



OPEN

A comparative survey between cascade correlation neural network (CCNN) and feedforward neural network (FFNN) machine learning models for forecasting suspended sediment concentration

Bhupendra Joshi¹, Vijay Kumar Singh², Dinesh Kumar Vishwakarma³✉, Mohammad Ali Ghorbani⁴, Sungwon Kim⁵, Shivam Gupta⁶, V. K. Chandola¹, Jitendra Rajput⁷, Il-Moon Chung⁸, Krishna Kumar Yadav^{9,10}, Ehsan Mirzania⁴, Nadhir Al-Ansari¹¹✉ & Mohamed A. Mattar¹²✉

Suspended sediment concentration prediction is critical for the design of reservoirs, dams, rivers ecosystems, various operations of aquatic resource structure, environmental safety, and water management. In this study, two different machine models, namely the cascade correlation neural network (CCNN) and feedforward neural network (FFNN) were applied to predict daily-suspended sediment concentration (SSC) at Simga and Jondhara stations in Sheonath basin, India. Daily-suspended sediment concentration and discharge data from 2010 to 2015 were collected and used to develop the model to predict suspended sediment concentration. The developed models were evaluated using statistical indices like Nash and Sutcliffe efficiency coefficient (N_{ES}), root mean square error (RMSE), Willmott's index of agreement (WI), and Legates–McCabe's index (LM), supplemented by a scatter plot, density plots, histograms and Taylor diagram for graphical representation. The developed model was evaluated and compared with CCNN and FFNN. Nine input combinations were explored using different lag-times for discharge (Q_{t-n}) and suspended sediment concentration (S_{t-n}) as input variables, with the current suspended sediment concentration as the desired output, to develop CCNN and FFNN models. The CCNN4 model with 4 lagged inputs (S_{t-1} , S_{t-2} , S_{t-3} , S_{t-4}) outperformed the other developed models with the lowest RMSE = 95.02 mg/l and the highest N_{ES} = 0.0.662, WI = 0.890 and LM = 0.668 for the Jondhara Station while the same CCNN4 model secure as the best with the lowest RMSE = 53.71 mg/l and the highest N_{ES} = 0.785, WI = 0.936 and LM = 0.788 for the Simga Station.

¹Department of Agricultural Engineering, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India. ²Department of Soil and Water Conservation Engineering, Acharya Narendra Deva University of Agriculture & Technology, Kumarganj, Ayodhya, Uttar Pradesh 224229, India. ³Department of Irrigation and Drainage Engineering, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, Uttarakhand 263145, India. ⁴Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz 5166616471, Iran. ⁵Department of Railroad Construction and Safety Engineering, Dongyang University, 36040 Yeongju, South Korea. ⁶Department of Irrigation and Drainage Engineering, Acharya Narendra Deva University of Agriculture & Technology, Kumarganj, Ayodhya, Uttar Pradesh 224229, India. ⁷Division of Agricultural Engineering, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India. ⁸Department of Water Resources and River Research, Korea Institute of Civil Engineering and Building Technology, Goyang-si 10223, Republic of Korea. ⁹Faculty of Science and Technology, Madhyanchal Professional University, Ratibad, Bhopal 462044, India. ¹⁰Environmental and Atmospheric Sciences Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Nasiriyah 64001, Iraq. ¹¹Department of Civil, Environmental, and Natural Resources Engineering, Lulea University of Technology, 97187 Luleå, Sweden. ¹²Department of Agricultural Engineering, College of Food and Agriculture Sciences, King Saud University, P.O. Box 2460, Riyadh 11451, Saudi Arabia. ✉email: dinesh.vishwakarma4820@gmail.com; nadhir.alansari@ltu.se; mmattar@ksu.edu.sa

The result shows the CCNN model was better than the FFNN model for predicting daily-suspended sediment at both stations in the Sheonath basin, India. Overall, CCNN showed better forecasting potential for suspended sediment concentration compared to FFNN at both stations, demonstrating their applicability for hydrological forecasting with complex relationships.

Keywords Cascade correlation neural network, Feedforward neural network, Suspended sediment concentration, Machine learning, Seonath basin

Abbreviations

SSC	Suspended sediment concentration
CCNN	Cascade correlation neural network
FFNN	Feed Forward neural network
N_{ES}	Nash and Sutcliffe efficiency coefficient
RMSE	Root mean square error
WI	Willmott's index of agreement
LM	Legates–McCabe's index
SRC	Sediment rating curve
MLR	Multiple linear regression
ML	Machine learning
ANN	Artificial neural network
FFBPNN	Feedforward backpropagation neural network
SVM	Support vector machine
R	Pearson's correlation coefficient
RSS	Random subspace
RF	Random forest
RBF	Radial basis function kernel
NPK	Normalized polynomial kernel
WANN	Wavelet based artificial neural network
SM-LSTM	Selective multimodal long short-term memory network
LSTM	Long short-term memory network
RNN	Recurrent neural network
MLP	Multilayer perceptron
DO	Dissolved oxygen
TDS	Total dissolved solid
pH	Potential of hydrogen
RBNN	Radial basis neural networks
WGRNN	Wavelet-generalized regression neural network
GRNN	Generalized regression neural network
Km	Kilometer
CWC	Central Water Commission

Suspended sediment is typically defined as sediment carried by a fluid in such a way that the force of turbulent eddies is stronger than the particles tendency to settle through the fluid¹. It affects in rivers significantly impact water quality^{2,3}. Precise prediction of suspended sediment load in rivers play a crucial role in both environmental science and the development of engineering infrastructure⁴. They are essential for effective watershed management strategies^{4,5}. Sediment outflow from the agricultural land due to rainfall and runoff action leads to a reduction in soil fertility^{6–9}. Sediment flows using two routes to reach the watershed outlet; the first is through suspension, and second is through rollover along the land surface as bed load^{10,11}. Since sedimentation can lead to floods as deposition of sediment in canal/stream/river, reservoir significantly decreases the depth of flow by virtue of rising in bed, and decrease in live storage capacity of reservoir⁴. In addition, sedimentation significantly affects the intakes of turbines for hydropower plants¹². Thus, accurate estimation of sediment outflow is desired for better planning, designing, and maintaining water resources structures for water supply, irrigation, drainage, flood control, soil and water conservation, and water quality control^{13–16}. In line with the requirement for effective tools for the prediction of sediment yield, it is becoming necessary to develop models capable of estimating sediment outflow¹⁷. Owing to the complex and nonlinearity of sediment models, it has always been challenging to develop model capable of forecasting exact amount of sediment outflow^{18,19}.

Many kinds of researches have been conducted for sediment modelling by using traditional mathematical models like sediment rating curve (SRC)^{20,21} and multiple linear regression (MLR)^{22,23}, and they concluded that these models were incapable of model sediment yield^{22–25}. Different conventional techniques were analysed to estimate discharge and suspended sediment concentration^{26–29}. The conventional models are less effective for sediment computation based on the previous researches. In recent years, machine learning (ML) techniques have been used to overcome problems faced when conventional modelling is attempted^{30–36}. Among various ML techniques, artificial neural network (ANN) is the most popular for estimating sediment load³⁷, and has provided good results compared to the traditional MLR and SRC methods^{38–44}.

Rahul et al.⁴⁵ compared feedforward backpropagation neural network (FFBPNN) and support vector machine (SVM) to forecast suspended sediment concentration at the Varanasi cross-section of the Ganga River. The results indicated that, for validation, the FFBPNN (RSME = 176.2, R = 0.955, and N_{ES} = 0.912) exhibited greater

precision in predicting suspended sediment load compared to SVM (RSME = 222.1, R = 0.930, and $N_{ES} = 0.864$). This study highlights the robustness of soft computing techniques for suspended sediment load prediction. The predictive capability of random subspace (RSS) for predicting suspended sediment load in the Haraz River, Iran, was compared with commonly used methods: random forest (RF) and two machine SVM models using radial basis function kernel (SVM-RBF) and normalized polynomial kernel (SVM-NPK)⁴⁶. The results revealed that the RSS model provided superior predictive accuracy ($N_{ES} = 0.83$) compared to SVM-RBF ($N_{ES} = 0.80$), SVM-NPK ($N_{ES} = 0.78$), and RF ($N_{ES} = 0.68$). Additionally, the RBF kernel showed better performance than the NPK kernel. Rajaei et al.⁴⁷ compared the wavelet based ANN (WANN), ANN, MLR and conventional sediment rating curve and found the performance of WANN better as compared to the ANN, MLR and conventional sediment RC techniques in the Yadkin River at Yadkin College, NC station in the USA. Sahoo et al.⁴⁸ compared selective multimodal Long Short-Term Memory network (SM-LSTM) framework with Long Short-Term Memory network (LSTM) and Recurrent Neural Network (RNN) models to forecast daily suspended sediment loads at two monitoring stations, namely Thebes on the Mississippi River and Omaha on the Missouri River. Comparative analysis of prediction accuracies highlighted that the SM-LSTM model significantly outperformed LSTM and RNN, showcasing its better ability to predict daily water level patterns. Sahoo et al.⁴⁸ emphasizes the potential of deep learning in environmental monitoring and management, particularly in predicting sediment dynamics, which is crucial for maintaining water quality and ecosystem health. Studies have shown the effectiveness of models like the radial M5 tree (RM5Tree) model⁴⁹, adaptive neuro-fuzzy models (ANFIS)^{50–53}, multilayer perceptron (MLP)⁵⁴, support vector machine (SVM) models^{55–59}, and the coupled Soil and Water Assessment Tool (SWAT) with long short-term memory (LSTM) model⁶⁰. These models have demonstrated accurate sediment yield estimations by utilizing hydro-meteorological variables like temperature, rainfall, discharge, and sediment data. The use of these advanced algorithms can provide reliable predictions even in data-scarce situations, as seen in various watershed studies, enhancing watershed management and engineering structure design, as evident from the research findings⁶¹.

In the current study, the potential of two different machine learning algorithms including cascade correlation neural network (CCNN) and feedforward neural network (FFNN) were investigated for forecasting daily-suspended sediment load in Sheonath basin, India. The CCNN model has potentially used to examine the capability for predicting/forecasting the different hydrological variables. Karunanithi et al.⁶² investigated the potential of CCNN model for discharge prediction at the Dexter station, Huron River. Alok et al.⁶³ predicted river flow of Bramani basin, India. Kim et al.⁶⁴ compared CCNN and multilayer perceptron (MLP) models for predicting daily evaporation, South Korea. Ghorbani et al.⁶⁵ applied CCNN and random forest (RF) models to predict daily river flow using stage-discharge at Dulhunty and Herbert stations, Australia. Also, Zounemat-Kermani et al.⁶⁶ examined the prediction of surface water quality parameters (e.g., water temperature, dissolved oxygen (DO), total dissolved solid (TDS), and pH etc.) in the St. Johns River, Florida. Similarly, FFNN model has been widely utilized by different researchers. Bilhan et al.⁶⁷ compared different models (FFNN and RBNN) with the conventional technique for simulating lateral outflow in channel and found FFNN model superior than RBNN (RMSE = 0.037). Kisi⁶⁸ applied WGRNN GRNN and FFNN models for the prediction of monthly streamflow in two different rivers and found WGRNN outperformed than the GRNN and FFNN model (RMSE = 5.31 m³/s, and R = 0.728). Zounemat-Kermani et al.⁶⁹ assessed the performance of FFNN model for predicting daily streamflow in Cahaba River, Alabama. Ehteram et al.⁷⁰ investigated FFNN model with an evolutionary algorithm for estimating suspended sediment concentration yield in the Atrek basin, Iran. Heddam et al.⁷¹ evaluated the potential of different machine learning models to predict phycocyanin pigment of surface water in the river basin.

