Hunza basin. Discharge of Hunza basin is measured at Danyior bridge of Hunza River. Astore basin is located in the western Himalayan ranges with 14% of glacial and snow coverage, which is 3% of the total UIB glacial coverage (Hasson et al. 2014). There is only one climate station (Astore) measuring climate and precipitation in the Astore basin, where the discharge of the basin is measured at Doyian. Gilgit basin is situated in the eastern part of Hindukush ranges, which drains south-east and joins the Indus River. The climate and precipitation data in the Gilgit basin is recorded at Gilgit, Yasin, Gupis, and Ushkore stations, where the basin's discharge is measured at the Gilgit station and Alam bridge (the confluence of Gilgit and Hunza rivers). Bisham Qila station is the final station used in the current study, located at the Indus River, and is considered an exit point.

Datasets and methodology

In-situ data from rain gauges, climate stations, and streamflow gauges

The daily *in-situ* precipitation and climate (both maximum and minimum temperature, minimum and maximum relative humidity, wind speed, and solar radiation) data are collected from Pakistan Meteorology Department (PMD), and Water and Power Development Authority (WAPDA). It is worthy of mentioning that the streamflow gauges' data is collected solely from WAPDA. Table 1 represents the names of RGs, climate stations, and streamflow gauges used in the current research. After the rigorous screening of collected data, a temporal span of 1995–2013 was selected to warm-up, calibrate, and validate the SWAT model across selected basins of UIB. It is ensured that all the gauges/stations have daily

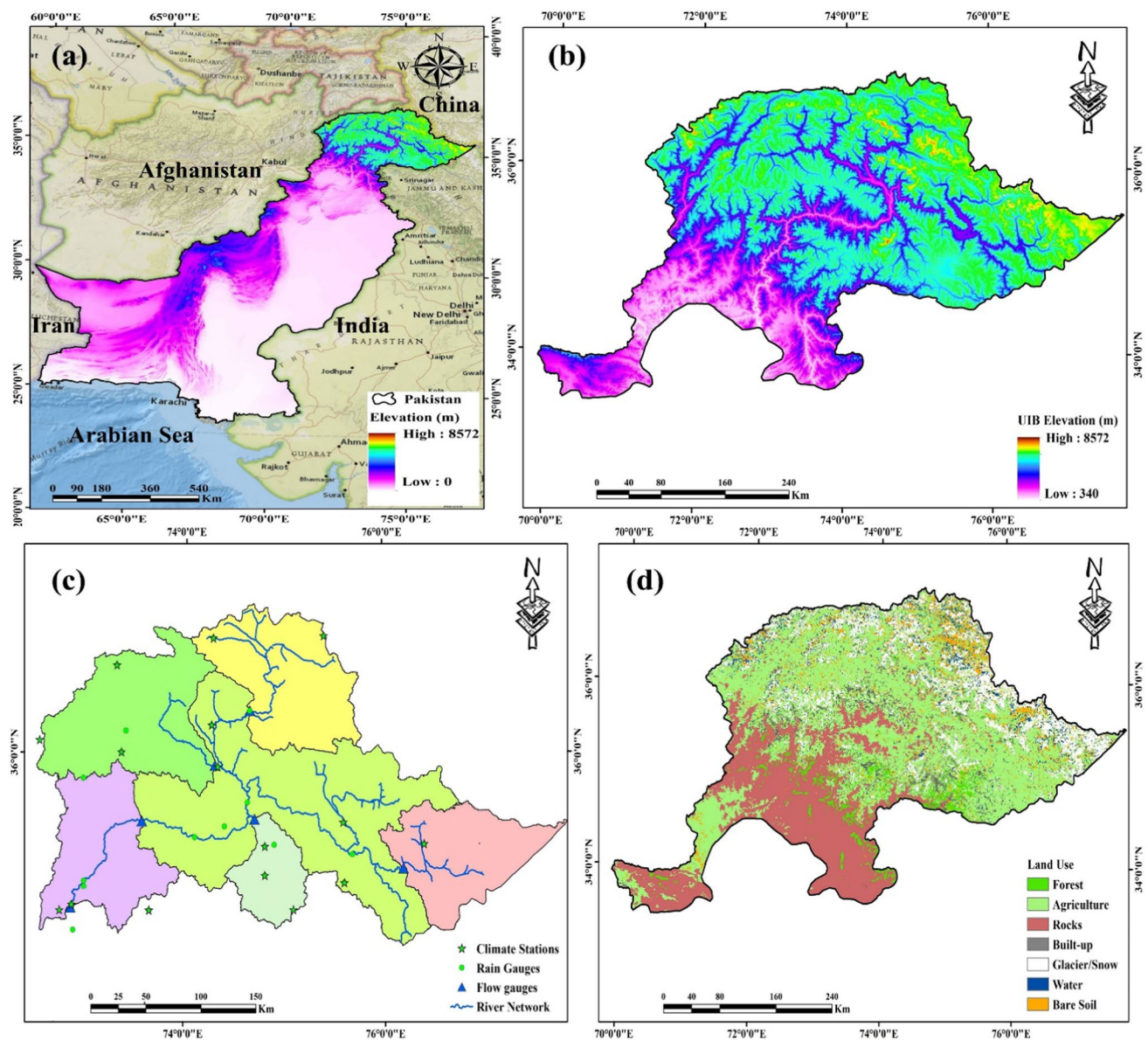


Fig. 1 Detailed information about the study area, **a** geographical location and elevation of Pakistan, **b** elevation of UIB, **c** selected basins and the distribution of rain gauges, climate and streamflow stations, **d** land use of UIB

data without any significant missing information. PMD and WAPDA perform the manual collection of obtained data, which might have several types of inevitable errors, including instrumental and human errors, splashing effects, snow impact, and wind errors. These factors might deteriorate the quality of *in-situ* data (Rahman et al. 2020b). Therefore, PMD and WAPDA monitor and improve the data quality using the Guide to Hydrological Practices suggested by the world meteorology organization (WMO, 1994).

SWAT model can automatically fill the missing meteorological input data by employing the weather generator. In order to fill out the data, more input observations and further efforts are needed (Rahman et al. 2020a). Furthermore,

the accuracy of SWAT model (output) is dependent on the accuracy of input data. Therefore, the zero-order method was employed to fill the missing data (if any) in precipitation and climate data before its integration into SWAT and MLP models. Moreover, Kurtosis and Skewness methods are employed to check the quality of input data (Rahman et al. 2018).

UIB has only 2% of the cultivable land and the irrigation system is a traditional one (ICIMOD, 2017). Around 15% of the cultivable land is not cultivated because of unavailability or limited access to irrigation water and irrigation system complexity (ICIMOD, 2017). Further, the agricultural farmlands are sparsely distributed and most of them are

Table 1 Input data (daily time scale) from RGs, climate stations, and streamflow gauges during 1995–2013

Precipitation	Temperature	Relative humidity	Wind speed	Solar radiation	Streamflow
Khunjrab	Khunjrab	Khunjrab	Khunjrab	Khunjrab	Gilgit
Ziarat	Ziarat	Ziarat	Ziarat	Ziarat	Hunza
Yasin	Yasin	Yasin	Yasin	Yasin	Astore river at Doyian
Gupis	Shendure	Shendure	Shendure	Shendure	Alam bridge
Ushkore	Ushkore	Ushkore	Ushkore	Ushkore	Shyok river at Yogo
Hunza	Naltar	Naltar	Naltar	Naltar	Indus River at Shatial bridge
Naltar	Gilgit	Gilgit	Gilgit	Gilgit	Indus River at Bisham Qila
Gilgit	Bisham Qila	Bisham Qila	Bisham Qila	Bisham Qila	
Kalam	Shangla	Shangla	Shangla	Shangla	
Bunji	Rama	Rama	Rama	Rama	
Chillas	Ratu	Ratu	Ratu	Ratu	
Pattan	Burzil	Burzil	Burzil	Burzil	
Bisham Qila	Shigar	Shigar	Shigar	Shigar	
Rama	Deosai	Deosai	Deosai	Deosai	
Astore	Hushy	Hushy	Hushy	Hushy	
Ratu	Naran	Naran	Naran	Naran	
Skardu					
Deosai					
Hushy					

above the level of flowing water in rivers. Irrigation in UIB is mainly rainfed irrigation due to its poor irrigation system and complex terrain (Malik and Azam, 2009; Parveen et al. 2015). Agriculture lands are irrigated from snow and glacier melt water through dug earthen irrigation channels, which are lengthy, rough, and crude, resulting in inadequate water supply for irrigation (Parveen et al. 2015; Khan et al. 2021). Therefore, we did not consider the irrigation and agricultural water use data, and cropping pattern in the current study due to its non-availability.

SWAT model

The SWAT model is a process-based semi-distributed, Hydrological Response Unit (HRU)-based, spatially explicit, and time-continuous hydrological model developed by the agricultural research service of the United States Department of Agriculture (Arnold et al. 1998). SWAT model divide the large basins into smaller sub-basins to provide more accurate spatial details, which make the model more reliable and accurate (Jha, 2004). SWAT model was designed to simulate as well as forecast the impacts of agricultural and land management decision/practices on water resources in terms of quantity and water quality across a range of basins sizes (Gassman et al. 2007). Further, the hydrological responses to land use and climate changes are mostly investigated using the scenario-based simulations through SWAT model (Yin et al. 2017). The computational efficiency of SWAT model makes the simulation across large basins or different types

of management strategies easy (Coutu and Vega, 2007). In this research, ArcSWAT version 2012, revision 664 was used to simulate streamflow in UIB. Generally, the SWAT model is extensively used in water quality assessment, simulation of rainfall-runoff processes, evapotranspiration, and soil erosion. SWAT has the potential to assess climate change impact on water resources, transport of nutrients and sediments under the various circumstances of land use land cover (LULC), meteorological, and soil data (Ali et al. 2020; Khan et al. 2018; Marahatta et al. 2021; Song et al. 2011; Tuo et al. 2016).

