University of Birjand

Water Harvesting Research Vol. 8, No. 1, Spring & Summer 2025, Original paper p. 78-92

10.22077/JWHR.2025.9156.1173

Evaluating the Performance of the SWAT Model in Hydrological Simulation Using Global and Local Land Use and Soil Data (Case Study: Chel-Chai Watershed)

Mohammad Mahdi Sahebi[a], Ali Mohammad Akhoond-Ali[b], Farshad Ahmadi[c&*], Amin Khoramian[4]

[a]Msc Student in Water Resources Engineering, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
[b]Professor, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
[c]Associate Professor, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
[d]Assistant professor, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

[*]Corresponding Author, E-mail address: f.ahmadi@scu.ac.ir
Received: 22 January 2025/ **Revised**: 15 April 2025/ **Accepted**: 30 April 2025

Abstract

Water, this vital element, plays an irreplaceable role in our lives. From drinking water supply to energy production, agriculture, industry, and numerous other sectors depend on it. However, human activities have profoundly impacted water resource availability. Under these circumstances, a deeper understanding of the hydrological cycle and river behavior becomes more crucial than ever. One key tool for hydrological cycle simulation is the SWAT (Soil and Water Assessment Tool) model. This study investigated the effects of using regional versus global soil and land use data on SWAT model performance in the Chel-Chai watershed. Monthly river discharge data from Lazoureh and Jangaldeh stations (2006-2020 for calibration; 1997-2005 for validation) were utilized. The results showed that at Lazoureh station, during the calibration phase, regional and global data showed similar performance; the NS and R indices for both data types were 0.60 and 0.77, respectively, and the MAE and RMSE errors were both 0.78 and 1.11. Consequently, no difference was observed between regional and global data in this phase. During the validation phase at Lazoureh station, regional data performed better than global data, reducing MAE and RMSE by 2.56% and 1.92%, respectively. At Jangaldeh station, during the calibration phase, regional data also outperformed global data. The NS and R indices for regional data were 0.73 and 0.90, respectively, while for global data they were 0.58 and 0.87. Regional data also showed better performance during the validation phase. The results demonstrate that regional data can provide more accurate river discharge estimates, particularly during validation phases. This study highlights the importance of spatial data resolution in hydrological modeling accuracy.

Keywords: Calibration, Performance criteria, Regional vs. Global Data, Validation.

1. Introduction

Human activities have profoundly impacted access to water resources. These impacts have not only accelerated climate change and glacier melting but have also disrupted the hydrological cycle, leading to consequences such as increased droughts and floods, alterations in watersheds, and consequent changes in river flow regimes. These changes have created significant challenges for water resource management worldwide. Under these circumstances, a deeper understanding of the hydrological cycle and river behavior has become more crucial than ever for optimal utilization of this vital resource.

Among hydrological simulation tools, the Soil and Water Assessment Tool (SWAT) model serves as one of the key instruments for hydrological cycle modeling. While numerous hydrological models with both unique and shared characteristics continue to be developed (Wang et al., 1996; DHI, 2004), SWAT has emerged as one of the most widely used watershed simulation models (Holisagar et al., 2023; Xiao et al., 2023), offering considerable flexibility in watershed configuration.

As a semi-distributed, physics-based model operating on a continuous daily time step, SWAT divides watersheds into smaller hydrological units called Hydrologic Response Units (HRUs) based on topography, land use, and slope. The model then calculates hydrological processes such as surface runoff and evapotranspiration within each HRU.

SWAT has been extensively applied for diverse simulations including river flow, phosphorus transport, pollutant movement, microorganisms, analysis of extreme precipitation and flood frequency trends, greenhouse gas studies, and water quality assessments (Abou Rafee et al., 2019; Machado et al., 2023; Neitsch et al., 2011; Nie et al., 2011; Peterson and Hamlett, 1998; Tullu, 2023; Wagena et al., 2017).

Numerous researchers worldwide have evaluated the performance of the SWAT model. Goodarzi et al. (2012) studied surface runoff which leads to soil erosion, sedimentation in reservoirs, and water quality degradation. In this research, runoff simulation in the Qarasu watershed was conducted using three models: SWAT, IHACRES, and SIMHYD. The evaluation criteria included the Nash-Sutcliffe coefficient (NS), coefficient of determination (R^2), and root mean square error (RMSE). The results indicated that during the calibration period, the SWAT model with an NS coefficient of 0.8 and the SIMHYD model with an NS coefficient of 0.68 showed the highest and lowest performance, respectively. Moreover, during the validation period, the SWAT model with an NS of 0.73 demonstrated the best performance in runoff simulation.

This research emphasizes the importance of selecting an appropriate model for water resources management. In a study by Shukla et al. (2023), the impact of land use changes on runoff in Germany's Rur River basin was investigated using the SWAT model. The SWAT model was calibrated based on observed flow and runoff data at three locations (Stah, Linnich, Monschau) between 2000-2010 and validated between 2011-2015. The hydrological model performance was evaluated using statistical parameters such as the coefficient of determination (R^2), p-value, r-value, and percent bias (PBIAS), showing that the mean R^2 values for model calibration and validation were 0.68 and 0.67, respectively.

The study evaluated the impact of three change scenarios on river runoff by replacing portions of forest with urban areas, agricultural lands, and grasslands compared to the 2006 land cover map. The SWAT model successfully simulated the patterns and effects of land use changes on river runoff. Results showed that changing land cover from deciduous forest to urban areas, agricultural lands, and grasslands increased total watershed runoff by 43%, 14%, and 4%, respectively.

In another study by Susiwidiyaliza et al. (2023) on the Batanghari River in West Sumatra, Indonesia, the impacts of predicted land use changes until 2040 and climate changes based on two different greenhouse gas concentration scenarios (RCP4.5 and RCP8.5) were analyzed. Results showed changing patterns in agricultural area from 1990, 1997, 2005, 2015 to 2040. The SWAT model was calibrated using SWAT-CUP. Model simulations showed that land use changes increased surface flow while decreasing lateral flow and baseflow, leading to increased sediment load over time. Climate change impacts on water and sediment increased the mean flow discharge ratio, resulting in more frequent droughts and floods.

Machado et al. (2023) examined the effects of climate change and land use changes on river flow in the Piracicaba basin using the SWAT model. They selected two time periods using the Pettitt non-parametric test to detect abrupt changes in annual flow: a pre-change period (1985-2000) and post-change period (2001-2015), using land use maps from 1990 and 2010. Results showed the SWAT model performed well, with Nash-Sutcliffe (NS) coefficients of 0.88 for both periods during calibration, and 0.92 and 0.94 respectively during validation. Simulation results showed river flow in the Piracicaba basin during 1985-2015 responded more to land use changes than to climate changes.

Liu et al. (2024) analyzed the impact of land use changes on river flow in the Danjiang River using land use data from three time periods (2000, 2010, and 2020). Results showed that from 2000 to 2020, cultivated land area decreased while forests, grasslands, and urban areas increased. The SWAT model based

on land use data showed good calibration and validation results. Future land use scenarios showed that expansion of farmland, grassland and urban areas increased river runoff, while increased forest area decreased runoff. This study created three SWAT models and combined different land use scenarios, which could serve as a scientific basis for rational water allocation and land use structure optimization in the Danjiang River source region.

In another study by Zewde et al. (2024), the SWAT model was used to examine runoff and sediment in the Jemma sub-basin of the Upper Nile in Ethiopia. The model estimated sediment yield with reasonable accuracy. Results showed that barren lands had the highest sensitivity and sediment load. Agricultural lands in lower, middle and upper sub-basin areas had higher sediment yields. Forested lands with dense vegetation and root density showed the lowest erosion sensitivity.

The SWAT model indicated maximum sediment yield at high elevations was 3685.14 t/ha, with annual average sediment yield of 78.1 t/ha. Effective vegetation cover and conservation practices could reduce sediment yield by 44.28% and 35.92% for RCP8.5 and RCP4.5 scenarios respectively. Designing and implementing appropriate soil and water management measures in the studied sub-basin is crucial for reducing soil erosion challenges.

