Journal of Hunan University

(Natural Sciences)

Vol. 51 No. 12

December 2024

Available online at

http://jonuns.com

Open Access Article https://doi.org/10.55463/issn.1674-2974.51.12.6

Development of Watershed Characteristics-Based Synthetic Unit Hydrograph

Abdul Ghoni Majdi [1*], Suhardjono[2], Lily Montarcih Limantara[2*], Ussy Andawayanti[2]

[1]Doctoral Student, Department of Water Resources Engineering, Faculty of Engineering, University of Brawijaya, Malang, Indonesia

[2]Department of Water Resources Engineering, Faculty of Engineering, University of Brawijaya, Malang, Indonesia

* Corresponding authors: lilymont@ub.ac.id, onny_majdi@yahoo.com

Article History:

Received: October 25, 2024
Revised: November 22, 2024
Accepted: December 15, 2024
Published: December 30, 2024

Abstract: This paper intends to build a Synthetic Unit Hydrograph based on watershed characteristics, not from rainfall and runoff data. The geomorphology parameter from the synthetic unit hydrograph model is the most useful approach for predicting run-off and the simplest method for understanding the different hydrology behaviors of watersheds, mainly in ungauged watersheds or those with a lack of data. This research was conducted in 10 sub-watersheds in Indonesia that were completed using an automatic water level recorder (AWLR) and an automatic rainfall recorder (ARR). The methodology consists of data collection of watershed characteristic parameter hydrology, feature selection using genetic algorithm (GA), processing the initial data, assessment of SVR model by optimal subset model, hyperparameter tuning using particle swarm optimization (PSO), building Model of Support Vector Regression (SVR), and evaluation model. The results are as follows: 1) formulation of peak discharge: $f(X) = 1.1627\ X_1 + (-0.5060\ X_2) + (-0.0315\ X_3) + (-0.1493\ X_4 + 0.0775\ X_5 + (-0.0340\ X_6)$; 2) formulation of time to peak: $f(X) = (-0.8013)\ X_1 + 0.9282\ X_2 + 0.0162\ X_3 + 0.5461\ X_4$; and 3)formulation of time base: $f(X) = 0.1771\ X_1 + 0.3323\ X_2 + 0.1128\ X_3 + 0.7149\ X_4$; explanation: X_1 = watershed area; X_2 = river length; X_3 = river slope; X_4 = river drainage density; X_5 = watershed shape; and X_6 = run-off coefficient.

Keywords: Synthetic unit hydrograph, model, watershed characteristic, peak discharge, time to peak, time base

基于流域特征的合成单位线开发

摘要：本研究旨在根据流域特征而非降雨和径流数据构建合成单位线。合成单位线模型中的地貌参数是预测径流最有用的方法，也是了解流域不同水文行为的最简单方法，主要用于未测量或缺乏数据的流域。本研究在印度尼西亚的 10 个子流域进行，这些子流域使用自动水位记录器（AWLR）和自动降雨记录器（ARR）完成。该方法包括流域特征参数水文数据收集、使用遗传算法（GA）进行特征选择、处理初始数据、通过最佳子集模型评估 SVR 模型、使用粒子群优化（PSO）进行超参数调整、建立支持向量回归（SVR）模型和评估模型。研究结果如下：1）峰值流量计算公式：f(X) = 1.1627 X1 + (-0.5060 X2) + (-0.0315 X3) + (-0.1493 X4 + 0.0775 X5 + (-0.0340 X6)；2）峰值时间计算公式：f(X) = (-0.8013) X1 + 0.9282 X2 + 0.0162 X3 + 0.5461 X4；3）时间基准计算公式：f(X) = 0.1771 X1 + 0.3323 X2 + 0.1128 X3 + 0.7149 X4；说明：X1 = 流域面积；X2 = 河流长度；X3 = 河流坡度；X4 = 河流排水密度；X5 = 流域形状；X6 = 径流系数

关键词：合成单位线，模型，流域特性，峰值流量，峰值时间，时间基准

1. Introduction

There are many constraints of data availability [1], so the synthetic unit hydrograph (SUH) based on the watershed characteristics is needed for expanding unit hydrograph theory in an ungauged watershed [2-3].

The concept of a synthetic unit hydrograph can be implemented based on the watershed characteristics, but not from rainfall and runoff data [4]. However, the geomorphology parameter from SUH is as the most useful approach for predicting runoff and as the simplest method for understanding the different hydrology behaviors of watersheds, mainly in ungauged watersheds or those with a lack of data [3, 5-6].

The morphometry parameters of watersheds have been studied by researchers using conventional methods, remote sensing methods, and geographical information systems (GIS) [2, 7-10]. The ratios between the time rise of hydrograph, duration, peak discharge, and volume of runoff with the watershed physical parameters like area, shape, slope, river network, and combination of the parameters were developed in [11].

Hydrological models are crucial for designing, developing, and managing water resources. However, geo-spatial technology that is rapidly developing is as follows: Geographical Information System (GIS, Digital Elevation Model (DEM), and Global Positioning System (GPS). These tools are used as effective and popular tools for solving the problem of water resource design and management due to the use of conventional methods in data processing [9, 12-15]. However, it can give the newest data in some scales [14], and calibrating it is very constructive in producing the model parameters spatially [16-18].

For some hydrology modeling scenarios, it is possible to use various watershed attributes such as elevation, slope, and land cover. The machine learning (ML) method significantly increased the prediction results in the watershed with limited data, mainly when the model parameter was optimized using calibrated watershed data. The ML method is also able to catch complex patterns in discharge data without explicit assumptions about the relation of input-output to make it a flexible solution for various scenarios.

Support Vector Regression (SVR) is a method in Machine Learning (ML) that can be used for regression analysis. SVR is often used in modeling hydrology data because of its reliability in handling complex datasets and its ability to provide good generalization even when the available data are limited and have a high dimensionality. SVR can build a nonlinear relation model between input and output using a kernel function that maps the data into high-dimensional space. It is possible for SVR to catch the complex pattern in hydrology data like the relation between rainfall and runoff, although the pattern is difficult to understand with the linear conventional statistical model. SVR is designed to work with a limited amount of data like there is often found in hydrology modeling, especially in ungauged watersheds. SVR can capture the nonlinear relationship between the watershed attribute and model parameter with small assessment data.

The traditional SUH often depends on some assumptions or limited historical data; however, the

SVR is possible to handle more complex data like climate variable that is changed or unstable hydrology conditions. The SVR can be accordance with more heterogeneous data and increase the prediction accuracy of peak discharge and time to peak, which are the focus of the SUH model.

