

The potential of ensemble WT-EEMD-kernel extreme learning machine techniques for prediction suspended sediment concentration in successive points of a river

Kiyoumars Roushangar, Nasrin Aghajani, Roghayeh Ghasempour and Farhad Alizadeh

ABSTRACT

Sediment transport is one of the most important issues in river engineering. In this study, the capability of the Kernel Extreme Learning Machine (KELM) approach for predicting the river daily Suspended Sediment Concentration (SSC) and Discharge (SSD) was assessed. Three successive hydrometric stations of Mississippi river were considered and based on the sediment and flow characteristics during the period of 2005–2008. Several models were developed and tested for SSC and SSD modeling. For improving the applied model efficiency, two post-processing techniques, namely Wavelet Transform (WT) and Ensemble Empirical Mode Decomposition (EEMD), were used. Also, two states of modeling based on stations' own data (state 1) and previous stations' data (state 2) were considered. The single and integrated KELM model results comparison indicated that the integrated WT and EEMD-KELM models resulted in more accurate outcomes. Results showed that data processing with WT was more effective than EEMD in increasing the models' efficiency. Data processing enhanced the models' capability by up to 15%. The results showed that the state 1 modeling led to better results, however, using the integrated KELM approaches the previous stations data could be applied successfully for SSC and SSD modeling when the stations' own data were not available.

Key words | EEMD, KELM, pre-processing, suspended load, suspended sediment discharge, WT

Kiyoumars Roushangar (corresponding author)
Nasrin Aghajani
Roghayeh Ghasempour
Farhad Alizadeh
Department of Water Resources Engineering,
Faculty of Civil Engineering,
University of Tabriz,
Tabriz,
Iran
E-mail: kroshangar@yahoo.com

Kiyoumars Roushangar
Center of Excellence in Hydroinformatics,
University of Tabriz,
Tabriz,
Iran

HIGHLIGHT

- The suspended sediment concentration (SSC) and suspended sediment discharge (SSD) were predicted via artificial intelligence approach in successive hydrometric stations. The data pre-processing impacts on models' efficiency improvement was assessed. The sensitivity analysis showed the most effective subseries was obtained from pre-processing models.

INTRODUCTION

Sediment transportation and accurate estimation of its rate is a significant issue for river engineers and researchers. So far,

various and complex relationships have been proposed to predict the suspended sediment transport rate, such as velocity and critical shear stress based equations. However, the complex nature of sediment transport and lack of validated models make it difficult to model the suspended sediment concentration and suspended sediment discharge carried by rivers. *Bhattacharya et al. (2004)* stated that it is

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

doi: 10.2166/hydro.2021.146

difficult to express the transport process through a deterministic mathematical framework. Based on laboratory experiments, [Vongvisessomjai *et al.* \(2010\)](#) studied the sediment transport for non-cohesive sediment in uniform flow at a no-deposition state. [Harrington & Harrington \(2013\)](#) evaluated the efficiency of the Sediment Rating Curve (SRC) method in modeling the suspended sediment load of the Bandon and Owenabue rivers in Ireland. [Rajaei *et al.* \(2009\)](#) and [Chen & Chau \(2016\)](#) indicated that the sediment rating curve and the auto-regressive integrated moving average model are inadequate to predict SSC under extreme hyperconcentrated flow conditions. Although the mentioned models led to promising results in sediment transport prediction, due to the importance of sediment transport and its impact on hydraulic structures it is necessary to use other methods with higher efficiency ([Lafdani *et al.* 2013](#); [Rahman & Chakrabarty 2020](#)).

In recent years, intelligence techniques such as Artificial Neural Networks (ANNs), Neuro-Fuzzy models (NF), Genetic Programming (GP), Multivariate Adaptive Regression Splines (MARS), Kernel Extreme Learning Machine (KELM), and Gaussian Process Regression (GPR) have been used in assessing the complex hydraulic and hydrological phenomena ([Roushangar & Ghasempour 2018](#)) such as estimation of reference evapotranspiration ([Yin *et al.* 2017](#)), daily suspended sediment concentration modeling ([Kaveh *et al.* 2017](#)), side weir discharge coefficient modeling ([Azamathulla *et al.* 2017](#)), prediction of roughness coefficient in sewer pipes ([Roushangar *et al.* 2020](#)), and modeling form resistance coefficient of movable bed channels ([Saghebian *et al.* 2020](#)). In artificial intelligence models we are looking for a learning machine capable of finding an accurate approximation of a natural phenomenon, as well as expressing it in the form of an interpretable equation. However, this bias towards interpretability creates several new issues. The computer-generated hypotheses should take advantage of the already existing body of knowledge about the domain in question. However, the method by which we express our knowledge and make it available to a learning machine remains rather unclear ([Babovic 2009](#)). Machine learning, a branch of artificial intelligence, deals with representation and generalization using data learning techniques. Representation of data instances and functions evaluated on these instances are part of all machine learning

systems. Generalization is the property that the system will perform well on unseen data instances; the conditions under which this can be guaranteed are a key object of study in the subfield of computational learning theory. There are a wide variety of machine learning tasks and successful applications ([Mitchell 1997](#)). In general, the task of a machine learning algorithm can be described as follows: Given a set of input variables and the associated output variable(s), the objective is learning a functional relationship for the input-output variables set. It should be noted that artificial intelligence models typically do not really represent the physics of a modeled process; they are just devices used to capture relationships between the relevant input and output variables. However, when the interrelationships among the relevant variables are poorly understood, finding the size and shape of the ultimate solution is difficult, and conventional mathematical analysis methods do not (or cannot) provide analytical solutions; these methods can predict the interest variable with more accuracy.

On the other hand, hybrid models based on signal decomposition can be effective in increasing the time series prediction methods' efficiency ([Pachori *et al.* 2015](#)). Wavelet analysis is one of the commonly used methods for signal decomposition. Additionally, the Empirical Mode Decomposition (EMD) method, which is suitable for nonlinear and non-stationary time series ([Huang *et al.* 1998](#)), has been used recently. Unlike wavelet decomposition, empirical mode decomposition extracts the data oscillatory mode components without a priori determining the basis functions or level of decomposition ([Labate *et al.* 2013](#)).

Therefore, in the current study, the Kernel Extreme Learning Machine (KELM) as a kernel-based approach was used for modeling Suspended Sediment Concentration (SSC) and Suspended Sediment Discharge (SSD) in three successive hydrometric stations. The KELM, as a kernel-based approach based on quadratic optimization of convex function, can easily switch from linear to nonlinear separation. This is realized by nonlinear mapping using so-called kernel functions. Kernel based approaches such as KELM are a relatively new and important method based on the different kernel types. Such models are based on statistical learning theory and are capable of adapting themselves to predict any variable of interest via sufficient inputs. The training of this method is fast, has high accuracy,

and the probability of occurrence of data overtraining in this method is less. Discrete Wavelet Transform (DWT) and EEMD were used as pre-processing methods to improve the models efficiency. In integrated pre-processing models, the inputs data were decomposed into sub-series by Wavelet Transform (WT) and EEMD. Then, these sub-series were used as inputs in the KELM method. In this regard, daily sediment and flow data of the Mississippi river in the period of 2005–2008 were used and under two scenarios various models were developed. In the first scenario, the intended parameters of each station were estimated using the stations' own data, and in the second scenario, the SSC and SSD parameters were estimated using the previous station's data. Also, sensitivity analysis was carried out to select the most effective sub-series obtained from WT and EEMD in the modeling process.

MATERIALS AND METHODS

Study area

The Mississippi river is the second longest river, and the most important river of the second-largest drainage system on the North American continent. From its traditional source of Lake Itasca in northern Minnesota, it flows generally south for 3,730 km to the Mississippi River Delta in the Gulf of Mexico. In the current study, daily data of stream-flow and suspended sediment discharge and concentration during the period of 2005–2008 were used. Three consecutive stations, namely station A (7010000), station B (7020500), and station C (7022000), were selected and suspended sediment discharge and suspended sediment concentration were investigated under two scenarios. In

the first scenario, modeling was performed based on each station's data and in the second scenario the previous station's data were used. Table 1 shows the statistical characteristics of the selected stations. In this table, parameters Q_{sc} , Q_{sd} , and Q_d are suspended sediment concentration, suspended sediment discharge, and flow discharge, respectively. Figure 1 shows the location of the selected stations.

