

Article

Superiority of Hybrid Soft Computing Models in Daily Suspended Sediment Estimation in Highly Dynamic Rivers

Tarate Suryakant Bajirao ¹, Pravendra Kumar ¹, Manish Kumar ^{1,*}, Ahmed Elbeltagi ^{2,3}  and Alban Kuriqi ^{4,*} 

¹ Department of Soil and Water Conservation Engineering, G. B. Pant University of Agriculture and Technology, Pantnagar 263145, India; ct51181d@gbpuat.ac.in (T.S.B.); tswpk2@gbpuat.ac.in (P.K.)

² Agricultural Engineering Department, Faculty of Agriculture, Mansoura University, Mansoura 35516, Egypt; ahmedelbeltagi81@mans.edu.eg

³ College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China

⁴ CERIS, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal

* Correspondence: ct52623d@gbpuat.ac.in (M.K.); alban.kuriqi@tecnico.ulisboa.pt (A.K.)

Abstract: Estimating sediment flow rate from a drainage area plays an essential role in better watershed planning and management. In this study, the validity of simple and wavelet-coupled Artificial Intelligence (AI) models was analyzed for daily Suspended Sediment (SSC) estimation of highly dynamic Koyna River basin of India. Simple AI models such as the Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) were developed by supplying the original time series data as an input without pre-processing through a Wavelet (W) transform. The hybrid wavelet-coupled W-ANN and W-ANFIS models were developed by supplying the decomposed time series sub-signals using Discrete Wavelet Transform (DWT). In total, three mother wavelets, namely Haar, Daubechies, and Coiflets were employed to decompose original time series data into different multi-frequency sub-signals at an appropriate decomposition level. Quantitative and qualitative performance evaluation criteria were used to select the best model for daily SSC estimation. The reliability of the developed models was also assessed using uncertainty analysis. Finally, it was revealed that the data pre-processing using wavelet transform improves the model's predictive efficiency and reliability significantly. In this study, it was observed that the performance of the Coiflet wavelet-coupled ANFIS model is superior to other models and can be applied for daily SSC estimation of the highly dynamic rivers. As per sensitivity analysis, previous one-day SSC (S_{t-1}) is the most crucial input variable for daily SSC estimation of the Koyna River basin.

Keywords: ANN; ANFIS; sediment deposition; sediment rating curve; wavelet transform



Citation: Bajirao, T.S.; Kumar, P.; Kumar, M.; Elbeltagi, A.; Kuriqi, A. Superiority of Hybrid Soft Computing Models in Daily Suspended Sediment Estimation in Highly Dynamic Rivers. *Sustainability* **2021**, *13*, 542. <https://doi.org/10.3390/su13020542>

Received: 16 December 2020

Accepted: 6 January 2021

Published: 8 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

To prevent soil degradation and improve the water quality, the soil and water management treatments must be carried out at the watershed level. Sediment loss measurement is crucial to assess soil and water conservation treatments on sediment flow through the river [1–3]. Sediment transport information is essential for designing and planning soil and water conservation structures on the river. The sediment movement's magnitude depends on rainfall volume and intensity, Land Use/Land Cover (LULC), topography, and soil physical properties [4]. Sediment concentration flow prediction is important because the accumulation of sediments reduces the reservoir capacity. It also reduces the productivity of such land over which sedimentation takes place. Sedimentation is responsible for increasing the flood hazard due to sediment mass trapping over the river bed [2,5].

Accumulated sediment also blocks the entrance of the intake structure. It also reduces the flow capacity and increases the maintenance and repair cost of irrigation and drainage channels. Increased sediment concentration reduces the available dissolved oxygen in water; hence it is responsible to retard water life [6]. Increased sediment concentration is responsible for increasing the domestic and drinking water purification cost. Higher

sediment flow also carries contaminants developed over upstream areas due to point and non-point pollution sources and increases environmental pollution. Sediment flow forms visual pollution, and hence it reduces the recreational potential of wetlands and lakes. It also reduces fish eggs proliferation by accumulating over the surroundings [6,7]. Sediment load information is essential for the design of reservoirs and dams, stable channels, the design of lakes and estuaries, the reduction of sediment and pollutants flow in rivers, environmental impact assessment, and the wildlife protection of fish habitats [4].

Suspended sediment load measurements are carried out by taking samples along the river section using different samplers. Such samplers truly represent the flowing sediment concentration at a given measuring point of river cross-section. It is still very time-consuming, expensive, and frequent sampling cannot be easily conducted [8,9]. Another method of suspended sediment load measurement is empirical and semi-empirical equations that depend on laboratory experiments conducted under uniform flow conditions with uniform bed material [1]. Still, it is contradictory to real practical conditions [7]. The existing empirical equations and theories available in the literature for sediment load computation are only approximate values because of the complex interaction between sediment transport and water movement. It is challenging to describe sediment flow mathematically because of such a complex relationship between discharge and sediment flow. Different empirical equations provide different results with significant errors [7,10]. The relationship between suspended sediment and discharge is significant in the estimation of sediment load. In the past, many studies have been conducted for the sediment modeling process.

Generally, mathematical models require many data with a long response time, which is rarely available [11]. The detailed information of the watershed's physical properties must be known [12]. The traditional methods are limitedly used and not preferable because of non-linear relationships among hydrological variables and dependence on many hydrological field data [13]. Hence, it is quite a tricky task to forecast a non-stationary time series by using statistical models. However, it is essential to select alternative models to disseminate non-stationary and non-linearity in the time series data. The Artificial Intelligence (AI) technique has an advantage over traditional methods. It can perform well with a large amount of noisy data resulting from the dynamic and non-linear system where the system's fundamental physical relationships are unknown [14–17]. AI techniques can solve the complex problems of different hydrologic processes [18]. A black box model such as an Artificial Neural Network (ANN) does not need watershed physical characteristics to transform inputs into an output and to recognize any hydrologic process [8,19]. AI-based data-driven models such as ANN do not need site-specific information that is rarely available but instead focuses on forming the relationship between input and output variables [20–22]. The modeling of different hydrological processes using ANN is the latest technique [23]. It considers both linearity and non-linearity concepts for model construction. It can perform satisfactorily for the memoryless or dynamic input–output system.

Therefore, ANNs have been used in most of the application areas. Some of the application areas are water resource and hydrology, including modeling of drought, rainfall-runoff, water quality, runoff-sediment, precipitation forecasting, climate change impacts on streamflow, sediment transport process, and groundwater quality and groundwater level forecasting, water level forecasting in aquifers and lakes [7,24–26]. •ANN is a modeling tool capable of identifying high complexity present in input-output data, which operates by considering previously recorded input-output data [12]. •ANN modeling is crucial for predicting suspended sediment concentration in the river flow [27]. •ANN is considered as a lumped parameter model to simulate rainfall, runoff, and sediment •ANN undergoes strong generalization ability; once the ANN architecture is adequately trained, ANN can give accurate results even for those that have never been seen before [28]. •Many researchers have recently applied the ANN model for rainfall-runoff and runoff-sediment modeling [4,19,20,24,28].

The time-series data used for rainfall-runoff-sediment modeling usually contains uncertainties. In such a case, the fuzzy theory can be adapted to model these uncertainties to

solve real-world problems. Adaptive Neuro-Fuzzy Inference System (ANFIS) can be made by combining ANN and fuzzy logic and subsequently receiving benefits from both [22,26]. The ANN model must perform a trial-and-error process to generate the optimal network architecture. Still, the ANFIS model is free from such a tedious procedure. The ANFIS is a multilayer self-organize network structure that adapts the fuzzy system's parameters to predict the system output. However, the fuzzy logic approach's main problem is that no systematic procedure is available to design a fuzzy controller [12]. The ANFIS technique has recently been applied successfully for rainfall–runoff–sediment modeling [7,21,23]. Generally, Fourier transform requires static data for investigation purposes, but signals of hydrological time series are highly non-stationary. The difficulty with Fourier transforms during non-stationary hydrologic modeling is that it only considers the central time-series' frequency feature. It is suitable only for static time series [29].

To overcome this weakness, a wavelet transform can be used. Wavelet transform comes into effect while modeling with non-stationary data. It can produce different series which identify possible trends, seasonal variations, and internal correlation among different components [20]. Wavelet analysis consists of the shifting and scaling of the original (mother) wavelet. Wavelet transform improves the model performance because it simultaneously considers both temporal and spectral information available within the signal; this feature overcomes the Fourier analysis's fundamental limitation. The Fourier spectrum provides only globally averaged information [26]. Wavelet transform decomposes primary time series data into different sub-components without losing any information and extracting required information from historical data using few coefficients. It reveals the hidden information available in the data. It provides a more concise form of original time series data with a timescale representation of processes relationships within them [30,31]. Hence, wavelet data-driven models use these different sub-components' data, which improves the model's performance by capturing the required information available in the central time series hydrological data [22,32]. Different researchers recently used the wavelet transform technique for hydrological time series modeling to improve models' performance accuracy. Grossmann and Morlet [33] introduced wavelet transform, which can reveal the different aspects of time series data such as breakdown points, trends, and discontinuities; this is the superiority of wavelet over other signal analysis techniques [31,32,34].

In a few last years, simple data-driven models, as well as the coupling of data-driven models and wavelet transform, have been applied successfully for hydrological modeling [29,35–37]. Accordingly, this study has been carried out with the following objectives: (a) Select the best input variables for reasonable daily SSC prediction of Koyna River basin using the Gamma test (GT). (b) Develop simple AI and hybrid wavelet-coupled AI models for daily SSC prediction of the study area. (c) Assess the reliability of the best-selected models using uncertainty analysis, and (d) analyze the sensitivity of selected input variables for daily SSC modeling using sensitivity analysis. Different AI techniques behave differently at varying climatic and topographic conditions. Hence, in this study, such a complex river basin having a highly dynamic nature was selected to assess different AI techniques' performance. In this study, different simple AI and hybrid wavelet-coupled AI models were developed to compare their suspended sediment estimation performances. The performance of the developed models was assessed using standard model performance criteria. The novelty of this study includes the finding that the coiflet wavelet-coupled ANFIS model could be used for daily suspended sediment estimation in highly dynamic rivers like the Koyna River basin.

2. Materials and Methods

2.1. Study Area and Data Collection

The Koyna River (a tributary of Krishna River) originates at Mahabaleshwar in the Satara district of Maharashtra, India. The Konya River flows North to South from its origin for about 65 km and then flows for about 56 km eastward to meet the Krishna River. Before taking an eastward turn, it is dammed at Koynanagar, called Koyna dam. The study area

lies between $17^{\circ}7'55''$ N to $17^{\circ}57'50''$ N latitude and $73^{\circ}33'15''$ E to $74^{\circ}11'10''$ E longitude, as shown in Figure 1a.

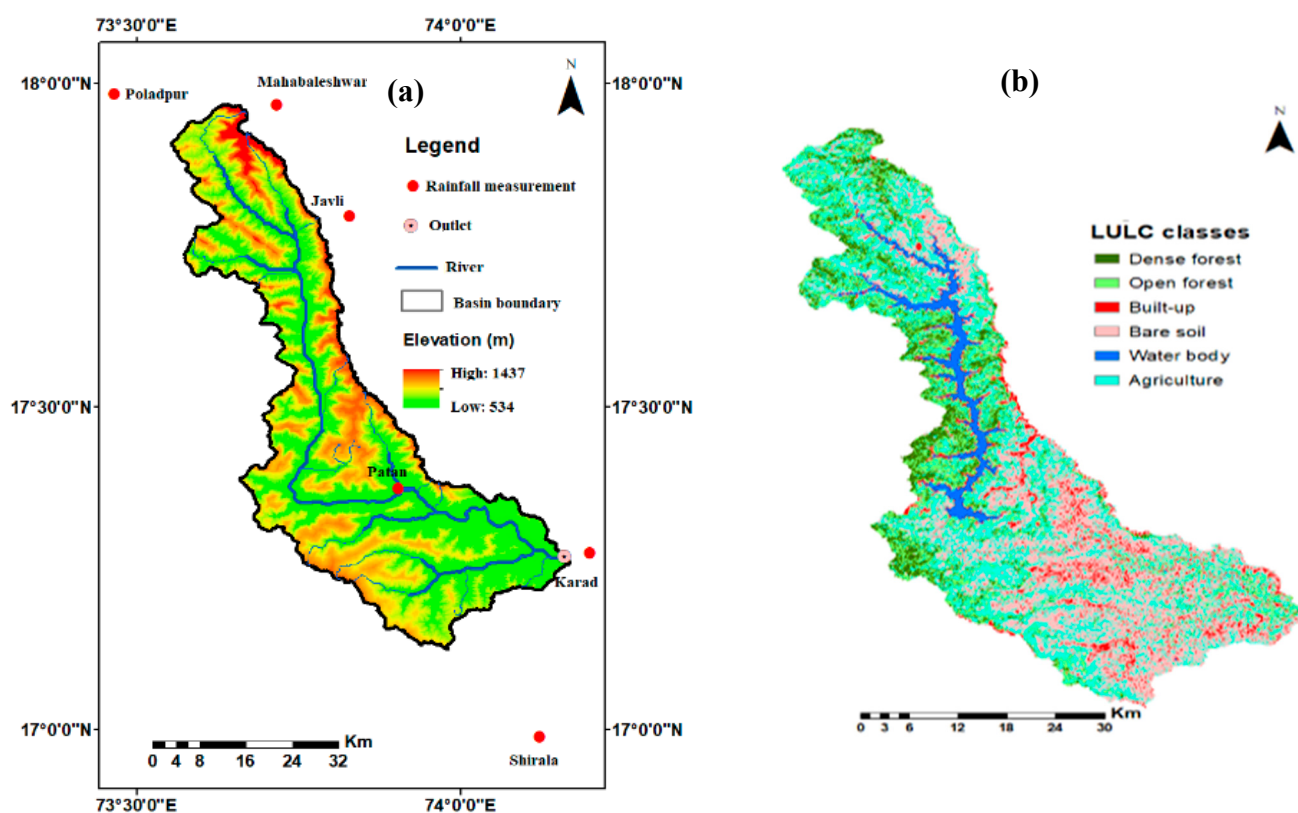


Figure 1. (a) The geographical location and the gauging stations (b) land use/land cover of the Koyna River basin.

