ISH Journal of Hydraulic Engineering

ISSN: 0971-5010 (Print) 2164-3040 (Online) Journal homepage: http://www.tandfonline.com/loi/tish20

Estimation of suspended sediment load using regression trees and model trees approaches (Case study: Hyderabad drainage basin in Iran)

Ali Talebi, Javad Mahjoobi, Mohammad Taghi Dastorani & Vahid Moosavi

ISH JOURNAL OF HYDRAULIC ENGINEERING, 2016
http://dx.doi.org/10.1080/09715010.2016.1264894

Taylor & Francis
Taylor & Francis Group

Estimation of suspended sediment load using regression trees and model trees approaches (Case study: Hyderabad drainage basin in Iran)

Ali Talebi[a], Javad Mahjoobi[b], Mohammad Taghi Dastorani[c] and Vahid Moosavi[a]

[a]Faculty of Natural Resources, Yazd University, Yazd, Iran; [b]Ministry of Energy, Yazd Regional Water Authority, Yazd, Iran; [c]Faculty of Natural Resources and Environment, Ferdowsi University of Mashhad, Mashhad, Iran

ABSTRACT

Estimation of suspended sediment load is one of the important topics in river engineering. Different methods are used for estimating the sediment rate. In recent years, different artificial intelligence (AI) methods, such as artificial neural network (ANN), have been used for the estimation of sediments in rivers. In this research, the suspended sediment load has been studied by using regression trees (RTs) and model trees (MTs). The study area has been located in Hyderabad watershed in west of Iran. The input data included the flow discharge, sum of three days discharge, sum of five days precipitation and the suspended sediment discharge were considered as output in the models. The numbers of total data of sediment discharge was 223 records. The obtained results were compared with ANN method (feed forward back propagation algorithm) and sediment rating curve (SRC). Results showed that RT and MT outperformed ANN method in the study area. The method of SRC had high accuracy for daily sediment discharge less than 100 ton per day in comparison with AI models, while the AI models had higher accuracy for high sediment discharge. Moreover, the combination of artificial intelligent models had high accuracy regarding to each model lonely.

ARTICLE HISTORY
Received 28 May 2016
Accepted 22 November 2016

KEYWORDS
Suspended sediment; CART; M5′ algorithm; ANN; sediment rating curve

1. Introduction

Estimation of sediment transport rate is one of the basic problems in river engineering. Several empirical methods have been developed to solve this problem. As these methods have been obtained based on climatic conditions of other parts of the world, they have high level of errors when are used in rivers of Iran. One of the common methods for estimating the suspended load in rivers is the rating curve method in which the relation between flow discharge and sediment discharge is presented as a power equation. In recent years, the methods based on artificial intelligence (AI) and machine learning have been used for the estimation and prediction of different phenomena in river engineering. The artificial neural network (ANN) is one of these methods that were used by many scientists for estimating the sediment rates in rivers (Abrahart and White 2001; Jain 2001; Nagy et al. 2002; Tayfur 2002; Merritt et al. 2003; Yitian and Gu 2003; Cigizoglu 2004; Kisi 2004; Agarwal et al. 2005; Cigizoglu and Alp 2006; Cigizoglu and Kisi 2006; Cigizoglu and Alp 2007; Dogan et al. 2007).

Decision trees (DT) are one of the other common and strong tools for prediction and classification. In contrast to ANN, DT produces the roles. This means that DT presents its prediction based on the role set, while in the procedure in ANN is not transparent and it is like a black box. Recently, application of regression trees (RTs) and model trees (MTs) have been presented in water resource engineering field. Mahjoobi and Etemad-Shahidi (2008) predicted the wave height due to wind in Lake Michigan using RT and applying classification and regression trees (CART) algorithm. Ayoubloo et al. (2010) investigated the regular wave scour around a circular pile using regression tress (CART algorithm). Moreover, Etemad-Shahidi and Mahjoobi (2009) predicted the significant wave height in Lake Superior using MT model and applying the M5′ algorithm. MTs have also been applied in rainfall–runoff modeling (Solomatine and Dulal 2003); flood forecasting (Solomatine and Yunpeng 2004); modeling water-level discharge relationship (Bhattacharya and Solomatine 2005), sediment transport (Bhattacharya et al. 2007), derivation of wave spectrum (Sakhare and Deo 2009), estimation of wind speed from wave measurements (Daga and Deo 2009) and prediction of suspended sediment load in rivers. Reddy and Ghimire (2009) applied the M5 MT and Gene Expression Programming to predict suspended sediment load. They also compared the obtained results with sediment rating curve (SRC) and multiple linear regressions (MLRs) and concluded that MT gives good performance as compared with other used models. Etemad-Shahidi and Ghaemi (2011) used MT method to predict pile groups scour due to waves. They presented new equations using MT method and demonstrated that the proposed equations were as accurate as other soft computing methods, such as ANN and SVM. Bonakdar and Etemad-Shahidi (2011), predicted wave run-up on rubble-mound structures using M5 MT. They stated that the main advantage of MTs, unlike the other soft computing tools, is their easier use and more importantly their understandable mathematical rules. They showed that the predictive accuracy of the MT approach was superior to that of Van der Meer and Stam's empirical formula. Wolfs and Willems (2014) developed discharge-stage curves using