The operating mechanism of SWAT model includes the division of entire basins into several sub-basins and finally into HRUs (which is a distinctive combination of slope, LULC, and soil type) based on digital elevation model (DEM). SWAT model generate HRUs using two methods, i.e., generate HRUs for individual sub-basin using the soil and LULC information and multiple HRUs based on threshold values (Arnold et al. 1998). As recommended by Setegn et al. (2009), the current study used 10%, 20%, and 10% threshold for land, soil and slope, respectively. After the successful overlaying of soil, slope, and LULC datasets, the number of sub-basins (HRUs) generated by SWAT model for Gilgit, Hunza, Shatial bridge, Yugo, Doyian, and Bisham Qila are 12 (22), 10 (18), 24 (45), 6 (10), 5 (8), and 9 (17), respectively.

SWAT model has two distinct phases, which are named as land and routing phases. The daily precipitation is used by SWAT model to simulate surface runoff for each HRU using

the Soil Conservation Service (SCS) technique during the land phase hydrological component (USDA, 1972). Green and Ampt infiltration method (Green and Ampt, 1911) is the alternative method to SCS in the SWAT model to simulate surface runoff. Green and Ampt infiltration method require precipitation inputs on a sub-daily scale. Simulated streamflow is routed during the routing phase through streams/river network to the basin outlet using Muskingum or Variable storage techniques.

Input data for SWAT model

SWAT model requires soil properties, LULC and soil maps, and elevation data as input before simulating streamflow. For the current research, DEM with a resolution of 30 m retrieved from shuttle radar topographic mission (SRTM) was downloaded from USGS earth explorer (<https://earthexplorer.usgs.gov/>). The basin delineation and retrieval of required topographic parameters for the SWAT model were acquired from DEM. The LULC map (shown in Fig. 1d) was developed for 2005 using the supervised classification method. Landsat-7 satellite imageries were used for the preparation of LULC map. Landsat 7 ETM+ images have eight spectral bands with 30 m spatial resolution for bands 1–7, while band 8 is a panchromatic band with 15 m spatial resolution. LULC map was developed using the SVM method. For a detailed description of the SVM method, readers are referred to Balkhair and Ur Rahman (2019). The soil data for the study region was extracted from a soil map prepared by the Food and Agriculture Organization (FAO) (<http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/>), which has a resolution of 1:5,000,000 following Rahman et al. (2020a). Harmonized World Soil Database v1.2, combined with the extracted soil map from FAO, was used to acquire the required soil properties for SWAT. Further, the default setting of the SWAT model, i.e., SCS method, Penman–Monteith equation, and variable storage method, is used in the current study to simulate streamflow. Streamflow using SWAT model is simulated by employing water balance approach, which depends temperature and precipitation inputs and can be presented as follows:

$$SW_t = SW_0 + \sum_{i=1}^t (P_{\text{day}} - R_{\text{Surf}} - ET_{\text{day}} - W_{\text{seep}} - R_{\text{gw}}) \quad (1)$$

where SW_t and SW_0 represents the final and initial soil moisture conditions, P_{day} depicts the daily precipitation, R_{Surf} is surface runoff, ET_{day} is the daily evapotranspiration (ET), W_{seep} is water seeped into the ground and R_{gw} is the groundwater return flow. The units of all the above-listed variables are in “mm”. Surface runoff is calculated using SCS-CN

method, while ET is calculated using Penman–Monteith (PM) equation.

Calibration and validation of SWAT model

The calibration and validation (parameter optimization) of SWAT model is performed using Sequential Uncertainty Fitting version 2 (SUFI-2) in SWAT Calibration and Uncertainty Program (SWAT-CUP) (Abbaspour et al. 2015). In order to alleviate the impact of initial conditions and allow a stable SWAT performance, the first three years (1995–1997) were considered as a warm-up period (Tuo et al. 2016). The model is calibrated and validated at multiple sites shown in Table 1 during 1998–2005 and 2006–2013. Besides the data quality, accuracy of SWAT output depends on the careful selection of sensitive parameters. In the current study, sensitive parameters (listed in Table 2) for calibrating and validating the SWAT model were selected from extensive literature review (Abbaspour et al. 2015; Ali et al. 2020; Arnold et al. 2012; Arnold et al. 1998; Bhatta et al. 2019; Duan et al. 2019; Garee et al. 2017; Rahman et al. 2020a; Shah et al. 2020; Shahid et al. 2021; Shrestha et al. 2016). The multi-site calibration of SWAT model is preferred, which produces high accuracy compared with single-site calibration (Rahman et al. 2020a; Shrestha et al. 2016). Therefore, the SWAT model in the current study is calibrated and validated at multiple sites (five interior stations and one outlet station) by following the recommendations of Lerat et al. (2012), Duan et al. (2019), and Rahman et al. (2020a).

The most sensitive parameters are selected by employing Global sensitivity analyses in SWAT-CUP. The initial values for selected parameters were based on physically practical intervals for each parameter suggested in the official documents of SWAT (Arnold et al. 2012) and various studies (Grusson et al. 2015; Tuo et al. 2016). Four iterations with 1000 simulations (total number of 4000 simulations during calibration phase) were performed to calibrate the model with Nash–Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) as an objective function. The parameters range (values) are narrowed down further after every iteration, based on the suggestions from SWAT-Cup (Abbaspour et al. 2004; Abbaspour et al. 2007) and their pre-defined physical ranges. Readers are referred to Abbaspour et al. (2015) for the detailed description of model calibration procedures. The performance of SWAT model was evaluated using several statistical indicators, including NSE, percent BIAS (PBIAS), coefficient of determination (R^2), and mean absolute percentage error (MAPE), as calculated below.

Table 2 List of sensitive parameters, their description, lower and upper bounds and units. “a,” “v,” and “r” represent an absolute increase to default values of selected parameters, replacing the actual

values by new selected values, and relative change to the initial values of selected parameters, respectively

Parameters	Description	Lower bound	Upper bound	Unit
r_CN2.mgt	SCS runoff curve number	35	98	Nil
r_SOL_AWC.sol	Available water capacity of soil layer	− 0.1	1	mm H ₂ O/mm soil
a_SOL_K.sol	Saturated hydraulic conductivity	− 0.5	1	mm/hr
v_CH_K1.sub	Effective hydraulic conductivity	0	300	mm/hr
v_CH_K2.rte	Hydraulic conductivity of the channel	− 0.01	500	mm/hr
v_CH_N2.rte	Manning’s n value for the main channel	0.016	0.2	Nil
r_CANMX.hru	Maximum storage in canopy	0	10	mm
v_SURLAG.bsn	Lag time for surface runoff	1	12	days
v_CH-N1.sub	Manning’s n value for channels	0.016	0.2	Nil
v_SNO50COV.bsn	Snow water equivalent corresponding to 50% snow cover	0	0.9	Nil
v_TIMP.bsn	Snowpack temperature lag factor	0.01	1	Nil
v_SMFMX.bsn	Maximum rate of snowmelt in the year	0	10	mm/C ⁰ -day
v_SMFMN.bsn	Minimum rate of snowmelt in the year	0	10	mm/C ⁰ -day
v_SMTMP.bsn	Snowmelt base temperature	− 5	5	C ⁰

sensitivity analysis, it can be observed that SWAT model is sensitivity to two parameters of each snow-cover, channel flow paths and soil characteristics

$$NSE = 1 - \frac{\sum (Q_{m,i} - Q_{s,i})^2}{\sum (Q_{m,i} - \bar{Q}_m)^2} \tag{2}$$

$$PBIAS = 100 \times \frac{\sum (Q_{m,i} - Q_{s,i})}{\sum Q_{m,i}} \tag{3}$$

$$R^2 = \frac{\left| \sum (Q_{m,i} - \bar{Q}_m)(Q_{s,i} - \bar{Q}_s) \right|^2}{\sum (Q_{m,i} - \bar{Q}_m)^2 \sum (Q_{s,i} - \bar{Q}_s)^2} \tag{4}$$

$$MAPE = \frac{\left| \frac{(Q_{m,i} - Q_{s,i})}{Q_{m,i}} \right|}{N} \times 100 \tag{5}$$

where $Q_{s,i}$ and $Q_{m,i}$ are the simulated and observed streamflow, while the average of observed and simulated streamflow is represented by \bar{Q}_m and \bar{Q}_s respectively.

NSE demonstrates the quantitative difference among observed and simulated streamflow, with the optimal value of 1. PBIAS depicts overestimation or underestimation of simulated streamflow, where the perfect value for PBIAS is 0, the positive and negative values illustrate overestimation and underestimation, respectively. The performance of SWAT model is categorized into

four classes according to criteria defined by Moriasi et al. (2007): unsatisfactory ($NS \leq 0.50$, $|PBIAS| \geq 25\%$), satisfactory ($0.50 < NS \leq 0.65$, $15\% \leq |PBIAS| < 25\%$), good ($0.65 < NS \leq 0.75$; $10\% \leq |PBIAS| < 15\%$), and very good ($NS > 0.75$, $|PBIAS| < 10\%$). According to Lewis (1982), $MAPE < 10$ shows high accurate streamflow simulation, between 10 and 20 shows good simulation, between 20 and 50 shows reasonable simulation, while greater than 50 shows inaccurate simulation.