Generally, many SUH models have been developed in many countries, including Indonesia, by using watershed morphometry parameters and to combine with the other parameters. According to [19], the Snyder, Mockus, and Soil Conservation Service (SCS) models are most often used SUHs. As stated in [20], besides Snyder, SCS, and Nakayasu SUH models, Gama I, Limantara, ABG, ITB-1 and ITB-2 synthetic hydrographs are generally used in Indonesia.

In some studies, the SUH methods still have limitations, mainly if they are applied in the watershed outside of the study locations because they often produce a high enough deviation of peak discharge. The performance of the SUHs is varied in a watershed to other watersheds, and there is no method that can work best in every watershed. The main factor that influences the accuracy of the SUH model is the accuracy in determining the watershed characteristics and the other parameters that are used in the method, so the SUH model can be close to the observed unit hydrograph.

This research is conducted using a new methodology as an alternative approach to SUH modeling that is expected to represent regional conditions in accordance with watershed characteristics in Indonesia. The development of SUH is carried out based on the watershed characteristics, which include watershed morphometry factors like watershed physical parameters, which are static (relatively not changed), and the parameters of the watershed land cover, which are dynamic (can be changed). To obtain accurate parameters of watershed characteristics and the reliability of the SUH model, and to consider that the hydrology process is complex and nonlinear, DEM analysis and SVR regression method so it is close to the real condition.

2. Materials and Methods

This research was carried out by collecting data on AWLR and ARR and watershed characteristics at 12 locations in the watersheds/sub-watersheds. The rainfall and AWLR data were analyzed to produce the observed unit hydrograph and peak discharge. However, the data on watershed characteristics are processed to produce the dominant watershed parameters. Based on the observed unit hydrograph, peak discharge, and dominant watershed characteristic parameters, a synthetic unit hydrograph was formulated. Then, the model was validated on the serial data of AWLR that are used for building the model. Figure 1 presents the research locations.

Figure 1. Research locations (by the authors)

The formulation of SUH in this research is carried out in parts of Indonesian watersheds with suitable availability of data as follows: 3 watersheds/sub-watersheds in Nusa Tenggara Barat, 4 watersheds/sub-watersheds in Jawa Timur, and 5 watersheds/sub-watersheds in Daerah Istimewa Yogyakarta with the limitations as follows:

(1) From the 12 watersheds, 10 were used to build the model and the others (2 watersheds) will be used for verification.

(2) The selected watersheds have an automatic rainfall recorder (ARR) and an automatic water level recorder (AWLR).

(3) Watershed area < 5000 km^2 (small watershed), remembering that watershed with the area > 5000 km^2 will not be influenced by a homogeneous intensity of rainfall.

(4) The observed unit hydrograph from AWLR was selected as the isolated one with a single peak and sufficient hourly rainfall distribution.

(5) Watershed characterization consists of morphometry parameters, land cover and type of soil that are obtained from spatial analysis results using data from DEM (2024), RBI map (2019), land cover map of KLHK (Kementerian Lingkungan Hidup dan Kehutanan, 2017 and 2019) and map of soil type from FAO (2024);

(6) The model formulation assumptions are as follows:

a) Separation of the direct runoff from the base flow in the hydrograph using the straight-line method

b) Effective rainfall was analyzed using the phi (Φ) index.

c) The observed unit hydrograph is differentiated from the observed hydrograph using the Collins method.

(7) Run-off coefficient or CN (curve number) in the condition of AMC II (Antecedent Moisture Condition II) that is average soil moisture condition.

(8) Model formulation of peak discharge, time to peak, and time base using the regression analysis of the SVR (support vector regression)

2.1. Observed Unit Hydrograph

Observed unit hydrograph was analyzed from 10 events in each watershed by using Collins Method with the analysis steps are as follows [1]:

1. The stage hydrograph is transformed into a discharge hydrograph by calibration.
2. The base flow is separated from the hydrograph by using one of empirical Methods such as Straight-Line Method.
3. The effective rainfall that causes a flood is analyzed using the Phi Index.
4. The trial unit hydrograph is determined by determining the ordinates with certain dimensions.
5. The initial unit hydrograph (trial) was multiplied by all effective rainfalls, with the exception of the largest effective rainfall.
6. The direct run-off hydrograph above is reduced by the measured direct run-off hydrograph; there is obtained the direct run-off hydrograph caused by the maximum rainfall, and then there is obtained the second unit hydrograph (trial).
7. The second unit hydrograph is compared with the initial unit hydrograph. If there is still a large difference (in accordance with the standard error that is determined), then the difference is returned in the fifth and sixth stages based on the last unit hydrograph.
8. Thus, it should be proceeded until the smallest possible difference between the last and previous unit hydrographs is obtained.
9. To produce the observed unit hydrograph for the whole watersheds, it is carried out with to average the ordinate of the unit hydrograph in the same hour, peak discharge, and time to peak with the steps as follows:

a) To calculate the average time to peak and discharge.

b) The dimensionless observed unit hydrographs (t/T_p and Q/Q_p) for each watershed.

c) To calculate the average of the dimensionless observed unit hydrograph.

d) To calculate the average of observed unit hydrograph.

2.2. The Formulation of the SUH Model

The formulation of the SUH model consists of the peak discharge model (Q_p), time-to-peak model (T_p), and time-based model (T_b), which is based on the 6 watershed characteristics, namely, watershed area (A), main river length (L), main river slope (S), watershed shape factor (F_b), drainage density (D_d), and runoff coefficient/curve number (CN). The machine learning of the SVR linear regression method is carried out as an approach of hydrology process that is complex and non-linear to obtain the SUH model, which is simple, accurate, and applicable for ungauged watersheds in Indonesia. However, the steps for developing the SUH model using SVR are as follows:

1. To collect data of watershed characteristic parameters and hydrology, such as watershed area (A), main river length (L), river slope (S), river drainage density (D_d), watershed shape factor (R_c), and run-off coefficient/curve number (CN), we input the independent variable, and the output variables are peak discharge (Q_p), time to peak (T_p), and time base (T_b).
2. To select features using Genetic Algorithm (GA).

The genetic algorithm (GA) is used to select the best subset feature/variable in the model. Feature selection is used for selecting the best input variable (watershed data) that will be used by SVR model (development of SUH model).

3. To pre-process the data.

Normalization/standardization of data: remembering that SVR is sensitive to the feature scale, the normalization is needed so the input variables have the same scale. In this study, we used the min-max scaling method to normalize the feature/input variable.