The Koyna River basin has a geographical area of 1917 km^2 on the Deccan plateau. The study area comes under the survey of India toposheets 47G/9, 47G/11, 47G/11, 47G/13, 47G/14, 47G/15, 47G/16, 47K/13, and 47K/14. The study area comes under varying climatic and topographic conditions. The annual rainfall at the upstream part of the basin is 5000 mm, reducing to 866 mm downstream. About 88% rainfall occurs in the monsoon (1 June to 30 September) season. In this study, the mean areal rainfall of the study area was determined using the Thiessen polygon method in ArcGIS 10.2 software. The study area comes under a subtropical climate. The winter season starts in October and extends up to January, while the summer season extends from February to May. The daily mean monthly maximum temperature varies between 31°C to 37°C , while the daily mean monthly minimum temperature varies between 10°C to 14°C . Agriculture is the primary source of income for people who live in the River basin. The soil at the upstream part of the basin is light laterite, while the central and downstream area is under black cotton soil. The elevation in the basin varies between 534 m to 1437 m above the mean sea level. The dominant part of the basin is under the steep sloping condition with varying topography and prone to soil loss. The entire basin area is covered by agriculture (717.13 km^2), bare soil (491.69 km^2), open forest (342.60 km^2), dense forest (159.84 km^2), built-up land (120.20 km^2), and water body (85.68 km^2), as shown in Figure 1b which was prepared using ArcGIS 10.2 software in this study runoff and sediment measurements were conducted at the basin's outlet, situated at Warunji village in the Satara district.

The monsoon season's daily rainfall data from 1999 to 2005 (7 years) were collected from the agricultural department of Maharashtra state, India. The daily runoff and suspended sediment concentration (SSC) data of monsoon season for the same duration reported for rainfall were collected from India's central water commission (CWC). Out of

the total data, the first 70% data were used for training. The remaining 30% of the data were used for testing the developed models.

2.2. ANN

ANN is inspired by the biological (brain) neuron system. Each neuron receives processes and sends the signal to make functional relationships between future and past events [38]. The most common structure of ANN is shown in Figure 2.

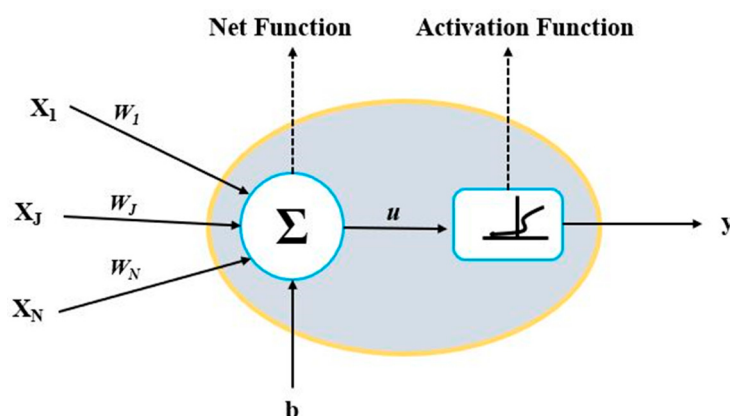


Figure 2. The basic structure of ANN.

ANN consists of the input layer, one or more hidden layers, and the output layer. In the neural network, each neuron (node) has input variables and output variables. A neuron determines the output variable's value after applying the net and activation (transfer) function on input variables. The net function (u) is determined by adding the product of input variables (x_i) with their corresponding connection weights (W_i) plus the bias or threshold value (b) of a neuron. Usually, the net function is in the linear form given as:

$$u = \sum_{i=1}^N x_i W_i + b \quad (1)$$

where x_i is an input variable, W_i is the connection weight from the i^{th} neuron in the input layer, and b is the bias/threshold value of the neuron [39]. The net function (u) at a hidden node is transformed into output (y) using a non-linear activation function. More details of ANN were added in the Appendix A.

2.3. ANFIS

The Fuzzy Inference System (FIS) is a popular computing framework that depends on fuzzy set theory. The building blocks of FIS are the fuzzy operators, fuzzy sets, and the knowledge base. It allows extracting fuzzy rules from expert knowledge or numerical data and constructs a rule base adaptively. Moreover, it can adapt to the complicated conversion of human intelligence to fuzzy systems. Fuzzy theory can be adapted to model the uncertainties to solve real-world problems. For prediction purposes, the ANFIS model is better than the ANN model in peak flow prediction accuracy, prediction error, and computational error [26]. ANFIS is functionally equivalent to FIS; it is a multilayer feed-forward network that uses ANN learning algorithms and fuzzy reasoning to characterize an input space to an output space, powerful in modeling numerous time series processes.

ANFIS is a universal approximator because it has set-off If-Then rules that have the learning potential to approximate a non-linear system. It can approximate any real continuous function. ANFIS structure consists of many nodes connected through directional links with each other. Each node has to perform a specific function with adjustable or fixed

parameters [40]. For first-order Takagi-Sugeno-Kang (TSK) FIS with two inputs (x and y) and one output (f), two typical rules can be expressed as:

$$\text{Rule 1: If } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ then } f_1 = p_1x + q_1y + r_1 \quad (2)$$

$$\text{Rule 2: If } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ then } f_2 = p_2x + q_2y + r_2 \quad (3)$$

where A_i and B_i are Membership Functions (MFs) for input x and y , respectively. Simultaneously, p_i , q_i , and r_i are the design (consequent) parameters estimated during the training process. The TSK model's fuzzy reasoning mechanism derives an output (f) with inputs (x and y). General ANFIS structure with two inputs, two rules, and a single output. The functioning of the ANFIS structure (five layers) is described below. The description of the ANFIS layers functions was presented in the Appendix B.

2.4. Subtractive Clustering

Data clustering is a process that puts similar data into groups. A clustering algorithm divides the whole data into various groups. The existence of similarity within a group is more than in other groups. Clustering algorithms are mostly used for data categorization as well as data compression and model construction. When there is no clear idea to select how many clusters in a given data set, then the subtractive clustering method can be used. It is a fast, one-pass algorithm to estimate the number of clusters and cluster centers in a given data set [37]. In subtractive clustering, the number of the fuzzy rule set is equal to the number of cluster centers.

Considering n data points ($x_1, \dots, x_n$), subtractive clustering assumes each data point act as a candidate for representing the cluster center. The subtractive clustering depends on data density. The density index at any point x_i is expressed as:

$$D_i = \sum_{j=1}^n e^{\left(-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right)} \quad (4)$$

D_i represents density index, and r_a is positive ($r_a > 0$), indicating the neighborhood radius in each cluster center. So, the data point which has more neighborhood points indicates more potential to represent as cluster center. Those data points which are located outside the radius create less impact on the density index. Selection of clustering radius receives greater importance while determining the number of clusters. The high value of r_a causes the minimum number of clusters and vice versa. After determining a data point with a high potential to act as a cluster center, say X_{c1} is a point act as a first cluster center determined by the D_{ci} density index. The expression recalculates the next density index for each data point X_i :

$$D_{i+1} = D_i - D_{ci} e^{\left(-\frac{\|x_i - x_j\|^2}{(r_b/2)^2}\right)} \quad (5)$$

where r_b is a positive constant ($r_b > 0$), which indicates the neighborhood radius for which the most significant reduction in the density will be achieved. To prevent closely spaced cluster centers, r_b is usually equal to 1.5 times of r_a . After this density measurement, the next cluster center X_{c2} is selected. The same process is repeated until sufficient numbers of cluster centers are achieved.

2.5. Wavelet Transform

In total, two types of wavelet transform, namely, Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT), are used. The DWT can capture the non-linear and non-stationary time series' dynamic properties using different wavelet coefficients [11]. DWT analysis is mostly preferred for forecasting water resource system

problems because it needs a short computational time [30]. The discrete mother wavelet can be expressed as [26]:

$$M_{a,b}(t) = \frac{1}{\sqrt{m_0^a}} M\left(\frac{t - bn_0 m_0^a}{m_0^a}\right) \quad (6)$$

where a and b are the integers that control the wavelet dilation and translation, respectively. The most commonly used value for these parameters is $m_0 = 2$ and $n_0 = 1$, where n_0 is the location parameter. It must be greater than zero and m_0 displays fined dilation step that is greater than one. The DWT scale and position are based on the power of two (dyadic scales and positions); this power of two logarithmic scalings of the translation and dilation is known as the dyadic grid arrangement. The dyadic wavelet function is defined as [41]:

$$M_{a,b}(t) = 2^{-\left(\frac{a}{2}\right)} M(2^{-a}t - b) \quad (7)$$

For a discrete-time, series, x_i , the dyadic wavelet transform becomes [26]:

$$T_{a,b} = 2^{-\left(\frac{a}{2}\right)} \sum_{i=0}^{L-1} M(2^{-a}i - b) x_i \quad (8)$$

$T_{a,b}$ is the wavelet coefficient for the discrete wavelet with scale $m = 2^a$ and location $n = 2^a b$. ($i = 0, 1, 2, \dots, L-1$; and L is an integer power of 2: $L = 2^A$). Also, the signal's smoothed component, which represents the time series' overall trend, is considered T . The discrete inverse transform can reconstruct the signal x_i as [42]:

$$x_i = T(t) + \sum_{a=1}^A T_{a,b}(t) \quad (9)$$

where $T(t)$ is the approximate sub-signal at level A and $T_{a,b}(t)$ are details sub-signals at levels $a = 1, 2, \dots, A$ and time dimension of t ($t = 1, 2, \dots, b$). The wavelet coefficients, $T_{a,b}(t)$ with ($a = 1, 2, \dots, A$), give the detailed sub-signals which can capture small features of interpretational value available in the time series data. The residual term $T(t)$ indicates an approximate sub-signal, representing background information available in the time series data. An approximate sub-signal represents the general trend of the original time series signal. In contrast, a detailed sub-signal represents high-frequency components of the original time series signal [42]. Using these approximate and detail sub-signals, the characteristics of time series like jump, period, hidden period, and dependence can be identified easily [22]. When the original signal passes through low and high pass filter at each decomposition level, it gets resolved into approximate and detail sub-signals; the decomposition at each level is satisfied by the condition given as:

$$X(n) = cA1 + cD1 = cA2 + cD1 + cD2 \quad (10)$$

In this study, original rainfall, runoff, and SSC time series data were decomposed through DWT into approximate and detail sub-signals using MATLAB (R2015a) wavelet toolbox.

2.6. Mother Wavelets

Different mother wavelets are characterized by their features, such as their support region and the corresponding number of vanishing moments. In this study, for investigation purposes, Haar, Daubechies, and Coiflets mother wavelets were selected for assessing their comparative prediction performance for daily SSC prediction. Some of the most-used mother wavelets are described below.

Haar wavelet: It is the first and the most straightforward form of all other wavelets. It is discontinuous and resembles a step function. It is suitable for such time series which

has sudden transitions. Nevertheless, this property may be a disadvantage of the Haar wavelet as it is not differentiable [31]. Daubechies wavelet (db): It is compactly supported (finite length) orthonormal wavelets and makes discrete wavelet transform possible. It is represented by dbN where N is the number of vanishing or zero moments. The ordinary members of this family are db2, db3, db4, db5, db6, db7, db8, db9, and db10.

Coiflets wavelet (coif): In the Coiflet wavelet, both the scaling function and wavelet function have vanishing moments. It facilitates wavelet transformation and provides an excellent approximation of polynomial function at different resolutions, increasing computational efficiency. This wavelet family has four members, namely coif1, coif2, coif3, and coif4. In this study, for investigation purposes, Haar, Daubechies (db2), and Coiflets (coif2) mother wavelets were selected to assess their comparative prediction performance for daily SSC prediction. Schematic representation of the selected mother wavelets.

2.7. Gamma Test (GT)

Hydrological processes are highly complex, dynamic, and non-linear. To select the best input combination, it needs to carry out a trial-and-error procedure. GT provides useful information to select the best input variables for constructing a reliable and smooth model. It also reduces the workload required to develop models by considering all input combinations [12]. GT is a non-parametric test. GT determines the variance of noise or the Mean Square Error (MSE) without any over-fitting. GT evaluates the non-linear correlation between two random variables like input and output pairs.

GT assumes that if two variables b and b' are close together in the input space, their corresponding output variables c and c' should also be close in the given output space. If the outputs are not close together, then it will indicate noise. Suppose the data set is given in the form as:

$$\{(b_i, c_i), 1 \leq i \leq N\} \quad (11)$$

where vector $b \in R^N$ denotes input, corresponding scalar $c \in R$ denotes output. Here, the assumption is that the input vector consists of useful information that can affect output c . The relationship among variables in the system is assumed in the form as:

$$c = f(b_1 \dots b_N) + s \quad (12)$$

where s denotes the noise. Here, it is also assumed that variance of the noise is bounded. The additional Gamma test forms were illustrated in the Appendix C.

In this study, the best input combination for daily SSC prediction was selected based on a minimum value of gamma (Γ) and V_{ratio} . To apply the GT, win GammaTM software was used.

2.8. Data Normalization

In this study, the original time series data were normalized between 0 and 1 for practical training of the developed models. It will give equal attention to all input parameters and helps in fast convergence during training. Interpretability of models improves due to normalization [30]. Here, the normalization of supplied input data was done for eliminating their dimensions using the expression as:

$$x_i = \frac{S_i - S_{min}}{S_{max} - S_{min}}, (i = 1, 2, \dots, n) \quad (13)$$

where S_{max} and S_{min} are the maximum and minimum values of original time series data, n is the number of data points, and x is the original input variable's normalized value.