Artificial Neural Network (ANN)

Artificial Neural Networks have mainly two architectures, i.e., recurrent and non-recurrent. It was demonstrated that for hydrological modeling, non-recurrent type architecture is well suited. Non-recurrent architecture has a single-layer perceptron and multi-layer perceptron (MLP) (Mohammadi et al. 2020). Time series based non-linear hydrological problems such as non-linear precipitation prediction, streamflow and sediment modeling, rainfall-runoff modeling, river stage-discharge modeling, etc., can be easily solved using the MLP. A more detailed description of the MLP method can be found in existing studies, e.g. (ASCE, 2000a, b; Lohani et al. 2012; Nourani, 2014, Kushwaha and Kumar, 2017).

Structural description of the multi-layer perceptron (MLP)

The basic structure of MLP can be broadly divided into three layers, i.e., the input layer followed by a hidden layer and

then the output layer (Fig. 2). The MLP layers are mainly represented by $n-n_{h1}-n_o$, in which n , n_{h1} , and n_o are the number of neurons in the input layer, first hidden layer, and output layer, respectively. The number of hidden layers can be increased or decreased depending on the requirement of suitable architecture, which has numerous processing elements and connections. Neurons in a typical MLP are designed to handle complex processes more efficiently and satisfactorily using several algorithms. The input layer does not have computational nodes, whereas hidden and output layers have computation nodes. The number of hidden layers and neurons in each hidden layer can be increased or decreased for a suitable network architecture requirement.

The crucial role of MLP is to convert a number of inputs into single or multiple outputs. If we consider, x_i ($i=1,2, \dots, m$) are inputs for a pre-organized model for which w_i ($i=1,2, \dots, m$) are the respective weights, which would be changed by error algorithms later on. Under such considerations, the net input to a single node can be expressed by Eq. 6. The activation function “ f ” converts the net input into an output (Eq. 7). The output of any node will behave like an input for the next computational node.

$$\text{net} = \sum_{i=1}^m x_i w_i \tag{6}$$

$$y = f(\text{net}) \tag{7}$$

To determine the weights of the network, two learning mechanisms including supervised and unsupervised have been involved. In the present study, supervised learning has been used in which the user completely knows the input as well as output pattern of the typical architecture of the network. In MLP, the back propagation algorithm (BPA) is used to adjust the weights in each iteration to minimize the mean square error between known and unknown output.

The known output is the observed values of the dataset, and the unknown output is the network computed values of the same dataset. Adjusting weights in each iteration or epoch is performed by two calculations: feed-forward calculations and the back-propagation of errors (also known as the mean square error (MSE)).

The feed-forward calculation occurs in each layer, which is represented by Eq. 8 through Eq. 11. The net input to j th node of the hidden layer is given by:

$$\text{neth}_j = \sum_{i=1}^n w h_{ji} x_i \tag{8}$$

Connection weight always lies between two nodes. For establishing connection weights, this is not necessary that each node should be a computational node. In Eq. 8, connection weight $w h_{ji}$ is a weight between i th node of the input layer and j th node of the hidden layer, although there is no computation on nodes of the input layer. Now, the output of this particular node in the hidden layer is given by Eq. 9.

$$h_j = f(\text{neth}_j) \tag{9}$$

Furthermore, net input to k th node of the output layer is given by Eq. 10.

$$\text{nety}_k = \sum_{j=1}^{n_h} w o_{kj} h_j \tag{10}$$

where $w o_{kj}$ is the connection weight between the computational j th node of the hidden layer and the computational k th node of the output layer. The final output (Eq. 11) from k th node of the output layer is;

$$y_k = f(\text{nety}_k) \tag{11}$$

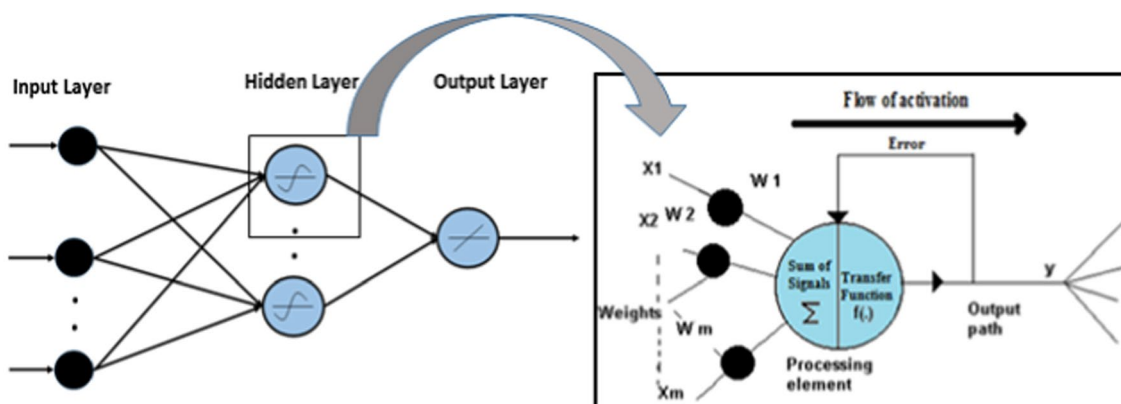


Fig. 2 Example of the MLP model by three layers, viz. input layer, hidden layer, and output layer (Left). A typical processing element with the flow of signals (Right)

After calculating this output, the error between known and unknown output is back propagated using BPA. For a single pair of input and output of a dataset, the sum square error E will be given by;

$$E = \frac{1}{2} \sum_{k=1}^{no} (y_k - t_k)^2 \tag{12}$$

where t_k is the known or observed output at the k th node and y_k is the unknown or calculated output at the same node. Levenberg–Marquardt (LM) learning algorithm is one of the most popular algorithms for minimizing the MSE of a typical MLP. This learning algorithm has been used in this study to train the network. The objective function in this algorithm for minimizing the error function is;

$$\Delta W = (J^T J + \mu I)^{-1} J^T e \tag{13}$$

where J is the Jacobian matrix, e is the error vector, ΔW is the increment of weights, T shows the target output, and μ is the parameter changed or learning rate which is to be updated using the β depending on the outcome. This study uses the Sigmoid activation function to convert the input signal to output signal at all computational nodes of hidden layers and the output layer.

In the current study, MLP has been applied individually for each sub-basin after selecting appropriate meteorological inputs from the gauging stations located inside a sub-basin.

In this context, five model inputs viz. maximum and minimum temperature, maximum and minimum RH, solar radiation, and precipitation have been selected on a lumped basis and used for each station of the sub-basin. If any sub-basin has four gauge stations, then its model inputs have been selected as five multiplied by four, i.e., twenty, to prepare a streamflow model for that particular sub-basin. So, there are 20 inputs and a single output, which is streamflow for the MLP model. Besides these selections, the number of neurons in each hidden layer, transfer function, and learning algorithms is extremely important to obtain a suitable MLP architecture (selection of the number of hidden layers). In this study, several trials have been performed to obtain an appropriate MLP architecture as a schematic diagram shown in Fig. 3.

Results

Streamflow simulation using SWAT model

Sensitivity analyses

The rank, t -test, and p -value for selected parameters achieved using the Global sensitivity analyses across each sub-basin is shown in Table 3. The t -test, which is calculated by dividing the coefficient by standard error, demonstrates the precision for measuring the regression coefficient. In

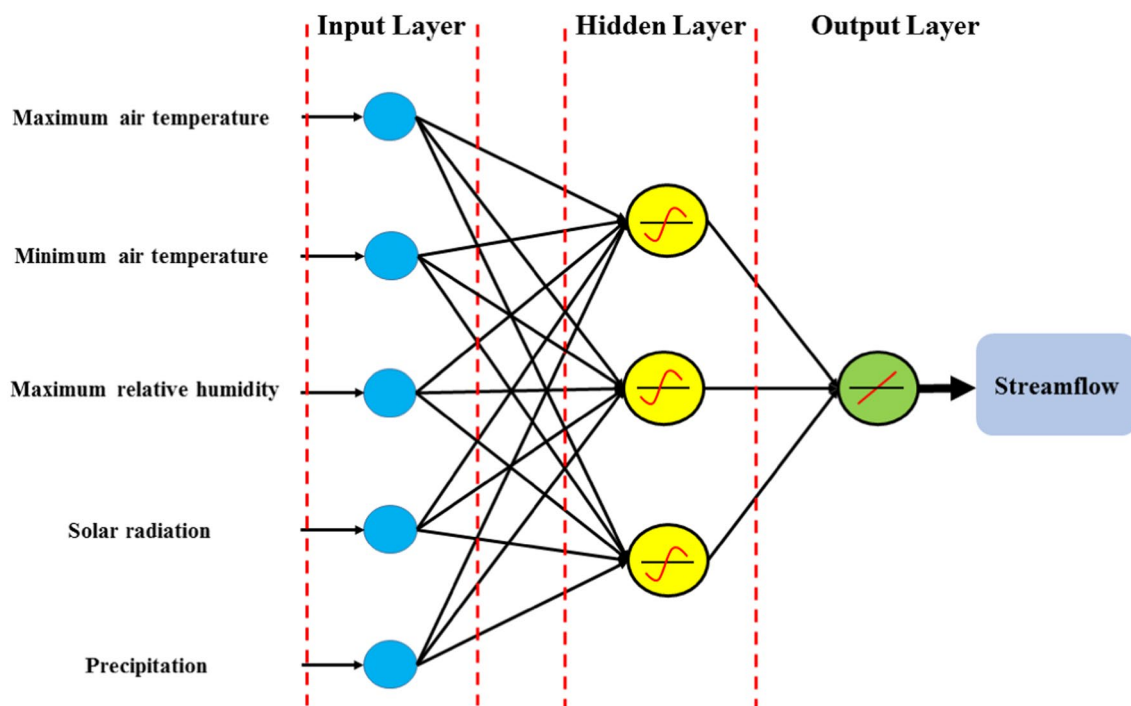


Fig. 3 Schematic diagram for the MLP model used in the current study

another way, *t*-test illustrates the importance of selected parameter (the high magnitude of the absolute *t*-test depicts the most sensitive parameter). On the other hand, *p*-values depict variations in the mean of sample (streamflow observations). The minimum *p*-value and maximum *t*-test value show that the parameter is very sensitive (Abbaspour et al. 2015). The most sensitive parameters (as listed in Table 3) for most sub-basins were CN2, SOL_AWC, SOL_K, TIMP, SMFMX, and CH_K2. The major source of water in UIB is the melting of snow-cover and glaciers; therefore, SWAT model was found sensitive to SMFMX (maximum rate of snowmelt in the year) and TIMP (snowpack temperature lag factor). On the other hand, soil characteristics of plays an important role in the infiltration and surface-runoff. Therefore, the SWAT model in UIB is sensitive to SOL_AWC (available water capacity of soil layer) and SOL_K (saturated hydraulic conductivity). Further, curve number (CN2, across the basins) and hydraulic conductivity (CH_K2, along

the streams) are the two important parameters in stream-flow simulation across the basin. For the sensitivity analysis, it can be observed that SWAT model is sensitivity to two parameters of each snow-cover, channel flow paths and soil characteristics.

