




An Ensemble Hydrologic Modeling System for Runoff and Evapotranspiration Evaluation over an Agricultural Watershed

Sushil Kumar Himanshu¹  · Ashish Pandey² · Mallappa J. Madolli³ · Santosh S. Palmate⁴ · Ashish Kumar⁵ · Nitesh Patidar⁶ · Basant Yadav²

Received: 18 May 2022 / Accepted: 20 November 2022 / Published online: 17 December 2022
© Indian Society of Remote Sensing 2022

Abstract

Modeling framework for simulation at a finer scale is important for long-term water resources planning for management. It has always been a challenge to select the appropriate model to simulate the hydrology of a watershed/river basin at a finer spatial resolution. Comparative evaluation of models based on field observations could help researchers to select the suitable model for their purpose. However, a single hydrologic model generally leads to simulation uncertainties due to poor input data, model structure, and model output uncertainty in large-scale exercises. The ensemble model approach could be a better decision-making tool to overcome uncertainty in modeling hydrological processes. In the present study, a widely used macroscale hydrologic model, the three-layer Variable Infiltration Capacity (VIC-3L), was employed to simulate runoff and evapotranspiration (ET) at $3' \times 3'$ grids (~ 5.5 km) resolution over an agriculture-based Marol watershed (5092 km²) of India. The VIC-simulated results were compared and assessed with the results obtained from the Hydrologic Response Unit (HRU)-based Soil and Water Assessment Tool (SWAT) hydrologic model. Further, the ensemble of VIC and SWAT outputs (EnSwaVi; averages of individual model-simulated datasets with equal weights) was also assessed. Simulated runoff and ET were evaluated using observed discharge data at the outlet of the watershed and the actual ET product (MOD16A2) of Moderate Resolution Imaging Spectroradiometer (MODIS), respectively. The simulated discharge values generated by the two models were closely matched with the observed flow. Conversely, ET simulated by VIC was found to be more precise as compared to SWAT. A minimal difference between two model results can be due to the difference in the model structure and runoff simulation method. In general, the ensembles of VIC and SWAT outputs (EnSwaVi) were found better than the individual model outputs. The ensemble modeling approach could provide more reliable assessments of hydrological processes for the planning and management of water resources.

Keywords VIC · SWAT · Ensemble model · Water Balance · Evaluation · MODIS

✉ Sushil Kumar Himanshu
sushil-kumar@ait.ac.th; sushilkumarhimanshu@gmail.com

¹ Department of Food, Agriculture and Bioresources, Asian Institute of Technology, Khlong Luang, Pathum Thani, Thailand

² Department of Water Resources Development and Management, Indian Institute of Technology Roorkee, Roorkee, India

³ Department of Agricultural Engineering, College of Agriculture, Chamarajanagara, Karnataka, India

⁴ Texas A&M AgriLife Research, Texas A&M University System, El Paso, TX, USA

⁵ Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India

⁶ National Institute of Hydrology Roorkee, Roorkee, India

Introduction

Water resources are the prime contributor to a developing economy, environmental protection, and sustainable development (Madolli et al., 2022). It helps in advancing economic growth if managed and planned properly (Dhami et al., 2018). Shortage and misuse of freshwater cause a serious and growing threat to the protection, management, and sustainable development of water resources. Unless the water and land resources are managed accurately, industrial expansion, the natural ecosystems on which they depend, human health, social well-being, and sustainable food production are all in danger (ICWE, 1992). Only a small fraction (about 2.53%) of the estimated total volume of water available on the earth is freshwater (Water facts,

2020). A considerable portion of this freshwater is not available for use, as they lie in inaccessible deep aquifers or frozen in polar regions. This causes a challenge to protect, manage and develop water resources in a sustainable manner considering the economic growth, climate change, and population increase (Amrit et al., 2019; Fan et al., 2022; Kumar et al., 2021a, 2021b; Shiklomanov, 1998; Swain et al., 2021). Hydrologic models have been widely used to assess and manage the sustainability of water resources (Paul et al., 2021).

Hydrologic modeling is an efficient way for consistent long-term behavioral studies of hydrologic and climatic variables (Tanmoyee et al., 2015). Initially, hydrologic models were focused on the development of theories, concepts, and models for a particular component of the hydrologic cycle, such as baseflow (Barnes, 1940), overland flow (Horton, 1939; Keulegan, 1944), channel flow (Manning, 1891), subsurface flow (Fair & Hatch, 1933; Jacob, 1943, 1944; Theis, 1935), depression storage (Horton, 1919; SCS-CN method, 1956), evapotranspiration (Cummings, 1935; Penman, 1948; Thornthwaite, 1948), infiltration (Green & Ampt, 1911) and interception (Horton, 1919). The first physical model capable of modeling the entire watershed with all hydrologic cycle components was most likely the Stanford Watershed Model (SWM), developed in 1966 (Crawford & Linsley, 1966). Further, many hydrologic models were developed to advance computational abilities and algorithms with recently available databases like space technology, remote sensing satellite data, high-resolution digital elevation models (DEMs), and radar rainfall (Pandey et al., 2016). There is vast variability in the capabilities and characteristics of these hydrologic models, such as a representation of processes, accountability of spatial–temporal scale, algorithms used, input requirements, and types of output they provide (Pandey et al., 2016; Paul et al., 2021).

A complex hydrological system has always been investigated by employing physically-based models and simulating the major components like streamflow and sediment yield (Himanshu et al., 2017, 2018a). Many literature studies have proven the robustness of the SWAT (Aadhar et al., 2019; Dhami et al., 2018; Gupta et al., 2020; Murty et al., 2014; Pandey & Palmate, 2019; Swain et al., 2022) and VIC (Narendra et al., 2017; Oubeidillah et al., 2014; Srivastava et al., 2017) hydrologic models in the evaluation of the water balance components. Kang and Sridhar (2018) found SWAT and VIC models reliable for short-term drought forecasting in the contiguous USA. Alvarenga et al. (2020) compared VIC and SWAT hydrologic models in their capabilities to simulate runoff in the Verde River Watershed, Brazil. They found both models suitable for streamflow simulation and suggested that the

integration of SWAT and VIC models can be useful in different water resource assessment studies.

With the development of advanced models and the availability of spatial–temporal data, modelers and stakeholders are now broadly depending on the information derived from hydrological models to make more sustainable choices. However, with the growing family of hydrological models and tools, it has become difficult for decision-makers to identify a plausible model for their intended application (Jajarmizadeh et al., 2012; Pandey et al., 2016). There are several queries related to the model's fitness for the intended application, model reliability, and uncertainties associated with the results. Moreover, because each model has a different modeling concept, algorithms, and input requirement, each would perform differently, and their performance could be non-unique in space and time. Further, to make an appropriate choice among various models, it is important to evaluate models with the available quantity and quality of the input data in the catchment.

Therefore, a comparative evaluation of the commonly used hydrological models (SWAT and VIC) was performed. Although both models perform catchment water balance, there are several characteristic differences between these models. The Soil and Water Assessment Tool (SWAT) was primarily developed by USDA's Agricultural Research Service (ARS) to assess the impacts of land management practices on water quantity, water quality, and sediment fluxes in a watershed (Arnold et al., 1998; Borah & Bera, 2003; Himanshu et al., 2019; Miller et al., 2007; Palmate & Pandey, 2021). It is a physically-based, continuous-time, long-term, semi-distributed watershed-scale hydrologic model (Arnold & Fohrer, 2005; Arnold et al., 1998; Garg et al., 2012; Pandey et al., 2016). On the other hand, the Variable Infiltration Capacity (VIC) is a physically-based semi-distributed macroscale model developed on the Land Surface Modelling scheme, primarily to link with the climate models (Liang et al., 1994). The VIC model explicates the sub-grid level spatial heterogeneity, vegetation phenological changes, soil textures, and terrain characteristics at different spatial resolutions (Kimball et al., 1997). The model can simulate several hydrologic and climatic variables such as snow depth, snowmelt, ET, surface runoff, soil moisture, frozen soil, and streamflow (Tanmoyee et al., 2015). Unlike SWAT, the VIC model can simulate energy balance in addition to water balance at a sub-daily time step (Liang et al., 1994).

Various hydrological models have their strengths and weakness in representing hydrological processes (Li et al., 2018). Due to insufficient input data, model structure, and model output uncertainty in large-scale exercises, relying on a single hydrological model generally leads to

simulation uncertainties (Dietrich et al., 2009; Kauffeldt et al., 2016; Li et al., 2018; Liu & Gupta, 2007; Palmate et al., 2021). To overcome uncertainty in modeling hydrological processes, several techniques have been used in the recent past (Liu et al., 2017; Kasiviswanathan and Sudheer, 2017; Gaur et al., 2022). Among these, the ensemble modeling technique has been gaining popularity in recent years in different sectors of water resources modeling (Doblas-Reyes et al., 2005; Gaur et al. 2021a, 2021b; Horan et al., 2021; Kumar & Nandagiri, 2015; Kumar et al., 2015; Li et al., 2018; Muhammad et al., 2018; Paul et al., 2021; Yadav et al., 2020). Multi-model ensembles, however, outperform individual models and tend to perform better than single-model ensembles in weather prediction and streamflow simulation (Gaur et al. 2021a, 2021b; Kumar et al., 2015; Mendoza et al., 2014; Paul et al., 2021). Different modeling approaches are used to simulate hydrological variables in the ensemble modeling technique. Out of which, the mean ensemble modeling approach that averages the individual hydrological model-simulated datasets with equal weights has been used in many studies to simulate hydrological variables for reducing errors with optimal bias (Doblas-Reyes et al., 2005; Baker & Ellison, 2008; Kumar & Nandagiri, 2015; Muhammad et al., 2018; Li et al., 2018; Horan et al., 2021).

SWAT and VIC models have their own strengths and weakness in representing hydrological processes. Many previous studies on SWAT and VIC comparison have shown that the results of these models may vary, and depending on the accuracy of inputs and parameters, these models overestimate, underestimate or contradict each other in assessing the hydrometeorological variables such as discharge and evapotranspiration. Seasonal simulation performance of the SWAT and VIC-3L models studied by Hu et al. (2007) showed underestimated runoff values than the observed values for spring and winter. The SWAT-simulated runoff values in summer were higher than the VIC-3L simulation but were smaller in winter. These hydrological models also provided useful insights into the impact of climate and anthropogenic activities on regional water security (Veetil et al., 2022). Kang et al. (2022) assessed the impacts of climate change on conventional and flash drought conditions using these models. The SWAT-driven drought indices showed an overall increase in drought occurrence; however, the VIC-driven drought indices showed a decrease in drought occurrence. Dash et al. (2021) revealed that the SWAT simulation-based standardized evapotranspiration drought index (SEDI) was consistent; in contrast, the VIC-3L simulation-based SEDI was continuously overestimated and underestimated with the benchmark satellite (MOD16A2-ET)-derived estimates. A study employing these models showed a remarkable decrease in drought predictions 36% for SWAT