This study aims to conduct a comprehensive hydrological analysis of the Chel-Chai watershed using the SWAT model. The primary focus of the research is to investigate the impact of global and regional-scale land use and soil data on river flow simulation. The accuracy of the SWAT model in predicting river discharge based on different datasets is evaluated, and the influence of these variables on the hydrological behavior of the watershed is analyzed.

2. Materials and Methods

2.1. Data and study area

The study watershed is located politically within Golestan Province, in the Minudasht County. This watershed is part of the Caspian Sea basin and the Gorganrud watershed. According to the classification of the Iran Water Resources Management Company, it belongs to the Golestan study area with code 1601. In terms of metric geographic coordinates (UTM), the study area is situated in zone 40 North, between eastings 346,888 to 379,700 meters and northings 4,093,983 to 4,122,514 meters. It covers an area of 469.9 square kilometers (equivalent to 46,998 hectares) with a perimeter of 108.9 kilometers. The elevation range of the area varies between 130 to 2,540 meters. The elevation map along with the study location, showing the watershed within the country, province, and county, is presented in Figure (1).

For hydrological analysis and simulation, precipitation data spanning a 30-year period (1992-2021) were utilized. Missing data were completed using nearby meteorological and hydrological stations, as these data play a crucial role in the accuracy of the analysis. The key data for this model include precipitation and temperature, where daily rainfall, maximum and minimum temperatures, and relative humidity were used for the study area. Data from eight stations -Ramian, Arazkuseh, Lazoureh, Tilabad, Galikesh, Gonbad, Nodeh Khormaloo, and Neraab - were employed to run the SWAT model (station details are presented in Table (1)).

Additionally, it should be noted that precipitation data from all eight stations were used, while for daily minimum temperature, maximum temperature, and relative humidity, data from Ramian station were utilized due to its most complete records.

In SWAT modeling, calibration and validation are essential steps following simulation. The required data for this phase include long-term runoff statistics at the watershed outlet. For this purpose, discharge data from the Jangaldeh and Lazoureh stations located on the river were used. The locations of these stations are shown in (Figure 1), with complementary information provided in (Table 1).

2.2. Soil data

For the soil component, SWAT requires input of various textural, physical and chemical soil properties including: Soil thickness, Bulk density, Available water, Hydraulic conductivity, Organic carbon, Clay percentage, Silt percentage, Sand percentage, Gravel percentage, Albedo coefficient, K-FACTOR, Soil EC (Electrical Conductivity) for each depth. The study area contains 16

different soil types in the regional dataset, designated as: AA, AB, AC, AD, AE, AF, AG, AH, AI, AJ, AK, AL, AM, AP, AQ, AR. Among these Soil type AD covers the largest area (23.93%) and Soil type AQ occupies the smallest area (0.19%).

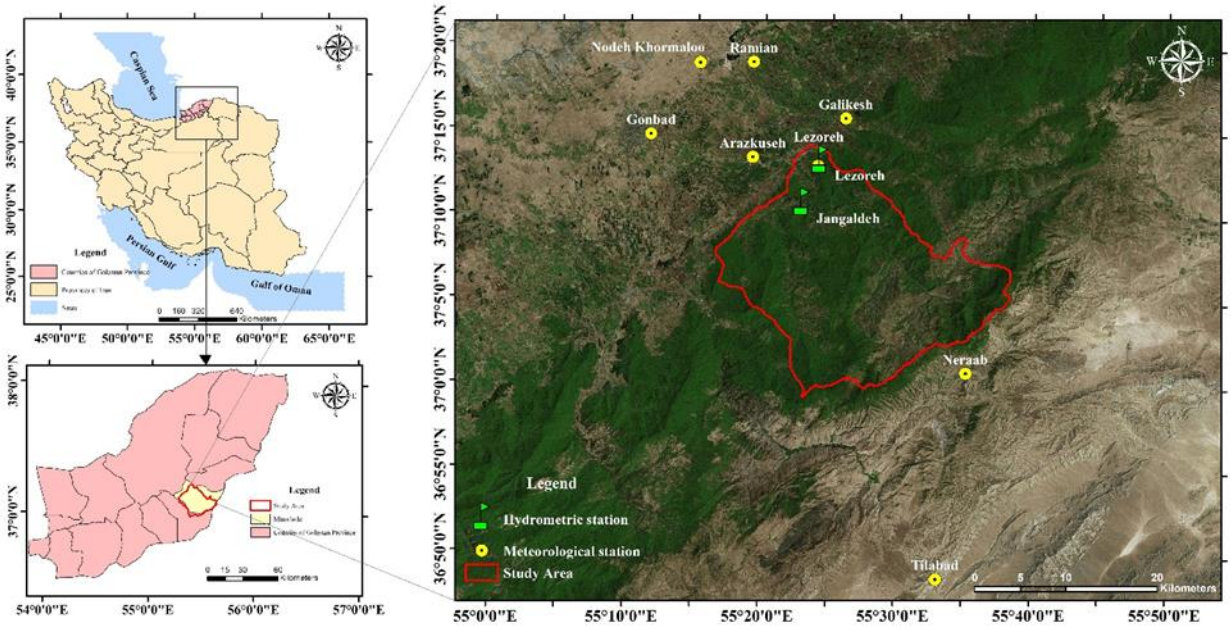

Fig. 1. Location of the study area In Iran and Golestan province Along with a map of elevation changes

Table 1. Characteristics of Meteorological Stations

Row	Station Name	Station Type	Longitude	Latitude	River	Elevation (m)
1	Ramian	Meteorological	55°19'07" E	37°19'01" N	Ghareh Chay	200
2	Arazkuseh	Meteorological	55°19'09" E	37°13'29" N	Gorganrud- Ghareh Soo	34
3	Lazoureh	Meteorological	55°24'00" E	37°13'00" N	Chehel Chay	200
4	Tilabad	Meteorological	55°33'07" E	36°48'41" N	Tilabad	1000
5	Galikesh	Meteorological	55°26'00" E	37°15'51" N	Oghan	222
6	Gonbad	Meteorological	55°11'35" E	37°14'46" N	Gorganrud	1000
7	Nodeh Khormaloo	Meteorological	55°15'09" E	37°19'04" N	Khormalu	1500
8	Neraab	Meteorological	55°35'08" E	37°00'52" N	Khormalu	280
9	Lazoureh	Hydrometric	55°24'09" E	37°13'28" N	Chehel Chay	196
10	Jangaldeh	Hydrometric	55°22'51" E	37°10'55" N	Narmab	180

The global soil data includes two classifications:

1. I-Rc-Yk-c-3508 (62.98% coverage)
   - Texture: Clay
   - Hydrologic group: High runoff potential
2. Xh38-3a-4056 (37.1% coverage)
   - Texture: Loam
   - Hydrologic group: High runoff potential

The complete regional soil information and spatial distribution are presented in Table (2) and Figure (2).

2.3. Land use information

The SWAT model requires land use data to define hydrological response units (HRUs). Based on this data, a land use map is prepared and each land use type is named according to the existing classifications. The land is divided into different subcategories and the percentage of area of each subcategory is calculated by the software. The major part of the area of the region is forest. In the following, the area of each section is given separately. Using regional data, 79.041% forest, 16.85% agriculture, 4.06% pasture, 0.04% riverside, which is given in Table (3). Using global data, it includes 63.24% broadleaf forest, 14.44% cropland, 12.60% savanna, 7.93% mixed forest, 0.89% irrigated pasture, 0.86% rainfed cropland and pasture, which is given in Table (3). The regional and global usage is shown in Figure (3).

2.4. SWAT model

The SWAT model represents the culmination of 30 years of modeling

experience by the United States Department of Agriculture (USDA) Agricultural Research Service, building upon previous USDA-ARS models including CREAMS, GLEAMS, and EPIC.