CONTACT Ali Talebi talebisf@yazd.ac.ir

several various approaches, i.e., single rating curves, rating curves with dynamic correction, ANNs and M5′ MTs. They showed that all abovementioned methods outperformed the traditional rating curve. Abolfathi et al. (2016), used M5′ DT algorithm to predict the wave run-up using existing laboratory data. They demonstrated that the M5′ MT algorithm had high precision in predicting the wave run-up. They also showed that a good agreement existed between the proposed run-up formulae and existing empirical relations. Zounemat-Kermani et al. (2016) used 8-year data series from hydrometric stations located in Arkansas, Delaware and Idaho (USA), to assess the ability of ANN and support vector regression (SVR) models to forecast/estimate daily suspended sediment concentrations and to compare the results with traditional MLR and SRC models. They tested three different ANN model algorithms, along with four different SVR model kernels. They showed that ANN and SVR outperformed traditional methods. Shamaei and Kaedi (2016) introduced stacking method to predict the suspended sediment. They used linear genetic programming and neuro-fuzzy methods as two successful soft computing methods to predict the suspended sediment. Then, they increased the accuracy of prediction by combining their results with the meta-model of neural network based on cross validation. The obtained results demonstrated that the stacking method greatly improved root mean square error (RMSE) and R2 statistics compared to use of linear genetic programming or neuro-fuzzy solitarily. Makarynskyy et al. (2015) used two numerical current and wave models in addition to AI technique of neural networks (ANNs) to reproduce values of sediment concentrations observed at two sites. They showed that ANN method provides accurate results. Nourani et al. (2016) used a two-stage modeling strategy in order to handle spatio-temporal variation of SSL. At temporal stage, they used support vector machine (SVM) to find the nonlinear relationship of SSL in time domain. In spatial modeling stage, they used semivariogram of monthly SSL data and then they fitted theoretical semivariogram model to the empirical variogram. The obtained results showed that the hybrid of SVM and Spatial statistics methods could predict and simulate SSL appropriately by enjoying unique features of both approaches. Chen and Chau (2016) used a hybrid double feedforward neural network (HDFNN) model for daily *SSL* estimation, by combining fuzzy pattern-recognition and continuity equation into a structure of double neural networks. They showed that HDFNN is appropriate for modeling the sediment transport process with nonlinear, fuzzy and time-varying characteristics. Shiau and Chen (2015) developed a probabilistic estimation scheme for daily and annual suspended sediment loads using quantile regression. They used daily suspended sediment load and discharge data to construct quantile-dependent SRCs. Their proposed approach was applied to the Laonung station located in southern Taiwan. The results indicated that the proposed approach provided not only the probabilistic description for daily and annual suspended sediment loads, but also the single estimations including the mean, median and mode of the derived probability distribution. The main purpose of this research is to apply the RT model (CART algorithm) and MTs (M5′ algorithm) for estimating the suspended sediment load in Hyderabad watershed, west Iran. In addition, the obtained results of these two methods will be compared with the SRC method and ANN model (feed forward back propagation algorithm).

2. Materials and methods

2.1. Study area and data

This research has been done on Hyderabad watershed in Kermanshah province in western part of Iran (Figure 1). The total area of watershed is 1719 km^2, mean height is 1871 m, maximum height 3300 m and minimum height is 1325 m. This watershed has been located in 47° 04′–47° 52′ longitudes and 34° 25′–34° 52′ latitude. The main river of the watershed is the Jamishan permanent river. The meteorological station of the watershed is Hyderabad station with 47° 27′ longitude and 34° 42′ latitude (Figure 2). The precipitation regime of the study area is rainy-snowy and mean annual precipitation is 420 mm which is mostly occurred in winter and spring. The used data in this research involve precipitation, flow discharge and sediment discharge. The length of data period is 21 years (from

Figure 1. The position of the study area in Iran. Source: The authors.

