Applied Water Science (2022) 12:87
https://doi.org/10.1007/s13201-022-01621-7

ORIGINAL ARTICLE

Modeling stage–discharge–sediment using support vector machine and artificial neural network coupled with wavelet transform

Manish Kumar[1] · Pravendra Kumar[1] · Anil Kumar[1] · Ahmed Elbeltagi[2] · Alban Kuriqi[3]

Received: 14 July 2021 / Accepted: 8 March 2022 / Published online: 4 April 2022

Abstract

Many real water issues involve rivers' sediment load or the load that rivers can bring without degrading the fluvial ecosystem. Therefore, the assessment of sediments carried by a river is also crucial in the planning and designing of various water resource projects. In the current study, five different data-driven techniques, namely artificial neural network (ANN), wavelet-based artificial neural network (WANN), support vector machine (SVM), wavelet-based support vector machine (WSVM), and multiple-linear regression (MLR) techniques, were employed for time-series modeling of daily suspended sediment concentration (SSC). Hydrological datasets containing the daily stage (h), discharge (Q), and SSC for 10 years (2004–2013) from June to October at Adityapur and Ghatshila station of Subernrekha river basin, Jharkhand, India, were considered for analysis. The Gamma test was used to determine the input variables in the first step. Various combinations were made by lagging the maximum three-day time step for predicting current-day SSC. The outcomes of ANN, SVM, WAAN, WSVM, and MLR models were evaluated with the actual values of SSC based on statistical metrics. Pearson correlation coefficient (PCC), root-mean-square error (RMSE), Nash–Sutcliffe efficiency (NSE), and Wilmot index (WI) as well as visual inspection of time variation, scatter plots, and Taylor diagrams. Our results stated that the WSVM model discovered the best trustworthy models among all existing models. PCC, RMSE, NSE, and WI values were 0.844 and 0.781, 0.096 g/l and 0.057 g/l, 0.711 and 0.591, 0.907 and 0.878, respectively, throughout the training and testing processes at the Adityapur site. Also, at the Ghatshila location, it was the most accurate model. During the training and testing stages, PCC, RMSE, NSE, and WI values were 0.928 and 0.751, 0.117 g/l and 0.095 g/l, 0.861 and 0.541, 0.962 and 0.859, respectively. Our findings showed that the WANN model was the second-best model during the testing phase for both sites. Hence, the WSVM technique can model SSC at this location and other similar (i.e., geomorphology and flow regime type) rivers.

Keywords Stage–discharge–sediment modeling · Gamma test · ANN · SVM · Wavelet · Adityapur · Ghatshila · Taylor diagram

✉ Alban Kuriqi
alban.kuriqi@tecnico.ulisboa.pt

Manish Kumar
ct52623d@gbpuat.ac.in

Pravendra Kumar
tswpk2@gbpuat.ac.in

Anil Kumar
anilkumar_swce61@rediffmail.com

Ahmed Elbeltagi
ahmedelbeltagy81@mans.edu.eg

[1] Department of Soil and Water Conservation Engineering, College of Technology, G.B.P.U.A.&T., Pantnagar, Uttarakhand, India

[2] Agricultural Engineering Dept., Faculty of Agriculture, Mansoura University, Mansoura 35516, Egypt

[3] CERIS, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Introduction

Sedimentation is one of the critical issues in the geomorphological processes within a basin. Due to sediment transport dynamics, various geological, hydrological, and hydraulic sedimentation issues are caused due to sediment transport dynamics (Adib and Mahmoodi 2017; Gholami et al. 2018; Jian et al. 2014). In the construction of hydraulic systems on various watershed sections, sediment content is often impacted. Thus, sediment loads are decisively calculated (Choubin et al. 2018; Kuriqi et al. 2020; Moeeni and Bonakdari 2018). It is essential to correctly estimate the amount of river sediment to design dams, storage structures, and canals, evaluate environmental effects, and decide the effectiveness of watershed management and other catchment treatments. Regression-based sediment rating curves are often used to estimate a river's sediment load. Generalization capability problems subject to multi-linear regression (MLR) and curve fitting techniques have proven insufficient (Kisi 2005). The high degree of scattering that can be reduced but not eliminated is an inherent problem in the rating curve technique (Jain 2001). Asselman (2000) used the sediment rating curve approach at four separate sites along the Rhino River and its major tributaries, showing that rating curves obtained from logarithmically transformed results are likely to underestimate sediment transport levels by 10 50%. Diverse studies contrast-suspended concentration, which has been carried out and predicted, shows that conventional rating curves can significantly underestimate existing sediment concentrations (Adib and Mahmoodi 2017; Asselman 2000; Hauser-Davis et al. 2010).

More advanced methods based on artificial neural networks (ANN) algorithms have been developed for sediment transport and accumulation estimation in different rivers and lakes (Banadkooki et al. 2020; Cigizoglu 2004; Cobaner et al. 2009; Jain 2012; Kumar et al. 2016; Partal and Cigizoglu 2008). ANN is an adaptive framework that can predict previously encountered datasets but with specific features connected to input datasets, learning input/output relationships (Gholami et al. 2018; Kisi 2005). ANN was used to model the stage–discharge that usually performs better than traditional ones (Ajmera and Goyal 2012; Hasanpour Kashani et al. 2015; Heggen 1999; Song et al. 2013; Sudheer and Jain 2003).

The support vector machine (SVM) algorithm is also successfully used in several hydraulic and hydrologic-related problems (Cherkassky and Ma 2004; Jain 2012). The SVM follows the principle of structural risk mitigation of upper binding to generalization mistake instead of minimizing a training error greater than the philosophy of methodological risk minimization (Jain 2012). This is an effective way to address nonlinear classifications, regression processes, and time series (Wang et al. 2008). The SVM stands for kernel-based learning utilizing a high-dimension linear theory space called feature space as a supervised machine learning environment that has become quite popular. The SVM works by mapping data to a higher-dimensional space using inferred kernel functions (Sivapragasam and Muttil 2005). SVMs are used effectively by many scientists in hydrological studies (Ghorbani et al. 2013; Jain 2012; Khan and Coulibaly 2006; Kisi and Cimen 2011; Malik and Kumar 2015; Rahgoshay et al. 2018; Wu et al. 2008).

In conjunction with ANN, many research types used wavelet techniques for water management and environmental engineering issues (Kişi 2010). Combined methods have recently gained growing popularity. Wavelet analysis (WA) is a standard analysis technique as spectral and temporal data can be seen simultaneously in a single signal (Nourani et al. 2009). Kim and Valdés (2003) predict droughts using wavelet-ANN in Mexico. Adib and Mahmoodi (2017) developed a hybrid wavelet-ANN model for the monthly rainfall–runoff research in Italy. Kişi (2008) investigated the precision in monthly streamflow prediction for wavelets-ANN and single ANN models and found that wavelets-ANN function much better than single ANNs. Nourani et al. (2009) examine the impact of data preprocessing on the outcomes of ANN models using continuous and discrete wavelet transformations. Partal and Cigizoglu (2008) suggested utilizing wavelets and neural networks in a study to estimate and forecast a load of rivers' suspended sediment. These studies showed that the preprocessed data with wavelet analysis performs better than the undecomposed raw data. It can be shown that the different features of the suspended sediment load prediction time series can be expressed by the sub-time series obtained using wavelets (Kuo et al. 2010). Short- and long-term forecasts' accuracy is enhanced (Nourani et al. 2018; Sharghi et al. 2018; Shiri and Kisi 2010).

However, while machine learning models have shown to be reliable in general, they are still not widely used for estimating the stage–discharge–sediment relationship in some situations. As a result, applying techniques to model, this dynamic relationship is motivated by previous use of effective learning strategies for various hydrological and hydraulic issues. Using specific actual datasets with ANN, SVM, WAAN, WSVM, and MLR, this work explores how current data-driven models are applied to explore the stage–discharge–sediment relationships Adityapur and Ghatshila sites of the Subernrekha river basin. According to the author's understanding, there has been less work to use modeling to estimate suspended sediment using the stage as input parameters. No work has been done using the given model for the given study site. This research aims to apply modern data-driven models to water management to solve

various complex problems in hydraulics and hydrology in the study area.

Materials and methods

Study area and data acquisition

This study was carried out in the Adityapur and Ghatshila sites in the Saraikela Kharsawan districts of Jharkhand, India. Adityapur site is at Kharkai River, a major tributary of the Subernrekha River, which lies at a latitude of 22° 47′29″ N and longitude 86° 10′06″ E. The Ghatshila site is situated on the main river course of Subernrekha, having a latitude of 22° 34′49″ N and a longitude of 86° 20′08″ E. Map for the studied area is shown in Fig. 1.

The area contributing runoff to the study site is 8335.25 km^2. It is located at the height of 140 m above sea level. The study site's estimated annual rainfall ranges from 1000 to 1400 mm. The southwest monsoon generally influences the study area, which has onset timing in June and extended up to October. The temperature variations during the summer season are from 35 to 40 °C, while during the winter season, it varies from 10 to 15 °C. The topography is generally flat with some undulations, small hillocks, and scattered ridges. The different rock types included in the study area are mica-schist, quartz-mica, quartzite, and schistose amphibolite of the Precambrian age. The vegetal cover is sparse, and the primary crop grown is wheat, rice, and maize, among others. The daily stage, discharge, and suspended sediment concentration (SSC) of the study area from June 1, 2004, to October 31, 2013, are considered in the analysis. Figure 2a–c presents the dataset of total data length of the period mentioned above as 1530 in which 70% (2004–2010) were used for training the dataset for model development, while the rest 30% (2011–2013) was used for the testing phase and validation purpose. The data