4. To assess the SVR model by optimal feature sub-set.

After sub-setting the best variable feature that was selected using GA, the SVR model was assessed by the sub-set.

5. To tune hyperparameters using Particle Swarm Optimization (PSO).

Hyperparameter tuning was used to find the optimal parameters C, epsilon, and gamma (RBF kernel). To realize this, particle swarm optimization is employed.

6. To build the model of Support Vector Regression (SVR).

After the hyperparameter has been obtained for the third target (Y), which is the peak discharge, time to peak, and time base for the parameters of C and epsilon, the next linear kernel is to build the SVR model using the hyperparameter.

7. To evaluate the model.

The model was evaluated using k-fold cross-validation to ensure the best model performance and generalizability on the new data.

3. Results and Discussion

3.1. Observed Unit Hydrograph

The average observed unit hydrographs in the 12 watersheds/sub-watersheds is calculated using Collins method. Ten watersheds were used for building the model formulation, and each was taken 20 maximum observed hydrographs that consisted of 10 events for model validation, which are presented in Table 1.

Table 1. Component Result of observed unit hydrograph by Collins method in 10 watersheds (by the authors)

NO	Name of watershed	Model calibration			Model verification		
		Dependent Variable (Y)			Dependent variable (Y)		
		Peak discharge	Time to peak	Time base	Peak discharge	Time to peak	Time base
		Q_p m^3/dsmm	T_p hour	T_b hour	Q_p m^3/s/mm	T_p hour	T_b hour
1	Sub DAS Lesti (Outlet AWLR Tawang Rejeni)	25.328	6.10	18.60	19.273	7.00	18.20
2	Sub DAS Brantas Hulu (Outlet AWLR Gadang)	51.885	3.70	18.30	52.260	3.80	18.00
3	DAS Welang (Outlet AWLR Welang)	21.806	5.20	17.30	22.082	4.20	17.10
4	DAS Brang Biji (Outlet AWLR Brang Biji)	13.297	4.40	14.20	13.252	4.20	13.60
5	DAS Dodokan (Outlet AWLR Karang Makam)	3.627	6.20	14.70	3.702	6.40	14.20
6	DAS Jangkok (Outlet AWLR Bug Bug)	9.581	5.40	14.30	5.724	7.00	18.50
7	Sub DAS Oyo (Outlet AWLR Bunder)	17.432	8.10	21.30	12.206	7.60	21.30
8	Sub DAS Oyo (Outlet AWLR Kedung Miri)	48.297	7.20	17.70	31.494	6.80	20.50
9	DAS Serang (Outlet AWLR Bendungan)	7.881	5.40	19.80	6.093	6.50	19.80
10	Sub DAS Winongo (Outlet AWLR Wonokromo)	8.503	5.80	19.90	8.562	5.60	19.70
	Average	**20.764**	**5.750**	**17.610**	**17.465**	**5.910**	**18.090**

3.2. Characteristic Parameters of Watershed in Developing Model

The characteristic parameters of the watershed are the watershed morphometry parameters that are watershed area (A), main river length (L), main river slope (S), watershed shape factor (F_b), drainage density (D_d), and runoff coefficient/curve number (CN) that are obtained from data spatial analysis by using DEM (Digital Elevation Model), RBI map, land cover map, and map of soil type by using ArcGIS software; the results are presented in Tables 2 and 3.

Table 2 Parameters of Watershed Morphometry and Spatial Analysis Result in 10 Watersheds (by the authors)

NO	Name of watershed	Independent variable (X)					
		watershed area	length of main river	slope of main river	river density	watershed shape factor	Curve number
		km^2	km	S	D_d	F_b	CN
1	Sub DAS Lesti (Outlet AWLR Tawang Rejeni)	378.850	47.954	0.0538	2.295	0.334	77.354
2	Sub DAS Brantas Hulu (Outlet AWLR Gadang)	796.331	35.924	0.0208	2.420	0.187	77.274
3	DAS Welang (Outlet AWLR Welang)	264.015	12.791	0.0212	2.318	0.152	75.798
4	DAS Brang Biji (Outlet AWLR Brang Biji)	167.278	25.364	0.0177	2.212	0.303	71.785
5	DAS Dodokan (Outlet AWLR Karang Makam)	67.658	16.770	0.0067	2.404	0.085	79.881
6	DAS Jangkok (Outlet AWLR Bug Bug)	153.424	34.158	0.0037	1.732	0.166	72.540
8	Sub DAS Oyo (Outlet AWLR Bunder)	426.909	63.268	0.0023	2.183	0.201	80.470
9	Sub DAS Oyo (Outlet AWLR Kedung Miri)	685.760	90.100	0.0024	1.558	0.206	78.984
10	DAS Serang (Outlet AWLR Bendungan)	185.483	26.673	0.0261	2.318	0.152	79.953
11	Sub DAS Winongo (Outlet AWLR Wonokromo)	86.206	23.848	0.0029	1.305	0.083	85.115
	Average	**321.191**	**37.685**	**0.016**	**2.074**	**0.187**	**77.915**

Table 3 Components of the observed unit hydrograph and morphometry parameters for model validation (by the authors)

NO	Name of watershed	Dependent variable (Y)			Area rainfall	Independent variable (X)					
		Peak discharge	Time to peak	Time base		Watershed area	Length of main river	Slope of main river	River density	watershed shape	Curve number
		Q_p $m^3/s/mm$	T_p hour	T_b hour	P mm	A km^2	L km	S	D_d	F_b	CN
1	DAS Samiran (Outlet AWLR Samiran)	5.862	6.70	18.40	31.77	109.339	21.553	0.0025	2.396	0.134	78.129
2	Sub DAS Progo (Outlet Kali Bawang)	83.968	5.50	18.30	22.54	1,736.143	82.214	0.0071	2.779	0.322	79.719

3.3. SUH Model Formulation

The SUH model formulation consists of the peak discharge model (Q_p), time-to-peak model (T_p), and time-based model (T_b), which is based on the 6 watershed characteristics, namely, watershed area (A), main river length (L), main river slope (S), watershed shape factor (F_b), drainage density (D_d), and runoff coefficient/curve number (CN). The machine learning of SVR linear regression method is carried out due to the steps as follows:

1. Preparing the watershed dataset for model formulation.

Data-set consists of the watershed morphometry parameters and the component of the observed unit hydrograph that is separated into variables/features of X and Y as follows:

a) X: this variable consists of columns that will be used as feature or input in the SVR model involving the watershed area (A), main river length (L), main river slope (S), watershed shape factor (F_b), drainage density (D_d), and runoff coefficient/curve number (CN).

b) Y: this variable consists of columns that become target/output that will be predicted by SVR model that involves peak discharge, time to peak, and time base.

