





Ensemble rainfall–runoff modeling of physically based semi-distributed models using multi-source rainfall data fusion

Tagesse Gichamo ^{a,b,*}, Vahid Nourani ^c, Hüseyin Gökçekuş ^b and Gebre Gelete ^{a,b,d}

^a College of Agriculture and Environmental Science, Arsi University, Assela 193, Ethiopia

^b Faculty of Civil and Environmental Engineering, Near East University, Nicosia/TRNC, Mersin-10 99138, Turkey

^c Center of Excellence in Hydroinformatics and Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

^d Environmental and Atmospheric Sciences Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Nasiriyah 64001, Iraq

*Corresponding author. E-mail: tagesseg@gmail.com

 TG, 0000-0002-4306-3867; VN, 0000-0002-7833-1032; HG, 0000-0001-5793-4937; GG, 0000-0001-5045-1094

ABSTRACT

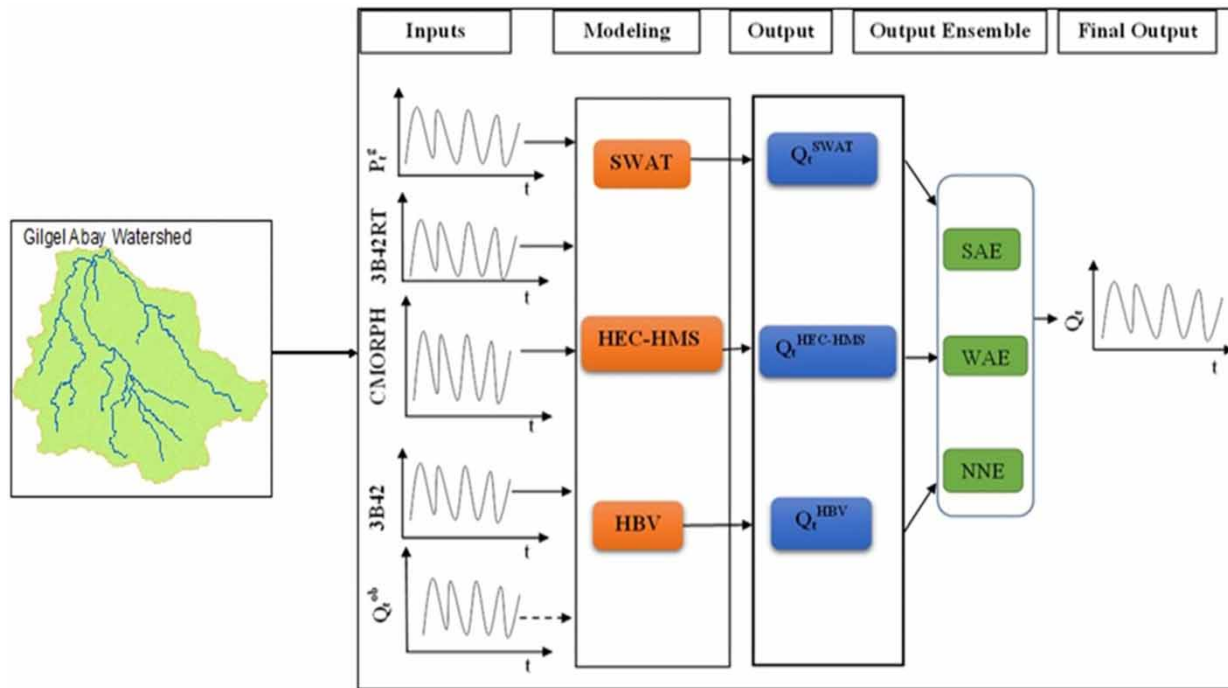
This study was aimed at ensemble rainfall–runoff modeling by Soil and Water Analysis Tool (SWAT), the Hydrologic Engineering Center's Hydraulic Modeling System, and Hydrologiska Byråns Vattenbalansavdelning of Gilgel-Abay watershed, Blue Nile basin, Ethiopia. For modeling, daily rainfall datasets of five gauges and three satellites, streamflow, and spatial data were used. In the modeling stage, first, the runoff was simulated separately using the rainfall data of gauges, satellites, and their fusion. Second, ensemble rainfall–runoff simulation of the rainfall dataset fusion-based runoff result was conducted via the proposed weighted average, simple average, and neural network (NNE) ensemble techniques. The results exhibited that all models are good in capturing the rainfall–runoff relationship; however, SWAT showed slight superiority by Nash–Sutcliffe efficiency of 0.807 and 0.821 for gauge and fusion data, respectively. The rainfall fusion ensemble model revealed significant improvement over modeling by satellite rainfall owing to the bias-correcting gauge rainfall over satellite rainfall. The NNE technique enhanced the efficiency of the low-performing satellite-rainfall-based model by 17.5% and the rainfall fusion-based model by 13.3% at the validation stage. In general, the result of this study points out that the rainfall datasets' fusion from multi-sources would be worthy for the rainfall–runoff simulation of ungauged basins.

Key words: ensemble modeling, Gilgel-Abay, HBV, HEC-HMS, rainfall–runoff modeling, SWAT

HIGHLIGHTS

- In this study, rainfall–runoff modeling was carried out by Soil and Water Analysis Tool, Hydrologic Engineering Center's Hydraulic Modeling System, and Hydrologiska Byråns Vattenbalansavdelning.
- Spatial, streamflow, satellite, and gauge rainfall datasets were used as inputs for modeling.
- Runoff was simulated separately using the rainfall data of gauges, satellites, and their fusion.
- Ensemble techniques were applied to further improve the efficiency of the models.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Rainfall-runoff modeling is a frequently used process to describe a watershed's hydrological response to what would occur as the result of precipitation. It has been widely used to realize regional hydrological characteristics and their spatial and temporal variations due to climatic and land use/land cover changes (Guzha *et al.* 2018; Leong & Yang 2020). Precise multiple-time-scale (hourly, daily, weekly, monthly, and seasonal) flow forecasting is valuable for the effective planning and operation of hydro-power generation, reservoirs, sediment transport, irrigation management, water supply and sanitation, and related hydrologic studies (Araghinejad *et al.* 2006; Solomatine & Shrestha 2009). Several models have been developed for rainfall-runoff simulation such as distributed physical-based, conceptual, and artificial intelligence-based black-box models (Nourani *et al.* 2021a, 2021b). Physically based models are usually applied to estimate the internal sub-processes and physical behavior of the hydrologic cycle. Physically based models can consider simple linear laws, and assume time-varying, non-linear, and deterministic parameters. Physically based models are further categorized as fully distributed and semi-distributed physically based models depending on input data requirements and the physical processes undergone. In fully distributed models, data with spatial variation and computational algorithms are used to evaluate the rainfall-runoff process (Gebre 2015). Moreover, they require a large quantity of input data (sometimes unavailable). Semi-distributed physically based models, however, are partially subjected to spatial variation by dividing the catchment into smaller sub-basins and they require fewer input data than fully distributed models (Orellana *et al.* 2008). Some examples of the semi-distributed physically based models are the Sacramento Soil Moisture Accounting Model (SAC-SMA), Soil and Water Assessment Tool (SWAT), Storm Water Management Model (SWMM), and flood hydrograph package of the Hydrologic Engineering Center of the US Army Corps of Engineers (HEC-1).

SWAT is a semi-distributed physically based model practiced for the hydrological simulation of streamflow at the basin level (Arnold *et al.* 1998; Arnold & Fohrer 2005). The SWAT model has been proved as a powerful and efficient tool not only for streamflow simulation but also for flood prediction, nutrient transporting, soil erosion, and land cover change modeling (Busico *et al.* 2020). The Hydrologic Engineering Center's Hydraulic Modeling System (HEC-HMS) characterizes rainfall-runoff processes in a dendritic watershed in time and space (Kwin *et al.* 2016). The model has been effectively applied in various catchments (Mandal *et al.* 2016; Young *et al.* 2017).

Hydrologiska Byråns Vattenbalansavdelning (HBV) is a semi-distributed conceptual rainfall-runoff model that uses a simple continuity equation to model rainfall-runoff and the physical relationships of variables dividing the catchment into smaller sub-basins (Ciupak *et al.* 2019). HBV has been effectively applied for several hydrological modelings such as stream-flow simulation, climate change analysis, and water level prediction (Uhlenbrook *et al.* 2010; Pervin *et al.* 2021).

In all such rainfall-runoff trends, precipitation is the basic element of the hydrological processes and it can play a substantial role in the modeling process. One of the main sources of uncertainty in hydrological modeling is related to inappropriate rainfall occurrences in space and time. Ground-based rain gauges, satellite-based precipitation measurement techniques, and radar are the common approaches to obtain precipitation for rainfall-runoff modeling. Ground-based rain gauges are the principal source of obtaining precipitation data, as they are a direct measurement method. However, gauge rainfall may not be realistic because of the sparse spatial resolution of rain gauge networks, especially in developing countries, orographic effects (Mondal *et al.* 2018; Nourani *et al.* 2021a, 2021b), and its continuous data availability that could also be affected by natural or man-made impacts. The rainfall estimating radar systems often have a limited spatial range and they are not affordable in developing countries because the installation and maintenance demands are high (Thorndahl *et al.* 2017). However, satellite-captured precipitation datasets have emerged as consistent, economical, and continuous rainfall sources for rainfall-runoff modeling (Gebremichael *et al.* 2014).

Several satellites have been launched for the purpose, for instance, the Tropical Rainfall Measuring Mission (TRMM) was sent to space in 1997, and the Global Precipitation Measurement (GPM) Core Observatory was launched in 2014 that estimates near-real-time precipitation and snowfall (Yong *et al.* 2015). The Climate Prediction Center (CPC) morphing technique (CMORPH) has been in operation since 1998 to estimate near-real-time precipitation (Gebremichael *et al.* 2014).

Uncertainties in rainfall-runoff simulations that might arise from the input data have been practically managed by calibrating and assimilating the input data (Kumar *et al.* 2015). In catchments where gauging stations could not adequately represent the area, the fusion of gauges and multi-source satellite precipitation data was revealed to have high performance in rainfall-runoff modeling because the bias in satellite rainfall estimation could be corrected by gauging rainfall during fusion. The models could not all equally perform on rainfall-runoff modeling, and they have their merits in some aspects and shortcomings in others. Therefore, ensemble modeling could optimize modeling by conjoining the benefits of each model in the calibration phase and improving the overall efficiency of rainfall-runoff modeling. The current study aims at ensemble rainfall-runoff modeling by physically based semi-distributed models (SWAT, HBV, and HEC-HMS) using the fusion of multi-source satellites and gauge rainfall for the Gilgel-Abay catchment. For this, three ensemble techniques, namely, neural network ensemble (NNE), weighted average ensemble (WAE), and simple average ensemble (SAE) techniques, were developed using the runoff results of the semi-distributed physically based models. The SWAT, HBV, and HEC-HMS models were chosen among different physically based models due to their free available documentation and software, which make them useful in poor countries where there are limited funds for commercial software. Moreover, these models are popular and provide reliable hydrologic modeling results in tropical regions (Tibangayuka *et al.* 2022; Gelete *et al.* 2023). Similarly, the SAE and WAE techniques were selected due to their straightforwardness and simplicity. The NNE technique was used as a nonlinear model combination technique in this study due to its popularity, compatibility, and high accuracy in previous studies (Nourani *et al.* 2021a, 2021b).

Gilgel-Abay is among the essential sub-catchments of the Ethiopian side of the Blue Nile River, which provides a great share of water to the river and is crucial to stabilizing the environment and the catchment hydrology socio-economic patterns of the people located in the regional countries. In the Gilgel-Abay catchment, the rain gauge meteorological stations' distribution is spatially uneven, and temporally, it has only a short period time series. Furthermore, the topography and landscape of the basin are inconstant and fluctuate from mountains to valleys, which might lead rainfall to be subjected to orographic influences, as well as to biased and inappropriate depiction of the rainfall features (Gebre 2015).