2.9. Training and Testing of Developed Models

MLP based feed-forward ANN is the most commonly used in hydrology while coupling with wavelet transform [32]. In this study, feed-forward MLP based ANN models were trained and tested in MATLAB (R2015a) software. The single hidden layer with (BP)

algorithm can model the input-output system's complex and non-linear behavior [30]. Hence, in this study, three layered (one input, one hidden, and one output layer) neural network models were developed for daily SSC prediction. The most accurate, reliable, and fast BP based Levenberg-Marquardt (LM) learning rule was used to train neural network models [30,39]. The selection of the number of neurons in the hidden layer is also a difficult task. The number of neurons in the input layer is equal to the number of input variables. In contrast, the number of neurons in the output layer is the same as that of the number of output variables (only one in this study). To date, no clear guideline is available to select the optimum number of neurons in the hidden layer. However, Olyaie, et al. [11] suggested that the hidden layer neurons should be increased from $(2(\sqrt{n}) + m)$ to $(2n + 1)$ for selecting the optimal number, where n is the number of input variables and m is the number of output variables. It can solve the problem of under- and over-fitting of models. Due to under- and over-fitting, the model cannot detect signals [30]. In this study, the number of neurons was increased from 1 to $2n + 1$ in the single hidden layer of both ANN and WANN models to avoid under- and over-fitting. The non-linear sigmoid transfer function is capable of solving a complex problem. Hence, the hyperbolic tangent sigmoid (tan sig) transfer function was used in this study, whose mathematical representation is given in Equation (21). Developed ANN and WANN models were trained with maximum iterations of 1000. In this study, feed-forward MLP based ANN/WANN models were developed using "nntool" in MATLAB (R2015a) software.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (14)$$

Generally, grid partitioning and subtractive clustering are used for FIS generation. Nevertheless, the problem with grid partitioning is that it creates m^n number of fuzzy rules (here, n is the number of input variables, and m is the number of MFs per input). Hence, when the inputs increase slightly, fuzzy rules increase rapidly [43]. Grid partitioning can be easily used for solving problems with less than 6 input variables [44]. Hence, in this study, subtractive clustering was employed to develop ANFIS/WANFIS models. In total, eight input MFs were used by changing the number of MFs per input from 2 to 4. The rule base was constructed with OR logical operation by changing the number of rules from 2 to 4. A total of two output MFs, such as constant and linear, were used. Hence, individually 48 ($8 \text{ input MFs} \times 2 \text{ output MFs} \times 3 \text{ rules}$) ANFIS models were developed by applying simple ANFIS or any WANFIS models technique by keeping the error tolerance of 0.001 and maximum iterations of 1000 in MATLAB (R2015a) software with a hybrid learning algorithm. In this study, different ANFIS/WANFIS models were developed using "anfisedit" tool in MATLAB (R2015a) software with hybrid learning algorithm and Takagi–Sugeno–Kang (TSK) FIS.

2.10. Performance Evaluation of Developed Models

2.10.1. Quantitative Evaluation

To avoid personal bias for selecting the best performing model with qualitative assessment method, different statistical and hydrological parameters were used quantitatively to evaluate the developed models' effectiveness. Statistical indices applied such as root mean squared error (RMSE), the correlation coefficient (r), and Willmott Index (WI) were included in Appendix D.

2.10.2. Hydrological Indices

Coefficient of Efficiency (CE)

Reference [45] introduced the coefficient of efficiency. It is used to evaluate the predictive power of developed models in the field of hydrology. Its value ranges between $-\infty$ to 1. The CE value of 1 indicates perfect matching of predicted and observed values. The CE value of 0 indicates that the model's prediction is as accurate as the mean value of the observed data values. When the CE value is less than 0, the observed data series is a

better predictor than the developed model. As the CE value is closer to 1, the model will be more accurate. It is expressed as:

$$CE = 1 - \frac{\sum_{i=1}^n (Q_o - Q_p)^2}{\sum_{i=1}^n (Q_o - \bar{Q}_o)^2} \quad (15)$$

Pooled Average Relative Error (PARE)

PARE can be used to judge the model's under-prediction and over-prediction performance [23]. PARE's positive value indicates over-prediction, and PARE's negative values indicate the under-prediction performance of the developed model. It can be expressed as:

$$PARE (\%) = \frac{1}{N} \left\{ \frac{\sum_{i=1}^n (Q_p - Q_o)}{\sum_{i=1}^n Q_o} \times 100 \right\} \quad (16)$$

2.11. Qualitative Evaluation

The qualitative and quantitative evaluation was carried out to select the best model for prediction purposes to assess developed models' performance. Hence, both qualitative and quantitative evaluation was performed to judge the goodness of fit between predicted and observed values. Here, the qualitative evaluation was completed from the graphical comparison of time series (sedimentographs) and scattered plots of observed and predicted SSC values. The best model was selected based on quantitative and qualitative evaluation criteria like maximum CE , r , WI , R^2 and minimum $RMSE$, and $PARE$.

2.12. Uncertainty Analysis

Due to the highly stochastic nature of hydrologic processes, the developed models may have too much uncertainty. The predicted values are not the same as observed values, while predicted values always have some uncertainties. Hence, in this study, to test the reliability of developed models, an uncertainty analysis was also carried out using the following indices.

2.12.1. Width of Uncertainty Band of Error Prediction (W_e)

The uncertainty band's width at a 95% confidence level is $\pm (1.96 \sigma / \sqrt{n})$. It indicates the margin of the prediction error, where σ is the standard deviation of prediction error, and it is expressed as:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (e - \bar{e})^2}{n}} \quad (17)$$

$$\bar{e} = \frac{1}{n} \sum_{i=1}^n e \quad (18)$$

where $\bar{e}$ is the mean error of prediction, and e is the individual prediction error, $e = Q_p - Q_o$.

2.12.2. 95% Confidence Interval of Error Prediction (CI_e)

It represents the 95% probability that the error incorporated in the model's prediction can lie within the specified interval. A wider interval indicates the presence of more uncertainty or vice versa. Its expression is given as:

$$CI_e = (\bar{e} \pm 1.96 \frac{\sigma}{\sqrt{n}}) \quad (19)$$

The flowchart of the methodology adopted for daily suspended sediment estimation is shown in Figure 3.

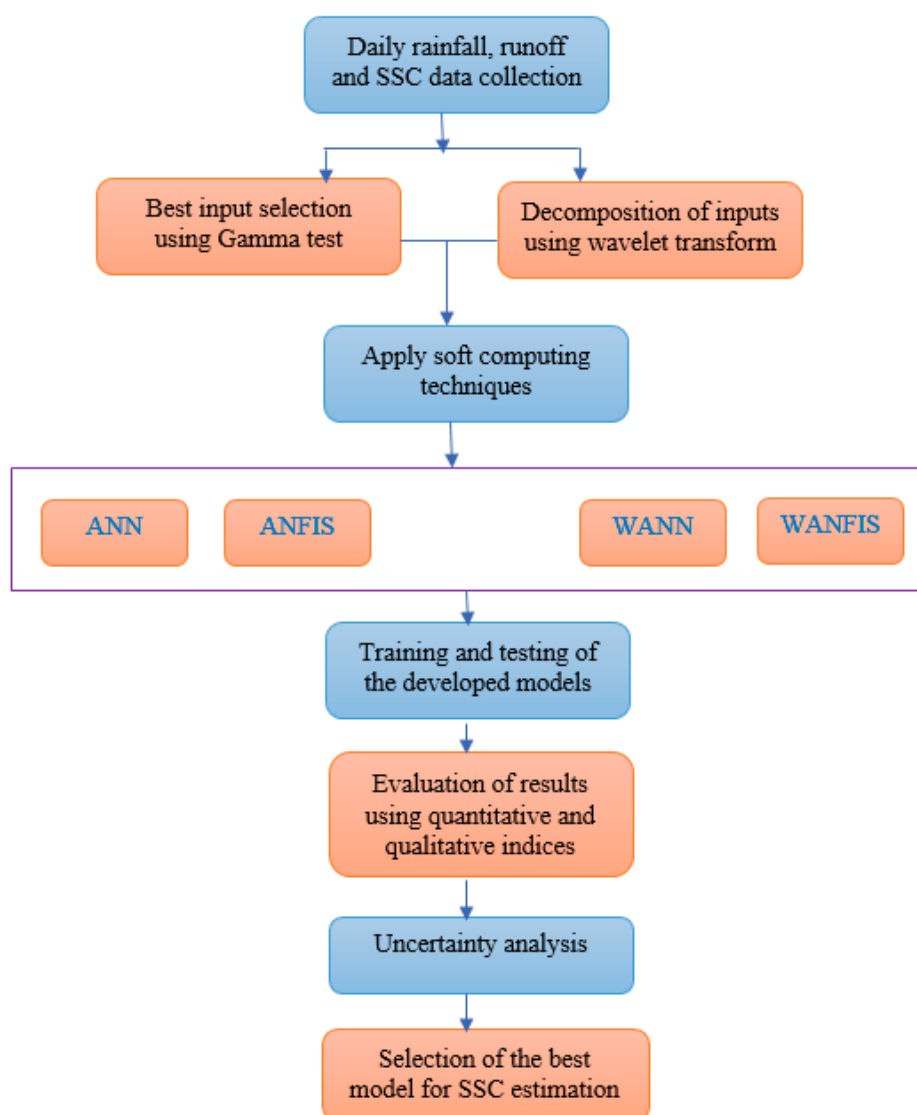


Figure 3. Flowchart of the methodology adopted for daily suspended sediment concentration (SSC) estimation.

3. Results

3.1. Data Analysis

Different statistical parameters of original rainfall, runoff, and SSC time series used in this analysis are presented in Table 1.

Table 1. Statistical parameters of the dataset used for daily SSC prediction.

Statistical Parameters	Whole Data			Training Data Set			Testing Data Set		
	R_t (mm)	Q_t (m ³ /s)	SSC_t (g/L)	R_t (mm)	Q_t (m ³ /s)	SSC_t (g/L)	R_t (mm)	Q_t (m ³ /s)	SSC_t (g/L)
Mean	14.60	215.65	0.059	12.00	141.03	0.048	20.61	388.02	0.084
Standard Deviation	22.91	405.67	0.088	17.33	191.73	0.069	31.53	646.81	0.118
Kurtosis	11.85	41.33	20.03	7.62	14.54	17.105	6.81	15.67	13.261
Skewness	3.01	5.38	3.643	2.51	3.41	3.175	2.47	3.48	3.180
Range	193.32	4641	0.841	119.51	1397.00	0.717	193.32	4640.50	0.839
Minimum	0.00	0.00	0.000	0.00	0.00	0.000	0.00	0.50	0.002
Maximum	193.32	4641	0.841	119.51	1397.00	0.717	193.32	4641.00	0.841
Count	854	854	854	596	596	596	258	258	258

The lower values of skewness of rainfall, runoff, and SSC were observed in both the training and testing data set. The low value of skewness is necessary for appropriate modeling. Table 2 represents the existence of a degree of correlation among rainfall (R), runoff (Q), and SSC (S) with three days lagged duration. The previous 3 days' rainfall, runoff, and SSC were considered for applying GT based on the degree of correlation.

Table 2. Cross-correlation matrix between different variables for daily SSC prediction.

Variable	S_t	R_t	R_{t-1}	R_{t-2}	R_{t-3}	Q_t	Q_{t-1}	Q_{t-2}	Q_{t-3}	S_{t-1}	S_{t-2}	S_{t-3}	p -Values
S_t	1.00												-
R_t	0.70	1.00											5.65×10^{-26}
R_{t-1}	0.68	0.74	1.00										0.73
R_{t-2}	0.61	0.58	0.74	1.00									0.97
R_{t-3}	0.51	0.45	0.58	0.74	1.00								0.27
Q_t	0.62	0.66	0.68	0.63	0.56	1.00							0.04
Q_{t-1}	0.54	0.48	0.66	0.68	0.63	0.89	1.00						0.78
Q_{t-2}	0.44	0.37	0.48	0.66	0.68	0.75	0.89	1.00					0.93
Q_{t-3}	0.35	0.29	0.37	0.48	0.66	0.65	0.75	0.89	1.00				0.10
S_{t-1}	0.72	0.51	0.70	0.68	0.61	0.59	0.62	0.54	0.44	1.00			1.82×10^{-27}
S_{t-2}	0.60	0.41	0.51	0.70	0.68	0.56	0.59	0.62	0.54	0.72	1.00		0.00016
S_{t-3}	0.49	0.33	0.41	0.51	0.70	0.52	0.56	0.59	0.62	0.60	0.72	1.00	0.18

3.2. GT for Input Selection

As the hydrologic processes are highly dynamic, the present response depends on the current response system and the past response in the hydrologic system's memory. Hence, it is considered that present-day SSC response depends on the present day's response of rainfall and runoff and the past three days' response of rainfall, runoff, and SSC. The relationship between input and output variables can be expressed mathematically as:

$$S_t = f(R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}, S_{t-3}) \quad (20)$$

where R_t and Q_t are the present day's rainfall and runoff, respectively. R_{t-1} , R_{t-2} , and R_{t-3} are the previous one, two- and three-days rainfall, respectively, Q_{t-1} , Q_{t-2} , and Q_{t-3} are the previous one, two- and three-days runoff, respectively, S_{t-1} , S_{t-2} , and S_{t-3} is the previous one, two- and three-days SSC, respectively. Here, initially, a total of 11 input variables (j) were selected for GT. Based on these 11 input variables, a total of $2^j - 1$ (i.e., 2047) input combinations can be possible, but, in this study, reliable 58 input combinations were made and analyzed as presented in Table 3.

Table 3. Selection of the best input combination using a Gamma test (GT).

Model	Model Input Combination	Mask	Gamma	V – Ratio
M1	R_t	1000000000	0.0047802000	0.5471600
M2	R_t, R_{t-1}	1100000000	0.0042625000	0.4879000
M3	R_t, R_{t-1}, R_{t-2}	1110000000	0.0041758000	0.4779800
M4	$R_t, R_{t-1}, R_{t-2}, R_{t-3}$	1111000000	0.0032813667	0.4219459
M5	$R_t, R_{t-1}, R_{t-2}, Q_t$	1111000000	0.0038298000	0.4383700
M6	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t$	1111100000	0.0030169776	0.3879486
M7	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}$	1111110000	0.0034506204	0.4437100
M8	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}$	1111111000	0.0034489278	0.4434924
M9	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}$	1111111100	0.0032920565	0.4233205
M10	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}$	1111111110	0.0032920565	0.4233205
M11	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}$	1111111111	0.0032920306	0.4233172
M12	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}, S_{t-3}$	1111111111	0.0032920306	0.4233172
M13	$R_t, R_{t-1}, R_{t-2}, Q_t, S_{t-1}, S_{t-2}$	11101000110	0.0034097067	0.4384490
M14	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, S_{t-1}$	11111000100	0.0030159990	0.3878227
M15	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, S_{t-1}, S_{t-2}$	11111000110	0.0030159998	0.3878228
M16	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, S_{t-1}, S_{t-2}, S_{t-3}$	11111000111	0.0030165903	0.3878988

Table 3. Cont.