2. Selecting the feature (variable X) using Genetic Algorithm for model formulation.

In the context of feature importance, no universal standard is determined for the value range that categorizes the contribution of features from low to high. However, many practitioners and researchers use a general approach for interpreting the value of feature importance based on the scale produced by the model as follows:

- *Low (0.0 - 0.1):*

A feature with a feature importance value of less than 0.1 is assumed to make a negligible contribution to the model. These features may not provide significant information, and they should be deleted.

- *Moderate (0.1 - 0.3):*

A feature with a value between 0.1 and 0.3 is assumed to make a moderate contribution. These features may provide several information, but they are not strong enough to be the main features in the model.

- *High (0.3 - 0.5):*

A feature with a value between 0.3 and 0.5 is assumed to make a significant contribution. These features play an important role in prediction and should be maintained in the model.

- *Very high (0.5 - 1.0):*

A feature with a value greater than 0.5 is considered critical. These features have a large influence on the model result, and in general, the main feature must be attended.

The results of feature selection described above are given in Tables 4, 5, and 6: features that are involved in the research for the development of the SUH model for the target (Y) include peak discharge, time to peak, and time base

Table 4. Values of feature contribution to the peak discharge target (compiled by the authors)

Feature	Contribution value	Category
Watershed Area	0.8199	Very high
River drainage density	0.0749	Low
River length	0.0541	Low
Watershed shape	0.0279	Low
Curve number	0.0119	Low
River slope	0.0111	Low

Table 5. Values of feature contribution to the time-to-peak target (compiled by the authors)

Feature	Contribution value	Category
Curve number	0.441212	High
River length	0.222656	Moderate
Watershed area	0.168769	Moderate
River slope	0.167364	Moderate
River Drainage Density	0	-
Watershed shape	0	-

Table 6. Values of feature contribution to the time base target (compiled by the authors)

Feature	Contribution value	Category
Curve number	0.490195	High
Watershed area	0.296802	Moderate
River slope	0.159667	Moderate
River length	0.053336	Low
River drainage density	0	-
Watershed area	0	-

3. Normalizing the data (data scale)

Method of data normalization is by using Z-score standardization, which is also known as Standard Scaling. Z-score standardization is a data normalization technique that changes the value in a dataset such that the mean is zero and the standard deviation is 1. This technique is used to determine whether all features in the dataset have equivalent scales to prevent features with large ranges from dominating the analysis or machine learning model [2]. The Z-score formulation is performed as follows:

$$z = \frac{x - \mu}{\sigma}$$

z = the value of the data that have been normalized;
x = the value of the original data;
μ = mean of the dataset;
σ = the dataset deviation standard.

4. Hyperparameter tuning for the SVR model

Hyperparameter tuning is used to find the parameter combination of SVR model by using the optimization model of Particle Swarm Optimization (PSO). The optimization method of PSO will find the combination of hyperparameters in the SVR model that is used in this research based on the determined range.

- The result of hyperparameter tuning for peak discharge:
  - C = 49.8455
  - Epsilon = 0.0010,
  - Best score (%) = 12.9350.
- The result of hyperparameter tuning for the time-to-peak:
  - C = 69.6559
  - Epsilon = 0.0010,
  - Best score (%) = 8.5058
- The result of hyperparameter tuning for the time base:
  - C = 32.5092,
  - Epsilon = 0.0010,
  - Best score (%) = 5.8222

The result of the hyperparameter tuning shows that the best score for peak discharge is 10.3447; the best score for time to peak is 8.7718, and for time base it is 5.9750. The value of C and epsilon parameter is assumed to be optimal and can be used to build the SVR model for the SUH model.

5. Building the Support Vector Regression (SVR) model

After hyperparameter has been obtained for the third target (Y), such as peak discharge, time to peak, and time base for C and epsilon parameters, and for linear kernel, it is possible to build SVR model using the hyperparameter. The SVR model was built through the following steps:

5.1 Prediction of the SVR model for the target of peak discharge

After the model is assessed, the prediction is carried out in the calibration test data, and the prediction result that has been scaled is changed back to the original value. Then, to calculate the error deviation for each data row and the average error deviation with the percentage of the mean absolute error (MAE) for evaluating the performance of the SVR model that is by finding the difference between the prediction target result (Y) and the target result (Y) of the actual peak discharge. Table 7 presents the prediction result of the SVR model calibration data for peak discharge. The percentage of the average error deviation of the SVR model for peak discharge is 12.9350%

Table 7. Prediction results of SVR model calibration data for peak discharge (compiled by the authors)

Watershed names	Actual	Prediction	Deviation	Deviation (%)
Sub-DAS Lesti (Outlet AWLR Tawang Rejeni)	25.328	19.288	6.040	23.848
Sub-DAS Citarum Hulu (Outlet AWLR Gadang)	51.885	52.275	0.390	0.752
DAS Welang (Outlet AWLR Welang)	21.806	21.110	0.696	3.194
DAS Brang Biji (Outlet AWLR Brang Biji)	13.297	13.271	0.026	0.195
DAS Dodokan (Outlet AWLR Karang Makam)	3.627	3.604	0.023	0.631
DAS Jangkok (Outlet AWLR Bug Bug)	9.581	10.400	0.819	8.548
Sub-DAS Oyo (Outlet AWLR Bunder)	17.432	17.450	0.018	0.102
Sub-DAS Oyo (Outlet AWLR Kedung Miri)	48.297	31.476	16.821	34.828
DAS Serang (Outlet AWLR Bendungan)	7.881	9.857	1.976	25.067
Sub-DAS Winongo (Outlet AWLR Wonokromo)	8.503	8.552	0.049	0.575

5.1.1 Visualization of SVR model prediction results for peak discharge

The prediction result and actual value are plotted as a curve to provide a visual illustration of the model performance. This curve shows how the prediction of the SVR model is equivalent to the actual value in the calibration dataset (Figure 2).