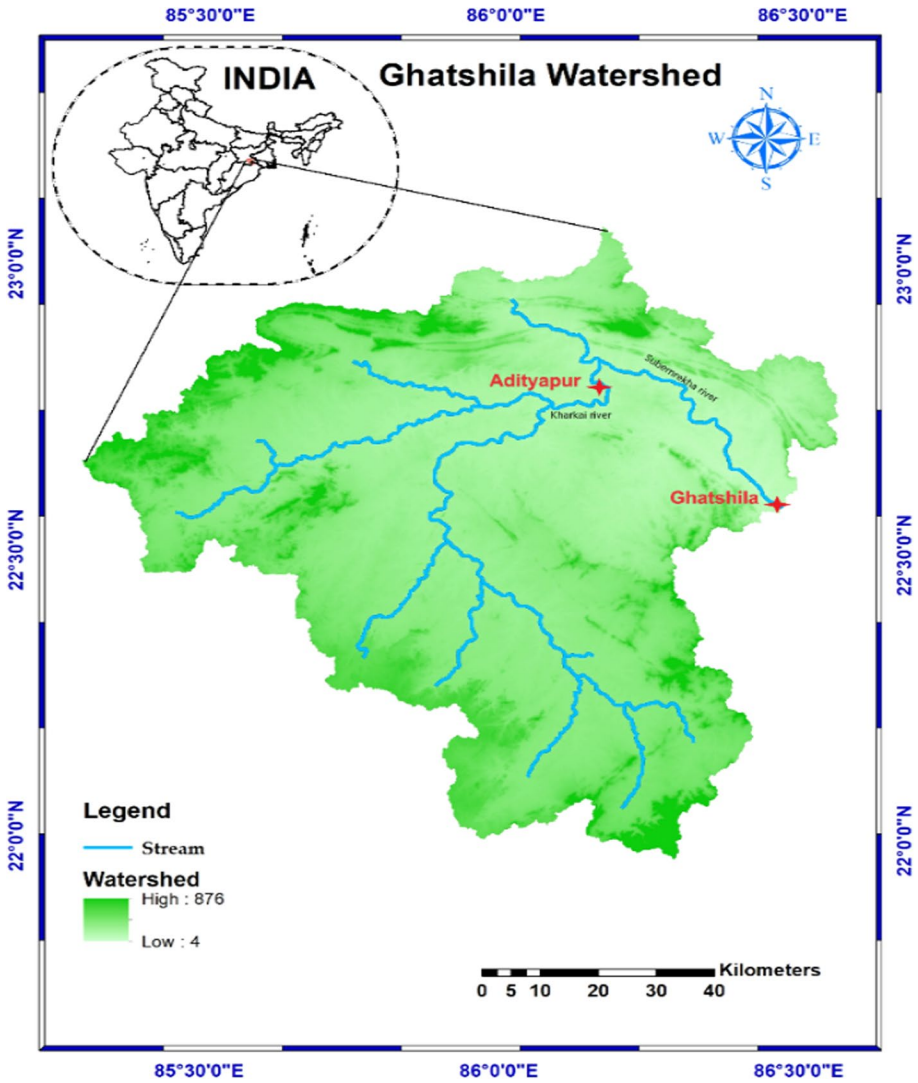

Fig. 1 Location map of the Ghatshila watershed

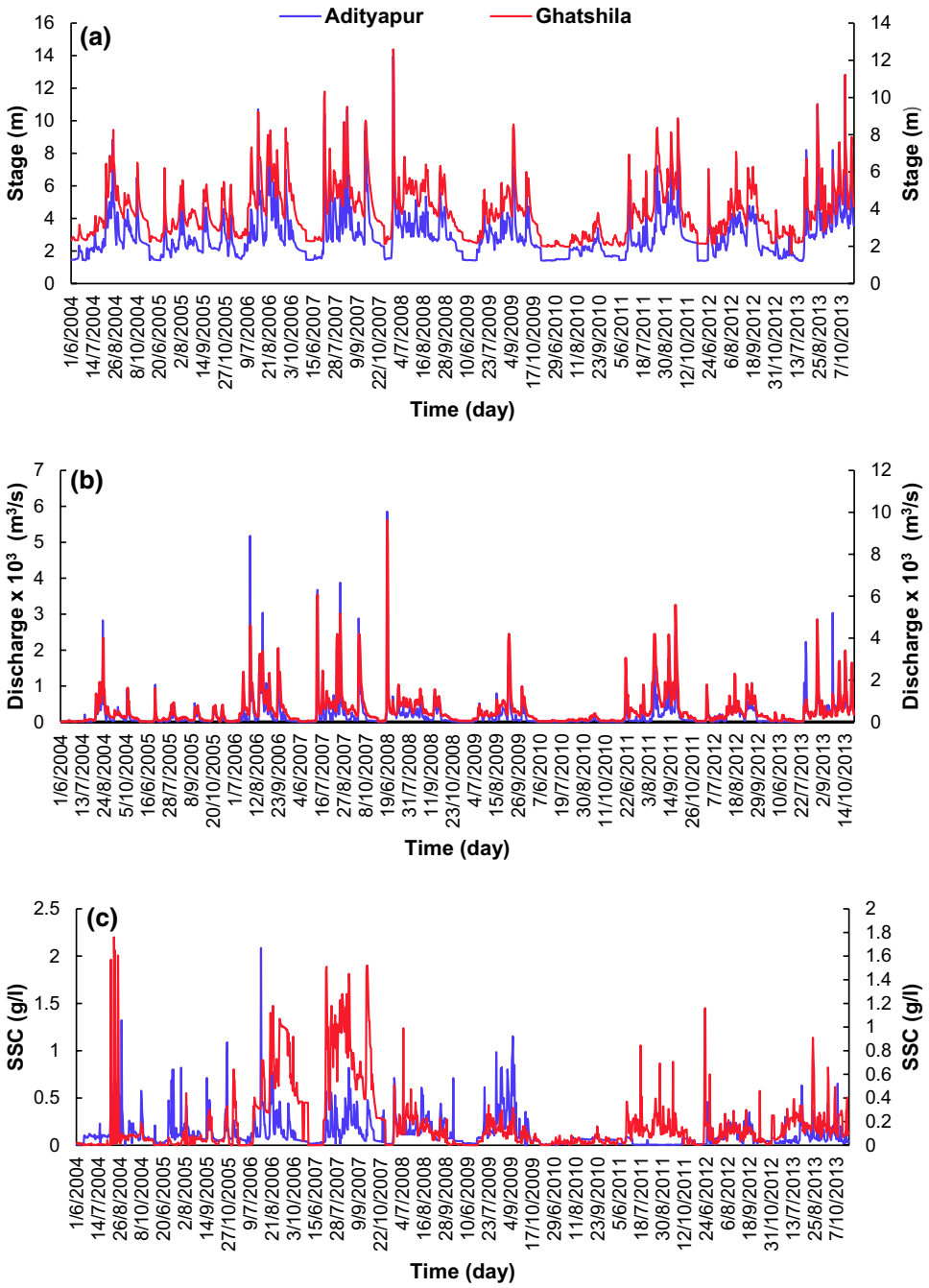

Fig. 2 Time-series representation: **a** stage, **b** discharge and **c** SSC for the Adityapur and Ghatshila site

are acquired from the government portal https://indiawris.gov.in.

Statistical analysis

The statistical analysis of daily stage (m), discharge (m^3/s), and suspended sediment concentration (SSC, g/l) for the Adityapur site and Ghatshila site (Jharkhand, India) is presented in Tables 1, 2. Statistical analysis for the datasets collected containing the training and the testing sets includes the meaning, median, minimum, maximum, standard deviation (Std. Dev.), coefficients of variations (C.V.), and skewness values. In general, Tables 1 and 2 showed statistical characteristics for all data, training sets, and testing sets that were more or less comparable in terms of mean, median, standard deviation, C.V., and skewness.

The mean values of stage and discharge are greater for testing data than all data and training sets, but SSC's value is lesser for the testing set than for the other two. The values of the stage range from 1.380 to 13.950 m and 1.60 to 12.59 m, respectively, for Adityapur and Ghatshila. The discharge values for Adityapur and Ghatshila range from 0.001 to 5.856

Table 1 Statistical investigation for all data, training data, and testing data of stage (m), discharge (m^3/s), and SSC (g/l) for the Adityapur site

Statistical parameters	Mean	Median	Minimum	Maximum	Std. dev.	C.V.	Skewness
Whole data							
Stage	2.911	2.600	1.380	13.950	1.417	8.425	2.189
Discharge	0.200	0.047	0.001	5.856	0.423	0.179	6.113
SSC	0.123	0.071	0.000	2.085	0.160	23.014	3.648
Training set							
Stage	2.830	2.540	1.400	13.950	1.389	9.491	2.361
Discharge	0.183	0.043	0.002	5.855	0.444	0.197	6.756
SSC	0.145	0.084	0.000	2.085	0.178	19.381	3.385
Testing set							
Stage	3.101	2.830	1.380	12.800	1.466	6.865	1.883
Discharge	0.241	0.132	0.001	3.0333	0.365	0.133	3.447
SSC	0.073	0.047	0.003	0.653	0.089	10.464	2.719

Table 2 Statistical investigation for all data, training data, and testing data of stage (m), discharge (m^3/s), and SSC (g/l) for the Ghatshila site

Statistical parameters	Mean	Median	Minimum	Maximum	Std. dev.	C.V.	Skewness
Whole data							
Stage	3.83	3.58	1.60	12.59	1.48	2.19	1.28
Discharge	0.532	0.285	0.007	9.609	0.790	0.624	4.166
SSC	0.18	0.08	0.00	1.76	0.27	0.08	2.59
Training set							
Stage	3.77	3.52	1.97	12.59	1.47	2.16	1.36
Discharge	0.481	0.241	0.007	9.609	0.775	0.601	4.976
SSC	0.20	0.05	0.00	1.76	0.31	0.10	2.24
Testing set							
Stage	3.96	3.65	1.60	11.22	1.49	2.23	1.13
Discharge	0.651	0.367	0.017	5.580	0.813	0.661	2.670
SSC	0.14	0.12	0.00	1.16	0.14	0.02	3.23

($\times 10^3$ m^3/s), and 0.007 to 9.609 ($\times 10^3$ m^3/s), respectively, and the SSC ranges from 0.0 to 2.085 g/l and 0.0 to 1.76 g/l, respectively. The skewness values in Tables 1 and 2 show that the distribution is positively skewed (Ghorbani et al. 2013; Liu et al. 2013; Rajaee et al. 2011). The skewness coefficients for discharge values are more significant, followed by SSC and stage values.

Artificial neural network (ANN)

The ANN technique is used to simulate a similar process as the human brain's problem-solving process. The ANN technique has received much attention in the last few decades to model and predict the nonlinear hydrologic and hydraulics processes' nonlinear behavior. Among ANN, the feed-forward back-propagation techniques have drawn much attention due to their less complexity (Choubin et al. 2018; Solomatine and Ostfeld 2008). The ANN technique has three layers: (i) the input layer, I (ii) the hidden layer (j), and (iii) the output layer (j) (k) (Fig. 3).

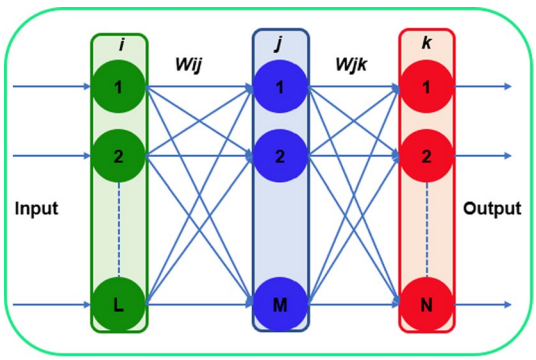

Fig. 3 Structure for the three-layer artificial neural network, adapted from Kişi (2010)

Between the layers of neurons (1, 2,…, L, M, N), entangled weight Wij and Wjk are used to link them. An input layer's neurons coordinate in a forward direction. The output for the given input value is computed during a nonlinear function called the activation function. The weight value is adjusted during training using the trial-and-error process

(Alp and Cigizoglu 2007; Kişi 2010). Overfitting is one of the biggest challenges during training processes. In this analysis, Levenberg–Marquardt was used to train the model. The hyperbolic tangent sigmoid transfer function was used to calculate the layer's output from its net input.

Support vector machine (SVM)

The SVM approach depends on the theory of statistical learning (Vapnik 1999). The SVM is a community of artificial networks notable for its overall success in classifying patterns and nonlinear regression (Cao and Tay 2003). The SVM is used for evaluating variable time series regression to estimate and simulate the same variables. The SVM model's relationship is as follows Kisi et al. (2017).

In the case of a training dataset, T, which is denoted by

$$T = \left\{\left(x_1, y_1\right), \left(x_2, y_2\right), \ldots, \left(x_m, y_m\right)\right\} \tag{1}$$

where $x \in X$ and $\mathbb{R}^n$ are the training inputs and $y \in Y$ and $\mathbb{R}^n$ are the training outputs. Assume that $f(x)$ is a nonlinear equation and given by:

$$f(x) = \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\Phi}(\boldsymbol{x_i}) + b \tag{2}$$

where $\boldsymbol{w}$ refers to the weight vector, b corresponds to the bias, and $\boldsymbol{\Phi}(\boldsymbol{x_i})$ denotes the high-dimensional feature space, linearly mapped from the input space x. SVM aims to reduce the gap between data from observations and simulations. Thus, SVM techniques reduce the objective function to minimize errors depending on the process of optimization. The error function ignores errors that are smaller than the threshold ε.

$$\begin{aligned} &\text{minimize} : \frac{1}{2}\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w} \\ &\text{subject to} : \begin{cases} y_i - \left(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\Phi}\left(\boldsymbol{x_i}\right) + b\right) \le \varepsilon \\ y_i - \left(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\Phi}\left(\boldsymbol{x_i}\right) + b\right) \ge \varepsilon \end{cases} \end{aligned} \tag{3}$$

where ε (≥ 0) represents the maximum acceptable deviation.