Model	Model Input Combination	Mask	Gamma	V – Ratio
M17	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, S_{t-1}, S_{t-2}, S_{t-3}$	11111100111	0.0034493755	0.4435500
M18	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}, S_{t-3}$	11111110111	0.0034489289	0.4434925
M19	$R_t, R_{t-1}, R_{t-2}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}, S_{t-3}$	11101111111	0.0033026471	0.4246824
M20	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}$	11111111110	0.0032920306	0.4233172
M21	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}, S_{t-3}$	11111110111	0.0034489289	0.4434925
M22	$R_t, R_{t-1}, R_{t-2}, Q_t, Q_{t-1}$	11101100000	0.0042706000	0.4888370
M23	$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, Q_{t-1}$	11111100000	0.0034506204	0.4437100
M24	$R_t, R_{t-1}, R_{t-2}, Q_t, Q_{t-1}, Q_{t-2}$	11111100000	0.0040742000	0.4663550
M25	$R_t, R_{t-1}, R_{t-2}, Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}$	11101110100	0.0040742000	0.4663550
M26	$R_t, R_{t-1}, R_{t-2}, Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}$	11101110110	0.0040742000	0.4663550
M27	$R_t, R_{t-1}, R_{t-2}, S_{t-2}$	11100000010	0.0041747000	0.4778580
M28	$R_t, R_{t-1}, R_{t-2}, S_{t-1}, S_{t-2}$	11100000110	0.0041719000	0.4775350
M29	$R_t, R_{t-1}, R_{t-2}, Q_t, S_{t-1}$	11101000100	0.0038292000	0.4383100
M30	$R_t, R_{t-1}, R_{t-2}, Q_t, S_{t-1}, S_{t-2}$	11101000110	0.0038289000	0.4382700
M31	$R_t, R_{t-1}, R_{t-2}, Q_t, Q_{t-1}, S_{t-1}, S_{t-2}$	11111100110	0.0042695000	0.4887090
M32	R_{t-1}	01000000000	0.0045396000	0.5196200
M33	R_{t-1}, R_{t-2}	01100000000	0.0046157000	0.5283400
M34	R_{t-1}, R_{t-2}, Q_t	01101000000	0.0048902000	0.5597500
M35	$R_{t-1}, R_{t-2}, Q_t, S_{t-1}$	01101000100	0.0048889000	0.5596100
M36	$R_{t-1}, R_{t-2}, Q_t, S_{t-1}, S_{t-2}$	01101000110	0.0048885000	0.5595600
M37	$R_{t-1}, R_{t-2}, Q_t, Q_{t-1}, S_{t-1}, S_{t-2}$	01101100110	0.0045635000	0.5223630
M38	$R_{t-1}, R_{t-2}, Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}$	01101110110	0.0044611000	0.5106410
M39	R_{t-2}	00100000000	0.0059649000	0.6827700
M40	R_{t-2}, Q_t	00101000000	0.0049465000	0.5662000
M41	R_{t-2}, Q_t, S_{t-1}	00101000100	0.0049457000	0.5661100
M42	$R_{t-2}, Q_t, S_{t-1}, S_{t-2}$	00101000110	0.0049452000	0.5660450
M43	Q_t	00001000000	0.0053284000	0.6099170
M44	Q_t, S_{t-1}	00001000100	0.0053036000	0.6070760
M45	Q_t, S_{t-1}, S_{t-2}	00001000110	0.0052017000	0.5954070
M46	S_{t-1}	00000000100	0.0045528000	0.5211380
M47	S_{t-1}, S_{t-2}	00000000110	0.0046496000	0.5322160
M48	S_{t-2}	00000000010	0.0060325000	0.6905120
M49	$R_t, R_{t-2}, Q_t, S_{t-1}, S_{t-2}$	10101000110	0.0041118000	0.4706530
M50	$R_t, Q_t, S_{t-1}, S_{t-2}$	10001000110	0.0042217000	0.4832370
M51	R_t, S_{t-1}, S_{t-2}	10000000110	0.0039567000	0.4529040
M52	R_t, S_{t-2}	10000000010	0.0042196000	0.4830030
M53	$R_t, R_{t-1}, Q_t, S_{t-1}, S_{t-2}$	11001000110	0.0038594000	0.4417640
M54	$R_t, R_{t-1}, S_{t-1}, S_{t-2}$	11000000110	0.0042129000	0.4822330
M55	R_t, R_{t-1}, S_{t-2}	11000000010	0.0042181000	0.4828270
M56	$R_t, R_{t-1}, R_{t-2}, S_{t-1}, S_{t-2}$	11100000110	0.0041719000	0.4775350
M57	$R_t, R_{t-1}, R_{t-2}, S_{t-2}$	11100000010	0.0041747000	0.4778580
M58	$R_t, R_{t-1}, R_{t-2}, Q_t, S_{t-2}$	11101000010	0.0038293000	0.4383140

The best input combination was selected based on minimum Gamma (Γ) and V_{ratio} value. As per GT, the model (M14) with six input variables ($R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, S_{t-1}$) were selected and used to develop different models.

3.3. Hydrological Model Development

3.3.1. ANN/ANFIS Models for Daily SSC Prediction

The original input variables selected by GT were directly used to construct simple ANN and ANFIS models. The schematic representation of different ANN/ANFIS models for daily SSC prediction is shown in Figure 4.

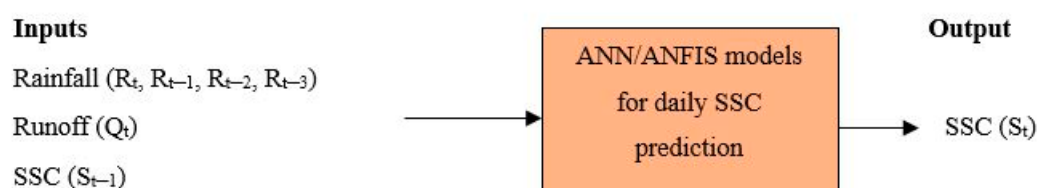


Figure 4. Schematic representation of ANN/ANFIS models for daily SSC prediction.

In simple ANN models, the number of neurons in the single hidden layer varied from 1 to $2n + 1$ ($2 \times 6 + 1$). Hence, 13 simple ANN models were developed to assess their performance for daily SSC prediction.

3.3.2. WANN/WANFIS Models for Daily SSC Prediction

For the prediction of daily SSC using WANN/WANFIS models, the original time series data selected by GT were decomposed by applying DWT and fed as input to ANN/ANFIS models. Here, wavelet transformed data were linked to the ANN/ANFIS network to develop WANN/WANFIS models, respectively. Haar, db2, and coif2 mother wavelets were used to decompose selected six input variables into different multi-frequency sub-signals at the appropriate decomposition level. The appropriate decomposition level was selected using the suggested empirical relation as [29,42]:

$$i = \text{int}[\log(N)] \quad (21)$$

where i is the appropriate decomposition level, N is the number of data points (Here, N was taken as 854 data points, and $\text{int}[\cdot]$ is the integer part function. Therefore, selected six input variables were decomposed at level (i) 2 using selected mother wavelets. Here, each of the selected input variables was decomposed at level 2, which produced 1 approximate (A2) and 2 detail (D1, D2) (total 3) sub-signals. Now, the best selected six input variables (i.e., R_t , R_{t-1} , R_{t-2} , R_{t-3} , Q_t , S_{t-1}) produced 18 (6×3) sub-signals. Hence, these total 18 sub-signals (input variables) were fed to ANN/ANFIS to develop WANN/WANFIS models, respectively, for daily SSC prediction, as shown in Figure 5 schematically.

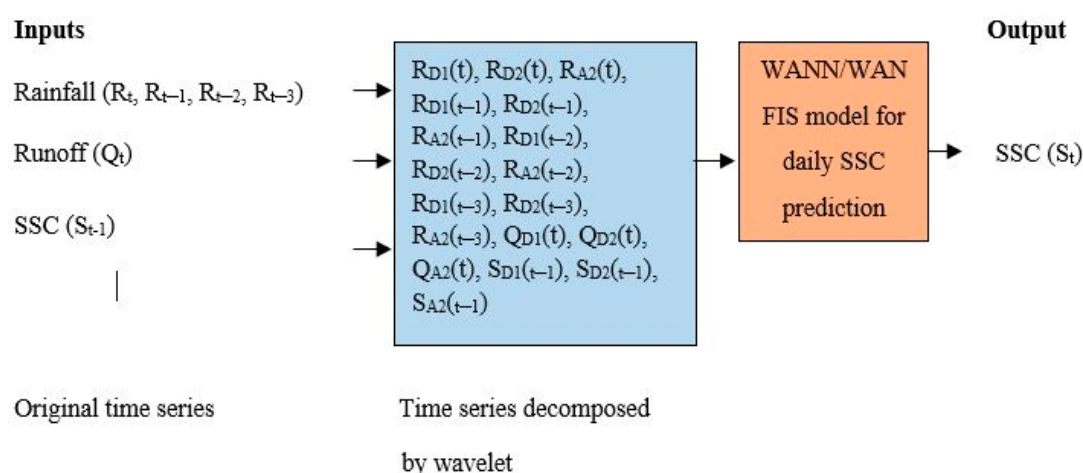


Figure 5. Schematic representation of wavelet coupled ANN/wavelet coupled ANFIS models for daily SSC prediction.

The decomposed sub-signals of normalized rainfall, runoff, and SSC time series with selected mother wavelets at level (i) 2 are shown in Figures 6–8, respectively.

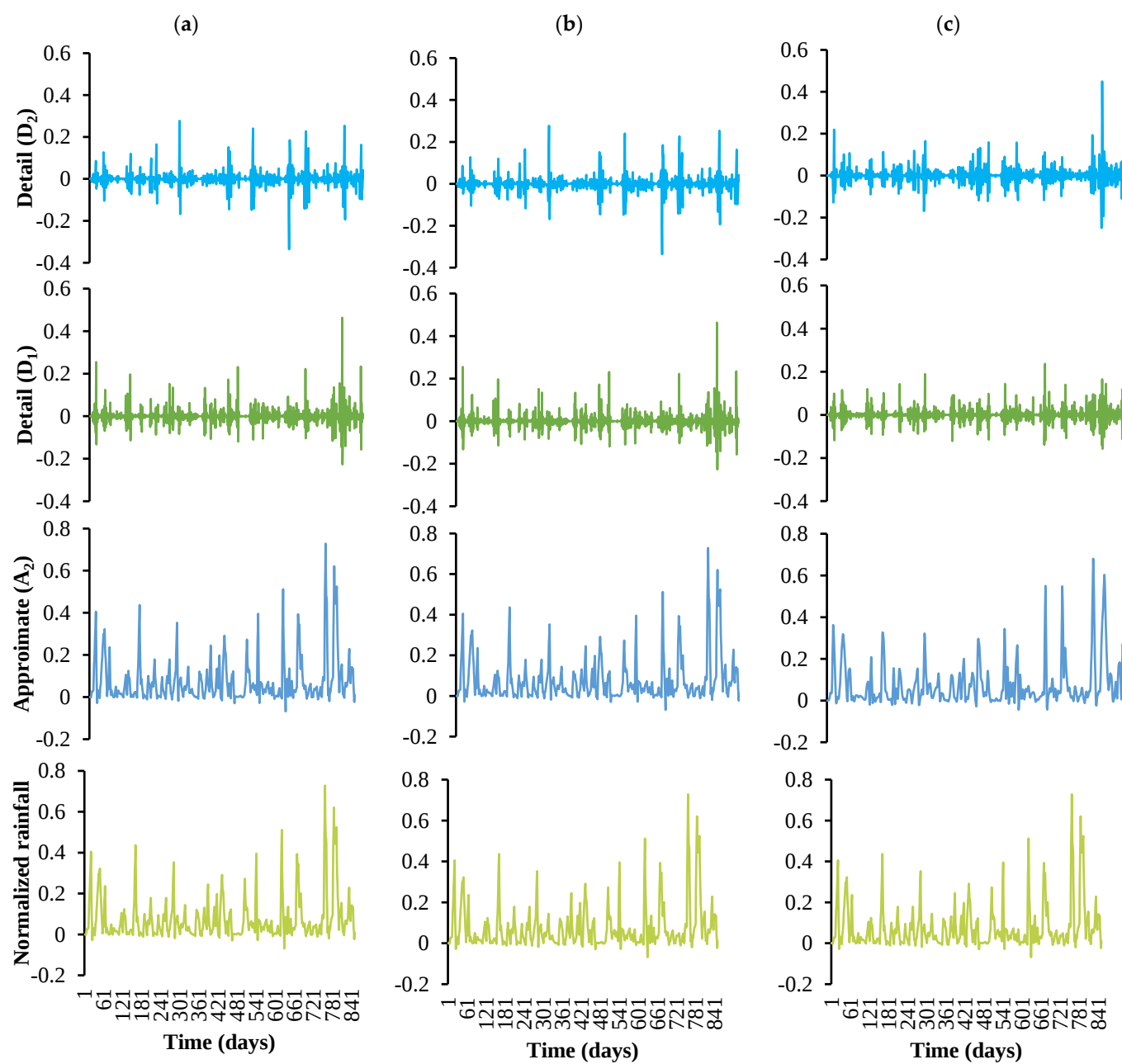


Figure 6. Decomposition of original rainfall time series using different mother wavelets. (a) Haar, (b) db2, (c) coif2.