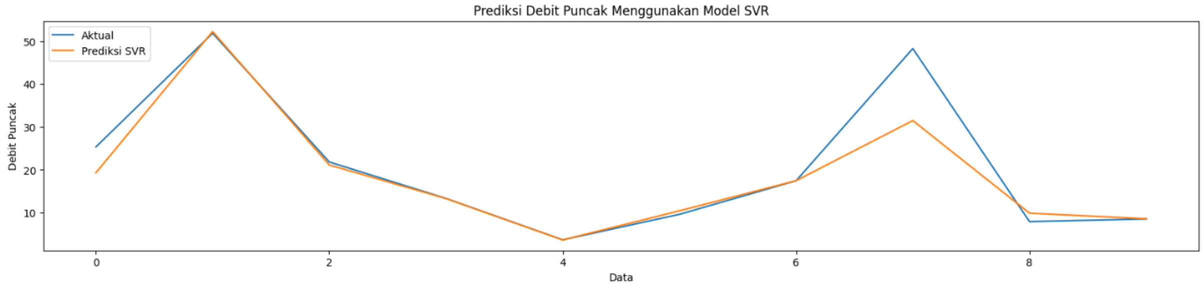

Figure 2. Curve plotting of the prediction results of SVR model calibration data for peak discharge (developed by the authors)

Based on the graphic above, it can be seen that the red line is the prediction result and the blue line is the target value (Y), which is the peak discharge. If the red and blue lines are close, then the error deviation is relatively small; however, if the lines are far away, then the error deviation is relatively large. However, overall, the line area is closer, which shows that the SVR model is good for predicting the peak discharge value.

5.1.2 Mathematical function of the SVR linear kernel for peak discharge

In the model of Support Vector Regression (SVR) with linear kernel, the regression function model the relation between the target (Y) of peak discharge and the features: X_1, X_2, X_3, X_4, X_5, X_6 with the mathematical notation is as follows:

$$f(X) = \omega^T x + b,$$

Where: ω = weight vector (coefficient);
x = feature vector (Input);
b = bias (Intercept)

The bias value of the SVR model for peak discharge = -0.0320

The bias value below 0 (negative) in the regression context indicates that the SVR model for peak discharge gives less estimation of the actual target. The weight vector (coefficient) represents the contribution of each feature (X) in the regression model. In the context of SVR with a linear kernel, this coefficient is produced from the assessment model process and reflects the influence of each feature on the output prediction [21]. The coefficient of ω shows the direction and strength of the influence of each feature on the prediction result. A larger coefficient (positive as well as negative) indicates a greater influence on the model output. If the coefficient of a feature is positive, then increasing the feature value will increase the prediction; however, if it is negative, increasing the feature value will decrease the prediction [21].

Table 8. Feature coefficients of the SVR model for peak discharge (compiled by the authors)

Name of the finite element	Variable	Coefficient
Watershed area	X_1	1.1627
River length	X_2	-0.5060
River slope	X_3	-0.0315
River drainage density	X_4	-0.1493
Watershed shape	X_5	0.0775
Curve number	X_6	-0.0340

From the bias and coefficient value above, a mathematical function can be built the mathematical function for SVR model (Support Vector Regression) with linear kernel for the SUH model about the target of peak discharge as follows:

The SVR model for the target of peak discharge:

$$f(X) = 1.1627X_1 + (-0.5060X_2) + (-0.0315X_3) + (-0.1493X_4) + 0.0775X_5 + (-0.0340X_6)$$

5.2 Prediction of the SVR model for the time to peak target

The prediction of the SVR model for the target time to peak is carried out in the same way as for the peak discharge. Table 9 presents the prediction result of the SVR model calibration data for the time to peak. The percentage of the average error deviation of the SVR model for the time to peak is 8.5058%.

Table 9. Prediction results of the SVR model with the data calibration for the time to peak (compiled by the authors)

Watershed names	Actual	Prediction	Deviation	Deviation (%)
Sub DAS Lesti (Outlet AWLR Tawang Rejeni)	6.100	6.101	0.001	0.024
Sub DAS Citarum Hulu (Outlet AWLR Gadang)	3.700	3.701	0.001	0.022
DAS Welang (Outlet AWLR Welang)	5.200	4.465	0.735	14.126
DAS Brang Biji (Outlet AWLR Brang Biji)	4.400	4.779	0.379	8.605
DAS Dodokan (Outlet AWLR Karang Makam)	6.200	6.199	0.001	0.012
DAS Jangkok (Outlet AWLR Bug Bug)	5.400	5.401	0.001	0.027
Sub DAS Oyo (Outlet AWLR Bunder)	8.100	7.176	0.924	11.404
Sub DAS Oyo (Outlet AWLR Kedung Miri)	7.200	7.201	0.001	0.009
DAS Serang (Outlet AWLR Bendungan)	5.400	6.253	0.853	15.792
Sub DAS Winongo (Outlet AWLR Wonokromo)	5.800	7.421	1.621	27.947

5.2.1 Visualization of SVR model prediction results for the time to peak

The prediction result and actual value are plotted as a curve to provide a visual illustration of the model performance. This curve shows how the prediction of the SVR model is equivalent to the actual value in the calibration dataset (Figure 3).

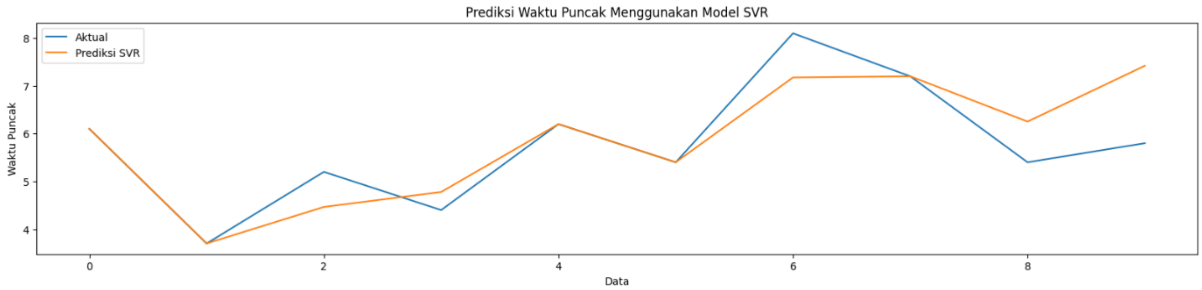

Figure 3. Curve plotting of prediction results of SVR model calibration data for time to peak (developed by the authors)

5.2.2 Mathematical function of SVR model-linear kernel for time to peak

Table 10 presents the feature coefficient results for the SVR model for time to peak. The value of the bias model for the time to peak = 0.688.