2. MATERIALS AND METHODS

2.1. Study area

The Gilgel-Abay watershed is positioned in northern Ethiopia at 10.92°–11.31° N latitude and 36.92°–37.14° E longitude and has an area of 1,635 km² (Figure 1). It is one of the Lake Tana basin sub-catchments that can drain over 60% of streamflow to the Lake Tana catchment (Wale *et al.* 2009). Most of the basin is dominated by highland topographic classes, and the altitude ranges from 1,866 to 3,543 m above the mean sea level (see Figure 1). The catchment is characterized by flat to mild slopes

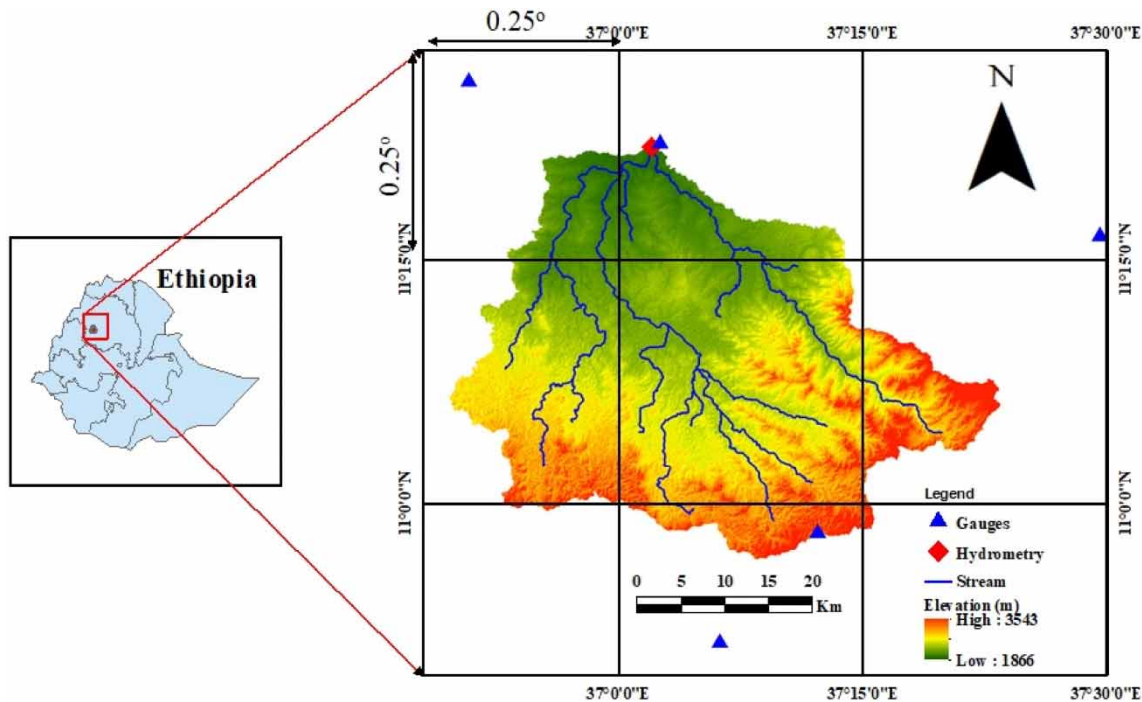


Figure 1 | Study area boundary with satellite rainfall grids and elevation.

ranging from 0% to 43%. The catchment is influenced by a cool and humid climate with an annual average temperature of 17–22 °C. Major rainfall occurs from early June to mid-September while October to May exhibits dry climatic conditions and the annual average rainfall is 1,420 mm. The catchment has one hydrometric station situated at the watershed outlet. The clay-loam, clay, and silt-loam soil textures are uniformly distributed in the catchment with 33% in each class. Haplic luvisols is the major soil type of the basin and mixed land use pattern has been practiced with 74% intensive agriculture, 15% grasslands, and 11% woodlands and forests.

2.2. Input datasets

All three models (SWAT, HEC-HMS, and HBV) prominently used both spatial and temporal data. In this study, the basic inputs for watershed delineation into sub-basins and hydrological response units (HRUs) are a digital elevation model (DEM), a land use land cover (LULC) map, and a soil map. The 12.5 m × 12.5 m DEM for the study area was downloaded from the Alaska satellite facility service (<https://search.asf.alaska.edu/#/>). The 2020 LULC map that was downloaded from ESA Sentinel-2 imagery is a 10 m × 10 m high-resolution raster and provides detailed land use categories (Karra *et al.* 2021).

Daily precipitation as well as the maximum and minimum temperatures of five gauging stations (Adet, Dangila, Gundel, Sekela, and Wetet Abay) for 19 years (2000–2018) were obtained from the National Meteorological Agency, and daily discharge time series measured at the outlet of the mainstream were collected from the Ministry of Water, Energy, and Irrigation of Ethiopia. The data series were checked for homogeneity by the double mass curve method. In addition to the gauge rainfall data, different satellite-based datasets for the same length of years were downloaded from the database and used as inputs for rainfall–runoff modeling.

Satellite-estimated precipitation data is an alternative source of data for ungauged catchments or ground-based spatial gauges distribution that is coarse. The TRMM was launched in 1997 by the joint effort of the Japan Aerospace Exploration Agency (JAXA) and the National Aeronautics and Space Administration (NASA) to estimate rainfall in tropical regions (Le *et al.* 2020). A TRMM precipitation sensor satellite includes precipitation radars, microwave imaging devices, as well as infrared and visible ray scanners. The TRMM precipitation estimation steps are as follows: first, the collected raw data are geo-located and calibrated; second, the area's physical and geographical features are updated with the resolution and location of raw data; third, the temporally averaged data are mapped uniformly in time and space grids.

CMORPH is another technique that has been operated since 1998 to estimate near-real-time rainfall that is commonly provided 18 h later for observation and has been developed by the United States National Oceanic and Atmospheric Administration (NOAA). In this study, two version 7 TRMM products (3B42RT and 3B42) and CMORPH satellite rainfall products were used for rainfall–runoff simulation. The 3B42RT rainfall dataset is available six hours after real-time and covers the globe at 60°N–60°S latitude, and the gauge adjusted 3B42 is the research post-real-time version that is available two weeks after the end of each month with global coverage of 50°N–50°S latitude (Li *et al.* 2018); 3B42RT uses the combined datasets from the TRMM precipitation radar and the microwave image sensor for the calibration of rainfall information derived from a low-Earth-orbit radiometer and microwave sensors.

CMORPH estimates rainfall from the amalgamation of low-orbit passive microwave satellite scans and infrared data derived from geostationary orbit at high spatial and temporal resolutions. CMORPH estimates relatively high-quality rainfall because the intensity and shape of the rainfall are morphed during the time between microwave scans, by applying time-weighted linear interpolating techniques between microwave-obtained information that has been propagating forward and backward in time from the previous and following microwave scans, respectively. CMORPH uses precipitation from passive microwave observation and then transmits the information in space by motion vectors originating from geostationary infrared information (Dinku *et al.* 2007).

CMORPH estimates precipitation using the information derived from low-orbit satellite microwave estimation exclusively and that information is conveyed by spatial propagation features derived entirely from geostationary infrared datasets. The infrared information serves as a transport medium for microwave-derived precipitation data when microwave data are not accessible at a particular position. The propagation vector matrices are formed by calculating spatial lag correlations on consecutive images from geostationary infrared satellites that are then used to propagate the microwave-derived precipitation data. First, the time sequences of data motion are controlled from infrared ray data; then this information is used to deliver displacement motion for morphing from one instantaneous microwave estimate to the other. Based on these processes, CMORPH is associated with retrieved super-precision of passive microwave and greater temporal and spatial resolutions of infrared ray data.

All satellite rainfall products used in this study are available at $0.25^\circ \times 0.25^\circ$ spatial and daily time resolutions and the whole study area is covered by eight satellite rainfall grids (see Figure 1). The three satellite rainfall products were picked because they have previously shown worthy results in earlier studies (Bitew & Gebremichael 2011; Bitew *et al.* 2012).

2.3. Proposed methodology

In this study, three different physically based models, namely, SWAT, HEC-HMS, and HBV models, were used for rainfall–runoff modeling of the Gilgel-Abay watershed. Each of the models requires different inputs for rainfall–runoff modeling. This study involved two stages of modeling for rainfall–runoff prediction. First, SWAT, HEC-HMS, and HBV physically based models were individually applied to model the rainfall–runoff process using satellite and gauge rainfall datasets separately and the fusion of them as well. Second, the runoff output of each model was applied as inputs for ensemble modeling via NNE, SAE, and WAE modeling techniques.

2.4. Used rainfall–runoff models

In the current study, the SWAT, HEC-HMS, and HBV models were proposed for rainfall–runoff modeling. These models are semi-distributed and could explore the spatially variable hydrological characteristics by discretizing the basin into smaller sub-basins and further into HRUs. In this way, the detailed internal sub-processes of rainfall–runoff for each hydrologic unit could be better visualized, accounted for, and aggregated at the final drainage point so that the output would be accurate.

SWAT is operated on a watershed scale and a semi-distributed physically based hydrological model that is applied for hydrological modeling on a daily or sub-daily time-scale (Arnold *et al.* 1998). It was developed by the United States Agricultural Department and Texas A&M University to simulate the effects of soil, land use, and farming practices in a long-time hydrological response over the catchment. SWAT divides the basin into several smaller sub-basins which are also further divided into numerous HRUs, each of which is characterized by unique land use, soil, topographic, and slope features (Leong & Yang 2020). Rainfall–runoff modeling in SWAT can be executed in two different phases. In the first phase all the available hydrological and climatic inputs are applied to simulate runoff in each sub-basin. The second phase is the routing step where the flow from each sub-basin is connected to the main channel and the cumulated flow moves to the final

outlet of the basin. The rainfall–runoff modeling in SWAT is based on the following water balance equation:

$$SW_f = SW + \sum_{i=1}^t [R_{\text{day}} - (Q_{\text{surface}} + ET + W + Q_{\text{ground}})] \quad (1)$$

where SW_f is the final soil moisture availability in a day (mm), SW is the initial soil moisture content in a day (mm), t is time (day), R_{day} is the daily precipitation (mm), Q_{surface} is the daily surface runoff (mm), ET is the combination of daily evaporation and transpiration (mm), W is daily percolation (mm), and Q_{ground} is the daily groundwater discharge (mm).

SWAT reproduces surface flow by utilizing either the soil conservation service (SCS) runoff curve number (CN) technique (Gassman *et al.* 2007) or the Green and Ampt infiltration process. In this study, the SCS runoff CN technique was applied to reproduce surface flow. The land use/cover and soil raster (Figure 2) were layered to develop a rainfall–runoff model for the basin. In the SWAT model setup, first, the watershed was delineated from the DEM, and then, LULC, soil, and slope maps were overlaid (see Figures 2 and 3) and HRUs were generated that provided details about soil attributes, slope, and land use. After creating HRUs and loading weather data, the rainfall–runoff simulation for the basin could be executed.

HEC-HMS is a semi-distributed conceptual model developed by the US Army Corps of Engineers that can model the rainfall–runoff process of dendritic watersheds (Feldman 2000). The model has been updated several times and different versions are available; hence, HEC-HMS version 4.7 was used in this study. The model is applied to simulate four hydrological processes such as loss, transform, baseflow, and routing models. The model determines effective rainfall in the catchment by characterizing the antecedent soil moisture conditions (Young *et al.* 2017), determined using Equation (2).