Calibration and validation of SWAT model

SWAT model was calibrated at the Gilgit sub-basin of UIB by fitting the initial values for each parameter. Several iterations with 1000 simulations each are performed to get the final values for each parameter (which were narrowed down after each iteration). After calibrating the SWAT model at Gilgit, the model is then calibrated at Hunza, Shatial bride, Yugo, Doyian, and Bisham Qila sub-basins using the initial values for the same set of parameters. Four parameters, i.e., CN2, SOL_AWC, SOL_K, and CH_K2 were initially selected for the calibration process. These parameters were

Table 3 The rank, *t*-test, and *p*-values for selected sensitive parameters across each of the six sub-basins (Gilgit, Hunza, Shatial bride, Yugo, Doyian, Bisham Qila)

Parameters	Gilgit			Shatial bridge			Hunza		
	Rank	<i>t</i> -test	<i>p</i> -value	Rank	<i>t</i> -test	<i>p</i> -value	Rank	<i>t</i> -test	<i>p</i> -value
CN2	1	2.20	0.02	1	2.45	0.01	1	2.3	0.01
SOL_AWC	2	1.76	0.15	2	2.18	0.17	2	2.16	0.10
SOL_K	3	1.51	0.20	3	2.09	0.21	3	2.02	0.18
CH_K1	7	-0.99	0.57	5	1.82	0.39	4	2.06	0.30
CH_K2	11	0.66	0.88	12	0.82	0.81	10	1.35	0.74
CH_N2	8	-0.82	0.71	7	-1.59	0.48	6	-1.88	0.45
CANMX	14	-0.33	0.97	13	-0.77	0.87	11	-1.16	0.83
SURLAG	12	-0.53	0.92	14	-0.61	0.93	12	-1.02	0.86
CH_N1	9	-0.78	0.79	8	-1.38	0.57	9	-1.46	0.66
SNO50COV	13	0.40	0.94	11	0.93	0.71	13	0.93	0.91
TMIP	4	1.34	0.26	4	1.98	0.32	5	1.93	0.41
SMFMX	5	1.22	0.37	6	1.77	0.45	7	-1.74	0.54
SMFMN	10	0.76	0.86	9	1.21	0.61	8	1.60	0.63
SMTMP	6	1.10	0.45	10	1.15	0.66	14	-0.78	0.94
	Yugo			Doyian			Bisham Qila		
Parameters	Rank	<i>t</i> -test	<i>p</i> -value	Rank	<i>t</i> -test	<i>p</i> -value	Rank	<i>t</i> -test	<i>p</i> -value
CN2	1	2.10	0.03	1	2.31	0.05	1	2.23	0.02
SOL_AWC	2	1.97	0.12	2	2.17	0.08	2	2.15	0.06
SOL_K	3	1.90	0.21	3	2.09	0.16	3	2.12	0.12
CH_K1	5	1.73	0.32	4	1.97	0.21	4	-2.01	0.20
CH_K2	11	0.98	0.79	12	0.79	0.85	14	0.73	0.98
CH_N2	8	-1.48	0.59	7	-1.52	0.47	6	-1.77	0.38
CANMX	13	-0.76	0.90	13	-0.73	0.91	12	-0.98	0.89
SURLAG	14	-0.72	0.95	14	-0.64	0.97	11	-1.13	0.84
CH_N1	9	-1.29	0.67	8	-1.37	0.54	8	-1.53	0.59
SNO50COV	12	-0.89	0.85	11	0.93	0.81	10	1.29	0.80
TMIP	6	1.67	0.45	5	1.73	0.30	5	-1.89	0.27
SMFMX	4	1.81	0.27	6	-1.60	0.43	7	1.65	0.45
SMFMN	10	-1.09	0.74	9	1.21	0.66	9	1.40	0.67
SMTMP	7	1.55	0.52	10	1.08	0.73	13	0.82	0.93

found very sensitive to streamflow simulation in a number of studies (Immerzeel and Droogers, 2008; Rahman et al. 2020a; Shen et al. 2008; Shrestha et al. 2016). After every simulation, the remaining parameters (shown in Table 2) are supplemented in groups comprised of five parameters each. The SWAT model was calibrated at each station in the six sub-basins using the final selected list of parameters. The final calibrated fitted parameter values are presented in Table 4.

Evaluation of SWAT model performance

The accuracy (performance) of SWAT model was evaluated using NSE, PBIAS, and R^2 based on the criteria suggested by Moriasi et al. (2007) and MAPE (Lewis, 1982). Table 5 presents the performance of SWAT model during calibration and validation periods. Results showed that the performance of SWAT model during the calibration period in selected sub-basins of UIB ranged from good to very good evaluated in terms of NSE and PBIAS. MAPE showed accurate streamflow simulation across all the sub-basins except for the Indus River at Shatial bridge and Bisham Qila sub-basins during the calibration period. However, the performance of SWAT during the validation

period assessed using NSE and PBIAS ranged from good to very good. MAPE depicted good streamflow simulation for all the sub-basin while reasonable streamflow simulation at Indus River at Shatial bridge. PBIAS shows that the simulated streamflow is overestimated at most sub-basins during the calibration and validation periods except for Gilgit and Bisham Qila (only during the validation period) sub-basins.

Figure 4 shows the daily scale calibration and validation of the SWAT model across Shatial bridge, Yugo, Doyian, and Bisham Qila sub-basins. Results from Gilgit and Hunza sub-basins (the remaining two sub-basins) are presented in the next section in comparison with MLP-based simulated streamflow. Overall, the performance of SWAT model ranged from good to very good during calibration and validation periods. PBIAS shows that SWAT model underestimated the streamflow across Bisham Qila sub-basin during the validation period. The simulated streamflow is mostly underestimated at the peaks during the monsoon season (May to September). Further, the analyses also demonstrated high overestimation during the validation period as compared with the calibration period.

Table 4 List of selected sensitive parameters with the final calibrated values in each sub-basin

Parameters	Gilgit	Hunza	Shatial bridge	Yugo	Doyian	Bisham Qila
CN2	-0.75	-0.71	-0.61	-0.67	-0.81	-0.91
SOL_AWC	0.76	0.72	0.79	0.66	0.71	0.81
SOL_K	0.50	0.49	0.66	0.34	0.46	0.65
CH_K1	276	266	229	262	232	254
CH_K2	148.48	128.48	158.48	134.54	142.34	145.23
CH_N2	0.20	0.23	0.13	0.22	0.18	0.28
CANMX	2.92	2.67	2.61	2.54	2.43	2.12
SURLAG	6.51	6.21	6.12	5.34	5.98	7.3
CH_N1	0.34	0.39	0.33	0.45	0.41	0.34
SNO50COV	0.43	0.41	0.49	0.46	0.55	0.51
TMIP	0.32	0.22	0.21	0.31	0.43	0.22
SMFMX	8.70	7.70	7.76	8.91	7.85	7.23
SMFMN	1.79	2.19	2.29	2.42	2.21	1.87
SMTMP	3.59	2.59	2.59	3.89	3.41	3.32

Table 5 Performance evaluation of SWAT model during calibration and validation period using NSE, PBIAS, R^2 , and MAPE statistical indices

Sub-basins	Calibration (1998–2005)				Validation (2006–2013)			
	NSE	PBIAS	R^2	MAPE	NSE	PBIAS	R^2	MAPE
Gilgit	0.84	-6.98	0.86	9.34	0.77	-7.90	0.79	13.51
Hunza	0.85	3.51	0.83	7.61	0.73	8.40	0.76	12.33
Shatial bridge	0.83	8.61	0.80	15.71	0.71	14.20	0.74	24.03
Yugo	0.84	4.79	0.87	8.50	0.73	9.88	0.74	12.87
Doyian	0.86	3.49	0.84	8.73	0.74	9.04	0.75	11.87
Bisham Qila	0.81	9.84	0.80	10.10	0.72	-8.70	0.73	16.34