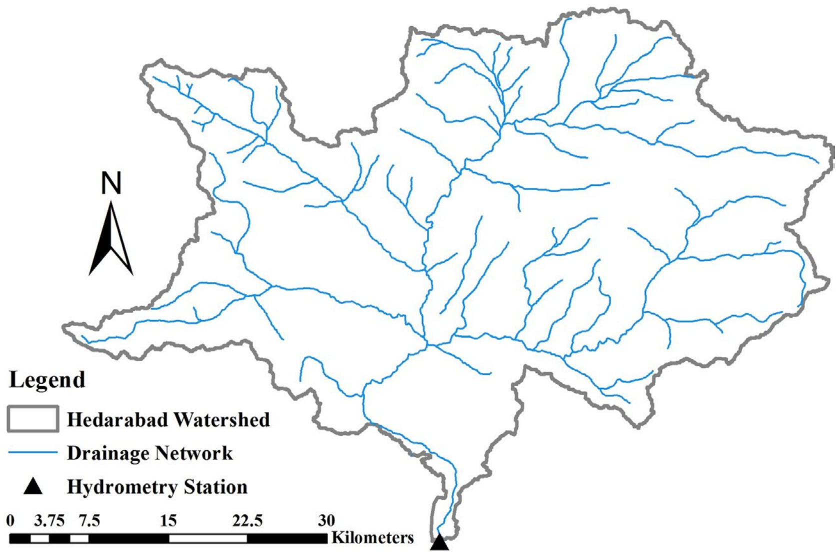

Figure 2. Drainage network and hydrometry station in Hydrabad watershed. Source: The authors.

Table 1. Ranges and average values of different parameters in training and test data sets.

Parameters	Training data set			Test data set		
	Minimum	Maximum	Average	Minimum	Maximum	Average
Water discharge (m^3/s)	0.04	285.41	17.95	0.25	188.67	23.35
Suspended sediment discharge (ton/day)	0.0001	215.897	2027.66	0.397	68,107.35	2696.08

1985 to 2006) with the total number of 223 samples. Eighty percent of these data have been used for training and 20% for testing and evaluating the models. One of the problems that occur during training process is called overfitting. The error on the training set is driven to a very small value, but when new data is presented to the network the error is large. The network has memorized the training examples, but it has not learned to generalize to new situations. One of the main ways to avoid overfitting (or recognize if occurs) is to separate data to training and test data sets. The training subset is composed of 60–80% of all the records. The remaining records are usually used as test data set. Gharagheizi (2007) showed that the percent of test set allocated from the main data set should be between 5% and 35%. If this percent is lower than 5%, the accuracy of the model over the training set is much greater than the test set. Also, if the percent is greater than 40%, the obtained model cannot predict the test set as well as the training set. Each record should be randomly chosen from the data set and placed in one of the two subsets. Therefore, the data were separated to training and test data sets using a common random method. The ranges and average values of water and sediment discharge for training and testing have been shown in Table 1.

2.2. RTs (CART algorithm)

The CART method developed by Breiman et al. (1984) generates binary DTs. CART is a nonparametric statistical methodology developed for analyzing classification issues either from categorical or continuous dependent variables. If the dependent variable is categorical, CART produces a classification tree. When the dependent variable is continuous, it produces a RT. The CART tree is constructed by splitting subsets of the data set using all predictor variables to create two child nodes repeatedly, beginning with the entire data set. The best predictor is chosen using a variety of impurity or diversity measures. The goal is to produce subsets of the data which are as homogeneous as possible with respect to the target variable. In CART algorithm for each split, each predictor is evaluated to find the best cut point (continuous predictors) or groupings of categories (nominal and ordinal predictors) based on improvement score or reduction in impurity (Breiman et al. 1984). Then, the predictors are compared and the predictor with the best improvement is selected for the split. The process repeats recursively until one of the stopping rules is triggered. RT building centers on three major components: (1) a set of questions of the form: is $X \le d$? where X is a variable and d is a constant. (2) Goodness of split criteria for choosing the best split on a variable and (3) the generation of summary statistics for terminal nodes. The least-squared deviation (LSD) impurity measure is used for splitting rules and goodness of fit criteria**.** The LSD measure $R(t)$ is simply the weighted within node variance for node t, and it is equal to the resubstitution estimate of risk for the node (Breiman et al. 1984). It is defined as:

$$R(t) = \frac{1}{N_W(t)} \sum_{i \in t} \omega_i f_i \left(y_i - \bar{y}(t)\right)^2 \tag{1}$$

$$\bar{y}(t) = \frac{1}{N_W(t)} \sum_{i \in t} \omega_i f_i y_i \tag{2}$$

$$N_W(t) = \sum_{i \in t} \omega_i f_i \tag{3}$$

where $N_W(t)$ is the weighted number of records in node t, ω_i is the value of the weighting field for record i (if any), f_i is the value of the frequency field (if any), y_i is the value of the target field, and $\bar{y}(t)$ is the mean of the dependent variable (target field) at node t. The *LSD* criterion function for split s at node t is defined as follows:

$$Q(s,t) = R(t) - R(t_L) - R(t_R) \quad (4)$$

where $R(t_R)$ is the sum of squares of the right child node and $R(t_L)$ is the sum of squares of the left child node. The split s is chosen to maximize the value of $Q(s, t)$. Stopping rules control how the algorithm decides when to stop splitting nodes in the tree. Tree growth proceeds until every leaf node in the tree triggers at least one stopping rule. Any of the following conditions will prevent a node from being split:

(1) All records in the node have the same value for all predictor fields used by the model.
(2) The number of records in the node is less than the minimum parent node size (user defined).
(3) If the number of records in any of the child nodes resulting from the node's best split is less than the minimum child node size (user defined).
(4) The best split for the node yields a decrease in impurity that is less than the minimum change in impurity (user defined).