The SCS-CN loss calculation method was selected to determine direct runoff because it has been widely applied in various watershed conditions and has produced better results as compared with the initial and constant loss methods:

$$P_e = \frac{(P - I_a)^2}{P - I_a + S} \quad (2)$$

where P_e is the excess rainfall and P is the accumulated rainfall depth at time t , I_a is the initial abstraction (e.g., infiltration loss), and S is the potential maximum retention. S (mm) and I_a can be determined by the SCS

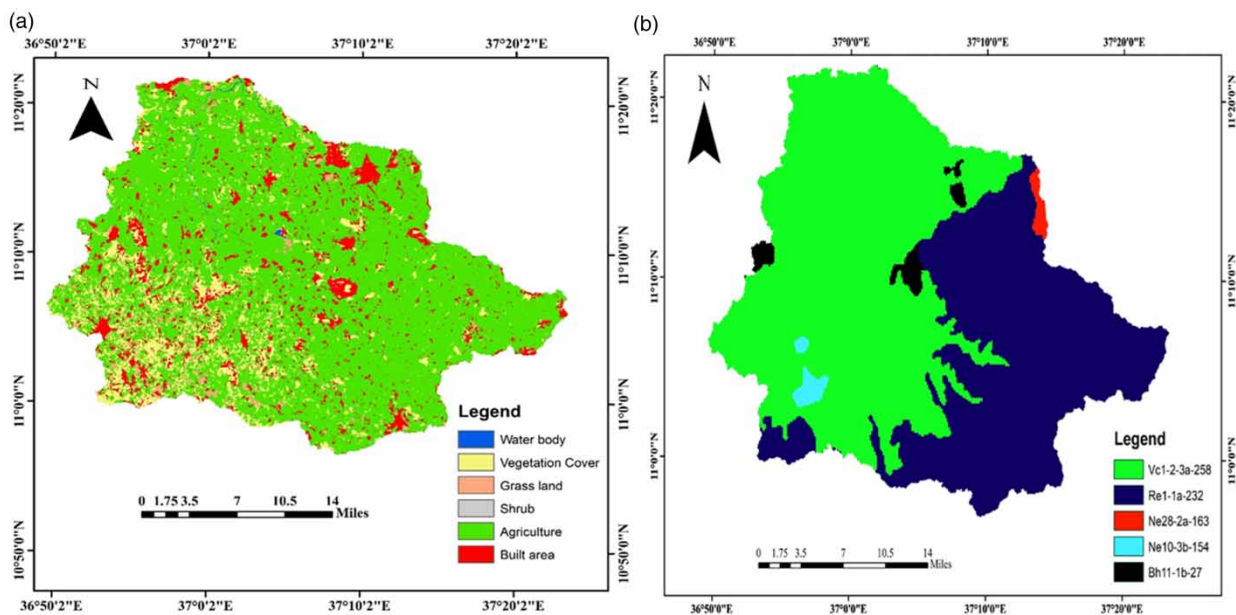


Figure 2 | (a) Land use and (b) soil map of the study area.

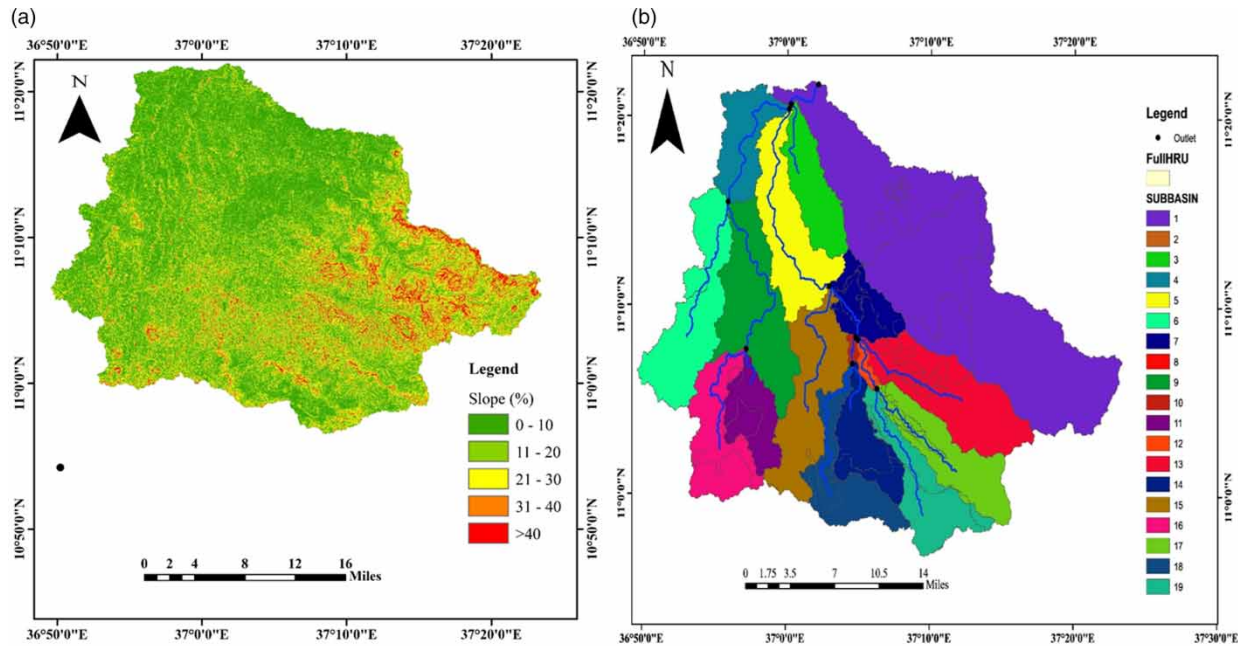


Figure 3 | The SWAT model set up parameters: (a) slope and (b) HRUs.

method based on CN as:

$$S = \frac{25,400}{CN} - 254 \quad (3)$$

$$I_a = 0.2S \quad (4)$$

The CN ranges from 1 to 100 and it is function of LULC and the soil hydrological group. For this study, CN was determined from the global hydrologic soil group and soil class texture grid by the zonal statistics tool on ArcGIS. The peak discharge (Q) is calculated as:

$$Q = \frac{(P - 0.2S)^2}{(P - I_a + S)} \quad (5)$$

HEC-HMS applies several transform methods to convert rainfall into runoff. In this study, the SCS unit hydrograph with a standard graph type was used as the transform method because it requires only lag time as an input parameter. In the current study, baseflow recession and Muskingum methods were applied as baseflow and routing models, respectively. The Muskingum method is a simple and lumped flow routing technique that determines the outflow hydrograph at the outlet point. The HEC-HMS basin model was developed for the basin (Figure 4), and runoff was simulated at the outlet point.

HBV is another conceptual semi-distributed hydrologic model, developed by the Swedish Meteorological and Hydrological Institute (SMHI) to simulate the rainfall-runoff process (Lindström *et al.* 1997). In this study, the HBV light version was used for rainfall-runoff simulation based on different elevation and vegetation zones. The model simulates daily discharge using daily rainfall, temperature, long-term potential monthly evapotranspiration, observed discharge, land use, vegetation zones, and elevation classes as inputs. The general water balance equation of the model is described below:

$$P - E - Q = \frac{d}{dt} [SP + SM + UZ + LZ + L], \quad (6)$$

where P is the precipitation, Q is the discharge, E is the evapotranspiration, SM is the soil moisture, SP is the snowpack, UZ is the discharge contribution from the upper reservoir, L is the lake volume, and LZ is the discharge contribution from the lower reservoir.

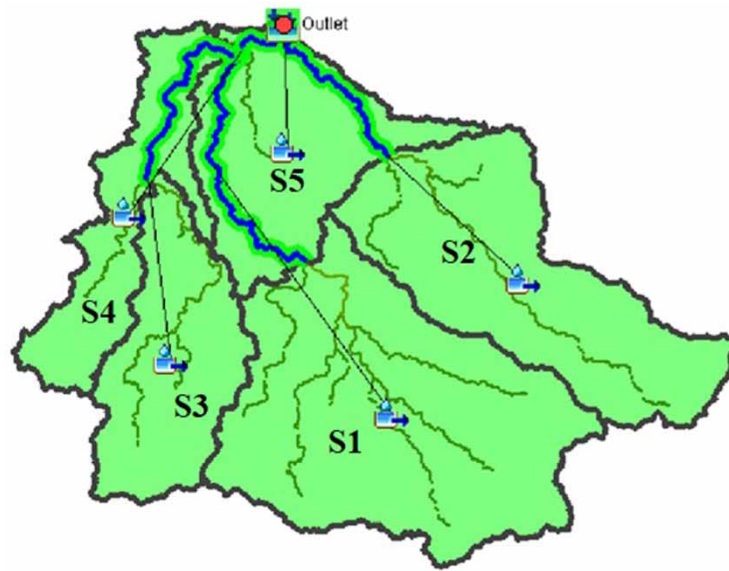


Figure 4 | The HEC-HMS basin model for the watershed.

The operation of the HBV model generally consists of various routines, each representing different hydrological processes (see Figure 5). The HBV rainfall–runoff modeling procedures consist of various hydrological routines representing snow, soil moisture, response, and routing (see Figure 5).

The runoff components are computed by three linear reservoir equations: Q_0 (direct runoff component), Q_1 (intermediate runoff component), and Q_2 (base runoff component) using recession coefficients K_0 , K_1 , and K_2 , respectively. The soil moisture subroutine is based on the parameters beta (β) (shape coefficient for nonlinear storage properties of the soil zone), maximum soil storage (FC), maximum percolation (PERC), and limit of potential evaporation (LP). Beta controls the influence of precipitation on the response function. MAXBAS (length of weighing function) is used as a transformation function to compute outflow from the catchment.

Snow routine represents the snowmelt process and its contribution to streamflow. The parameters involved in this routine are threshold temperature (TT), degree-day factor (CFMAX), a snowfall correction factor (SFCF), water-holding capacity (CWH), and refreezing coefficient (CFR). However, snow routine is not considered in this study because there is no snow in the study area. The changes in soil moisture (SM) and groundwater contribution are measured by soil moisture routine based on the quantity of flow coming from the preceding routine (P) and FC. In the response function, storage from an upper reservoir (S_{UZ}) and a lower reservoir (S_{LZ}) are aggregated to the final runoff.

The elevation zones of the area were analyzed from high-resolution SENTINEL-1 DEM (downloaded from <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1>). The area has been divided into five elevation zones at appropriate intervals (Figure 6). The vegetation classes of the area are defined into three, based on the land use map: grassland, mixed farm, and forest.

SWAT can model water quantity, quality, agricultural management, and climate change in combination; however, it needs extensive input datasets, and its representation of HRUs in each sub-basin is non-spatial (Glavan & Pintar 2012). The SWAT model has a groundwater modeling routine and algorithm that estimates the contribution of groundwater in rainfall–runoff modeling. The spatial distribution of land use, soil, topographic, and climatic datasets in SWAT modeling could make the model a powerful tool for capturing land use and climate change effects in rainfall–runoff modeling (Jaiswal *et al.* 2020).

Similar to the other conceptual models, HBV requires minimal input data, needs minimum computational time, and often comes up with comparable results to complex models. Nevertheless, the model is not suitable to simulate runoff in changing climatic conditions (Fowler *et al.* 2016). The model structures need further improvements to extensively quantify the effects of climate and LULC that could affect long-term catchment water balance (Huang *et al.* 2019).

The HEC-HMS model has the advantages of simple structure and comprehensive contemplation of climatic conditions, and can select various computational techniques adapted to various watersheds and datasets. HEC-HMS is very effective

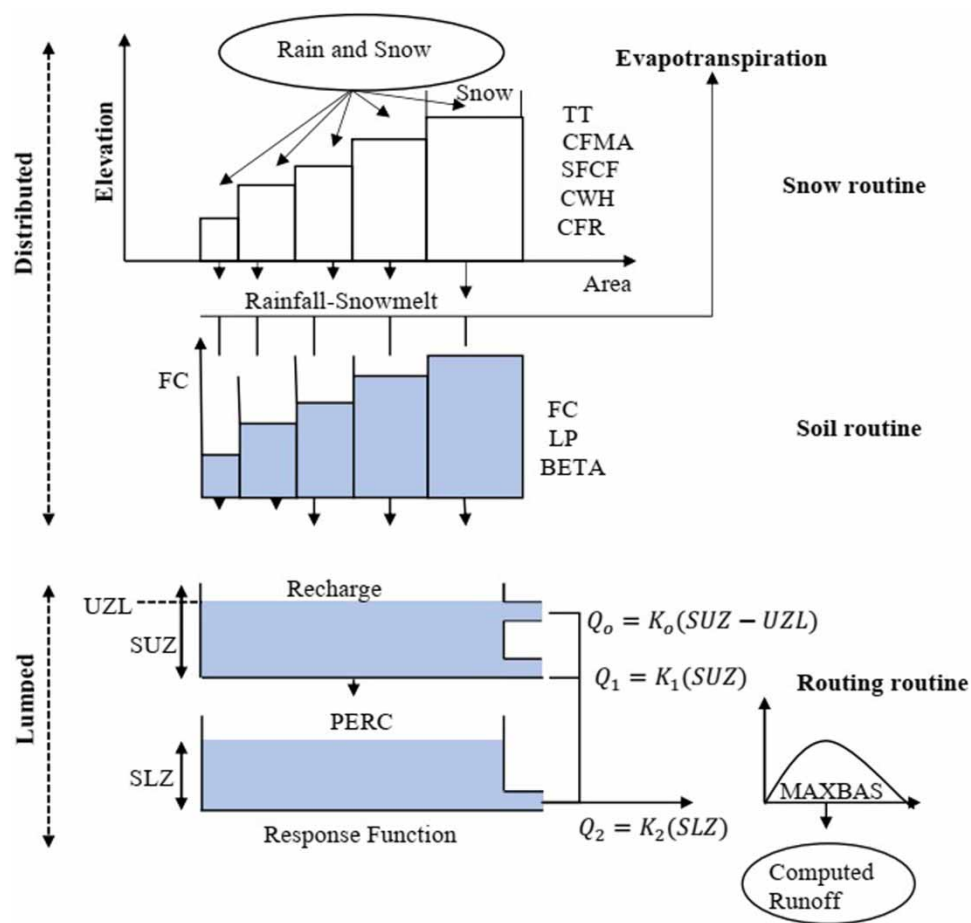


Figure 5 | The modified schematic representation of the HBV light model (Driessen *et al.* 2010).