For solving Eq. (3), the slack variables account for possible infeasible optimization problems. This may further lead to the following formulation as given by Vapnik (1995):

$$\begin{aligned} &\text{minimize} : \frac{1}{2}\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w} + \boldsymbol{C}\sum_{i=1}^{m}\left(\xi_i^+ + \xi_i^-\right) \\ &\text{subject to} : \begin{cases} y_i - \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\Phi}\left(\boldsymbol{x_i}\right) - b \le \varepsilon + \xi_i^+ \\ \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\Phi}\left(\boldsymbol{x_i}\right) + b - y_i \le \varepsilon + \xi_i^- \\ \xi_i^+, \xi_i^- \ge 0 \end{cases} \end{aligned} \tag{4}$$

where C is the penalty coefficient represents the weight loss function. The term $\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w}$ represented the regularization term and makes them as 'flat' as possible; second term $\boldsymbol{C}\sum_{i=1}^{m}\left(\xi_i^+ + \xi_i^-\right)$ is called a practical term and measures ε-incentive loss function. The slack variables, *i.e.*, ξ_i^+, ξ_i^- represents upper and lower deviations, respectively. The highest deviation represents the ε-tube. Since all of the data points in this tube are equal to 0, they do not refer to the regression model (Fig. 4).

The values of the above parameters are then substituted in Eq. 2 to obtain $f(x)$. Nonlinear time series can be predicted and analyzed using the SVM model. Thus, the final expansion of support vector regression is given by:

$$f(x) = \sum_{i=1}^{m}\left(\alpha_i^+ - \alpha_i^-\right)K\left(x_i, x_j\right) + b \tag{5}$$

where α_i^+, α_i^- are Lagrangian multipliers which are used to remove a few primary variables, and the term $K\left(x_i, x_j\right)$ is the kernel function. It has the advantage of independent of both dimensionalities of the input space X and the sample size m. The kernel function of the SVM technique allows nonlinear approximations. Linear function was the kernel function that was used in this analysis (L.F.). The most basic kernel function is as follows Han et al. (2007):

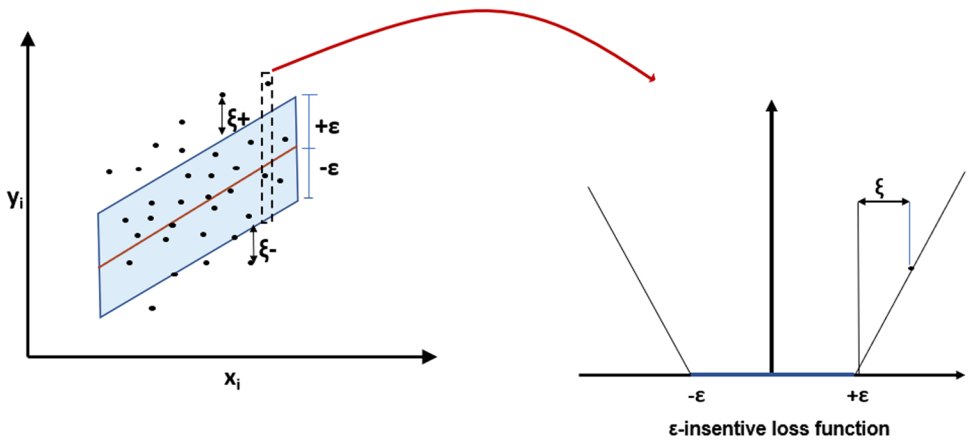

Fig. 4 Schematic presentation of the support vector machine structure

$$K(x_i, x_j) = (x_i, x_j) \tag{6}$$

The efficiency of the SVM techniques depends on the environment for an ε-insensitive loss of the function of three parameters of the training process (kernel, C, γ, and ε). For each kernel type, though, the values of C and ε affect the complexity of the final model. This value measures the number of support vectors (S.V.) used for projections. The greater value of ε intuitively results in fewer supporting vectors leading to less complex regression estimates. On the other hand, the value of C is a trade-off between model complexity and the variance allowed within the optimization formulation. As a result, a higher C value reduces model complexity (Cherkassky and Ma 2004). The optimal values for these training parameters (C and ε) ensure fewer complex models. This is an active research field.

Wavelet transform

Wavelet transform overcomes conventional solution issues by delivering the most potent way to dismantle signals into two-dimensional space: time or space (Sharghi et al. 2018). Like the Fourier transform, the wavelet transform allows for time conversion of the different frequency parts of a data set. However, with one crucial difference, the short-term Fourier transform produces a more accurate window width operation. Therefore, both the resolution of the time and the resulting transformations' frequency must be provisionally established. Still, in wavelet transform, the study will adjust its time width to the frequency. Higher-frequency waves become very narrow, while lower-frequency waves become very broad (Khan and Coulibaly 2006). The wavelet transform's ability to concentrate on brief intervals for high-frequency components and extended periods for low-frequency components improves signal processing of concentrated impulses and oscillation. As a result, wavelet decomposition is an excellent choice for evaluating transient signals and obtaining a more accurate comparison and discrimination process (Wang et al. 2008; Youssef 2003).

A continuous-time signal, $x(t)$, is transformed by the wavelet time-scale as Addison et al. (2001):

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} g^*\left(\frac{t-b}{a}\right) x(t) \cdot dt \tag{7}$$

where 'a' and 'b' denote the function's dilatation factor and temporal localization $g(t)$, respectively, to enable the study of the signal around 'b,' * denotes equivalent to the complex conjugation, and $g(t)$ denotes the wavelet or mother wavelet function (Youssef 2003).

Equation (7) discretization is perhaps the simplest discretization based on the trapezoidal law of a continuous wavelet transformation (CWT). From the given data set of length N, the above transformation method yields N^2 coefficients; thus, obsolete information is plugged inside the coefficients, which might or might not be desirable (Kişi 2010; Rajaee et al. 2011). For overcoming this complexity, uniform logarithmic spacing can be used for a correspondingly coarser resolution of b positions, allowing a complete definition of a signal length N by N transforming coefficients. A discrete wavelet of this kind has the following shape:

$$g_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} g\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \tag{8}$$

WANN and WSVM model

The present study used a discrete wavelet transformation to hybridize the model. Hybridized approaches for model creation and validation were used for the wavelet-based artificial neural network (WANN) (Kumar et al. 2016; Liu et al. 2013) and wavelet-based support vector machine (WSVM). To begin, the time-series data for the measured stage (h), discharge (Q), and sediment (S) were decomposed into many frequencies. Figure 5 represents the decomposed time series by wavelet transform (Sudheer and Jain 2003).

The decomposed components of time series by DWT like $h_{Di}(t)$, …, $h_{Di}(t-n)$, $h_{Ai}(t)$,.., $h_{Ai}(t-n)$, $Q_{Di}(t)$,…,

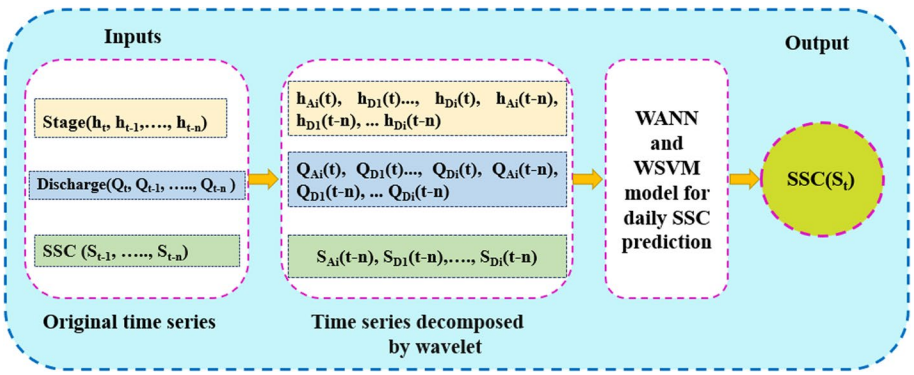

Fig. 5 Construction of proposed WANN and SVM model

$Q_{Di}(t-n)$, $Q_{Ai}(t)$,.., $Q_{Ai}(t-n)$, and $S_{Di}(t-1)$,..,$S_{Di}(t-n)$, $S_{Ai}(t-1)$,..,$S_{Ai}(t-n)$ were used for stage–discharge–sediment modeling, where, $h_{Di}(t)$, …, $h_{Di}(t-n)$ and $h_{Ai}(t)$,.., $h_{Ai}(t-n)$ are the details and approximate sub-signals of stage time series. $Q_{Di}(t)$,…, $Q_{Di}(t-n)$ and $Q_{Ai}(t)$,.., $Q_{Ai}(t-n)$ are the details and approximate sub-signals of discharge time series, respectively. $S_{Di}(t-1)$, …, $S_{Di}(t-n)$ and $S_{Ai}(t-1)$, …, $S_{Ai}(t-n)$ are the detail and approximation coefficients of SSC time series, respectively (Bajirao et al. 2021). The original h, Q, and SSC time series selected per the Gamma test were decomposed using H_{aar} á t_{rous} mother wavelet at appropriate decomposition levels. Afterward, these decomposed time-series values act as input for ANN and SVM techniques to predict the output value.

Multiple linear regression (MLR)

MLR is a form of linear regression analysis that involves more than one independent variable. The advantage of MLR is that it is simple, which shows how dependent variables are with independent variables (Choubin et al. 2018). The overall model of the MLR is:

$$y = c_0 + c_1x_1 + c_2x_2 + \ldots + c_nx_n \tag{9}$$

Y represents the dependent variable, and x_1, x_2,..,x_n refer to independent variables, c_1, c_2, …,c_n correspond to regression coefficients, and c_0 is intercepted. The least-square rule or regression rule is used to measure these values, representing the local actions (Kisi and Cimen 2011).

Gamma test

The Gamma test is a versatile and impartial method for assessing each input parameter's significant potential. Stefánsson et al. (1997) pioneered gamma testing in modeling, later adopted by other researchers (Kumar et al. 2016; Malik and Kumar 2015; Nourani et al. 2009). The Gamma test was used for any input–output dataset to estimate a minimum standard error for continuous nonlinear models. A linear regression line is built to measure gamma as:

$$\boldsymbol{Y} = \boldsymbol{A}\Delta + \Gamma \tag{10}$$

Y denotes the regression line's output vector, A represents the gradient, and Γ corresponds to the intercept (Δ = 0). The value of Γ corresponds to the output at Δ = 0. The smaller value of Γ (close to zero) is acceptable. The gamma test was processed in 'winGamma' software (Hassangavyar et al. 2020). The flowchart of the adopted methodology in this study is shown in Fig. 6.

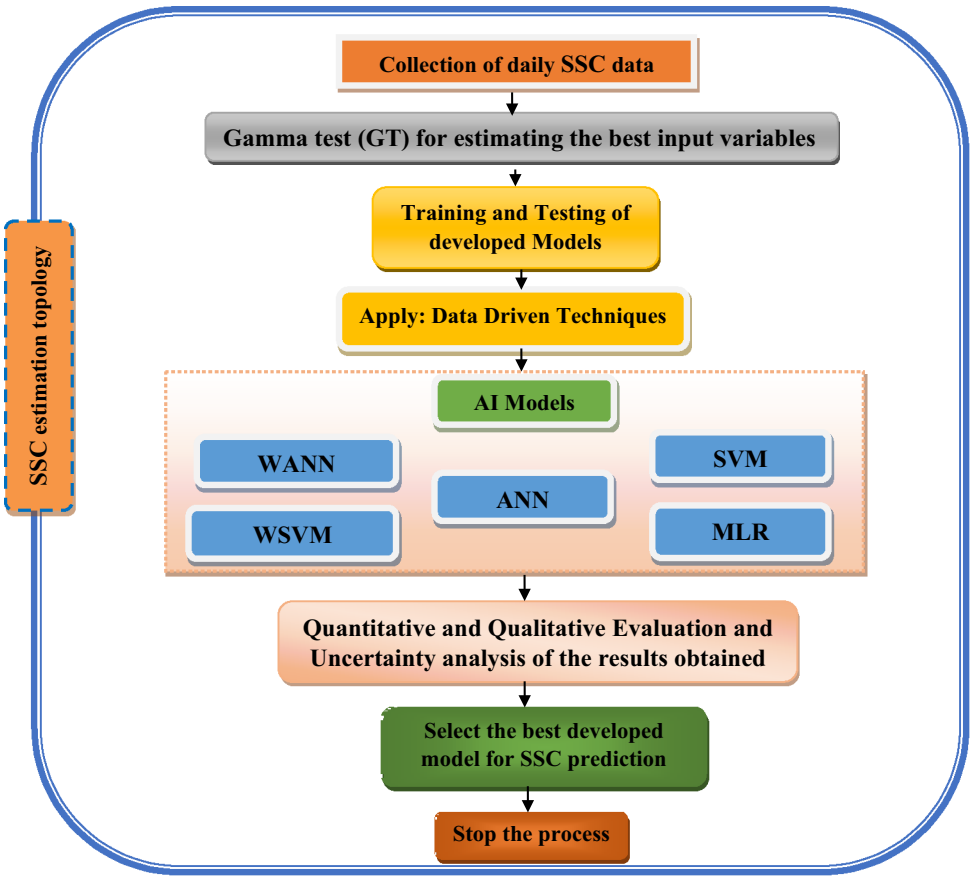

Fig. 6 Flowchart of SSC estimation methodology in the study area

Model development and performance evaluation

This research was undertaken to establish the relationship between stage–discharge–sediment for the Adityapur site of Jharkhand, India. The modeling included ANN, SVM, WAAN, WSVM, and MLR techniques to develop and validate the model. ANN and wavelet decomposed data were developed using MATLAB (R2015a) software. SVM models were developed in R-Studio and MLR models constructed in MS-Excel 2019 software. The model's performance was assessed using quantitative metrics (RMSE, PCC, and WI) and qualitative metrics (time variance map, scatter plot, and Taylor diagram) between observed and expected SSC (g/l) values. The input variables for the ANN, SVM, WAAN, WSVM, and MLR models developed were selected by gamma test based on minimum gamma value.