Kernel Extreme Learning Machine (KELM)

Extreme Learning Machines (ELM) are feedforward neural networks for classification, regression, clustering, sparse approximation, compression and feature learning with a single layer or multiple layers of hidden nodes, where the parameters of hidden nodes (not just the weights connecting inputs to hidden nodes) do not need to be adjusted. These hidden nodes can be randomly assigned and never updated (i.e. random projections can be sorted into these nodes to nonlinearly reduce their dimensionality), or can be inherited from their ancestors without being changed. In most cases, the output weights of hidden nodes are usually learned in a single step, which essentially amounts to learning a linear model. The name 'extreme learning machine' (ELM) was given to such models. In ELM, the main idea involves the hidden layer weights. Furthermore, the biases are randomly generated and the calculation of the output weights is carried out using the least-squares solution. Furthermore, they have been defined by the outputs of the targets and the hidden layer. ELM is a Single Layer Feed Forward Neural Network (SLFFNN) preparing method initially introduced by Huang *et al.* (2006). SLFFNN is a straight framework where information weights linked to hidden neurons and hidden layer biases are haphazardly chosen, while the weights among the hidden nodes are resolved logically.

Table 1 | Characteristics of the Mississippi River consecutive hydrometric stations

Hydrometric station number	Q_{sc} (mg/L)		Q_d (ft ³ /s)		Q_{sd} (ton/day)	
	Max	Min	Max	Min	Max	Min
7010000	716,000	63,000	1,510	59.6	200,000	12,600
7020500	695,000	64,100	1,650	44.8	2,150,000	10,900
7022000	710,000	68,300	1,260	40.3	1,740,000	11,300

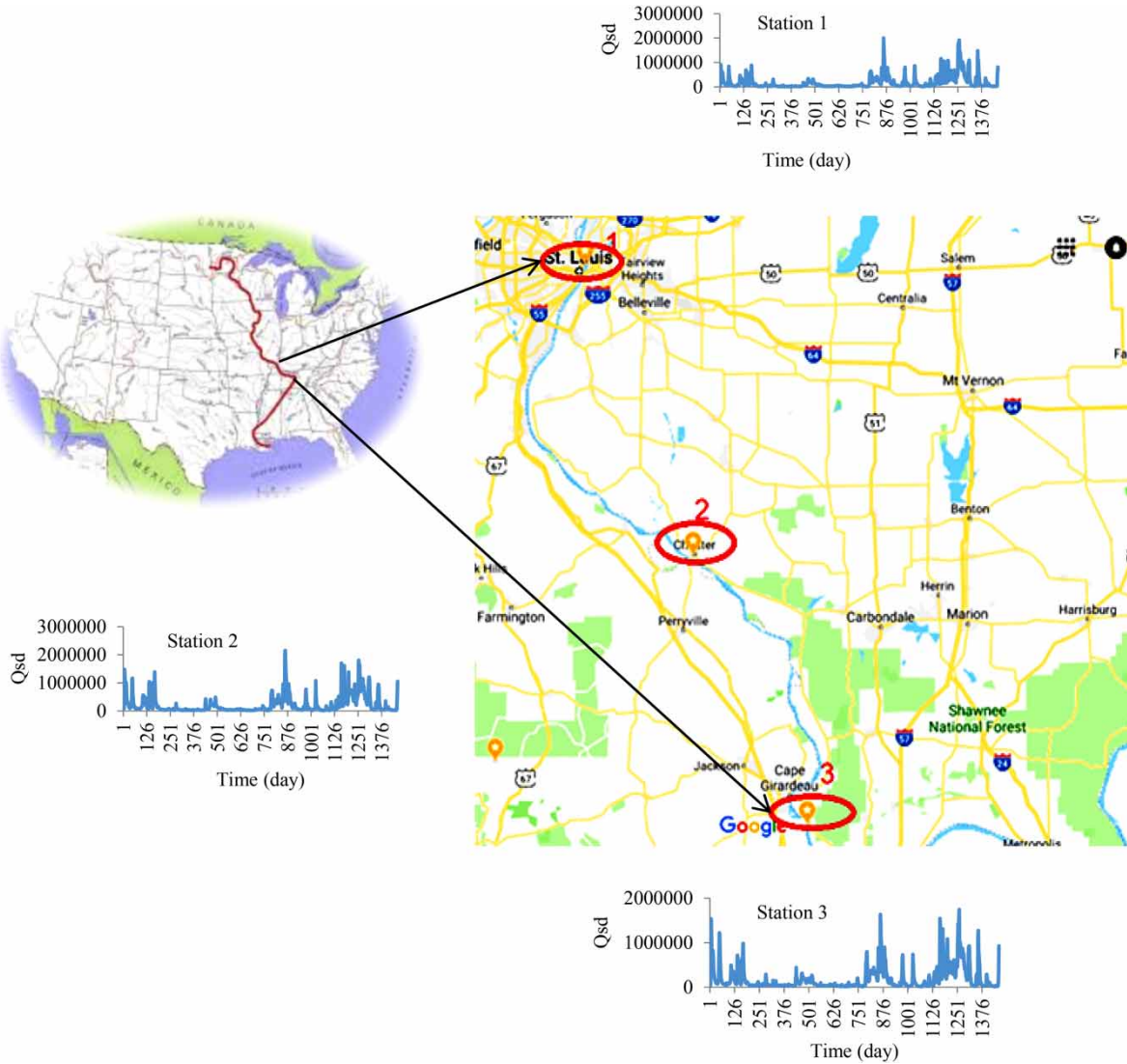


Figure 1 | The location of the selected consecutive stations of Mississippi River.

This strategy likewise has preferred execution and adapts progressively over the bygone era’s learning methods (Huang et al. 2006). In ELM, there is no need for tuning the initial parameters of the hidden layer and almost all nonlinear piecewise continuous functions can be used as the hidden neurons. The standard single-layer neural system with N random information (a_i, b_i) , M hidden neurons, and the active function $f(a)$ are shown as follows:

$$a_i = [a_{i1}, a_{i2}, \dots, a_{iN}]^T \in R^n \tag{1}$$

$$b_i = [b_{i1}, b_{i2}, \dots, b_{in}]^T \in R^m \tag{2}$$

$$\sum_{i=1}^N \alpha_i f(w_i \alpha_j + c_i) = X_j \quad j = 1, 2, \dots, N \tag{3}$$

where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the weight vector that joins the input layer to the hidden layer, $\alpha_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{im}]^T$ is the weight vector that joins the hidden layer to the target layer. c_i shows the hidden neuron biases. The general SLFFNN network with the M hidden neurons and the activation function $f(a)$ can predict N information with an average zero error $(\sum_{i=1}^N ||X_j - X_i|| = 0)$. The SLFFNN network’s aim is to minimize the difference between the

predicted (X_j) and target (Y_j) values which can be expressed as below:

$$\sum_{i=1}^N \alpha_i f(w_i \alpha_j + c_i) = Y_j \quad j = 1, 2, \dots, N \quad (4)$$

Equation (4) can be summarized as:

$$A\alpha = K \quad (5)$$

$$A(w_1, \dots, w_M, c_1, \dots, c_M, a_1, \dots, a_M) = \begin{bmatrix} f(w_1 a_1 + c_1) \dots f(w_N a_1 + c_M) \\ \dots \\ f(w_1 a_N + c_1) \dots f(w_M a_N + c_M) \end{bmatrix}_{N \times M} \quad (6)$$

$$K = \begin{bmatrix} p_1^T \\ \dots \\ p_N^T \end{bmatrix}_{N \times L}, \quad \alpha = \begin{bmatrix} \alpha_1^T \\ \dots \\ \alpha_N^T \end{bmatrix}_{M \times L} \quad (7)$$

The matrix K is identified as the target matrix of the hidden layers of the neural network. Huang et al. (2012) also introduced kernel functions in the design of ELM. Now, a number of kernel functions are used in the design of ELM such as linear, radial basis, normalized polynomial, polynomial kernel functions. Kernel function based ELM design is known as Kernel Extreme Learning Machine (KELM). For more details about KELM, readers and researchers are referred to Huang et al. (2012).