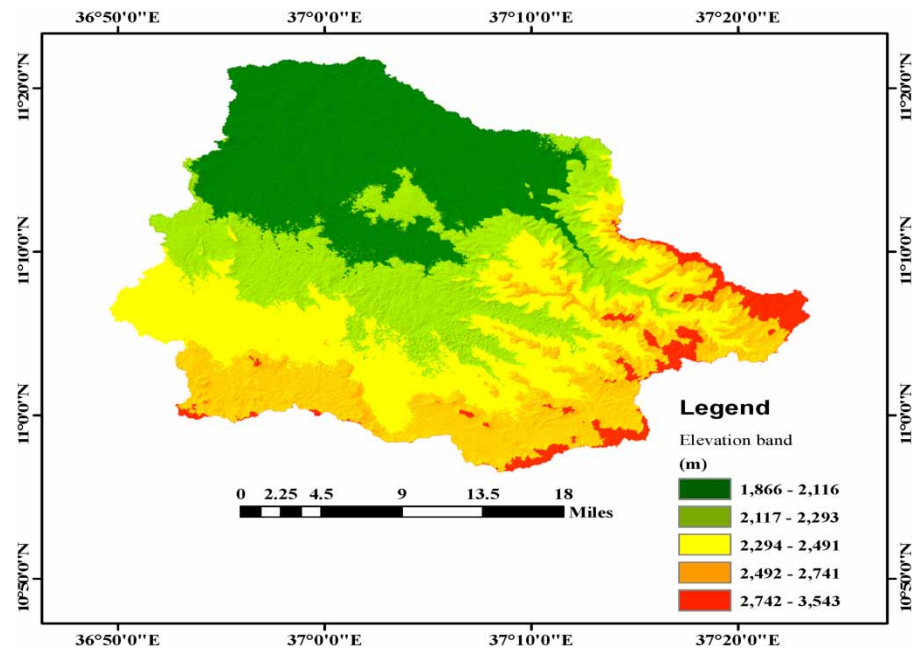


Figure 6 | Elevation zones of the catchment for the HBV model.

in short-time-event-based simulations of peak flow hydrographs characterized by rapid rise and recession (Xin *et al.* 2019) but it tends to overestimate flows for long-time-series streamflow simulations.

An ensemble of multiple models could produce a good robust simulation via a better representation of model structures and by reducing associated uncertainties (Velázquez *et al.* 2011). Therefore, ensemble modeling is important to reduce uncertainties resulting from individual model structures, input data, and model parameters, and it takes over the merits of each model.

2.5. Performance evaluation criteria

The performance of the proposed modeling could be evaluated using the standard evaluation criteria such as the Nash–Sutcliffe coefficient of efficiency (NSE) and the root mean square error (RMSE) as follows:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (Q_o - Q_s)^2}{\sum_{i=1}^N (Q_o - \bar{Q}_o)^2} \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_o(t) - Q_s(t))^2} \quad (8)$$

where NSE is the Nash–Sutcliffe criterion, RMSE is the root mean square error, Q_o is the observed discharge, N is the number of observations, \bar{Q}_o is the average of the observed discharge, and Q_s is the predicted discharge at time t .

2.6. Ensemble unit

For the same input datasets, one model may be more efficient than the others and when several input data are used, the outputs of different models can be totally different. To take advantage of each model without losing the general character of the data, the ensemble method can be used for rainfall–runoff simulation. In ensemble modeling, the individual model's result is used as the input with a confident significance level assigned to each with the support of an arbitrator to offer the output (Kiran & Ravi 2008). The output amalgamation of numerous models provides outputs that are more accurate than individual models and improve the final model performance. In ensemble modeling, the output from each model is believed to best describe data sources that may be different from the other models, and the fusion of information from different sources may allow augmenting of all valuable input characteristics in the modeling. To boost modeling efficiency, numerous ensemble techniques have been applied, for instance, SAE, NNE, WAE, least-square, and random forest regression. These ensemble methods have been applied for the prediction of rainfall–runoff, dissolved oxygen concentration (Abba *et al.* 2020), ground-water level (Sharafati *et al.* 2020), and streamflow simulation (Shiru & Park 2020). In this study, SAE, WAE, and NNE techniques are employed to enhance the modeling capability of individual models. The employed ensemble techniques were selected because they consume less time to model and produce better results.

2.6.1. SAE

In SAE, the arithmetic average of runoff outputs of SWAT, HEC-HMS, and HBV are used to compute the final runoff output as:

$$\bar{Q}_o = \frac{\sum_{i=1}^n Q_{oi}}{n} \quad (9)$$

where \bar{Q}_o is SAE runoff, Q_{oi} is runoff from the i th model, and n is the number of individual models involved (in this study, n is 3).

2.6.2. WAE

WAE assigns a weight factor (w_i) for each individual model output based on the relative significance of the outputs as:

$$\bar{Q}_o = \sum_{i=1}^n w_i Q_{oi} \quad (10)$$

where w_i can be determined from a model's performance evaluation criteria, e.g., NSE as:

$$w_i = \frac{NSE_i}{\sum_{i=1}^n NSE_i} \quad (11)$$

2.6.3. NNE

NNE is the nonlinear ensemble technique that can be trained by a feed-forward neural network (FFNN) with a back-propagation (BP) algorithm using outputs of individual models as inputs; each of them is dispensed to a neuron of the input layer. In this ensemble technique, the best model topology and maximum epoch are obtained by several trial-and-error methods and the sigmoid might be deliberated as a hidden and output activation function. Other nonlinear ensemble kernels such as gene expression programming (GEP) and adaptive neuro-fuzzy inference systems (ANFIS) can also be used, but FFNN was chosen for this study because of its simplicity and rapid training capability and because it provides comparable results to other non-linear ensemble techniques.

3. RESULTS AND DISCUSSION

The proposed modeling in the current study includes two stages: (i) separate rainfall-runoff modeling by SWAT, HEC-HMS, and HBV using gauge and satellite rainfall separately, and gauge and satellite rainfall data fusion; and (ii) ensemble modeling via a nonlinear (NNE) and two linear (SAE, WAE) ensemble techniques, to boost the efficiency of individual models. The outputs of single models are used as inputs for ensemble modeling.

3.1. Results of sensitivity analysis

Sensitivity analysis is a vital process in modeling because it detects the crucial parameters and their importance level that are used for model calibration and validation.

For SWAT, the sequential uncertainty fitting (SUFI-2) algorithm of the SWAT-CUP standalone tool was applied for global sensitivity analysis and the calibration of SWAT model parameters (Abbaspour 2015). The SUFI-2 algorithm accounts for different types of uncertainties that arise from model conceptualization, observed datasets, and parameters (Singh *et al.* 2013).

The global sensitivity of the runoff parameters was determined by Latin hypercube regression analyses, and the minimum and maximum ranges of parameters were then fitted for calibration using the SUFI-2 uncertainty methods. In this algorithm, each parameter range should be changed for all iterations until the objective function is achieved. The parameters sensitive to runoff are listed (Table 1) and used to deploy the proper range of parameters that could produce the preeminent result as compared with the observed flow.

For HEC-HMS, the most common sensitivity analysis methods are partial derivation and changing the values of parameters one at a time (Hamby 1994). Perturbing the parameter values $\pm 25\%$ to $\pm 30\%$ with a 5% interval indicated a good result. In

Table 1 | The SWAT hydrologic parameters used for sensitivity analysis

Parameter	Description	Lower bound	Upper bound	Optimal value	Process
CN2*	*Initial SCS runoff curve number for moisture condition II	-0.2	0.2	0.04	Runoff
GW_DELAY	Groundwater delay time (days)	35	450	45	Groundwater
ALPHA_BF	Baseflow alpha factor-baseflow recession constant	0.01	1.2	0.68	Runoff
SOL_AWC	Available water capacity of soil layer (mm H ₂ O/mm soil)	-0.8	0.8	0.16	Soil channel
HRU_SLP	Average slope steepness	0	1	0.18	HRU
GW_REVAP	Groundwater 'revap' coefficient	0.02	0.2	0.05	Groundwater
Surlag	Surface runoff lag coefficient	1	24	1.21	Runoff
SOL_K	Saturated hydraulic conductivity	0.7	0.8	0.72	Soil
SOL_BD	Moist bulk density	0.9	2.5	0.12	Soil

*This parameter is more sensitive than the others.

this study, the sensitivity of parameters for HEC-HMS modeling was analyzed by altering the values of the parameters in the range of $\pm 25\%$ with 5% intervals until the best agreement was achieved between observed and simulated values. This method tested and updated one parameter at a time while the other parameters were kept constant. The basic parameters are Muskingum coefficients (k and x), infiltration coefficient, CN, baseflow, and initial abstraction values.

For HBV, the Monte Carlo automatic method was applied to the best objective function that generates random parameters in a range within predefined model parameters' sensitivity analysis. Before calibration, the model parameters were defined based on upper and lower bounds that could represent catchment characteristics. The parameters were calibrated in a specified range until the objective function was achieved (Table 2).

Automatic calibration was applied for global sensitivity analysis based on a minimum p -value and an absolute maximum t -stat for SWAT model sensitivity analysis. The parameters for calibration were selected based on previous studies (Khelifa *et al.* 2017; Chen *et al.* 2019). Accordingly, nine parameters were selected based on their significance in streamflow simulation and were ranked (see Table 3). The parameters were identified as the most sensitive if their t -stat was higher and the p -value was lower. The gauge and satellite rainfall-based rainfall were calibrated for the same initial parameter setups. As the result indicated, the parameters are ranked from the highest to the least according to their impact on streamflow for both datasets (see Table 3). The sensitivity analysis result indicates that the CN is the most sensitive parameter for runoff simulation.

Sensitivity analysis for HEC-HMS rainfall-runoff modeling was carried out to trigger input parameters that highly affected the model's performance. The sensitivity of parameter analysis was deployed by changing the values of each parameter one at a time between -25% and $+25\%$, until the objective function fitted. Thereafter, the sensitive parameters were identified based on their effects on the simulated runoff. The most sensitive parameters detected were CN, lag time, initial abstraction, and Muskingum coefficients (k and x) (Table 4). As the values of these parameters varied, the best agreement was achieved

Table 2 | HBV parameters' upper and lower ranges

Parameters	Description	Unit	Lower	Upper
UZL	Reservoir threshold	mm	0	100
MAXBAS	Base of weight function	day	13	24
LP	Soil moisture threshold for evaporation reduction	–	0.4	0.7
FC	Soil moisture storage	mm	100	1,000
PERC	Percolation to groundwater	mm/day	0	0.25
K_0	Recession coefficient	day ⁻¹	0.05	0.5
K_1	Recession coefficient (upper storage)	day ⁻¹	0.01	0.1
K_2	Recession coefficient (lower storage)	day ⁻¹	0.001	0.1
β	Shape coefficient	–	1	6

Table 3 | SWAT parameters' sensitivity analysis and their ranks

Parameter	P-value	t-stat	Rank
CN2	0	–56	1
ALPHA_BF	0	11	2
Sol_K	0	6.2	3
HRU_SLP	0	3.72	4
SOL_BD	0	–3.62	5
SOL_AWC	0.24	–1.3	6
GW_REVAP	0.3	–1.25	7
GW_DELAY	0.41	–0.95	8
SURLAG	0.55	0.62	9

Table 4 | The sensitive parameters of the HEC-HMS model

Parameters	Range	Optimum	Rank
CN	62–92	82	1
Initial abstraction (mm)	2–6	4.2	2
Lag time (hours)	2–4	3	3
Muskingum_ <i>k</i>	0.5–2	1.15	4
Muskingum_ <i>x</i>	0.25–2	1	5

between observed and simulated values. The reason behind this could be that the most runoff-causing parameters such as topographic, soil, and LULC information are lumped together in a single CN.