The objective of this study is to investigate the use of machine learning models, specifically CCNN and FFNN, for forecasting daily-suspended sediment concentration in the Sheonath basin, India, with a focus on short time-series data. The study also compares the performance of these models and evaluates their suitability for practical application in hydrological organizations within the Sheonath basin. The novelty of current research work is to develop a suspended sediment concentration model based on short time-series data. A comparison of the CCNN models and the FFNN models was also made with the data generated by the corresponding CCNN and FFNN models, and the results were compared in the end. Its performance is assessed statistically and compared with observed data.

Methodology

Study area

The Sheonath River involves in the Rajnandgaon district, Chhattisgarh, India. The basin is bounded by latitude 20°15' N to 22° 02' N and longitude 80° 26' E to 81° 36' E. The total catchment area is about 30,858 km² (Fig. 1). The length of the river is approximate 379 km. The small tributaries including Arpa, Agrar, Tandula, Kharun, and Hump are associated with the mainstream of Sheonath river. The basin is situated in a tropical climate region. The southwest monsoon is responsible for most of the precipitation in the region. It starts in June and ends on October. Winter or cold season begins from November to February, and the lowest temperature can be found in January. Summer season starts from March to June, and the highest temperatures are measured in the last week of May and June months. The mean rainfall of Sheonath basin is around 1298.60 mm, 51.10 mm, 1132.40 mm, 75.40 mm, and 56.50 mm during the annual, pre-monsoon season, monsoon season, post-monsoon season and winter season, respectively. The area of land use is covered by forest (18.44%), agriculture (72.66%), urban area (2.94%), water (2.04), and barren land (3.92%).

Data collection

The daily hydrological data (i.e., streamflow and suspended sediment concentration) from 2010–2015 is gathered from the Central Water Commission (CWC), India. Simga and Jondhara stations are located in Sheonath basin.

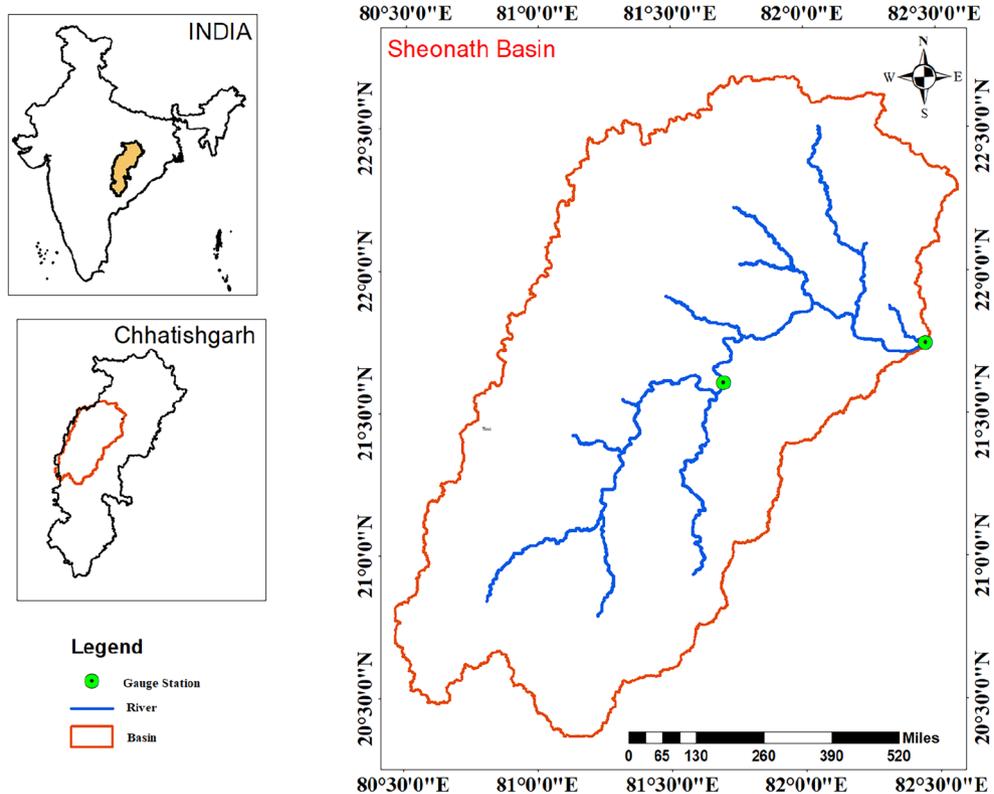


Figure 1. Map of Sheonath basin.

Simga station is situated at 21°37'37"N and 81°41'30"E in Raipur district. Jondhara station is an outlet of Sheonath basin, which is located at 21°42'47"N and 82°21'30"E in Bilaspur district. The suspended sediment concentration samples are collected using the observation of discharge every morning on 08:00 am. While, suspended sediment concentration samples are collected at 0.6 m depth of discharge where the velocity of discharge is measured. The observed data was obtained from a gauging station which was equipped with modern technology.

Machine learning models

In this study, two of the most common neural network/machine learning structures were chosen for modeling: the cascade correlation neural network (CCNN) and feedforward neural network (FFNN). Moreover, several pair of input combination were also used for forecasting daily suspended sediment concentration in Sheonath basin, India.

Cascade correlation neural network (CCNN)

A cascade correlation neural network comprises a cascade network, where hidden neurons are added to the hidden layer and do not change after they have been included^{65,72,73}. It is known as a cascade because the provision from all neuron's feeds into new neurons (Fig. 2a).

As new neurons are included to the hidden layer, the learning process expands the extent of connection between the new neurons, and the leftover error of system limits. The goal of this expands the extent of connection between the new neurons OU, the added total output units zero of the correlation degree between value of candidate units (U) and O_e, the output error (O_e) observed at unit zero. We describe OU as:

$$OU = \sum_o \left| \sum_v (U_v - \bar{U})(O_{e,o} - \bar{O}_o) \right| \tag{1}$$

where O = system output at which the inaccuracy is observed; e = the calibration array. Also, the amounts \bar{U} and \bar{O}_o are total mean arrays corresponding to the values of U and O_e.

To maximize OU, we must compute $\frac{\partial OU}{\partial w_i}$, the fractional derived of OU with regard to candidate unit's received weights (Wi). We can develop and separate the equation for OU to find.

$$\frac{\partial OU}{\partial w_i} = \varphi_o \sum_{e,o} (O_{e,o} - \bar{O}_e) d'_v I_{i,v} \tag{2}$$

where φ_o = the correlated signal between output O and candidate's value, d'_v = array v of activation functions the candidate units; $I_{i,v}$ = the input the candidate unit collects by I unit for array v.

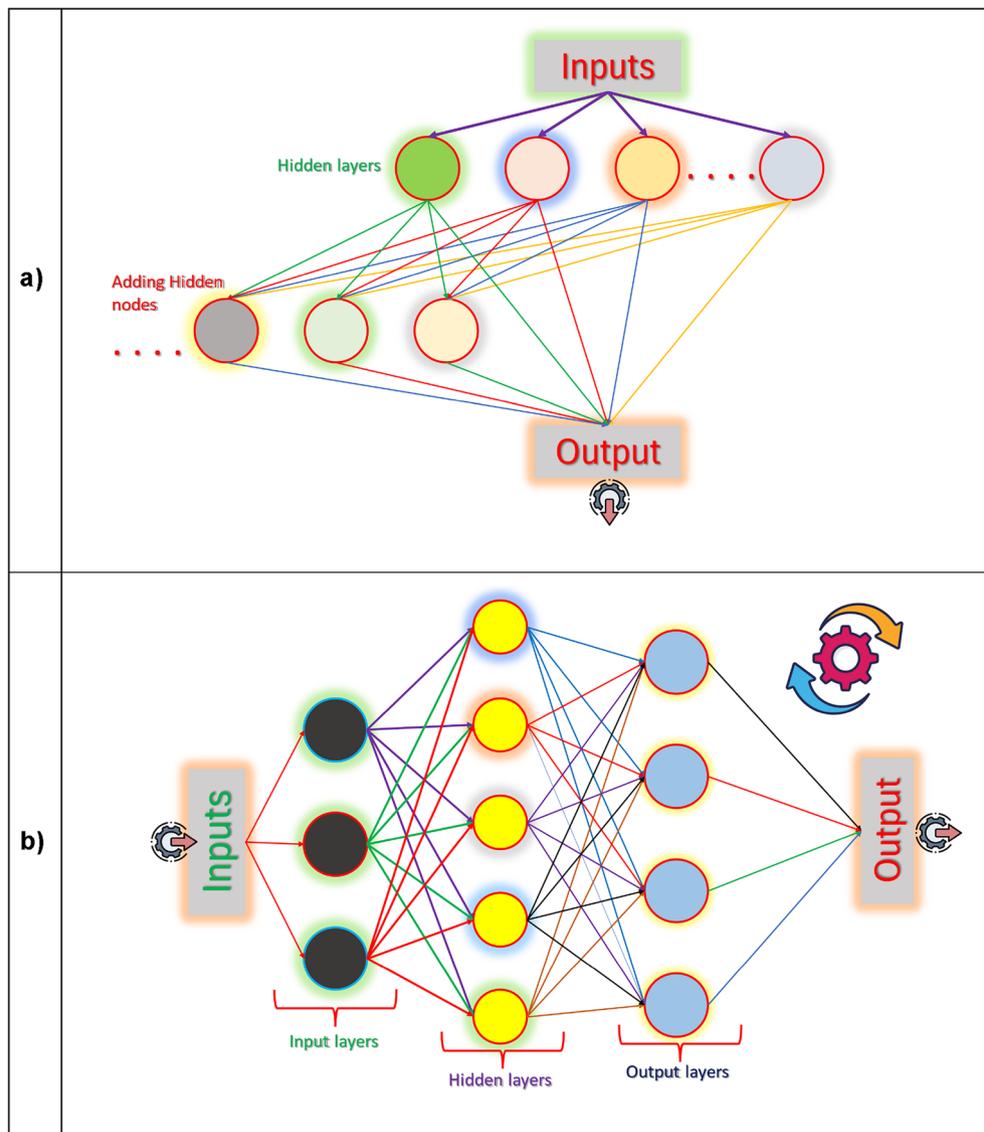


Figure 2. Structure of the models: (a) Network of cascade correlation neural network (CCNN) and (b) feedforward neural network (FFNN).

Feedforward neural network (FFNN)

The neural systems, as the name infers, are motivated by their natural partners, the organic mind, and the sensory system. Natural cerebrum is altogether not the same as the customary computerized digital computer as far as its structure and the manner in which it forms data^{74,75}. The essential structure of neural systems is a "neuron". A neuron can be seen as a handling unit. In a neural system, neurons are associated with each other through "synaptic weight"s, or "weight"s in short. Every neuron in a system gets "weighted" data by means of these synaptic associations from the neurons that it is associated with and produces a yield bypassing the weighted total of those input signals (either outside contributions from nature or the yields of other neurons) by an "activation function".

In the feedforward system, the hubs in the info (Input) layer get the information signals which are passed to the covered (hidden) layer and afterward to the yield layer (Fig. 2b). The signs are duplicated by the present estimations of loads, and afterward, the weighted information sources are added to yield the net contribution to every neuron of the following layer. The net contribution of a neuron is gone through an enactment or move capacity to deliver the yield of the neuron. The large numbers of literature are already published, therefore we cited only some literatures^{75–84}.

Statistical metrics for performance evaluation

The present study attempts to estimate of suspended sediment concentration by using a variety of discharge and suspended sediment concentration for different time lags variables as inputs to conduct the study. To evaluate the developed models, four distinct statistical performance metrics have been utilized in this study to assess their performance, error, and accuracy ability of predictive models. These three indices are: root mean squared

error (RMSE), Nash–Sutcliffe Efficiency Coefficient (N_{ES}), Willmott's Index (WI), and the Legates–McCabe's index (LM), calculated as follows:

Root Mean Square Error (RMSE)

RMSE is a measure of the average magnitude of the errors between predicted and observed values. It calculates the square root of the average squared differences between predicted and observed values over the entire dataset.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (SC_{Obs} - SC_{Pre})^2}{n}} \quad (3)$$

where SC_{Obs} is observed sediment concentration, SC_{Pre} is predicted sediment concentration, n is number of observations.