In RTs, each terminal node's predicted category is the weighted mean of the target values for records in the node ($\bar{y}(t)$).

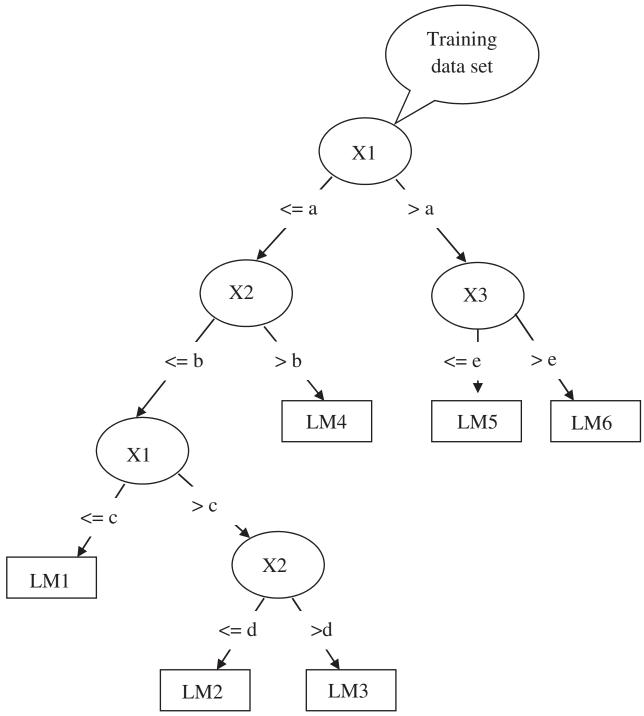

Figure 3. MT used to split input space (X_i: inputs, LM_i: linear model).

2.3. MTs (M5′ algorithm)

MTs (Quinlan 1992) are an extension of RTs in the sense that they associate leaves with multivariate linear models. MTs are a technique for dealing with continuous class problems that provide a structural representation of the data and a piecewise linear fit of the class. They have a conventional DT structure but use linear function at the leaves instead of discrete class labels (Figure 3). M5 MTs were first introduced by Quinlan (1992), and then, the idea was reconstructed and improved in a system called M5′ by Wang and Witten, (1997). An M5′ MT is an effective learning method for predicting real values. M5′ MT algorithm first constructs a RT by recursively splitting the instance space. The splitting criterion is used to minimize the intrasubset variability in the values down from the root through the branch to the node. The variability is measured by the standard deviation of the values that reach that node from the root through the branch with calculating the expected reduction in error as a result of testing each attribute at that node. The attribute that maximizes the expected error reduction is chosen. The splitting stops if the values of all instances that reach a node vary slightly or only a few instances remain. The standard deviation reduction (SDR) is calculated by:

$$\mathrm{SDR} = \mathrm{sd}(T) - \sum_i \frac{|T_i|}{|T|} \times \mathrm{sd}(T_i) \quad (5)$$

where T is the set of examples that reach the node, T_i is the sets that are resulted from splitting the node according to the chosen attribute and SD is the standard deviation (Wang and Witten 1997). After the tree has been grown, M5′ computes a linear multiple regression model for every interior node. The data associated with that node and only the attributes tested in the subtree rooted at that node are used in the regression. The attributes will be dropped one by one if they lower the estimated error. Then the tree is pruned from the leaves if that results in a lower expected estimated error. In Wang and Witten, (1997)'s implementation, the expected error is calculated by averaging the absolute difference between the predicted value and the actual value for each of the training examples that reach that node. This results in underestimation of the expected error outside the calibrating data. The expected error is multiplied by $(n + v)/(n - v)$, where n is the number of training instances that reach the node and the v is the number of parameters in the model that represent the value at that node (Wang and Witten 1997). After pruning, the adjacent linear models will be sharply discontinuous at the leaves of the pruned tree. M5 applies smoothing process combining the model at a leaf with the models on the path to the root to form the final model that is placed at the leaf. In the smoothing process, the estimated value of the leaf model is filtered along the path back to the root. At each node, that value is combined with the value predicted by the linear model for that node as follows:

$$P' = \frac{np + kq}{n + k} \quad (6)$$

where P' is the prediction passed up to the next higher node, p is the prediction passed to this node from the below, q is the value predicted by the model at this node, n is the number of training instances that reach the node below, and k is a constant (Wang and Witten 1997). Experiments of Wang and Witten, (1997) have showed that smoothing substantially increases the accuracy of predictions.