The automatic Monte Carlo optimization method was deployed to detect the most sensitive parameters of HBV rainfall–runoff modeling. The upper and lower limits of the model parameters were set for the Monte Carlo simulation of 700,000 runs for the study watershed. The model parameters in the predefined range were standardized and the minimum values of ranges were considered as the sensitive parameters for rainfall–runoff simulation. The Monte Carlo modeling output was utilized to detect the least degree of model objective function variations. The parameter value was considered for the model efficiency ($NSE \geq 0.5$ – 0.95) of this study area. The values of the given parameters corresponding to the defined range of model efficiency were standardized. Hence, the higher range of parameter values with the standardized scope was regarded as less sensitive and the lower range of parameter values with the standardized scope was regarded as highly sensitive to model efficiency. The ranges of sensitive parameters are ranked and presented in Table 5 for HBV-based modeling.

The result indicated that CN is the most dominant parameter and that it strongly influenced the runoff generated from each HRU of the catchment. The CN value depends on soil type, hydrological group of soil, and land use trend. The land use in this study area is predominantly generic agricultural, and it is subjected to destructive agricultural practices that can lead to land use changes. Perhaps, this is the main reason why CN is the most sensitive parameter in the watershed.

3.2. Results of single models

All three models (SWAT, HEC-HMS, and HBV) were applied using both gauge and satellite rainfall datasets. The models were calibrated and validated based on 19 years of daily rainfall time series and discharge data. The results of calibration and validation for each model as well as ensemble modeling are described here.

The SWAT model was set up with watershed delineation using high-resolution DEM, and then HRUs and sub-basin features were generated. The predefined spatial datasets such as LULC and soil data for the catchment were used to execute HRU definitions. The SWAT divided the catchment into sub-basins and then further divided it into smaller HRUs to lump the watershed into small units with similar LULC, soil, and slope characteristics of the watershed. The catchment was discretized into sub-basins based on watershed features' similarity, and defining HRUs to reasonable quantity was done as suggested by Setegn *et al.* (2008). Accordingly, the whole catchment was configured into six sub-basins and 19 HRUs (see Figure 3(b)).

Climatic data were loaded on the SWAT database and the model automatically calculated the Thiessen polygon average of the parameters for each sub-basin. After the model configuration, two SWAT models (one gauge and one for each of the satellite-based rainfall data) were run from 2000 to 2018 considering the first two years as a warm-up period. For calibration, 2008–2018 rainfall and discharge data were utilized and 2002–2008 data were used for the validation of the model. To obtain

Table 5 | HBV parameters' sensitivity to rainfall–runoff modeling

Parameter	Standardized range	Rank
FC	0.42–0.83	1
K_2	0.67–0.23	2
β	0.71–0.27	3
LP	0.72–0.23	4

the best fitting between observed and simulated data, calibration was conducted both manually and automatically by the SUFI-2 algorithm on SWAT-CUP. The parameter values were iterated within a permissible range until the observed and simulated flows were well-fitted, and the ranks are presented in Table 3. The performance of daily streamflow modeling by SWAT is presented in Table 6 driven by various rainfall datasets. In terms of NSE (Table 6 and Figure 7), the gauge-based simulation exhibited better performance with NSE values of 0.834 and 0.807 in the calibration and validation phases, respectively. Among the satellite-driven datasets, CMORPH slightly outperformed 3B42RT and 3B42 daily streamflow simulations. The reason behind the lower performance of streamflow simulation using satellite-based datasets could be the bias of rainfall magnitude estimation captured by satellites. As indicated by Nourani *et al.* (2021a, 2021b), 3B42 and 3B42RT underestimate most of the peak rainfall, which could reduce the performance of the simulation. The runoff simulated by the models using both data sources is presented in Figure 8. For SWAT modeling, the validation result with $NSE > 0.5$ is generally an acceptable performance (Whittemore 2002). In this study, the SWAT performance for both gauge and satellite datasets is far better than the acceptable range in terms of NSE. Therefore, it is rational to conclude that SWAT can successfully simulate rainfall-runoff by using satellite datasets.

For modeling by HEC-HMS, the model was configured, the watershed was delineated, and then, the sub-watersheds were reproduced from DEM (Figure 4). The basin, meteorological, and control specification models were created before running the model. The watershed was divided into five sub-basins and the physiographic characteristics of each sub-basin were determined (see Table 7). For the simulation of HEC-HMS, soil CN was determined from the global hydrologic soil group and soil-class texture grid by the zonal statistics tool on ArcGIS. The soil hydrologic groups for the watershed were identified as A, B, and D and the portion of area for each was 33%, 12%, and 55%, respectively. The weighted CN value was found to be 83–85. Sensitivity analysis was conducted to detect the influence of each parameter on the runoff simulation performance of the model and the results are as shown in Table 4. Similar to SWAT, HEC-HMS was run for gauge and satellite datasets. The daily runoff simulation performance measures of the models, in terms of NSE and RMSE, are given in Table 6 for both datasets. In terms of both performance measures, the model well-simulated runoff for gauge datasets having 0.777 NSE at the validation stage (see Figure 9(b)). The model performance was also good for CMORPH datasets with an NSE of 0.750 at the validation phase. However, the result exhibits a slight overestimation of high flows and an underestimation of low flows (see Figure 8); this is the usual limitation of hydrological models (Zhang & Savenije 2005) because of uncertainties resulting from model structures and input data. Scatter plots of observed versus simulated runoff using gauge and satellite data for SWAT, HEC-HMS, and HBV are shown in Figure 7.

The HBV model was configured for five elevation groups (Figure 6) and three vegetation classes by using nine model parameters (see Table 2) and the Monte Carlo runs were applied to simulate the streamflow. The automatic calibration and validation of the model were carried out using 2000–2018 daily rainfall (gauge and satellite) and temperature data, in addition to the spatial datasets. The first two years of data were used to spin up the model, 2002–2007 and 2008–2018 data were used

Table 6 | Rainfall-runoff results of SWAT, HEC-HMS, and HBV for gauge and satellite datasets

Model	Rainfall source	NSE		RMSE	
		Calibration	Validation	Calibration	Validation
SWAT	Gauge	0.834	0.807	31.624	34.312
	CMORPH	0.815	0.785	32.542	35.929
	3B42RT	0.794	0.761	48.042	51.672
	3B42	0.785	0.754	49.524	52.432
HEC-HMS	Gauge	0.801	0.777	33.714	36.834
	CMORPH	0.796	0.750	35.342	38.436
	3B42RT	0.776	0.752	37.434	42.652
	3B42	0.764	0.746	39.092	45.382
HBV	Gauge	0.814	0.782	32.654	36.482
	CMORPH	0.803	0.762	34.748	38.085
	3B42RT	0.788	0.758	36.647	41.247
	3B42	0.775	0.751	38.541	44.262

Note: RMSE is in m^3/s .

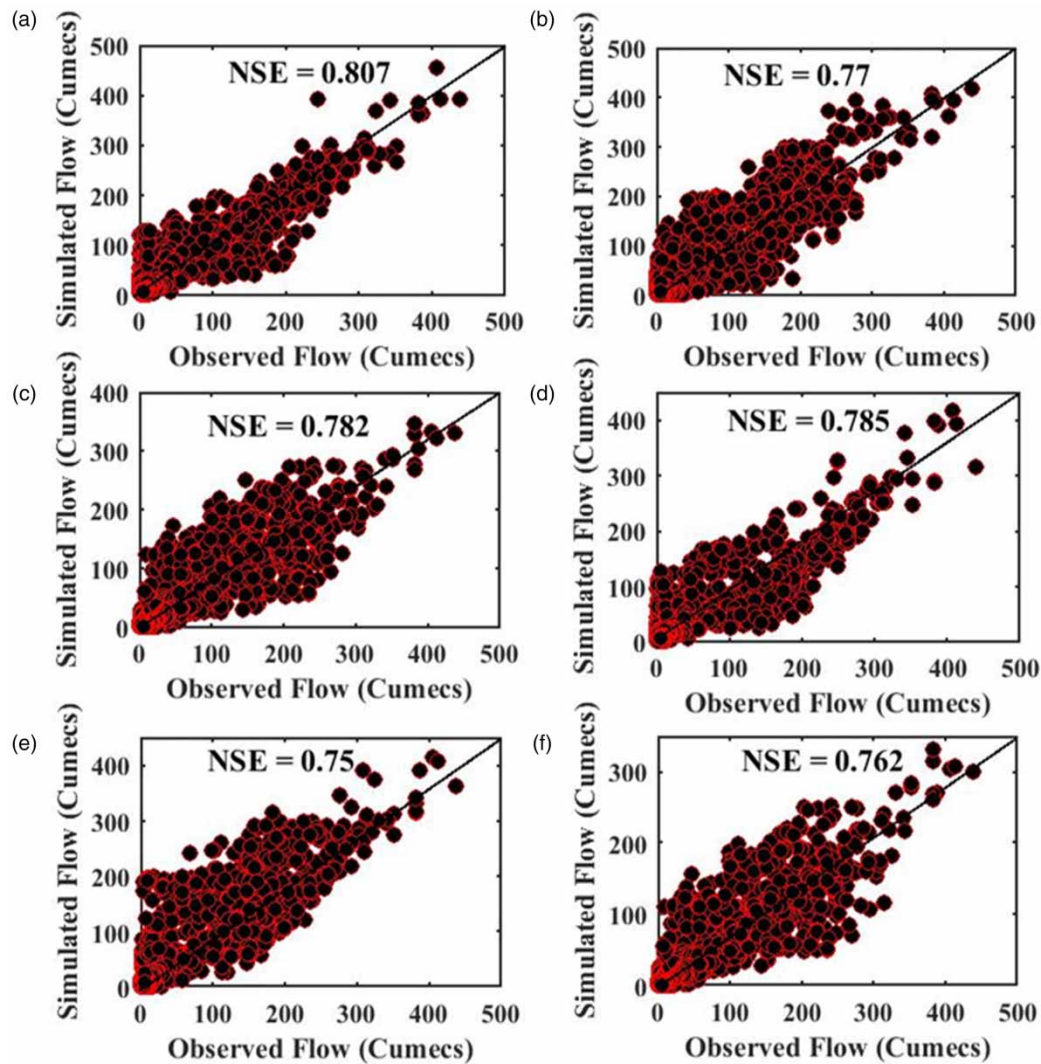


Figure 7 | Scatter plot of the observed and simulated flows: (a) SWAT-gauge, (b) HEC-HMS-gauge, (c) HBV-gauge, (d) SWAT-CMORPH, (e) HEC-HMS-CMORPH, and (f) HBV-CMORPH, at the validation stage.

for validation and calibration, respectively. The watershed average rainfall and temperature were computed using the recorded datasets from the five gauging stations by the Thiessen polygon method and evapotranspiration was calculated by the Hargreaves method using maximum and minimum daily temperatures. The physical characteristics of the watershed were classified into elevation–vegetation classes to better understand the semi-distributed behavior of the watershed (Table 8). NSEs of the model at validation were 0.782 and 0.762 when using gauge and CMORPH datasets, respectively (Table 6). The runoff simulated by HBV by both gauge and satellite-driven datasets at the validation phase showed a good agreement with the measured runoff. Hence, it is worth mentioning that the HBV is well-calibrated and its parameter optimization is physically meaningful and the satellite dataset could be used for hydrological modeling in the catchment. As depicted in Figure 8, the HBV model was capable of simulating low and mean flows, whereas it slightly underestimated peak flows. These indicate the model's soil and runoff generation routine capability and well represent the capability of the catchment's physical processes, and reproduce a reliable rainfall–runoff relationship.