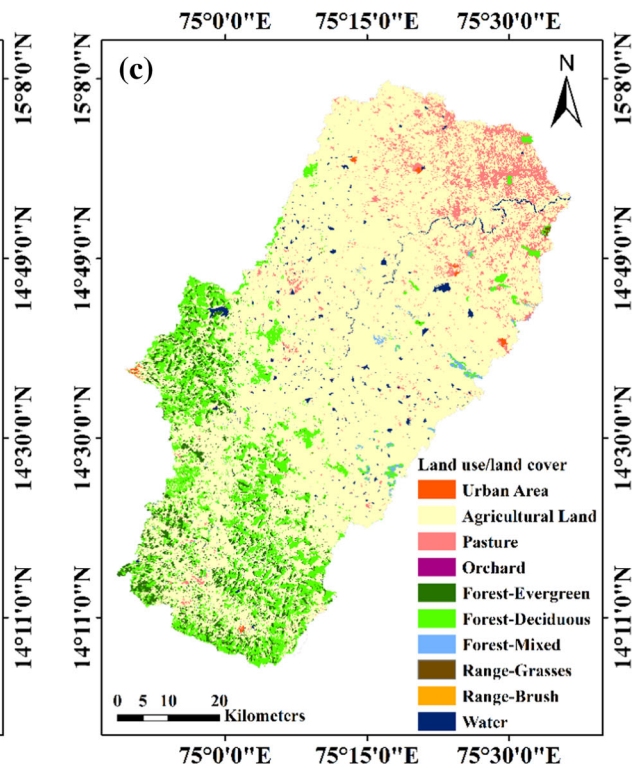
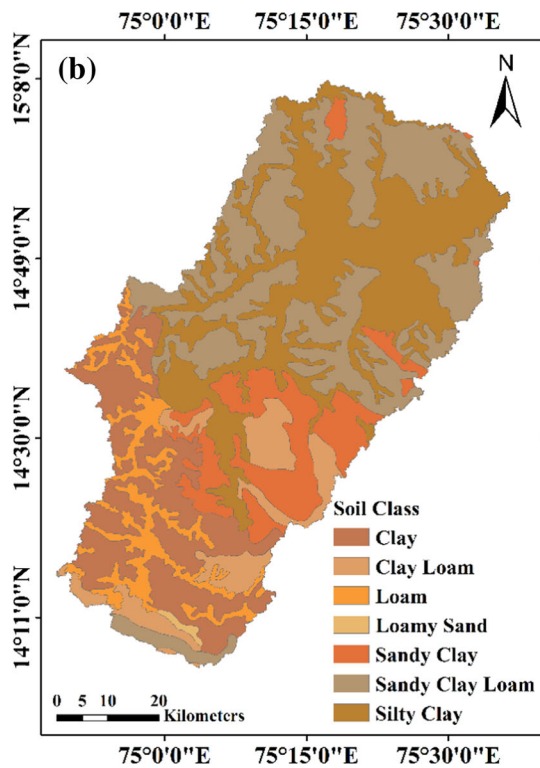
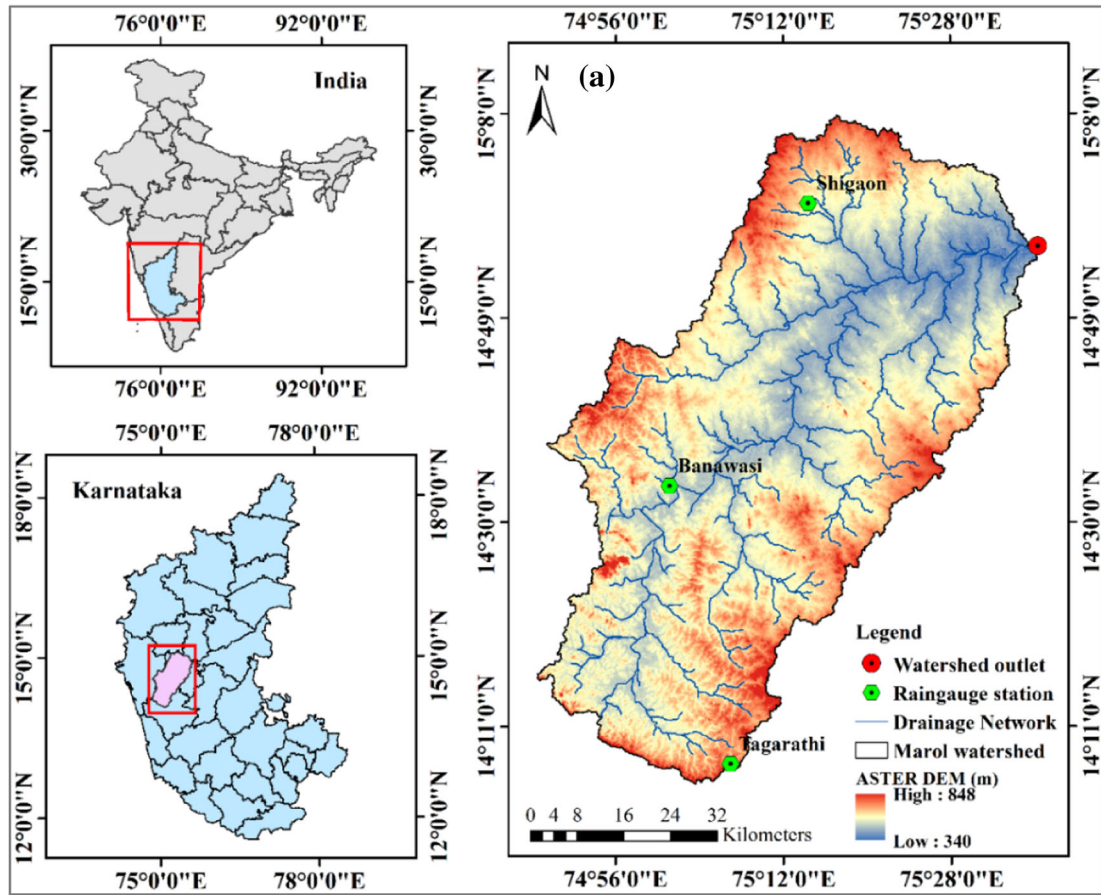
and 38% for VIC, due to uncertainties associated with the meteorological variables (Kang & Sridhar, 2018). An ensemble modeling approach leads to offset uncertainty in input data as well as poor reservoir operation functionality, if any, within the models (Horan et al., 2021).

In a large watershed, the application of a single model can lead to simulation uncertainties if detailed data are not available. Ensemble modeling combines multiple model predictions to create a single prediction that generally tends to perform better than the individual model. Ensemble modeling can be utilized to better simulate components of the hydrologic cycle, and provide a range of possible outcomes and uncertainty. Looking at the above mentioned, this study explores the applicability of an ensemble modeling approach for hydrological variables (runoff and evapotranspiration) over the study area. The study could help in developing a finer spatial resolution modeling framework to simulate the hydrology of a watershed that can contribute to policy and decision-making processes for sustainable water resource management. The SWAT and VIC models were selected to investigate the predictive capability of individual models and the performance of the mean ensemble of these two models (EnSwaVi) to improve the accuracy of the simulation of runoff and evapotranspiration in the study area. To evaluate the models for their ability to address spatial heterogeneity in soil, land cover, and topography in the semiarid region, a heterogeneous watershed named Marol watershed (5092 km²), India, was identified for the study.

Materials and Methods

Study Area

The Marol watershed is part of the upper Krishna River basin, which covers a geographical area of about 5092 km² between longitude from 74°48'30" E to 75°36'38" E and latitude from 14°05'18" N to 15°07'48" N. The Krishna River is an important eastside flowing river in the peninsular region of India (Himanshu et al., 2018b). The study area is positioned along a sub-tributary Varada River of Tungabhadra River in the State of Karnataka, India (Fig. 1). The elevation of the watershed above the mean sea level varies from 340 to 848 m. An average slope varies from 0 to 8.9%, which majorly consists of a gently undulating plain area. However, because of some western hilly areas, the maximum slope of the study area goes up to 31%. Topographic elevation, land use/land cover, and soil textures of the watershed are given in Fig. 1. The average annual rainfall of the watershed is 1330 mm, with a variation in temperature between 16 and 38 °C. The availability of observed hydrometeorological data,



◀ **Fig. 1** Details of the Marol watershed: **a** location map, **b** land use/cover map, and **c** soil map

heterogeneous land use, and absence of any large storage structure makes the watershed appropriate for the present case study.

Data

Different types of datasets, including meteorological data, hydrological data, and thematic data, summarized in Table 1 were used in this study. Daily IMD gridded precipitation data available at a spatial resolution of $0.25^\circ \times 0.25^\circ$ grid (Pai et al., 2014, 2015) were used as inputs to the model. The Marol watershed covers fifteen precipitation grid points. The daily IMD gridded temperature data available at $1^\circ \times 1^\circ$ were also used in the present study (Srivastava et al., 2009). In addition to this, other important data variables, namely relative humidity, solar radiation, and wind speed, not available for the study area, were obtained from the Global Weather Database for SWAT (Dile & Srinivasan, 2014) website at $0.25^\circ \times 0.25^\circ$ spatial resolution.

Daily hydrologic data, i.e., streamflow, measured at the Marol gauge and discharge (G&D) site, were obtained for the years 2000–2010 from the India Water Resources Information System (WRIS) WebGIS portal, Government of India. The Marol G&D site of the Varada River is located at the longitude of $75^\circ 36' 38''$ E and a latitude of $14^\circ 55' 04''$ N. The period between 2005 and 2007 was not considered in the evaluation as no discharge data were available. Also, due to inconsistency in the data for

November to May, the model evaluation was performed only for the months from June to October.

The freely available digital elevation model (DEM) of advanced space-borne thermal emission and reflection radiometer (ASTER) at 30 m spatial resolution was used to delineate the watershed and sub-watershed boundaries and generate drainage networks. In this study, the soil data were procured from the “National Bureau of Soil Survey and Land Use Planning (NBSS & LUP), Government of India” (Shivaprasad et al., 1998). The study area covers seven soil textural classes, as presented in Fig. 1. The spatial land use/land cover map was procured from the “National Remote Sensing Centre (NRSC) Hyderabad, Government of India”. The study area covers regionally important ten land use/land cover classes (NRSC, 2014) (Fig. 1).

The vegetation parameters were defined in the models based on the land use/cover map. The ET and vegetation parameters, including Leaf Area Index (LAI) and Albedo, were obtained from MODIS 1 km 8-day composite product (MOD16A2). The MODIS onboard the Aqua and Terra satellites makes available reliable ET estimates at different spatial–temporal resolutions (Anderson et al., 2011; Senay et al., 2013). The downloaded MODIS products were pre-processed to filter out poor-quality pixels utilizing MODIS Quality Control (QC) band. Finally, the ET values are resampled at the model grid/hru scale for the analysis. These pre-processing steps were performed on the MODIS data using the model builder and batch processing tools of ArcGIS. The study used the MODIS ET estimates at 8-day and monthly temporal resolutions to validate the model simulation-based ET values.

Table 1 Details of datasets used in the present study

Data type	Data name	Resolution	Duration	Data source	Online data reference
Meteorological Forcing	Precipitation	$0.25^\circ \times 0.25^\circ$	2000–2010	IMD Gridded	Pai et al., (2014, 2015)
	Temperature	$1^\circ \times 1^\circ$	2000–2010	IMD Gridded	Srivastava et al. (2009)
	Humidity	$0.25^\circ \times 0.25^\circ$	2000–2010	NCEP–CFSR	http://globalweather.tamu.edu/
	Wind Speed	$0.25^\circ \times 0.25^\circ$	2000–2010	NCEP–CFSR	http://globalweather.tamu.edu/
	Solar Radiation	$0.25^\circ \times 0.25^\circ$	2000–2010	NCEP–CFSR	http://globalweather.tamu.edu/
Hydrological data	River discharge	–	2000–2010	CWC Gauge	http://india-wris.nrsc.gov.in/
Thematic Data	Topography	30 m	2008	ASTER GDEM	http://earthexplorer.usgs.gov/
	Land use	1:250,000	2007–08	NRSC, ISRO	NRSC (2014)
	Soil	1:250,000	1998	NBSS&LUP	Shivaprasad et al. (1998)
	Leaf Area Index	1 km	2000–2010	MODIS	http://modis.gsfs.nasa.gov
	Albedo	1 km	2000–2010	MODIS	http://modis.gsfs.nasa.gov
	Evapotranspiration	1 km	2000–2010	MODIS	http://modis.gsfs.nasa.gov

Model Performance Evaluation

In this study, the model simulation performance was evaluated using the four statistical measures, namely coefficient of correlation (CC), root-mean-square error (RMSE) observations, standard deviation ratio (RSR), percent error (PBIAS), and index of agreement (d-index). In addition, the Nash–Sutcliffe model efficiency (NSE) was also used to evaluate the model for discharge simulation. The d-index, which varies between 0 (no agreement) and 1 (perfect agreement), measures the degree of model simulation error (Willmott et al., 1985) (Eq. 1). The CC, which ranges from -1 to $+1$, measures the direction and strength of a linear relationship between observed and estimated data (Eq. 2). The CC value of 1 represents the perfect correlation, while 0 represents no correlation, and— and + signs indicate negative and positive linear correlations between the observed and simulated values. The RSR, which ranges from the optimal value of 0 to a large positive value, is estimated as the ratio of RMSE and a standard deviation of the measured data (Eq. 3). The PBIAS was used to assess systematic over- or under-prediction and varies between -100 and ∞ (Xu et al., 2010) (Eq. 4). The PBIAS value close to 0 shows a perfect agreement between observed and simulated data. The NSE is a normalized statistic that determines the relative magnitude of the residual variance compared to the measured data variance (Nash & Sutcliffe, 1970) (Eq. 5). NSE ranges between $-\infty$ and 1.0, with $NSE = 1.0$ being the optimal value.