2.4. ANNs and SRC

ANNs are powerful nonlinear modeling approaches based on the function of human brain. They can identify and learn correlated patterns between input data sets and target values. Neural networks can be described as a network of simple processing nodes or neurons, interconnected to each other in a specific order, performing simple numerical manipulations (See and Openshaw 1999). A three-layered neural network is consists of several elements namely nodes. These networks are made up of an input layer consisting of nodes representing different input variables, the hidden layer consisting of many hidden nodes and an output layer consisting of output variables (Haykin 1999). ANN is widely applied in hydrology and water resource studies as a forecasting tool. Feed-forward neural networks are applied successfully in many different problems. This network architecture and the corresponding learning algorithm can be viewed as a generalization of the popular least-mean-square algorithm (Haykin 1999). Several training algorithms, such as Gradient descent with momentum and adaptive learning rate back-propagation (GDX), Levenberg–Marquardt (LM) and Bayesian regularization, were used in this study to train the networks. The association between rate of sediment discharge and rate of water discharge at a cross section of a stream is frequently expressed by an average curve. This curve is the SRC. It has been widely used in the computation of average sediment discharge from water discharge for periods when sediment samples were not collected.

Equation (7) provides the power relationship between the sediment load Q_s (t/d) and discharge Q (m^3/s), where α is the rating coefficient ($t\ s^{\beta} d^{-1}\ m^{-3\beta}$) and β is the rating exponent (dimensionless).

$$Q_s = \alpha Q^{\beta} \tag{7}$$

The rating curve parameters, α and β, are each influenced by different physical processes; α represents an index of soil erodibility (Horowitz 2003), while β is influenced by river erosive and transport power (Heng and Suetsugi 2014).

2.5. Evaluation criteria

For statistical comparison of predicted and observed values discrepancy ratio (Dr), correlation coefficient (R), RMSE, scatter index (SI) and mean absolute error (MAE) were used. These statistical measures are defined as follows:

$$\bar{D}r = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i} \tag{8}$$

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum (x_i - y_i)^2} \tag{10}$$

$$\mathrm{SI} = \frac{RMSE}{\bar{x}} \tag{11}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i| \tag{12}$$

In all formulas, x_i is an observed value, y_i is a predicted value and n is the number of observations, finally, $\bar{x}$ is the mean of x and $\bar{y}$ is the mean of y. R can be between −1 and 1. Zero indicates no relationship. +1 indicates a perfect positive relationship and −1 indicates a perfect negative linear relationship. The best value for Dr is one. Higher and lower values show that the observed and estimated values are far from each other. The best value for MAE and RMSE is zero. The higher values show the higher amounts of error. It presents the amount of RMS difference with respect to mean observation. The best value for SR is zero.

3. Results and discussion

By using the training data (178 data), RT (CART algorithm) and MT (M5′ algorithm) were built. The used input data are including the water instantaneous discharge (related flow discharge with sediment discharge), cumulative water discharge (3 days) and cumulative rainfall (5 days). The suspended sediment has also been considered as the output data. The developed tree and produced laws by using M5′ algorithm are shown in Figure 4. As the figure shows, the produced branches and laws are only based on sediment discharge and cumulative 3 days discharge. This means that the cumulative 5 days precipitation has the less important level. After applying the MT and RT, the obtained results have been compared with the results of ANN model (feed forward back propagation algorithm) and SRC. Deo et al. (2001) implied that any nonlinear mathematical dependency structure can be approximated using a three layered feed-forward neural network. To prevent overfitting during the training of the ANN, the number of nodes of the hidden layer was chosen using expression given by Huang and Foo (2002):

$$M \le 2Z + 1 \tag{13}$$

where M and Z are number of the nodes in hidden and input layers, respectively. The number of the neurons of the input and output layers were 3 and 1, respectively. One hidden layer with seven neurons was found to be the best topology.

Table 2 shows the values of statistical indices for the different models. As can be seen, the RMSE, SI and MAE parameters in CART algorithm have the least values. This means that for estimating the suspended sediment load, the accuracy of RTs and MTs is more than ANN model. The least value of Dr is related to SRC. Although R value for ANN and MT is 0.99, but other error indices are high, the R index is not good criteria for judgment. It should be mentioned that the R value can be 1 with 100% error; for instance, if the obtained value of ANN is two times of observed values (100% error in prediction), the R value will also be 1. Figure 5 shows the SRC for training data. Using the obtained equation, the sediment values for test data have been computed and have been compared with the observed data. These results are shown in Table 2.