Table 2. Local soil characteristics in the SWAT model

Row	Soil Type in SWAT	Soil Texture	Soil Depth (cm)	Hydrologic Group
1	AA	Loam	65	B - Moderate Runoff Potential
2	AB	Loam	30	D - High Runoff Potential
3	AC	Loam	65	B - Moderate Runoff Potential
4	AD	Loamy Sand	80	B - Moderate Runoff Potential
5	AE	Loamy Sand	72	B - Moderate Runoff Potential
6	AF	Loam	110	B - Moderate Runoff Potential
7	AG	Loamy Sand	100	B - Moderate Runoff Potential
8	AH	Sandy Loam	90	C - Moderately High Runoff Potential
9	AI	Sandy Loam	114	B - Moderate Runoff Potential
10	AJ	Clay Loam	95<	C - Moderately High Runoff Potential
11	AK	Loam	100	B - Moderate Runoff Potential
12	AL	Clay Loam	60	D - High Runoff Potential
13	AM	Sandy Loam	90<	B - Moderate Runoff Potential
14	AP	Loam	65	D - High Runoff Potential
15	AQ	Loam	40	B - Moderate Runoff Potential
16	AR	Loamy Sand	105	B - Moderate Runoff Potential

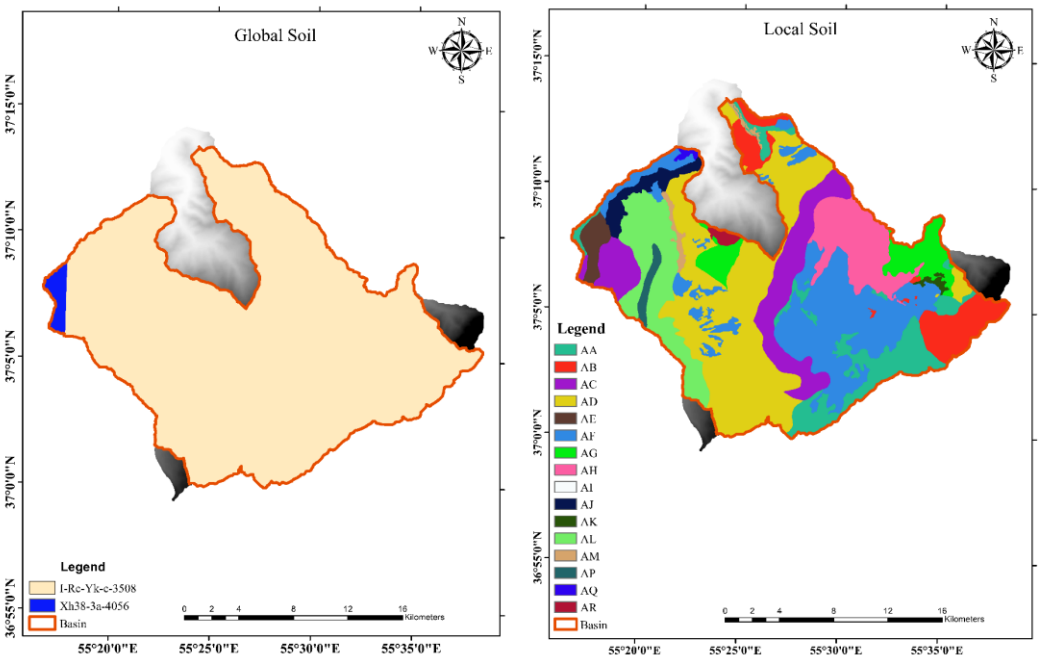

Fig. 2. Global and local soil maps

Table 3. Local and global land use characteristics in SWAT model

Type	Land Use Type in SWAT	Code Used	Percent
Global	MIXED FOREST	FOMI	7.93
	DECIDUOUS BROADLEAF FOREST	FODB	
	SAVANNA	SAVA	
	CROPLAND/GRASSLAND MOSAIC	CRGR	
	IRRIGATED CROPLAND AND PASTURE	CRIR	
	DRYLAND CROPLAND AND PASTURE	CRDY	
	MIXED FOREST	FOMI	
Local	Forest-Mixed	FRST	
	Range-Grasses	RNGE	
	Agricultural Land-Generic	AGRL	
	Wetlands-Non-Forested	WETN	

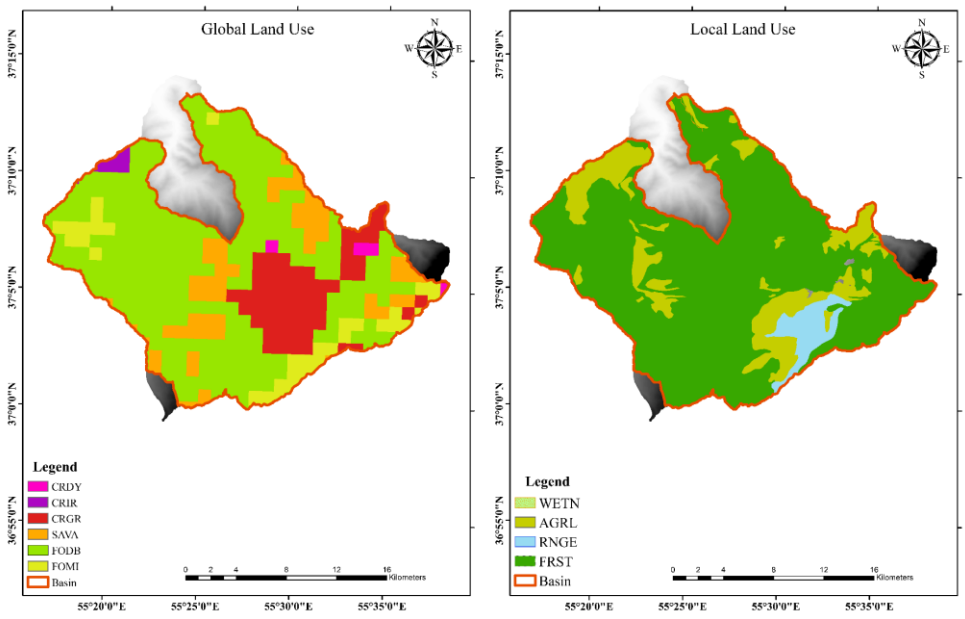

Fig. 3. Global and local land use maps

Essentially, SWAT constitutes the next generation of the SWRRB model. Its development began in the 1980s with modifications to the CREAMS daily rainfall hydrology model. By the late 1980s, further enhancements were incorporated, including the optional USDA Soil Conservation Service (SCS) technology for estimating peak runoff rates and newly developed sediment yield equations. These improvements expanded the model's capability to address water quality management issues in watersheds.

During the 1990s, Arnold and colleagues developed the ROTO model to assess watershed impacts across large areas of Arizona and New Mexico. This model operated by linking outputs from multiple SWRRB simulations and then routing flows through channels and reservoirs in ROTO. While this approach overcame SWRRB's limitation of only allowing 10 sub-watersheds, managing multiple SWRRB input/output files proved cumbersome and computationally intensive.

To resolve these issues, the SWRRB and ROTO models were merged to create SWAT, which retained all SWRRB features while enabling simulation of much larger areas. Since its inception in the 1990s, SWAT has undergone continuous refinement and capability expansion. Key improvements to earlier model versions were described by Arnold and Fohrer (2005) and Neitsch (2005), with theoretical documentation for previous versions provided by Arnold et al. (1998) (Arnold et al., 1993; Arnold and Williams, 1987; Gassman et al., 2004; Izaurralde et al., 2006; Knisel, 1980; Leonard et al., 1987; Williams et al., 1985).

2.4.1. Model structure and functionality

SWAT is a continuous-time, watershed-scale model operating on a daily time step, designed to predict the impacts of land management practices on water, sediment, and agricultural chemical yields in ungauged watersheds. This physically-based, computationally efficient model can simulate long-term continuous periods.