$$d - \text{index} = 1 - \frac{\sum_{i=1}^n (Y_i^{\text{sim}} - Y_i^{\text{obs}})^2}{\sum_{i=1}^n \left(\left| Y_i^{\text{sim}} - \overline{Y^{\text{obs}}} \right| + \left| Y_i^{\text{obs}} - \overline{Y^{\text{obs}}} \right| \right)^2} \quad (1)$$

$$CC = \left[\frac{\sum_{i=1}^n (Y_i^{\text{obs}} - \overline{Y^{\text{obs}}}) (Y_i^{\text{sim}} - \overline{Y^{\text{sim}}})}{\sqrt{\sum_{i=1}^n (Y_i^{\text{obs}} - \overline{Y^{\text{obs}}})^2} \sqrt{\sum_{i=1}^n (Y_i^{\text{sim}} - \overline{Y^{\text{sim}}})^2}} \right] \quad (2)$$

$$RSR = \left[\frac{\sqrt{\sum_{i=1}^n (Y_i^{\text{obs}} - Y_i^{\text{sim}})^2}}{\sqrt{\sum_{i=1}^n (Y_i^{\text{obs}} - \overline{Y^{\text{obs}}})^2}} \right] \quad (3)$$

$$PBIAS = \left[\frac{\sum_{i=1}^n (\overline{Y^{\text{sim}}} - \overline{Y^{\text{obs}}}) * (100)}{\sum_{i=1}^n \overline{Y^{\text{obs}}}} \right] \quad (4)$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (Y_i^{\text{obs}} - Y_i^{\text{sim}})^2}{\sum_{i=1}^n (Y_i^{\text{obs}} - \overline{Y^{\text{obs}}})^2} \right] \quad (5)$$

where Y_i^{sim} , Y_i^{obs} , $\overline{Y^{\text{sim}}}$ and $\overline{Y^{\text{obs}}}$ are the simulated, observed, average simulated, and average observed values, respectively.

SWAT Model Setup

The hydrological SWAT model is governed by water mass balance. The model processes depend on the discretized hydrological response units (HRUs) and are simulated at daily time steps using the following soil water balance equation (Eq. 6) (Neitsch et al., 2011).

$$SW_t = SW_o + \sum_{i=1}^n (R_{\text{day}} - Q_{\text{surf}} - E_a - w_{\text{seep}} - Q_{\text{gw}}) \quad (6)$$

where SW_t = final soil water content (mm); t = time (days); SW_o = initial soil water content on day i (mm); R_{day} = amount of precipitation on day i (mm); Q_{surf} = amount of surface runoff on day i (mm); E_a = amount of ET on day i (mm); w_{seep} = amount of percolation and bypass exiting the soil profile bottom on day i (mm); Q_{gw} = amount of return flow on day i (mm).

Modified rational method and modified soil conservation service curve number (SCS-CN) method (USDA, 1972) are used to compute peak runoff rate and surface runoff, respectively. The actual ET, as well as potential transpiration, is calculated using the Penman–Monteith method. In the present study, Muskingum method approach (Cunge, 1969) was adopted for flood routing.

The required weather and spatial datasets were prepared using the ArcGIS interface. The whole Marol watershed was discretized into several smaller sub-watersheds, which were further sub-divided into HRUs representing homogeneous combinations of land use/land cover, soil texture, and slope class. In this study, 1% threshold for each land use, soil, and the slope was considered and generated 647 HRUs. The ASTER DEM data were used to delineate watershed, sub-watershed, and drainage networks. By specifying the initial threshold on the drainage area, the ArcSWAT interface allows the user to fix the number of sub-watersheds. This study uses a threshold value of 8000 ha to delineate the drainage network and define outlet points for discretizing the Marol watershed into 31 sub-watershed. The threshold value of 8000 ha was considered to discretize the watershed so that each sub-watershed has a drainage area smaller than the precipitation grid area. This ensures minimum spatial degradation of precipitation data for capturing temporal variations over the watershed. The

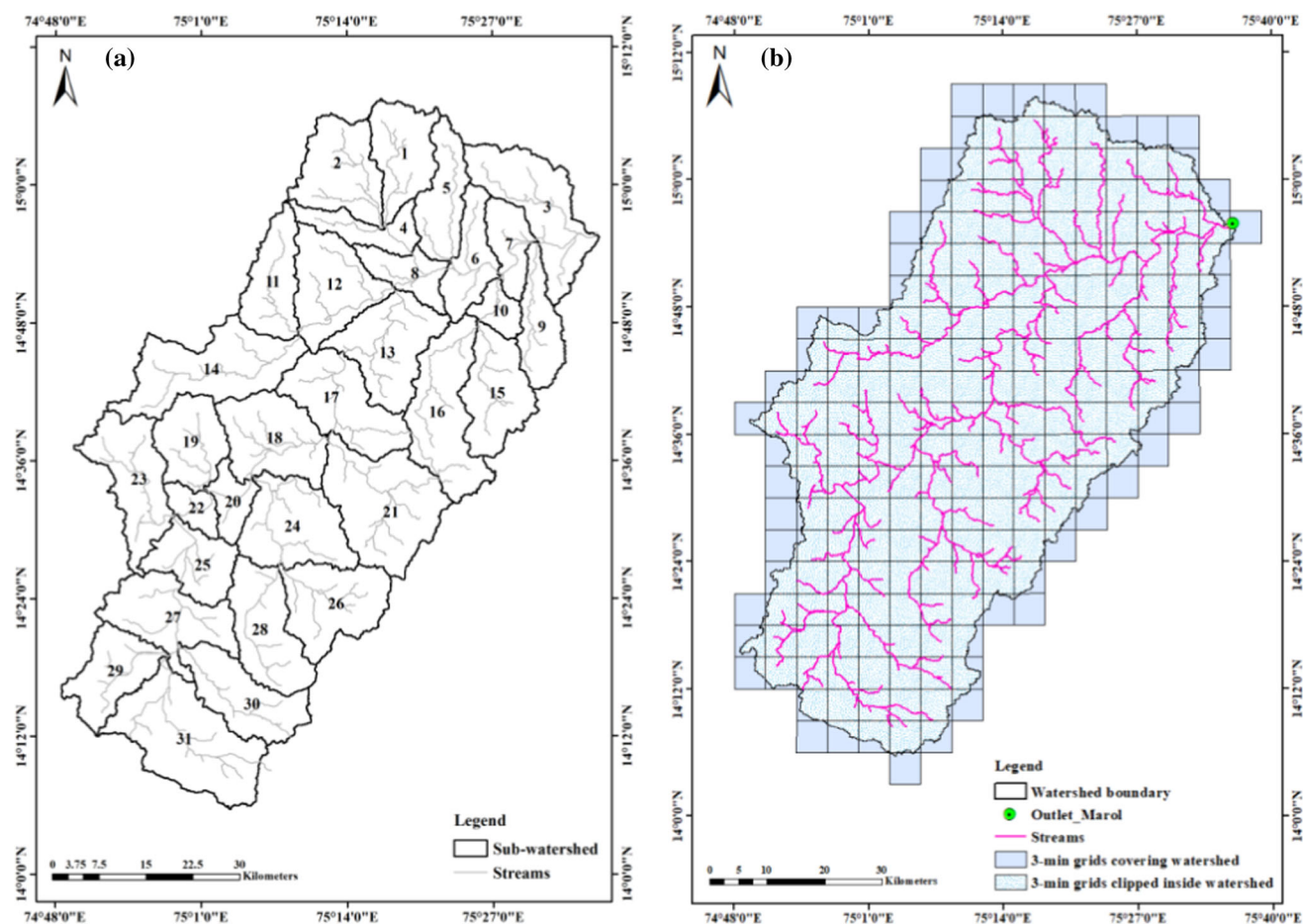


Fig. 2 a Delineated sub-watersheds and reach map of the study area, and b 3-min grids covering the Marol watershed

minimum and maximum area of a particular sub-watershed are estimated as 48.25 km² and 311.35 km², respectively, with an average area of 164.26 km². The delineated sub-watersheds and reach map of the study area are presented in Fig. 2. The SWAT model was simulated at daily time steps in this study.

VIC Model Setup

The VIC model accounts for sub-grid scale land use fractions of transpiration from vegetation, canopy layer evaporation, and soil evaporation for partitioning the grid-scale ET (Liang et al., 1994). The VIC model computes the water balances and surface energy budgets within the specified grid using the water budget equation (Eq. 7) (Narendra et al., 2017). The budget balance equation for the canopy layer is expressed in Eq. (8).

$$\frac{\partial S}{\partial t} = PR - ET - RF \tag{7}$$

$$\frac{\partial W_i}{\partial t} = PR - E_c - P_t \tag{8}$$

where $\frac{\partial S}{\partial t}$ = change in water storage, PR = precipitation, ET = Evapotranspiration, RF = runoff, $\frac{\partial W_i}{\partial t}$ = change in canopy intercepted water, E_c = evaporation from the canopy layer, and P_t = throughfall.

The total evapotranspiration is computed as the summation of canopy layer evaporation (E_c), evaporation from bare soil (E_b), and transpiration from vegetation (E_t), as follows (Eq. 9) (Liang et al., 1994):

$$ET = \sum_{n=1}^N C_v[n] \cdot (E_c[n] + E_t[n]) + C_v[N + 1] \cdot E_b \tag{9}$$

$C_v[n]$ = fraction of vegetation cover for the nth surface cover class (vegetation tile). $C_v[N + 1]$ = fraction of area covered with bare soil.

To calculate the runoff, the VIC model utilizes the variable infiltration curve accounting for the spatial variation. It states that runoff generates from two upper layers of soil when received precipitation and soil moisture from the initial time step exceeds the storage capacity of the soil.

The VIC model computes the fluxes at each cell, consisting of the discharge, base flow, evapotranspiration, soil moisture, and other outputs. These outputs are routed using a separate routing model to obtain the discharge at the outlet locations. The routing model developed by Lohmann et al. (1998) was used for this present study. The routing model states that the water flowing outwards of any grid cell does not flow back toward the same grid cell, and the water after entering the river channel no longer remains part of the water balance. The routing model comprises routing within the grid cell and routing in the channel (river routing). Impulse response function within each grid cell, which is depicted by the Linear transfer function, is used for routing within the grid cell, whereas after streamflow reaches the channel, Saint–Venant’s equation-based channel routing is used to generate the discharge at the required outlet.

$3' \times 3'$ grid (~ 5.5 km) resolution was used to estimate water balance components for the Marol watershed of Karnataka, India. The Marol watershed is covered under 216 grid points. The average rainfall, ET, and runoff fluxes were calculated based on clipped gridded portion coming inside the watershed boundary by the weighted area method. The dominant land-use type in the Marol watershed is agricultural land (77.56%, covering 168 grids) followed by deciduous forest (10.74%, covering 23 grids). The model was run in water balance computation mode from January 2000 to December 2010 using various geospatial datasets (Table 1). Soil hydrologic and hydraulic characteristics for 3-layer depth were calculated from NBSS & LUP soil map based on USDA soil texture classification. The three soil layer depths at 0–15 cm, 15–35 cm, and 35–100 cm intervals were set for the top, bottom, and deep layers, respectively. A soil texture ID was assigned to each grid cell for its use in the VIC model. One of the strengths of the VIC model is its capability to compute variable infiltration through the definition of multiple soil layers. Each grid fraction covering specific soil was given to the input file of the VIC model. The major surface soil types in the Marol watershed categorized under USDA textural classification are silty clay (31.11%, covering 68 grids) followed by sandy clay loam (28.01%, covering 61 grids). The delineated sub-watersheds and reach map of the study area and 3-min grids covering the Marol watershed are presented in Fig. 2.