Three performance standards were used in the present study to assess the model's performance. These are Pearson correlation coefficient (PCC), root mean square error (RMSE), Nash–Sutcliffe efficiency, and Wilmot index (WI). The combined use of RMSE (Bajirao et al. 2021; Kumar et al. 2016; Malik and Kumar 2015) and WI (Willmott 1984) provides an adequate evaluation of the results. It compares the exactness of the various measurement and modeling techniques used in this study, further discussed by Ghorbani et al. (2013).

The PCC value ranges from -1 to $+1$, and the value close to $+1$ represents the best fit. Its aim in hydrological studies is to determine the degree of collinearity between observed and predicted variables. It is oversensitive to extreme value (Liu et al. 2013), from 0 to infinity, the RMSE value ranges. The value close to zero represents the model's better performance. The RMSE value has the same unit as the model output and reports the typical error size. The NSE was initially proposed by McCuen et al. (2006) and frequently used to assess the hydrologic model's performance. It determines the relative magnitude of residual variance compared to measured data variance. NSE values vary from $-\infty$ to 1. The WI value ranges from 0 to 1. The values close to 1 represent the best fit, while 0 means disagreement between observed and predicted data. It is also known as the index of agreement.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(S_{p_{\mathbf{obs},i}} - S_{p_{\mathbf{pre},i}}\right)^2}{N}} \tag{11}$$

$$\mathrm{PCC} = \frac{\sum_{i=1}^{N}\left(S_{p_{\mathrm{obs},i}} - \overline{S}_{p_{\mathrm{obs},i}}\right)\left(S_{p_{\mathrm{pre},i}} - \overline{S}_{p_{\mathrm{pre},i}}\right)}{\sqrt{\sum_{i=1}^{N}\left(S_{p_{\mathrm{obs},i}} - \overline{S}_{p_{\mathrm{obs},i}}\right)^2 \sum_{i=1}^{N}\left(S_{p_{\mathrm{pre},i}} - \overline{S}_{p_{\mathrm{pre},i}}\right)^2}} \tag{12}$$

$$\mathrm{NSE} = 1 - \left[\frac{\sum_{i=1}^{N}\left(S_{p_{\mathbf{obs},i}} - S_{p_{\mathbf{pre},i}}\right)^2}{\sum_{i=1}^{N}\left(S_{p_{\mathbf{obs},i}} - \overline{S}_{p_{\mathbf{pre},i}}\right)^2}\right] \tag{13}$$

$$\mathrm{WI} = 1 - \frac{\sum_{i=1}^{N}\left(S_{p_{\mathbf{obs},i}} - S_{p_{\mathbf{pre},i}}\right)^2}{\sum_{i=1}^{N}\left(\left|S_{p_{\mathrm{pre},i}} - \overline{S}_{p_{\mathrm{obs},i}}\right| + \left|\overline{S}_{p_{\mathrm{obs},i}} - \overline{S}_{p_{\mathrm{obs},i}}\right|\right)^2} \tag{14}$$

Results and discussion

This section discusses the outcomes of dividing the stage–discharge–sediment model into two sections. The first section contains the results of the gamma test used to pick input variables, and the second section contains the results of model creation and output for both the Adityapur and Ghatshila sites.

Input selection: gamma test

The first step in every modeling process is to choose input variables. Many scientists have stated that the present-day suspended sediment concentration (SSC) can be estimated more accurately by the simultaneous current day stage (h), discharge (Q), along with previous day h, Q, and SSC values (Cigizoglu 2004; Jain 2012). Therefore, current day (t), one-day lag ($t-1$), two-day lag ($t-2$), and three-day lag ($t-3$) time steps h_t, h_{t-1}, h_{t-2}, h_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}, and S_{t-3} which is represented by model 48 (mask-11111111111) and 0 represents the absence of that variables in other combinations (Table 2). A total of forty-eight combinations were created using different time steps of the stage, discharge, and SSC data, as seen in Tables 3 and 4 for the Adityapur and Ghatshila sites. The gamma value (Γ), variance ratio (V_{ratio}), and mask for various combinations of input variables for model creation are shown in Tables 3 and 4.

The selection process depends on the smallest value of Γ and V_{ratio} (Jain 2012; Malik et al. 2019; Nourani et al. 2018). V_{ratio} tests its predictability with the available inputs for the specified output. V_{ratio} near 1 indicates that the fundamental model is not quite close to being smooth. However, V_{ratio} near 0 demonstrates that the results are generated from the smooth model (Malik et al. 2019). As shown in Table 3, the integration of $h_t + h_{t-1} + Q_t + Q_{t-1} + Q_{t-2} + S_{t-1}$ (Model no.-19) showed the most negligible value of Γ and V_{ratio} as 0.0813 and 0.3253, respectively.

Therefore, the combination (mask-11001110100) of $h_t + h_{t-1} + Q_t + Q_{t-1} + Q_{t-2} + S_{t-1}$ was selected as input variables for ANN, SVM, WAAN, WSVM, and MLR models for

Table 3 Findings of the gamma test obtained for different combinations of input variables for the Adityapur site

Model no.	Mask (combination)	Gamma	V-ratio
M1	10001000100	0.0945	0.3781
M2	10001000110	0.0938	0.3753
M3	10001000111	0.0867	0.3467
M4	10001100100	0.0901	0.3603
M5	10001100110	0.0909	0.3638
M6	10001100111	0.0861	0.3444
M7	10001110100	0.0911	0.3643
M8	10001110110	0.0891	0.3564
M9	10001110111	0.0842	0.3367
M10	10001111100	0.0897	0.3587
M11	10001111110	0.0912	0.3647
M12	10001111111	0.0827	0.3308
M13	11001000100	0.0833	0.3334
M14	11001000110	0.0901	0.3604
M15	11001000111	0.0877	0.3510
M16	11001100100	0.0911	0.3646
M17	11001100110	0.0915	0.3660
M18	11001100111	0.0904	0.3618
M19	**11001110100**	**0.0813**	**0.3253**
M20	11001110110	0.0904	0.3615
M21	11001110111	0.0914	0.3657
M22	11001111100	0.0847	0.3390
M23	11001111110	0.0972	0.3887
M24	11001111111	0.0863	0.3454
M25	11101000100	0.0900	0.3599
M26	11101000110	0.0843	0.3370
M27	11101000111	0.0840	0.3360
M28	11101100100	0.0844	0.3376
M29	11101100110	0.0932	0.3729
M30	11101100111	0.0908	0.3633
M31	11101110100	0.0901	0.3606
M32	11101110110	0.0920	0.3682
M33	11101110111	0.0897	0.3589
M34	11101111100	0.0900	0.3599
M35	11101111110	0.1031	0.4125
M36	11101111111	0.0875	0.3500
M37	11111000100	0.0920	0.3678
M38	11111000110	0.1012	0.4050
M39	11111000111	0.0846	0.3386
M40	11111100100	0.0858	0.3433
M41	11111100110	0.1026	0.4103
M42	11111100111	0.0947	0.3787
M43	11111110100	0.0890	0.3560
M44	11111110110	0.1035	0.4138
M45	11111110111	0.0905	0.3620
M46	11111111100	0.0985	0.3942
M47	11111111110	0.1085	0.4339
M48	11111111111	0.0926	0.3704

$S_t = f(h_t, h_{t-1}, h_{t-2}, h_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}, S_{t-3}) = f(1111111111-1)$

Bold represents the best input combination

Adityapur site. Likewise, Table 4 showed the combination $h_t + Q_t + S_{t-1} + S_{t-2} + S_{t-3}$ (Model-3) observed the minimum values of Γ and V_{ratio} as 0.0046 and 0.1853, respectively. As a result, for the Ghatshila location, the combination (mask-10001000111) of $h_t + Q_t + S_{t-1} + S_{t-2} + S_{t-3}$ variables was chosen as input variables for ANN, SVM, WAAN, WSVM, and MLR models. As shown in Fig. 7, the correlation between output *St* and other input variables was satisfactory for all datasets (*a*, *b*).

Trials of models

The ANN, WANN, SVM, WSVM, and MLR were analyzed in two phases to select the best model. The first phase involves developing the model during the training phase—the second phase checks to validate the model. The model's performance was evaluated based on the lower value of RMSE (0: + ∞: good: poor), a higher value of PCC, NSE, and WI (close to + 1) for selections of the best model (Kumar et al. 2020). Several trials were conducted for single output on the best model for ANN, WANN, SVM, and WSVM (Tables 5 and 6). The number of neurons in hidden layers was varied in ANN trials. An input layer, a hidden layer, and an output layer make up the ANN architecture. Considering architecture 6-4-1, 6 represents the number of input parameters, and the number of neurons in the un-seen layer is 4. The output is 1.

The 24 represents the input parameters in the WANN architecture, 9 represents the number of hidden layer neurons, and 1 represents the output. Simultaneously, SVM trails were run using a variety of SVM-γ, SVM-*c*, and SVM-ε parameter values. All sites' cost parameters (SVM-*c*) were taken as 10 based on separate trials, whereas ε is an insensitive loss feature. Training effects were not taken into account in this analysis to prevent the biases and overfitting of data.