RMSE values range from zero to infinity. RMSE provides insight into the overall model accuracy, with lower values indicating better performance. However, RMSE does not distinguish between systematic and random errors^{85–87}.

Nash and Sutcliffe Efficiency Coefficient (N_{ES})

N_{ES} is a widely used metric for assessing the predictive accuracy of hydrological and environmental models. It compares the observed data to the model predictions and evaluates how well the model captures the variability of the observed data. It can be calculate using the following formula⁸⁸:

$$N_{ES} = 1 - \frac{\sum_{i=1}^n (SC_{Obs} - SC_{Pre})^2}{\sum_{i=1}^n (SC_{Obs} - \overline{SC_{Obs}})^2} \quad (4)$$

where SC_{Obs} is observed sediment concentration, SC_{Pre} is predicted sediment concentration, n is number of observations, $\overline{SC_{Obs}}$ is average of the observed sediment concentration and $\overline{SC_{Pre}}$ is average of predicted sediment concentration.

N_{ES} values range from negative infinity to 1, where a value of 1 indicates perfect agreement between the observed and predicted values, while values closer to 0 indicate poorer performance. N_{ES} is sensitive to errors in both magnitude and timing, making it a comprehensive measure of model performance^{86,89,90}.

Willmott's Index of Agreement (WI)

WI is a metric that evaluates the similarity between observed and predicted values relative to the range of variability in the observed data. WI is particularly useful for comparing models across different datasets and variable ranges. It can be calculate using the following formula^{90,91}:

$$WI = 1 - \frac{\sum_{i=1}^n (SC_{Obs} - SC_{Pre})^2}{\sum_{i=1}^n (|SC_{Pre} - \overline{SC_{Obs}}| + |SC_{Obs} - \overline{SC_{Obs}}|)^2} \quad (5)$$

where SC_{Obs} is observed sediment concentration, SC_{Pre} is predicted sediment concentration, n is number of observations, $\overline{SC_{Obs}}$ is average of the observed sediment concentration.

It ranges from 0 to 1, where 1 indicates perfect agreement and 0 indicates no agreement beyond the mean of the observed data. WI considers both systematic and random errors, making it a robust measure of model performance^{90,92}.

Legates–McCabe's Index (LM)

LM is another index that assesses the agreement between observed and predicted values. It provides insight into how well the model captures the variability and distribution of the observed data. It can be calculate using the following formula⁹³:

$$LM = 1 - \left[\frac{\sum_{i=1}^n |SC_{Obs} - SC_{Pre}|}{\sum_{i=1}^n |SC_{Obs} - \overline{SC_{Obs}}|} \right] \quad (6)$$

where SC_{Obs} is observed sediment concentration, SC_{Pre} is predicted sediment concentration, n is number of observations, $\overline{SC_{Obs}}$ is average of the observed sediment concentration.

It ranges from negative infinity to 1, with 1 indicating perfect agreement and values closer to 0 indicating poorer performance. LM is sensitive to systematic errors but less sensitive to random errors compared to other indices.

The following reference values for N_{ES} and WI statistical indices as: very good ($0.75 < N_{ES} \leq 1.0$); good ($0.65 < N_{ES} \leq 0.75$); satisfactory ($0.50 < N_{ES} \leq 0.65$); acceptable ($0.40 < N_{ES} \leq 0.50$), and unsatisfactory ($N_{ES} \leq 0.40$) describes how the N_{ES} and WI results were analyzed. Considering that RMSE near to zero, and the N_{ES} , WI, and LM values would be expected to be a unity for a perfect estimation model^{90,94–96}.

Model development

Daily discharge and suspended sediment concentration data (11/05/2010–10/31/2015) were divided into training (11/05/2010–11/01/2014) and testing (11/02/2014–10/31/2014) data at both stations. Training data contains 1458 data which are about 80% of total data and testing data involves 364 data which are nearly 20% of whole

data and explored using CCNN and FFNN models. Statistical analysis of observed data for training and testing phases has been carried out to determine the behaviour of data characteristics using mean, minimum, maximum, median, standard deviation, skewness, and kurtosis as given in Table 1.

The minimum values showed that there was a period when no discharge and suspended sediment concentration condition prevailed, while maximum values provided discharge and suspended sediment concentration values was fluctuating considerably during training and testing phases at Jondhara station (i.e., maximum discharge (training phase) and maximum suspended sediment concentration (testing phase) and Simga station (i.e., maximum discharge and suspended sediment concentration (testing phase)). Median values of discharge and suspended sediment concentration at Jondhara station were found zero which shows that half of the data are zero. Although, at Simga station, the median values are found as non-zero positive value (e.g., except for suspended sediment concentration during testing phase). Maximum and minimum mean values of discharge and suspended sediment concentration were found during training (Simga station) and testing phases (Jondhara station), respectively. Also, the values of standard deviation were calculated at both stations, and highest deviation in discharge was found during testing phase (Jondhara station), while highest deviation in suspended sediment concentration was found during testing phase (Jondhara station).

Also, the highest skewness in discharge was found during training phase at Simga station, while Jondhara station during training phase showed the highest skewness in suspended sediment concentration. Based on the statistical survey for discharge and suspended sediment concentration, it can be found that applying data is highly fluctuating and is not normally distributed.

Different combinations of lag-times discharge Q_{t-n} (i.e., 1- and 2-days) and suspended sediment concentration S_{t-n} (i.e., 1-, 2-, 3, and 4-days) as input variables and current suspended sediment concentration (S_t) as the desired variable were investigated to develop CCNN and FFNN models. Therefore, nine inputs combinations were developed based on correlation of lag-times discharge and suspended sediment concentration which is presented in Table 2.

First four combinations (i.e., combinations 1–4) were developed based on only lag-times suspended sediment concentration data while other combinations (i.e., combinations 5–9) developed using both discharge and suspended sediment concentration data.

Daily discharge and suspended sediment concentration data are plotted separately with time scale (training and testing phases) on X-axis and corresponding discharge or suspended sediment concentration on Y-axis at both stations. Time series plotting of suspended sediment concentration at Jondhara station clearly showed there was considerable suspended sediment concentration with maximum suspended sediment concentration in monsoon period (testing phase). However, the rest year were found as negligible suspended sediment concentration in Fig. 3. Also, at Jondhara station, the discharge was found only in monsoon period with maximum discharge (training phase). At Simga station, the peak discharge and suspended sediment concentration were found during testing phase. Figure 3 illustrates that many values in the dataset were zero due to the non-perennial

Station	Dataset	Data type	No of data	Minimum	Maximum	Median	Mean	Std deviation	Skewness	Kurtosis
Jondhara	Training	Sediment concentration (mg/L)	1458	0.00	988.00	0.00	80.19	168.90	2.20	4.07
		Discharge (m ³ /s)	1458	0.00	9193.26	0.00	284.51	703.93	4.76	37.52
	Testing	Sediment concentration (mg/L)	364	0.00	1220.00	0.00	82.79	168.92	3.04	11.67
		Discharge (m ³ /s)	364	0.00	6528.58	0.00	547.43	994.81	2.71	9.67
Simga	Training	Sediment concentration (mg/L)	1458	0.00	890.00	1.00	39.39	90.48	3.48	16.04
		Discharge (m ³ /s)	1458	0.00	7358.73	7.71	144.13	462.15	7.42	79.02
	Testing	Sediment concentration (mg/L)	364	0.00	863.00	0.00	73.58	116.07	1.99	6.18
		Discharge (m ³ /s)	364	0.00	6844.26	7.37	293.04	735.65	6.02	45.44

Table 1. The daily statistical parameters of sediment and discharge data sets.

No	Input combination	Output	FFNN	CCNN
1	S_{t-1}	S_t	FFNN1	CCNN1
2	S_{t-1}, S_{t-2}	S_t	FFNN2	CCNN2
3	$S_{t-1}, S_{t-2}, S_{t-3}$	S_t	FFNN3	CCNN3
4	$S_{t-1}, S_{t-2}, S_{t-3}, S_{t-4}$	S_t	FFNN4	CCNN4
5	S_{t-1}, Q_{t-1}	S_t	FFNN5	CCNN5
6	$S_{t-1}, S_{t-2}, Q_{t-1}$	S_t	FFNN6	CCNN6
7	$S_{t-1}, Q_{t-1}, Q_{t-2}$	S_t	FFNN7	CCNN7
8	$S_{t-1}, S_{t-2}, Q_{t-1}, Q_{t-2}$	S_t	FFNN8	CCNN8
9	$S_{t-1}, S_{t-2}, S_{t-3}, Q_{t-1}, Q_{t-2}$	S_t	FFNN9	CCNN9

Table 2. Input combinations.

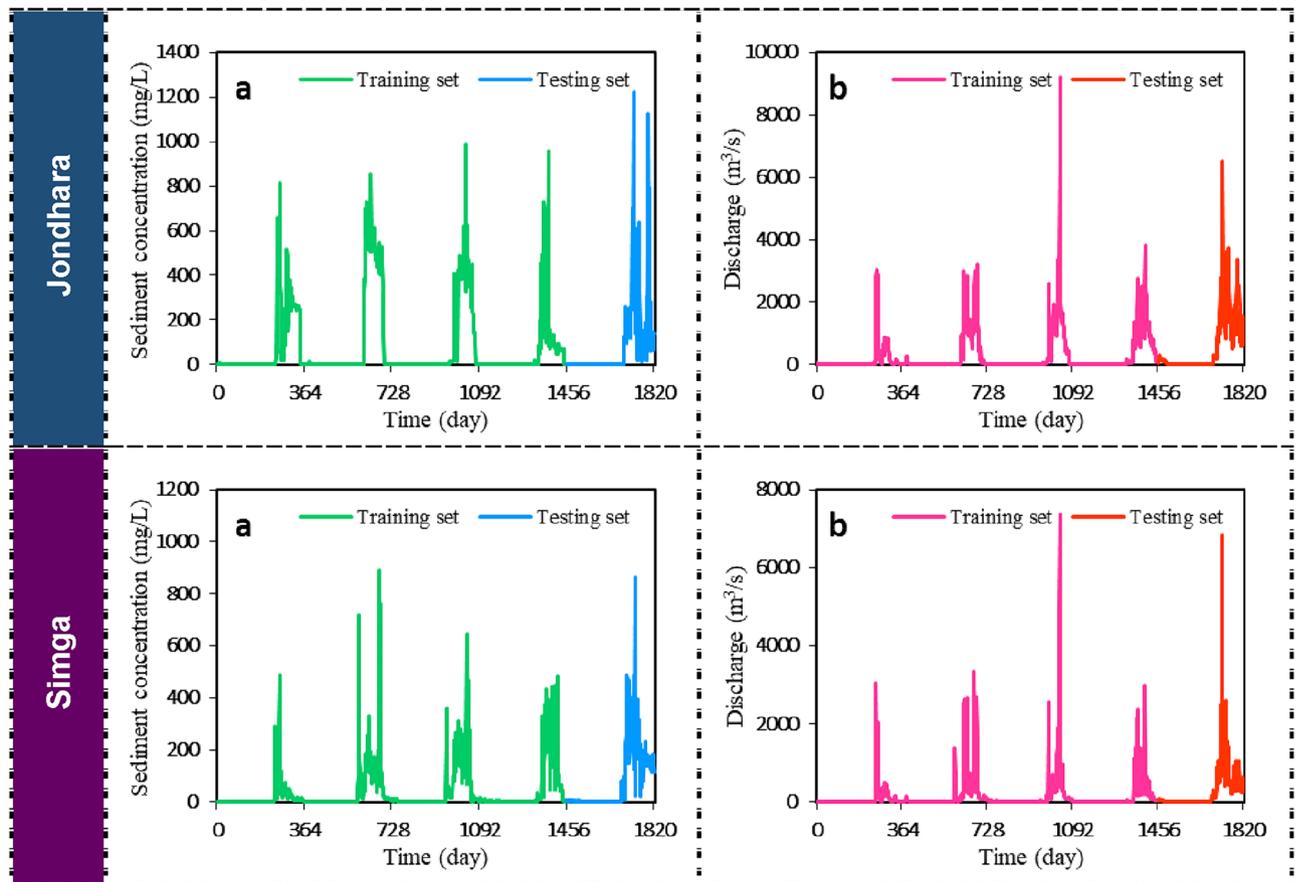


Figure 3. Time series plot for the data period of (2010/11/05–2015/10/31): (a) sediment; (b) discharge.

nature of the river, where continuous water flow was absent during certain periods. Consequently, sediment concentrations were also zero during these periods. This resulted in a mix of zero values and non-zero data in the dataset. The presence of zero values is further attributed to the validation data, which also captured periods of zero stream flow.