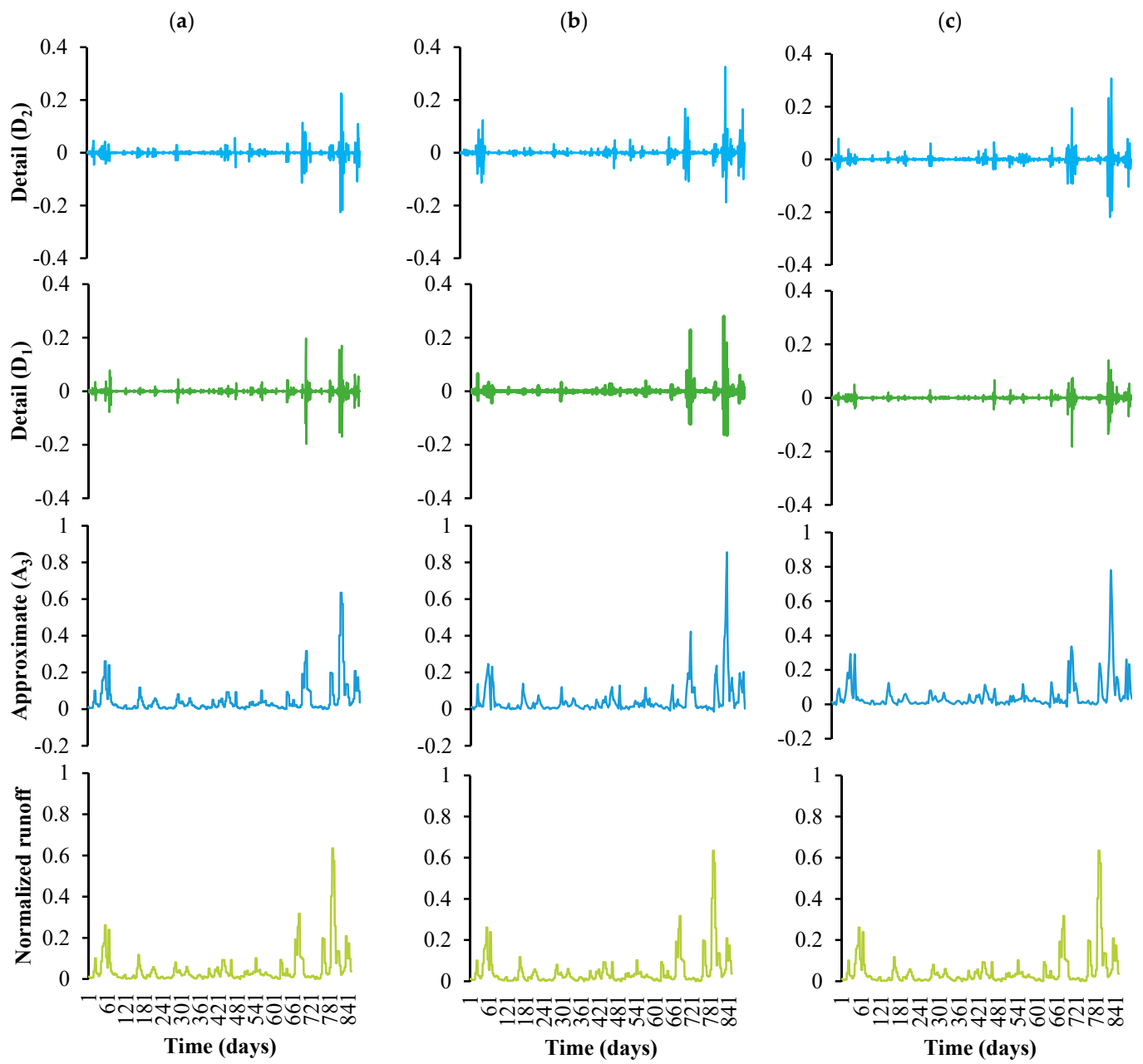


Figure 7. Decomposition of original runoff time series using different mother wavelets. (a) Haar, (b) db2, (c) coif2.

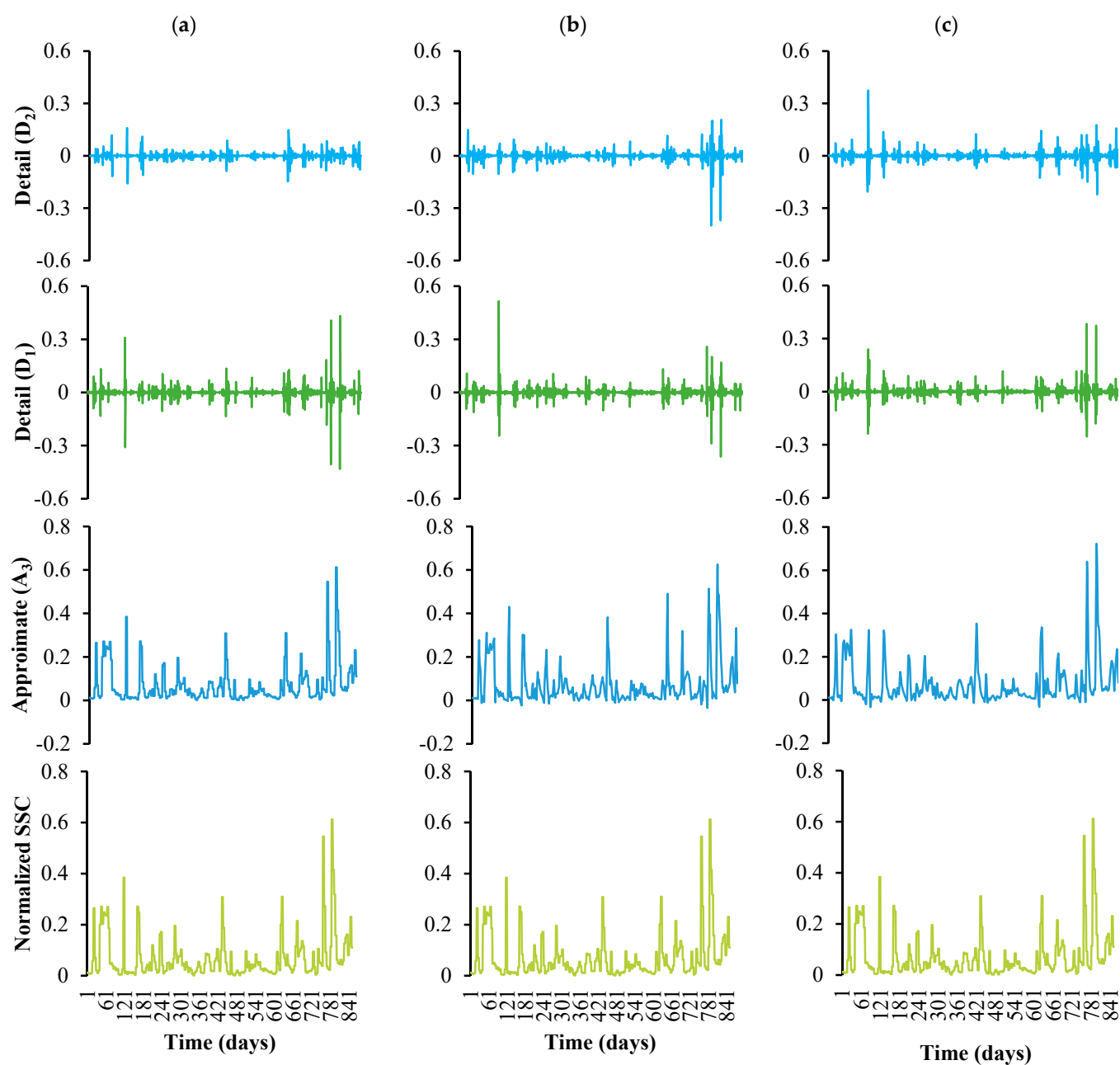


Figure 8. Decomposition of original SSC time series using different mother wavelets. (a) Haar, (b) db2, (c) coif2.

Each sub-signal plays a different role during prediction [30]. In WANN models, the number of neurons in the hidden layer varied from 1 to $2n + 1$ ($2 \times 18 + 1$). Hence, a total of 37 WANN models per single mother wavelet were developed to assess their daily SSC prediction performance.

3.4. Quantitative Performance Evaluation of Developed Models for Daily SSC Prediction

Quantitative performance evaluation indicators like RMSE, r , WI, CE, and PARE were used to assess developed models' predictive performance. Different AI techniques were employed for that purpose. The single best model was selected from each technique presented in Tables 4 and 5.

Table 4. Quantitative performance evaluation indices of the best selected Artificial Neural Network (ANN) and wavelet coupled ANN (WANN) models.

Model	Architecture	Training					Testing				
		RMSE (g/L)	r	WI	CE	PARE (%)	RMSE (g/L)	r	WI	CE	PARE (%)
ANN-3	6-3-1	0.040	0.81	0.78	0.655	0.0017	0.078	0.75	0.73	0.560	−0.013
Haar-WANN-21	18-21-1	0.042	0.80	0.73	0.633	0.0041	0.069	0.83	0.74	0.661	0.006
db2-WANN-21	18-21-1	0.031	0.90	0.72	0.801	0.0236	0.061	0.86	0.75	0.734	−0.012
coif2-WANN-11	18-11-1	0.021	0.95	0.81	0.903	0.0134	0.065	0.84	0.76	0.696	−0.017

Table 5. Quantitative performance evaluation indices of the best selected ANFIS and WANFIS models.

Model	Training					Testing				
	RMSE (g/L)	R	WI	CE	PARE (%)	RMSE (g/L)	r	WI	CE	PARE (%)
ANFIS-29	0.041	0.80	0.78	0.638	6.6×10^{-9}	0.080	0.78	0.73	0.545	0.060
Haar-WANFIS-27	0.032	0.88	0.83	0.782	1.7×10^{-9}	0.074	0.82	0.76	0.605	0.017
db2-WANFIS-25	0.023	0.94	0.85	0.892	2.1×10^{-10}	0.068	0.84	0.77	0.666	0.051
coif2-WANFIS-43	0.029	0.91	0.83	0.821	3×10^{-11}	0.060	0.87	0.80	0.745	0.015

After comparing the quantitative performance among 13 simple ANN models, it was revealed that RMSE, r, WI, CE, and PARE varies from 0.04 g/L to 0.067 g/L and 0.078 g/L to 0.119 g/L, 0.48 to 0.83 and 0.32 to 0.75, 0.56 to 0.78 and 0.56 to 0.73, 0.063 to 0.67 and −0.016 to 0.560, −0.09% to 0.032% and −0.189% to 0.089% during training and testing, respectively. Among 13 simple ANN models, the ANN-3 model with architecture (6-3-1) was the superior model compared to others based on quantitative performance criteria. The results obtained were extremely acceptable and agree with those suggested by Rajae [46], who applied the ANN for predicting daily SS under different time series and found that determination coefficients (R²) ranged from 0.12 to 0.87 for training and from 0.10 to 0.83 for testing datasets. Moreover, these results coincide with Sudhishri, et al. [47] observation, who investigated that ANN produced high correlations for predicting SS varied from 0.81 to 0.83. From 0.74 to 0.75 for training and validation periods, respectively. Similarly, from all 37 Haar-WANN models, it was observed that RMSE, r, WI, CE, and PARE ranges from 0.026 to 0.083 g/L and 0.069 g/L to 0.120 g/L, 0.22 to 0.93 and 0.55 to 0.85, 0.38 to 0.87 and 0.48 to 0.75, −0.454 to 0.852 and −0.039 to 0.661, −0.166% to 0.142% and −0.181% to 0.185% during training and testing, respectively. Out of 37 Haar-WANN models, the Haar-WANN-21 model with architecture (18-21-1) was selected as the best among others. The findings were parallel to Nourani and Komasi [26] results, who predicted daily multi-step ahead SS based on WANN. Their results confirmed high correlations from 0.6 to 0.89 for calibration, and from 0.55 to 0.85 for verification periods were obtained. These findings were similar to those revealed by Rajae [46], who used several mother wavelets for daily SS simulation and stated that Haar-WANN at ANN structure (6-4-1) generated R² 0.63 RMSE of 4321 ton/day. Additionally, research outcomes are in line with those who evaluated wavelet-based ANN models for estimating daily SS and obtained a positive relationship (R² = 0.70) between the observed–predicted data. By comparing prediction performance among 37 db2-WANN models, it was revealed that RMSE, r, WI, CE, and PARE varies from 0.015 g/L to 0.042 g/L and 0.061 g/L to 0.132 g/L, 0.81 to 0.98 and 0.43 to 0.86, 0.65 to 0.90 and 0.62 to 0.76, 0.628 to 0.952 and −0.025 to 0.734, −0.018% to 0.054% and −0.071% to 0.138% during training and testing, respectively and db2-WANN-21 model with architecture (18-21-1) was observed to be the best model for SSC prediction. These models generated more favorable outcomes compared with the results of Rajae [46], who observed that db2-WANN at ANN structure (6-1-1) achieved R² of 0.55 and RMSE of 4751.8 ton/day. Similarly, from all 37 coif2-WANN models, it was observed that RMSE, r,

WI, CE, and PARE varies between 0.020 g/L to 0.041 g/L and 0.065 g/L to 0.120 g/L, 0.86 to 0.96 and 0.49 to 0.84, 0.69 to 0.88 and 0.62 to 0.76, 0.650 to 0.915 and -0.033 to 0.696 , -0.033% to 0.056% and -0.10% to 0.123% during training and testing, respectively. By comparing all coif2-WANN models' predictive performance, the coif2-WANN-11 model with architecture (18-11-1) was observed to be the best. These results agree with Rajaei [46], who found that the coif-WANN gave satisfactory results ($R^2 = 0.74$, and $RMSE = 3601.1$ ton/day) at ANN structure 4-1-1. After analyzing the quantitative prediction performance among all 48 simple ANFIS models, it was revealed that RMSE, r , WI, CE, and PARE varies between 0.033 g/L to 0.044 g/L and 0.080 g/L to 1.23 g/L, 0.77 to 0.88 and -0.43 to 0.78 , 0.71 to 0.80 and 0.27 to 0.74, 0.587 to 0.772 and -107.9 to 0.545 , $-1 \times 10^{-8}\%$ to $2.3 \times 10^{-8}\%$ and -0.736% to 0.975% during training and testing, respectively. Among all 48 simple ANFIS models, the ANFIS-29 model (triangular input MF, constant output MF and 3 MFs/input) performed better during the training and testing period. Similarly, for all 48 Haar-WANFIS models, it was found that RMSE, r , WI, CE, and PARE ranges from 0.016 g/L to 0.050 g/L and 0.074 g/L to 0.288 g/L, 0.69 to 0.97 and 0.1 to 0.82, 0.63 to 0.88 and 0.5 to 0.76, 0.47 to 0.947 and -4.955 to 0.605 , $-4 \times 10^{-9}\%$ to $2.5 \times 10^{-9}\%$ and -0.105% to 0.094% during training and testing, respectively. Here, the Haar-WANFIS-27 model (triangular input MF, linear output MF, and 4 MFs/input) performed better during both the training and testing period. After comparing the prediction performance of all 48 db2-WANFIS models, it was observed that RMSE, r , WI, CE, and PARE varies between 0.012 g/L to 0.087 g/L and 0.068 g/L to 0.808 g/L, 0.11 to 0.98 and -0.51 to 0.84 , 0.00 to 0.91 and 0.24 to 0.77, -0.607 to 0.969 and -45.92 to 0.666 , $-5 \times 10^{-9}\%$ to 0.029% and -0.96% to 0.115% during training and testing, respectively. Here, the db2-WANFIS-25 model (triangular input MF, linear output MF and 2 MFs/input) performed better during both the training and testing period. Similarly, among all 48 coif2-WANFIS models, it was revealed that RMSE, r , WI, CE, and PARE ranges between 0.013 g/L to 0.045 g/L and 0.06 g/L to 0.392 g/L, 0.76 to 0.98 and -0.40 to 0.87 , 0.69 to 0.92 and 0.41 to 0.80, 0.573 to 0.964 and -10.04 to 0.745 , $-5 \times 10^{-9}\%$ to $8.9 \times 10^{-8}\%$ and -0.398% to 0.07% during training and testing, respectively. Among all 48 coif2-WANFIS models, the coif2-WANFIS-43 model (z input MF, linear output MF and 2 MFs/input) performed better for both the training and testing period. Our results are in line with Rajaei, et al. [48], who found high linear relationships from R^2 of 0.72 to 0.87 and RMSE ranged from 1805.3 ton/day to 2459.6 ton/day applying ANFIS. While the R^2 was from 0.62 to 0.67, and RMSE varied from 2543 ton/day to 2838 ton/day by using WANFIS.