Results at Adityapur site

At the Adityapur location, Table 5 displayed the quantitative results of all produced models. For the training period, PCC values ranged from 0.783 to 0.801, RMSE values ranged from 0.106 to 0.111 g/l, and NSE values were found in the range of 0.613 to 0.644 WI values ranged from 0.866 to 0.894. During the testing process, PCC values ranged from 0.562 to 0.632, RMSE values ranged from 0.097 to 0.106 g/l, NSE values ranged from − 0.425 to − 0.216, and WI values ranged from 0.690 to 0.729 for ANN techniques. During the training phase, the performance of WANN models showed that the PCC values were obtained in the range of 0.835–0.863. RMSE values were obtained in the range 0.090–0.099, the values of NSE obtained in the range of 0.691–0.745 while for WI were 0.899–0.921. The PCC

Table 4 Findings of the gamma test obtained for different combinations of input variables for the Ghatshila site

Model no.	Mask (combination)	Gamma	V-ratio
M1	10001000100	0.0473	0.1891
M2	10001000110	0.0407	0.1626
M3	**10001000111**	**0.0046**	**0.1853**
M4	10001100100	0.0416	0.1662
M5	10001100110	0.0501	0.2003
M6	10001100111	0.0485	0.1938
M7	10001110100	0.0506	0.2025
M8	10001110110	0.5008	0.2003
M9	10001110111	0.0485	0.1938
M10	10001111100	0.0506	0.2025
M11	10001111110	0.0505	0.2019
M12	10001111111	0.0486	0.1944
M13	11001000100	0.0472	0.1888
M14	11001000110	0.0477	0.1906
M15	11001000111	0.0488	0.1950
M16	11001100100	0.0442	0.1769
M17	11001100110	0.0503	0.2010
M18	11001100111	0.0458	0.1833
M19	11001110100	0.0424	0.1695
M20	11001110110	0.0472	0.1887
M21	11001110111	0.0467	0.1887
M22	11001111100	0.0430	0.1720
M23	11001111110	0.0512	0.2050
M24	11001111111	0.0463	0.1853
M25	11101000100	0.0469	0.1878
M26	11101000110	0.0456	0.1826
M27	11101000111	0.0444	0.1776
M28	11101100100	0.5112	0.2045
M29	11101100110	0.0492	0.1969
M30	11101100111	0.0467	0.1870
M31	11101110100	0.0456	0.1822
M32	11101110110	0.0404	0.1618
M33	11101110111	0.0504	0.2014
M34	11101111100	0.0472	0.1889
M35	11101111110	0.0489	0.1956
M36	11101111111	0.0443	0.1773
M37	11111000100	0.0427	0.1708
M38	11111000110	0.0430	0.1719
M39	11111000111	0.0437	0.1747
M40	11111100100	0.0486	0.1945
M41	11111100110	0.0463	0.1851
M42	11111100111	0.0416	0.1663
M43	11111110100	0.0457	0.1826
M44	11111110110	0.0433	0.1731
M45	11111110111	0.0410	0.1639
M46	11111111100	0.0368	0.1474
M47	11111111110	0.0402	0.1607
M48	11111111111	0.0434	0.1735

$S_t = f(h_t, h_{t-1}, h_{t-2}, h_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}, S_{t-3}) = f(1111111111-1)$

Bold represents the best input combination

values for WANN models ranged from 0.634 to 0.718, the RMSE values ranged from 0.071 to 0.078 g/l, the NSE values ranged from 0.232 to 0.356, and the WI ranged values from 0.767 to 0.812 during the testing process. During the training period of SVM models, PCC values ranged from 0.760 to 0.768, RMSE values ranged from 0.114 to 0.117, NSE values ranged from 0.568 to 0.586, and WI values ranged from 0.857 to 0.859. During the SVM models' performance testing phase, the values of PCC ranged from 0.572 to 0.610, the values of RMSE were obtained in the range of 0.086–0.094, the NSE values ranged from − 0.139 to 0.046, and WI values ranged from 0.724 to 0.760. The performance of the hybrid techniques of wavelet support vector machine (WSVM) during training phase, the values of PCC ranged from 0.844 to 0.847. RMSE was obtained around 0.095 to 0.096 g/l, the values of NSE ranged from 0.711 to 0.714, and WI values ranged from 0.906 to 0.907. During the testing period, PCC values ranged from 0.745 to 0.781, RMSE ranged from 0.057 to 0.062, NSE ranged from 0.516 to 0.591, and WI varied from 0.856 to 0.878.

Table 5 shows that the WSVM model found the most reliable models out of all the existing models. During the training and testing processes, PCC, RMSE, NSE, and WI values were 0.844 and 0.781, 0.096 g/l and 0.057 g/l, 0.711 and 0.591 0.907 and 0.878, respectively. In contrast to other models, the NSE values for the WSVM model during the testing process significantly increased. The MLR model did not do well.

Figures 8a–e and 9a–e show the line diagram and scatter plots for all created models. These figures qualitatively represent the results of all developed models. On nearly all simulations, the expected values were over-predicted for lower SSC values and under-predicted for higher SSC values. The determination coefficient was the highest for the WSVM model (0.6096), followed by WANN values (0.5159). The R^2 value was poorly obtained for the MLR model (0.2890).

The sequence of models results from best to poor in order WSVM > WANN > ANN > SVM > MLR for Adityapur site. Hence, the WSVM model can be used to estimate SSC for the Adityapur site.

Results at Ghatshila site

Table 6 shows the results of the various performance metrics that were used to choose the best model. Table 6 shows that during training and testing, the values of r, RMSE (g/l), NSE, and WI ranged from 0.906 to 0.919 and 0.548 to 0.580, 0.054 to 0.131 and 0.131 to 0.139, 0.819 to 0.843 and 0.030 to 0.125, and 0.946 to 0.956 and 0.722 to 0.746, respectively. A-3 model with architecture (5-9-1) was observed to be superior model as compared to other ANN models.

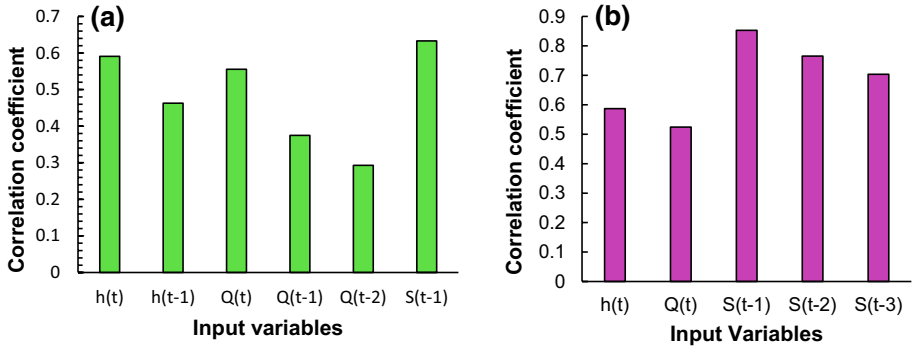

Fig. 7 Correlation graph of input variables for all dataset with output S_t. at **a** Adityapur, **b** Ghatshila

Likewise, the values of *r*, RMSE (g/l), NSE, and WI for WANN model varied from 0.940 to 0.950 and 0.703 to 0.725, 0.099 to 0.107 and 0.099 to 0.115, 0.884 to 0.902 and 0.333 to 0.500, and 0.968 to 0.971 and 0.827 to 0.845, respectively, during training and testing phases. W-1 model with architecture (20-3-1) found best among all WANN models. From Table 5, it is clear that the values of r, RMSE, NSE, and WI for SVM ranged from 0.886 to 0.891 and 0.579 to 0.586, 0.144 to 0.148 g/l and 0.125 to 0.128 g/l, 0.779 to 0.791 and 0.177 to 0.206, and 0.940 to 0.942 and 0.748 to 0.753, respectively, during training and testing. S-1 with structure ($C = 10$, $\gamma = 0.2$, $\varepsilon = 0.1$) found superior among SVM. During the training and testing phases of the hybridized wavelet SVM model, the values for r, RMSE, NSE, and WI ranged from 0.928 to 0.929 and 0.749 to 0.751, 0.116 to 0.117 g/l and 0.095 to 0.096, 0.861 to 0.862 and 0.538 to 0.543, and 0.962 and 0.859, respectively. Of all WSVM models, WS-1 with structure ($C = 10, = 0.05$, $\varepsilon = 0.1$) was found to be superior.

Table 6 reveals that the WSVM model observed the most accurate models among all developed models at the Ghatshila site. PCC, RMSE, NSE, and WI values were obtained as 0.928 and 0.751, 0.117 g/l and 0.095 g/l, 0.861 and 0.541, and 0.962 and 0.859, respectively, during training and testing phases. The results also showed that the NSE values for the WSVM model during the testing phase significantly improved compared to other models.

The line diagram and scatter plots at the Ghatshila site for all created models are shown in Figs. 10a–e and 11a–e. The qualitative results from the figures showed that the predicted values were over-predicted and under-predicted for SSC values for almost all models. The R^2 values obtained maximum for WSVM model (0.5639) followed by WANN values (0.5250) and then followed by MLR model (0.2890).

Thus, based on the obtained results discussed above for the Ghatshila site, the wavelet hybridized model (WSVM and WANN) outperformed all other models by a large margin. The sequence of the best to poor performance of models given as WSVM > WANN > MLR > SVM > ANN.

The comparative results of each of the best-developed Adityapur and Ghatshila sites are shown in Table 7. This table shows that the wavelet hybridized model was found to be superior to all other models. It is because of the application of the wavelet transform that may find various sub-series of the primary time series data that have extra information obscured by the original time series data. Wavelet transform improves the model performance because it simultaneously considers both time and frequency information available within the signal (Nourani et al. 2009).

Our results are similar to Nourani et al. (2018), who applied the SVM technique with different input combinations to predict monthly suspended sediment load and stated that the correlation coefficient values ranged from 0.49 to 0.91 RMSE varied from 0.015 to 0.9. Furthermore, for the SVM in sediment yield prediction, our findings agree with Kumar et al. (2016), who found that correlation coefficients ranged from 0.66 to 0.90 for training models and from 0.24 to 0.93 for the testing phase. For SVM, the findings of this model are close and in line with Choubin et al. (2018), who concluded that the SVM gave correlations varied from 0.43 to 0.67 under different combinations for forecasting the suspended sediment. Moreover, Sharafati et al. (2020) used WSVM and SVM models in sediment yield (SY) modeling. They observed that the WSVM model produced better results than other algorithms. Their outcomes are acceptable and agree with our results. They showed that the use of WSVM is a more reliable alternative to conventional SY models. Besides, the efficiency of a wavelet-based model in the prediction of suspended sediment load was investigated by Nourani et al. (2009). They used WSVM and WANN wavelet complementary versions, respectively. Their outputs pointed out that the WSVM integrated model generated more reliable results than WANN. Liu et al. (2013) developed a WANN complement model. When the findings

Table 5 Performance indicators of ANN, WANN, SVM, WSVM, and MLR models during the training and testing at the Adityapur site

Model	Structure	Dataset	PCC	RMSE (g/l)	NSE	WI
ANN-A						
A-1	6-4-1	Train	0.783	0.111	0.613	0.866
		Test	0.568	0.103	-0.357	0.699
A-2	6-8-1	Train	0.799	0.107	0.636	0.884
		Test	0.567	0.098	-0.216	0.693
A-3	6-13-1	Train	0.810	0.106	0.644	0.894
		Test	0.562	0.106	-0.425	0.690
A-4	6-14-1	Train	0.849	0.094	0.721	0.913
		Test	0.591	0.099	-0.210	0.721
A-5	**6-15-1**	**Train**	**0.801**	**0.109**	**0.628**	**0.888**
		Test	**0.632**	**0.097**	**-0.252**	**0.729**
WANN-A						
W-1	**24-9-1**	**Train**	**0.853**	**0.094**	**0.722**	**0.918**
		Test	**0.718**	**0.071**	**0.356**	**0.812**
W-2	24-10-1	Train	0.848	0.095	0.717	0.908
		Test	0.634	0.078	0.232	0.772
W-3	24-15-1	Train	0.863	0.090	0.745	0.921
		Test	0.645	0.072	0.339	0.776
W-4	24-16-1	Train	0.835	0.099	0.691	0.899
		Test	0.663	0.075	0.278	0.767
SVM-A						
S-1	$\gamma = 0.9$, $\varepsilon = 0.01$	Train	0.760	0.117	0.568	0.859
		Test	0.609	0.087	0.043	0.759
S-2	$\boldsymbol{\gamma = 0.1667}$, $\varepsilon = 0.1$	**Train**	**0.760**	**0.117**	**0.569**	**0.859**
		Test	**0.610**	**0.086**	**0.046**	**0.760**
S-3	$\gamma = 0.1667$, $\varepsilon = 0.01$	Train	0.768	0.114	0.586	0.857
		Test	0.572	0.094	-0.139	0.724
S-4	$\gamma = 0.1667$, $\varepsilon = 0.001$	Train	0.767	0.115	0.584	0.857
		Test	0.590	0.091	-0.026	0.741
WSVM-A						
WS-1	$\boldsymbol{\gamma = 0.04167}$, $\boldsymbol{\varepsilon = 0.1}$	**Train**	**0.844**	**0.096**	**0.711**	**0.907**
		Test	**0.781**	**0.057**	**0.591**	**0.878**
WS-2	$\gamma = 0.04167$, $\varepsilon = 0.01$	Train	0.844	0.095	0.712	0.907
		Test	0.777	0.057	0.582	0.875
WS-3	$\gamma = 0.04167$, $\varepsilon = 0.001$	Train	0.847	0.095	0.714	0.906
		Test	0.745	0.062	0.516	0.856
WS-4	$\gamma = 0.1$, $\varepsilon = 0.1$	Train	0.845	0.095	0.713	0.906
		Test	0.764	0.059	0.556	0.867
MLR-A		Train	0.897	0.112	0.606	0.862
		Test	0.592	0.111	-0.563	0.675

Bold represents the best model among developed models

from WANN and ANN were compared, it was discovered that the WANN model could better forecast the extremely nonlinear time series than ANN. Also, as reported by, Partal and Cigizoglu (2008), the wavelet-ANN model demonstrated higher levels of accuracy than both the conventional ANN and the SRC. The results show that wavelet-ANN is capable of capturing better approximations for peak values. Jain (2012) applied ANN, fuzzy logic, and evolutionary algorithms in river stage–discharge–sediment rating modeling.