Pre-processing approaches

One of the most popular approaches in time series processing is Wavelet Transform (WT) (Farajzadeh & Alizadeh 2017). The WT uses a flexible window function (mother wavelet) in signal processing. The flexible window function can be changed over time according to the signal shape and compactness (Mehr et al. 2013). After using WT, the signal will decompose into two approximations (large-scale or low-frequency component) and detailed (small-scale component) components. An illustration of a three-level WT is shown in Figure 2. In the first level, the original signal (x) is decomposed to two components of approximation (cA_1) and detailed (cD_1). In the second level, cA_1 is again decomposed to approximation (cA_2) and detailed (cD_2) components. Finally, in the third level, cA_2 is decomposed to cA_3 approximation and cD_3 detailed components. The sum of all detailed sub-series and approximation series obtained from the third level will be the original signal (i.e. $x = cD_1 + cD_2 + cD_3 + cA_3$). The other approach for time series processing is

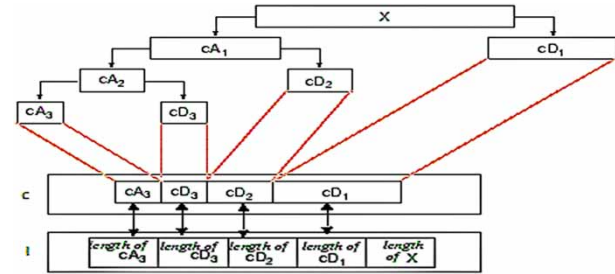


Figure 2 | The steps of a time series decomposition into detail (D) and approximation (A) sub-series.

Empirical Mode Decomposition (EMD). The EMD method is an effective self-adaptive dyadic filter bank which is applied to the white noise (a random signal which has equal intensity at different frequencies). By applying this method, each signal can be decomposed into a number of Inherent Mode Functions (IMFs) which can be used to process nonlinear and non-stationary signals. One of the advantages of this method is the ability to determine the instantaneous frequency of the signal. At each step of the signal decomposition into its frequency components, the high frequency components are separated first and this process must continue until the component with the lowest frequency remains (see Lei et al. (2009) for more details). EEMD is developed based on EMD. The main benefit of EEMD is solving the mode mixing problem of EMD which determines the true IMF as the mean of an ensemble of trials (Wu & Huang 2009). For selecting the most effective IMFs and using them as inputs in the modeling process, their energy values can be calculated and the IMFs with higher energy can be used as inputs.

Performance criteria

In the current study, the proposed model's efficiency was assessed via Correlation Coefficient (R), Determination Coefficient (DC), and Root Mean Square Errors (RMSE) criteria as follows:

$$DC = 1 - \frac{\sum_{i=1}^N (l_o - l_p)^2}{\sum_{i=1}^N (l_o - \bar{l}_p)^2}, \quad R = \frac{\sum_{i=1}^N (l_o - \bar{l}_o) \times (l_p - \bar{l}_p)}{\sqrt{\sum_{i=1}^N (l_o - \bar{l}_o)^2 \times (l_p - \bar{l}_p)^2}}, \quad RMSE = \sqrt{\frac{\sum_{i=1}^N (l_o - l_p)^2}{N}} \quad (8)$$

where l_o , l_p , \bar{l}_o , \bar{l}_p , N are the observed values, predicted values, mean observed values, mean predicted values, and number of data samples, respectively. The DC describes the relative assessment of the model performance in dimensionless measures; R indicates the linear dependence between observation and predicted values and should not be applied alone as a performance criterion (Legates & McCabe 1999). The RMSE describes the average difference between predicted and measured values corresponding to the predicted values. These measurements are not oversensitive to extreme values (outliers), but are rather sensitive to additive and proportional differences between model predictions and observations. Therefore, correlation-based measures (e.g. R and DC) can indicate that a model is a good predictor (Legates & McCabe 1999). Evidently, a high value for R and DC (up to one) and a small value for RMSE indicate the high efficiency of the model. Usually the R and DC of 0.70 are considered desirable. The model with R and DC values above 0.7 will predict the value of the intended parameter successfully. It should be noted that in this study all input variables were scaled between 0 and 1 in order to eliminate the input and output variables dimensions.

Simulation and model development

Selection of appropriate variables as inputs is the most important step in modeling via intelligence methods. In

this research, previous values of daily suspended sediment discharge and concentration over the period of 2005–2008 were used as inputs to model the SSC and SSD values. In the modeling process, two states were considered: in the first state, the SSC and SSD parameters of selected stations were predicted based on the data of each station, and in the second state, modeling was done based on the data from previous stations. Also, the impact of data pre-processing on improving the models' accuracy was assessed using the WT and EEMD methods. According to Aussem et al. (1998), the minimum decomposition level in the WT method can be obtained as follows:

$$L = \lceil \log N \rceil \quad (9)$$

where L is the decomposition level and N is the number of time series (Farajzadeh & Alizadeh 2017). In this study, $L = 5$ was used as the decomposition level value. Table 2 indicates the developed models in this study. Also, in Figure 3, the considered modeling process is shown. The first 75% of data was employed to train the models and rest of the data was employed to test the models.

Table 2 | KELM developed models

Output variable	Model	Input variable	Output variable	Model	Input variable
SSC modeling					
$Q_{SC}(t)$	SC(I)	$Q_{SC}(t-1)$	$Q_{SC2}(t)$ or $Q_{SC3}(t)$	H(I)	$Q_{SC1}(t-1)$ or $Q_{SC2}(t-1)$
	SC(II)	$Q_{SC}(t-1), Q_d(t)$		H(II)	$Q_{SC1}(t-1), Q_{SD1}(t-1)$ or $Q_{SC2}(t-1), Q_{SD2}(t-1)$
	SC(III)	$Q_{SC}(t-1), Q_{SC}(t-2)$		H(III)	$Q_{SC1}(t-1), Q_{SC1}(t-2)$ or $Q_{SC2}(t-1), Q_{SC2}(t-2)$
	SC(IV)	$Q_{SC}(t-1), Q_d(t-1)$		H(IV)	$Q_{SC1}(t), Q_{SC1}(t-1)$ or $Q_{SC2}(t), Q_{SC2}(t-1)$
SSD modelling					
$Q_{SD}(t)$	SD(I)	$Q_{SD}(t-1)$	$Q_{SD2}(t)$ or $Q_{SD3}(t)$	D(I)	$Q_{SD1}(t-1)$ or $Q_{SD2}(t-1)$
	SD(II)	$Q_{SD}(t-1), Q_{SD}(t-2)$		D(II)	$Q_{SD1}(t-1), Q_{SD1}(t-2)$ or $Q_{SD2}(t-1), Q_{SD2}(t-2)$
	SD(III)	$Q_{SC}(t-1), Q_{SC}(t-2), Q_d(t-1), Q_d(t-2)$		D(III)	$Q_{SD1}(t), Q_{SD1}(t-1)$ or $Q_{SD2}(t), Q_{SD2}(t-1)$
				D(IV)	$Q_{SD1, 2}(t-1), Q_{SD1, 2}(t-2), Q_{SD1, 2}(t-3)$

Note: Q_{SC} : suspended sediment concentration, Q_{SD} : suspended sediment discharge, Q_d : water discharge. In parameters Q_{SCi} or Q_{SDi} , i shows the station number.