Table 10. Feature coefficients of the SVR model for the time to peak (compiled by the authors)

Feature	Variable	Coefficient
Watershed area	X_1	-0.8013
River length	X_2	0.9282
River slope	X_3	0.0162
Curve number	X_6	0.5461

Using the bias and coefficient value above, a mathematical function can be built the mathematical function for SVR model (Support Vector Regression) with linear kernel for the SUH model about the target time to peak as follows:

The SVR model for the target time to peak:

$$f(X) = (-0.8013)X_1 + 0.9282X_2 + 0.0162X_3 + 0.5461X_4$$

5.3 Prediction of SVR model for the time base target

The SVR model for the time-to-peak target is predicted using the same technique as for the time-to-peak prediction. Table 11 presents the prediction results of the SVR model calibration data for the time base. The percentage of the average error deviation of the SVR model for the time base was 5.8222%.

Table 11. Prediction results of the SVR model with the data calibration for the time base (compiled by the authors)

Watershed names	Actual	Prediction	Deviation	Deviation (%)
Sub DAS Lesti (Outlet AWLR Tawang Rejeni)	18.600	18.598	0.002	0.012
Sub DAS Citarum Hulu (Outlet AWLR Gadang)	18.300	18.297	0.003	0.018
DAS Welang (Outlet AWLR Welang)	17.300	15.840	1.460	8.442
DAS Brang Biji (Outlet AWLR Brang Biji)	14.200	14.197	0.003	0.019
DAS Dodokan (Outlet AWLR Karang Makam)	14.700	17.260	2.560	17.416
DAS Jangkok (Outlet AWLR Bug Bug)	14.300	14.588	0.288	2.011
Sub DAS Oyo (Outlet AWLR Bunder)	21.300	19.760	1.540	7.231
Sub DAS Oyo (Outlet AWLR Kedung Miri)	17.700	20.497	2.797	15.804
DAS Serang (Outlet AWLR Bendungan)	19.800	18.202	1.598	8.070
Sub DAS Winongo (Outlet AWLR Wonokromo)	19.900	19.899	0.001	0.006

5.3.1 Visualization of SVR model prediction results for time base

The prediction results and actual values are plotted in a curve to provide a visual illustration of the model performance. This curve shows how the prediction of the SVR model is equivalent to the actual value in the calibration dataset (Figure 4).

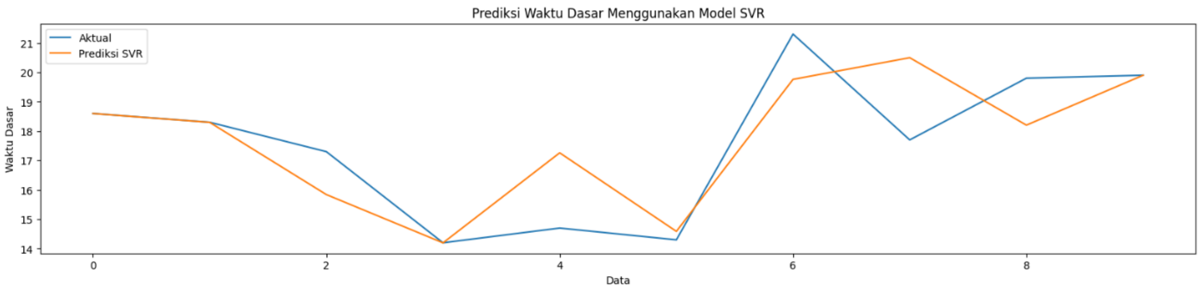

Figure 4. Curve plotting of prediction results of SVR model calibration data for time base (developed by the authors)

5.3.2 Mathematical function of SVR model-linear kernel for time base

Table 12 presents the feature coefficient results for the SVR model on the time base. The value of the bias model for time to peak = -0.0171

Table 12. Feature coefficient results for the SVR model on the time base (compiled by the authors)

Feature	Variable	Coefficient
Watershed area	X_1	0.1771
River length	X_2	0.3323
River slope	X_3	0.1128
Curve number	X_6	0.7149

From the bias and feature coefficient values above, a mathematical function can be built for the Support Vector Regression (SVR) model with linear kernel for the SUH model about the target time base as follows:

The SVR model for the target of time base:

$$f(X) = 0.1771X_1 + 0.3323X_2 + 0.1128X_3 + 0.7149X_4$$

6. Predicting data validation

The results of each prediction were used to determine the performance of the SVR model if it was tested using data validation.

6.1. Predicting data validation of the SVR model for the peak discharge target

Validation was carried out by calculating the error deviation in each data row and the average error deviation by the percentage of the mean absolute error (MAE) for evaluating the performance of the SVR model, that is, to find the difference between the prediction target result (Y) and actual target result of peak discharge. Table 13 presents the prediction results of the data validation of the SVR model for peak discharge. The MAE for peak discharge is 43.4532%.

Table 13. Prediction results of SVR model based on data validation for peak discharge (compiled by the authors)

Watershed names	Actual	Prediction	Deviation	Deviation (%)
DAS Samiran (Outlet AWLR Samiran)	5.142	6.216	1.074	20.896
Sub-DAS Progo (Outlet Kali Bawang)	73.013	105.899	32.886	45.042

6.2. Predicting data validation of the SVR model for the time to peak target

Validation was carried out by calculating the error deviation in each data row and the average error deviation by the percentage of the mean absolute error (MAE) for evaluating the performance of the SVR model, that is, to find the difference between the prediction target result (Y) and actual target result of the time to peak. Table 14 presents the prediction results of the SVR model data validation for the time to peak. The MAE for the time to peak was 34.3782%.

Table 14. Prediction results of the SVR model based on data validation for time to peak (compiled by the authors)

Watershed names	Actual	Prediction	Deviation	Deviation (%)
DAS Samiran (Outlet AWLR Samiran)	7.000	5.949	1.051	15.020
Sub-DAS Progo (Outlet Kali Bawang)	6.000	2.582	3.418	56.962

6.3 Predicting data validation of the SVR model for the time base target

Data were validated by calculating the error deviation in each data row and the average error deviation by the percentage of the mean absolute error (MAE) for evaluating the performance of the SVR model, that is, to find the difference between the prediction target result (Y) and actual target result of the time base. Table 15 presents the prediction results of the SVR model data validation for the time base. The MAE for the time base is 17.8851%.

Table 15. Prediction results of the SVR model based on data validation for the time base (compiled by the authors)

Watershed names	Actual	Prediction	Deviation	Deviation (%)
DAS Samiran (Outlet AWLR Samiran)	19.800	16.621	3.179	16.054
Sub-DAS Progo (Outlet Kali Bawang)	18.800	22.525	3.725	19.813

4. Conclusion

In some studies, the SUH methods still have limitations, mainly if they are applied in the watershed outside of the study locations because they often produce a high enough deviation of peak discharge. The main factor that influences the accuracy of the SUH model is the accuracy in determining the watershed characteristics and the other parameters that are used in the method; therefore, the SUH model can be close to the observed unit hydrograph.