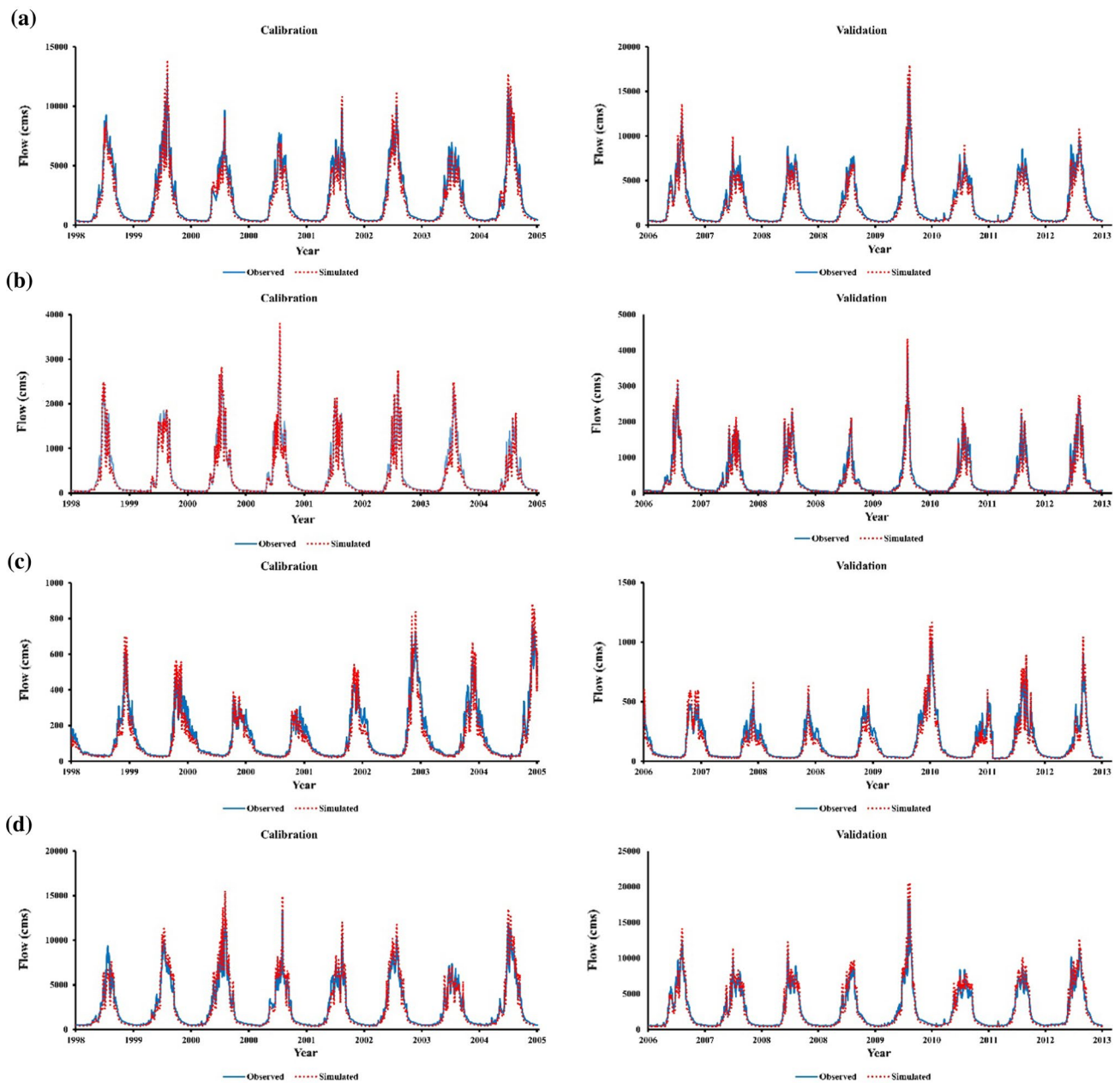


Fig. 4 Calibration (left) and validation (right) of SWAT model across **a** Shatial bridge, **b** Yugo, **c** Doyian, **d** Bisham Qila sub-basins

Streamflow simulation using MLP and its comparison with the SWAT model

Recent studies show that the application of data-driven approaches in hydrological modeling increases, and these models are proved more robust and accurate. Streamflow in the current research was also simulated using MLP across UIB, and the results are compared with the SWAT model (shown in Table 6 and Fig. 5). Table 6 shows that the MLP simulated streamflow with high accuracy as compared with the SWAT model. The performance of MLP is very good

during the calibration and validation periods across all sub-basins evaluated based on criteria specified by Moriasi et al. (2007). In the MLP model, a training dataset was used to tune the parameters of MLP, and the remaining dataset has been applied to check model performance with the unknown testing dataset.

Figure 5 shows the comparison of MLP and SWAT simulated streamflow across Gilgit and Hunza sub-basins. Figure 5a shows that SWAT underestimated the streamflow across the Gilgit sub-basin while MLP has accurately captured the streamflow. A contrasting trend is observed for

Table 6 Evaluation of MLP model performance in the calibration and validation periods using NSE, PBIAS, R^2 , and MAPE statistical indices

Sub-basins	Network architecture	Calibration (1998–2005)				Validation (2006–2013)			
		NSE	PBIAS	R^2	MAPE	NSE	PBIAS	R^2	MAPE
Gilgit	20–5–1	0.991	0.024	0.990	0.102	0.982	0.033	0.984	0.167
Hunza	20–4–1	0.985	–0.037	0.986	0.008	0.982	–0.046	0.984	0.027
Shatial bridge	20–9–1	0.904	0.095	0.925	0.216	0.877	0.105	0.911	0.228
Yugo	20–4–1	0.968	0.049	0.948	0.112	0.952	0.088	0.930	0.201
Doyian	20–5–1	0.981	0.032	0.974	0.088	0.963	0.040	0.943	0.135
Bisham Qila	20–4–1	0.901	0.115	0.890	0.267	0.887	0.131	0.885	0.301

the Hunza sub-basin, where the SWAT model overestimated while MLP underestimated the streamflow. Maximum over/underestimation is observed during the monsoon season having peak flows. The network architecture (shown in Table 6), for example, 20–5–1 in Gilgit sub-basin, represents that there are 20 neurons in the input layer (one neuron per input), five neurons in one hidden layer, and one neuron in the output layer (one neuron per output).

The comparison of SWAT with MLP against the observed flow (considered as a reference) in the validation period is presented with the Tylor diagram (shown in Fig. 6). The results illustrate high performance of the MLP model (minimum standard deviation and maximum correlation coefficient values) across all sub-basins in UIB. The performance of both the models also varies with the area of sub-basins and magnitude of streamflow, i.e., relatively poor performance across Indus River at Shatial bridge and Bisham Qila sub-basins while significantly higher across the remaining sub-basins. Overall, the results demonstrate relatively accurate streamflow estimates across the UIB with a correlation coefficient greater than 0.90 (while 0.85 to 0.90 for Indus River at Shatial bridge and Bisham Qila sub-basins). Moreover, the standard deviation ranges from 169 to 3132 for SWAT and 138 to 2908 for the MLP model.

Discussion

It is an extremely arduous task to perform hydrological modeling (e.g., streamflow simulation) in poorly gauged basins characterized by complex topography and permanent glacier cover. Accurate hydrological modeling often requires a dense network of *in-situ* gauges or stations that can provide *in-situ* observations with high quality as input to hydrological models. However, the *in-situ* gauges/stations are sparsely distributed, especially in developing countries like Pakistan (Rahman et al. 2019), particularly across the complex topographic and diverse climatic regions of UIB. There are several factors associated with the relatively poor performance of different hydrological models in the glacial regions; including the unavailability of enough *in-situ* observations for calibration and validation of hydrological

models (Rahman et al. 2020a), complex topography, the seasonal impact of snow, and glaciers on streamflow and river discharge (Tuo et al. 2018), and climate change (Huss and Hock, 2018; Lettenmaier et al. 2015).

Streamflow simulation in the glacial basin is extremely difficult and important from the perspective of effective water management, and it is considered an unsolved problem, which deserves the attention of hydrologists. Numerous studies have appraised the performance of different hydrological models to simulate streamflow in glacial basins (Chen et al. 2019; Shrestha et al. 2016; Slezziak et al. 2020; Wang et al. 2019; Wortmann et al. 2019). Some studies have employed the SWAT model in basins characterized by glacier and snow cover (Bhatta et al. 2019; Debele et al. 2010; Garee et al. 2017; Grusson et al. 2015; Khan et al. 2018; Luo et al. 2013; Rahman et al. 2013; Shah et al. 2020; Shahid et al. 2021; Troin and Caya, 2014; Tuo et al. 2018). However, these studies reported that the performance of hydrological models varies from satisfactory to good, which is subjected to several factors.

The relatively poor performance of the physical or distributed hydrological model across different basins has shifted the paradigm towards data-driven approaches. In the recent few decades, the application of machine learning approaches, e.g., ANN, ANFIS, and SVM, etc., have been significantly increased due to their high accuracy and robustness. Based on the results obtained, it was found that machine learning models (such as the MLP structure of ANN used in the current study) are time and computationally efficient, which do not require extensive investigations and have no restrictions for the type and number of input selection. To the best of our knowledge, very few studies extensively evaluated the performance of machine learning-based models in glacial regions like UIB. In the current study, the performance SWAT model is comprehensively compared with MLP (ANN-based data-driven model) to simulate streamflow in UIB.