Its core components include weather, hydrology, soil temperature, plant growth, nutrients, pesticides, and land management. Within SWAT, a watershed is divided into multiple sub-watersheds, which are further partitioned into Hydrologic Response Units (HRUs) based on land use, management, and soil characteristics. HRUs represent percentage areas of sub-watersheds and are not spatially identified within a SWAT simulation. Alternatively, a watershed can be divided solely into sub-watersheds characterized by dominant land use, soil type, and management (Gassman et al., 2007).

2.4.2. Hydrological Processes

The hydrological cycle in SWAT incorporates numerous factors influencing water movement and water balance simulation,

including solar radiation and temperature that affect evaporation, snowfall, and snowmelt. These factors are accounted for in the model as critical elements affecting simulation accuracy. Thus, SWAT bases its simulations on the hydrological cycle and water balance equation (Neitsch et al., 2011). The SWAT model equation is expressed as follows:

$$sw_t = sw_o + \sum_{i-1}^{t}\left(R_{\text{day}} - Q_{\text{surf}} - E_a - W_{\text{seep}} - Q_{gw}\right) \quad (1)$$

sw_t: Final soil water content (mm), sw_o: Initial soil water content (mm), R_{day}: Daily rainfall amount (mm water), Q_{surf}: Surface runoff amount on day i (mm), E_a: Evapotranspiration amount on day i (mm), W_{seep}: Amount of water entering from soil profile to unsaturated zone on day i (mm), Q_{gw}: Return flow amount on day i (mm).

2.4.3. SWAT CUP

SWAT-CUP (SWAT Calibration and Uncertainty Programs) is a specialized software tool designed to calibrate and analyze uncertainties in the Soil and Water Assessment Tool (SWAT) model. It provides a user-friendly interface for optimizing model parameters, assessing sensitivity, and quantifying uncertainties in hydrological simulations. The tool incorporates multiple advanced algorithms, including SUFI-2 (Sequential Uncertainty Fitting), PARASOL, PSO (Particle Swarm Optimization), GLUE (Generalized Likelihood Uncertainty Estimation), and MCMC (Markov Chain Monte Carlo), to improve simulation accuracy. SWAT-CUP helps researchers identify critical parameters, refine model inputs, and generate reliable predictions by calculating prediction uncertainty ranges (e.g., 95PPU) and performance metrics like P-factor and R-factor (Zewde et al., 2024).

The software is widely used in hydrological and environmental studies to enhance the reliability of SWAT model outputs for watershed management, climate impact assessments, and pollution control. It supports various calibration techniques, from automated parameter optimization to manual fine-tuning, and works with different versions of SWAT, including SWAT+. By integrating uncertainty analysis directly into the calibration process, SWAT-CUP allows users to evaluate model confidence and improve decision-making in water resource planning. Its continuous updates, such as machine learning-enhanced parameter screening and 3D visualization, ensure it remains a robust tool for researchers and practitioners in hydrology and environmental modeling.

2.5. Model performance evaluation

In order to evaluate the performance of the SWAT model, Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), correlation (R), root mean square error (RMSE), and mean absolute error (MAE) were used as statistical indicators. To perform the calibration and validation process of the SWAT model, the SWAT-CUP software (Soil and Water Assessment Tool: Calibration and Uncertainty Programs) was used.

2.5.1. Nash-Sutcliffe Efficiency (NSE)

The Nash-Sutcliffe efficiency (NSE) ranges from infinite to one and is calculated by comparing the fit of the model to the variation of the observed data. The efficiency is equivalent to a perfect match between the predicted flow and the measured data. It is defined as the minus the sum of the absolute squares of the differences between the predicted and observed values, normalized by the variance of the observed values over the period under consideration. Its formula and various ranges are given in Table (4).

2.5.2. Correlation coefficient (R)

Correlation coefficient (R) measures the strength and direction of a linear relationship between two variables. It is between -1 and +1:

+1: Perfect positive linear relationship (as one variable increases, the other increases).

-1: Perfect negative linear relationship (as one variable increases, the other decreases).

0: There is no linear relationship between the variables. Its formula and different ranges are given in Table (4).

2.5.3. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is another common metric used to evaluate the accuracy of models, especially in regression tasks. Unlike RMSE, which squares the errors before averaging, MAE takes the absolute difference between the predicted and observed values and

then averages them. Its formula is given in Table (4).

2.5.4. Root Mean Square Error (RMSE)

The root mean square error (RMSE) is a metric commonly used to evaluate the accuracy of models, especially in fields such as regression, forecasting, and georeferencing. It measures the average magnitude of the errors between the predicted and observed values. Its formula is given in Table (4).

Table 4. Statistics used for evaluation of model results; Q_{sim} : predicted runoff, Q_{obs} : observed runoff, Q_{mean} : average observed runoff and n : number of data

Equation	Index
$r = \frac{\sum_{i=1}^{n} (Q_{obs} - \bar{Q}_{obs})(Q_{sim} - \bar{Q}_{sim})}{\sqrt{\sum_{i=1}^{n} (Q_{obs} - \bar{Q}_{obs})^2}\sqrt{\sum_{i=1}^{n} (Q_{sim} - \bar{Q}_{sim})^2}}$	Correlation Coefficient
$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} \lvert Q_{obs} - Q_{sim} \rvert$	Mean Absolute Error
$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Q_{sim} - Q_{obs})^2}$	Root Mean Square Error
$\text{NSE} = 1 - \frac{\sum_{i=1}^{n} (Q_{obs} - Q_{sim})^2}{\sum_{i=1}^{n} (Q_{obs} - Q_{mean})^2}$	Nash-Sutcliffe coefficient of Efficiency

Unsatisfactory if		Satisfactory if		Very good if
NSE ≤0.36		0.36 < NSE ≤0.75		0.75< NSE ≤1.00
Negligible	Weak	Moderate	Strong	Very strong
0< r ≤0.3	0.3< r ≤0.5	0.5< r ≤0.7	0.7< r ≤0.9	0.9< r ≤1.00
0< r ≤-0.3	-0.3< r ≤-0.5	-0.5< r ≤-0.7	-0.7< r ≤-0.9	-0.9< r ≤-1.00

3. Results and Discussion

3.1. Model calibration and validation

For model calibration and validation, SWAT-CUP software and the SUFI-2 algorithm were employed. Observed data from the Lazoureh and Jangaldeh stations were used for calibration (2006–2020) and validation (1997–2005). The Nash-Sutcliffe Efficiency (NS) index was adopted as the objective function for model optimization. Additionally, the uncertainty of the model simulation results was estimated within a 95% confidence interval. To evaluate model performance during calibration and validation, the following metrics were applied: Nash-Sutcliffe Efficiency (NS), Correlation Coefficient (R), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

3.2. Simulation of river flow using global data for lazoureh and jangaldeh stations

The SWAT model was used to simulate river flow by utilizing global and regional land use and soil data. The global data included six land use categories and two soil types. For hydrological flow analysis, time-series and violin plots were used for the calibration (2006–2020) and validation (1997–2005) periods at the Lazoureh and Jangaldeh stations.

Figures 4 and 5 present the time-series and violin plots of observed and simulated flow for these stations. The results indicate that the SWAT model has a suitable capability in simulating the overall trend of flow variations, particularly during wet months. However, in dry months, a greater discrepancy is observed between the observed and simulated data, which is likely due to input data errors, limited accuracy of model parameters, or unexpected hydrological events. The Nash-Sutcliffe Efficiency (NS) index for the Lazoureh station was calculated as 0.60 and 0.42 for the calibration and validation periods, respectively, while for the Jangaldeh station, it was 0.58 and 0.41, indicating the model's acceptable performance. The violin plots, which combine box plots and scatter plots, illustrated the distribution and comparison of observed and simulated data.