The comparison of predicted and observed values of suspended load in ANN, M5′ and CART models are shown in Figures 6–8, respectively. To evaluate the behavior of the used models in low and high values of sediment load, the error indices were computed for values less than 100, 100–1000 and more than 1000 ton/day. The obtained results are presented in Table 3. As can be seen in Figure 6, the CART algorithm underestimates the suspended load when this rate is more than 1000 ton/day. For instance, the value of 68,107 ton/day has been estimated equal to 17,383 ton/day (0.25 of real value). Moreover, this algorithm overestimates the suspended load when the value lies between 1000 and 10,000 ton/day as the observed value of 7476 ton/day has been estimated equal to 19,046 ton/day (almost 2.5 times more). The sediment rate more than 10,000 ton/day has been overestimated (two times more) by M5′ algorithm and ANN model. The SRC also underestimates the sediment rate more than 10,000 ton/day (0.4 times of real

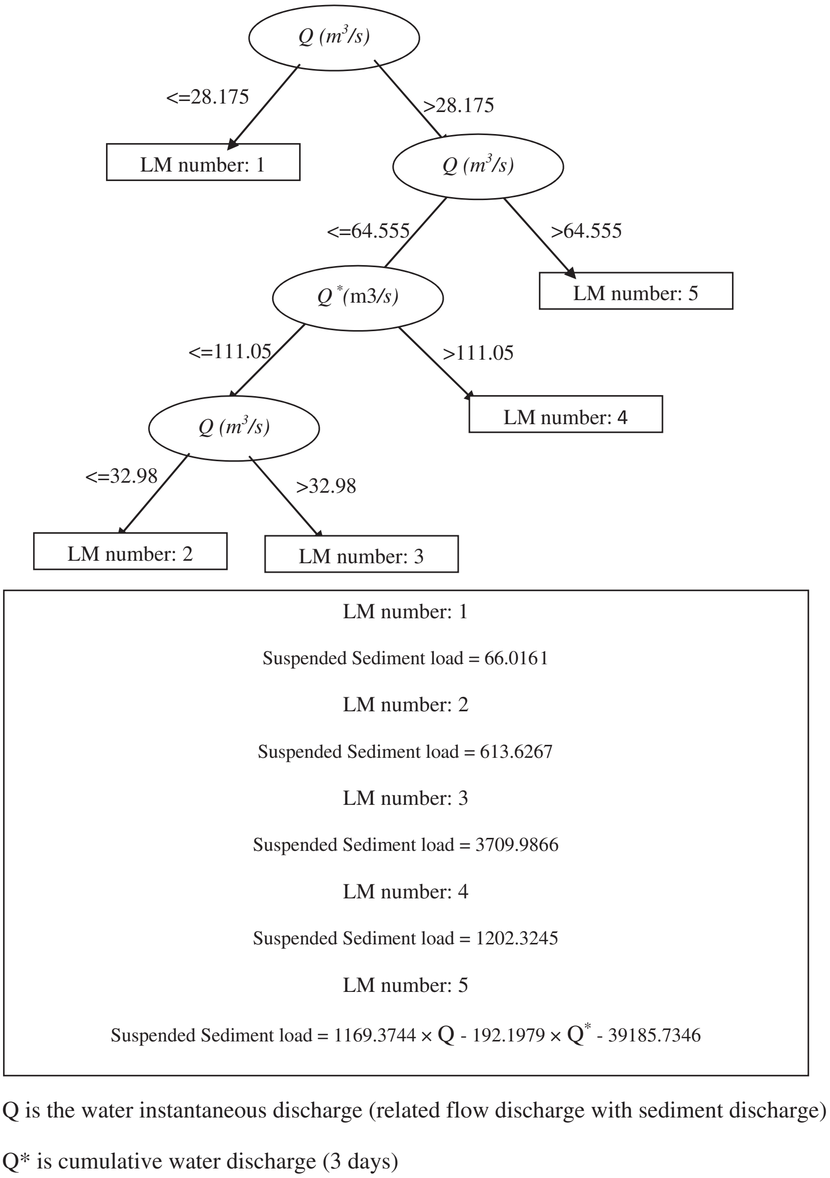

Figure 4. MT generated by M5′ algorithm, LM is linear model and Q is water discharge and Q^* is sum of 3 days water discharge.
Notes: Q is the water instantaneous discharge (related flow discharge with sediment discharge); Q^* is cumulative water discharge (3 days).

Table 2. The statistical measures for the different methods (for test data set).

Method	RMSE	SI (%)	R	MAE	$\bar{Dr}$
CART	7952	294.94	0.75	1958.3	4.44
MT	10,773.2	399.59	0.99	2545.3	37.43
ANN	13,831	513.002	0.99	2929.3	37.5
SRC	9530.36	353.49	0.91	2227.48	2.3

value). For example, the observed value of 68,107 has been estimated equal to 7990 ton/day (0.12 time of real value). As Table 3 shows, the SRC method has higher accuracy in comparison with the AI models for sediment discharge less than 100 ton/day, and for sediment discharge more than 100 ton/day, the CART algorithm has more accuracy than the other models (Figure 9).