Ensemble of Model Outputs

Hydrological models—VIC and SWAT—although provide satisfactory outputs, the uncertainties caused by inadequate data and assumptions/simplifications made in modeling are obvious. A variety of techniques has been developed to minimize modeling uncertainties, such as by incorporating

in situ data through data assimilation and by merging outputs of multiple models by generating ensembles. A multi-model ensemble can provide a more skillful and reliable system of hydrological simulation by combining the strengths of multiple models. A classic argument to support the use of a multi-model approach has been that it allows “compensatory effects” that control the excess spread coming from individual model errors. However, it should also be regarded that the verification metrics used to compare the single best model with several multi-model configurations might make a big difference when deciding what approach should be used (Mendoza et al., 2014). There are several approaches to combining the outputs of different models, while a simple and effective way of developing a model ensemble is to take the arithmetic mean of output variables. This approach has been used widely by researchers to simulate hydrological variables (Williams, 1969; Doblas-Reyes et al., 2005; Baker & Ellison, 2008; Kumar & Nandagiri, 2015; Muhammad et al., 2018; Li et al., 2018; Horan et al., 2021), and the same approach has also been adopted in this study. The HRU-based SWAT-simulated runoff is compared with grid-based VIC-simulated runoff at the watershed outlet. However, for ET comparison, an average of SWAT- and VIC-simulated ET over the watershed was considered. A schematic outlining the procedure used to generate the ensemble model is presented in Fig. 3.

Results and Discussion

Sensitivity and Uncertainty Analysis

The sensitivity and uncertainty analyses of the SWAT model parameters were performed using the Sequential Uncertainty Fitting (SUFI-2) algorithm of the SWAT calibration and uncertainty program (SWAT-CUP) (Abbaspour et al., 2007). The analysis showed a p-factor between 0.6 and 1 and an r-factor between zero and 0.3, i.e., the model simulation values correspond to observed values. Hence, the uncertainty associated with the model simulation was considered lower and acceptable. This study considered 17 most sensitive parameters (Table 2). The analysis showed that streamflow is most sensitive to CH_N2 (Manning’s ‘n’ value for the main channel), followed by CH_K2 (Effective hydraulic conductivity in main channel alluvium).

Further, to calibrate the VIC model, seven model parameters, namely infiltration parameter (b-infilt), sub-surface flow parameters (Ds, Dsmax, and Ws), and three soil layers (d_1 , d_2 , and d_3), were considered (Table 3). The b-infilt parameter was altered to a low and high value to match the observed peak flows. A lower value was given to

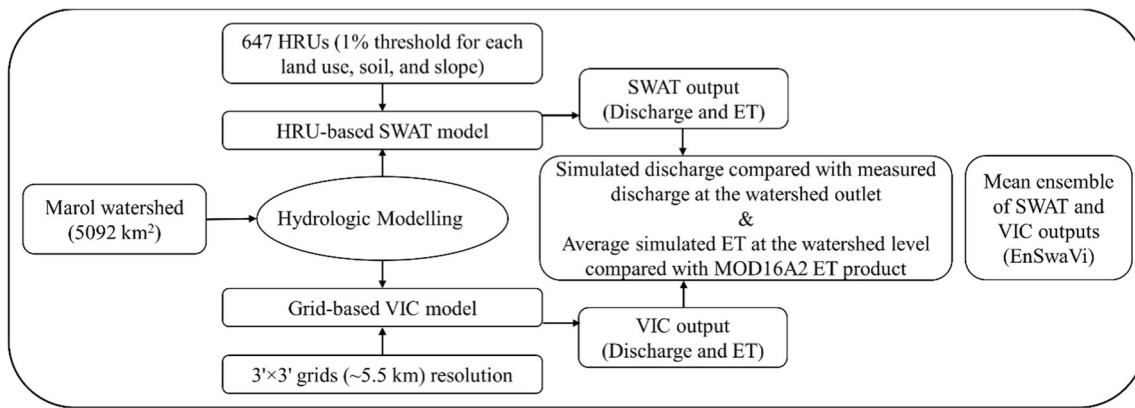


Fig. 3 A schematic outlining the procedure used to generate the ensemble model

Table 2 Sensitivity order and calibrated values of the SWAT model parameters for Marol Watershed

Sensitivity order	Parameters [#]	Parameter's description	Range used for calibration	Calibrated value
1	v__CH_N2.rte	Manning's "n" value for the main channel	0.01 to 0.3	0.20
2	v__CH_K2.rte	Effective hydraulic conductivity in main channel alluvium	100 to 500	377.20
3	r__CN2.mgt	SCS runoff curve number	- 20 to + 20%	0.11%
4	v__ALPHA_BF.gw	Base-flow alpha factor (days)	0.8 to 1	0.81
5	v__RCHRG_DP.gw	Deep aquifer percolation fraction	0 to 1	0.86
6	r__SOL_K.sol	Saturated hydraulic conductivity	- 20% to + 20%	- 0.17%
7	v__EPCO.hru	Plant uptake compensation factor	0 to 1	0.98
8	r__SLSUBBSN.hru	Average slope length	- 20% to + 20%	14.68%
9	a__GWQMN.gw	Threshold depth of water required for return flow to occur in the shallow aquifer (mm)	- 1000 to + 1000	- 798.00
10	r__SOL_AWC.sol	Available water capacity of the soil layer	- 20% to + 20%	- 0.09%
11	v__GW_REVAP.gw	Groundwater "revap" coefficient	0.02 to 0.2	0.15
12	a__GW_DELAY.gw	Groundwater delay (days)	0 to 470	67.21
13	v__SURLAG.bsn	Surface runoff lag time	0.05 to 24	18.23
14	a__REVAPMN.gw	Threshold depth of water in the shallow aquifer for "revap" to occur (mm)	- 100 to + 300	- 13.99
15	v__ESCO.hru	Soil evaporation compensation factor	0 to 1	0.93
16	v__CH_S1.sub	Average slope of tributary channels	0.0001 to 10	4.19
17	v__CH_COV2.rte	Channel cover factor	0.5 to 1	0.91

[#] The extension in the name of the parameter file indicates the process controlled by the parameter; rte = water routing, mgt = crop cover management, gw = groundwater, sol = soil water dynamics, hru = water dynamics at HRU level, bsn = entire watershed scale. The initials before the name of the parameter file indicate the method used for defining the range of parameters in auto-calibration; v = replacement of value within a given range, r = relative change to an initial value, a = absolute change with respect to the default value

lower the peak, and a higher value to increase the peak. The Dsmax and Ds parameters were adjusted to fit the baseflow, while parameter Ws was adjusted to fit the soil moisture. Detailed information on sensitivity and uncertainty analysis adopted for both SWAT and VIC models can be found in the supplementary file.

Evaluation of the SWAT, VIC, and EnSwaVi Models

The SWAT, VIC, and EnSwaVi models were evaluated for the period from 2000 to 2010 on a daily and monthly basis for discharge simulation and an 8-daily and monthly basis for ET simulation. The model results were evaluated using observed discharge data for the Varada River at the Marol

Table 3 Sensitivity order of the VIC model parameters for Marol Watershed

Sr. no.	Parameter	Parameter's Description	Range used for calibration	Calibrated value
01	b-infiltr	variable infiltration curve parameter	0.125 to 0.15	0.128
02	Ds	Fraction of DsMax parameter	0.170 to 0.475	0.348
03	Dsmax	Maximum velocity of baseflow, mm/day	0.102 to 1.000	0.352
04	Ws	Fraction of maximum soil moisture	0.500 to 0.664	0.565
05	d₁ (cm)	The thickness of the first (top thin) soil moisture layer	–	15
06	d₂ (cm)	The thickness of the second soil moisture layer	–	35
07	d₃ (cm)	The thickness of the third soil moisture layer	–	100

G & D site in the state of Karnataka, India, and the reference ET dataset from the MODIS. The total available observed data series were divided into two parts, 2000–2004 for calibration and 2008–2010 for validation, out of which the year 2000 was used as the model warm-up period. The performance evaluation of the SWAT, VIC, and EnSwaVi models for discharge and ET simulations is presented in Tables 4 and 5, respectively.

The discharge simulation analysis based on PBIAS for the SWAT and VIC model indicates that the SWAT model captured the physical processes accurately on a daily basis during both calibration and validation stages. However, it marginally underestimated the discharge values in the monthly simulation. On the other hand, the VIC model overestimated discharge values during the calibration and validation stage at daily and monthly time intervals (Table 5). Further, the coefficient of correlation (CC) for the SWAT model-simulated discharge was consistently better than the VIC model-simulated discharge, especially for daily simulations. Similarly, the RSR values for the SWAT-simulated discharge were comparatively better than the VIC model-simulated discharge for both daily and monthly simulations. In general, different performance statistics (PBIAS, CC, RSR, d index) indicated that the EnSwaVi model's performance was marginally better than

the SWAT and VIC models for discharge simulation on a daily and monthly scale.

Table 5 presents the performance evaluation of the SWAT and VIC model while predicting the ET at 8-daily and monthly time intervals. Significant underestimation by the SWAT model (negative PBIAS) and overestimation by VIC model (Positive PBIAS) was observed while simulating the evapotranspiration at 8-daily and monthly intervals during the calibration stage. In general, the VIC model performance was comparatively better than the SWAT model, especially for monthly simulation (Table 5). Overall, the performance of both SWAT and VIC models for ET simulation was good. However, the EnSwaVi model-simulated ET values were considerably better than SWAT, and VIC-simulated ET on daily and the monthly timescale. Hence, it can be inferred that the ensemble model provides a better ET estimate than the individual one (Table 5).

Evaluation of Discharge

The observed and simulated daily discharges for the calibration and validation period using SWAT, VIC, and EnSwaVi models are presented in Fig. 4. Similarly, the observed and simulated monthly discharges for the

Table 4 Performance evaluation of the SWAT, VIC, and EnSwaVi for discharge simulation

Sr. no.	Parameter	SWAT				VIC				EnSwaVi			
		Daily		Monthly		Daily		Monthly		Daily		Monthly	
		Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.
1	PBIAS	0.3	2.1	– 6.1	– 2.3	32.7	2.2	36.5	11.3	10.1	2.1	15.0	4.5
2	CC	0.87	0.90	0.94	0.95	0.78	0.83	0.96	0.95	0.89	0.90	0.98	0.96
3	RSR	0.51	0.49	0.33	0.33	1.28	0.71	0.58	0.45	0.49	0.49	0.28	0.34
4	d-index	0.93	0.94	0.97	0.97	0.78	0.89	0.93	0.96	0.94	0.94	0.98	0.97
5	NSE	0.82	0.81	0.83	0.84	0.75	0.78	0.81	0.79	0.91	0.92	0.93	0.92

Cal., calibration; Val., validation

Table 5 Performance evaluation of the SWAT and the VIC and EnSwaVi for evapotranspiration simulation

Sr. no.	Parameter	SWAT				VIC				EnSwaVi			
		8-Daily		Monthly		8-Daily		Monthly		8-Daily		Monthly	
		Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.
1	PBIAS	– 11.8	– 4.7	– 6.8	– 8.2	12.3	14.9	1.8	4.7	1.2	5.1	0.5	1.8
2	CC	0.68	0.60	0.77	0.62	0.62	0.64	0.70	0.69	0.71	0.75	0.82	0.71
3	RSR	1.93	1.94	1.08	1.14	1.04	1.24	0.94	0.71	0.75	0.84	0.73	0.76
4	d index	0.67	0.63	0.71	0.64	0.72	0.66	0.72	0.71	0.83	0.83	0.88	0.86

Cal., calibration; Val., validation

calibration and validation period are presented in Fig. 5. The scatter plot between observed and simulated discharge for daily and monthly calibration and validation using the SWAT, VIC, and EnSwaVi outputs are presented in supplementary Figs. 1 and 2. The graphical results show that the observed and simulated discharges using the SWAT model closely matched for the most part except for some high-flow events, which were slightly underestimated. Similarly, a good agreement between the observed and simulated hydrographs was observed using the VIC model. However, in general, the high flow events were overestimated. The SWAT simulation using monthly discharge data has performed better, which reveals that in comparison to short-term or single storm simulation, the SWAT model performs better for long-term simulation, and such observations were also reported previously (Borah et al., 2007). In general, the EnSwaVi-simulated discharge values were found more accurate than the SWAT- and VIC-simulated discharge values on both daily and monthly scales, specifically, over low and high flows. These results are also reflected through different performance statistics (Table 4).