For the Lazoureh station, the highest dispersion of observed data during the calibration period was between 0.5 and 2.25 (m^3/s), while the simulated data ranged from 0.75 to 2.25 (m^3/s). During the validation period, this range was 0.75 to 2.3 (m^3/s) for observed data and 0.75 to 2.2 (m^3/s) for simulated data. For the Jangaldeh station, the highest dispersion of observed data during the

calibration period was between 0 and 2.5 (m^3/s), while the simulated data ranged from 0 to 1.5 (m^3/s). In the validation period, this range was 0.4 to 2.7 (m^3/s) for observed data and 0.2 to 2.0 (m^3/s) for simulated data.

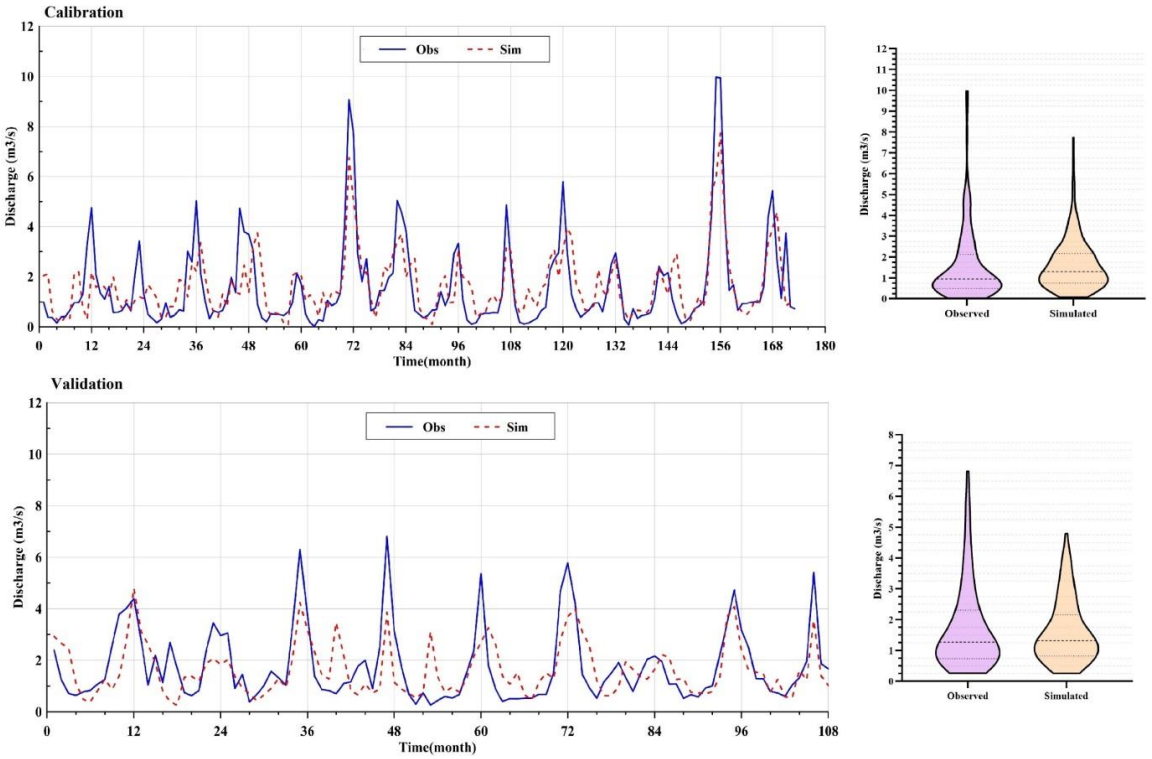

Fig. 4. Time-series and violin plots of observed versus simulated streamflow during calibration (2006-2020) and validation (1997-2005) periods at the Lazoureh station, using global input datasets

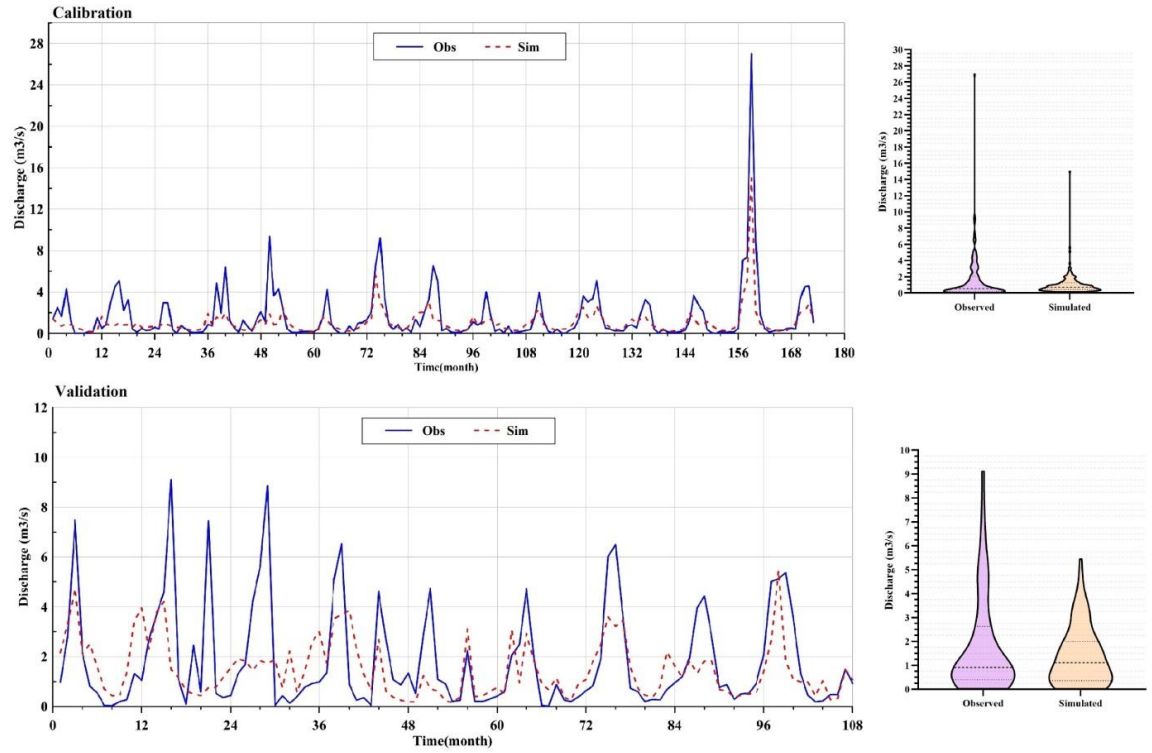

Fig. 5. Time-series and violin plots of observed versus simulated streamflow during calibration (2006-2020) and validation (1997-2005) periods at the Jangaldeh station, using global input datasets

Table 5. Key calibration parameters for global-data simulations at Lazoure station

Parameter	Range		Description
	Min	Max	
*r_CN2.mgt	-0.3	0.3	Curve Number
r_HRU_SLP.hru	-0.2	0.2	Average Slope Gradient
R__LAT_TTIME.hru	-0.2	0.2	Lateral Flow Travel Time
R__SLSOIL.hru	-0.5	-0.1	Slope Length for Lateral Subsurface Flow
v_SLSUBBSN.hru	-0.2	0.2	Average Slope Length
r_SOL_AWC.sol	-0.4	0.4	Available Water Capacity
r_SOL_BD.sol	-0.4	0.4	Bulk Density of Soil
r_SOL_K.sol	-0.4	0.4	Saturated Hydraulic Conductivity of Soil
R__SUB_SFTMP (..).sno	8	12	Snowfall Temperature
R__SUB_SMFMN (..).sno	5	8	Snowmelt Degree-Day Factor on December 21
R__SUB_SMFMX (..).sno	2	4	Snowmelt Degree-Day Factor on June 21
R__SUB_SMTMP (..).sno	0	1	Snowmelt Base Temperature
R__SUB_TIMP (..).sno	8	10	Snowpack Temperature Lag Factor
v_ALPHA_BF.gw	0	0.5	Baseflow Recession Constant
V__ALPHA_BNK.rte	0.8	0.1	Baseflow Alpha Factor for Bank Storage
v_CH_N2.rte	-0.1	0.1	Manning's Roughness Coefficient for Main Channel Flow
v_EPCO.hru	0	1	Plant Uptake Compensation Factor
v_ESCO.hru	0	1	Soil Evaporation Compensation Factor
v_GW_DELAY.gw	15	20	Groundwater Recharge Delay Time
v_GW_REVAP.gw	0.02	0.2	Return Flow from Shallow Aquifer
v_GWQMN.gw	1200	1300	Threshold Water Depth in Shallow Aquifer for Return Flow
v_OV_N.hru	0.01	0.5	Manning's Roughness Coefficient for Overland Flow
V__PLAPS.sub	-60	-45	Rainfall Infiltration Rate
v_RCHRG_DP.gw	-0.1	0.3	Aquifer Percolation Coefficient
v_REVAPMN.gw	400	500	Threshold Water Depth in Shallow Aquifer for Deep Aquifer Percolation
v_SFTMP.bsn	1.8	2.5	Snowfall Temperature
v_SMTMP.bsn	1.86	2.5	Snowmelt Base Temperature
v_SURLAG.bsn	13	15	Surface Runoff Lag Time
V__TLAPS.sub	9.5	11	Temperature Lapse Rate