To obtain the high accuracy, the geometric average of outputs of AI models (committee model) can be used as follows:

$$O_c = \sqrt[3]{(O_{ANN} \times O_{RT} \times O_{MT})} \qquad (14)$$

where O_c is output of committee model, O_{ANN} is output of ANN model, O_{RT} is the output of RT model and O_{MT} is the output of tree model. The obtained results by computing the error parameters have been shown in Table 3. As it is shown, the accuracy of geometric average model is more than each model individually.

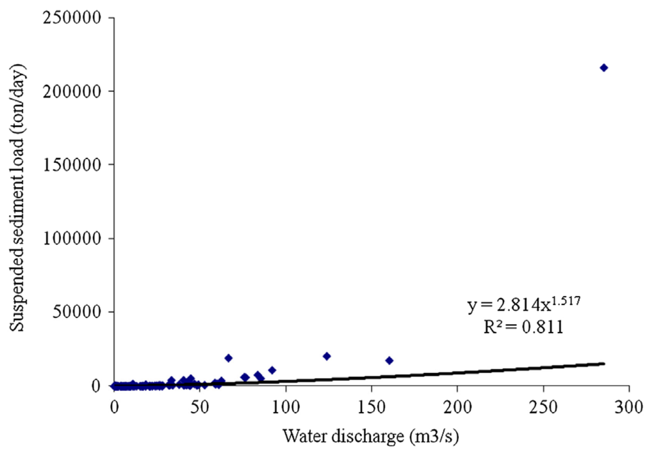

Figure 5. The SRC for training data.

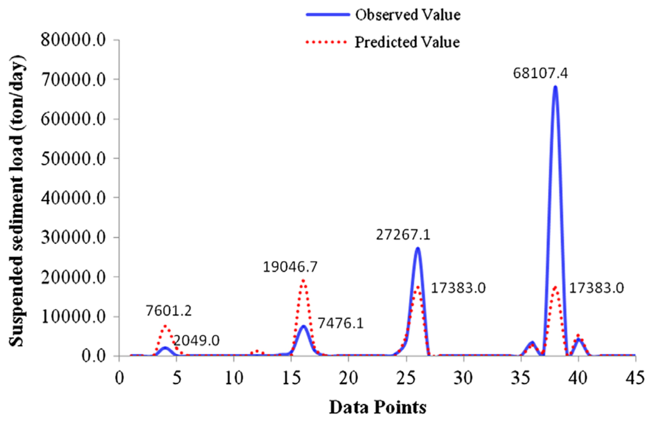

Figure 6. Comparison between observed and predicted suspended sediment load by CART model for test data set.

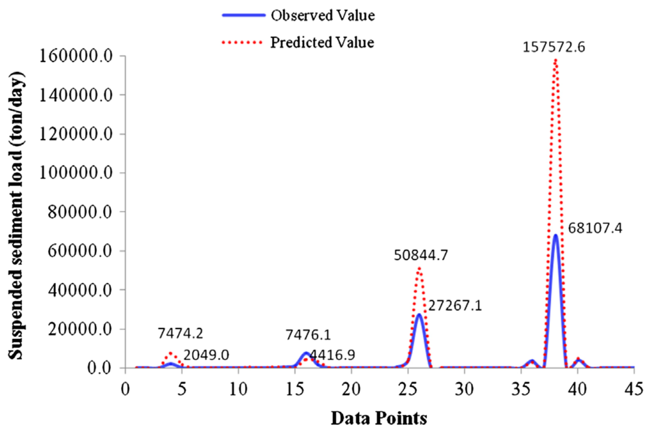

Figure 8. Comparison between observed and predicted suspended sediment load by ANN model for test data set.

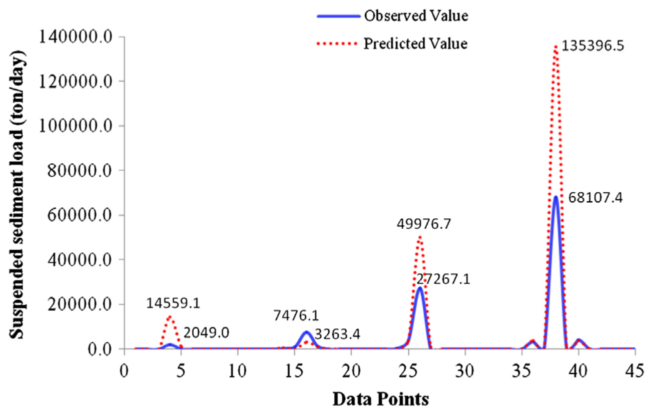

Figure 7. Comparison between observed and predicted suspended sediment load by MTs for test data set.