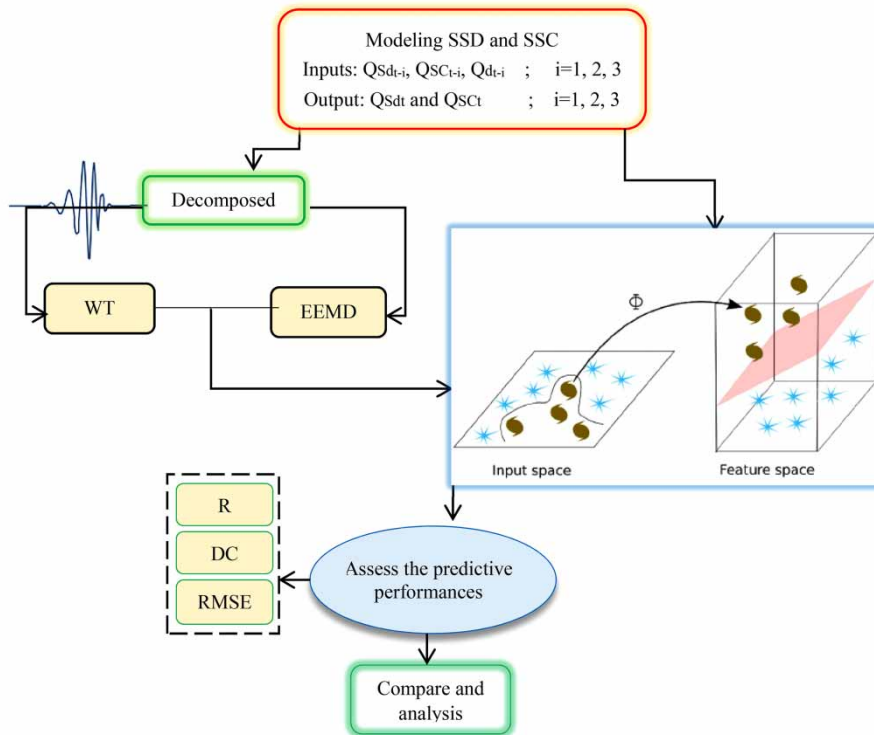


Figure 3 | Considered modeling process in the study.

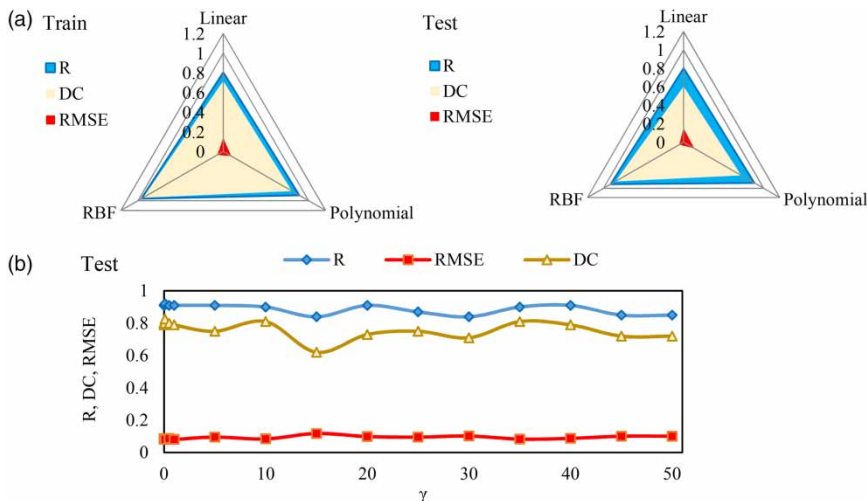


Figure 4 | The statistical parameters of KELM method (a) with different kernel functions, and (b) with different γ values.

RESULTS AND DISCUSSION

KELM models development

It should be noted that each artificial intelligence method has its own parameters for achieving the desired results,

and the optimized amount of these parameters should be determined. For example, in designing the KELM approach, the selection of appropriate type of kernel function is needed. There are various kernel functions which can be used based on the nature of the studied phenomenon. In this research, for selecting the best kernel function, the

model SC(III) was used for SSC prediction in station 1 via various kernels. Figure 4(a) indicates the results of statistical parameters of different kernels for this model. According to Figure 4(a), the RBF kernel function $[e^{-\gamma|x-y|^2}]$ in which γ is kernel parameter] was fined as the best kernel function. Figure 4(b) shows the RMSE statistic parameter via γ values for comparing the impact of RBF kernel parameter of γ on the performance of the employed algorithm for the testing set of model SC(III) for station 1. In this study, optimization of γ was performed by a systematic grid search of the parameter using cross-validation.

The results of SSC modeling

Modeling based on raw data

For evaluating the suspended sediment concentration in three selected stations, several models were developed based on the suspended sediment concentration, suspended sediment discharge, and flow discharge data. The models

were analyzed with KELM model to carry out the SSC prediction. Table 3 and Figure 5 show the results of KELM models. From the obtained results of statistical parameters (RMSE, R and DC) it could be stated that in the first

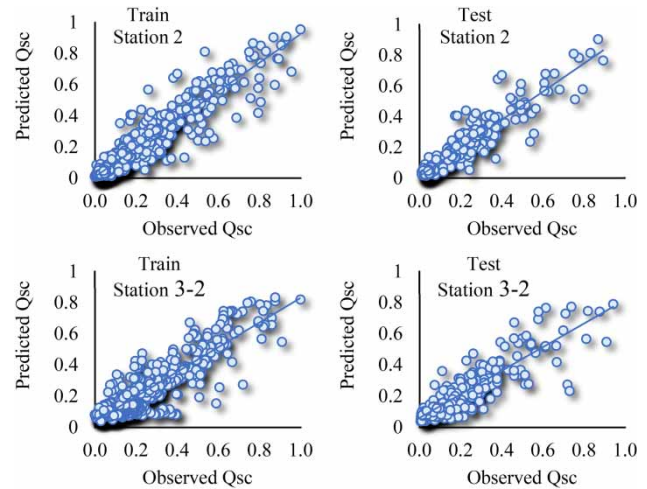


Figure 5 | Comparison of observed and predicted SSC for superior KELM model.

Table 3 | Statistical parameters of the KELM models for SSC modeling: State 1 without data processing

		Performance criteria												
		Train			Test			Train			Test			
Station/model		R	DC	RMSE	R	DC	RMSE	R	DC	RMSE	R	DC	RMSE	
State 1														
1	SC(I)	0.92	0.85	0.047	0.91	0.81	0.068	2	0.93	0.87	0.053	0.92	0.83	0.065
	SC(II)	0.83	0.86	0.046	0.83	0.62	0.099		0.94	0.88	0.051	0.88	0.69	0.089
	SC(III)	0.93	0.83	0.049	0.90	0.82	0.069		0.95	0.91	0.044	0.93	0.86	0.059
	SC(IV)	0.93	0.85	0.047	0.89	0.81	0.071		0.93	0.87	0.052	0.88	0.75	0.081
3	SC(I)	0.96	0.94	0.069	0.92	0.85	0.081							
	SC(II)	0.93	0.87	0.052	0.86	0.64	0.104							
	SC(III)	0.90	0.89	0.042	0.89	0.87	0.056							
	SC(IV)	0.93	0.86	0.054	0.87	0.83	0.071							
State 2														
2-1	H(I)	0.9	0.81	0.064	0.84	0.63	0.099	3-2	0.91	0.83	0.061	0.91	0.82	0.074
	H(II)	0.91	0.83	0.059	0.87	0.61	0.102		0.92	0.84	0.057	0.85	0.71	0.092
	H(III)	0.83	0.69	0.082	0.75	0.55	0.109		0.88	0.77	0.071	0.88	0.69	0.096
	H(IV)	0.92	0.87	0.058	0.88	0.71	0.087		0.92	0.86	0.06	0.91	0.82	0.072
3-1	H(II)	0.89	0.82	0.062	0.89	0.62	0.107							
	H(III)	0.91	0.82	0.062	0.90	0.71	0.095							
	H(IV)	0.89	0.82	0.061	0.89	0.72	0.094							

Note: in state 2, 2-1 means that the SSC values of station 2 are predicted based on station 1' data. 3-1 means that the SSC values of station 3 are predicted based on station 1' data. 3-2 means that the SSC values of station 3 are predicted based on station 2' data.

state the model SC(III) with input parameters of $Q_s(t-1)$, $Q_s(t-2)$ performed better than the others. Based on the results, it could be seen that in estimation of the SSC, using the previous suspended sediment concentration led to more accurate results and the use of sediment discharge had no significant impact on modeling. In the second state, the model H(IV) with input parameters of $Q_s(t)$, $Q_s(t-1)$ was selected as the superior model. A comparison between the results of the two states showed that modeling

based on each station's own data led to more desirable results. However, using the previous station's data in the modeling process yielded relatively accurate results, therefore, via the KELM kernel-based approach, the previous station's data could be used when the station's own data were unavailable. In fact, in the second state, this issue was investigated whereby the existing sub-basins between the consecutive stations may have noticeable impacts on the flow regime of the downstream station. The distance