The geomorphology parameter from SUH is the most useful approach for predicting runoff and the simplest method for understanding the different hydrology behaviors of watersheds, mainly in ungauged watersheds or those with a lack of data. This research builds a Synthetic Unit Hydrograph based on the watershed characteristics, not from rainfall and runoff data.

The results are as follows: 1) formulation of peak discharge: $f(X) = 1.1627 X_1 + (-0.5060 X_2) + (-0.0315 X_3) + (-0.1493 X_4 + 0.0775 X_5 + (-0.0340 X_6)$; 2) formulation of time to peak: $f(X) = (-0.8013) X_1 + 0.9282 X_2 + 0.0162 X_3 + 0.5461 X_4$; and 3)formulation of time base: $f(X) = 0.1771 X_1 + 0.3323 X_2 + 0.1128 X_3 + 0.7149 X_4$; explanation: X_1 = watershed area; X_2 = river length; X_3 = river slope; X_4 = river drainage density; X_5 = watershed shape; and X_6: run=off coefficient.

Declarations

Author Contributions

Conceptualization, A.G.M. and L.M.L.; methodology, A.G.M.; software, S.; investigation, A.G.M., S., L.M.L. and U.A.; validation, A.G.M., S., L.M.L. and U.A.; formal analysis, A.G.M. and L.M.L..; resources, A.G.M.; data curation, A.G.M. and

L.M.L.; writing—original draft preparation, all authors contributed equally; writing—review and editing, A.G.M., and L.M.L.; visualization, S., and U.A.; supervision, A.G.M.; project administration, L.M.L. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

Funding information is not available.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

References

[1] MAJDI A.G., SUHARDJONO, LIMANTARA L.M., and ANDAWAYANTI U. Flood observation in Indonesian Regions to Support the development of synthetic unit hydrograph based on watershed characteristic. *Journal of Hunan University (Natural Sciences)*, 2024, 51(10): 41-49, https://doi.org/10.55463/issn.1674-2974.51.10.5

[2] SINGH P., GUPTA A., and SINGH M. Hydrological inferences from watershed analysis for water resource management using remote sensing and GIS techniques. *The Egyptian Journal of Remote Sensing and Space Sciences*, 2014, 17(2): 111–121, https://doi.org/10.1016/j.ejrs.2014.09.003

[3] PATIL S.J. and BHAGWAT T.N. A review of synthetic hydrograph methods for design storm. *International Research Journal of Engineering and Technology*, 2019, 6 (11): 2413-2418.

[4] MISHRA N., KHARE D., GUPTA K.K., and SHUKLA R. Impact of land use change on groundwater - A review. *Advances in Water Resource and Protection*, 2014, 2(2): 28-41.

[5] NOURANI V., SINGH V.P., DELAFROUZ H. Three geomorphological rainfall–runoff models based on the linear reservoir concept. *Journal of Catena*, 2009, 76: 206–214. https://doi.org/10.1016/j.catena.2008.11.008

[6] KHALEGHI M.R., GHOLAMI V., GHODUSI J., and HOSSEINI H. Efficiency of the geomorphologic instantaneous unit hydrograph method in flood hydrograph simulation. *Journal of Catena*, 2011, 87(2): 163-171. https://doi.org/10.1016/j.catena.2011.04.005

[7] NARENDRA K, and NAGESWARA R. K. Morphometry of the Mehadrigedda watershed, Visakhapatnam district, Andhra Pradesh using GIS and Resourcesat data. *Journal of the Indian Society of Remote Sensing*, 2006, 34(2):101–110, http://doi.org/10.1007/BF02991815

[8] SREEDEVI P.D., OWAIS S., KHAN H.H., & AHMED S. Morphometric analysis of a watershed of south India using SRTM data and G.I.S. *Journal of the Geological Society of India*, 2009, 73(4): 543–552, http://doi.org/10.1007/s12594-009-0038-4

[9] AHER P.D., ADINARAYANA J., & GORANTIWAR S.D. Quantification of morphometric characterization and prioritization for management planning in semi-arid tropics of India: A remote sensing and G.I.S. approach. *Journal of Hydrology*, 2014, 511: 850–860, https://doi.org/10.1016/j.jhydrol.2014.02.028

[10] ALBAROOT M., AL-AREEQ N.M., ALDHARAB H.S., ALSHAYEF M., and GHAREB S.A. Quantification of morphometric analysis using remote sensing and G.I.S. techniques in the Qa'Jahran Basin, Thamar Province, Yemen. *International Journal of New Technology and Research*, 2018, 4(8): 12–22.

[11] MURPHEY, J.B., WALLACE, D.E., & LANE, L.J. Geomorphic parameters predict hydrograph characteristics in the southwest. *Water Resources Bulletin*, 1977, 13(1): 25-38. https://doi.org/10.1111/j.1752-1688.1977.tb01987.x

[12] TRIPATHI S., SONI S.K., and MAURYA A.K. Morphometric characterization and prioritization of sub-watershed of Seoni River in Madhya Pradesh through remote sensing and GIS technique. *International Journal of Remote Sensing & Geoscience*, 2013, 2(3): 46–54

[13] SONI S.K., TRIPATHI S, and MAURYA A.K. GIS based morphometric characterization of mini-watershed—Rachhar Nala of Anuppur District Madhya Pradesh. *International Journal of Advanced Technology & Engineering Research*, 2013, 3(3): 32–38

[14] TIGNATH S., KAPOOR M., JHA M., and SHARMA S.K. Morphometric analysis of part of the Hiran River, District Jabalpur, M.P., India using Remote sensing and GIS. *International Journal of Environmental Sciences*, 2014, 5: 181–196.

[15] BANERJEE A., SINGH P., and PRATAP K. Morphometric evaluation of Swarnrekha watershed, Madhya Pradesh, India: an integrated GIS-based approach. *Applied Water Scencei*, 2015, 7: 1807-1815, https://doi.org/10.1007/s13201-015- 0354-3

[16] FAROOQ M., SINGH S.K., and KANGA S. Inherent vulnerability profiles of agriculture sector in temperate Himalayan region: a preliminary assessment. *Indian Journal of Ecology*, 2021, 48(2): 434–441.