Results

Quantitative assessment of developed models based on statistical indices

All the input combinations were investigated using CCNN and FFNN models with one hidden layer and different numbers of neurons in hidden layer. FFNN models with selected input combinations were developed for determining a single hidden layer at both stations. The number of neurons was increased to improve the model performance. When the model performance was not improved by adding the neurons, the neurons were not added in the hidden layer. The best model was determined based on the results of performance during testing phase. The model with a minimum value of RMSE and maximum values of N_{ES} , WI and LM was selected as the best model and given in Table 3 at both stations.

All eighteen developed FFNN models developed precise daily suspended sediment concentration estimations during training with a range of RMSE (mg/L), N_{ES} , WI and LM between 46.610 to 59.930 (mean = 52.239) (mg/L), 0.700 to 0.900 (mean = 0.802), 0.902 to 0.973 (mean = 0.940) and 0.644 to 0.878 (mean = 0.775) respectively, and for testing with a range of 54.040 to 114.280 (mean = 82.986) (mg/L), 0.541 to 0.783 (mean = 0.671), 0.847 to 0.936 (mean = 0.898) and 0.441 to 0.794 (mean = 0.683). Based on the values of N_{ES} for the Jondhara station were found satisfactory (0.50 < N_{ES} ≤ 0.65) while for the Simga station N_{ES} were found good as 0.75 < N_{ES} ; and WI for both stations were higher than 0.75, showing high precision in predicting daily suspended sediment concentration.

FFNN8 model with four inputs and six neurons in the hidden layer (4-6-1 structure) at Jondhara station was carefully determined as the best on after comparing the performance of all models given in Table 3. For training phase, the values of RMSE, N_{ES} , WI, and LM were found as 58.77 mg/L, 0.879, 0.966, and 0.966, respectively, while for testing phase, the values provided as 107.86 mg/L, 0.591, 0.867, and 0.588, respectively, for the selected FFNN8 model. While, FFNN5 model performed worst with RMSE, N_{ES} , WI and LM values of 56.58 mg/L, 0.888, 0.969, 0.846 during training phase and 114.28 mg/L, 0.541, 0.872, and 0.576 during testing phase of the developed models. After comparison, FFNN3 model was chosen as the best model at Simga station with the values of RMSE, N_{ES} , WI, and LM as 48.58 mg/L, 0.712, 0.910, and 0.741 during training phase and 54.04 mg/L, 0.783, 0.936, and 0.780, during the testing phase, respectively. While, FFNN1 model performed worst with RMSE, N_{ES} , WI and LM values of 49.54 mg/L, 0.700, 0.902, 0.715 during training phase and 57.09 mg/L, 0.757, 0.924,

Station	Model	Model structure	Training data				Testing data			
			RMSE (mg/L)	N _{ES}	WI	LM	RMSE (mg/L)	N _{ES}	WI	LM
Jondhara	FFNN1	1-8-1	53.49	0.900	0.973	0.878	112.95	0.552	0.859	0.649
	FFNN2	2-2-1	54.36	0.896	0.972	0.859	113.68	0.546	0.847	0.635
	FFNN3	3-5-1	54.08	0.897	0.972	0.877	108.33	0.588	0.861	0.663
	FFNN4	4-13-1	59.93	0.874	0.965	0.813	108.15	0.589	0.863	0.598
	FFNN5	2-11-1	56.58	0.888	0.969	0.846	114.28	0.541	0.872	0.576
	FFNN6	3-8-1	58.68	0.879	0.966	0.722	110.80	0.569	0.860	0.441
	FFNN7	3-15-1	56.45	0.888	0.969	0.812	109.28	0.580	0.867	0.573
	FFNN8	4-6-1	58.77	0.879	0.966	0.822	107.86	0.591	0.867	0.588
	FFNN9	5-12-1	53.73	0.899	0.973	0.867	109.77	0.577	0.886	0.624
Simga	FFNN1	1-10-1	49.54	0.700	0.902	0.715	57.09	0.757	0.924	0.775
	FFNN2	2-10-1	49.11	0.705	0.907	0.728	54.23	0.781	0.932	0.794
	FFNN3	3-12-1	48.58	0.712	0.910	0.741	54.04	0.783	0.936	0.780
	FFNN4	4-19-1	49.05	0.706	0.910	0.728	55.23	0.773	0.934	0.777
	FFNN5	2-10-1	48.23	0.716	0.907	0.644	56.24	0.765	0.928	0.734
	FFNN6	3-14-1	47.47	0.725	0.913	0.709	55.00	0.775	0.933	0.760
	FFNN7	3-5-1	46.61	0.734	0.916	0.719	56.32	0.764	0.930	0.764
	FFNN8	4-4-1	48.57	0.712	0.907	0.722	55.09	0.774	0.930	0.791
	FFNN9	5-6-1	47.07	0.729	0.915	0.743	55.40	0.772	0.932	0.776

Table 3. The result of FFNN model for different input combinations of Jondhara and Simga. Bold values show the best model structure with lowest error and highest model efficiency.

and 0.775 during testing phase of the developed models. The training performance of the best model was better compared to testing phase.

Quantitative evaluation of CCNN model is given in Table 4. Selected input combinations as given in Table 1 were used for CCNN model development; model with single hidden layer and different input and hidden neurons. To get a better performance among individual CCNN models, a comparison has been made to obtain the best CCNN model at both stations. All eighteen developed CCNN models developed precise daily suspended sediment concentration estimations during training with a range of RMSE (mg/L), N_{ES}, WI and LM between 46.350 to 58.410 (mean = 52.383) (mg/L), 0.692 to 0.895 (mean = 0.798), 0.900 to 0.971 (mean = 0.9380) and 0.692 to 0.861 (mean 0.781) respectively, and for testing with a range of 53.710 to 108.990 (mean = 80.361) (mg/L), 0.582

Station	Model	Model structure	Training data				Testing data			
			RMSE (mg/L)	N _{ES}	WI	LM	RMSE (mg/L)	N _{ES}	WI	LM
Jondhara	CCNN1	1-5-1	56.40	0.888	0.969	0.861	107.18	0.596	0.868	0.669
	CCNN2	2-3-1	56.48	0.888	0.969	0.855	108.14	0.589	0.866	0.658
	CCNN3	3-1-1	57.32	0.885	0.968	0.850	107.57	0.593	0.869	0.647
	CCNN4	4-2-1	57.02	0.886	0.969	0.845	98.02	0.662	0.890	0.668
	CCNN5	2-6-1	54.84	0.895	0.971	0.859	99.62	0.651	0.889	0.657
	CCNN6	3-5-1	56.88	0.887	0.968	0.849	99.47	0.652	0.893	0.669
	CCNN7	3-2-1	58.18	0.881	0.966	0.850	108.67	0.585	0.872	0.647
	CCNN8	4-2-1	58.41	0.880	0.967	0.844	105.67	0.608	0.876	0.642
	CCNN9	5-0-1	57.97	0.882	0.967	0.836	108.99	0.582	0.866	0.618
Simga	CCNN1	1-5-1	50.21	0.692	0.900	0.695	57.07	0.758	0.926	0.781
	CCNN2	2-4-1	48.84	0.709	0.907	0.723	54.96	0.775	0.933	0.783
	CCNN3	3-5-1	47.91	0.719	0.911	0.732	54.68	0.777	0.933	0.776
	CCNN4	4-3-1	48.82	0.709	0.908	0.722	53.71	0.785	0.936	0.788
	CCNN5	2-4-1	48.44	0.713	0.908	0.692	58.20	0.748	0.925	0.747
	CCNN6	3-2-1	49.81	0.697	0.902	0.709	56.15	0.765	0.929	0.773
	CCNN7	3-4-1	46.35	0.737	0.916	0.724	58.45	0.746	0.926	0.747
	CCNN8	4-2-1	49.52	0.700	0.904	0.697	56.03	0.766	0.929	0.779
	CCNN9	5-3-1	47.68	0.722	0.911	0.722	53.91	0.784	0.937	0.778

Table 4. The result of CCNN model for different input combinations of Jondhara and Simga. Bold values show the best model structure with lowest error and highest model efficiency.

to 0.785 (mean = 0.690), 0.866 to 0.937 (mean = 0.904) and 0.618 to 0.788 (mean = 0.713). Based on the testing data values of N_{ES} for the Jondhara station were found satisfactory ($0.50 < N_{ES} \leq 0.65$) while for the Simga station N_{ES} were found good as $0.75 < N_{ES}$; and WI for both stations were higher than 0.75, showing high precision in predicting daily suspended sediment concentration.

CCNN4 model was selected as the best model for forecasting suspended sediment concentration at Jondhara station, and the values of RMSE, N_{ES} , WI, and LM as 57.02 mg/L, 0.886, 0.969 and 0.845, respectively during training phase and 95.02 mg/L, 0.662, 0.890, and 0.668, respectively during testing phase. The CCNN9 model performed worst for Jondhara station. For Simga station, CCNN4 model with four input variables and three neurons in hidden layer was selected as the best model among individual CCNN models. The values of RMSE, N_{ES} , WI, and LM were found 48.82 mg/L, 0.709, 0.908, and 0.722, respectively during training phase and 53.71 mg/L, 0.785, 0.936, and 0.788, respectively during testing phase. For Simga station, CCNN7 model performed worst compared to all other developed models.

Comparison between the best models (i.e., FFNN8 and CCNN4) at Jondhara station explained that CCNN4 model performed better compared to FFNN8 model during testing phase clearly. Therefore, CCNN4 model was selected as the best model for forecasting suspended sediment concentration at Jondhara station. In addition, comparison between the best models (i.e., FFNN3 and CCNN4) at Simga station revealed that CCNN4 model was a little accurate compared to FFNN3 model during testing phase. CCNN4 model, therefore, was determined as the best model for forecasting suspended sediment concentration at Simga station.