Table 4 | Statistical parameters of the WT-KELM or EEMD-KELM models for SSC modeling; State 2 with data processing

Station/model/method		Performance criteria														
		Train			Test			Train			Test					
		R	DC	RMSE	R	DC	RMSE	R	DC	RMSE	R	DC	RMSE			
State 1																
1	SC(I)	EEMD	0.91	0.88	0.047	0.87	0.84	0.067	2	EEMD	0.91	0.9	0.045	0.88	0.85	0.062
		WT	0.95	0.88	0.041	0.92	0.85	0.061		WT	0.91	0.9	0.045	0.88	0.85	0.062
	SC(II)	EEMD	0.92	0.91	0.035	0.83	0.81	0.069		EEMD	0.92	0.92	0.041	0.8	0.82	0.087
		WT	0.98	0.95	0.025	0.93	0.88	0.055		WT	0.98	0.97	0.022	0.96	0.91	0.044
	SC(III)	EEMD	0.95	0.94	0.028	0.96	0.90	0.039		EEMD	0.97	0.95	0.029	0.95	0.92	0.045
		WT	0.96	0.97	0.023	0.97	0.93	0.037		WT	0.98	0.96	0.028	0.98	0.94	0.041
	SC(IV)	EEMD	0.92	0.91	0.036	0.82	0.82	0.065		EEMD	0.91	0.88	0.049	0.87	0.79	0.074
		WT	0.96	0.92	0.034	0.92	0.85	0.062		WT	0.96	0.91	0.042	0.93	0.85	0.062
3	SC(I)	EEMD	0.96	0.93	0.049	0.88	0.82	0.064								
		WT	0.95	0.95	0.046	0.94	0.89	0.056								
	SC(II)	EEMD	0.94	0.82	0.049	0.91	0.89	0.055								
		WT	0.98	0.92	0.027	0.97	0.91	0.033								
	SC(III)	EEMD	0.97	0.95	0.029	0.94	0.91	0.038								
		WT	0.97	0.96	0.027	0.96	0.93	0.033								
	SC(IV)	EEMD	0.92	0.91	0.036	0.82	0.82	0.065								
		WT	0.95	0.9	0.045	0.93	0.84	0.064								
State 2																
2-1	H(I)	EEMD	0.91	0.85	0.059	0.85	0.72	0.088	3-2	EEMD	0.92	0.87	0.058	0.92	0.85	0.066
		WT	0.95	0.89	0.054	0.88	0.76	0.085		WT	0.96	0.91	0.053	0.96	0.88	0.064
	H(II)	EEMD	0.92	0.87	0.054	0.88	0.70	0.091		EEMD	0.93	0.88	0.057	0.86	0.73	0.082
		WT	0.96	0.91	0.050	0.91	0.74	0.088		WT	0.97	0.92	0.052	0.89	0.76	0.079
	H(III)	EEMD	0.84	0.72	0.075	0.76	0.63	0.097		EEMD	0.89	0.80	0.064	0.89	0.71	0.086
		WT	0.87	0.76	0.069	0.79	0.67	0.094		WT	0.92	0.84	0.059	0.92	0.74	0.083
	H(IV)	EEMD	0.93	0.91	0.053	0.89	0.82	0.078		EEMD	0.93	0.90	0.055	0.92	0.85	0.064
		WT	0.97	0.95	0.049	0.92	0.86	0.075		WT	0.97	0.94	0.051	0.96	0.88	0.062
3-1	H(II)	EEMD	0.90	0.86	0.057	0.90	0.71	0.095								
		WT	0.93	0.90	0.052	0.93	0.75	0.092								
	H(III)	EEMD	0.92	0.86	0.057	0.91	0.82	0.085								
		WT	0.96	0.90	0.052	0.95	0.86	0.082								
	H(IV)	EEMD	0.90	0.86	0.056	0.90	0.83	0.084								
		WT	0.93	0.90	0.051	0.93	0.87	0.081								

between the stations used in this research was 50 km and since much sediment was not carried between stations, therefore a relationship was found between the flow regime of the upstream and downstream. However, if there are special conditions between the stations (such as diversion dams, intake structures, etc.) this connection may be less.

Modeling based on pre-processing data

In this section, the effect of time series pre-processing on increasing the model's accuracy was investigated. Therefore, the time series were decomposed using WT and EEMD methods. To decompose the time series by WT, a mother wavelet which is more similar to the signal should be selected. In this study, the daubechies (db2 and db4) and symlet (sym2 and sym4) mother wavelets were trained and it was found that the db4 mother wavelet led to better outcomes. Therefore, the db4 mother wavelet was used for time series decomposition. Also, in the second step, data was decomposed via EEMD. The principle of EEMD is decomposition of signal to different IMFs and one residual signal. The sum of these signals will be the same original signal. The formation of IMFs is based on subtracting the basic function from the original signal. This process continues until the residual signal remains almost constant. In this study, time series were decomposed into 10 IMFs and one residual signal. Then, the obtained sub-series were used as inputs in the KELM model to predict the SSC. The results of the integrated pre-processing models are listed in Table 4 and shown in Figure 6. According to the results presented in Tables 2 and 3, it could be induced that data pre-processing significantly improved the results accuracy and integrated models were more accurate than a single-KELM model. In fact, the use of WT and EEMD led to an improvement in the outcomes. Sadeghpoor (2014) tried to design and evaluate the efficiency of the wavelet-SVM model for daily SSC forecasting and showed that the integrated models (wavelet SVM) provide acceptable predictions of the SSC. It was found that the wavelet transform is a powerful tool which has a great ability to extract useful information from time series. Consequently, it increases the SVM models' performances significantly. In this study, in the integrated pre-processing method, the model's

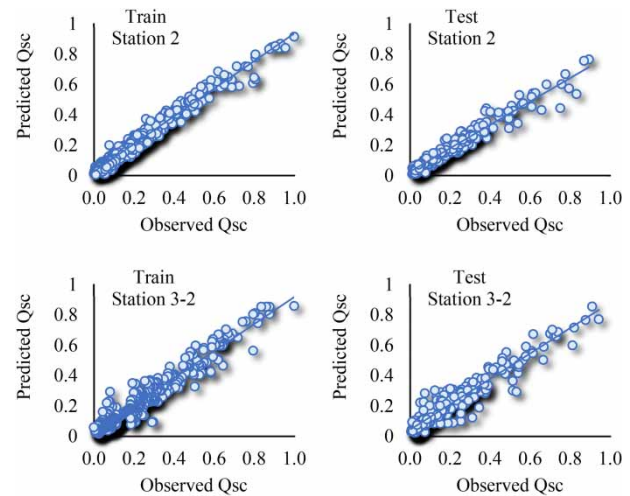


Figure 6 | Comparison of observed and predicted SSC for superior WT-KELM model.

accuracy increased between 10 and 12% in training sets, and 8 and 20% in testing sets. According to the results, it could be seen that between two pre-processing methods, the EEMD had higher RMSE error criteria in comparison with the WT. Therefore, it could be stated that in enhancing the predictions accuracy, the WT method performed more successfully than the EEMD method. Also, it was found that the model SC(III) with input parameters of $Q_s(t-1)$, $Q_s(t-2)$ in modeling based on each station's own data and the model H(IV) with input parameters of $Q_s(t)$, $Q_s(t-1)$ in modeling based on the previous station's data performed more successfully.