For the SWAT model simulation on a daily scale, the coefficient of correlation (CC) values were estimated as 0.87 and 0.90; however, on the monthly timescale, the CC values were estimated as 0.94 and 0.95 during the calibration and validation period, respectively. Similarly, for the VIC model simulation on a daily scale, the CC values were estimated as 0.78 and 0.83; however, on a monthly timescale, the CC values were estimated as 0.96 and 0.95 during the calibration and validation periods, respectively (Table 4). For EnSwaVi-simulated discharge, the CC values were estimated as 0.89 and 0.90; however, on a monthly timescale, the CC values were estimated as 0.98 and 0.96 during the calibration and validation period, respectively. The CC values in EnSwaVi-simulated discharge were improved by approximately 10–15% on a daily scale and 5–6% on a monthly scale as compared to the VIC-simulated discharge.

The RSR values for daily simulation were estimated as 0.51 and 0.49 using the SWAT model, while 1.28 and 0.71 using the VIC model during the calibration and validation period, respectively. Similarly, the RSR values for monthly simulation were estimated as 0.33 and 0.33 using the SWAT model, while 0.58 and 0.45 using the VIC model during the calibration and validation period, respectively (Table 4). The RSR values for daily simulation using the EnSwaVi model were estimated as 0.49 during both calibration and validation period, which showed a considerable/clear improvement over VIC-simulated discharge. Similarly, a significant improvement was seen for monthly simulation. However, improvements were not substantial as compared to SWAT-simulated discharge.

The different performance evaluation criteria showed a good agreement between observed and simulated hydrographs on daily and monthly timescales, indicating the good performance of the SWAT and the VIC models (Moriasi et al., 2007). PBIAS of 0.3 and 2.1 for daily calibration and validation, respectively, using the SWAT model, indicated that on average, the SWAT model overestimated discharge by 0.3% and 2.1% during daily calibration and validation, respectively (Fig. 4). Similarly, PBIAS of 32.7 and 2.2 for daily calibration and validation, respectively, using the VIC model, indicated that, on average, the VIC model overestimated discharge by 32.7% and 2.2% during daily calibration and validation, respectively (Fig. 5). A similar trend was observed for monthly simulation; the VIC model overestimated the discharge during both calibration and validation period; however, negligible overestimation/underestimation was observed using the SWAT model. On the other hand, the EnSwaVi model overestimated the discharge by 10.1% and 2.1% during the calibration and validation periods, respectively, on a daily scale. Moreover, it overestimated the discharge by 15% and 4.5% on a monthly scale (Table 4). Based on the PBIAS values, it can be inferred that overestimation was significantly lower in the EnSwaVi model compared to the VIC model at daily and monthly timescale.

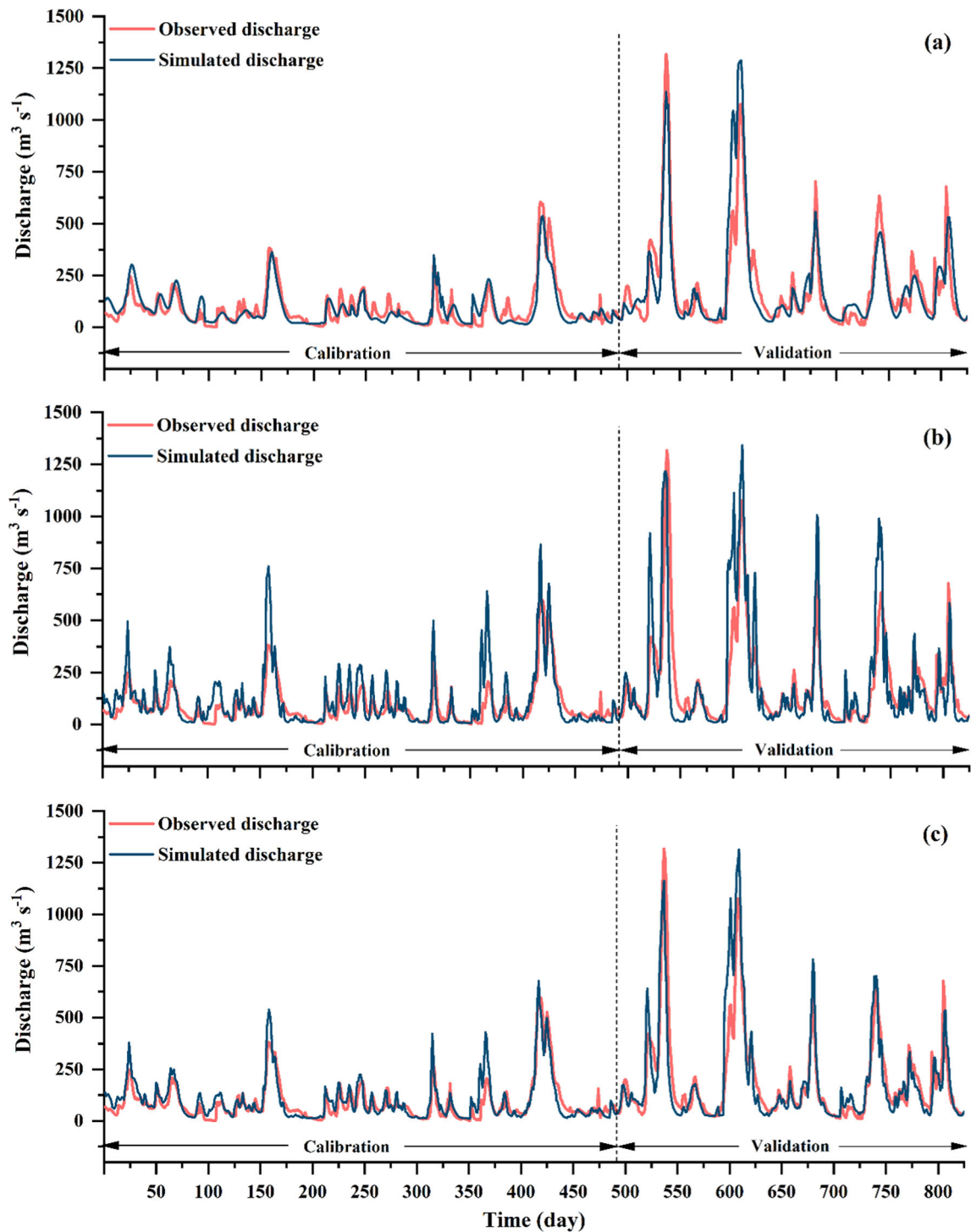


Fig. 4 Comparison of the observed and simulated discharge for daily calibration (2001–2004) and validation (2008–2010) using the **a** SWAT model, **b** VIC model, and **c** EnSwaVi output

Evaluation of Evapotranspiration

The observed and simulated 8-daily ET for the evaluation period using the SWAT, VIC, and EnSwaVi hydrologic models are presented in Fig. 6. Similarly, the observed and

simulated monthly ET for the calibration and validation period are presented in Fig. 7. The scatter plot between observed and simulated evapotranspiration for 8-daily and monthly calibration and validation using the SWAT, VIC, and EnSwaVi outputs are presented in supplementary

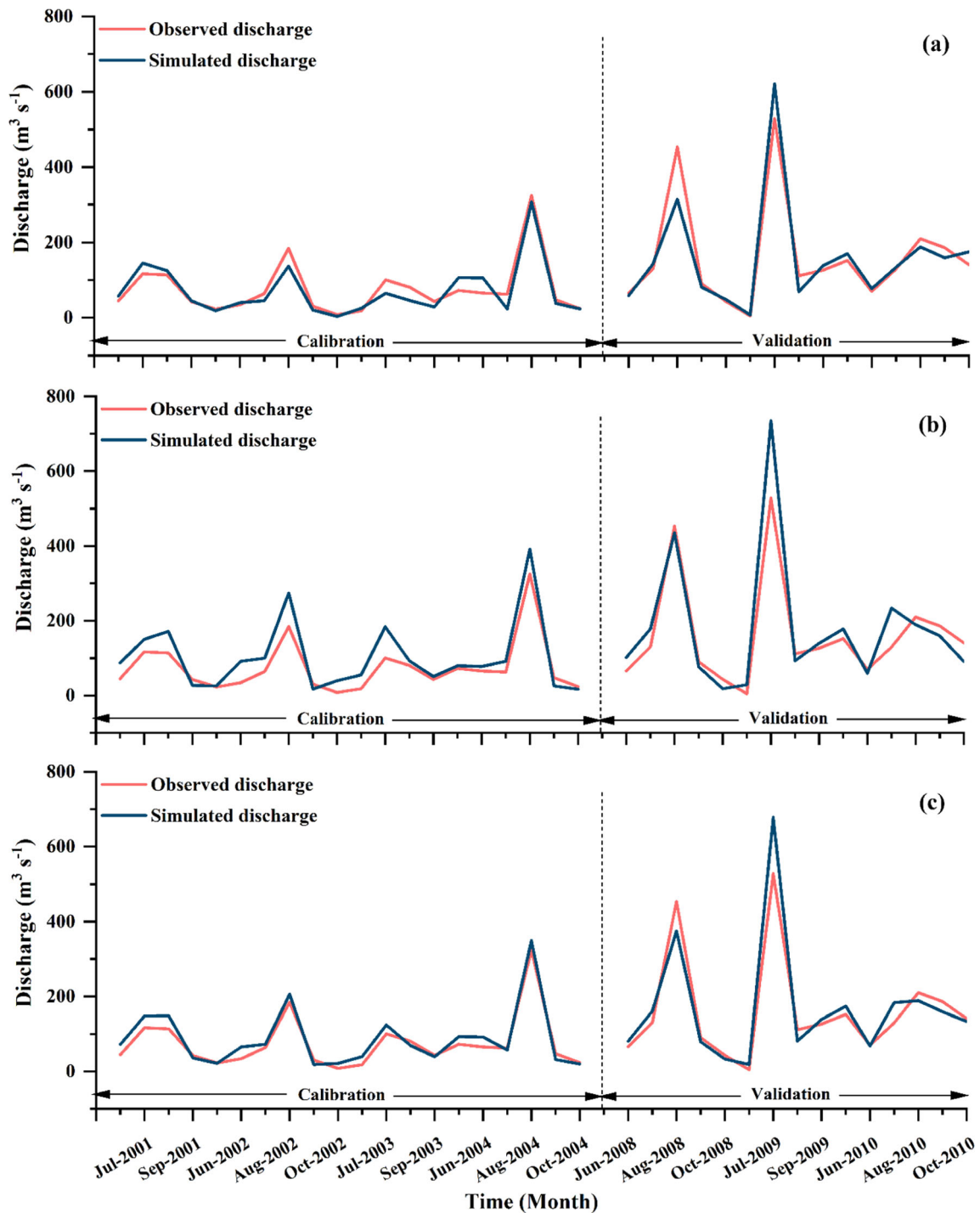


Fig. 5 Comparison of the observed and simulated discharge for monthly calibration (2001–2004) and validation (2008–2010) using the **a** SWAT model, **b** VIC model, and **c** Ensemble output

Figs. 3 and 4. The graphical results show that the observed and simulated ET were mostly matched during the simulation period using both SWAT and VIC models. However, the VIC simulation results were comparatively matching better with the reference ET dataset than the SWAT simulation results. It can also be seen that the EnSwaVi

simulation results matched closely with the reference ET. Interestingly, these EnSwaVi-simulated ET estimates were better than the SWAT and VIC model's simulated ET.