Qualitative assessment of developed models based on visual interpretation

Qualitative evaluation of selected models was carried out based on the line diagram and scatter plots as depicted in Fig. 4. The forecasted suspended sediment concentration using FFNN8 and CCNN4 models during testing phase was plotted compared to observed suspended sediment concentration at Jondhara station. The forecasted values at both stations showed the superior agreement compared to observed ones during low observed data. However, when the values of suspended sediment concentration increase, the clear fluctuation in the forecasted

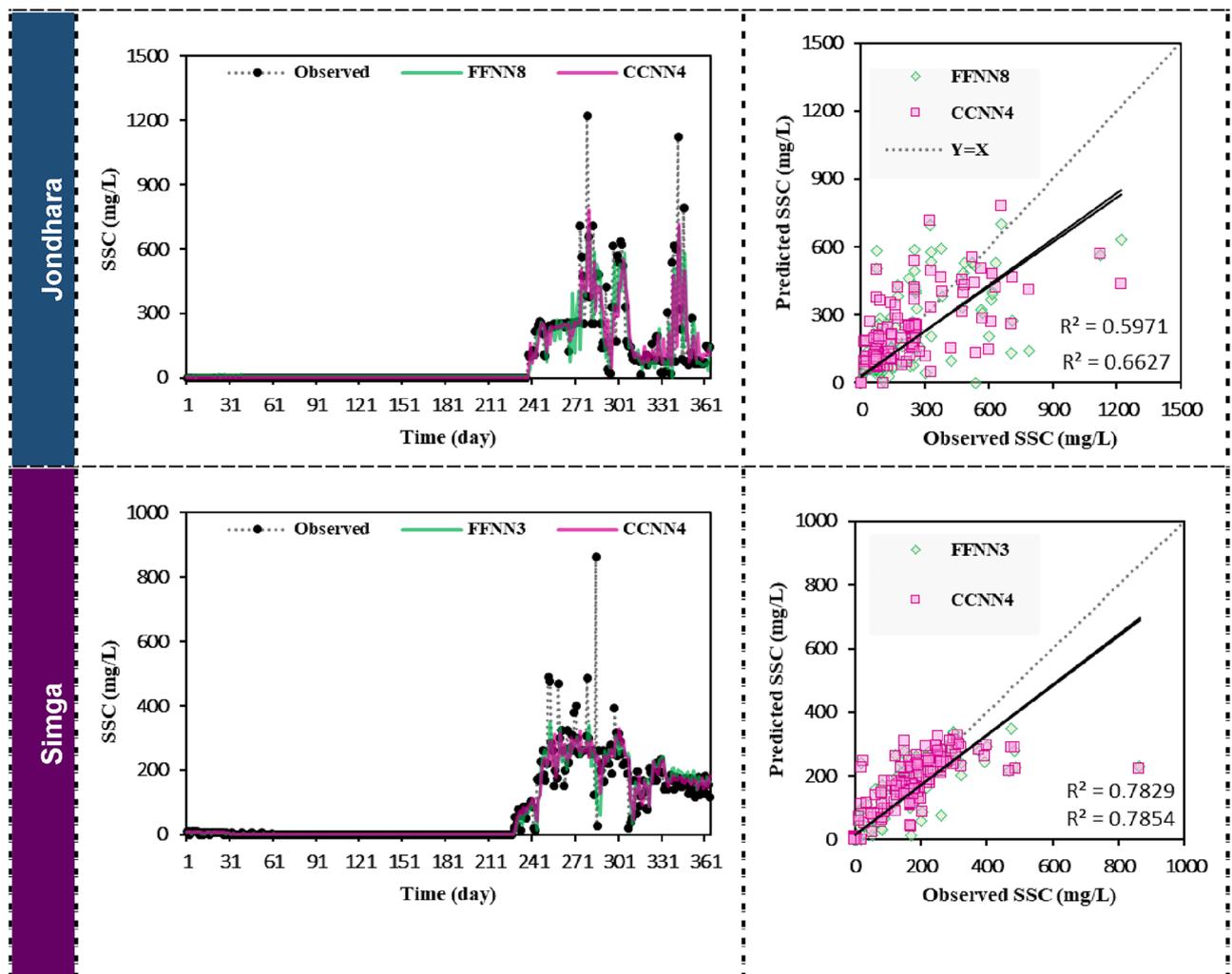


Figure 4. Forecasted suspended sediment concentration (SSC) by the best FFNN and CCNN models (left side) and scatter diagrams (right side) at Jondhara and Simga stations in testing phase.

values was provided. In addition to fluctuation, FFNN8 and CCNN4 model gave the under-forecasted values compared to peak observed ones. The relationship between forecasted values between FFNN8 and CCNN4 models was judged using the scatter plot. From the plotting, it could be seen that there was a good agreement between FFNN8 and CCNN4 models.

Line diagram at Simga station provided accurate agreement until 215 days. After that, the values of forecasted suspended sediment concentration for FFNN3 and CCNN4 models were fluctuated with under-forecasted values compared to the peak ones (Fig. 4). While the very good agreement was found between FFNN3 and CCNN4 models at Simga station, the agreement is better compared to Jondhara station.

Distribution of discharge and suspended sediment concentration on time was carried out using density plots and histograms as represented in Fig. 5. Observed and forecasted values using FFNN and CCNN models were plotted in density plots and histograms. From the density plot and histogram based on Jondhara station, the observed values less than 150 mg/L were found as maximum data, while FFNN8 and CCNN4 models provided the maximum data between 100 and 300 mg/L. In case of Simga station, the similarity in the data distribution between observed and forecasted suspended sediment concentration using FFNN3 and CCNN4 models was found as shown in Fig. 5. The maximum data point of all three sets were found among values 0–100 mg/L, 100–300 mg/L, and 300–500 mg/L. Since the similarity was outstanding at Simga station, the best models (i.e., FFNN3 and CCNN4) forecasted suspended sediment concentration accurately.

Taylor diagram is a single-window for comparing the performance of different models based on three different statistical indices, including Pearson's correlation coefficient (R), RMSE, and standard deviation (SD)^{97–101}. Taylor diagram, therefore, was applied to assess performance of FFNN and CCNN models based on observed suspended sediment concentration graphically using CC, RMSE, and SD during training and testing phases at both stations. For all the models at Jondhara station, the value of CC was found nearly 0.95, the value of SD was provided below 200, and the value of RMSE was put below 100 mg/L during training phase as given in Fig. 6, while the value of CC was found above 0.7, the value of SD was provided near to 150 except for FFNN5 and FFNN9, and the value of RMSE was put below 150 mg/l.

In case of Simga station, the value of CC was showed between 0.80 and 0.90 for all models, the value of SD was produced between 50 and 100, and the value of RMSE was yielded nearly 50 mg/l for all the models during the training period as depicted in Fig. 6. Although the value of CC was showed slightly less than 0.9 for all the models, the value of SD was produced nearly 100, and the values of RMSE was yielded slightly greater than

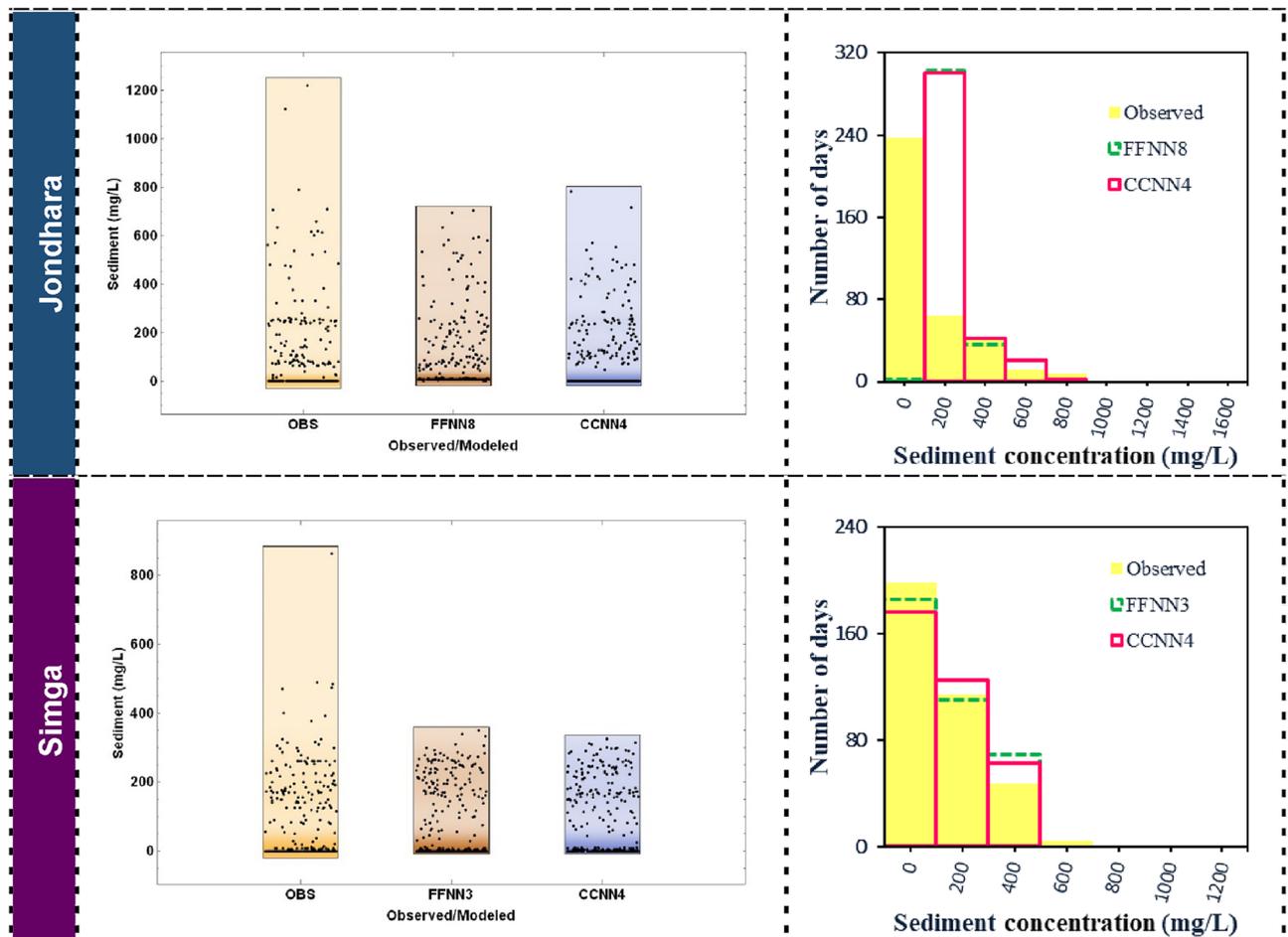


Figure 5. Density plots (left side) and histograms (right side) of observed and the best FFNN and CCNN models for testing phase at Jondhara and Simga stations.

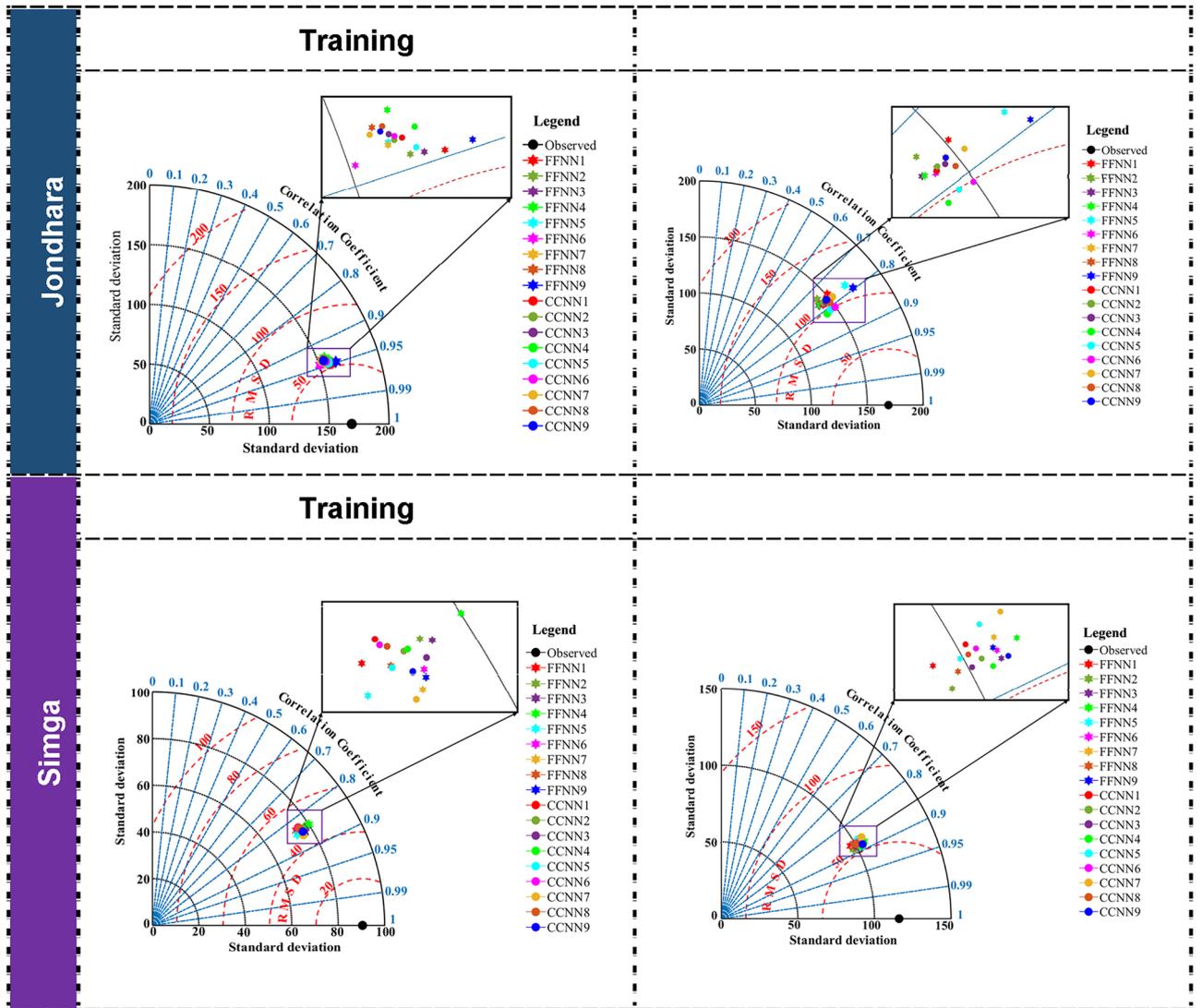


Figure 6. Taylor diagram of the models in the training and testing phase for best model of each model type at Jondhara and Simga stations.