Results of SSD modeling

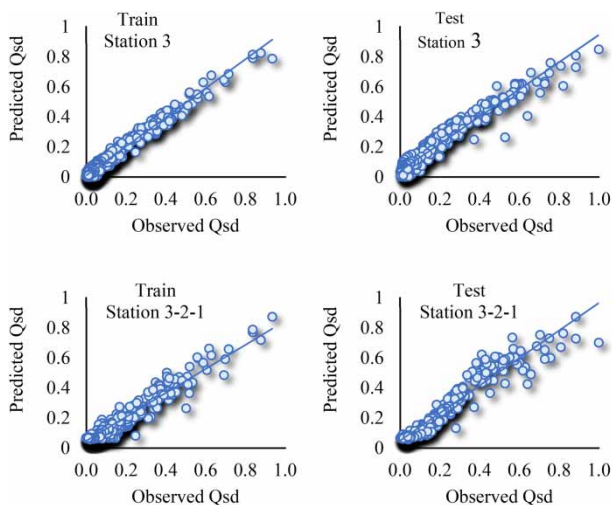
Modeling based on raw data

Accurate prediction of the suspended sediment discharge in rivers or streams is crucial for sustainable water resources and environmental systems. In this study, the suspended sediment discharge in selected stations was assessed via a KELM kernel-based approach. The previous parameters of flow and sediment discharge were used for development of the models. Table 5 and Figure 7 show the results of KELM models. The obtained results indicated that in the state of modeling based on each station's own data, the model SD(III) with input parameters of $Q_{sc}(t-1)$, $Q_{sc}(t-2)$, $Q_d(t-1)$, $Q_d(t-2)$ performed more successfully than other

Table 5 | Statistical parameters of the KELM models for SSD modeling: State 1 without data processing

Station/model		Performance criteria												
		Train			Test			Train			Test			
		R	DC	RMSE	R	DC	RMSE	R	DC	RMSE	R	DC	RMSE	
State 1														
1	SD(I)	0.89	0.83	0.048	0.88	0.80	0.057	2	0.86	0.84	0.045	0.85	0.79	0.065
	SD(II)	0.90	0.86	0.038	0.89	0.83	0.056		0.88	0.85	0.031	0.87	0.82	0.058
	SD(III)	0.92	0.88	0.032	0.91	0.84	0.054		0.88	0.87	0.028	0.87	0.84	0.053
3	SD(I)	0.89	0.85	0.042	0.86	0.82	0.062							
	SD(II)	0.92	0.88	0.035	0.88	0.85	0.052							
	SD(III)	0.94	0.90	0.028	0.90	0.86	0.048							
State 2														
2-1	D(I)	0.87	0.78	0.038	0.84	0.73	0.094	3-2	0.87	0.81	0.034	0.86	0.77	0.066
	D(II)	0.87	0.81	0.038	0.83	0.74	0.092		0.88	0.83	0.032	0.86	0.80	0.062
	D(III)	0.89	0.84	0.032	0.85	0.75	0.090		0.89	0.85	0.033	0.87	0.84	0.055
3-1	D(II)	0.88	0.83	0.035	0.88	0.83	0.061							
	D(III)	0.87	0.81	0.038	0.87	0.80	0.066							
3-2-1	D(IV)	0.90	0.88	0.031	0.89	0.87	0.054							

Note: in state 2, 3-2-1 means that the SSD values of station 3 are predicted based on data from stations 1 and 2.

**Figure 7** | Comparison of observed and predicted SSD for superior KELM model.

models. Kisi *et al.* (2012) showed that in suspended sediment modeling via genetic programming, the model whose inputs were the current water discharge and one previous water discharge and sediment load performed better. Also, the models whose inputs were current and one immediate

previously recorded water discharge and one and two previous sediment loads, as well as models whose input were current water discharge and one previous sediment load, were ranked as second and third best models, respectively. In this study, from the obtained results, it could be stated that the model D(II) with two input parameters of $Q_{sd}(t-1)$, $Q_{sd}(t-2)$ yielded the desired accuracy. Therefore, the sediment discharge can be predicted using only the previous 1 and 2 days' sediment discharge variables.

In the second state, it could be seen that between three stations the third station SSD modeling based on data from both first and second stations performed more successfully. However, modeling based on the station's own data led to more desirable predictions. Artificial intelligence methods are very powerful tools and when the interrelationships among the relevant variables are difficult to understand and conventional mathematical analysis methods cannot provide analytical solutions, these methods can be used successfully. Choubin *et al.* (2018) evaluated the use of a Classification and Regression Tree (CART) model to estimate SSD based on hydro-meteorological data. They indicated that the CART as an artificial intelligence model

can be a helpful tool in basins where hydro-meteorological data are readily available. The scatter plots of the KELM-best model for each state are shown in Figure 7. The term 3–2–1 in this figure means that the SSD values of station 3 are predicted based on the data from stations 1 and 2.

Modeling based on pre-processing data

The impact of pre-processing of data on predicting the SSD variable was assessed. The input combinations were decomposed using WT and EEMD methods. It was found

that the db4 mother wavelet is more similar to the SSD signals and it led to better outcomes. Also, a decomposition level of 5 was used. In the EEMD method, time series were decomposed into 10 IMFs and one residual signal. The results of the integrated pre-processing models are listed in Table 6 and shown in Figure 8. According to the results, it could be stated that data pre-processing significantly improved the SSD prediction accuracy. It was observed that the applied pre-processing methods improved the models' efficiency between approximately 8 and 12% in training sets and between 10 and 18% in testing sets. From

Table 6 | Statistical parameters of the WT-KELM or EEMD-KELM models for SSD modeling: State 2 with data processing

Station/model/method			Performance criteria														
			Train			Test			Train			Test					
			R	DC	RMSE	R	DC	RMSE	R	DC	RMSE	R	DC	RMSE			
State 1																	
1	SD(I)	EEMD	0.94	0.88	0.037	0.92	0.84	0.044	2	EEMD	0.91	0.87	0.035	0.89	0.83	0.042	
		WT	0.95	0.93	0.036	0.93	0.89	0.042		WT	0.92	0.89	0.034	0.90	0.87	0.040	
	SD(II)	EEMD	0.95	0.92	0.031	0.93	0.88	0.042		EEMD	0.93	0.91	0.025	0.91	0.87	0.037	
		WT	0.96	0.94	0.030	0.94	0.91	0.040		WT	0.93	0.92	0.024	0.92	0.90	0.031	
	SD(III)	EEMD	0.97	0.94	0.025	0.95	0.92	0.039		EEMD	0.95	0.95	0.022	0.91	0.92	0.027	
		WT	0.98	0.96	0.024	0.95	0.94	0.037		WT	0.96	0.96	0.021	0.95	0.94	0.024	
3	SD(I)	EEMD	0.94	0.90	0.033	0.90	0.86	0.048									
		WT	0.95	0.93	0.031	0.91	0.91	0.046									
	SD(II)	EEMD	0.96	0.94	0.029	0.92	0.91	0.040									
		WT	0.98	0.95	0.028	0.95	0.93	0.038									
	SD(III)	EEMD	0.98	0.95	0.022	0.94	0.95	0.034									
		WT	0.99	0.97	0.021	0.95	0.96	0.033									
State 2																	
2-1	D(I)	EEMD	0.92	0.87	0.030	0.88	0.77	0.071	3-2	EEMD	0.92	0.87	0.027	0.90	0.81	0.050	
		WT	0.93	0.89	0.028	0.89	0.81	0.067		WT	0.93	0.89	0.025	0.91	0.85	0.047	
	D(II)	EEMD	0.92	0.91	0.031	0.87	0.79	0.069		EEMD	0.93	0.91	0.025	0.90	0.86	0.048	
		WT	0.93	0.92	0.030	0.88	0.81	0.066		WT	0.94	0.92	0.024	0.91	0.87	0.046	
	D(III)	EEMD	0.97	0.91	0.025	0.89	0.82	0.065		EEMD	0.97	0.93	0.026	0.91	0.89	0.039	
		WT	0.98	0.93	0.024	0.95	0.84	0.062		WT	0.98	0.94	0.025	0.95	0.92	0.037	
3-1	D(II)	EEMD	0.93	0.87	0.027	0.92	0.87	0.046									
		WT	0.94	0.89	0.026	0.93	0.88	0.044									
	D(III)	EEMD	0.92	0.91	0.031	0.91	0.85	0.051									
		WT	0.93	0.92	0.029	0.92	0.87	0.046									
3-2-1	D(IV)	EEMD	0.93	0.92	0.025	0.92	0.91	0.038									
		WT	0.97	0.96	0.022	0.93	0.93	0.035									

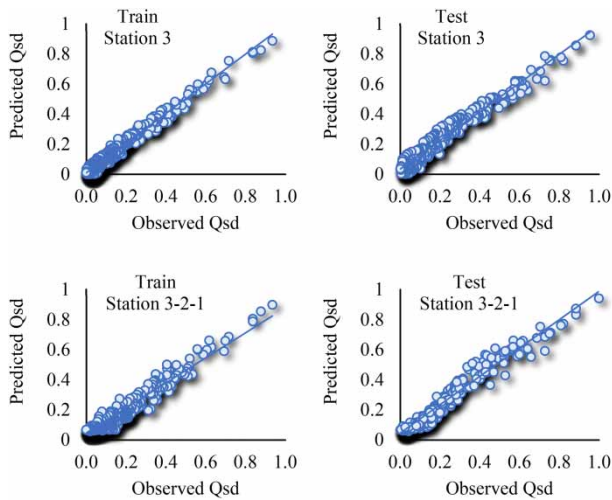


Figure 8 | Comparison of observed and predicted SSD for superior WT-KELM model.

the results, it could be indicated that in the case of modeling based on each station data, the model SD(III) with input parameters of $Q_{sc}(t-1)$, $Q_{sc}(t-2)$, $Q_d(t-1)$, $Q_d(t-2)$ led to more accurate results. In the state of investigating the relationship between stations, modeling the station 2 sediment discharge based on the first station data in the term of $Q_{sd1}(t)$, $Q_{sd1}(t-1)$ performed more successfully.