Table 6. Key calibration parameters for global-data simulations at Jangaldeh station

Parameter	Range		Description
	Min	Max	
*R__CN2.mgt	-0.24	0.19	Curve Number
R__HRU_SLP.hru	-0.13	0.36	Slope Gradient
R__PCPMM (..).wgn	-0.21	0.38	Average Monthly Precipitation
R__SOL_AWC (..).sol	-0.04	0.89	Available Water Capacity
R__SOL_BD (..).sol	-0.17	0.29	Soil Bulk Density
R__SOL_K (..).sol	-0.05	1.45	Saturated Soil Hydraulic Conductivity
R__SUB_SFTMP (..).sno	8.17	10.72	Snowfall Temperature
R__SUB_SMFMN (..).sno	6.28	8.83	Snowmelt Degree-Day Factor on December 21
R__SUB_SMFMX (..).sno	2.32	3.44	Snowmelt Degree-Day Factor on June 21
R__SUB_SMTMP (..).sno	0.05	0.68	Base Snowmelt Temperature
R__SUB_TIMP (..).sno	8.90	10.72	Snowpack Temperature Delay Factor
V__ALPHA_BF.gw	0.47	1.42	Base Flow Decline Constant
V__ALPHA_BNK.rte	0.24	0.75	Alpha Base Flow Coefficient for Bank Storage
V__CH_K2.rte	2.16	100.74	Effective Hydraulic Conductivity in Main Channel
V__EPCO.hru	0.05	0.68	Vegetation Infiltration Compensation Factor
V__ESCO.hru	-0.12	0.63	Soil Evaporation Compensation Factor
V__GW_DELAY.gw	-23.44	32.20	Aquifer Recharge Delay Time
V__LAT_TTIME.hru	-54.11	101.99	Lateral Flow Travel Time
V__OV_N.hru	0.31	0.94	Manning's Roughness Coefficient for Surface Flow
V__RCHRG_DP.gw	0.05	0.68	Aquifer Infiltration Coefficient
V__REVAPMN.gw	78.48	235.52	Shallow Aquifer Threshold Depth for Deep Aquifer Infiltration
V__SFTMP.bsn	-4.68	1.78	Snowfall Temperature
V__SLSUBBSN.hru	45.34	116.06	Average Slope Length
V__SMFMN.bsn	3.49	10.49	Snowmelt Factor on December 21
V__SMTMP.bsn	-8.34	0.56	Base Snowmelt Temperature
V__TIMP.bsn	0.38	1.15	Snowpack Temperature Delay Factor

The results of this analysis demonstrate that the SWAT model successfully established a suitable agreement between observed and simulated data in both calibration and validation phases, providing a more accurate estimation of hydrological data. In this study,

the examined parameters and their variation ranges during the model calibration phase using global datasets for the studied stations are presented in Tables (5 and 6). To select the optimal ranges, 2000 simulations were performed to determine the best parameters and evaluate model uncertainties.

3.3. River Flow Analysis Using Regional Data at Lazoureh and Jangaldeh Stations

As mentioned earlier, the regional data used in this analysis included 15 soil types with codes AA to AR and four land-use classifications. Figures 6 and 7 present the time-series and violin plots of observed and simulated flow for these stations.

The evaluation of the SWAT model's performance showed that the Nash-Sutcliffe Efficiency (NSE) coefficient for Lazoureh station during calibration and validation periods was 0.60 and 0.44, respectively, and for Jangaldeh station was 0.73 and 0.51, respectively. These values confirm the model's appropriate performance in simulating hydrological flow using regional data. For more detailed analysis, violin plots were used to examine the dispersion of observed and simulated data. At Lazoureh station, the mean dispersion of observed and simulated data during the calibration period was estimated at approximately 1 and 1.25 (m^3/s), respectively. During the validation period, this value decreased to about 1.25 (m^3/s) for both data types.

At Jangaldeh station, the mean data dispersion during the calibration period was about 0.5 (m^3/s) for both observed and simulated data, while during the validation period these values increased to 1 and 1.2 (m^3/s), respectively. The SWAT model was able to establish significant agreement between observed and simulated data in both calibration and validation phases. These results confirm the model's efficiency in using regional data for hydrological flow simulation.

In this study, the examined parameters and their variation ranges during the model calibration phase using regional data for the studied stations are presented in Tables 7 and 8. To select the optimal ranges, 2000 simulations were performed to choose the best parameters and evaluate model uncertainties.

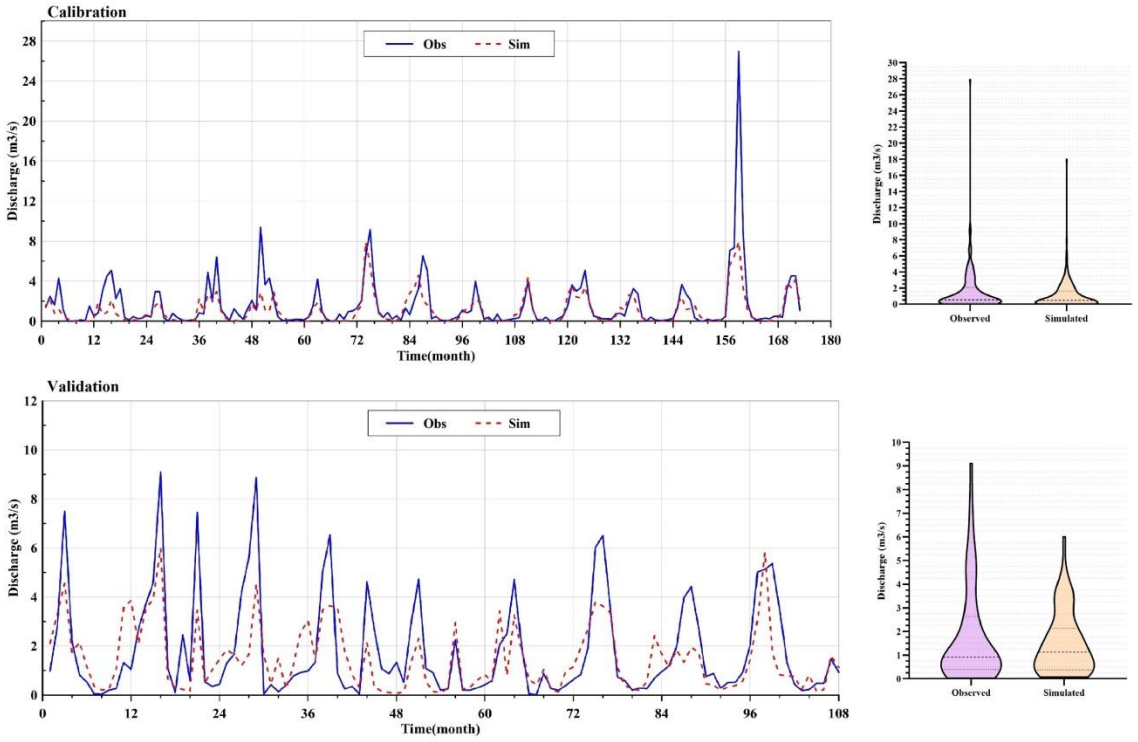

Fig. 6. Time-series and violin plots of observed versus simulated streamflow during calibration (2006-2020) and validation (1997-2005) periods at the Lazoureh station, using local input datasets