50 mg/L for all the models. After analysing Taylor diagram of different models, it is not clear that which model is the best model. After comparing the results of different developed models based on RMSE, N_{ES} , WI, and LM values, it can be concluded that the potential of CCNN model for forecasting suspended sediment concentration was better compared to FFNN model at both stations. Based on the visual comparisons (i.e., line diagram, scatter diagram, density plot, histograms, and Taylor diagram) during testing phase, CCNN model was more accurate based on density plot and histograms. In addition, the scatter diagram indicated that CCNN model showed less deviation from $Y = X$ line compared to CCNN model for forecasting suspended sediment concentration during testing phase at both stations. In finding of the best models, all models indicated the different results for forecasting suspended sediment concentration at both stations. It was also found that the model parameters can be considered as the main factors to find the best input arrangements. It can be concluded that that CCNN and FFNN models could be forecasted suspended sediment concentration within satisfactory and accurate category.

Discussion

The performance of the FFNN and CCNN based models was influenced by the choice of the input considered for both the stations. The best performance of the FFNN model was achieved when one day and two days lagged values of suspended sediment concentration and discharge were considered for Jondhara station. The model structure corresponds for the best suspended sediment model at Jondhara station was 4-6-1 (Input-hidden layer and output layer)). While, for the Simga station, the best model was FFNN3 having inputs of one day, two days and three days legged values of suspended sediment load. The addition of discharge data as input variables had reduced the performance of the FFNN model for Simga station. This model structure of the FFNN3 was 3-12-1. The comparative performance of the models FFNN at two stations for predicting suspended sediment load showed the inputs plays a key role in predicting capability of the model. The effects of the ANN architecture

greatly influenced the models performance and this funding is align with Shukla et al.¹⁰² and Rajput et al.¹⁰³. Essem et al.⁵⁷ compared SVM, ANN, and LSTM to forecast suspended sediment load in Malaysia. Among these models, the ANN3 model, formulated using the ANN algorithm and input scenario 3 (comprising current-day sediment flow, previous-day sediment flow, and previous-day suspended sediment load), emerged as the most effective model for prediction. Our findings using the FFNN model are in line with the results reported by Essem et al.⁵⁷. The efficacy of the models in predicting suspended sediment load is better with multiple inputs as compared to single input variable.

The best performance of the CCNN model was achieved when one day, two days, three days and four days lagged values of suspended sediment load were considered for Jondhara station. The model structure corresponds for the best suspended sediment model at Jondhara station was 4-2-1 (Input-hidden layer and output layer)). While, for the Simga station, the best model was CCNN4 having inputs of one day, two days and three days and four days lagged values of suspended sediment load. The addition of discharge data as input variables had reduced the performance of the CCNN model for both the stations. This model structure of the CCNN4 was 4-3-1. The comparative performance of the models CCNN at two stations for predicting suspended sediment load showed the inputs plays a key role in predicting capability of the model. Our study findings using the CCNN model disagreed with the Jimeno-Sáez¹⁰⁴ which reported that the performance of the machine learning models found best considering all the inputs. However, we found that the performance of the CCNN reduced while adding the discharge data with the suspended sediment load. Elbisy et al.¹⁰⁵ compared feed-forward back propagation neural network (FFNN) and cascade correlation neural network (CCNN), and found CCNN model produces slightly more accurate results, and it also learns almost as fast as the BP model when compared to the FFNN model. The present study finding align with this finding.

Recent study also showed the accuracy and effectiveness of using a different input with different architecture for removing error in the time series data and, thus, enhancement of model forecasting accuracy in assessment to a standalone model^{38,47,106-115}. Overall, our study demonstrates the model structure (input-hidden-output) layers, the suspended sediment load carrying the stream/river and the algorithm used influences the predicting capability of the models. The FFNN model showed better performance considering the discharge flow data along with the suspended sediment load data, however, CCNN model showed optimum performance with suspended sediment load data alone. As seen from the result, it is clear that for both the stations, different models with different architectures, giving different results. But if what comparison is to be made between the two models, then CCNN4 models for both the stations are giving more accurate results. The CCNN4 model is capable of giving better results by understanding the hydrological complexity well. The quantification of the suspended sediment load is essential for planning of desilting of the reservoirs, water availability assessment and ascertaining the capacity of the reservoirs. Our study results could play significant role in accurate prediction of the suspended sediment load and the developed methodology may be evaluated at other places for its accuracy.

Shortcomings and future study

Accurate forecasting of suspended sediment concentration (SSC) is crucial tasks for water resource management, flood prediction, erosion control and ecosystem conservation. The proposed input combination to compare Cascade Correlation Neural Network (CCNN) and Feedforward Neural Network (FFNN) model offers enhanced precision and reliability in predicting SSC, facilitating informed decision-making for policymakers and stakeholders. Additionally, integrating hybrid algorithms can notably boost the predictive accuracy of suspended sediment concentration models. Therefore, upcoming research should explore the integration of hybrid models to improve prediction accuracy further. Ultimately, this research underscores the promise of CCNN and FFNN approaches in suspended sediment concentration prediction with small data sets, emphasizing the importance of continued exploration in this field.

Employing model ensemble techniques, such as combining predictions from multiple machine learning models or integrating machine learning with physical-based models, can potentially improve predictive accuracy and reliability by leveraging the strengths of different modeling approaches.

Conclusions

In the current study, CCNN and FFNN models were used to forecast daily suspended sediment concentration at Jondhara and Sigma stations, India. The suspended sediment concentration forecasting was carried out for both stations with nine input combinations which contained the previous one- and two-day discharge and one-, two-, three-, and four-day suspended sediment concentration. The total data was divided into training data and testing data. The performance of developed models was examined using statistical indices based on RMSE, N_{ES} , WI, and LM values. The model has the lowest value of RMSE and is close to zero and the highest value and is close to one of N_{ES} , WI, and LM values, were the best-chosen the best input combination model. Based on quantitative and visual observation, FFNN8 model at Jondhara station and FFNN3 model at Simga station were found the best models among different model architectures explored in FFNN technique. The values of RMSE, N_{ES} , WI, and LM during the training and testing phases indicated that FFNN8 with input $(S_{t-1}, S_{t-2}, Q_{t-1}, Q_{t-2})$ and FFNN3 $(S_{t-1}, S_{t-2}, S_{t-3})$ models have the best performance out of nine FFNN models at both stations. The architectures 4-2-1 and 4-3-1 of CCNN model with input $(S_{t-1}, S_{t-2}, S_{t-3}, S_{t-4})$ combination were considered as the best models for forecasting suspended sediment concentration at both stations. Owing to the deficiency of overfitting during the training period, the model was selected based on performance during the testing period to select the model with stable results. Based on the comparison of FFNN and CCNN models performance, CCNN model was found to have a good proximity with observed values at Jondhara station, while CCNN4 provided slightly better performance than FFNN3 model for Simga station. After comparing the results of different developed models based on RMSE, N_{ES} , WI, and LM values, it can be concluded that the potential of CCNN model for

forecasting suspended sediment concentration was better compared to FFNN model algorithm at both stations in Sheonath basin, India. It can be confirmed from the current study that CCNN and FFNN models can be applied to perform better forecasting of hydrological variables with non-linear and complex relationship. Every station has a specific networked model which could model the data more precisely.

The sources of uncertainty in predicting Suspended Sediment Load (SSL) are multifaceted and can stem from various factors, influencing the reliability and accuracy of predictive models. Some of the key sources of uncertainty include: (a) Data Variability (b) Model Complexity (c) Parameter Estimation (d) Input Data Quality and (e) Model Selection. In summary, while both CCNN and FFNN have their strengths and weaknesses, the choice between them depends on the specific requirements of the sediment concentration prediction task, including data complexity, computational resources, and desired prediction accuracy. A comparative analysis can help researchers and practitioners understand the trade-offs between these models and select the most suitable one for their applications.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Code availability

The codes used during the current study are available from the corresponding author upon reasonable request.