Table 6 Performance indicators of ANN, WANN, SVM, WSVM, and MLR models during the training and testing at the Ghatshila site

Model	Structure	Dataset	PCC	RMSE (g/l)	NSE	WI
ANN-G						
A-1	5-4-1	Train	0.919	0.054	0.843	0.956
		Test	0.558	0.136	0.066	0.733
A-2	5-5-1	Train	0.906	0.065	0.819	0.946
		Test	0.562	0.131	0.125	0.736
A-3	**5-9-1**	**Train**	**0.909**	**0.131**	**0.825**	**0.948**
		Test	**0.580**	**0.137**	**0.055**	**0.746**
A-4	5-10-1	Train	0.914	0.128	0.834	0.952
		Test	0.548	0.139	0.030	0.722
WANN-G						
W-1	**20-3-1**	**Train**	**0.945**	**0.103**	**0.892**	**0.971**
		Test	**0.725**	**0.103**	**0.462**	**0.845**
W-2	20-5-1	Train	0.950	0.099	0.902	0.974
		Test	0.703	0.115	0.333	0.827
W-3	20-10-1	Train	0.944	0.105	0.889	0.969
		Test	0.718	0.099	0.500	0.833
W-4	20-15-1	Train	0.940	0.107	0.884	0.968
		Test	0.716	0.100	0.491	0.829
SVM-G						
S-1	$\boldsymbol{\gamma = 0.2, \varepsilon = 0.1}$	**Train**	**0.891**	**0.144**	**0.791**	**0.942**
		Test	**0.586**	**0.125**	**0.206**	**0.753**
S-2	$\gamma = 0.2, \varepsilon = 0.01$	Train	0.888	0.147	0.782	0.941
		Test	0.582	0.125	0.180	0.751
S-3	$\gamma = 0.2, \varepsilon = 0.001$	Train	0.886	0.148	0.779	0.940
		Test	0.579	0.128	0.177	0.748
S-4	$\gamma = 0.5, \varepsilon = 0.1$	Train	0.886	0.148	0.779	0.940
		Test	0.579	0.128	0.177	0.748
WSVM-G						
WS-1	$\boldsymbol{\gamma = 0.05, \varepsilon = 0.1}$	**Train**	**0.928**	**0.117**	**0.861**	**0.962**
		Test	**0.751**	**0.095**	**0.543**	**0.859**
WS-2	$\gamma = 0.05, \varepsilon = 0.01$	Train	0.928	0.117	0.861	0.962
		Test	0.751	0.095	0.540	0.859
WS-3	$\gamma = 0.05, \varepsilon = 0.001$	Train	0.929	0.116	0.862	0.962
		Test	0.749	0.096	0.538	0.858
WS-4	$\gamma = 0.5, \varepsilon = 0.1$	Train	0.928	0.117	0.861	0.962
		Test	0.751	0.095	0.543	0.859
MLR-G		Train	0.897	0.139	0.805	0.944
		Test	0.592	0.129	0.163	0.752

Bold represents the best model among developed models

When the findings of SVR were compared to those of ANNs, it was discovered that SVR outperformed ANNs.

The results for both the sites are also verified from the Taylor diagram. The Taylor diagram consists of a line designated as a straight line and standard deviation and root-mean-square difference (RMSD) designated as curvilinear (Fig. 12). Because of the highest correlation values and lower standard deviation and RMSD values for both locations, the findings of the wavelet hybridized model were found to be superior. The Taylor diagram also shows that the WSVM model got closer to the observed SSC values for both locations. On both sites, the Taylor diagram yields the same series of models as previously mentioned.

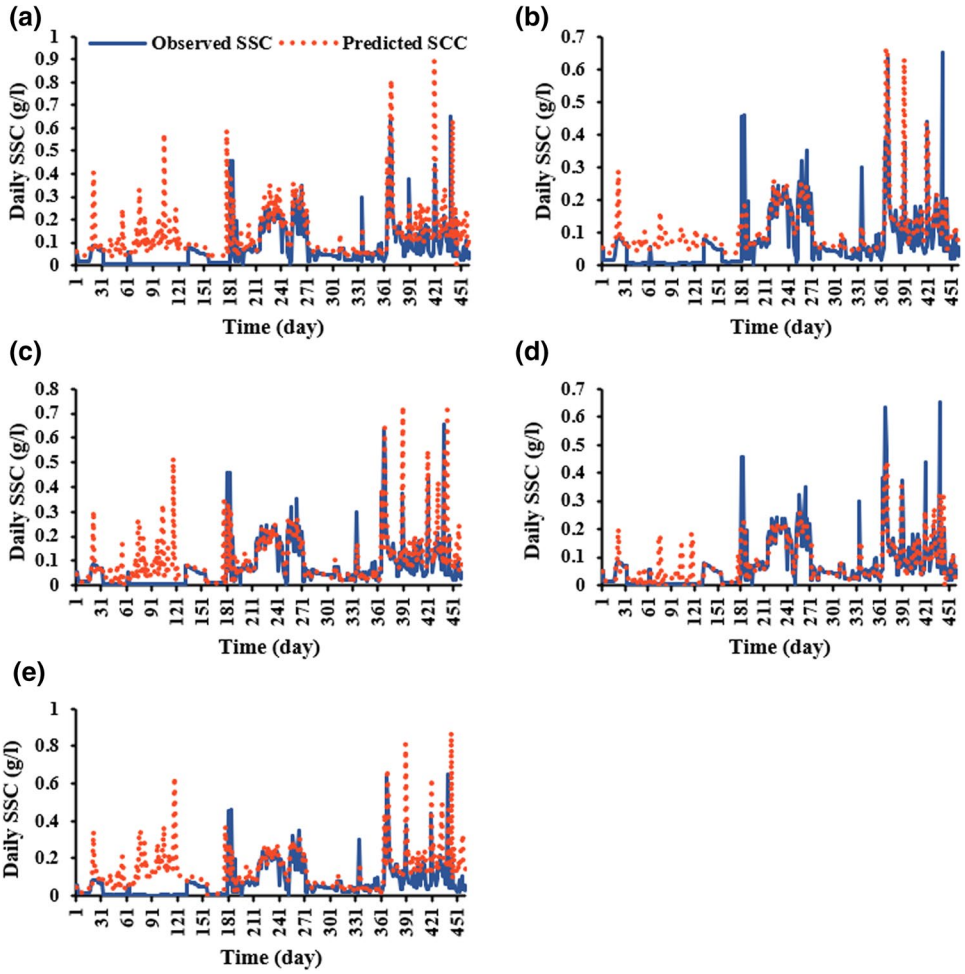

Fig. 8 Line diagram of developed models **a** ANN, **b** WANN, **c** SVM, **d** WSVM and **e** MLR during the testing phase for Adityapur sites

Conclusions

Estimating river sediment volumes is vital for measuring river sediment flow, designing dams, storage structures, and canals, evaluating environmental effects, and deciding the effectiveness of watershed management and other catchment treatments. In the present analysis, daily SSC model estimation was studied at the Adityapur site and Ghatshila site in the Saraikela Kharsawan district of Jharkhand, India. Hydrological datasets containing the daily stage (h), discharge (Q), and SSC for 10 years (2004–2013) period from June to October were taken for analysis. Five data-driven approaches, namely artificial neural network (ANN), support vector machine (SVM), wavelet-based artificial neural network (WANN), wavelet-based support vector machine (WSVM), and multi-linear regression (MLR) techniques were employed for modeling SSC for the study area. The gamma test was used for selecting input variables for the model, as mentioned earlier. The combination showed the most negligible value of Γ and V_{ratio} as 0.0813 and 0.3253, respectively, for input combinations based on the gamma test $h_t + h_{t-1} + Q_t + Q_{t-1} + Q_{t-2} + S_{t-1}$ (mask-11001110100). Likewise, the combination $h_t + Q_t + S_{t-1} + S_{t-2} + S_{t-3}$ (mask-10001000111) observed the minimum values of Γ, and V_{ratio} as 0.0046 and 0.1853, respectively. Therefore, it was considered as input variables for modeling. The performance

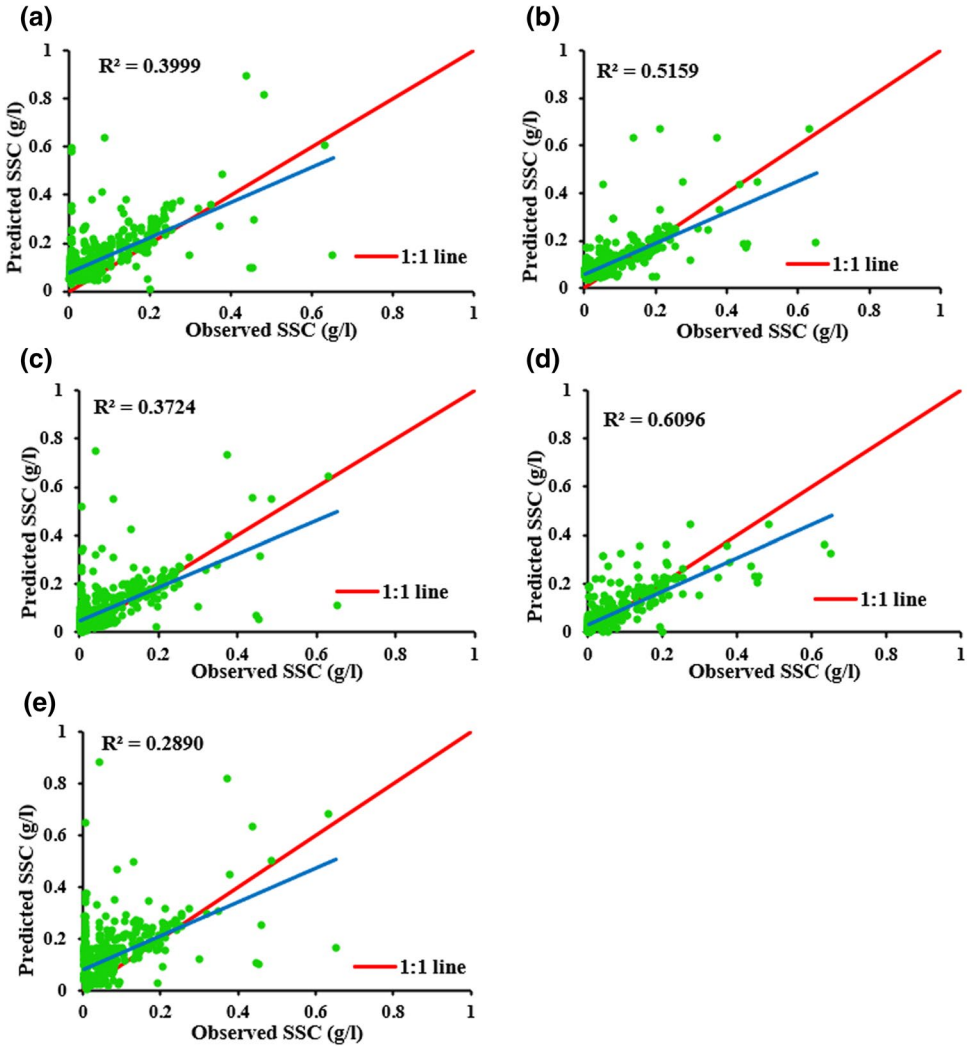

Fig. 9 Scatter plots of developed models **a** ANN, **b** WANN, **c** SVM, **d** WSVM and **e** MLR during the testing phase for Adityapur sites

of the model was evaluated through quantitative indicators (RMSE, PCC, and WI) and qualitative indicators (time variance map, scatter plot, and Taylor diagram) between actual and expected SSC (g/l) values. According to our findings, the WSVM model was the most reliable model among all existing models. Throughout the training and testing operations at the Adityapur location, PCC, RMSE, NSE, and WI values were 0.844 and 0.781, 0.096 g/l and 0.057 g/l, 0.711 and 0.591, and 0.907 and 0.878, respectively. It was also the most precise model on the Ghatshila site. During the training and testing stages, the PCC, RMSE, NSE, and WI values were 0.928 and 0.751, 0.117 g/l and 0.095 g/l, 0.861 and 0.541, and 0.962 and 0.859, respectively. The WSVM model outperformed the ANN, WANN, SVM, and model MLR models. The wavelet hybridized model (WSVM and WANN) performed better at both locations than the non-wavelet hybridized model. Also, the WSVM and WANN models' best performance can assist researchers' in the future in using extremely variable SSC data for such modeling.