3.3. Results of modeling by input data fusion

In the next step, rainfall driven from gauge and satellite sources was combined to deploy rainfall–runoff modeling to evaluate the combined effects on streamflow simulation (Table 9). Commonly, before using satellite-driven rainfall for hydrological

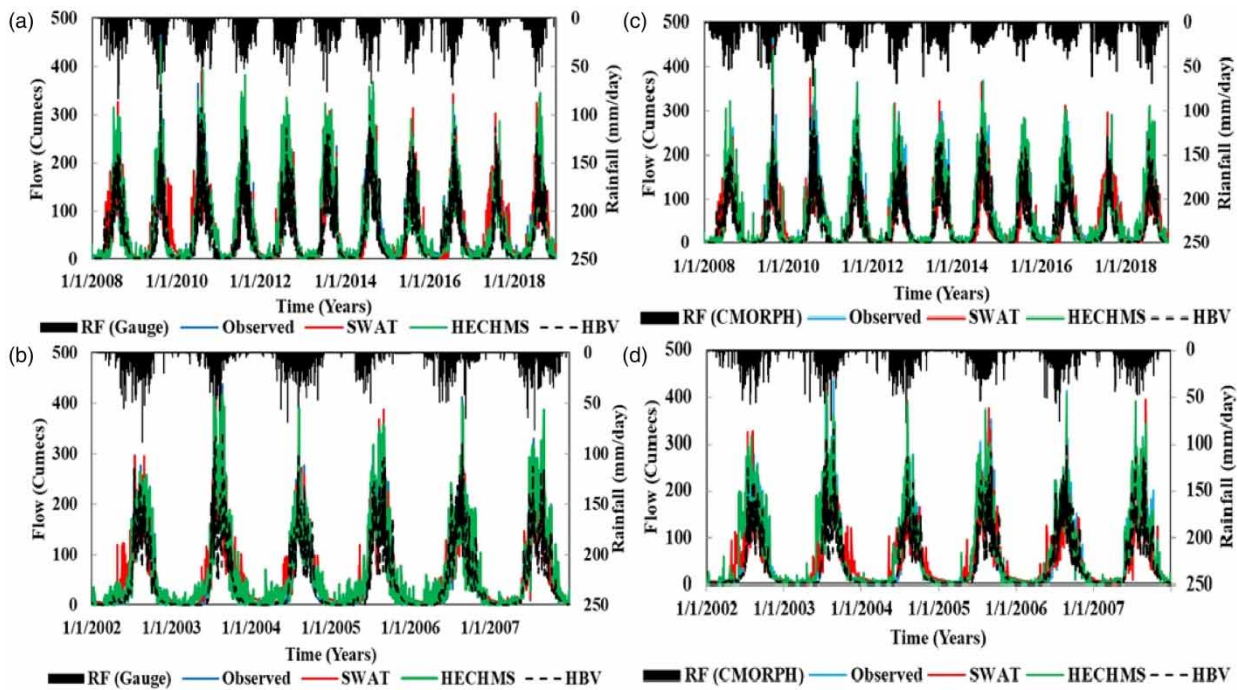


Figure 8 | Simulated runoff by SWAT, HEC-HMS, and HBV: (a) at calibration, (b) at validation phases for gauge dataset, (c) at calibration, and (d) at validation phases for satellite (CMORPH) dataset.

Table 7 | Physiographic features of the watershed

Sub-basin	Area (km ²)	Slope (%)	CN	Lag (hours)
Sub-basin 1 (S1)	561.34	22.5	85.2	4.43
Sub-basin 2 (S2)	386.12	15.18	83.87	3.12
Sub-basin 3 (S3)	233.58	14.45	84.26	2.71
Sub-basin 4 (S4)	188.97	19.51	83.75	2.42
Sub-basin 5 (S5)	268.24	17.33	84.57	1.92

simulation, it should be 'bias-corrected' via a statistical relation with gauge rainfall datasets (Biteu *et al.* 2012). The most common approaches for satellite rainfall datasets' bias correction are based on the following steps. First, the bias on the rainfall satellite datasets is computed as the ratio of daily average satellite rainfall products on the specific grid that covers the gauge station to the corresponding rainfall recorded by the gauge. Second, the satellite rainfall dataset that was originally obtained is multiplied by the bias determined in the first step and the bias is removed. Nevertheless, in the current study, the raw satellite dataset has been imposed into the models alongside gauge datasets, which could perform as a bias correction strategy for satellite datasets. For HBV, the weighted average of the rainfall datasets from both data sources was used as the combined input, since HBV accepts only single climate data. The modeling performance in rainfall data fusion (Table 9) reveals that the simulation efficiency of modeling could be significantly improved when compared with using individual satellite dataset results but it only slightly improved the modeling by gauge dataset results (Table 6). In the particular case of using a satellite dataset, the simulation enhancement of the modeling performance could be the bias rectification in input fusion. The result depicted that the runoff simulation capacity capability of the gauge dataset is superior to that of the satellite dataset. The logic behind this could be the fact that gauge-based rainfall could capture the most accurate and valid hydrological information that can represent physical processes at the basin scale. The quality and accuracy of the satellite rainfall dataset depended on factors such as sky cloud coverage, revisit time of satellites, and corresponding orbital positions. The temporal and spatial variability of the mentioned factors could cause bias in satellite rainfall estimation. Combining rainfall products from both sources boosted the simulation

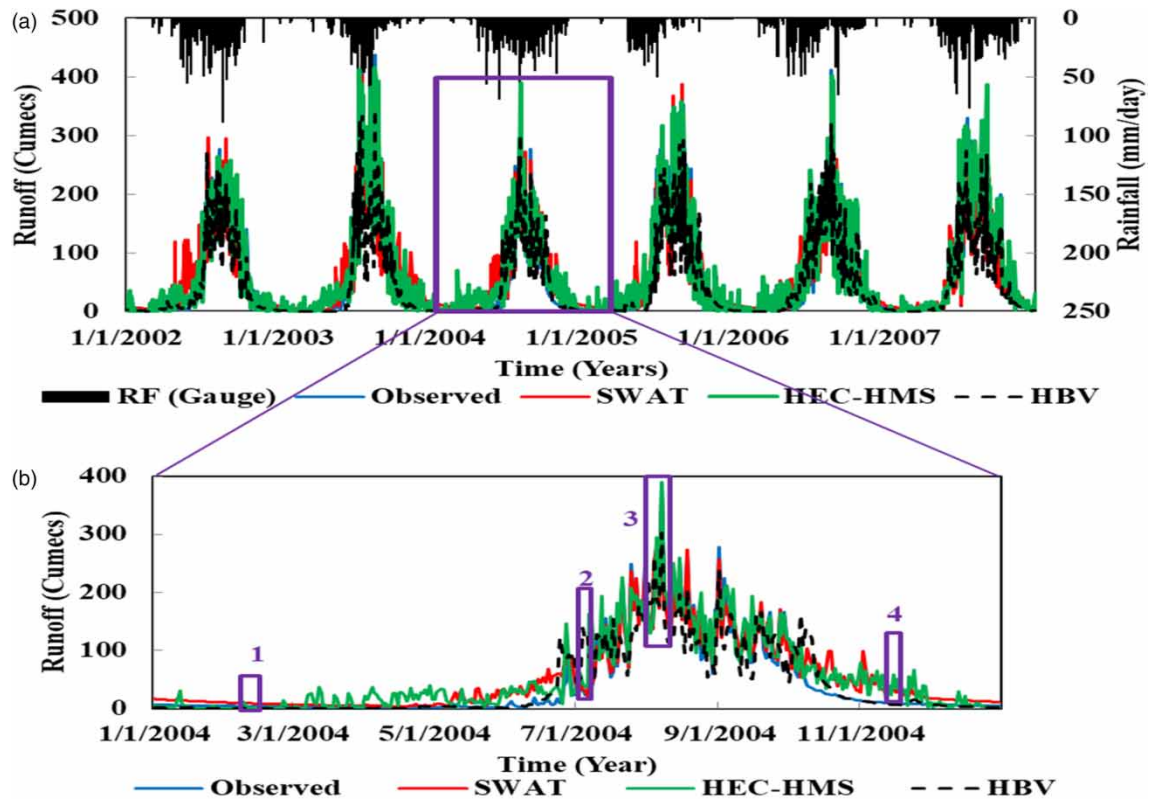


Figure 9 | Simulated runoff using gauge rainfall via SWAT, HEC-HMS, and HBV: (a) at the validation step and (b) details for the year 2004.

Table 8 | The HBV sensitive parameters and their optimized values for each vegetation zone

Parameter	Range	Vegetation class 1	Vegetation class 2	Vegetation class 3
FC	100–1,000	625.7	957.3	566.2
K_2	0.01–0.1	0.04	0.02	0.05
β	1–6	0.91	0.54	0.47
LP	0.3–0.7	0.37	0.41	0.53

Table 9 | Modeling performance for rainfall datasets' fusion

Model	NSE		RMSE	
	Calibration	Validation	Calibration	Validation
SWAT	0.852	0.821	28.436	32.618
HECH-HMS	0.816	0.785	32.672	35.657
HBV	0.824	0.794	31.842	34.675

Note: RMSE is in m^3/s .

performance as it is compared with results that were simulated using rainfall from single sources. Therefore, it is logical to mention that gauge datasets have amended the bias of satellite datasets and improved the rainfall–runoff simulation capacity of the models.

All applied models simulated the catchment runoff with good performance at both calibration and validation stages. However, the modeling result of satellite-based datasets over/underestimated the simulated runoff and showed slightly lower

performance. The reason could be due to the accuracy of satellite sensors in retrieving data and topography fluctuations of the study area. As indicated by Gebremichael *et al.* (2014), three satellite rainfall datasets (CMORPH, 3B42RT, and 3B42) may overestimate daily rainfall in lowland areas and underestimate it in highland areas. Hence, the study area has a high elevation that deviates between 1,866 and 3,543 m above sea level (see Figure 6); therefore, it could be more exposed to topographic influences. The satellite rainfall estimation precision also depends on the algorithms that were used to convert the sensed information into rainfall. The performance of the satellite rainfall dataset that uses microwave algorithms surpasses the one that utilizes infrared waves (Bitew & Gebremichael 2010). As the result depicts, the microwave-based CMORPH outperformed the infrared–microwave combination-based 3B42RT and infrared-based 3B42 rainfall products (see Table 6).

Based on the given performance measures, the runoff simulation of all models led to promising results using both datasets (see Figure 9); nevertheless, all models could not equally perform and capture the existing physical relationship of the watershed. For instance, SWAT surpassed the HEC-HMS and HBV models in both calibration and validation phases for all data sources in the overall result (see Table 6). If the details of simulated runoff for a specific season by each model are accounted for, the strength and weaknesses of each model would be better visualized. Hence, to further evaluate modeling capability in rainy and dry seasons, two different points for each season were randomly selected, and the simulated runoff values by each model were compared with the observed runoff. For the dry season, February 8 and November 16, 2004 (indicated as points 1 and 4 in Figure 9), respectively, were selected. For the rainy season, July 9 and August 7, 2004 (points 2 and 3), respectively, were picked up. The corresponding runoff values of observed and predicted runoff by different models at each selected point are shown in Table 10.

As shown in Table 10, different models showed different performances in predicting runoff at different points in the time series. The details in Table 10 indicate that HBV simulated closer runoff to the observed value in the dry season whereas HEC-HMS overestimated dry-season flow but well-captured peak flows and SWAT is good at simulating average flows. The results at the selected dates indicate that each model at different seasons could infer different outcomes. Hence, the combination of different models' outputs via the ensemble method could improve the simulation capacity of modeling and could lead to more accurate results. To this end, outputs from single models were utilized as inputs, and two linear (SAE and WAE) and one nonlinear (NNE) ensemble modeling were carried out, as discussed in the following section.