Received: 12 March 2024; Accepted: 4 May 2024

Published online: 09 May 2024

References

- Parsons, A. J., Cooper, J. & Wainwright, J. What is suspended sediment?. *Earth Surf. Process. Landforms* **40**, 1417–1420 (2015).
- Loperfido, J. V. Surface water quality in streams and rivers: Scaling and climate change. In *Comprehensive Water Quality and Purification* (ed. Ahuja, S.) 87–105 (Elsevier, 2014). <https://doi.org/10.1016/B978-0-12-382182-9.00064-5>.
- Pandion, K. *et al.* Seasonal influence on physicochemical properties of the sediments from Bay of Bengal coast with statistical approach. *Environ. Res.* **235**, 116611 (2023).
- Samantaray, S. & Sahoo, A. Prediction of suspended sediment concentration using hybrid SVM-WOA approaches. *Geocarto Int.* **37**, 5609–5635 (2022).
- Darabi, H. *et al.* Prediction of daily suspended sediment load (SSL) using new optimization algorithms and soft computing models. *Soft Comput.* **25**, 7609–7626 (2021).
- Haddadchi, A., Movahedi, N., Vahidi, N., Omid, M. H. & Dehghani, A. A. Evaluation of suspended load transport rate using transport formulas and artificial neural network models (Case study: Chelchay Catchment). *J. Hydrodyn.* **25**, 459–470 (2013).
- Kumar, A. & Saha, A. Effect of polyacrylamide and gypsum on surface runoff, sediment yield and nutrient losses from steep slopes. *Agric. Water Manag.* **98**, 999–1004 (2011).
- Zuazo, V. H. D. & Pleguezuelo, C. R. R. Soil-erosion and runoff prevention by plant covers: A review. In *Sustainable Agriculture* (eds Lichtfouse, E. *et al.*) 785–811 (Springer Netherlands, 2009). https://doi.org/10.1007/978-90-481-2666-8_48.
- Girmay, G., Singh, B. R., Nyssen, J. & Borrosen, T. Runoff and sediment-associated nutrient losses under different land uses in Tigray, Northern Ethiopia. *J. Hydrol.* **376**, 70–80 (2009).
- Vafakhah, M. Application of artificial neural networks and adaptive neuro-fuzzy inference system models to short-term stream-flow forecasting. *Can. J. Civ. Eng.* **39**, 402–414 (2012).
- Tsai, Z.-X., You, G.-Y., Lee, H.-Y. & Chiu, Y.-J. Modeling the sediment yield from landslides in the Shihmen Reservoir watershed, Taiwan. *Earth Surf. Process. Landforms* **38**, 661–674 (2013).
- Demirci, M. & Baltaci, A. Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches. *Neural Comput. Appl.* **23**, 145–151 (2013).
- Tayfur, G. Modern optimization methods in water resources planning, engineering and management. *Water Resour. Manag.* **31**, 3205–3233 (2017).
- Ayele, G. T. *et al.* Sediment yield and reservoir sedimentation in highly dynamic watersheds: The Case of Koga Reservoir, Ethiopia. *Water* **13**, 3374 (2021).
- Sadiqi, S. S. J., Hong, E. & Nam, W. Identification of priority management practices for soil erosion control through estimation of runoff and sediment yield using soil and water assessment tool on Salma watershed in Afghanistan. *Irrig. Drain.* **71**, 804–822 (2022).
- Mohanta, N. R., Biswal, P., Kumari, S. S., Samantaray, S. & Sahoo, A. Estimation of sediment load using adaptive neuro-fuzzy inference system at Indus River Basin, India. In *Intelligent Data Engineering and Analytics. Advances in Intelligent Systems and Computing* (eds Satapathy, S. C. *et al.*) 427–434 (Springer Singapore, 2021). https://doi.org/10.1007/978-981-15-5679-1_40.
- Rai, R. K. & Mathur, B. S. Event-based sediment yield modeling using artificial neural network. *Water Resour. Manag.* **22**, 423–441 (2008).
- Kumar, A. & Tripathi, V. K. Capability assessment of conventional and data-driven models for prediction of suspended sediment load. *Environ. Sci. Pollut. Res.* **29**, 50040–50058 (2022).
- El Bilali, A., Taleb, A., El Idrissi, B., Brouziyne, Y. & Mazigh, N. Comparison of a data-based model and a soil erosion model coupled with multiple linear regression for the prediction of reservoir sedimentation in a semi-arid environment. *Euro-Mediterranean J. Environ. Integr.* **5**, 64 (2020).
- Malik, A., Kumar, A. & Piri, J. Daily suspended sediment concentration simulation using hydrological data of Pranhita River Basin, India. *Comput. Electron. Agric.* **138**, 20–28 (2017).
- Liu, Q.-J., Shi, Z.-H., Fang, N.-F., Zhu, H.-D. & Ai, L. Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet-ANN approach. *Geomorphology* **186**, 181–190 (2013).
- 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.* **77**, 708 (2018).
- Zounemat-Kermani, M., Kişi, Ö., Adamowski, J. & Ramezani-Charmahineh, A. Evaluation of data driven models for river suspended sediment concentration modeling. *J. Hydrol.* **535**, 457–472 (2016).
- 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* **84**, 111–125 (2007).
- Rajae, T., Mirbagheri, S. A., Zounemat-Kermani, M. & Nourani, V. Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Sci. Total Environ.* **407**, 4916–4927 (2009).

26. Sadeghi, S. H. R. *et al.* Determinant factors of sediment graphs and rating loops in a reforested watershed. *J. Hydrol.* **356**, 271–282 (2008).
27. Ganju, N. K., Knowles, N. & Schoellhamer, D. H. Temporal downscaling of decadal sediment load estimates to a daily interval for use in hindcast simulations. *J. Hydrol.* **349**, 512–523 (2008).
28. Arabkhedri, M., Lai, F. S., Ibrahim, N.-A. & Mohamad-Kasim, M.-R. Effect of adaptive cluster sampling design on accuracy of sediment rating curve estimation. *J. Hydrol. Eng.* **15**, 142–151 (2010).
29. Singh, V. K., Singh, B. P. & Vivekanand, B. Basin suspended sediment prediction using soft computing and conventional approaches in India. *Environ. Sci.* **7**, 459–468 (2016).
30. Chakravorti, T., Patnaik, R. K. & Dash, P. K. Detection and classification of islanding and power quality disturbances in microgrid using hybrid signal processing and data mining techniques. *IET Signal Process.* **12**, 82–94 (2018).
31. Chakravorti, T. & Dash, P. K. Multiclass power quality events classification using variational mode decomposition with fast reduced kernel extreme learning machine-based feature selection. *IET Sci. Meas. Technol.* **12**, 106–117 (2018).
32. Chakravorti, T., Priyadarshini, L., Dash, P. K. & Sahu, B. N. Islanding and non-islanding disturbance detection in microgrid using optimized modes decomposition based robust random vector functional link network. *Eng. Appl. Artif. Intell.* **85**, 122–136 (2019).
33. Salem, S. *et al.* Applying multivariate analysis and machine learning approaches to evaluating groundwater quality on the Kairouan Plain, Tunisia. *Water* **15**, 3495 (2023).
34. Lemaoui, T. *et al.* Machine learning approach to map the thermal conductivity of over 2,000 neoteric solvents for green energy storage applications. *Energy Storage Mater.* **59**, 102795 (2023).
35. Pinthong, S. *et al.* Imputation of missing monthly rainfall data using machine learning and spatial interpolation approaches in Thale Sap Songkhla River Basin, Thailand. *Environ. Sci. Pollut. Res.* <https://doi.org/10.1007/s11356-022-23022-8> (2022).
36. Sahoo, A., Barik, A., Samantaray, S. & Ghose, D. K. Prediction of sedimentation in a watershed using RNN and SVM. In *Communication Software and Networks. Lecture Notes in Networks and Systems* (eds Satapathy, S. C. *et al.*) 701–708 (Springer, 2021). https://doi.org/10.1007/978-981-15-5397-4_71.
37. Samantaray, S. & Ghose, D. K. Evaluation of suspended sediment concentration using descent neural networks. *Procedia Comput. Sci.* **132**, 1824–1831 (2018).
38. 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.* **98**, 855–866 (2011).
39. Yadav, A., Chatterjee, S. & Equeenuddin, S. M. Prediction of suspended sediment yield by artificial neural network and traditional mathematical model in Mahanadi river basin, India. *Sustain. Water Resour. Manag.* **4**, 745–759 (2018).
40. Gupta, D., Hazarika, B. B., Berlin, M., Sharma, U. M. & Mishra, K. Artificial intelligence for suspended sediment load prediction: A review. *Environ. Earth Sci.* **80**, 346 (2021).
41. Yadav, A., Chatterjee, S. & Equeenuddin, S. M. Suspended sediment yield estimation using genetic algorithm-based artificial intelligence models: Case study of Mahanadi River, India. *Hydrol. Sci. J.* **63**, 1162–1182 (2018).
42. Khosravi, K., Mao, L., Kisi, O., Yaseen, Z. M. & Shahid, S. Quantifying hourly suspended sediment load using data mining models: Case study of a glacierized Andean catchment in Chile. *J. Hydrol.* **567**, 165–179 (2018).
43. Kisi, O. & Yaseen, Z. M. The potential of hybrid evolutionary fuzzy intelligence model for suspended sediment concentration prediction. *CATENA* **174**, 11–23 (2019).
44. Ghose, D. K. & Samantaray, S. Modelling sediment concentration using back propagation neural network and regression coupled with genetic algorithm. *Procedia Comput. Sci.* **125**, 85–92 (2018).
45. Rahul, A. K., Shivhare, N., Kumar, S., Dwivedi, S. B. & Dikshit, P. K. S. Modelling of daily suspended sediment concentration using FFBPNN and SVM algorithms. *J. Soft Comput. Civ. Eng.* **5**, 120–134 (2021).
46. Nhu, V.-H. *et al.* Monthly suspended sediment load prediction using artificial intelligence: Testing of a new random subspace method. *Hydrol. Sci. J.* **65**, 2116–2127 (2020).
47. Rajaei, T. Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. *Sci. Total Environ.* **409**, 2917–2928 (2011).
48. Sahoo, B. B., Sankalp, S. & Kisi, O. A novel smoothing-based deep learning time-series approach for daily suspended sediment load prediction. *Water Resour. Manag.* **37**, 4271–4292 (2023).
49. Keshtegar, B. *et al.* Prediction of sediment yields using a data-driven radial M5 tree model. *Water* **15**, 1437 (2023).
50. Moradinejad, A. Suspended load modeling of river using soft computing techniques. *Water Resour. Manag.* **38**, 1965–1986 (2024).
51. Tulla, P. S. *et al.* Daily suspended sediment yield estimation using soft-computing algorithms for hilly watersheds in a data-scarce situation: A case study of Bino watershed, Uttarakhand. *Theor. Appl. Climatol.* <https://doi.org/10.1007/s00704-024-04862-5> (2024).
52. Idrees, M. B., Jehanzaib, M., Kim, D. & Kim, T.-W. Comprehensive evaluation of machine learning models for suspended sediment load inflow prediction in a reservoir. *Stoch. Environ. Res. Risk Assess.* **35**, 1805–1823 (2021).
53. Sahoo, A., Mohanta, N. R., Samantaray, S. & Satapathy, D. P. Application of hybrid ANFIS-CSA model in suspended sediment load prediction. In *Advanced Computing and Intelligent Technologies. Lecture Notes in Electrical Engineering* (eds Shaw, R. N. *et al.*) 295–305 (Springer Singapore, 2022). https://doi.org/10.1007/978-981-19-2980-9_24.
54. Mohanta, N. R., Panda, S. K., Singh, U. K., Sahoo, A. & Samantaray, S. MLP-WOA is a successful algorithm for estimating sediment load in Kalahandi Gauge Station, India. In *Proceedings of International Conference on Data Science and Applications. Lecture Notes in Networks and Systems* (eds Saraswat, M. *et al.*) 319–329 (Springer, 2022) https://doi.org/10.1007/978-981-16-5120-5_25.
55. Yadav, A., Alam, M. A. & Suryavanshi, S. Daily sediment yield prediction using hybrid machine learning approach. *Int. J. Environ. Clim. Chang.* **13**, 143–157 (2023).
56. Kumar, M., Kumar, P., Kumar, A., Elbeltagi, A. & Kuriqi, A. Modeling stage–discharge–sediment using support vector machine and artificial neural network coupled with wavelet transform. *Appl. Water Sci.* **12**, 87 (2022).
57. Essam, Y., Huang, Y. F., Birima, A. H., Ahmed, A. N. & El-Shafie, A. Predicting suspended sediment load in Peninsular Malaysia using support vector machine and deep learning algorithms. *Sci. Rep.* **12**, 302 (2022).
58. Samantaray, S., Sahoo, A. & Prakash Satapathy, D. Improving accuracy of SVM for monthly sediment load prediction using Harris hawks optimization. *Mater. Today Proc.* **61**, 604–617 (2022).
59. Sahoo, A., Samantaray, S. & Sathpathy, D. P. Prediction of sediment load through novel SVM-FOA approach: A case study. In *Data Engineering and Intelligent Computing. Lecture Notes in Networks and Systems* (eds Bhateja, V. *et al.*) 291–301 (Springer, 2022). https://doi.org/10.1007/978-981-19-1559-8_30.
60. Chen, S., Huang, J. & Huang, J.-C. Improving daily streamflow simulations for data-scarce watersheds using the coupled SWAT-LSTM approach. *J. Hydrol.* **622**, 129734 (2023).
61. Samantaray, S., Sahoo, A., Paul, S. & Ghose, D. K. Prediction of bed-load sediment using newly developed support-vector machine techniques. *J. Irrig. Drain. Eng.* [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.000168](https://doi.org/10.1061/(ASCE)IR.1943-4774.000168) (2022).
62. Karunanithi, N., Grenney, W. J., Whitley, D. & Bovee, K. neural networks for river flow prediction. *J. Comput. Civ. Eng.* **8**, 201–220 (1994).