3.5. Qualitative Performance Evaluation of Developed Models for Daily SSC Prediction

Qualitative performance evaluation was carried out by comparing the scatter plots and time series (sediment) graphs of predicted versus observed SSC during the testing period, as shown in Figures 9 and 10, respectively.

The best models selected from each technique as per quantitative evaluation were also assessed qualitatively to analyze their capability to capture the observed SSC time-series graph. The capability of the best-selected models for capturing low, medium, and high (peak) SSC values was examined using a time-series graph and observing the regression line's shifting from a 1:1 line in scatter plots. Here, it was revealed that the best selected coif2-WANFIS-43 model is more accurate to capture the overall shape of the observed SSC time series. The selected coif2-WANFIS-43 model closely predicted low, medium, and high SSC values with a strong R^2 of 0.76 during the testing period compared to other selected models.

All the observed and predicted data points are concentrated near the 1:1 line. The regression line is not too much shifted above or below from the 1:1 line than other models. Hence, it is revealed that the coif2-WANFIS-43 model can predict low, medium, and peak SSC values of the Koyna River basin.

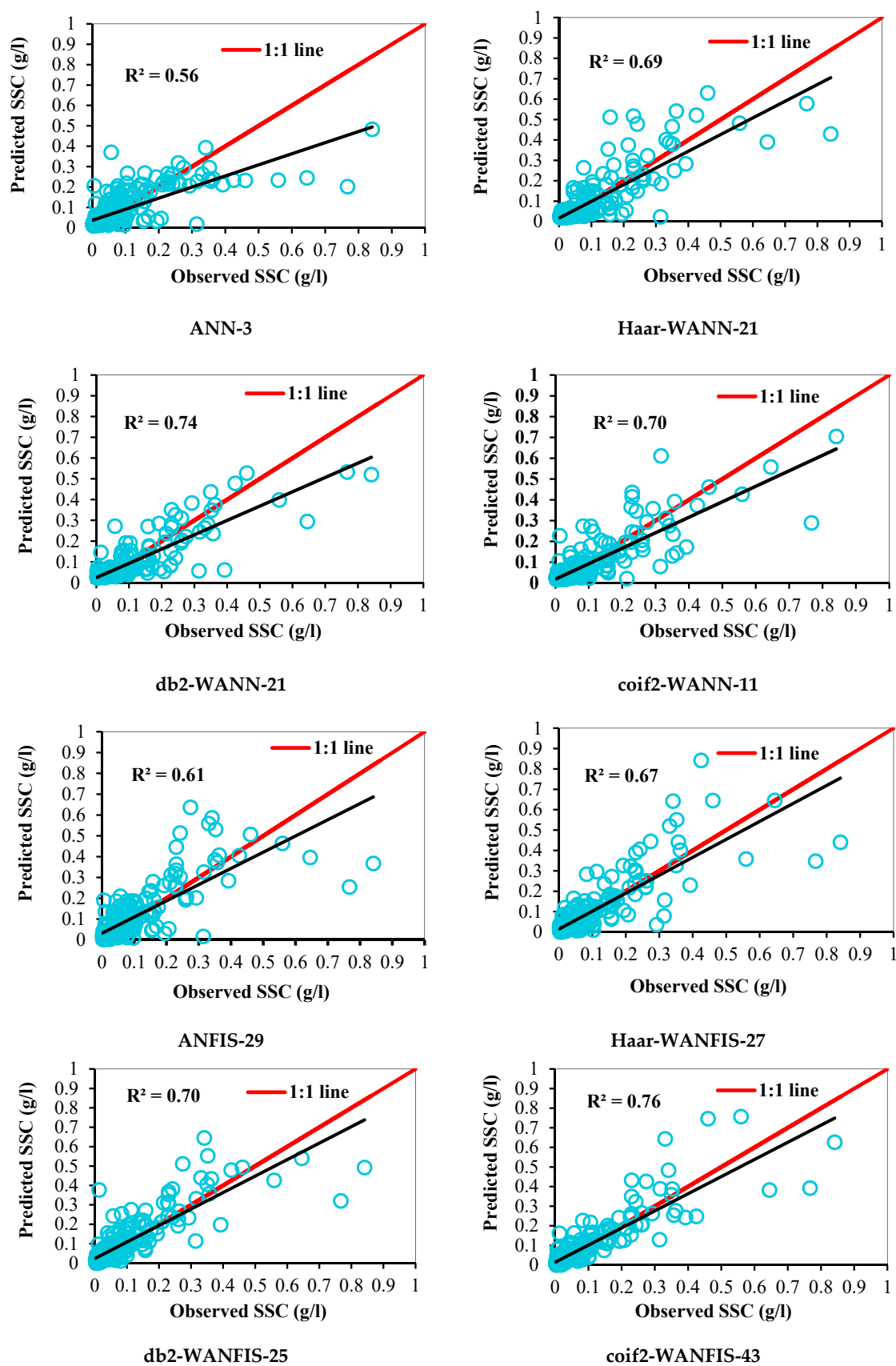


Figure 9. Scatter plots showing the predictive performance of different models during testing.

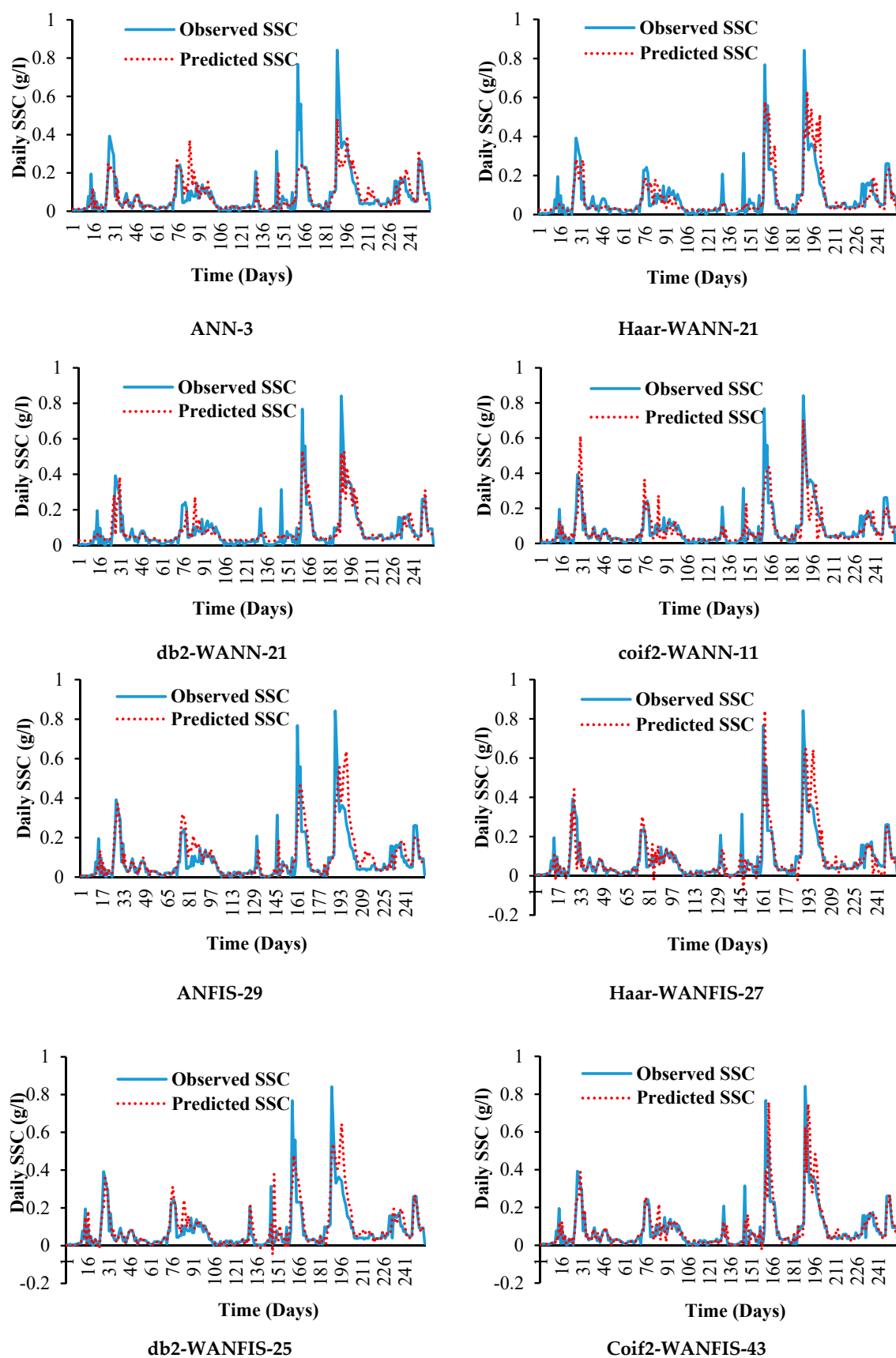


Figure 10. Actual daily SSC versus the predicted of different models.

3.6. Uncertainty Analysis

The strength of correlation between inputs and output determines the amount of uncertainty present in any prediction model. The amount of uncertainty incorporated in the best-selected models during the testing period is presented in Table 6.

Table 6. Uncertainty analysis of the best-selected models during the testing period for daily SSC predictions.

Model	We (g/L)	Cle(g/L)
ANN-3	±0.0096	−0.0125 to 0.0067 (0.0192)
Haar-WANN-121	±0.0084	−0.0071 to 0.0097 (0.0168)
db2-WANN-21	±0.0074	−0.0100 to 0.0049 (0.0149)
coif2-WANN-11	±0.0080	−0.0118 to 0.0042 (0.0160)
ANFIS-29	±0.0096	0.0226 to 0.0034 (0.0260)
Haar-WANFIS-27	±0.0091	−0.0054 to 0.0128 (0.0182)
db2-WANFIS-25	±0.0083	0.0028 to 0.0193 (0.0221)
coif2-WANFIS-43	±0.0073	−0.0041 to 0.0105 (0.0146)

Better prediction models must have small We and Cle. In this study, the coif2-WANFIS-43 model represented significantly less uncertainty in prediction, indicated by smaller We (± 0.0073 g/L) and narrower Cle (0.0146 g/L) other best-selected models. All WANN models have less uncertainty than a simple ANN model for prediction purposes.

Similarly, all WANFIS models have less uncertainty than the simple ANFIS model for daily SSC prediction. Finally, it is concluded that the coif2-WANFIS-43 model is less uncertain than others. Hence, it could provide accurate daily SSC predictions. Hence, in this study, it is revealed that data pre-processing using DWT is essential to improve the model predictive accuracy.

3.7. Sensitivity Analysis

In this study, the sensitivity analysis was completed to determine the most effective hydrologic variable of the original time series data for daily SSC prediction. The best selected simple ANN-3 model was selected for detecting the critical variable. In this study, the approach given by Azamathulla, et al. [49] was used for sensitivity analysis. In this approach, a total of 10 scenarios were considered. During each scenario, the developed model performance was tested by varying the input variables one by one. During the first scenario, the developed model performance was tested by using all the input variables. During the second scenario, the developed model performance was tested by removing only one input variable and keeping all other input variables. This process of removing a single input variable and keeping all other input variables as it was continued until the removal of all the input variables one by one and tested its performance simultaneously. The model's predictive performance will be reduced by removing a single input variable from the selected input vector. After studying the degree of reduction in any model's prediction performance during the absence of any single input variable, the degree of sensitivity of such an absent input variable can be determined. In this way, the essential input variable for daily SSC prediction can be determined. Hence, this approach was used here. The prediction performance observed after removing all input variables, one by one, is presented in Table 7.

Table 7. Sensitivity analysis for detection of the most crucial variable for daily SSC prediction.