For the SWAT model simulation on an 8-daily scale, the CC values were estimated as 0.68 and 0.60, however, on a monthly timescale, the CC values were estimated as 0.77

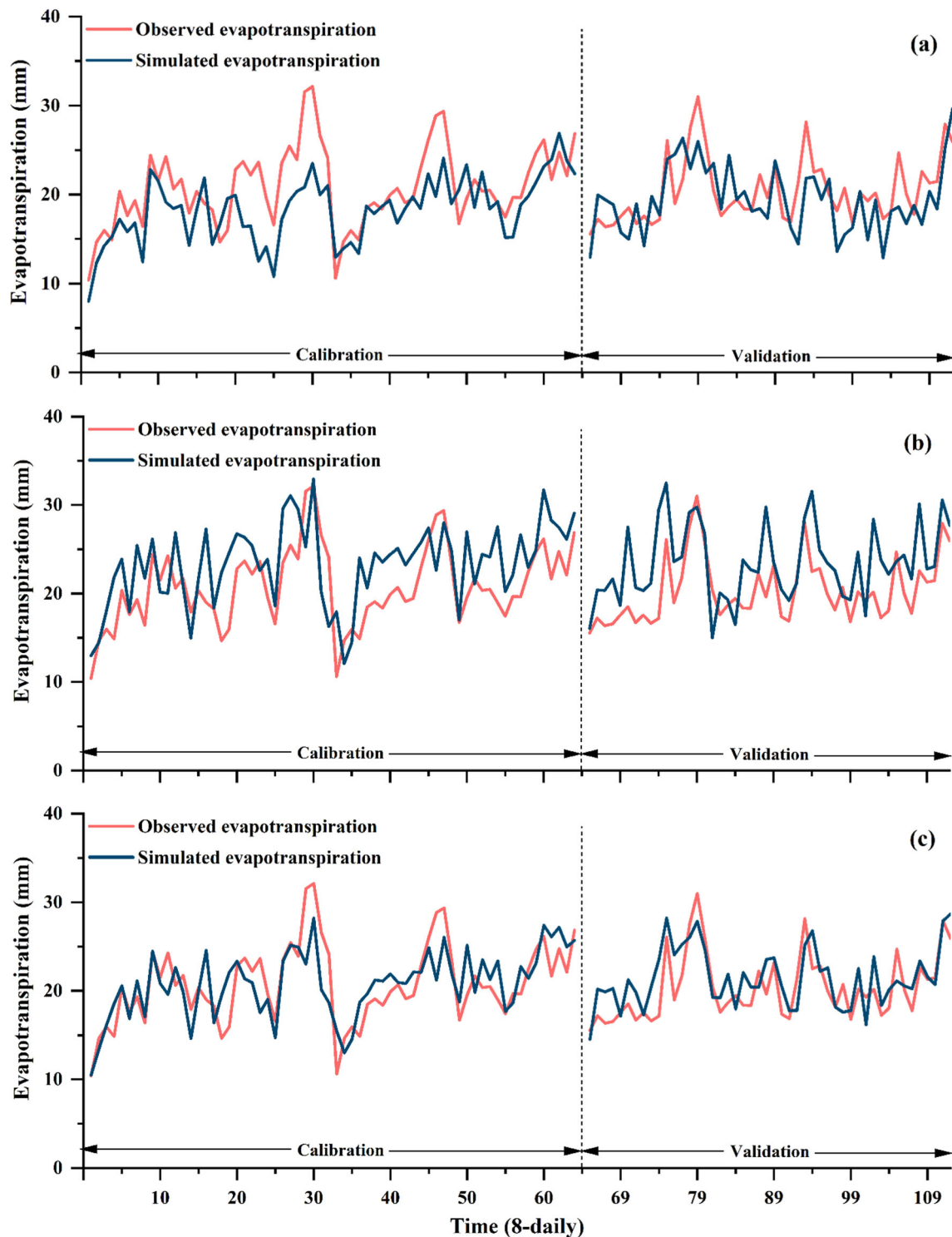


Fig. 6 Comparison of the observed and simulated evapotranspiration for 8-daily calibration (2000–2004) and validation (2008–2010) using the a) SWAT model, b) VIC model, and c) ensemble results

and 0.62 during the calibration and validation period, respectively. Similarly, for the VIC model simulation on an 8-daily scale, the CC values were estimated as 0.62 and 0.64; however, on a monthly timescale, the CC values were

estimated as 0.70 and 0.69 during calibration and validation period, respectively. It is interesting to note that the CC values were 0.71 and 0.75 for the EnSwaVi model simulation on an 8-daily scale, however, the values were

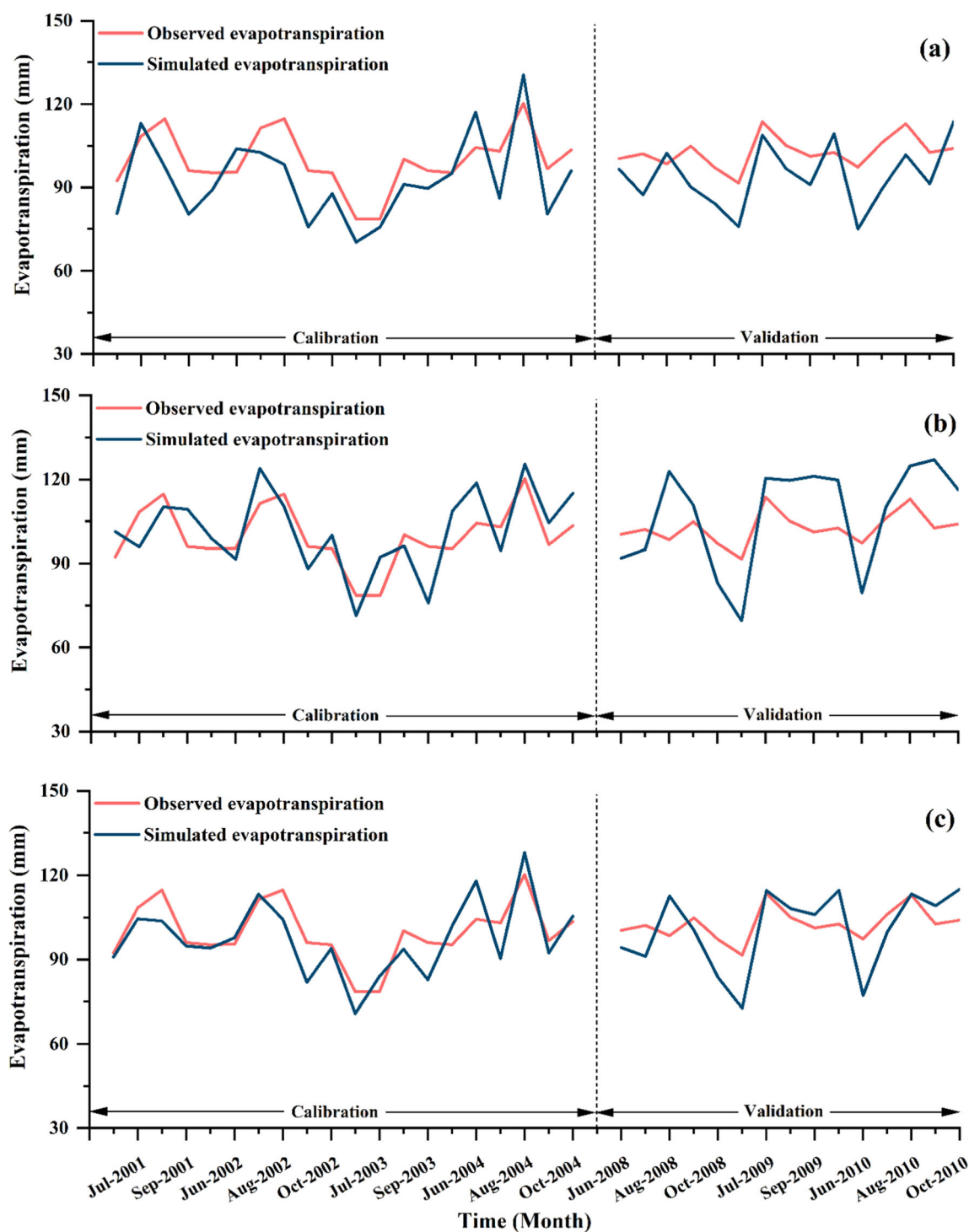


Fig. 7 Comparison of the observed and simulated evapotranspiration for monthly calibration (2000–2004) and validation (2008–2010) using the **a** SWAT model, **b** VIC model, and **c** ensemble results

0.82 and 0.71 on a monthly scale during calibration and validation period, respectively (Table 5). The RSR values for an 8-daily simulation were estimated as 1.93 and 1.94 using the SWAT model, while 1.04 and 1.24 using the VIC model during calibration and validation period,

respectively. Similarly, the RSR values for monthly simulation were estimated as 1.08 and 1.14 using the SWAT model, while 0.94 and 0.71 using the VIC model during calibration and validation period, respectively. On the other hand, the RSR values for 8-daily simulation were 0.75 and

0.84 using the EnSwaVi model on an 8-daily scale; however, these values are 0.73 and 0.76 on a monthly scale during calibration and validation period, respectively (Table 5).

PBIAS of -11.8 and -4.71 for an 8-daily calibration and validation, respectively, using the SWAT model, indicated that on average, the SWAT model underestimated ET by 11.8% during calibration and 4.71% during the validation period (Fig. 6). Similarly, PBIAS of 12.3 and 14.9 for 8-daily calibration and validation, respectively, using the VIC model, indicated that on average, the VIC model overestimated ET by 12.3% during calibration and 14.9% during the validation period (Fig. 7). Performance of the VIC model for ET simulation was observed to be very good on a monthly timescale (PBIAS of 1.8 and 4.7 for calibration and validation, respectively). However, the performance of the SWAT model for ET simulation was observed relatively poor on a monthly timescale (PBIAS of -6.8 and -8.2 for calibration and validation, respectively). One can note that the PBIAS of 1.2 and 5.1 was observed for the EnSwaVi model on an 8-daily scale for ET simulation, while these values were 0.5 and 1.8 on a monthly scale for calibration and validation, respectively. Overall, different performance evaluation criteria showed relatively better performance of the VIC model compared to the SWAT model; however, the performance of the EnSwaVi model was marginally better than the SWAT and VIC models.

Discussion on the Performance of the SWAT and VIC Models

Both the hydrologic model SWAT and VIC can be efficiently applied to carry out water balance analysis and for planning and management of water resources. The SWAT model simulated the discharge more accurately than the VIC model. The results were in conformity with Hu et al. (2007), Dash et al. (2021), and Kang et al. (2022). In general, overestimation was observed using the VIC model during both calibration and validation period; this may be due to inconsideration of the upstream abstraction. However, the accuracy of the VIC model in simulating the ET was found better than the SWAT model. Study results contradict Dash et al. (2021), which revealed that the SWAT simulation-based standardized evapotranspiration drought index (SEDI) was consistent; in contrast, the VIC-3L simulation-based SEDI was continuously overestimated and underestimated with the benchmark satellite (MOD16A2-ET)-derived estimates. In general, underestimation was observed using the SWAT model during the calibration and validation periods. The SWAT hydrologic model lump the soil characteristics and land use in each

grid cell without considering the sub-grid scale variability of LULC and soil moisture resulting in more bias in the ET estimates (Rathjens et al., 2015). It computes water balance components over the HRUs and averages it for a sub-watershed. Conversely, the VIC modeling framework can be advantageous over the SWAT model as it accounts for the sub-grid scale variability of soil types, soil moisture, and vegetation; hence, it can simulate the ET more closely to reality (Srivastava et al., 2017). One can note that the EnSwaVi-simulated discharges were considerably better than VIC-simulated discharges; however, the improvement was not much as compared to the SWAT-simulated discharges. On the other hand, the EnSwaVi-simulated ET was found superior to both SWAT and VIC-simulated ET. These results revealed that the ensemble model performs better as compared to the individual model for ET as well as discharge simulations (Horan et al., 2021). Our study outcomes are consistent with Horan et al. (2021) and Muhammad et al. (2018), and suggest that an ensemble reduces the noise, bias, and variance of simulations and can potentially create a more in-depth understanding of the data. However, ensemble modeling results can suffer from a lack of interpretability and are dependent on the prediction accuracy of the ensemble members.