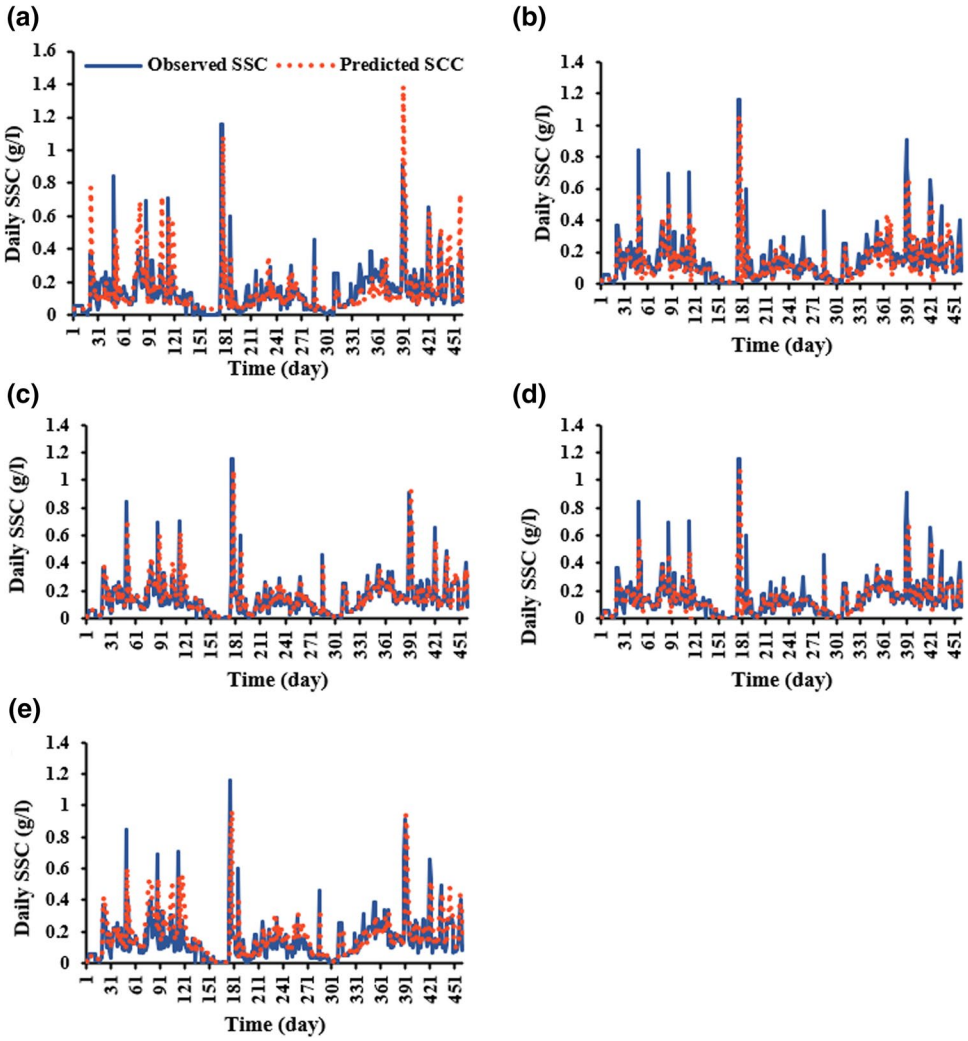

Fig. 10 Line diagram of developed models **a** ANN, **b** WANN, **c** SVM, **d** WSVM and **e** MLR during the testing phase for Ghatshila sites

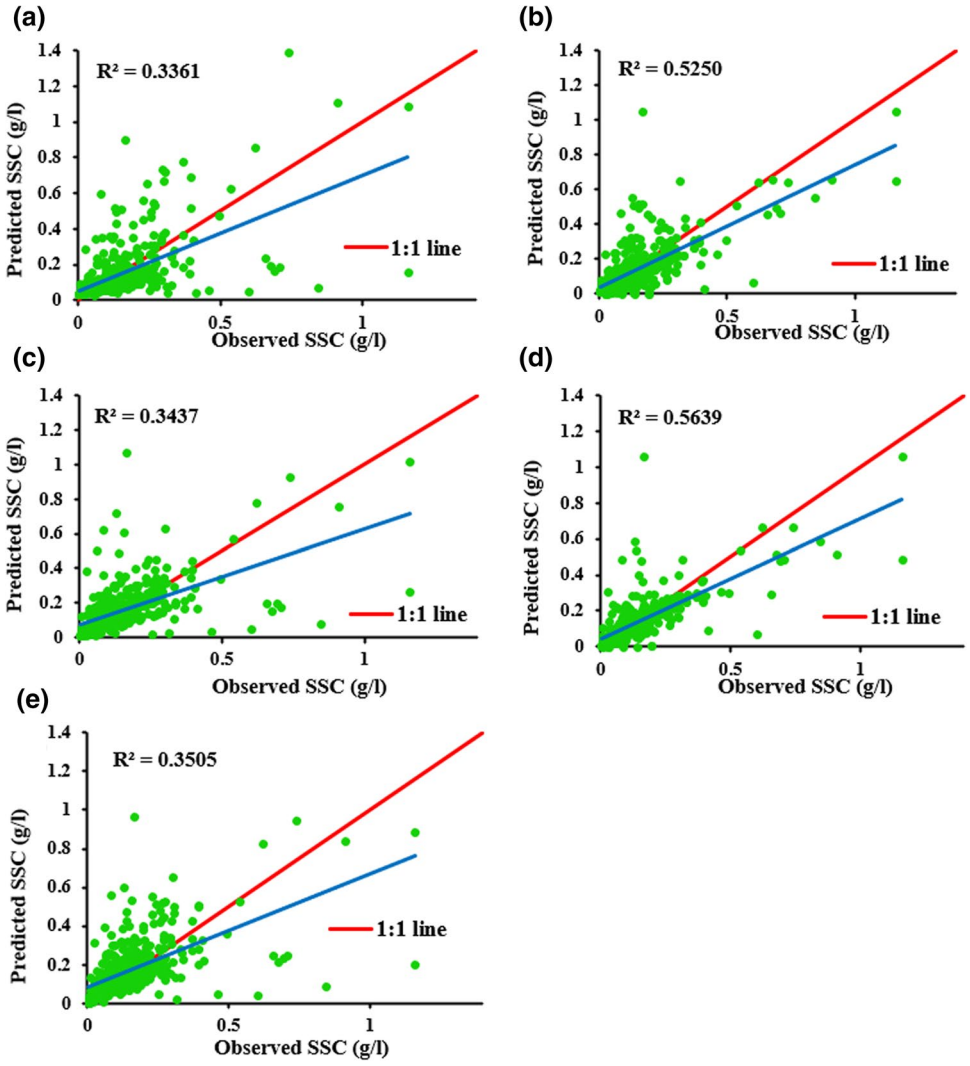

Fig. 11 Scatter plots of developed models **a** ANN, **b** WANN, **c** SVM, **d** WSVM and **e** MLR during testing for Adityapur sites

Table 7 Comparative results of models during testing for both sites

Model	PCC	RMSE (g/l)	NSE	WI
Adityapur				
A-5	0.632	0.097	−0.252	0.729
W-1	0.718	0.071	0.356	0.812
S-2	0.610	0.086	0.046	0.760
WS-1	0.781	0.057	0.591	0.878
MLR	0.592	0.111	−0.563	0.675
Ghatshila				
A-3	0.580	0.137	0.055	0.746
W-1	0.725	0.103	0.462	0.845
S-1	0.586	0.125	0.206	0.753
WS-1	0.751	0.095	0.543	0.859
MLR	0.592	0.129	0.163	0.752

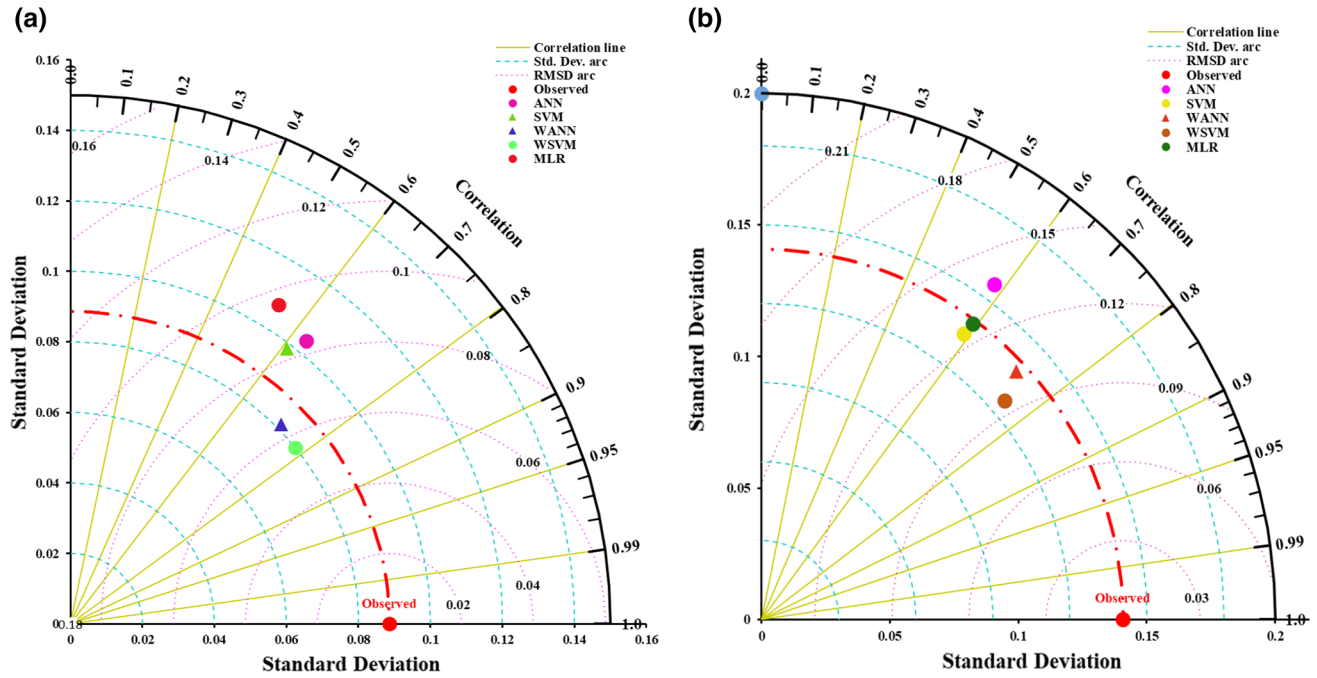

Fig. 12 Taylor diagram of ANN, SVM, WANN, WSVM, and MLR models during the testing period at **a** Adityapur site and **b** Ghatshila site

Acknowledgements The author would like to sincerely thank the Technical Education Quality Improvement Programme (TEQIP-III) for continuous financial assistance during the research work. Alban Kuriqi acknowledges the Portuguese Foundation for Science and Technology (FCT) support through PTDC/CTA-OHR/30561/2017 (WinTherface).

Funding The authors received no specific funding for this work.

Data availability The data and materials of analysis should be available from the corresponding author.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

Consent to participate The manuscript has been read and approved for submission by all the named authors.

Consent to publish Yes, there is consent to publish this paper.