While in station 3, using both stations 1 and 2 data led to better predictions. This issue showed the impact of the previous station's information on the modeling process. It was also observed that in the case of pre-processing data, the maximum and minimum amounts of time series were predicted more accurately.

The RMSE error criterion was used to graphically compare the performance of single and integrated KELM models. The results are shown in Figure 9. As can be seen, in both SSC and SSD modeling processes, the RMSE values were smaller for integrated methods, and the WT-KELM model led to more accurate results.

Uncertainty analysis results

In this part of the study the uncertainty analysis (UA) was carried out to determine the uncertainty of the best-KELM model. In this study, the Monte Carlo uncertainty analysis method was used. In the UA method, two elements are used to test the robustness and analyse the model's uncertainty. The first one is the percentages of the studied outputs which are in the range of 95 PPU and the next

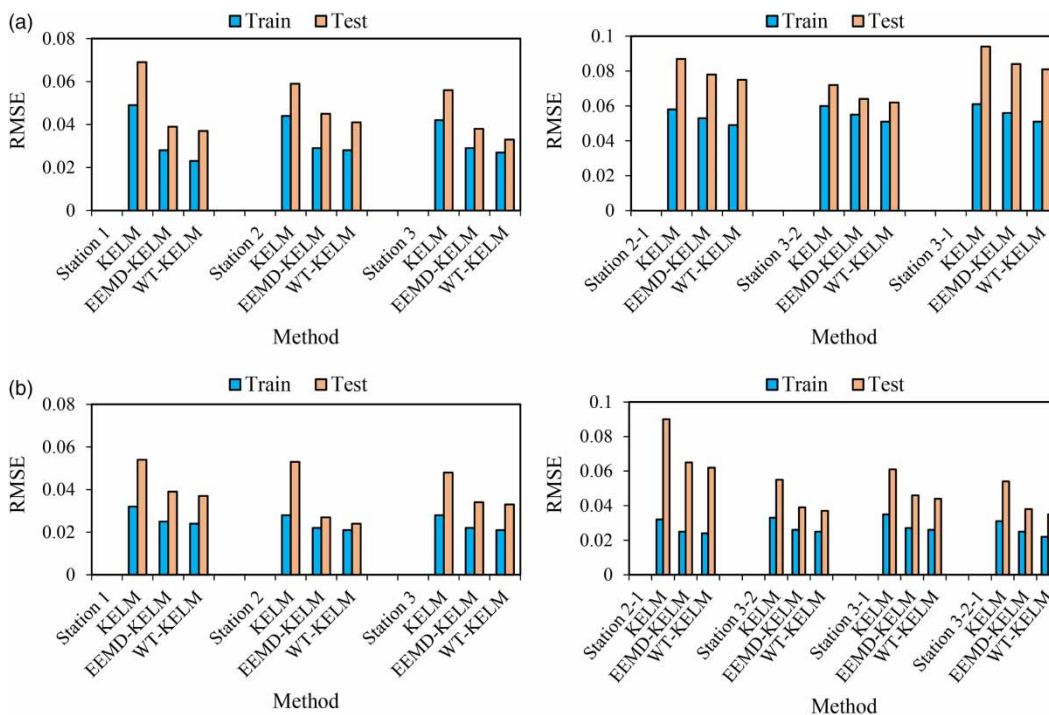


Figure 9 | Comparison of the values of the RMSE criterion for superior models of used methods: (a) SSC and (b) SSD modeling.

one is the average distance between the upper (XU) and lower (XL) uncertainty bands (Noori et al. 2015). In this regard, the considered model should be run many times (1,000 times in the research), and the empirical cumulative distribution probability of the models calculated. The upper and lower bands are considered 2.5 and 97.5% probabilities of the cumulative distribution, respectively. In the proper confidence level two important indices should be considered. First, the 95 PPU band brackets most of the observations. Second, the average distance between the upper and lower parts of the 95 PPU (d-Factor) should be smaller than the observed data standard deviation (Abbaspour et al. 2007). The mentioned indices were applied for accounting input uncertainties. According to Abbaspour et al. (2007), the average width of the confidence interval band can be calculated as follows:

$$d - \text{factor} = \frac{\overline{dx}}{\sigma_x} \quad (10)$$

where σ_x and \overline{dx} are the observed data standard deviation and the confidence band's average width, respectively. The percentage of the data within the confidence band of 95% is calculated as:

$$\text{Bracketed by 95PPU} = \frac{1}{k} \text{Cont}(j|X_L < X_{reg} < X_U) \quad (11)$$

where 95 PPU shows 95% predicted uncertainty; k shows the number of observed data and X_{reg} shows the current registered data. The obtained results for the uncertainty analysis are shown in Table 7. Based on the values obtained for the d-Factor and 95 PPU, it could be indicated that in both SSD and SSC modeling the observed and predicted values were within the 95 PPU band in most of the cases. Also, it was found that the amount of d-Factors for train and test datasets were smaller than the standard deviation of the observed data. Therefore, based on the results, it could be deduced that the SSD and SSC modeling via integrated WT-KELM model led to an allowable degree of uncertainty.

Investigating the most effective sub-series

Sensitivity analysis is used to evaluate the effect of each sub-series obtained from WT and EEMD on the modeling process. For evaluating the impact of each sub-series, the model SC(III) in the SSC predicting of station 2 was selected and run with all sub-series and then, one of the input sub-series was eliminated and the integrated KELM model was re-run. DC error criterion was used as an indication of the significance of each parameter. Figure 10 shows the sensitivity analysis results. Based on the results, it could be deduced that the IMF9 in the EEMD method and A5 approximation sub-series in the WT method were the most important sub-series in the prediction process.

Table 7 | Uncertainty indices of the KELM and WT-KELM models

Timescale	Station	Performance criteria							
		95 PPU		d-Factor		95 PPU		d-Factor	
KELM									
SSC	1	70.48%	0.235	2	70.48%	0.211	3	70.7%	0.257
	2-1	77.7%	0.214	3-2-1	74.3%	0.129			
SSD	1	76.41%	0.215	2	74.40%	0.223	3	72.7%	0.224
	2-1	74.6%	0.105	3-2-1	73.9%	0.187			
WT-KELM									
SSC	1	89.2%	0.105	2	80.1%	0.108	3	81.5%	0.216
	2-1	84.2%	0.108	3-2-1	80.9%	0.095			
SSD	1	89.51%	0.108	2	87.1%	0.102	3	83.5%	0.204
	2-1	87.32%	0.115	3-2-1	84.9%	0.109			

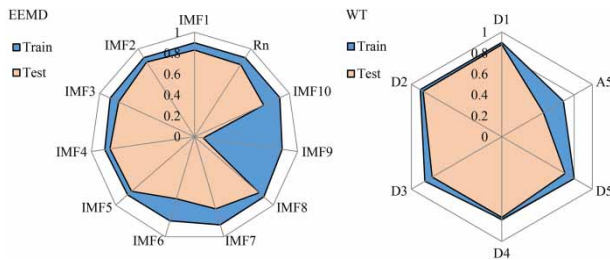


Figure 10 | The impact of each sub-series based on DC performance criterion obtained from sensitivity analysis.