Table 5 and Fig. 4 illustrated that set of parameters obtained through the calibrating and validating SWAT model has a strong influence on streamflow simulation. Sensitivity analyses in hydrological modeling demonstrate the share of each individual parameter in the propagation

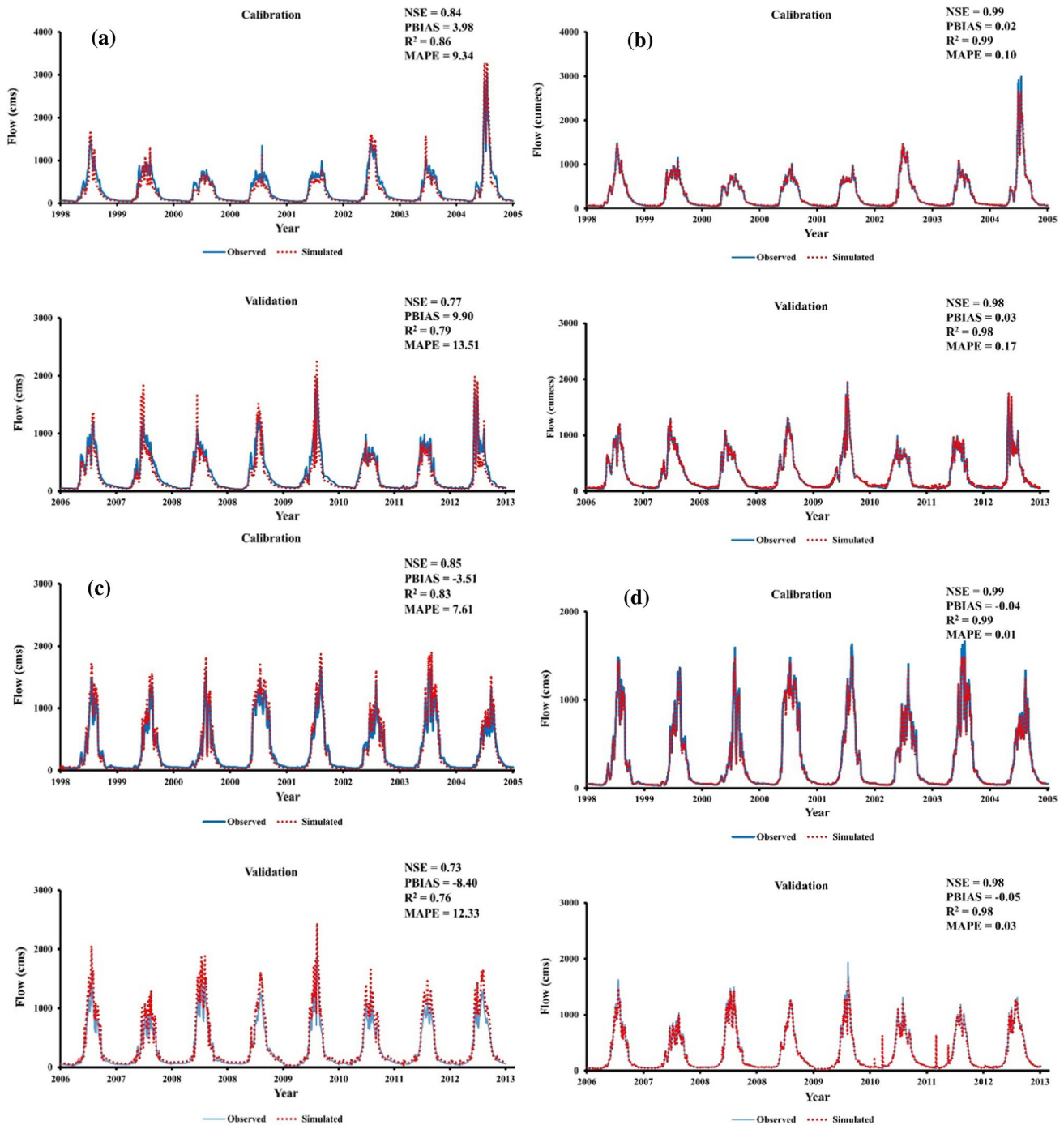


Fig. 5 Comparison of streamflow simulated by SWAT (left panel) across **a** Gilgit and **c** Hunza sub-basins with MLP (right panel) across **b** Gilgit and **d** Hunza sub-basins

of uncertainties in model output. Hence, highly sensitive parameters will result in high shares of model uncertainty compared with less sensitive parameters. Therefore, sensitivity analysis is the first step that must be performed in model calibration. The regionalization procedure (one basin at a time) adopted in the current research helped in selecting sub-basin specific SWAT parameters, which

resulted in good to very good (good to reasonable) performance in streamflow simulation evaluated using NSE, PBIAS, and R^2 (MAPE). On the other hand, the MLP model depicted better performance during both calibration and validation periods as compared to SWAT by matching the observed streamflow, peak flows and presented better statistical indices ($NSE > 0.90$, $PBIAS < 1$, $R^2 > 0.90$,

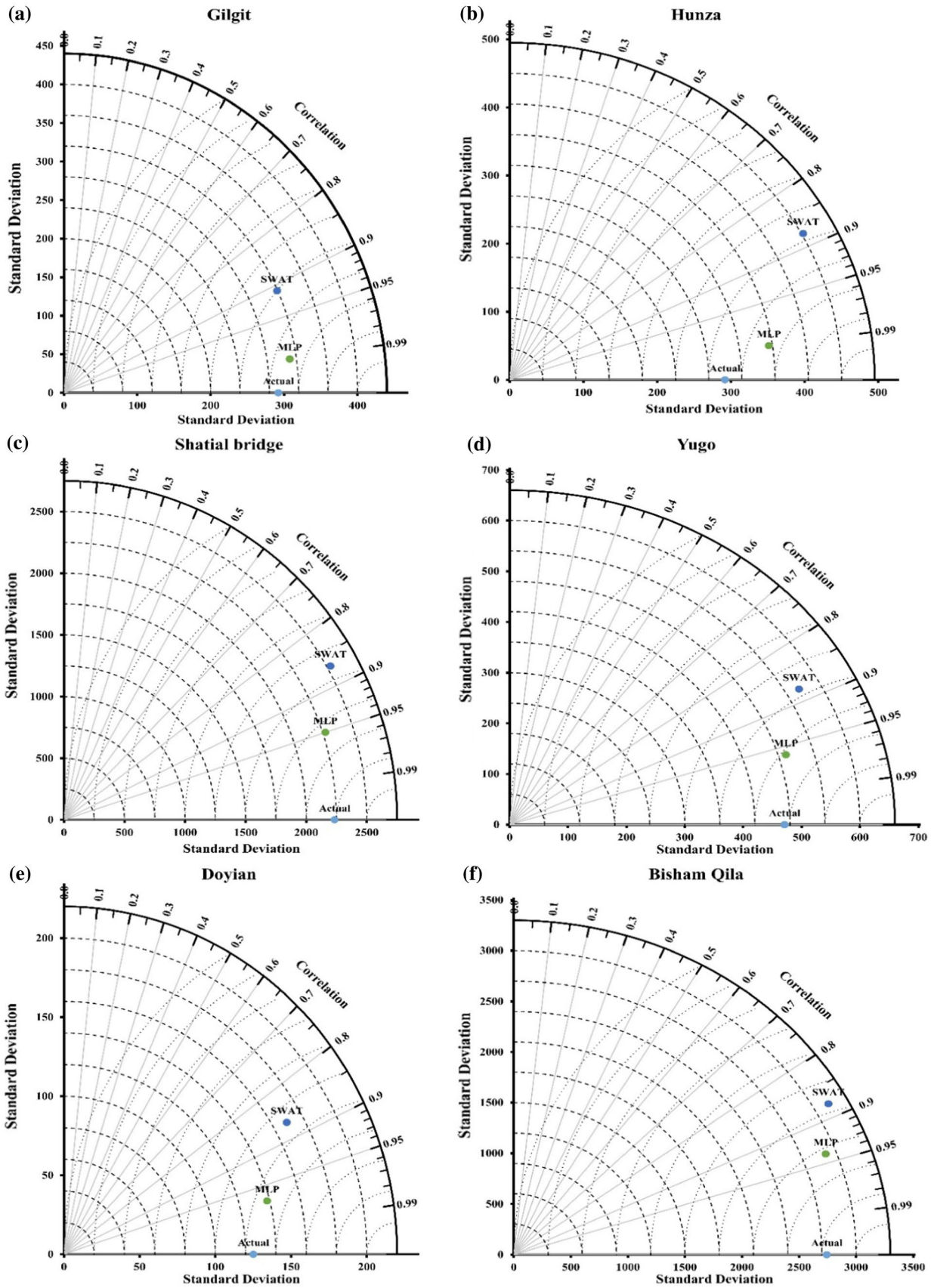


Fig. 6 Performance comparison of SWAT and MLP models with Taylor diagram across all sub-basins

and MAPE < 10%) across all sub-basin except for Shatial bridge and Bisham Qila (shown in Table 6).

The relatively poor performance of the SWAT model as compared with the MLP model might be attributed to several factors: (a) SWAT model may suffer from the identification of the most sensitive parameters (Cibin et al. 2010; Shen et al. 2008), (b) selected set of snow specific parameters for each sub-basin could not fit the snow conditions of each sub-basin, (c) the identified sensitive snow specific parameters might also be influenced by different sources of uncertainties, (d) the potential limitation of SCS-CN method used in SWAT model to simulate streamflow, which produces relatively poor results when there is significant proportion of impermeable land surface as in the case of UIB (Tas-dighi et al. 2018), (e) discharge is accumulating from each sub-basin to the Indus River (Shatial bridge and Bisham Qila) and hydropeaking of discharge cannot be accurately reproduced by SWAT model. Therefore, the performance of the SWAT model in simulating streamflow is relatively poor compared with the MLP model.

During the analyses, few limitations associated with MLP were observed that cannot be ignored. These limitations include the transformation of MLP from input to output can be affected by the techniques having no physical bases. Since MLP is a lumped approach, there may be errors in averaging the sub-catchment parameters. Moreover, empirical models like MLP cannot make spatial predictions within the watershed for processes (e.g., runoff generation, soil moisture, or nutrient export). On the other hand, process-based models like the SWAT model requires many kinds of input data (e.g., elevation, land use/land cover, soil type, drainage, geology, climate data, etc.). However, some of these data are not fully available in many regions because of inevitable problems such as poor distributions observed stations, social and political issues, and restricted data sharing among countries in transboundary basins. In this case, machine learning models like MLP is useful as it is not a data-intensive approach. This study shows that with significantly less modeling effort and resources, still the performance of MLP is better than that of the SWAT model.