63. Alok, A., Patra, K. C. & Das, S. K. Prediction of discharge with Elman and cascade neural networks. *Res. J. Recent Sci.* **2**, 279–284 (2013).
64. Kim, S., Singh, V. P. & Seo, Y. Evaluation of pan evaporation modeling with two different neural networks and weather station data. *Theor. Appl. Climatol.* **117**, 1–13 (2014).
65. Ghorbani, M. A. *et al.* Development and evaluation of the cascade correlation neural network and the random forest models for river stage and river flow prediction in Australia. *Soft Comput.* **24**, 12079–12090 (2020).
66. Zounemat-Kermani, M. *et al.* Can decomposition approaches always enhance soft computing models? Predicting the dissolved oxygen concentration in the St. Johns River, Florida. *Appl. Sci.* **9**, 2534 (2019).
67. Bilhan, O., Emin Emiroglu, M. & Kisi, O. Application of two different neural network techniques to lateral outflow over rectangular side weirs located on a straight channel. *Adv. Eng. Softw.* **41**, 831–837 (2010).
68. Kişi, Ö. A combined generalized regression neural network wavelet model for monthly streamflow prediction. *KSCE J. Civ. Eng.* **15**, 1469–1479 (2011).
69. Zounemat-kermani, M., Kisi, O. & Rajaei, T. Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff. *Appl. Soft Comput.* **13**, 4633–4644 (2013).
70. Ehteram, *et al.* Investigation on the potential to integrate different artificial intelligence models with metaheuristic algorithms for improving river suspended sediment predictions. *Appl. Sci.* **9**, 4149 (2019).
71. Heddam, S., Sanikhani, H. & Kisi, O. Application of artificial intelligence to estimate phycocyanin pigment concentration using water quality data: A comparative study. *Appl. Water Sci.* **9**, 164 (2019).
72. Fahlman, S. & Lebiere, C. The cascade-correlation learning architecture. *Adv. Neural Inf. Process. Syst.* **2**, 524–532 (1989).
73. Nobandegani, A. S. & Shultz, T. R. Converting cascade-correlation neural nets into probabilistic generative models. *arXiv Prepr. <https://doi.org/10.48550/arXiv.1701.05004>* (2017).
74. Kumar, A., Kumar, P. & Singh, V. K. Evaluating different machine learning models for runoff and suspended sediment simulation. *Water Resour. Manag.* **33**, 1217–1231 (2019).
75. Zakhrouf, M., Bouchelkia, H., Stamboul, M. & Kim, S. Novel hybrid approaches based on evolutionary strategy for streamflow forecasting in the Chellif River, Algeria. *Acta Geophys.* **68**, 167–180 (2020).
76. Bebis, G. & Georgiopoulos, M. Feed-forward neural networks. *IEEE Potentials* **13**, 27–31 (1994).
77. Lima, A. R., Cannon, A. J. & Hsieh, W. W. Forecasting daily streamflow using online sequential extreme learning machines. *J. Hydrol.* **537**, 431–443 (2016).
78. Deo, R. C., Samui, P. & Kim, D. Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models. *Stoch. Environ. Res. Risk Assess.* **30**, 1769–1784 (2016).
79. Luk, K. C., Ball, J. E. & Sharma, A. An application of artificial neural networks for rainfall forecasting. *Math. Comput. Model.* **33**, 683–693 (2001).
80. Annayat, W., Gupta, A., Prakash, K. R. & Sil, B. S. Application of artificial neural networks and multiple linear regression for rainfall-runoff modeling BT—communication software and networks. In (eds Satapathy, S. C. *et al.*) 719–727 (Springer Singapore, 2021).
81. Daliakopoulos, I. N., Coulibaly, P. & Tsanis, I. K. Groundwater level forecasting using artificial neural networks. *J. Hydrol.* **309**, 229–240 (2005).
82. Faris, H., Aljarah, I., Al-Madi, N. & Mirjalili, S. Optimizing the learning process of feedforward neural networks using lightning search algorithm. *Int. J. Artif. Intell. Tools* **25**, 1650033 (2016).
83. Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366 (1989).
84. Meshram, S. G. *et al.* New approach for sediment yield forecasting with a two-phase feedforward neuron network-particle swarm optimization model integrated with the gravitational search algorithm. *Water Resour. Manag.* **33**, 2335–2356 (2019).
85. Samantaray, S., Sahoo, A. & Ghose, D. K. Prediction of sedimentation in an arid watershed using BPNN and ANFIS. In *ICT Analysis and Applications. Lecture Notes in Networks and Systems* (eds Fong, S. *et al.*) 295–302 (Springer, 2020). https://doi.org/10.1007/978-981-15-0630-7_29.
86. Samantaray, S. & Ghose, D. K. Sediment assessment for a watershed in arid region via neural networks. *Sādhanā* **44**, 219 (2019).
87. Ghose, D. K. & Samantaray, S. Sedimentation process and its assessment through integrated sensor networks and machine learning process. In *Computational Intelligence in Sensor Networks. Studies in Computational Intelligence* (eds Mishra, B. *et al.*) 473–488 (Springer, 2019). https://doi.org/10.1007/978-3-662-57277-1_20.
88. Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **10**, 282–290 (1970).
89. Samantaray, S. & Ghose, D. K. Assessment of suspended sediment load with neural networks in arid watershed. *J. Inst. Eng. Ser. A* **101**, 371–380 (2020).
90. Samantaray, S., Sahoo, A. & Ghose, D. K. Assessment of sediment load concentration using SVM, SVM-FFA and PSR-SVM-FFA in arid watershed, India: A case study. *KSCE J. Civ. Eng.* **24**, 1944–1957 (2020).
91. Willmott, C. J. & Wicks, D. E. An empirical method for the spatial interpolation of monthly precipitation within California. *Phys. Geogr.* **1**, 59–73 (1980).
92. Samantaray, S. & Sahoo, A. Assessment of sediment concentration through RBNN and SVM-FFA in arid watershed, India. In (eds Satapathy, S. *et al.*) 701–709 (Springer, 2020). https://doi.org/10.1007/978-981-13-9282-5_67.
93. Legates, D. R. & McCabe, G. J. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **35**, 233–241 (1999).
94. Vishwakarma, D. K. *et al.* Methods to estimate evapotranspiration in humid and subtropical climate conditions. *Agric. Water Manag.* **261**, 107378 (2022).
95. Tao, H. *et al.* Hybridized artificial intelligence models with nature-inspired algorithms for river flow modeling: A comprehensive review, assessment, and possible future research directions. *Eng. Appl. Artif. Intell.* **129**, 107559 (2024).
96. Sahoo, G. K., Sahoo, A., Samantara, S., Satapathy, D. P. & Satapathy, S. C. Application of adaptive neuro-fuzzy inference system and salp swarm algorithm for suspended sediment load prediction. In *Intelligent System Design. Lecture Notes in Networks and Systems* (eds Bhatija, V. *et al.*) 339–347 (Springer Singapore, 2023). https://doi.org/10.1007/978-981-19-4863-3_32.
97. Kumar, D. *et al.* Multi-ahead electrical conductivity forecasting of surface water based on machine learning algorithms. *Appl. Water Sci.* **13**, 192 (2023).
98. Vishwakarma, D. K. *et al.* Forecasting of stage-discharge in a non-perennial river using machine learning with gamma test. *Heliyon* **9**, e16290 (2023).
99. Mirzania, E., Vishwakarma, D. K., Bui, Q.-A.T., Band, S. S. & Dehghani, R. A novel hybrid AIG-SVR model for estimating daily reference evapotranspiration. *Arab. J. Geosci.* **16**, 301 (2023).
100. Saroughi, M. *et al.* A novel hybrid algorithms for groundwater level prediction. *Iran. J. Sci. Technol. Trans. Civ. Eng.* <https://doi.org/10.1007/s40996-023-01068-z> (2023).
101. Samantaray, S., Sahoo, A. & Satapathy, D. P. Prediction of groundwater-level using novel SVM-ALO, SVM-FOA, and SVM-FFA algorithms at Purba-Medinipur, India. *Arab. J. Geosci.* **15**, 723 (2022).
102. Shukla, R. *et al.* Modeling of stage-discharge using back propagation ANN-, ANFIS-, and WANN-based computing techniques. *Theor. Appl. Climatol.* <https://doi.org/10.1007/s00704-021-03863-y> (2021).

103. Rajput, J. *et al.* Performance evaluation of soft computing techniques for forecasting daily reference evapotranspiration. *J. Water Clim. Chang.* **14**, 350–368 (2023).
104. Jimeno-Sáez, P., Martínez-España, R., Casalí, J., Pérez-Sánchez, J. & Senent-Aparicio, J. A comparison of performance of SWAT and machine learning models for predicting sediment load in a forested Basin, Northern Spain. *CATENA* **212**, 105953 (2022).
105. Elbisy, M. S., Ali, H. M., Abd-Elall, M. A. & Alaboud, T. M. The use of feed-forward back propagation and cascade correlation for the neural network prediction of surface water quality parameters. *Water Resour.* **41**, 709–718 (2014).
106. Elbeltagi, A. *et al.* Modelling daily reference evapotranspiration based on stacking hybridization of ANN with meta-heuristic algorithms under diverse agro-climatic conditions. *Stoch. Environ. Res. Risk Assess.* <https://doi.org/10.1007/s00477-022-02196-0> (2022).
107. Kumar, A. R. S., Goyal, M. K., Ojha, C. S. P., Singh, R. D. & Swamee, P. K. Application of artificial neural network, fuzzy logic and decision tree algorithms for modelling of streamflow at Kasol in India. *Water Sci. Technol.* **68**, 2521–2526 (2013).
108. Kakaei Lafdani, E., Moghaddam Nia, A. & Ahmadi, A. Daily suspended sediment load prediction using artificial neural networks and support vector machines. *J. Hydrol.* **478**, 50–62 (2013).
109. Nourani, V. A review on applications of artificial intelligence-based models to estimate suspended sediment load. *Int. J. Soft Comput. Eng.* **3**, 121–127 (2014).
110. Nourani, V. & Andalib, G. Daily and monthly suspended sediment load predictions using wavelet based artificial intelligence approaches. *J. Mt. Sci.* **12**, 85–100 (2015).
111. Sharafati, A., Haji Seyed Asadollah, S. B., Motta, D. & Yaseen, Z. M. Application of newly developed ensemble machine learning models for daily suspended sediment load prediction and related uncertainty analysis. *Hydrol. Sci. J.* **65**, 2022–2042 (2020).
112. Markuna, S. *et al.* Application of innovative machine learning techniques for long-term rainfall prediction. *Pure Appl. Geophys.* **180**, 335–363 (2023).
113. Singh, A. K. *et al.* An integrated statistical-machine learning approach for runoff prediction. *Sustainability* **14**, 8209 (2022).
114. 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* **13**, 1–29 (2021).
115. Elbeltagi, A. *et al.* Modelling daily reference evapotranspiration based on stacking hybridization of ANN with meta-heuristic algorithms under diverse agro-climatic conditions. *Stoch. Environ. Res. Risk Assess.* **36**, 3311–3334 (2022).

Acknowledgements

The authors would like to thank Researchers Supporting Project number (RSPD2024R958), King Saud University, Riyadh, Saudi Arabia. The work was supported by the KICT Research Program (project no. 20230166-001, Development of Coastal Groundwater Management Solution) funded by the Ministry of Science and ICT.

Author contributions

Conceptualization, B.J., V.K.S. and D.K.V.; methodology, V.K.S. and D.K.V.; software, V.K.S. and D.K.V.; validation, V.K.S., B.J., and D.K.V.; formal analysis, V.K.S., S.A.A., and D.K.V.; investigation, V.K.S. and D.K.V.; resources, V.K.S.; data curation, S.G., D.K.V.; writing—original draft preparation, B.J., S.A.A., V.K.C., S.G., M.A.G., S.K. and D.K.V.; writing—review and editing, D.K.V., M.A.G., S.K., J.R., M.C., K.K.Y., E.M., N.A.-A. and M.A.M.; visualization, V.K.S. and D.K.V., E.M.; supervision, D.K.V. N.A.-A. and M.A.M.; project administration, B.J., V.K.S. and D.K.V.; funding acquisition, M.C., N.A.-A. and M.A.M. All authors have read and agreed to the published version of the manuscript.

Funding

Open access funding provided by Lulea University of Technology. Researchers Supporting Project number (RSPD2024R958), King Saud University, Riyadh, Saudi Arabia.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.K.V., N.A.-A. or M.A.M.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024