Inputs	Training					Testing				
	RMSE (g/L)	r	WI	CE	PARE (%)	RMSE (g/L)	r	WI	CE	PARE (%)
$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t, S_{t-1}$	0.040	0.81	0.78	0.655	0.0017	0.078	0.75	0.73	0.560	−0.013
$R_{t-1}, R_{t-2}, R_{t-3}, Q_t, S_{t-1}$	0.043	0.78	0.72	0.612	0.0037	0.091	0.71	0.67	0.409	0.032
$R_t, R_{t-2}, R_{t-3}, Q_t, S_{t-1}$	0.040	0.82	0.77	0.666	0.0015	0.080	0.73	0.72	0.539	0.008
$R_t, R_{t-1}, R_{t-3}, Q_t, S_{t-1}$	0.043	0.79	0.68	0.613	0.0138	0.089	0.66	0.68	0.427	−0.054
$R_t, R_{t-1}, R_{t-2}, Q_t, S_{t-1}$	0.045	0.77	0.63	0.570	0.0327	0.096	0.68	0.67	0.342	0.054
$R_t, R_{t-1}, R_{t-2}, R_{t-3}, S_{t-1}$	0.041	0.80	0.73	0.639	0.0207	0.087	0.69	0.69	0.455	−0.046
$R_t, R_{t-1}, R_{t-2}, R_{t-3}, Q_t$	0.045	0.75	0.69	0.563	0.0071	0.108	0.45	0.59	0.160	−0.027

Here, it was revealed that in the absence of the previous 1-day SSC (S_{t-1}), the selected ANN-3model's predictive performance decreased dramatically. Therefore, it is concluded that S_{t-1} is the most critical hydrologic variable in the daily SSC prediction. The order of sensitivity for daily SSC prediction was observed to be S_{t-1} followed by R_{t-3} , R_t , R_{t-2} , Q_t , R_{t-1} .

4. Discussion

Comparison of ANN, WANN, ANFIS, and WANNFIS Models for Daily SSC Prediction

The best models selected after applying different techniques were compared with each other based on quantitative and qualitative performance evaluation criteria to select the single most accurate model with high SSC prediction performance. Different performance indicators like RMSE, r , WI, CE, PARE, and R^2 were used for that purpose. It was observed that WANN models performed better than simple ANN models. Among WANN models, db2-WANN models performed better than Haar- and coif2-WANN models. Similarly, WANNFIS models outperformed simple ANFIS models. Among all WANNFIS models, coif2-WANNFIS models performed better than Haar- and db2-WANNFIS models (Figure 11).

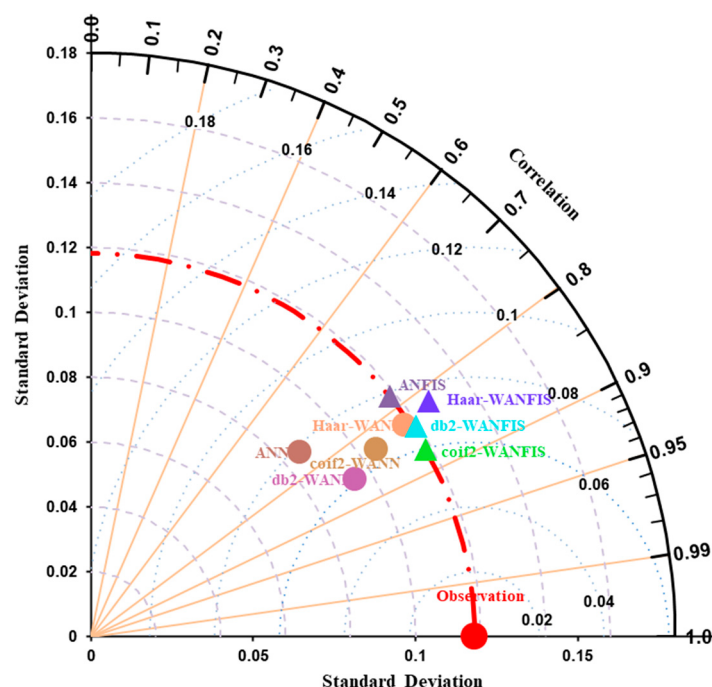


Figure 11. Taylor diagram of the developed models for daily SSC prediction.

Here, it was revealed that the coif2-WANNFIS model performed better than all other models. Here, it is observed that the wavelet-coupled hybrid models performed better than simple data-driven models. Based on sensitivity analysis, it was observed that the previous one day SSC is an essential input variable for daily SSC estimation. Koyna River basin is a very complex River basin having varying climatic and topographic conditions. Due to the complex nature of the basin, the sediment flow is highly dynamic. In this study, it is found that the coiflet mother wavelet-coupled ANFIS model is capable of predicting the highly dynamic behavior of the Koyna River basin. Hence, it can be concluded that the coiflet mother wavelet, coupled with the ANFIS model, can be able to predict the dynamic nature of the suspended sediment flow of any other complex river basin. Hence, it can be applied to other areas that have highly varying climatic and topographic conditions. Hence, the coiflet mother wavelet, coupled with the ANFIS model, is highly recommended for the river basins that come under varying climatic and topographic conditions.

5. Conclusions

In this study, the predictive performances of simple and wavelet-coupled AI models were analyzed for daily SSC prediction. In this study, 124 ANN/WANN and 192 ANFIS/WANFIS models were developed to predict daily SSC and tested their predictive performance. The developed models' capability to capture the observed sediment graph at low, medium, and peak SSC was evaluated. Finally, it is revealed that the hybrid wavelet-coupled AI models can model non-linear behavior between inputs and output variables.

Hybrid AI models outperformed simple AI models. Hydrologic processes such as runoff flow and sediment flow are highly complex and non-stationary; hence data pre-processing with DWT receives too much importance. Here, it was observed that simple AI models like ANN or ANFIS could not precisely model the hydrologic process without pre-processing original time series data. Out of 316 models, the coif2-WAFIS-43 model performed better based on quantitative and qualitative performance evaluation criteria. Hence, it could be applied for daily SSC prediction of the Koyna River basin. The final selected model is better, consistent, and accurate for daily SSC predictions among all other models. The reliability of the developed models was evaluated using uncertainty analysis. Here, it was revealed that simple AI models' reliability increased after coupling it with wavelet transform. The uncertainty analysis also indicated that the coif2-WAFIS-43 model is more reliable than others for daily SSC prediction. St-1 input variable was observed to be the most critical input variable, followed by Rt-3, Rt, Rt-2, Qt, Rt-1 for daily SSC modeling based on sensitivity analysis.

Here, in this study, it was found that the coiflet wavelet-coupled ANFIS model can predict sediment flow of highly complex river basin with varying climatic and topographic conditions. Hence, it is concluded that a coiflet mother wavelet, coupled with the ANFIS model, is more capable of modeling the dynamic behavior of complex river basins like the Koyna River basin. The future study should be focused on investigating the validity of this approach for multiple time steps (1, 2, 3 days/months ahead) SSC prediction. The impact of some additional inputs such as rainfall intensity, soil moisture content, maximum and minimum temperature, wind speed, relative humidity, bright sunshine hours, etc., should be assessed to improve the model's predictive efficiency. The applicability of other AI models like a genetic algorithm (GA), support vector machine (SVM), and wavelet-coupled SVM should be investigated to improve the prediction performance. Further, the result of the hybrid mentioned above approach should also be verified by selecting only those sub-signals with a higher correlation to output.

Author Contributions: Conceptualization, T.S.B. and P.K.; methodology, T.S.B. and P.K.; software, T.S.B., M.K., and P.K.; validation, T.S.B. and P.K.; formal analysis, T.S.B., P.K., A.E., and A.K.; investigation, T.S.B., P.K., A.E., and A.K.; writing—original draft preparation, T.S.B. and M.K.; writing—review and editing, T.S.B., P.K., M.K., and A.K.; visualization, T.S.B., P.K., A.E., and A.K.; supervision, P.K., A.E., and A.K.; project administration, A.K. funding acquisition, A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve this manuscript further. Alban Kuriqi was supported by a Ph.D. scholarship granted by Fundação para a Ciência e a Tecnologia, I.P.P (FCT), Portugal the Ph.D. Program FLUVIO—River Restoration and Management, grant number: PD/BD/114558/2016.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The activation function f reduces the amplitude of output and show its non-linearity in the form given as:

$$yf = (u) \quad (A1)$$

The neurons' performance can be altered by varying the transfer functions and changing the parameter like thresholds or gains. The processing at particular neurons occurs independently of all the other neurons of a given network. The processing completed at particular neurons can affect the whole neural network at the same time because the output of that particular neuron becomes an input to many other neurons of a neural network. Similarly, the neural network learns by changing the connection weights between the neurons. By using a suitable learning algorithm, the connection weights are altered using the training data set. The weights are frozen after the completion of learning. A feed-forward Multilayer Perceptron (MLP) neural network is the most widely used neural network in hydrology [11]. In the feed-forward network, the input vector is first forwarded to the output layer through a hidden layer using a non-linear transfer function that may be differentiable, continuous, or bounded. The output layer's error between simulated output and target output is propagated to adjust weights through some training mechanism. This process of "feed-forward" and "error back-propagation" is repeated until there is an acceptable reduction in error.

Appendix B

Layer 1: Every node in this layer creates a membership grade for an input variable. Every node i in this layer is a square node (or adaptive node), whose node function is defined as:

$$O_i^1(x) = \mu_{A_i}(x) \text{ for } i = 1, 2 \quad (A2)$$

$$O_i^1(x) = \mu_{B_{i-2}}(y) \text{ for } i = 3, 4 \quad (A3)$$

where $(x \text{ or } y)$ is the input to the i^{th} node, A_i (or B_{i-2}) is the fuzzy set associated with this node, and it is characterized by MFs shape. The MF may be any appropriate functions (continuous and piecewise differentiable) like Gaussian, Trapezoidal, Generalized bell, and Triangular shaped functions. Different researchers apply different MFs for the search for the solution to any problem. Considering, generalized bell function, the output of the first layer (i^{th} node) is determined as:

$$O_i^1(x) = \mu_{A_i}(x) = \frac{1}{1 + |(x - c_i)/a_i|^{2b_i}} \quad (A4)$$

where (a_i, b_i, c_i) is the premise parameter set, and by varying this premise parameter, the shape of MF can change. For a Gaussian function, the output of the first layer (i^{th} node) is determined as:

$$O_i^1(x) = \mu_{A_i}(x) = e^{-\frac{1}{2}(\frac{x-c_i}{\sigma_i})^2} \quad (A5)$$

where $(c_i \text{ or } \sigma_i)$ is the premise parameter set, and by varying this premise parameter, the shape of MF can change. Here, MFs center is represented by c while MFs width is represented by σ .

Layer 2: Every node in this layer is a fixed node labeled as II, whose output multiplies all incoming signals. The output O_i^2 indicates the firing strength of a rule, and it can be determined as:

$$O_i^2 = W_i = \mu_{A_i}(x)_i \mu_{B_i}(x)_i \text{ (} i = 1, 2 \text{)} \quad (A6)$$

Layer 3: Every node in this layer is a fixed node labeled as N. The i^{th} node in this layer determines the normalized firing strengths as:

$$O_i^3 = \bar{W}_i = \frac{W_i}{(W_1 + W_2)} \text{ (} i = 1, 2 \text{)} \quad (A7)$$

Layer 4: Every node i in this layer is a square node with a node function given as:

$$O_i^4 = \bar{W}_i f_i = \bar{W}_i (p_i x + q_i y + r_i) \quad (A8)$$

where $\bar{W}_i$ is a normalized firing strength obtained in layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set of this node (consequent parameters).

Layer 5: The single node in this layer is a fixed node labeled as sigma that computes the overall output as the summation of all incoming signals,

$$O_1^5 = \sum_i \bar{W}_i f_i = \frac{\sum_i W_i f_i}{\sum_i W_i} \quad (A9)$$

The single fixed node of this layer computes the final output by adding the entire incoming signal. In this manner, the input vector can be fed layer by layer into the ANFIS structure. The ANFIS uses a hybrid algorithm as a learning model for training, representing a combination of the least square and gradient descent method. The consequent parameters $\{p_i, q_i, r_i\}$ and premise parameters $\{a_i, b_i, c_i\}$ are required to be optimized. Consequent parameters are identified using the least square method during the forward pass of the hybrid learning approach when the node outputs move forward. The error signals are propagated backward during the backward pass. The premise parameters are adjusted by using the gradient descent method [22].

Appendix C

Gamma statistic Γ defines the determination of such a part of output variance, which cannot be accounted for in the smooth data model. Let, $b_{T[i,k]}$ represents k th ($1 \leq k \leq p$) nearest neighbor, in terms of Euclidean distance to b_i ($1 \leq i \leq N$). GT derived from Delta function of the input vector is given as:

$$\delta_N(k) = \frac{1}{N} \sum_{i=1}^N |b_{T[i,k]} - b_i|^2 \quad (A10)$$

where $|\dots|$ is the Euclidean distance and corresponding output values Gamma function is denoted as:

$$\gamma_N(k) = \frac{1}{2N} \sum_{i=1}^N |c_{T[i,k]} - c_i|^2 \quad (A11)$$

After applying the least square regression analysis, Γ value can be determined through p points (i.e., $\delta_N(k)$, $\gamma_N(k)$) which is given as:

$$\gamma = F\delta + \Gamma \quad (A12)$$

The intercept on the vertical axis ($\delta = 0$) gives the Γ value. It can be indicated as:

$$\Gamma_N(k) \text{ Var}(s) \text{ in probability as } \delta_N(k) \rightarrow 0$$

To standardize the results, another term, V_{ratio} , can be used. It ranges from 0 to 1, and it is defined as:

$$V_{\text{ratio}} = \frac{\Gamma}{\sigma^2(c)} \quad (A13)$$

where $\sigma^2(c)$ represents the variance of output (c). The value of V_{ratio} closer to 1 represents a higher degree of predictability.

Appendix D

Appendix D.1 Root Mean Squared Error (RMSE)

It is an indicator of determining the prediction accuracy of any model. It compares the values of observation versus prediction and finds the difference between them. RMSE represents how the data points are spread around the line of the best fit. It also measures

the average magnitude of the error. It gives positive value ranges between 0 to ∞ . For a perfect fit between observed and predicted values, the RMSE is zero. In contrast, the value of RMSE increases as the deviation increases between observed and predicted values [50]. RMSE can be calculated using the relationship as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_{oi} - Q_{pi})^2}{n}} \quad (A14)$$

Q_o is the observed value, Q_p is the predicted value, and n is the total number of values.