While dividing the data into calibration and validation subsets, it is important to check the data which present the same statistical population (Masters, 1993). The model performs better if it does not extrapolate beyond the range of the data used for model calibration (Tokar & Johnson, 1999). Although, in general, the calibration and validation datasets have relatively similar statistical characteristics, it has been observed that a few higher-value peaks are there in the validation period, which may be not well-calibrated, which resulted in higher overestimation/underestimation of these peak values. The SWAT and VIC models were calibrated and validated using observed discharge data at the watershed outlet and the average reference ET at the watershed scale only. Though model calibration performance seems quite good for the calibrated gauging station, multi-site evaluation of the models should be carried out to achieve a better representation of the physical parameters and to improve the model's predictability. But due to the availability of observed data at watershed outlet only, single-site calibration was carried out in this study. The model's simulation capability could also be improved if standardized MODIS-derived ET estimates are used since the MODIS-derived ET estimates are generally not free from bias.

The average annual water balance for simulation has also been estimated for the entire 31 sub-watersheds using the SWAT model. It has been inferred that about 39.75% flows out as surface runoff from the watershed, out of annual average precipitation of 1330.90 mm. ET has been

found predominant and accounts for about 38.46% of the annual average precipitation falling over the area. It was observed that almost all the sub-watersheds flow out more than 25% of annual precipitation as surface runoff, indicating the need for implementing suitable soil and water management programs to decrease the runoff volume by increasing in-watershed application of water, in turn minimizing soil erosion.

The SWAT and VIC hydrologic models are useful platforms extensively applied for water resource assessment and management worldwide. The ensemble of VIC and SWAT outputs, i.e., EnSwaVi model, can be used by policymakers to make decisions regarding water resource management in the study area. However, in this study, the water balance was carried out assuming that land use/land cover and other parameters remain constant with time. In reality, several parameters change with time/season. Therefore, a water balance study incorporating the variability of these parameters with time/season in the GIS environment can be a scope for future research. Further, it is recommended that additional studies should be conducted over other river basins/watersheds to evaluate the long-term capabilities of hydrologic models in simulating the water balance components.

Conclusions

In the present study, the SWAT and VIC hydrologic models were used to simulate the water balance components (runoff and ET) over an agriculture-based watershed. Further, the ensemble of VIC and SWAT outputs (EnSwaVi; averages of individual model-simulated datasets with equal weights) were also simulated for hydrological variables to assess whether modeling uncertainties could be minimized. Following major conclusions were drawn from the present study:

1. The results revealed that discharge had been simulated well using both SWAT (d-index 0.93 and 0.94 during daily calibration and validation; d-index 0.97 during both monthly calibration and validation) and VIC (d-index 0.78 and 0.89 during daily calibration and validation; d-index 0.93 and 0.96 during monthly calibration and validation) models. However, the discharge simulated by the SWAT model was found more accurate than the VIC model.
2. The performance of the VIC model (d-index 0.72 and 0.66 during 8-daily calibration and validation; d-index 0.72 and 0.71 during monthly calibration and validation) in simulating ET was found better as compared to the SWAT model (d-index 0.67 and 0.63 during

8-daily calibration and validation; d-index 0.71 and 0.64 during monthly calibration and validation).

3. The EnSwaVi (ensemble of VIC and SWAT) model-simulated runoff (d-index 0.94 during both daily calibration and validation; d-index 0.98 and 0.97 during monthly calibration and validation) and ET (d-index 0.83 during both 8-daily calibration and validation; d-index 0.88 and 0.86 during monthly calibration and validation) were more accurate than individual SWAT and VIC outputs.
4. Based on the results, it can be concluded that both SWAT and VIC models can be efficiently applied to carry out water balance analysis and for planning and management of water resources. However, the EnSwaVi model could marginally improve the results.
5. ET has been found predominant and accounts for about 38.46% of the annual average precipitation falling over the area. It has been inferred that about 39.75% flows out as surface runoff from the watershed, out of annual average precipitation of 1330.90 mm.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aadhar, S., Swain, S., & Rath, D. R. (2019). Application and performance assessment of SWAT hydrological model over Kharun river basin, Chhattisgarh, India. *World environmental and water resources congress 2019: Watershed management, irrigation and drainage, and water resources planning and management* (pp. 272–280). American Society of Civil Engineers.
- Abbaspour, K. C., Vajdani, M., & Haghghat, S. (2007). SWAT-CUP calibration and uncertainty programs for SWAT. In MODSIM 2007- international congress on modelling and simulation, modelling and simulation society of Australia and New Zealand, pp. 1603–1609.
- Alvarenga, L. A., Carvalho, V. S. O., Oliveira, V. A. D., Mello, C. R. D., Colombo, A., Tomasella, J., & Melo, P. A. (2020). Hydrological simulation with SWAT and VIC Models in the Verde River Watershed, Minas Gerais. *Revista Ambiente & Água*, 15.
- Amrit, K., Mishra, S. K., Pandey, R. P., Himanshu, S. K., & Singh, S. (2019). Standardized precipitation index-based approach to predict environmental flow condition. *Ecohydrology*, 12(7), e2127.
- Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., Gonzalez-Dugo, M. P., Cammalleri, C., D'Urso, G., Pimstein, A., & Gao, F. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrology and Earth System Sciences*, 15, 223–239.

- Arnold, J. G., & Fohrer, N. (2005). SWAT 2000: 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 1. *Journal of the American Water Resources Association*, 34(1), 73–89.
- Baker, L., & Ellison, D. (2008). Optimisation of pedotransfer functions using an artificial neural network ensemble method. *Geoderma*, 144(1–2), 212–224. <https://doi.org/10.1016/j.geoderma.2007.11.016>
- Barnes, B. S. (1940). Discussion on analysis of runoff characteristics by O H. Meyer. *Transactions of the American Society of Civil Engineers*, 105, 104–106.
- Borah, D. K., Arnold, J. G., Bera, M., Krug, E. C., & Liang, X. Z. (2007). Storm event and continuous hydrologic modeling for comprehensive and efficient watershed simulations. *Journal of Hydrologic Engineering*. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:6\(605\),605-616](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:6(605),605-616)
- Borah, D. K., & Bera, M. (2003). Watershed-scale hydrologic and nonpoint-source pollution models: Review of mathematical bases. *Transactions of the ASAE*, 46(6), 1553.
- Crawford, N.H., & Linsley, R.K. (1966). Digital simulation in hydrology: Stanford watershed model IV. Technical Report No. 39. Department of Civil Engineering, Stanford University, p. 210.
- Cummings, N. W. (1935). Evaporation from water surfaces: Status of present knowledge and need for further investigations. *Transactions, American Geophysical Union*, 16(2), 507–510.
- Cunge, J. A. (1969). On the subject of a flood propagation computation method (Muskingum method). *Journal of Hydraulic Research*, 7(2), 205–230.
- Dash, S. S., Sahoo, B., & Raghuvanshi, N. S. (2021). How reliable are the evapotranspiration estimates by Soil and Water Assessment Tool (SWAT) and Variable Infiltration Capacity (VIC) models for catchment-scale drought assessment and irrigation planning? *Journal of Hydrology*, 592, 125838. <https://doi.org/10.1016/j.jhydrol.2020.125838>
- Dhami, B., Himanshu, S. K., Pandey, A., & Gautam, A. K. (2018). Evaluation of the SWAT model for water balance study of a mountainous snowfed river basin of Nepal. *Environmental Earth Sciences*, 77(1), 21.
- Dietrich, J., Schumann, A. H., Redetzky, M., Walther, J., Denhard, M., Wang, Y., Pfützner, B., & Büttner, U. (2009). Assessing uncertainties in flood forecasts for decision making: prototype of an operational flood management system integrating ensemble predictions. *Natural Hazards and Earth System Sciences*, 9(4), 1529–1540. <https://doi.org/10.5194/nhess-9-1529-2009>
- Dile, Y. T., & Srinivasan, R. (2014). Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: An application in the Blue Nile river basin. *Journal of the American Water Resources Association*, 50, 1226–1241. <https://doi.org/10.1111/jawr.12182>
- Doblas-Reyes, F. J., Hagedorn, R., & Palmer, T. N. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting—II. Calibration and combination. *Tellus A: Dynamic Meteorology and Oceanography*, 57(3), 234–252. <https://doi.org/10.3402/tellusa.v57i3.14658>
- Fair, G. M., & Hatch, L. P. (1933). Fundamental factors governing the streamline flow of water through sand. *Journal American Water Works Association*, 25, 1551–1565.
- Fan, Y., Himanshu, S. K., Ale, S., DeLaune, P. B., Zhang, T., Park, S. C., Colaizzi, P. D., Evett, S. R., & Baumhardt, R. L. (2022). The synergy between water conservation and economic profitability of adopting alternative irrigation systems for cotton production in the Texas High Plains. *Agricultural Water Management*, 262, 107386. <https://doi.org/10.1016/j.agwat.2021.107386>
- Garg, K. K., Bharati, L., Gaur, A., George, B., Acharya, S., Jella, K., & Narasimhan, B. (2012). Spatial mapping of agricultural water productivity using the SWAT model in Upper Bhima Catchment India. *Irrigation and Drainage*, 61(1), 60–79.
- Gaur, S., Bandyopadhyay, A., & Singh, R. (2021a). From changing environment to changing extremes: Exploring the future streamflow and associated uncertainties through integrated modelling system. *Water Resources Management*, 35(6), 1889–1911. <https://doi.org/10.1007/s11269-021-02817-3>
- Gaur, S., Bandyopadhyay, A., & Singh, R. (2021b). Projecting land use growth and associated impacts on hydrological balance through scenario-based modelling in the Subarnarekha basin. *India. Hydrological Sciences Journal*, 66(14), 1997–2010. <https://doi.org/10.1080/02626667.2021.1976408>
- Gaur, S., Singh, B., Bandyopadhyay, A., Stisen, S., & Singh, R. (2022). Spatial pattern-based performance evaluation and uncertainty analysis of a distributed hydrological model. *Hydrological Processes*. <https://doi.org/10.1002/hyp.14586>
- Green, W. H., & Ampt, G. A. (1911). Studies on soil physics, 1. The flow of air and water through soils. *Journal of Agricultural Sciences*, 4, 11–24.
- Gupta, A., Himanshu, S. K., Gupta, S., & Singh, R. (2020). Evaluation of the SWAT model for analysing the water balance components for the upper Sabarmati Basin. In R. AIKhaddar, R. K. Singh, S. Dutta, & M. Kumari (Eds.), *Advances in Water Resources Engineering and Management: Select Proceedings of TRACE 2018* (pp. 141–151). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-8181-2_11
- Hengade Narendra, T. I., & Eldho, G. S. (2017). Hydrological simulation of a large catchment using the variable infiltration capacity model. In V. Garg, V. P. Singh, & V. Raj (Eds.), *Development of water resources in India* (pp. 19–30). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-55125-8_2
- Himanshu, S. K., Pandey, A., & Dayal, D. (2018a). Evaluation of satellite-based precipitation estimates over an agricultural watershed of India. *World environmental and water resources congress 2018: watershed management, irrigation and drainage, and water resources planning and management* (pp. 308–320). American Society of Civil Engineers.
- Himanshu, S. K., Pandey, A., & Patil, A. (2018b). Hydrologic evaluation of the TMPA-3B42V7 Precipitation data set over an agricultural watershed using the SWAT model. *Journal of Hydrologic Engineering*, 23(4), 05018003.
- Himanshu, S. K., Pandey, A., & Shrestha, P. (2017). Application of SWAT in an Indian river basin for modeling runoff, sediment and water balance. *Environmental Earth Sciences*, 76(1), 3.
- Himanshu, S. K., Pandey, A., Yadav, B., & Gupta, A. (2019). Evaluation of best management practices for sediment and nutrient loss control using SWAT model. *Soil and Tillage Research*, 192, 42–58. <https://doi.org/10.1016/j.still.2019.04.016>
- Horan, R., Gowri, R., Wable, P. S., Baron, H., Keller, V. D., Garg, K. K., Mujumdar, P. P., Houghton-Carr, H., & Rees, G. (2021). A comparative assessment of hydrological models in the Upper Cauvery catchment. *Water*, 13(2), 151. <https://doi.org/10.3390/w13020151>
- Horton, R. E. (1919). Rainfall interception. *Monthly Weather Review*, 147, 603–623.
- Horton, R. E. (1939). Analysis of runoff-plat experiments with varying infiltration-capacity. *Transactions, American Geophysical Union*, 20(4), 693. <https://doi.org/10.1029/TR020i004p00693>
- Hu, H., Wang, G., Bi, X., Yang, F., & Chongyi, E. (2007). Application of two hydrological models to Weihe River basin:

- a comparison of VIC-3L and SWAT. In *Geoinformatics 2007: Geospatial Information Technology and Applications* (Vol. 6754, p. 67541T). International Society for Optics and Photonics. <https://doi.org/10.1117/12.764920>
- ICWE: International Conference on Water and the Environment. (1992). Dublin, Ireland. <http://www.wmo.int/pages/prog/hwrrp/documents/english/icwedece.html>
- Jacob, C. E. (1943). Correlation of groundwater levels and precipitation on Long Island, New York: 1. Theory. *Transactions, American Geophysical Union*, 24, 564–573.
- Jacob, C. E. (1944). Correlation of groundwater levels and precipitation on Long Island, New York: 2. Correlation of data. *Transaction, American Geophysical Union*, 24, 321–386.
- Jajarmizadeh, M., Harun, S., Ghahraman, B., & Mokhtari, M. H. (2012). Modeling daily stream flow using plant evapotranspiration method. *International Journal of Water Resources and Environmental Engineering*, 4(6), 218–226.
- Kang, H., & Sridhar, V. (2018). Improved drought prediction using near real-time climate forecasts and simulated hydrologic conditions. *Sustainability*, 10(6), 1799. <https://doi.org/10.3390/su10061799>
- Kang, H., Sridhar, V., & Ali, S. A. (2022). Climate change impacts on conventional and flash droughts in the Mekong River Basin. *Science of The Total Environment*. <https://doi.org/10.1016/j.scitotenv.2022.155845>
- Kasiswathanan, K. S., & Sudheer, K. P. (2017). Methods used for quantifying the prediction uncertainty of artificial neural network based hydrologic models. *Stochastic Environmental Research and Risk Assessment*, 31(7), 1659–1670. <https://doi.org/10.1007/s00477-016-1369-5>
- Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., & Thielen, J. (2016). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software*, 75, 68–76. <https://doi.org/10.1016/j.envsoft.2015.09.009>
- Keulegan, G. H. (1944). Spatially variable discharge over a sloping plane. *Transactions, American Geophysical Union*, 25(6), 956. <https://doi.org/10.1029/TR025i006p00956>
- Kimball, J. S., Running, S. W., & Nemani, R. (1997). An improved method for estimating surface humidity from daily minimum temperature. *Agricultural and Forest Meteorology*, 85(1–2), 87–98.
- Kumar, A., Kumar, M., Pandey, R., ZhiGuo, Y., & Cabral-Pinto, M. (2021a). Forest soil nutrient stocks along altitudinal range of Uttarakhand Himalayas: An aid to nature based climate solutions. *CATENA*, 207, 105667.
- 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>
- Kumar, A., Taxak, A. K., Mishra, S., & Pandey, R. (2021b). Long term trend analysis and suitability of water quality of River Ganga at Himalayan hills of Uttarakhand. *India. Environmental Technology & Innovation*, 22, 101405.
- Kumar, R., & Nandagiri, L. (2015). *Evaluating uncertainty of the soil and water assessment tool (SWAT) model in the upper Cauvery basin*. Karnataka.
- Li, Z., Yu, J., Xu, X., Sun, W., Pang, B., & Yue, J. (2018). Multi-model ensemble hydrological simulation using a BP neural network for the upper Yalongjiang River Basin, China. *Proceedings of the International Association of Hydrological Sciences*, 379, 335–341. <https://doi.org/10.5194/piahs-379-335-2018>
- Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, 99(D7), 14415–14428.
- Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*. <https://doi.org/10.1029/2006WR005756>
- Liu, Y. R., Li, Y. P., Huang, G. H., Zhang, J. L., & Fan, Y. R. (2017). A Bayesian-based multilevel factorial analysis method for analyzing parameter uncertainty of hydrological model. *Journal of Hydrology*, 553, 750–762. <https://doi.org/10.1016/j.jhydrol.2017.08.048>
- Lohmann, D., Raschke, E., Nijssen, B., & Lettenmaier, D. P. (1998). *Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model*. *Hydrological sciences journal*, 43(1), 131–141.
- Madolli, M. J., Himanshu, S. K., Patro, E. R., & De Michele, C. (2022). Past, present and future perspectives of seasonal prediction of Indian summer monsoon rainfall: A review. *Asia-Pacific Journal of Atmospheric Sciences*, 58(4), 591–615. <https://doi.org/10.1007/s13143-022-00273-6>
- Manning, R. (1891). On the flow of water in open channels and pipes. *Transactions of the Institution of Civil Engineers of Ireland*, 20, 161–207.
- Masters, T. (1993). *Practical neural network recipes in C++*. Academic Press.
- Mendoza, P. A., Rajagopalan, B., Clark, M. P., Cortés, G., & McPhee, J. (2014). A robust multi-model framework for ensemble seasonal hydroclimatic forecasts. *Water Resources Research*, 50(7), 6030–6052. <https://doi.org/10.1002/2014WR015426>
- Miller, S. N., Semmens, D. J., Goodrich, D. C., Hernandez, M., Miller, R. C., Kepner, W. G., & Guertin, D. P. (2007). The automated geospatial watershed assessment tool. *Environmental Modelling & Software*, 22(3), 365–377.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885–900.
- Muhammad, A., Stadnyk, T. A., Unduche, F., & Coulibaly, P. (2018). Multi-model approaches for improving seasonal ensemble streamflow prediction scheme with various statistical post-processing techniques in the Canadian Prairie region. *Water*, 10(11), 1604. <https://doi.org/10.3390/w10111604>
- Murty, P. S., Pandey, A., & Suryavanshi, S. (2014). Application of semi-distributed hydrological model for basin level water balance of the Ken basin of Central India. *Hydrological Processes*, 28(13), 4119–4129.
- Nash, J., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.
- Neitsch, S. L., Arnold, J.G., Kiniry, J. R., & Williams, J. R. (2011). Soil and water assessment tool theoretical documentation version 2009. Texas Water Resources Institute Technical Report No. 406, Texas A & M University System, College Station, Texas.
- NRSC. (2014). Land Use/Land Cover database on 1:50,000 scale, Natural Resources Census Project, LUCMD, LRUMG, RSAA, National Remote Sensing Centre, ISRO, Hyderabad.
- Oubeidillah, A. A., Kao, S. C., Ashfaq, M., Naz, B. S., & Tootle, G. (2014). A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US. *Hydrology and Earth System Sciences*, 18(1), 67–84.
- Pai, D. S., Sridhar, L., Badwaik, M. R., & Rajeevan, M. (2015). Analysis of the daily rainfall events over India using a new long period (1901–2010) high resolution (0.25 × 0.25) gridded rainfall data set. *Climate Dynamics*, 45(3–4), 755–776.
- Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). Development of a new high spatial

- resolution (0.25 × 0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, 65(1), 1–18.
- Palmate, S. S., & Pandey, A. (2021). Effectiveness of best management practices on dependable flows in a river basin using hydrological SWAT Model. In S. K. Ashish Pandey, M. L. Mishra, R. D. Kansal, & V. P. S. Singh (Eds.), *Water management and water governance: hydrological modeling* (pp. 335–348). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-58051-3_22
- Palmate, S. S., Wagner, P. D., Fohrer, N., & Pandey, A. (2021). Assessment of uncertainties in modelling land use change with an integrated cellular automata–Markov chain model. *Environmental Modeling & Assessment*. <https://doi.org/10.1007/s10666-021-09804-3>
- Pandey, A., Himanshu, S. K., Mishra, S. K., & Singh, V. P. (2016). Physically based soil erosion and sediment yield models revisited. *CATENA*, 147, 595–620.
- Pandey, A., & Palmate, S. S. (2019). Assessing future water–sediment interaction and critical area prioritization at sub-watershed level for sustainable management. *Paddy and Water Environment*, 17(3), 373–382. <https://doi.org/10.1007/s10333-019-00732-3>
- Paul, P. K., Zhang, Y., Ma, N., Mishra, A., Panigrahy, N., & Singh, R. (2021). Selecting hydrological models for developing countries: Perspective of global, continental, and country scale models over catchment scale models. *Journal of Hydrology*, 600, 126561.
- Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A*, 193, 120–145.
- Rathjens, H., Oppelt, N., Bosch, D. D., Arnold, J. G., & Volk, M. (2015). Development of a grid-based version of the SWAT landscape model. *Hydrological Processes*, 29(6), 900–914.
- Senay, G. B., Bohms, S., Singh, R. K., Gowda, P. H., Velpuri, N. M., Alemu, H., & Verdin, J. P. (2013). Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. *Journal of the American Water Resources Association*, 49(3), 577–591.
- Shiklomanov, I. A. (1998). World water resources. A new appraisal and assessment for the twenty first century. UNESCO, Paris.
- Shivaprasad, C. R., Reddy, R. S., Sehgal, J., & Velayutham, M. (1998). Soils of Karnataka for optimizing land use. NBSS Publ. 47b (Soils of India Series). Nagpur, India: National Bureau of Soil Survey and Land Use Planning.
- Soil Conservation Service (SCS). (1956). Supplement A, Section 4, Chapter 10, Hydrology. National engineering handbook, USDA, Washington D.C.
- Srivastava, A. K., Raajeevan, M., & Kshirsagar, S. R. (2009). Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region. *Atmospheric Science Letters*, 10(October), 249–254.
- Srivastava, A., Sahoo, B., Raghuwanshi, N. S., & Singh, R. (2017). Evaluation of variable-infiltration capacity model and MODIS-terra satellite-derived grid-scale evapotranspiration estimates in a river basin with tropical monsoon-type climatology. *Journal of Irrigation and Drainage Engineering*, 143(8), 04017028.
- Swain, S., Mishra, S. K., & Pandey, A. (2021). A detailed assessment of meteorological drought characteristics using simplified rainfall index over Narmada River Basin India. *Environmental Earth Sciences*, 80(6), 221.
- Swain, S., Mishra, S. K., Pandey, A., Pandey, A. C., Jain, A., Chauhan, S. K., & Badoni, A. K. (2022). Hydrological modelling through SWAT over a Himalayan catchment using high-resolution geospatial inputs. *Environmental Challenges*, 8, 100579.
- Tanmoyee, B., Raju, P. V., & Hakeem, A. (2015). Climate change impact on snowmelt runoff modelling for Alaknanda river basin. *Journal of Environment and Earth Science*, 5(11), 56–67.
- Theis, C. V. (1935). The relation between the lowering of the piezometric surface and the rate and duration of discharge of well using ground-water storage. *Transactions, American Geophysical Union*, 16, 519–524.
- Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical Review*, 38, 55–94.
- Tokar, A. S., & Johnson, P. A. (1999). Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering*, 4(3), 232–239.
- USDA Soil Conservation Service. (1972). Hydrolog. In V. Mockus (Ed.), *National engineering Handboo*. Washington, DC: US Department of Agriculture-Soil Conservation Service.
- Veettil, A. V., Mishra, A. K., & Green, T. R. (2022). Explaining Water security indicators using hydrologic and agricultural systems models. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2022.127463>
- Water facts. (2020). Worldwide water supply. Bureau of Reclamation California-Great Basin. <https://www.usbr.gov/mp/arwec/water-facts-ww-water-sup.html> (Accessed 20 July 2022).
- Willmott, C. J., Ackleson, S. G., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., O'Donnell, J., & Rowe, C. M. (1985). Statistics for the evaluation and comparison of models. *Journal of Geophysical Research: Oceans*, 90(C5), 8995–9005.
- Xu, H., Taylor, R. G., Kingston, D. G., Jiang, T., Thompson, J. R., & Todd, M. C. (2010). Hydrological modeling of River Xiangxi using SWAT2005: A comparison of model parameterizations using station and gridded meteorological observations. *Quaternary International*, 226(1–2), 54–59.
- Yadav, B., Gupta, P. K., Patidar, N., & Himanshu, S. K. (2020). Ensemble modelling framework for groundwater level prediction in urban areas of India. *Science of the Total Environment*, 712, 135539. <https://doi.org/10.1016/j.scitotenv.2019.135539>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.