Conclusions

Streamflow simulation is extremely important in snow and glacier-dominated Upper Indus Basin (UIB), Pakistan, which serves as a water tower for domestic, agriculture, and industrial use downstream of the Indus River. In this study, streamflow across UIB was simulated using the SWAT model, and its performance was compared with the machine learning-based MLP (Multi-Layered Perceptron) model. The main findings of the current study are listed below:

1. Evaluation with multiple statistical indicators showed SWAT model performed reasonably well in simulating daily streamflow across different sub-basins of UIB, with model performance ranging from “good” to “very good”. However, the performance of SWAT is relatively poor as compared with the MLP model.
2. MLP model captured the streamflow dynamics and peak flows with extremely high accuracy. Evaluation with multiple statistical indicators showed that MLP performed better than SWAT and yielded very good and high accurate streamflow simulation with NSE > 0.90, PBIAS < 1%, R^2 > 0.90, and MAPE < 10% for all the six sub-basins of UIB.
3. The comparatively poor performance of the SWAT model might be associated with several factors, including issues in the identification of sensitive parameters, selected snow parameters that might not fit the snow conditions in sub-basins, and the potential limitation of the SCS-CN method employed to simulate streamflow.
4. The poor performance of the SWAT model in Shatial bridge and Bisham Qila is due to the large size of the sub-basin and accumulation of sub-basin discharges to the Indus River resulting in hydropeaking, which cannot be accurately captured by hydrological models.
5. The results demonstrated that the development of a local hydrological model, e.g., MLP, might suit better in simulating streamflow, which considers the sub-basin specific characteristics.

Keeping in view the high performance and robustness of machine learning-based models, this study recommends the development and evaluation of further machine learning models across UIB. Moreover, in view of the advantages and disadvantages of SWAT and machine learning models, the hybrid models are expected to improve streamflow simulation and our understanding of the hydrological processes in snow-glacier-dominated regions.

Acknowledgements The authors extend their gratitude to the Water and Power Development Authority (WAPDA) for providing streamflow and climate data. The authors are also thankful to Pakistan Meteorology Department (PMD) for providing *in-situ* precipitation data.

Authors contributions KUR, QBP: Project administration, Conceptualization, Writing- original draft, Software, Formal analysis, Visualization. KZJ, MS, and DPK: Formal analysis; Writing- original draft, Visualization. ZD, BM: Data curation, Writing, Review and editing. KMK, DTA: Supervision, Writing, Review, Editing.

Funding This research work was supported by the Shuimu Scholar Program of Tsinghua University (Grant number 2020SM072).

Data availability statement The data that support the findings of this study are available from the author, [Quoc Bao Pham, phambaoquoc@tdmu.edu.vn], upon reasonable request.

Declarations

Conflict of interest None

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent to publish Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbaspour KC, Johnson C, Van Genuchten MT (2004) Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J* 3(4):1340–1352
- Abbaspour KC et al (2007) Modelling hydrology and water quality in the pre-alpine/alpine thur watershed using SWAT. *J Hydrol* 333(2–4):413–430
- Abbaspour KC et al (2015) A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. *J Hydrol* 524:733–752
- Ali WRM, Chen N, Umar WRM, Sundas A, Mahfuzur R (2020) Assessment of runoff, sediment yields and nutrient loss using the swat model in Upper Indus Basin of Pakistan. *J Geosci Environ Prot* 8(9):62–81
- Archer D (2003) Contrasting hydrological regimes in the Upper Indus Basin. *J Hydrol* 274(1–4):198–210
- Arnold JG, Srinivasan R, Muttiah RS, Williams JR (1998) Large area hydrologic modeling and assessment part I: model development 1. *JAWRA J Am Water Resour Assoc* 34(1):73–89
- Arnold JG et al (2012) SWAT: model use, calibration, and validation. *Trans ASABE* 55(4):1491–1508
- ASCE (2000a) Task committee on application of artificial neural networks in hydrology artificial neural networks in hydrology, I: preliminary concepts. *J Hydrol Eng ASCE* 5(2):124–137
- ASCE (2000b) Task committee on application of artificial neural networks in hydrology artificial neural networks in hydrology, II: hydrologic application. *J Hydrol Eng ASCE* 5(2):115–123
- Baffaut C et al (2015) Hydrologic and water quality modeling: spatial and temporal considerations. *Trans ASABE* 58(6):1661–1680
- Balkhair KS, Rahman KU (2019) Development and assessment of rainwater harvesting suitability map using analytical hierarchy process, GIS and RS techniques. *Geocarto Int* 36(4):421–448. <https://doi.org/10.1080/10106049.2019.1608591>
- Beniston M et al (2018) The European mountain cryosphere: a review of its current state, trends, and future challenges. *Cryosphere* 12(2):759–794
- Bhambri R et al (2013) Heterogeneity in glacier response in the upper Shyok valley, Northeast Karakoram. *Cryosphere* 7(5):1385–1398
- Bhatta B, Shrestha S, Shrestha PK, Talchabhadel R (2019) Evaluation and application of a SWAT model to assess the climate change impact on the hydrology of the Himalayan river basin. *CATENA* 181:104082
- Cheema MJM, Bastiaanssen WG (2012) Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. *Int J Remote Sens* 33(8):2603–2627
- Chen Y et al (2019) Uncertainty in simulation of land-use change impacts on catchment runoff with multi-timescales based on the comparison of the HSPF and SWAT models. *J Hydrol* 573:486–500
- Choong CE, Ibrahim S, El-Shafie A (2020) Artificial neural network (ANN) model development for predicting just suspension speed in solid-liquid mixing system. *Flow Meas Instrum* 71:101689
- Cibin R, Sudheer K, Chaubey I (2010) Sensitivity and identifiability of stream flow generation parameters of the SWAT model. *Hydrol Process Int J* 24(9):1133–1148
- Coutu GW, Vega C (2007) Impacts of land use changes on runoff generation in the east branch of the brandy wine creek watershed using a Gis-based hydrologic model. *Middle States Geographer* 40:142–149
- Debele B, Srinivasan R, Gosain A (2010) Comparison of process-based and temperature-index snowmelt modeling in SWAT. *Water Resour Manage* 24(6):1065–1088
- Duan Z et al (2019) Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly Gauged basin in Ethiopia. *J Hydrol* 569:612–626
- Fan H et al (2020) Comparison of long short term memory networks and the hydrological model in runoff simulation. *Water* 12(1):175
- Francesconi W, Srinivasan R, Pérez-Miñana E, Willcock SP, Quintero M (2016) Using the soil and water assessment tool (SWAT) to model ecosystem services: a systematic review. *J Hydrol* 535:625–636
- Garee K, Chen X, Bao A, Wang Y, Meng F (2017) Hydrological modeling of the Upper Indus Basin: a case study from a high-altitude glacierized catchment Hunza. *Water* 9(1):17
- Gassman PW, Reyes MR, Green CH, Arnold JG (2007) The soil and water assessment tool: historical development, applications, and future research directions. *Trans ASABE* 50(4):1211–1250
- Golmohammadi G, Rudra R, Dickinson T, Goel P, Veliz M (2017) Predicting the temporal variation of flow contributing areas using SWAT. *J Hydrol* 547:375–386
- Green WH, Ampt G (1911) Studies on soil physics. *J Agric Sci* 4(1):1–24
- Grusson Y et al (2015) Assessing the capability of the SWAT model to simulate snow, snow melt and streamflow dynamics over an alpine watershed. *J Hydrol* 531:574–588
- Hasson S et al (2014) Early 21st century snow cover state over the Western river basins of the Indus River system. *Hydrol Earth Syst Sci* 18(10):4077–4100
- Hasson S, Böhner J, Lucarini V (2017) Prevailing climatic trends and runoff response from Hindukush–Karakoram–Himalaya, Upper Indus Basin. *Earth Syst Dynam* 8(2):337–355
- He Z, Wen X, Liu H, Du J (2014) A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *J Hydrol* 509:379–386
- Hock R (2003) Temperature index melt modelling in mountain areas. *J Hydrol* 282(1–4):104–115
- Hock R et al (2019) GlacierMIP—a model intercomparison of global-scale glacier mass-balance models and projections. *J Glaciol* 65(251):453–467