Ethical approval All the contents of this paper are unique and with minimum plagiarism. This manuscript is original, has not been published before, and is not currently considered elsewhere.

References

Addison PS, Murray KB, Watson JN (2001) Wavelet transform analysis of open channel wake flows. J Eng Mech 127:58–70. https://doi.org/10.1061/(ASCE)0733-9399(2001)127:1(58)

Adib A, Mahmoodi A (2017) Prediction of suspended sediment load using ANN GA conjunction model with Markov chain approach at flood conditions. KSCE J Civ Eng 21:447–457. https://doi.org/10.1007/s12205-016-0444-2

Ajmera TK, Goyal MK (2012) Development of stage–discharge rating curve using model tree and neural networks: an application to Peachtree Creek in Atlanta. Expert Syst Appl 39:5702–5710. https://doi.org/10.1016/j.eswa.2011.11.101

Alp M, Cigizoglu HK (2007) Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ Model Softw 22:2–13. https://doi.org/10.1016/j.envsoft.2005.09.009

Asselman NEM (2000) Fitting and interpretation of sediment rating curves. J Hydrol 234:228–248. https://doi.org/10.1016/S0022-1694(00)00253-5

Bajirao TS, Kumar P, Kumar M, Elbeltagi A, Kuriqi A (2021) Superiority of hybrid soft computing models in daily suspended sediment estimation in highly dynamic rivers. Sustainability 13:542

Banadkooki FB et al (2020) Suspended sediment load prediction using artificial neural network and ant lion optimization algorithm. Environ Sci Pollut Res 27:38094–38116. https://doi.org/10.1007/s11356-020-09876-w

Cao LJ, Tay FEH (2003) Support vector machine with adaptive parameters in financial time series forecasting. IEEE Trans Neural Netw 14:1506–1518. https://doi.org/10.1109/TNN.2003.820556

Cherkassky V, Ma Y (2004) Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw 17:113–126. https://doi.org/10.1016/S0893-6080(03)00169-2

Choubin B, Darabi H, Rahmati O, Sajedi-Hosseini F, Kløve B (2018) River suspended sediment modelling using the CART model: a comparative study of machine learning techniques. Sci Total Environ 615:272–281. https://doi.org/10.1016/j.scitotenv.2017.09.293

Cigizoglu HK (2004) Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. Adv Water Resour 27:185–195. https://doi.org/10.1016/j.advwatres.2003.10.003

Cobaner M, Unal B, Kisi O (2009) Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. J Hydrol 367:52–61. https://doi.org/10.1016/j.jhydrol.2008.12.024

Gholami V, Booij MJ, Nikzad Tehrani E, Hadian MA (2018) Spatial soil erosion estimation using an artificial neural network (ANN) and field plot data. CATENA 163:210–218. https://doi.org/10.1016/j.catena.2017.12.027

Ghorbani MA, Khatibi R, Hosseini B, Bilgili M (2013) Relative importance of parameters affecting wind speed prediction using artificial neural networks. Theoret Appl Climatol 114:107–114. https://doi.org/10.1007/s00704-012-0821-9

Han D, Chan L, Zhu N (2007) Flood forecasting using support vector machines. J Hydroinf 9:267–276. https://doi.org/10.2166/hydro.2007.027%JJournalofHydroinformatics

Hasanpour Kashani M, Daneshfaraz R, Ghorbani MA, Najafi MR, Kisi O (2015) Comparison of different methods for developing a stage–discharge curve of the Kizilirmak River. J Flood Risk Manag 8:71–86. https://doi.org/10.1111/jfr3.12064

Hassangavyar MB, Damaneh HE, Pham QB, Linh NTT, Tiefenbacher J, Bach Q-V (2020) Evaluation of re-sampling methods on performance of machine learning models to predict landslide susceptibility. Geocarto Int. https://doi.org/10.1080/10106049.2020.1837257

Hauser-Davis RA, Oliveira TF, Silveira AM, Silva TB, Ziolli RL (2010) Case study: comparing the use of nonlinear discriminating analysis and artificial neural networks in the classification of three fish species: acaras (Geophagus brasiliensis), tilapias (Tilapia rendalli) and mullets (Mugil liza). Ecol Inform 5:474–478. https://doi.org/10.1016/j.ecoinf.2010.08.002

Heggen RJ (1999) Hysteresis sensitive neural networks for modeling rating curves. J Comput Civ Eng 13:56–57. https://doi.org/10.1061/(ASCE)0887-3801(1999)13:1(56)

Jain SK (2001) Development of Integrated sediment rating curves using ANNs. J Hydraul Eng 127:30–37. https://doi.org/10.1061/(ASCE)0733-9429(2001)127:1(30)

Jain SK (2012) Modeling river stage–discharge–sediment rating relation using support vector regression Hydrology Research. J Hydrol Res 43:851–861

Jian L, Zhongwu J, Wenjun Y (2014) Numerical modeling of the Xiangxi River algal bloom and sediment-related process in China. Eco Inform 22:23–35. https://doi.org/10.1016/j.ecoinf.2014.03.002

Khan MS, Coulibaly P (2006) Application of support vector machine in lake water level prediction. J Hydrol Eng 11:199–205. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:3(199)

Kim T-W, Valdés JB (2003) Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. J Hydrol Eng 8:319–328. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:6(319)

Kisi O (2005) Suspended sediment estimation using neuro-fuzzy and neural network approaches/estimation des matières en suspension par des approches neurofloues et à base de réseau de neurons. Hydrol Sci J. https://doi.org/10.1623/hysj.2005.50.4.683

Kişi Ö (2008) Stream flow forecasting using neuro-wavelet technique hydrological processes. Int J 22:4142–4152. https://doi.org/10.1002/hyp.7014

Kişi Ö (2010) Daily suspended sediment estimation using neuro-wavelet models. Int J Earth Sci 99:1471–1482. https://doi.org/10.1007/s00531-009-0460-2

Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly streamflow forecasting. J Hydrol 399:132–140. https://doi.org/10.1016/j.jhydrol.2010.12.041

Kisi O, Parmar KS, Soni K, Demir V (2017) Modeling of air pollutants using least square support vector regression, multivariate adaptive regression spline, and M5 model tree models. Air Qual Atmos Health 10:873–883. https://doi.org/10.1007/s11869-017-0477-9

Kumar D, Pandey A, Sharma N, Flügel W-A (2016) Daily suspended sediment simulation using machine learning approach. CATENA 138:77–90. https://doi.org/10.1016/j.catena.2015.11.013

Kumar M, Kumari A, Kushwaha DP, Kumar P, Malik A, Ali R, Kuriqi A (2020) Estimation of daily stage–discharge relationship by using data-driven techniques of a perennial river, India. Sustainability 12(19):7877. https://doi.org/10.3390/su12197877

Kuo C-C, Gan TY, Yu P-S (2010) Wavelet analysis on the variability, teleconnectivity, and predictability of the seasonal rainfall of Taiwan. Mon Weather Rev 138:162–175. https://doi.org/10.1175/2009MWR2718.1

Kuriqi A, Koçileri G, Ardiçlioğlu M (2020) Potential of Meyer-Peter and Müller approach for estimation of bed-load sediment transport under different hydraulic regimes. Model Earth Syst Environ 6:129–137. https://doi.org/10.1007/s40808-019-00665-0

Liu Q-J, Shi Z-H, Fang N-F, Zhu H-D, Ai L (2013) Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet–ANN approach. Geomorphology 186:181–190. https://doi.org/10.1016/j.geomorph.2013.01.012

Malik A, Kumar A (2015) Pan evaporation simulation based on daily meteorological data using soft computing techniques and multiple linear regression. Water Resour Manag 29:1859–1872. https://doi.org/10.1007/s11269-015-0915-0

Malik A, Kumar A, Kisi O, Shiri J (2019) Evaluating the performance of four different heuristic approaches with Gamma test for daily suspended sediment concentration modeling. Environ Sci Pollut Res 26:22670–22687. https://doi.org/10.1007/s11356-019-05553-9

McCuen RH, Knight Z, Cutter AG (2006) Evaluation of the Nash–Sutcliffe efficiency index. J Hydrol Eng 11:597–602

Moeeni H, Bonakdari H (2018) Impact of normalization and input on ARMAX-ANN Model performance in suspended sediment load prediction. Water Resour Manag 32:845–863. https://doi.org/10.1007/s11269-017-1842-z

Nourani V, Alami MT, Aminfar MH (2009) A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. Eng Appl Artif Intell 22:466–472. https://doi.org/10.1016/j.engappai.2008.09.003

Nourani V, Tajbakhsh AD, Molajou A (2018) Data mining based on wavelet and decision tree for rainfall-runoff simulation. Hydrol Res 50:75–84

Partal T, Cigizoglu HK (2008) Estimation and forecasting of daily suspended sediment data using wavelet–neural networks. J Hydrol 358:317–331. https://doi.org/10.1016/j.jhydrol.2008.06.013

Rahgoshay M, Feiznia S, Arian M, Hashemi SAA (2018) Modeling daily suspended sediment load using improved support vector machine model and genetic algorithm. Environ Sci Pollut Res 25:35693–35706. https://doi.org/10.1007/s11356-018-3533-6

Rajaee T, Nourani V, Zounemat-Kermani M, Kisi O (2011) River suspended sediment load prediction: application of ANN and wavelet conjunction model. J Hydrol Eng 16:613–627. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000347

Sharafati A, Haghbin M, Haji Seyed Asadollah SB, Tiwari NK, Al-Ansari N, Yaseen ZM (2020) Scouring Depth assessment downstream of weirs using hybrid intelligence models. Appl Sci 10:3714

Sharghi E, Nourani V, Molajou A, Najafi H (2018) Conjunction of emotional ANN (EANN) and wavelet transform for rainfall-runoff modeling. J Hydroinform 21:136–152. https://doi.org/10.2166/hydro.2018.054

Shiri J, Kisi O (2010) Short-term and long-term streamflow forecasting using a wavelet and neuro-fuzzy conjunction model. J Hydrol 394:486–493. https://doi.org/10.1016/j.jhydrol.2010.10.008

Sivapragasam C, Muttil N (2005) Discharge rating curve extension: a new approach. Water Resour Manag 19:505–520. https://doi.org/10.1007/s11269-005-6811-2

Solomatine DP, Ostfeld A (2008) Data-driven modelling: some past experiences and new approaches. J Hydroinf 10:3–22. https://doi.org/10.2166/hydro.2008.015%JJournalofHydroinformatics

Song K, Park Y-S, Zheng F, Kang H (2013) The application of artificial neural network (ANN) model to the simulation of denitrification rates in mesocosm-scale wetlands. Ecol Inform 16:10–16. https://doi.org/10.1016/j.ecoinf.2013.04.002

Stefánsson A, Končar N, Jones AJ (1997) A note on the Gamma test. Neural Comput Appl 5:131–133. https://doi.org/10.1007/BF01413858

Sudheer KP, Jain SK (2003) Radial basis function neural network for modeling rating curves. J Hydrol Eng 8:161–164. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:3(161)

Vapnik VN (1995) Introduction: four periods in the research of the learning problem. In: Vapnik VN (ed) The nature of statistical learning theory. Springer New York, New York, NY, pp 1–14. https://doi.org/10.1007/978-1-4757-2440-0_1

Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 10:988–999. https://doi.org/10.1109/72.788640

Wang W, Men C, Lu W (2008) Online prediction model based on support vector machine. Neurocomputing 71:550–558. https://doi.org/10.1016/j.neucom.2007.07.020

Willmott CJ (1984) On the evaluation of model performance in physical geography. In: Gaile GL, Willmott CJ (eds) Spatial statistics and models. Springer Netherlands, Dordrecht, pp 443–460. https://doi.org/10.1007/978-94-017-3048-8_23

Wu CL, Chau KW, Li YS (2008) River stage prediction based on a distributed support vector regression. J Hydrol 358:96–111. https://doi.org/10.1016/j.jhydrol.2008.05.028

Youssef OAS (2003) A wavelet-based technique for discrimination between faults and magnetizing inrush currents in transformers. IEEE Trans Power Deliv 18:170–176. https://doi.org/10.1109/TPWRD.2002.803797