Appendix D.2 The Correlation Coefficient (r)

It shows the degree of closeness between observed and predicted values. It will be zero only if predicted, and observed values are entirely independent of each other. It can be determined using the following formula:

$$r = \left[\frac{\sum_{i=1}^n \{ (Q_o - \bar{Q}_o) (Q_p - \bar{Q}_p) \}}{\sqrt{\sum_{i=1}^n (Q_o - \bar{Q}_o)^2} \sqrt{\sum_{i=1}^n (Q_p - \bar{Q}_p)^2}} \right] \quad (A15)$$

where $\bar{Q}_o$ is the average of the observed values, $\bar{Q}_p$ is the average of the predicted values.

Appendix D.3 Willmott Index (WI)

Willmott (1981) first invented the Willmott index (index of agreement) to overcome the insensitivity of the coefficient of determination (R^2) and Nash-Sutcliffe efficiency. WI lies between 0 to 1; the higher WI values indicate that predicted values show better agreement than observed values. Legates and McCabe Jr. [51] introduced a modified WI followed by a generic form of WI proposed by Willmott [52] to overcome the limitations of original WI against extreme values, which is expressed as:

$$WI = 1 - \frac{\sum_{i=1}^n |Q_o - Q_p|}{\sum_{i=1}^n (|Q_p - \bar{Q}_o| + |Q_o - \bar{Q}_o|)} \quad (A16)$$

where $\bar{Q}_o$ is the average of the observed values. The advantage of modified WI is that the errors and differences are not inflated. Their squared values and differences and weights are provided with their appropriate weights. The modified WI also varies from 0 to 1, and the higher values indicate a better fitting model.

References

1. Kuriqi, A.; Koçileri, G.; Ardiçlioğlu, M. Potential of Meyer-Peter and Müller approach for estimation of bed-load sediment transport under different hydraulic regimes. *Model. Earth Syst. Environ.* **2020**, *6*, 129–137. [\[CrossRef\]](#)
2. Ardiçlioğlu, M.; Kuriqi, A. Calibration of channel roughness in intermittent rivers using HEC-RAS model: Case of Sarımsaklı creek, Turkey. *SN Appl. Sci.* **2019**, *1*, 1080. [\[CrossRef\]](#)
3. Adnan, R.M.; Liang, Z.; El-Shafie, S.; Zounemat, K.; Kisi, O. Prediction of Suspended Sediment Load Using Data-Driven Models. *Water* **2019**, *11*, 2060. [\[CrossRef\]](#)
4. Melesse, A.M.; Ahmad, S.; McClain, M.E.; Wang, X.; Lim, Y.H. Suspended sediment load prediction of river systems: An artificial neural network approach. *Agric. Water Manag.* **2011**, *98*, 855–866. [\[CrossRef\]](#)
5. Towfiqul Islam, A.R.M.; Talukdar, S.; Mahato, S.; Kundu, S.; Eibek, K.U.; Pham, Q.B.; Kuriqi, A.; Linh, N.T.T. Flood susceptibility modelling using advanced ensemble machine learning models. *Geosci. Front.* **2020**. [\[CrossRef\]](#)
6. Kuriqi, A.; Pinheiro, A.N.; Sordo-Ward, A.; Garrote, L. Water-energy-ecosystem nexus: Balancing competing interests at a run-of-river hydropower plant coupling a hydrologic-ecohydraulic approach. *Energy Convers. Manag.* **2020**, *223*, 113267. [\[CrossRef\]](#)
7. Buyukyildiz, M.; Kumcu, S.Y. An Estimation of the Suspended Sediment Load Using Adaptive Network Based Fuzzy Inference System, Support Vector Machine and Artificial Neural Network Models. *Water Resour. Manag.* **2017**, *31*, 1343–1359. [\[CrossRef\]](#)
8. Kisi, O. 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 neurones. *Hydrol. Sci. J.* **2005**, *50*, 696. [\[CrossRef\]](#)

9. Kumari, A.; Kumar, M.; Kumar, A.; Amitabh, A. Daily Gauge-Discharge Simulation using ANN and Wavelet-ANN Models for Muri Station, Jharkhand. *Indian J. Ecol.* **2020**, *47*, 645–649.
10. Zhu, Y.-M.; Lu, X.X.; Zhou, Y. Suspended sediment flux modeling with artificial neural network: An example of the Longchuanjiang River in the Upper Yangtze Catchment, China. *Geomorphology* **2007**, *84*, 111–125. [[CrossRef](#)]
11. Olyaie, E.; Banejad, H.; Chau, K.-W.; Melesse, A.M. A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: A case study in United States. *Environ. Monit. Assess.* **2015**, *187*, 189. [[CrossRef](#)] [[PubMed](#)]
12. Güldal, V.; Tongal, H. Comparison of Recurrent Neural Network, Adaptive Neuro-Fuzzy Inference System and Stochastic Models in Eğirdir Lake Level Forecasting. *Water Resour. Manag.* **2010**, *24*, 105–128. [[CrossRef](#)]
13. Valizadeh, N.; El-Shafie, A. Forecasting the Level of Reservoirs Using Multiple Input Fuzzification in ANFIS. *Water Resour. Manag.* **2013**, *27*, 3319–3331. [[CrossRef](#)]
14. Samantaray, S.; Ghose, D.K. Assessment of Suspended Sediment Load with Neural Networks in Arid Watershed. *J. Inst. Eng. Ser. A* **2020**, *101*, 371–380. [[CrossRef](#)]
15. Rajaei, T.; Jafari, H. Two decades on the artificial intelligence models advancement for modeling river sediment concentration: State-of-the-art. *J. Hydrol.* **2020**, *588*, 125011. [[CrossRef](#)]
16. Meshram, S.G.; Singh, V.P.; Kisi, O.; Karimi, V.; Meshram, C. Application of Artificial Neural Networks, Support Vector Machine and Multiple Model-ANN to Sediment Yield Prediction. *Water Resour. Manag.* **2020**, *34*, 4561–4575. [[CrossRef](#)]
17. Kumar, M.; Kumar, P. Daily Suspended-sediment Concentration simulation using ANN and Wavelet ANN models. *Int. Arch. Appl. Sci. Technol.* **2019**, *11*, 60–69.
18. Hussan, W.; Khurram Shahzad, M.; Seidel, F.; Nestmann, F. Application of Soft Computing Models with Input Vectors of Snow Cover Area in Addition to Hydro-Climatic Data to Predict the Sediment Loads. *Water* **2020**, *12*, 1481. [[CrossRef](#)]
19. Talebizadeh, M.; Morid, S.; Ayyoubzadeh, S.A.; Ghasemzadeh, M. Uncertainty Analysis in Sediment Load Modeling Using ANN and SWAT Model. *Water Resour. Manag.* **2010**, *24*, 1747–1761. [[CrossRef](#)]
20. Malik, A.; Kumar, A.; Kisi, O.; Shiri, J. Evaluating the performance of four different heuristic approaches with Gamma test for daily suspended sediment concentration modeling. *Environ. Sci. Pollut. Res.* **2019**, *26*, 22670–22687. [[CrossRef](#)]
21. Yin, D.; Shu, L.; Chen, X.; Wang, Z.; Mohammed, M.E. Assessment of Sustainable Yield of Karst Water in Huaibei, China. *Water Resour. Manag.* **2011**, *25*, 287–300. [[CrossRef](#)]
22. Seo, Y.; Kim, S.; Kisi, O.; Singh, V.P. Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. *J. Hydrol.* **2015**, *520*, 224–243. [[CrossRef](#)]
23. Singh, V.K.; Kumar, D.; Kashyap, P.S.; Kisi, O. Simulation of suspended sediment based on gamma test, heuristic, and regression-based techniques. *Environ. Earth Sci.* **2018**, *77*, 708. [[CrossRef](#)]
24. Agarwal, A.; Mishra, S.K.; Ram, S.; Singh, J.K. Simulation of Runoff and Sediment Yield using Artificial Neural Networks. *Biosyst. Eng.* **2006**, *94*, 597–613. [[CrossRef](#)]
25. Dadaser-Celik, F.; Cengiz, E. A neural network model for simulation of water levels at the Sultan Marshes wetland in Turkey. *Wetl. Ecol. Manag.* **2013**, *21*, 297–306. [[CrossRef](#)]
26. Nourani, V.; Komasi, M. A geomorphology-based ANFIS model for multi-station modeling of rainfall–runoff process. *J. Hydrol.* **2013**, *490*, 41–55. [[CrossRef](#)]
27. Samantaray, S.; Ghose, D.K. Evaluation of suspended sediment concentration using descent neural networks. *Procedia Comput. Sci.* **2018**, *132*, 1824–1831. [[CrossRef](#)]
28. Lohani, A.K.; Goel, N.K.; Bhatia, K.K.S. Improving real time flood forecasting using fuzzy inference system. *J. Hydrol.* **2014**, *509*, 25–41. [[CrossRef](#)]
29. Alizadeh, M.J.; Kavianpour, M.R.; Kisi, O.; Nourani, V. A new approach for simulating and forecasting the rainfall-runoff process within the next two months. *J. Hydrol.* **2017**, *548*, 588–597. [[CrossRef](#)]
30. Ravansalar, M.; Rajaei, T. Evaluation of wavelet performance via an ANN-based electrical conductivity prediction model. *Environ. Monit. Assess.* **2015**, *187*, 366. [[CrossRef](#)]
31. Shoaib, M.; Shamseldin, A.Y.; Melville, B.W. Comparative study of different wavelet based neural network models for rainfall–runoff modeling. *J. Hydrol.* **2014**, *515*, 47–58. [[CrossRef](#)]
32. Shoaib, M.; Shamseldin, A.Y.; Khan, S.; Khan, M.M.; Khan, Z.M.; Sultan, T.; Melville, B.W. A Comparative Study of Various Hybrid Wavelet Feedforward Neural Network Models for Runoff Forecasting. *Water Resour. Manag.* **2018**, *32*, 83–103. [[CrossRef](#)]
33. Grossmann, A.; Morlet, J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *Fundam. Pap. Wavelet Theory* **1984**, *15*, 723–736. [[CrossRef](#)]
34. Kisi, O. Wavelet Regression Model as an Alternative to Neural Networks for River Stage Forecasting. *Water Resour. Manag.* **2011**, *25*, 579–600. [[CrossRef](#)]
35. Partal, T.; Kişi, Ö. Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *J. Hydrol.* **2007**, *342*, 199–212. [[CrossRef](#)]
36. Nayak, P.C.; Venkatesh, B.; Krishna, B.; Jain, S.K. Rainfall-runoff modeling using conceptual, data driven, and wavelet based computing approach. *J. Hydrol.* **2013**, *493*, 57–67. [[CrossRef](#)]
37. Prahlada, R.; Deka, P.C. Forecasting of Time Series Significant Wave Height Using Wavelet Decomposed Neural Network. *Aquat. Procedia* **2015**, *4*, 540–547. [[CrossRef](#)]

38. Hassan, M.; Ali Shamim, M.; Sikandar, A.; Mehmood, I.; Ahmed, I.; Ashiq, S.Z.; Khitab, A. Development of sediment load estimation models by using artificial neural networking techniques. *Environ. Monit. Assess.* **2015**, *187*, 686. [[CrossRef](#)]
39. Kişi, Ö. Daily suspended sediment estimation using neuro-wavelet models. *Int. J. Earth Sci.* **2010**, *99*, 1471–1482. [[CrossRef](#)]
40. Genç, O.; Kişi, Ö.; Ardiçlıoğlu, M. Determination of Mean Velocity and Discharge in Natural Streams Using Neuro-Fuzzy and Neural Network Approaches. *Water Resour. Manag.* **2014**, *28*, 2387–2400. [[CrossRef](#)]
41. Sharghi, E.; Nourani, V.; Najafi, H.; Molajou, A. Emotional ANN (EANN) and Wavelet-ANN (WANN) Approaches for Markovian and Seasonal Based Modeling of Rainfall-Runoff Process. *Water Resour. Manag.* **2018**, *32*, 3441–3456. [[CrossRef](#)]
42. Nourani, V.; Kisi, Ö.; Komasi, M. Two hybrid Artificial Intelligence approaches for modeling rainfall-runoff process. *J. Hydrol.* **2011**, *402*, 41–59. [[CrossRef](#)]
43. Ebtehaj, I.; Bonakdari, H. Performance Evaluation of Adaptive Neural Fuzzy Inference System for Sediment Transport in Sewers. *Water Resour. Manag.* **2014**, *28*, 4765–4779. [[CrossRef](#)]
44. Jang, J.R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]
45. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
46. Rajaei, T. Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. *Sci. Total Environ.* **2011**, *409*, 2917–2928. [[CrossRef](#)]
47. Sudhishri, S.; Kumar, A.; Singh, J.K. Comparative Evaluation of Neural Network and Regression Based Models to Simulate Runoff and Sediment Yield in an Outer Himalayan Watershed. *J. Agric. Sci. Technol.* **2016**, *18*, 681–694.
48. Rajaei, T.; Mirbagheri, S.A.; Nourani, V.; Alikhani, A. Prediction of daily suspended sediment load using wavelet and neurofuzzy combined model. *Int. J. Environ. Sci. Technol.* **2010**, *7*, 93–110. [[CrossRef](#)]
49. Azamathulla, H.M.; Haghiabi, A.H.; Parsaie, A. Prediction of side weir discharge coefficient by support vector machine technique. *Water Supply* **2016**, *16*, 1002–1016. [[CrossRef](#)]
50. Kumar, M.; Kumari, A.; Kushwaha, D.P.; Kumar, P.; Malik, A.; Ali, R.; Kuriqi, A. Estimation of Daily Stage–Discharge Relationship by Using Data-Driven Techniques of a Perennial River, India. *Sustainability* **2020**, *12*, 7877. [[CrossRef](#)]
51. Legates, D.R.; McCabe, G.J., Jr. Evaluating the use of “goodness-of-fit” Measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [[CrossRef](#)]
52. Willmott, C.J. On the Evaluation of Model Performance in Physical Geography. In *Spatial Statistics and Models*; Gaile, G.L., Willmott, C.J., Eds.; Springer: Dordrecht, The Netherlands, 1984; pp. 443–460. [[CrossRef](#)]