CONCLUSIONS

The accurate prediction of SSC and SSD of rivers is an important factor in improving water management. This study assessed the capability of time series pre-processing methods for the SSC and SSD modeling. In this regard, in the first step, the raw time series (without any data processing) were imposed to the KELM model. Then, time series were decomposed to several sub-series using WT and EEMD and used as inputs of KELM. According to the results, it was found that using both WT and EEMD pre-processing methods increased the model's accuracy. The applied pre-processing method enhanced the KELM model performance between approximately 10 and 18%. It was observed that in estimation of the SSC, using previous suspended sediment concentration led to more accurate results and the use of sediment discharge had no significant impact on the modeling process. It showed that modeling based on each station's own data led to more desirable results. In this state, the model with inputs $Q_{sc}(t-1)$, $Q_{sc}(t-2)$ in SSC modeling and the model with inputs $Q_{sc}(t-1)$, $Q_{sc}(t-2)$, $Q_d(t-1)$, $Q_d(t-2)$ in SSD modeling were superior. However, using the integrated KELM approaches, the previous station data could be used when the station's own data were unavailable. Sensitivity analysis results suggested that the IMF9 in EEMD method and A5 sub-series in WT method were the most effective sub-series in SSC prediction process. Also, it was found that the maximum and minimum values of SSC and SSD variables were well predicted using the integrated models. Therefore, the integration of the KELM model with pre-processing models could be a suitable solution for more accurate prediction of hydrological variables such as suspended sediment concentration and suspended sediment

discharge. It should, however, be noted that the KELM is a data-driven model and the KELM-based model is data sensitive, so further studies using data ranges out of this study should be carried out in the future to determine the merits of the applied model in the SSC and SSD modeling.

DATA AVAILABILITY STATEMENT

All relevant data are available from <https://waterdata.usgs.gov/nwis/sw>

REFERENCES

- Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J. & Srinivasan, R. 2007 *Modelling hydrology and water quality in the prealpine/alpine Thur watershed using SWAT*. *Journal of Hydrology* **333** (2), 413–430.
- Aussem, A., Campbell, J. & Murtagh, F. 1998 Wavelet-based feature extraction and decomposition strategies for financial forecasting. *Journal of Computational Finance* **6** (2), 5–12.
- Azamathulla, H. M., Haghiabi, A. H. & Parsaie, A. 2017 Prediction of side weir discharge coefficient by support vector machine technique. *Water Science and Technology: Water Supply* **16** (4), 1002–1016.
- Babovic, V. 2009 *Introducing knowledge into learning based on genetic programming*. *Journal of Hydroinformatics* **11** (3–4), 181–193.
- Bhattacharya, B., Price, R. K. & Solomatine, D. P. 2004 A data mining approach modeling sediment transport. In: *6th International Conference on Hydroinformatics*. World Scientific, Singapore, pp. 1663–1670.
- Chen, X. Y. & Chau, K. W. 2016 A hybrid double feedforward neural network for suspended sediment load estimation. *Water Resources Management* **30**, 2179–2194.
- Choubin, B., Darabi, H., Rahmati, O., Sajedi-Hosseini, F. & Kløve, B. 2018 *River suspended sediment modelling using the CART model: a comparative study of machine learning techniques*. *Science of the Total Environment* **615**, 272–281.
- Farajzadeh, J. & Alizadeh, F. 2017 A hybrid linear-nonlinear approach to predict the monthly rainfall over the Urmia Lake watershed using Wavelet-SARIMAX-LSKELM conjugated model. *Hydroinformatics* **20** (1), 246–262.
- Harrington, S. T. & Harrington, J. R. 2013 *An assessment of the suspended sediment rating curve approach for load estimation on the Rivers Bandon and Owenabue, Ireland*. *Geomorphology* **185**, 27–38.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. & Liu, H. H. 1998 *The empirical mode decomposition and the Hilbert spectrum for nonlinear*

- and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **454**, 903–995.
- Huang, G. B., Zhu, Q. Y. & Siew, C. K. 2006 **Extreme learning machine: theory and applications**. *Neurocomputing* **70** (1–3), 489–501.
- Huang, G. B., Zhou, H., Ding, X. & Zhang, R. 2012 **Extreme learning machine for regression and multiclass classification**. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **42** (2), 513–529.
- Kaveh, K., Bui, M. D. & Rutschmann, P. 2017 **A comparative study of three different learning algorithms applied to ANFIS for predicting daily suspended sediment concentration**. *International Journal of Sediment Research* **32** (3), 340–350.
- Kisi, O., Dailr, A. H., Cimen, M. & Shiri, J. 2012 **Suspended sediment modeling using genetic programming and soft computing techniques**. *Journal of Hydrology* **450**, 48–58.
- Labate, D., La Foresta, F., Occhiuto, G., Morabito, F. C., Lay-Ekuakille, A. & Vergallo, P. 2013 **Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ECG: a comparison**. *IEEE Sensors Journal* **13** (7), 2666–2674.
- Lafdani, E. K., Nia, A. M. & Ahmadi, A. 2013 **Daily suspended sediment load prediction using artificial neural networks and support vector machines**. *Journal of Hydrology* **478**, 50–62.
- Legates, D. R. & McCabe, G. J. 1999 **Evaluating the use of ‘goodness of fit’ measures in hydrologic and hydroclimatic model validation**. *Water Resources Research* **35** (1), 233–241.
- Lei, Y., He, Z. & Zi, Y. 2009 **Application of the EEMD method to rotor fault diagnosis of rotating machinery**. *Mechanical Systems and Signal Processing* **23** (4), 1327–1338.
- Mehr, A. D., Kahya, E. & Olyaie, E. 2013 **Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique**. *Journal of Hydrology* **505**, 240–249.
- Mitchell, T. M. 1997 *Machine Learning*. McGraw-Hill, New York.
- Noori, R., Deng, Z., Kiaghadi, A. & Kachoosangi, F. T. 2015 **How reliable are ann, anfis, and svm techniques for predicting longitudinal dispersion coefficient in natural rivers?** *Hydraulic Engineering* **142** (1), 04015039.
- Pachori, R. B., Avinash, P., Shashank, K., Sharma, R. & Acharya, U. R. 2015 **Application of empirical mode decomposition for analysis of normal and diabetic RR-interval signals**. *Expert Systems with Applications* **42** (9), 4567–4581.
- Rahman, S. A. & Chakrabarty, D. 2020 **Sediment transport modelling in an alluvial river with artificial neural network**. *Journal of Hydrology* **12**, 50–56.
- Rajae, T., Mirbagheri, S. A., Zounemat-Kermani, M. & Nourani, V. 2009 **Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models**. *Science of the Total Environment* **407**, 4916–4927.
- Roushangar, K. & Ghasempour, R. 2018 **Explicit prediction of expanding channels hydraulic jump characteristics using gene expression programming approach**. *Hydrology Research* **49** (3), 815–830.
- Roushangar, K., Ghasempour, R. & Biukaghazadeh, S. 2020 **Evaluation of the parameters affecting the roughness coefficient of sewer pipes with rigid and loose boundary conditions via kernel based approaches**. *International Journal of Sediment Research* **35** (2), 171–179.
- Sadeghpour, M. 2014 **A wavelet support vector machine combination model for daily suspended sediment forecasting**. *International Journal of Engineering* **27** (6), 855–864.
- Saghebian, S. M., Roushangar, K., Ozgur Kirca, V. S. & Ghasempour, R. 2020 **Modeling total resistance and form resistance of movable bed channels via experimental data and a kernel-based approach**. *Journal of Hydroinformatics* **22** (3), 528–540.
- Vongvisessomjai, N., Tingsanchali, T. & Babel, M. S. 2010 **Non-deposition design criteria for sewers with part-full flow**. *Urban Water Journal* **7** (1), 61–77.
- Wu, Z. H. & Huang, N. E. 2009 **Ensemble empirical mode decomposition: a noise assisted data analysis method**. *Advances in Adaptive Data Analysis* **1**, 1–41.
- Yin, Z., Wen, X., Feng, Q., He, Z., Zou, S. & Yang, L. 2017 **Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area**. *Hydrology Research* **48** (5), 1177–1191.

First received 20 September 2020; accepted in revised form 6 January 2021. Available online 22 February 2021