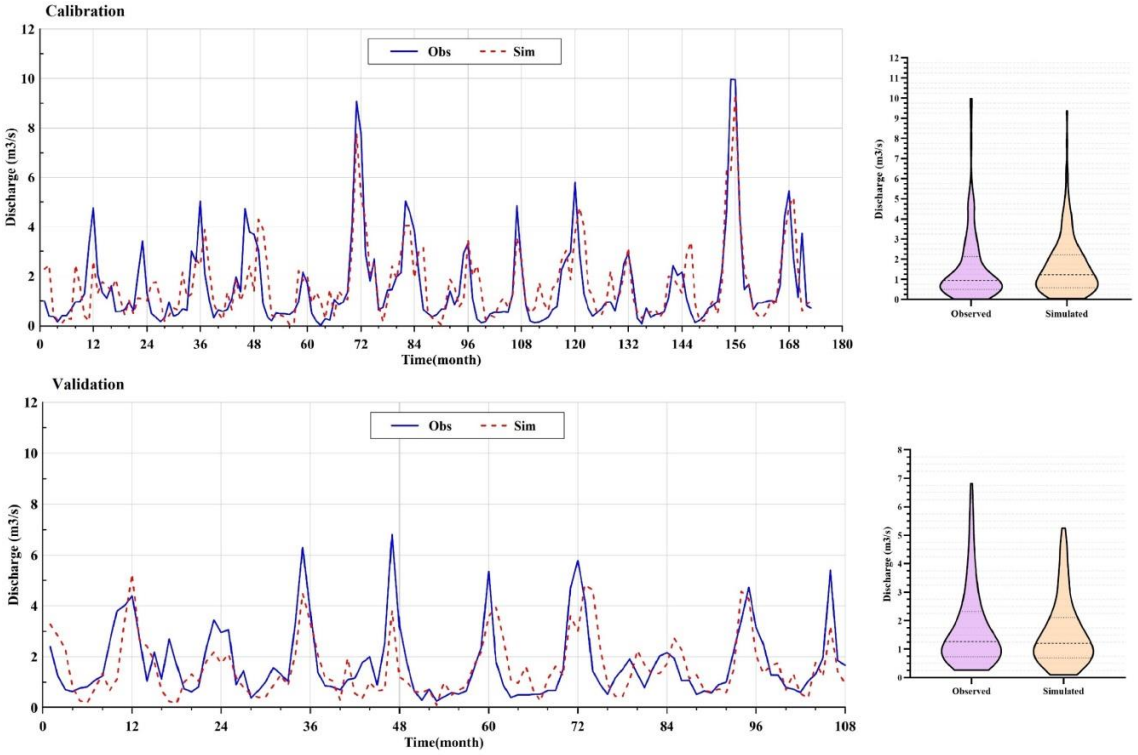

Fig. 7. Time-series and violin plots of observed versus simulated streamflow during calibration (2006-2020) and validation (1997-2005) periods at the Jangaldeh station, using local input datasets

Table 7. Key calibration parameters for local-data simulations at Lazoure station

Parameter	Range		Description
	Min	Max	
*r_CN2.mgt	-0.3	0.3	Curve Number
r_HRU_SLP.hru	-0.2	0.2	Average Slope Gradient
R__LAT_TTIME.hru	-0.2	0.2	Lateral Flow Travel Time
R__SLSOIL.hru	-0.5	-0.1	Slope Length for Lateral Subsurface Flow
v_SLSUBBSN.hru	-0.2	0.2	Average Slope Length
r_SOL_AWC.sol	-0.4	0.4	Available Water Capacity
r_SOL_BD.sol	-0.4	0.4	Bulk Density of Soil
r_SOL_K.sol	-0.4	0.4	Saturated Hydraulic Conductivity of Soil
R__SUB_SFTMP (..).sno	8	12	Snowfall Temperature
R__SUB_SMFMN (..).sno	5	8	Snowmelt Degree-Day Factor on December 21
R__SUB_SMFMX (..).sno	2	4	Snowmelt Degree-Day Factor on June 21
R__SUB_SMTMP (..).sno	0	1	Snowmelt Base Temperature
R__SUB_TIMP (..).sno	8	10	Snowpack Temperature Lag Factor
v_ALPHA_BF.gw	0	0.5	Baseflow Recession Constant
V__ALPHA_BNK.rte	0.8	0.1	Baseflow Alpha Factor for Bank Storage
v_CH_N2.rte	-0.1	0.1	Manning's Roughness Coefficient for Main Channel Flow
v_EPCO.hru	0	1	Plant Uptake Compensation Factor
v_ESCO.hru	0	1	Soil Evaporation Compensation Factor
v_GW_DELAY.gw	15	20	Groundwater Recharge Delay Time
v_GW_REVAP.gw	0.02	0.2	Return Flow from Shallow Aquifer
v_GWQMN.gw	1200	1300	Threshold Water Depth in Shallow Aquifer for Return Flow
v_OV_N.hru	0.01	0.5	Manning's Roughness Coefficient for Overland Flow
V__PLAPS.sub	-60	-45	Rainfall Infiltration Rate
v_RCHRG_DP.gw	-0.1	0.3	Aquifer Percolation Coefficient
v_REVAPMN.gw	400	500	Threshold Water Depth in Shallow Aquifer for Deep Aquifer Percolation
v_SFTMP.bsn	1.8	2.5	Snowfall Temperature
v_SMTMP.bsn	1.86	2.5	Snowmelt Base Temperature
v_SURLAG.bsn	13	15	Surface Runoff Lag Time
V__TLAPS.sub	9.5	11	Temperature Lapse Rate

Table 8. Key calibration parameters for local-data simulations at Jangaldeh station

Parameter	Range		Description
	Min	Max	
*R__CN2.mgt	-0.24	0.19	Curve Number
R__HRU_SLP.hru	-0.13	0.36	Slope Gradient
R__PCPMM (..).wgn	-0.21	0.38	Average Monthly Precipitation
R__SOL_AWC (..).sol	-0.04	0.89	Available Water Capacity
R__SOL_BD (..).sol	-0.17	0.29	Soil Bulk Density
R__SOL_K (..).sol	-0.05	1.45	Saturated Soil Hydraulic Conductivity
R__SUB_SFTMP (..).sno	8.17	10.72	Snowfall Temperature
R__SUB_SMFMN (..).sno	6.28	8.83	Snowmelt Degree-Day Factor on December 21
R__SUB_SMFMX (..).sno	2.32	3.44	Snowmelt Degree-Day Factor on June 21
R__SUB_SMTMP (..).sno	0.05	0.68	Base Snowmelt Temperature
R__SUB_TIMP (..).sno	8.90	10.72	Snowpack Temperature Delay Factor
V__ALPHA_BF.gw	0.47	1.42	Base Flow Decline Constant
V__ALPHA_BNK.rte	0.24	0.75	Alpha Base Flow Coefficient for Bank Storage
V__CH_K2.rte	2.16	100.74	Effective Hydraulic Conductivity in Main Channel
V__EPCO.hru	0.05	0.68	Vegetation Infiltration Compensation Factor
V__ESCO.hru	-0.12	0.63	Soil Evaporation Compensation Factor
V__GW_DELAY.gw	-23.44	32.20	Aquifer Recharge Delay Time
V__LAT_TTIME.hru	-54.11	101.99	Lateral Flow Travel Time
V__OV_N.hru	0.31	0.94	Manning's Roughness Coefficient for Surface Flow
V__RCHRG_DP.gw	0.05	0.68	Aquifer Infiltration Coefficient
V__REVAPMN.gw	78.48	235.52	Shallow Aquifer Threshold Depth for Deep Aquifer Infiltration
V__SFTMP.bsn	-4.68	1.78	Snowfall Temperature
V__SLSUBBSN.hru	45.34	116.06	Average Slope Length
V__SMFMN.bsn	3.49	10.49	Snowmelt Factor on December 21
V__SMTMP.bsn	-8.34	0.56	Base Snowmelt Temperature
V__TIMP.bsn	0.38	1.15	Snowpack Temperature Delay Factor

4. Conclusion

In this study, the performance of the SWAT model in simulating discharge at various stations within the basin was evaluated, and the effect of data type selection on the objective function was analyzed. The main objective of this research was to simulate and more accurately estimate flow discharge in the watersheds of Lazoureh and Jangaldeh stations using both global and regional data.