- Huss M, Hock R (2018) Global-scale hydrological response to future glacier mass loss. *Nat Clim Chang* 8(2):135–140
- ICIMOD. (2017) An innovative approach to agricultural water management in the upper Indus basin; the water-energy-food nexus at the local level. In: Proceedings International centre for integrated mountain development (ICIMOD).
- Immerzeel W, Droogers P (2008) Calibration of a distributed hydrological model based on satellite evapotranspiration. *J Hydrol* 349(3–4):411–424
- Immerzeel WW, Droogers P, De Jong S, Bierkens M (2009) Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing. *Remote Sens Environ* 113(1):40–49
- Immerzeel WW, Van Beek LP, Bierkens MF (2010) Climate change will affect the Asian water towers. *Science* 328(5984):1382–1385
- IPCC, I.P.o.C.C., (2018). Summary for policymakers of IPCC special report on global warming of 1.5 °C approved by governments.
- Khan AJ, Koch M (2018) Correction and informed regionalization of precipitation data in a high mountainous region (Upper Indus Basin) and its effect on SWAT-modelled discharge. *Water* 10(11):1557
- Khan MZ, Abbas H, Khalid A (2021) Climate vulnerability of irrigation systems in the Upper Indus Basin: insights from three Karakoram villages in Northern Pakistan. *Clim Dev*. <https://doi.org/10.1080/17565529.2021.1944839>
- Koycegiz C, Buyukyildiz M (2019) Calibration of SWAT and two data-driven models for a data-scarce mountainous headwater in semi-arid Konya closed basin. *Water* 11(1):147
- Krajewski WF et al (2017) Real-time flood forecasting and information system for the state of Iowa. *Bull Am Meteor Soc* 98(3):539–554
- Kumar S, Roshni T, Himayoun D (2019) A comparison of emotional neural network (ENN) and artificial neural network (ANN) approach for rainfall-runoff modelling. *Civil Eng J* 5(10):2120–2130
- Kushwaha DP, Kumar D (2017) Suspended sediment modeling with continuously lagging input variables using artificial intelligence and physics based models. *Int J Curr Microbiol App Sci* 6(10):1386–1399
- Latif Y, Yaoming M, Yaseen M, Muhammad S, Wazir MA (2020) Spatial analysis of temperature time series over the Upper Indus Basin (UIB) Pakistan. *Theoret Appl Climatol* 139(1):741–758
- Lerat J, Andréassian V, Perrin C, Vaze J, Perraud JM, Ribstein P, Loumagne C (2012) Do internal flow measurements improve the calibration of rainfall-runoff models?: Rainfall-runoff models calibration with internal flow data. *Water Resour Res*. <https://doi.org/10.1029/2010WR010179>
- Lettenmaier DP et al (2015) Inroads of remote sensing into hydrologic science during the WRR era. *Water Resour Res* 51(9):7309–7342
- Lewis CD (1982) Industrial and business forecasting methods: a practical guide to exponential smoothing and curve fitting. Butterworth-Heinemann
- Liu R, Xu F, Zhang P, Yu W, Men C (2016) Identifying non-point source critical source areas based on multi-factors at a basin scale with SWAT. *J Hydrol* 533:379–388
- Lohani AK, Kumar R, Singh RD (2012) Hydrological time series modeling: a comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques. *J Hydrol* 442:23–35
- Luo Y, Arnold J, Liu S, Wang X, Chen X (2013) Inclusion of glacier processes for distributed hydrological modeling at basin scale with application to a watershed in Tianshan Mountains, Northwest China. *J Hydrol* 477:72–85
- Malagò A et al (2016) Regional scale hydrologic modeling of a karst-dominant geomorphology: the case study of the Island of Crete. *J Hydrol* 540:64–81
- Malik MA and Azam M (2009). Impact evaluation of existing irrigation and agronomic practices on irrigation efficiency and crop yields in Northern areas of Pakistan. Pakistan council of research in water resources
- Marahatta S, Devkota LP, Aryal D (2021) Application of SWAT in hydrological simulation of complex Mountainous river basin (part I: model development). *Water* 13(11):1546
- Mohammadi B et al (2020) Developing novel robust models to improve the accuracy of daily streamflow modeling. *Water Resour Manage* 34(10):3387–3409
- Moradkhani H, Hsu K-L, Gupta HV, Sorooshian S (2004) Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *J Hydrol* 295(1–4):246–262
- Moriiasi DN et al (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 50(3):885–900
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—a discussion of principles. *J Hydrol* 10(3):282–290
- Nguyen HH et al (2019) Comparison of the alternative models SOURCE and SWAT for predicting catchment streamflow, sediment and nutrient loads under the effect of land use changes. *Sci Total Environ* 662:254–265
- Nourani V (2014) A review on applications of artificial intelligence-based models to estimate suspended sediment load. *Int J Soft Comput Eng (IJSCE)* 3(6):121–127
- Parveen S, Winiger M, Schmidt S, Nüsser M (2015) *Erdkunde* 69(1):69–85. <https://doi.org/10.3112/erdkunde.2015.01.05>
- Pradhan P, Tingsanchali T, Shrestha S (2020) Evaluation of soil and water assessment tool and artificial neural network models for hydrologic simulation in different climatic regions of Asia. *Sci Total Environ* 701:134308
- Rahman K et al (2013) Streamflow modeling in a highly managed mountainous glacier watershed using SWAT: the Upper Rhone River watershed case in Switzerland. *Water Resour Manage* 27(2):323–339
- Rahman KU, Shang S, Shahid M, Li J (2018) Developing an ensemble precipitation algorithm from satellite products and its topographical and seasonal evaluations over Pakistan. *Remote Sensing* 10(11):1835
- Rahman KU, Shang S, Shahid M, Wen Y (2019) Performance assessment of SM2RAIN-CCI and SM2RAIN-ASCAT precipitation products over Pakistan. *Remote Sensing* 11(17):2040
- Rahman KU, Shang S, Shahid M, Wen Y (2020) Hydrological evaluation of merged satellite precipitation datasets for streamflow simulation using SWAT: a case study of Potohar Plateau, Pakistan. *J Hydrol* 587:125040. <https://doi.org/10.1016/j.jhydrol.2020.125040>
- Rahman KU, Shang S, Shahid M, Wen Y, Khan AJ (2020b) Development of a novel weighted average least squares-based ensemble multi-satellite precipitation dataset and its comprehensive evaluation over Pakistan. *Atmos Res* 246:105133
- Rostamian R et al (2008) Application of a SWAT model for estimating runoff and sediment in two Mountainous basins in central Iran. *Hydrol Sci J* 53(5):977–988
- Senent-Aparicio J, Jimeno-Sáez P, Bueno-Crespo A, Pérez-Sánchez J, Pulido-Velázquez D (2019) Coupling machine-learning techniques with SWAT model for instantaneous peak flow prediction. *Biosys Eng* 177:67–77
- Shah MI, Khan A, Akbar TA, Hassan QK, Khan AJ, Dewan A (2020) Predicting hydrologic responses to climate changes in highly glacierized and mountainous region Upper Indus Basin. *R Soc Open Sci* 7(8):191957
- Shahid M, Rahman KU, Haider S, Gabriel HF, Khan AJ, Pham QB, Pande CB, Linh NTT, Anh DT (2021) Quantitative assessment

- of regional land use and climate change impact on runoff across Gilgit watershed. *Environ Earth Sci* 80(22):1–18
- Shen Z, Hong Q, Yu H, Liu R (2008) Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region. *China Sci Total Environ* 405(1–3):195–205
- Shrestha MK, Recknagel F, Frizenschaf J, Meyer W (2016) Assessing SWAT models based on single and multi-site calibration for the simulation of flow and nutrient loads in the semi-arid Onkaparinga catchment in South Australia. *Agric Water Manag* 175:61–71
- Sleziak P, Szolgay J, Hlavčová K, Danko M, Parajka J (2020) The effect of the snow weighting on the temporal stability of hydrologic model efficiency and parameters. *J Hydrol* 583:124639
- Solomatine D, See LM, Abrahart RJ (2008) Data-driven modelling: concepts, approaches and experiences. In: Abrahart RJ, See LM, Solomatine DP (eds) *Practical hydroinformatics*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 17–30. https://doi.org/10.1007/978-3-540-79881-1_2
- Song X, Duan Z, Kono Y, Wang M (2011) Integration of remotely sensed C factor into SWAT for modelling sediment yield. *Hydrol Process* 25(22):3387–3398
- Tasdighi A, Arabi M, Harmel D (2018) A probabilistic appraisal of rainfall-runoff modeling approaches within SWAT in mixed land use watersheds. *J Hydrol* 564:476–489
- Troin M, Caya D (2014) Evaluating the SWAT's snow hydrology over a Northern Quebec watershed. *Hydrol Process* 28(4):1858–1873
- Tuo Y, Duan Z, Disse M, Chiogna G (2016) Evaluation of precipitation input for SWAT modeling in alpine catchment: A case study in the Adige river basin (Italy). *Sci Total Environ* 573:66–82
- Tuo Y, Marcolini G, Disse M, Chiogna G (2018) A multi-objective approach to improve SWAT model calibration in alpine catchments. *J Hydrol* 559:347–360
- Ul Hassan S, Pascale S, Lucarini V, Böhner J (2016) Seasonal cycle of precipitation over major river basins in South and Southeast Asia: a review of the CMIP5 climate models data for present climate and future climate projections. *Atmos Res* 180:42–63
- USDA (1972) *National engineering handbook, section 4: hydrology*. US Department of Agriculture, Washington, DC
- Walter MT et al (2005) Process-based snowmelt modeling: Does it require more input data than temperature-index modeling? *J Hydrol* 300(1–4):65–75
- Wang R et al (2019) A review of pesticide fate and transport simulation at watershed level using SWAT: current status and research concerns. *Sci Total Environ* 669:512–526
- WMO (1994) *Guide to hydrological practices: data acquisition and processing, analysis, forecasting and other applications*
- Wortmann M, Bolch T, Buda S, Krysanova V (2019) An efficient representation of glacier dynamics in a semi-distributed hydrological model to bridge glacier and river catchment scales. *J Hydrol* 573:136–152. <https://doi.org/10.1016/j.jhydrol.2019.03.006>
- Yang S et al (2020) A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. *J Hydrol* 590:125206
- Yao T et al (2012) Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nat Clim Chang* 2(9):663–667
- Yaseen M, Ahmad I, Guo J, Azam MI, Latif Y (2020) Spatiotemporal variability in the hydrometeorological time-series over Upper Indus River Basin of Pakistan. *Adv Meteorol* 2020:1–18. <https://doi.org/10.1155/2020/5852760>
- Yin J, He F, Xiong YJ, Qiu GY (2017) Effects of land use/land cover and climate changes on surface runoff in a semi-humid and semi-arid transition zone in Northwest China. *Hydrol Earth Syst Sci* 21(1):183–196
- You Q, Min J, Kang S (2016) Rapid warming in the Tibetan Plateau from observations and CMIP5 models in recent decades. *Int J Climatol* 36(6):2660–2670
- Zhang X, Srinivasan R, Debele B, Hao F (2008) Runoff simulation of the headwaters of the yellow river using The SWAT model with three snowmelt algorithms 1. *JAWRA J Am Water Resour Assoc* 44(1):48–61
- Zhuhua LV, Zuo J, Rodriguez Dr (2020) Predicting of runoff using an optimized swat-ann: a case study. *J Hydrol Reg Stud* 29:100688. <https://doi.org/10.1016/j.ejrh.2020.100688>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.