[17] BERA A. and SINGH S.K. Comparative assessment of livelihood vulnerability of climate induced migrants: a micro level study on Sagar Island, India. *Sustainability, Agri, Food and Environmental Research*, 2021, 9(2): 1–15. https://doi.org/10.7770/safer-V9N2-art2324.

[18] SINGH P., GUPTA A., and SINGH M. Hydrological inferences from watershed analysis for water resource management using remote sensing and GIS techniques. *The Egyptian Journal of Remote Sensing and Space Sciences*, 2014, 17(2): 111–121, https://doi.org/10.1016/j.ejrs.2014.09.003

[19] SAFARINA A. B. Modified Nakayasu Synthetic Unit Hydrograph Method for Meso Scale Ungauge Watersheds. *International Journal of Engineering, Research and Applications*, 2012, 2(4): 649-654

[20] TUNAS I G. Performance of Flood Prediction Model in Tropical River Basin, Indonesia: A Synthetic Unit Hydrograph-based Evaluation. *Disaster Advances*, 2019, 12: 26-37.
[21] XIE M., XIE L., & ZHU P. An efficient feature weighting method for support vector regression. *Mathematical Problems in Engineering*, 2021, 2021: 6675218, http://doi.org/10.1155/2021/6675218

参考文:

[1] MAJDI A.G., SUHARDJONO, LIMANTARA L.M., 和 ANDAWAYANTI U. 印度尼西亚地区洪水观测支持基于流域特征的合成单位线的开发。湖南大学学报（自然科学版），2024，51（10）：41-49，https://doi.org/10.55463/issn.1674-2974.51.10.5
[2] SINGH P., GUPTA A., 和 SINGH M. 利用遥感和 GIS 技术从流域分析中进行水文推断以进行水资源管理。埃及遥感与空间科学杂志，2014，17（2）：111–121，https://doi.org/10.1016/j.ejrs.2014.09.003
[3] PATIL S.J.和 BHAGWAT T.N. 设计风暴的合成水文图方法综述。国际工程技术研究杂志，2019，6 (11): 2413-2418。
[4] MISHRA N.、KHARE D.、GUPTA K.K. 和 SHUKLA R. 土地利用变化对地下水的影响——综述。水资源与保护进展，2014，2(2): 28-41。
[5] NOURANI V.、SINGH V.P.、DELAFROUZ H. 基于线性水库概念的三种地貌降雨径流模型。系列 杂志，2009，76: 206–214。https://doi.org/10.1016/j.catena.2008.11.008
[6] KHALEGHI M.R.、GHOLAMI V.、GHODUSI J. 和 HOSSEINI H. 地形瞬时单位线法在洪水过程线模拟中的效率。卡泰纳杂志，2011，87(2)：163-171。https://doi.org/10.1016/j.catena.2011.04.005
[7] NARENDRA K 和 NAGESWARA R. K. 使用 GIS 和 Resourcesat 数据对安得拉邦维沙卡帕特南区 Mehadrigedda 流域进行形态测量。《印度遥感学会杂志》，2006，34(2):101–110，http://doi.org/10.1007/BF02991815
[8] SREEDEVI P.D.、OWAIS S.、KHAN H.H. 和 AHMED S. 使用 SRTM 数据和 G.I.S. 对印度南部流域进行形态分析。《印度地质学会杂志》，2009，73(4): 543–552，http://doi.org/10.1007/s12594-009-0038-4
[9] AHER P.D.、ADINARAYANA J. 和 GORANTIWAR S.D. 对印度半干旱热带地区管理规划的形态特征和优先排序进行量化：一种遥感和 G.I.S. 方法。《水文学杂志》，2014，511：850–860，https://doi.org/10.1016/j.jhydrol.2014.02.028
[10] ALBAROOT M.、AL-AREEQ N.M.、ALDHARAB H.S.、ALSHAYEF M. 和 GHAREB S.A. 使用遥感和 G.I.S. 技术对也门 Thamar 省 Qa'Jahran 盆地的形态分析进行量化。《国际新技术与研究杂志》，2018，4(8): 12–22。
[11] MURPHEY，J.B.、WALLACE，D.E. 和 LANE，L.J. 地貌参数预测西南地区的水文特征。《水资源公报》，1977，13(1)：25-38。https://doi.org/10.1111/j.1752-1688.1977.tb01987.x
[12] TRIPATHI S.、SONI S.K. 和 MAURYA A.K. 通过遥感和 GIS 技术对中央邦 Seoni 河子流域进行形态特征描述和优先排序。国际遥感与地球科学杂志，2013，2(3)： 46–54
[13] SONI S.K.、TRIPATHI S 和 MAURYA A.K. 基于 GIS 的微型流域形态特征——中央邦 Anuppur 区的 Rachhar Nala。国际先进技术与工程研究杂志，2013，3(3)： 32–38
[14] TIGNATH S.、KAPOOR M.、JHA M. 和 SHARMA S.K.使用遥感和 GIS 对印度中央邦贾巴尔普尔区希兰河部分河段进行形态分析。《国际环境科学杂志》，2014, 5：181-196。
[15] BANERJEE A.、SINGH P. 和 PRATAP K. 印度中央邦 Swarnrekha 流域的形态评估：一种基于 GIS 的综合方法。《应用水科学》，2015， 7：1807-1815，https://doi.org/10.1007/s13201-015-0354-3
[16] FAROOQ M.、SINGH S.K. 和 KANGA S. 温带喜马拉雅地区农业部门固有脆弱性概况：初步评估。《印度生态学杂志》，2021， 48(2)： 434-441。
[17] BERA A. 和 SINGH S.K. 气候诱发移民生计脆弱性的比较评估：印度萨加尔岛的微观层面研究。可持续性、农业、食品和环境研究，2021，9(2)：1-15。https://doi.org/10.7770/safer-V9N2-art2324。
[18] SINGH P.、GUPTA A. 和 SINGH M. 利用遥感和 GIS 技术从流域分析中进行水文推断，以进行水资源管理。埃及遥感与空间科学杂志，2014，17（2）：111–121，https://doi.org/10.1016/j.ejrs.2014.09.003
[19] SAFARINA A. B. 改进的 Nakayasu 合成单位线法用于中尺度无量纲流域。国际工程、研究与应用杂志，2012，2（4）：649-654
[20] TUNAS I G. 印度尼西亚热带河流流域洪水预测模型性能：基于合成单位线的评估。灾害进展，2019，12： 26-37。
[21] XIE M.，XIE L.，和 ZHU P. 一种有效的支持向量回归特征加权方法。 工程中的数学问题， 2021，2021： 6675218，http://doi.org/10.1155/2021/6675218