3.4. Results of ensemble modeling

Ensemble modeling could enhance the runoff simulation of the single model (SWAT, HEC-HMS, and HBV). The simulation results of single models obtained from the input fusion stage were imposed into the SAE, WAE, and NNE ensemble techniques. The weight for WAE was obtained from NSE at the validation phase according to Equation (11). The NNE ensemble technique was developed via FFNN by BP training with one hidden layer and changing the number of neurons until the optimum epoch was achieved. The ensemble modeling results are presented in Table 11 and according to NSE and RMSE values, the ensemble techniques significantly improved the simulation performance of single models as compared with both steps 1 and 2 modeling results (see Tables 6, 9 and 11). Among the ensemble techniques employed, the superiority of NNE was indicated over the SAE and WAE techniques. The NNE technique improved the performance of low-performing (3B42) satellite dataset-based modeling by 16.6%, 17.5%, and 17% for SWAT, HEC-HMS, and HBV, respectively, at the validation phase. The NNE technique also increased the performance of gauge-satellite fusion modeling of SWAT, HEC-HMS, and HBV by 9.2%, 13.3%, and, 12.2%, respectively, at the validation step. From the findings, it is worth mentioning that the

Table 10 | Observed and predicted runoff values at different points

Selected points	Runoff values (m ³ /s)			
	Observed	SWAT	HEC-HMS	HBV
1	4.03	9.36	12.2	2
2	60.66	55.26	83.2	44
3	387.94	363.67	390.6	302.8
4	5.15	13.9	21.2	3.3

Table 11 | The ensemble modeling results of rainfall data fusion

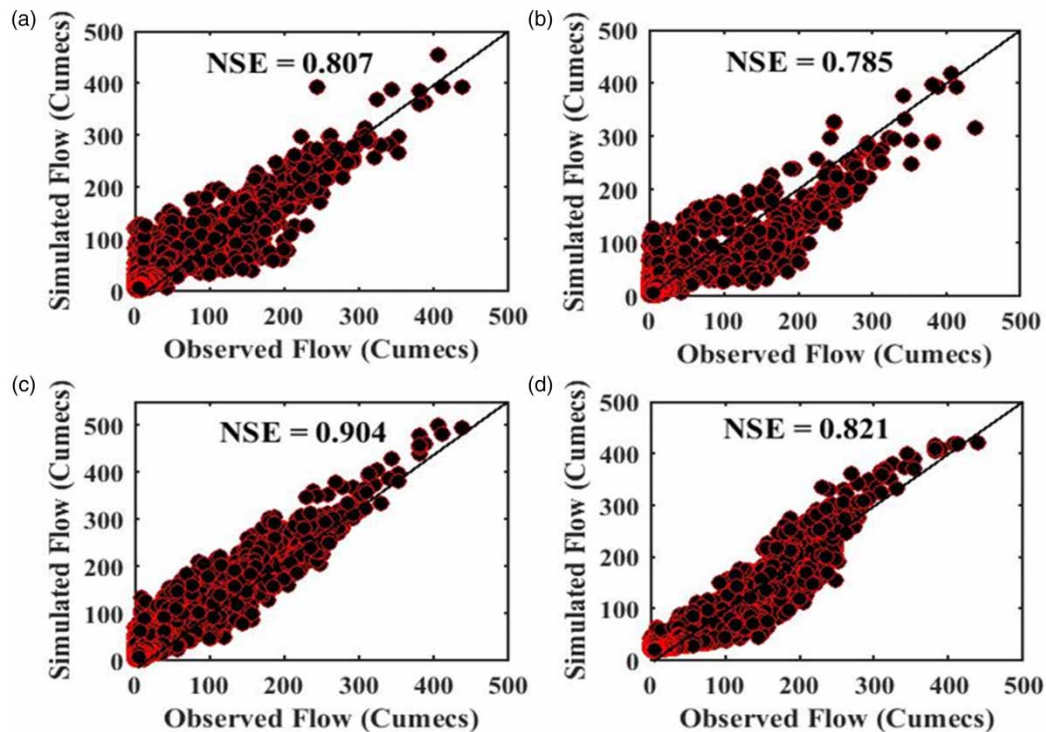
Ensemble methods	NSE		RMSE (m ³ /s)	
	Calibration	Validation	Calibration	Validation
SAE	0.875	0.825	26.473	31.311
WAE	0.891	0.854	24.582	29.163
NNE	0.952	0.904	20.864	25.601

Note: RMSE is m³/s.

ensemble of runoff from different sources could significantly improve the simulation performance of single models that used separate rainfall data from different sources.

The scatter plots of simulated runoff at the validation stage for a single model (SWAT) using both datasets, data fusion of both sources and NNE model versus observed flow, are indicated in Figure 10. As is seen in the scatter plots, NNE significantly improved the performance of the models. The superiority of NNE over SAE and WAE could be the nonlinear models' ability to well understand the nonlinear and complex physical relationship between rainfall and runoff. The linear models have shown less performance than the NNE model for the following reasons. Unlike NNE, the SAE and WAE methods are linear and they could only perform well at exploring the direct relationship between the inputs and output of the models. Hence, the weakness of single models might propagate and combine through linear ensemble models because the models directly amalgamate outputs from the single models.

The performance of SAE, WAE, and NNE are also indicated by the two-dimensional explicit transparency Taylor diagram (see Figure 11) that can visibly display the simulated and measured flow for precise comparisons. In the Taylor diagram, NSE and standard deviation were combined to form a multi-performance measure matrix in a single diagram and it could elaborate on the statistical relationship between the simulated and observed flow. This diagram aims to recapitulate multiple performances in a single arrangement that evaluates how close the simulated flow is to the observed flow. The Taylor diagram for the three ensemble techniques is given in Figure 11. In this method, if the ensemble technique result is closer to the

**Figure 10** | Scatter plots for (a) SWAT-gauge, (b) SWAT-CMORPH, (c) NNE ensemble, (d) SWAT-combination in the validation phase.

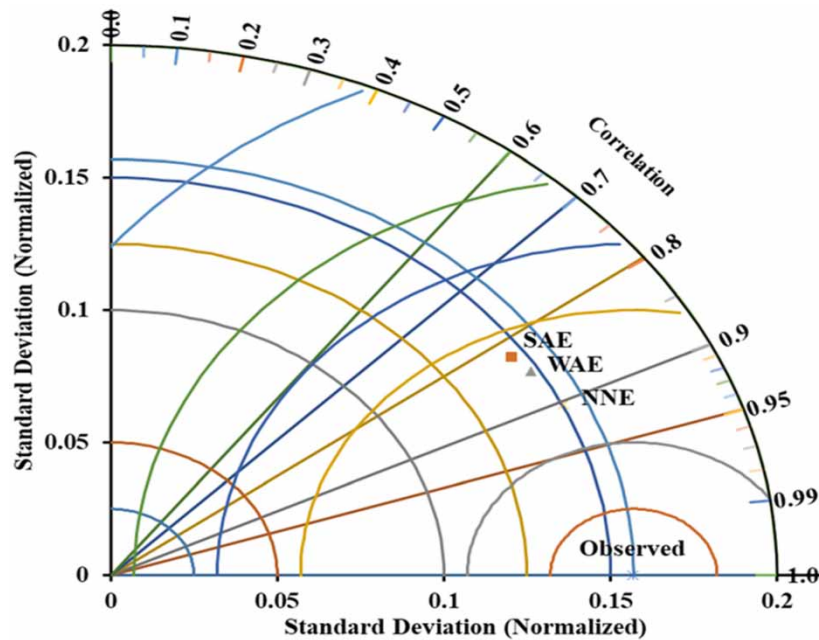


Figure 11 | Taylor diagram showing performance of SAE, WAE, and NNE ensemble techniques.

observed flow, the model is more efficient. As is indicated in Figure 11, NNE is closer to the observed flow than SAE and WAE; hence, it is more accurate than the rest.

4. CONCLUSIONS

The current study was motivated by rainfall–runoff modeling by SWAT, HEC-HMS, and HBV using five station gauges and three satellite-driven rainfall datasets in the Gilgel-Abay catchment situated in the upper Blue Nile basin, Ethiopia. The most sensitive parameters on rainfall–runoff modeling for each model were analyzed and CN2 and ALPHA_BF for SWAT, initial abstraction and lag time for HEC-HMS, and FC and K_2 for HBV were identified. First, rainfall–runoff modeling by each model using each satellite and gauge as well as rainfall dataset fusion was simulated, separately. Second, the runoff simulated by each model using rainfall dataset fusion was imposed to ensemble modeling via SAE, WAE, and NNE techniques to improve modeling accuracy. From the satellite datasets, CMORPH provided better performance for all models; nevertheless, it tends to overestimate the low flow. The result using 3B42 and 3B42RT underestimates the peak flow, and particularly, 3B42 produced arbitrary false pseudo-spikes in the dry season. Among the applied models, SWAT revealed the best performance for all data sources, which could be due to the ability of the model to well distinguish the physical relationship at each HRU since it uses multiple layers of spatial input data. The fusion of rainfall datasets from different sources indicated a substantial improvement over the results using only satellite datasets; it also showed a slight improvement over gauge-based runoff simulations. To further enhance the modeling accuracy, two linear (SAE and WAE) and nonlinear (NNE) ensemble techniques were employed using rainfall data fusion. Among the ensemble techniques, NNE was superior over the others and enhanced the performance of low-performing (3B42) satellite-based modeling by 16.6%, 17.5%, and 17% for SWAT, HEC-HMS, and HBV, respectively, at the validation stage. Similarly, it increased the performance of gauge–satellite combination modeling of SWAT, HEC-HMS, and HBV by 9.2%, 13.3%, and, 12.2%, respectively, at the validation stage. From the employed ensemble methods, NNE was a vigorous and accurate ensemble method for rainfall–runoff simulation as the model could examine the nonlinear relation of the hydrologic process. In general, the result of the current study would be a pioneering step to exploit rainfall datasets combined from multiple satellite sources in data-scarce ungauged and irregularly gauged catchments. This could be a good option to get reliable rainfall datasets in developing countries. Furthermore, future studies should focus on validating satellite rainfall products using the local rain-gauge time series data as the reference data; hence, the relevance and credibility of the satellite datasets would be better verified.

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design. Material preparation, data collection, analysis, and first draft manuscript writing were performed by TG. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analyzed during the current study are not publicly available due to third-party data provision policy of the Ethiopian Meteorological Agency but are available from the corresponding author on reasonable request.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Abba, S. I., Linh, N. T. T., Abdullahi, J., Ali, S. I. A., Pham, Q. B., Abdulkadir, R. A., Costache, R., Nam, V. T. & Anh, D. T. 2020 [Hybrid machine learning ensemble techniques for modeling dissolved oxygen concentration](#). *IEEE Access* **8**, 157218–157237. <https://doi.org/10.1109/ACCESS.2020.3017743>.
- Abbaspour, K. C. 2015 *SWAT-CUP: SWAT Calibration and Uncertainty Programs – A User Manual*. Eawag (Swiss Federal Institute of Aquatic Science and Technology), Dübendorf, Switzerland.
- Araghinejad, S., Burn, D. H. & Karamouz, M. 2006 [Long-lead probabilistic forecasting of streamflow using ocean-atmospheric and hydrological predictors](#). *Water Resources Research* **42** (3), W03431. <https://doi.org/10.1029/2004WR003853>.
- Arnold, J. G. & Fohrer, N. 2005 [SWAT2000: current capabilities and research opportunities in applied watershed modelling](#). *Hydrological Processes* **19** (3), 563–572.
- Arnold, J. G., Srinivasan, R., Mutiah, R. S. & Williams, J. R. 1998 [Large area hydrologic modeling and assessment part I: model development](#). *Journal of the American Water Resources Association* **34**, 73–89.
- Bitew, M. M. & Gebremichael, M. 2010 [Assessment of high-resolution satellite rainfall for streamflow simulation in medium watersheds of the East African highlands](#). *Hydrology and Earth System Sciences Discussions* **7**, 8213–8232. <https://doi.org/10.5194/hessd-7-8213-2010>.
- Bitew, M. M. & Gebremichael, M. 2011 [Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model](#). *Water Resources Research* **47** (6), W06526. <https://doi.org/10.1029/2010WR009917>.
- Bitew, M. M., Gebremichael, M., Ghebremichael, L. T. & Bayissa, Y. A. 2012 [Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia](#). *Journal of Hydrometeorology* **13** (1), 338–350. <https://doi.org/10.1175/2011JHM1292.1>.
- Busico, G., Colombani, N., Fronzi, D., Pellegrini, M., Tazioli, A. & Mastrocicco, M. 2020 [Evaluating SWAT model performance, considering different soils data input, to quantify actual and future runoff susceptibility in a highly urbanized basin](#). *Journal of Environmental Management* **266**, 110625. <https://doi.org/10.1016/j.jenvman.2020.110625>.
- Chen, Y., Xu, C. Y., Chen, X., Xu, Y., Yin, Y., Gao, L. & Liu, M. 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](#). *Journal of Hydrology* **573**, 486–500. <https://doi.org/10.1016/j.jhydrol.2019.03.091>.
- Ciupak, M., Ozga-Zielinski, B., Adamowski, J., Deo, R. C. & Kochanek, K. 2019 [Correcting satellite precipitation data and assimilating satellite-derived soil moisture data to generate ensemble hydrological forecasts within the HBV rainfall-runoff model](#). *Water* **11** (10), 2138. <https://doi.org/10.3390/w11102138>.
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J. & Ropelewski, C. F. 2007 [Validation of satellite rainfall products over East Africa's complex topography](#). *International Journal of Remote Sensing* **28** (7), 1503–1526. <https://doi.org/10.1080/01431160600954688>.
- Driessen, T. L. A., Hurkmans, R. T. W. L., Terink, W., Hazenberg, P., Torfs, P. J. J. F. & Uijlenhoet, R. 2010 [The hydrological response of the Ourthe catchment to climate change as modelled by the HBV model](#). *Hydrology and Earth System Sciences* **14** (4), 651–665. <https://doi.org/10.5194/hess-14-651-2010>.
- Feldman, A. D. 2000 *Hydrologic Modeling System HEC-HMS: Technical Reference Manual*. Hydrologic Engineering Center, US Army Corps of Engineers, Davis, CA, USA.
- Fowler, K. J. A., Peel, M. C., Western, A. W., Zhang, L. & Peterson, T. J. 2016 [Simulating runoff under changing climatic conditions: revisiting an apparent deficiency of conceptual rainfall-runoff models](#). *Water Resources Research* **52** (3), 1820–1846. <https://doi.org/10.1002/2015WR018068>.
- Gassman, P. W., Reyes, M. R., Green, C. H. & Arnold, J. G. 2007 [The Soil and Water Assessment Tool: historical development, applications, and future research directions](#). *Transactions of the ASABE* **50**, 1211–1250.
- Gebre, S. L. 2015 [Application of the HEC-HMS model for runoff simulation of Upper Blue Nile River Basin](#). *Hydrology: Current Research* **6** (2), 1000199. <https://doi.org/10.4172/2157-7587.1000199>.