In the calibration phase for Lazoureh station, both regional and global data produced nearly similar results. The error indices MAE and RMSE were 0.78 and 1.11 respectively for both datasets, with no significant difference observed. However, during the validation phase, regional data showed better performance, reducing MAE error by 2.56% and RMSE error by 1.92%.

At Jangaldeh station, regional data outperformed global data in both calibration and validation phases. During calibration, the accuracy indices NS and R for regional data were higher at 0.73 and 0.90 respectively, while the percentage change in MAE and RMSE errors for global data was 25.68% and 22.92% higher than regional data.

Furthermore, in the validation phase, regional data maintained their superiority with lower errors and higher accuracy, where MAE and RMSE errors for global data were 12% and 10.34% higher than regional data respectively.

Overall, this study shows that the use of local data, especially in the validation phase, can reduce modeling errors and provide more accurate representation of streamflow. Under conditions of constrained local data availability, global datasets offer a feasible alternative. Despite potential compromises in accuracy, they can still yield informative and actionable analyses. Therefore, the choice between local and global data should be made

according to data availability and required accuracy. This study emphasizes that both types of data can be effectively used in hydrological modeling, depending on the needs and conditions.

5. Conflict of Interest

No potential conflict of interest was reported by the authors

6. References

Abou Rafee, S. A., Uvo, C. B., Martins, J. A., Domingues, L. M ,.Rudke, A. P., Fujita, T., & Freitas, E. D. (2019). Large-scale hydrological modelling of the Upper Paraná River Basin. *Water, 11*(5), 882.

Arnold, J. G., & Fohrer, N. (2005). SWAT2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes: An International Journal, 19*(3), 563-572.

Arnold, J. G., & Williams, J. R. (1987). Validation of SWRRB—simulator for water resources in rural basins. *Journal of Water Resources Planning and Management, 113*(2), 243-256.

Arnold, J. G., Allen, P. M., & Bernhardt, G. (1993). A comprehensive surface-groundwater flow model. *Journal of Hydrology, 142*(1-4), 47-69.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment part I: model development 1. *JAWRA Journal of the American Water Resources Association, 34*(1), 73-89.

DHI. (2004). The MIKE SHE User Reference. Horsholm.

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*(4), 1211-1250.

Gassman, P. W., Williams, J. R., Benson, V. W., Izaurralde, R. C., Hauck, L. M., Jones, C. A., ... & Flowers, J. D. (2004). Historical development and applications of the EPIC and APEX models. In *2004 ASAE Annual Meeting* (p. 1). American Society of Agricultural and Biological Engineers.

Goodarzi, M. R., Zahabiyoun, B., Massah Bavani, A. R., & Kamal, A. R. (2012). Performance comparison of three hydrological models SWAT, IHACRES and SIMHYD for the runoff simulation of Gharesou basin. *Water and Irrigation Management*, *2*(1), 25-40.

Holisagar, P., Kambar, M., Pattepur, M., & Ahmed, M. F. (2023). prediction of streamflow in ungauged catchment using hydrological modelling. *International Research Journal of Education and Technology, 5*(6): 1-7.

Izaurralde, R., Williams, J. R., McGill, W. B., Rosenberg, N. J., & Jakas, M. Q. (2006). Simulating soil C dynamics with EPIC: Model description and testing against long-term data. *Ecological Modelling, 192*(3-4), 362-384.

Knisel, W. G. (1980). CREAMS: A field scale model for chemicals, runoff, and erosion from agricultural management systems. Department of Agriculture, Science and Education Administration.

Leonard, R. A., Knisel, W. G., & Still, D. A. (1987). GLEAMS: Groundwater loading effects of agricultural management systems. *Transactions of the ASAE, 30*(5), 1403-1418.

Lillesand, T. M., Kiefer, R. W., Dulbahri, Suharsono, P., Hartono, Suharyadi, & Sutanto. (1993). *Penginderaan jauh dan interpretasi citra*. Gadjah Mada University.

Liu, W., Wu, J., Xu, F., Mu, D., & Zhang ,P. (2024). Modeling the effects of land use/land cover changes on river runoff using SWAT models: A case study of the Danjiang River source area, China. *Environmental Research, 242*, 117810.

Machado, R., Sentelhas, P., Leite, R., & Paulino, J. (2023). Impacts of Land Use Change and Climate Variability on Streamflow in the Piracicaba Basin, Brazil. https://doi.org/10.21203/rs.3.rs-3136986/v1

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology, 10*(3), 282-290.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Williams, J. R., & King, K. W. (2005, January). *Soil and water assessment tool: theoretical documentation: version 2005*.

Nie, W., Yuan, Y., Kepner, W., Nash, M. S., Jackson, M., & Erickson, C. (2011). Assessing impacts of Landuse and Landcover changes on hydrology for the upper San Pedro watershed. *Journal of Hydrology, 407*(1-4), 105-114.

Peterson, J., & Hamlett, J. (1998). Hydrologic calibration of the Swat model in a watershed containing fragipan soils 1. *JAWRA Journal of the American Water Resources Association, 34*(3), 531-544.

Shukla, S., Meshesha, T. W., Sen, I. S., Bol, R., Bogena, H., & Wang, J. (2023). Assessing Impacts of Land Use and Land Cover (LULC) Change on Stream Flow and Runoff in Rur Basin, Germany. *Sustainability, 15*(12), 9811.

Susiwidiyaliza, U. H., SETIAWAN, F., & UTAMI, N. (2023). The Impact of Land-Use and Climate Change on Water and Sediment Yields in Batanghari Watershed, Sumatra, Indonesia. *Sains Malaysiana, 52*(3), 705-721.

Tullu, K. T. (2023). Response of Climate Change Impact on Streamflow in Sululta Catchment, Abay Basin, Ethiopia. *International Journal of Earth Sciences Knowledge and Applications, 5*(1), 79-91.

Wagena, M. B., Bock, E. M., Sommerlot, A. R., Fuka, D. R., & Easton, Z. M. (2017). Development of a nitrous oxide routine for the SWAT model to assess greenhouse gas emissions from agroecosystems. *Environmental modelling & software, 89*, 131-143 .

Wang, Z.-M., Batelaan, O., & De Smedt, F. (1996). A distributed model for water and energy transfer between soil, plants and atmosphere (WetSpa). *Physics and Chemistry of the Earth, 21*(3), 189-193.

Williams, J. R., Nicks, A., & Arnold, J. G. (1985). Simulator for water resources in rural basins. *Journal of Hydraulic Engineering, 111*(6), 970-986 .

Xiao, F., Wang, X., & Fu, C. (2023). Impacts of land use/land cover and climate change on hydrological cycle in the Xiaoxingkai Lake Basin. *Journal of Hydrology: Regional Studies, 47*, 101422.

Zewde, N. T., Denboba, M. A., Tadesse, S. A., & Getahun, Y. S. (2024). Predicting runoff and sediment yields using soil and water assessment tool (SWAT) model in the Jemma Subbasin of Upper Blue Nile, Central Ethiopia. *Environmental Challenges*, *14*, 100806.