- Gebremichael, M., Bitew, M. M., Hirpa, F. A. & Tesfay, G. N. 2014 Accuracy of satellite rainfall estimates in the Blue Nile Basin: lowland plain versus highland mountain. *Water Resources Research* **50**, 8775–8790. <https://doi.org/10.1002/2013WR014500>.
- Gelete, G., Nourani, V., Gokcekus, H. & Gichamo, T. 2023 Ensemble physically based semi-distributed models for the rainfall-runoff process modeling in the data-scarce Katar catchment, Ethiopia. *Journal of Hydroinformatics* **25** (2), 567–592. <https://doi.org/10.2166/hydro.2023.197>.
- Glavan, M. & Pintar, M., 2012 Strengths, weaknesses, opportunities and threats of catchment modelling with Soil and Water Assessment Tool (SWAT) model. In: *Water Resources Management and Modeling* (Nayak, P., ed.), InTech, Rijeka, Croatia, pp. 39–64.
- Guzha, A. C., Rufino, M. C., Okoth, S., Jacobs, S. & Nóbrega, R. L. B. 2018 Impacts of land use and land cover change on surface runoff, discharge and low flows: evidence from East Africa. *Journal of Hydrology: Regional Studies* **15**, 49–67.
- Hamby, D. M. 1994 A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment* **32** (2), 135–154.
- Huang, S., Eisner, S., Magnusson, J. O., Lussana, C., Yang, X. & Beldring, S. 2019 Improvements of the spatially distributed hydrological modelling using the HBV model at 1 km resolution for Norway. *Journal of Hydrology* **577**, 123585. <https://doi.org/10.1016/j.jhydrol.2019.03.051>.
- Jaiswal, R. K., Ali, S. & Bharti, B. 2020 Comparative evaluation of conceptual and physical rainfall-runoff models. *Applied Water Science* **10** (1), 48. <https://doi.org/10.1007/s13201-019-1122-6>.
- Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M. & Brumby, S. P. 2021 Global land use/land cover with Sentinel-2 and deep learning. In: *IGARSS 2021-2021 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, Piscataway, NJ, USA, pp. 4704–4707.
- Khelifa, W. B., Hermassi, T., Strohmeier, S., Zucca, C., Ziadat, F., Boufaroua, M. & Habaieb, H. 2017 Parameterization of the effect of bench terraces on runoff and sediment yield by SWAT modeling in a small semi-arid watershed in northern Tunisia. *Land Degradation and Development* **28** (5), 1568–1578. <https://doi.org/10.1002/ldr.2685>.
- Kiran, R. N. & Ravi, V. 2008 Software reliability prediction by soft computing techniques. *Journal of Systems and Software* **81** (4), 576–583. <https://doi.org/10.1016/j.jss.2007.05.005>.
- Kumar, A., Singh, R., Jena, P. P., Chatterjee, C. & Mishra, A. 2015 Identification of the best multi-model combination for simulating river discharge. *Journal of Hydrology* **525**, 313–325. <https://doi.org/10.1016/j.jhydrol.2015.03.060>.
- Kwin, C. T., Talei, A., Alaghmand, S. & Chua, L. H. C. 2016 Rainfall-runoff modeling using dynamic evolving neural fuzzy inference system with online learning. *Procedia Engineering* **154**, 1103–1109. <https://doi.org/10.1016/j.proeng.2016.07.518>.
- Le, M.-H., Lakshmi, V., Bolten, J. & Bui, D. D. 2020 Adequacy of satellite-derived precipitation estimate for hydrological modeling in Vietnam basins. *Journal of Hydrology* **586**, 124820. <https://doi.org/10.1016/j.jhydrol.2020.124820>.
- Leong, M. & Yang, X. 2020 Effect of rainfall station density, distribution and missing values on SWAT outputs in tropical region. *Journal of Hydrology* **584**, 124660. <https://doi.org/10.1016/j.jhydrol.2020.124660>.
- Li, D., Christakos, G., Ding, X. & Wu, J. 2018 Adequacy of TRMM satellite rainfall data in driving the SWAT modeling of Tiaoxi catchment (Taihu lake basin, China). *Journal of Hydrology* **556**, 1139–1152. <https://doi.org/10.1016/j.jhydrol.2017.01.006>.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M. & Bergström, S. 1997 Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* **201**, 272–288.
- Mandal, A., Stephenson, T. S., Brown, A. A., Campbell, J. D., Taylor, M. A. & Lumsden, T. L. 2016 Rainfall-runoff simulations using the CARIWIG Simple Model for Advection of Storms and Hurricanes and HEC-HMS: implications of Hurricane Ivan over the Jamaica Hope River watershed. *Natural Hazards* **83** (3), 1635–1659. <https://doi.org/10.1007/s11069-016-2380-3>.
- Mondal, A., Lakshmi, V. & Hashemi, H. 2018 Intercomparison of trend analysis of multisatellite monthly precipitation products and gauge measurements for river basins of India. *Journal of Hydrology* **565**, 779–790. <https://doi.org/10.1016/j.jhydrol.2018.08.083>.
- Nourani, V., Gokcekus, H. & Gelete, G. 2021a Estimation of suspended sediment load using artificial intelligence-based ensemble model. *Complexity* **2021**, 6633760. <https://doi.org/10.1155/2021/6633760>.
- Nourani, V., Gökçekuş, H. & Gichamo, T. 2021b Ensemble data-driven rainfall-runoff modeling using multi-source satellite and gauge rainfall data input fusion. *Earth Science Informatics* **14** (4), 1787–1808. <https://doi.org/10.1007/s12145-021-00615-4>.
- Orellana, B., Pechlivanidis, I. G., McIntyre, N., Wheeler, H. S. & Wagener, T. 2008 A toolbox for the identification of parsimonious semi-distributed rainfall-runoff models: application to the Upper Lee catchment. In: *Proceedings of the iEMSs Fourth Biennial Meeting: International Congress on Environmental Modelling and Software (iEMSs 2008)* (Sánchez-Marrè, M., Béjar, J., Comas, J., Rizzoli, A. E. & Guariso, G., eds), International Environmental Modelling and Software Society, Lugano-Viganello, Switzerland, pp. 670–677.
- Pervin, L., Gan, T. Y., Scheepers, H. & Islam, M. S. 2021 Application of the HBV model for the future projections of water levels using dynamically downscaled global climate model data. *Journal of Water and Climate Change* **12** (6), 2364–2377. <https://doi.org/10.2166/wcc.2021.302>.
- Setegn, S. G., Srinivasan, R. & Dargahi, B. 2008 Hydrological modelling in the Lake Tana Basin, Ethiopia using SWAT model. *The Open Hydrology Journal* **2**, 49–62. <https://doi.org/10.2174/1874378100802010049>.
- Sharafati, A., Asadollah, S. B. H. S. & Neshat, A. 2020 A new artificial intelligence strategy for predicting the groundwater level over the Rafsanjan aquifer in Iran. *Journal of Hydrology* **591**, 125468. <https://doi.org/10.1016/j.jhydrol.2020.125468>.
- Shiru, M. S. & Park, I. 2020 Comparison of ensembles projections of rainfall from four bias correction methods over Nigeria. *Water* **12** (11), 3044. <https://doi.org/10.3390/w12113044>.

- Singh, V., Bankar, N., Salunkhe, S. S., Bera, A. K. & Sharma, J. R. 2013 Hydrological stream flow modelling on Tungabhadra catchment: parameterization and uncertainty analysis using SWAT CUP. *Current Science* **104** (9), 1187–1199.
- Solomatine, D. P. & Shrestha, D. L. 2009 A novel method to estimate model uncertainty using machine learning techniques. *Water Resources Research* **45** (12), W00B11.
- Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M. R. & Molnar, P. 2017 Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences* **21** (3), 1359–1380. <https://doi.org/10.5194/hess-21-1359-2017>.
- Tibangayuka, N., Mulungu, D. M. M. & Izdori, F. 2022 Evaluating the performance of HBV, HEC-HMS and ANN models in simulating streamflow for a data scarce high-humid tropical catchment in Tanzania. *Hydrological Sciences Journal* **67** (14), 2191–2204. <https://doi.org/10.1080/02626667.2022.2137417>.
- Uhlenbrook, S., Mohamed, Y. & Gagne, A. S. 2010 Analyzing catchment behavior through catchment modeling in the Gilgel Abay, Upper Blue Nile River Basin, Ethiopia. *Hydrology and Earth System Sciences* **14** (10), 2153–2165. <https://doi.org/10.5194/hess-14-2153-2010>.
- Velázquez, J. A., Anctil, F., Ramos, M. H. & Perrin, C. 2011 Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. *Advances in Geosciences* **29**, 33–42. <https://doi.org/10.5194/adgeo-29-33-2011>.
- Wale, A., Rientjes, T. H. M., Gieske, A. S. M. & Getachew, H. A. 2009 Ungauged catchment contributions to Lake Tana's water balance. *Hydrological Processes* **23** (26), 3682–3693.
- Whittemore, R. C. 2002 'Validation of the SWAT model on a large river basin with point and nonpoint sources,' by C. Santhi, J. G. Arnold, J. R. Williams, W. A. Dugas, R. Srinivasan, and L. M. Hauck. *Journal of the American Water Resources Association* **38** (6), 1767–1768. <https://doi.org/10.1111/j.1752-1688.2002.tb04380.x>.
- Xin, Z., Shi, K., Wu, C., Wang, L. & Ye, L. 2019 Applicability of hydrological models for flash flood simulation in small catchments of hilly area in China. *Open Geosciences* **11** (1), 1168–1181. <https://doi.org/10.1515/geo-2019-0089>.
- Yong, B., Liu, D., Gourley, J. J., Tian, Y., Huffman, G. J., Ren, L. & Hong, Y. 2015 Global view of real-time TRMM multisatellite precipitation analysis: implications for its successor Global Precipitation Measurement mission. *Bulletin of the American Meteorological Society* **96**, 283–296.
- Young, C. C., Liu, W. C. & Wu, M. C. 2017 A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. *Applied Soft Computing Journal* **53**, 205–216. <https://doi.org/10.1016/j.asoc.2016.12.052>.
- Zhang, G. P. & Savenije, H. H. G. 2005 Rainfall-runoff modelling in a catchment with a complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach. *Hydrology and Earth System Sciences* **9** (3), 243–261. <https://doi.org/10.5194/hess-9-243-2005>.

First received 20 February 2023; accepted in revised form 24 November 2023. Available online 20 December 2023