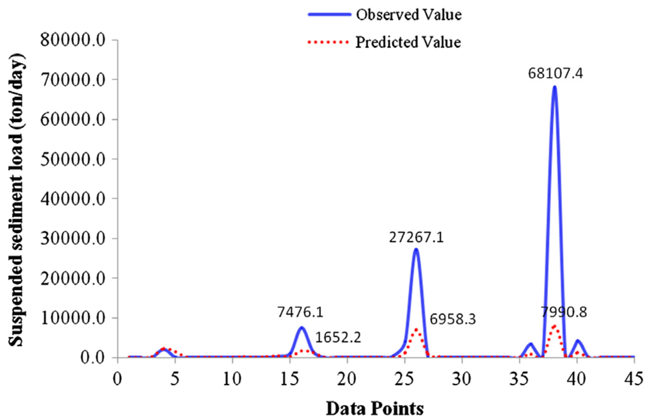

Figure 9. Comparison between observed and predicted suspended sediment load by SRC for test data set.

Table 3. Error statistics of suspended sediment load by CART, M5′, ANN, SRC and committee algorithms in different suspended sediment load ranges.

Method	Suspended sediment load ranges (*qs*)	Number of data	RMSE	SI (%)	MAE	$\bar{D}r$
CART algorithm	0 < *qs* ≤ 100	29	403.14	2671.36	131.7	6.11
	100 < *qs* ≤ 1000	7	360.7	158	271	1.49
	qs > 1000	9	17,763.6	134	9156.6	1.35
M5′ algorithm	0 < *qs* ≤ 100	29	305.4	2023.7	130.7	57.2
	100 < *qs* ≤ 1000	7	289.9	127	232.9	1.42
	qs > 1000	9	24,081.2	181.7	12,124.2	1.68
ANN	0 < *qs* ≤ 100	29	370.18	2452.95	161.86	57.25
	100 < *qs* ≤ 1000	7	240.2	105.2	183.68	2.03
	qs > 1000	9	30,919.1	233.27	13,982	1.55
Sediment rating curve	0 < *qs* ≤ 100	29	291.7	1933.29	91.59	3.01
	100 < *qs* ≤ 1000	7	266.7	116.8	235.9	1.8
	qs > 1000	9	21,302	160.7	10,658.8	0.4
Committee model	0 < *qs* ≤ 100	29	318.4	2109.7	105.7	21.05
	100 < *qs* ≤ 1000	7	196.2	85.96	136.3	1.3
	qs > 1000	9	3885.4	29.3	2536.9	1.37

4. Conclusions

In this research, the accuracy of RTs and MTs for predicting the suspended sediment load have been investigated in Hyderabad watershed in west of Iran (Jamishan river). The obtained results have compared with results of ANN model (feed forward back propagation algorithm) and SRC method. The results showed that the accuracy of RT and MT were more than the ANN model. Comparison of different methods of AI and SRC methods indicated that the SRC method has more accuracy for estimating the sediment load when the sediment discharge is less than 100 ton/day, while for values more than 100 ton/day; the SRC method underestimates the suspended sediment load. This means that AI models have high accuracy in comparison with SRC method for high discharges. Therefore, as in most problems in river engineering, the peak discharge has more importance; the AI models (like RT and MT) are more applicable. Moreover, results showed that by using the combination of outputs of AI models (geometric average of outputs); more accurate results can be obtained rather than each model separately.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abolfathi, S., Yeganeh-Bakhtiary, A., Hamze-Ziabari, S.M., and Borzooei, S. (2016). "Wave runup prediction using M5′ model tree algorithm." *Ocean Eng.*, 112, 76–81.

Abrahart, R.J., and White, S.M. (2001). "Modelling sediment transfer in Malawi: comparing backpropagation neural network solutions against a multiple linear regression benchmark using small data sets." *Phys. Chem. Earth Part B*, 26(1), 19–24.

Agarwal, A., Singh, R.D., Mishra, S.K., and Bhunya, P.K. (2005). "ANN-based sediment yield models for Vamsadhara river basin (India)." *Water SA*, 31(1), 95–100.

Ayoubloo, M.K., Etemad-Shahidi, A., and Mahjoobi, J. (2010). "Evaluation of regular wave scour around a circular pile using data mining approaches." *Appl. Ocean Res.*, 32(1), 34–39.

Bhattacharya, B., Price, R.K., and Solomatine, D.P. (2007). "Machine learning approach to modeling sediment transport." *J. Hydraul. Eng.*, 133(4), 440–450.

Bhattacharya, B., and Solomatine, D.P. (2005). "Neural networks and M5 model trees in modelling water level–discharge relationship." *Neurocomputing*, 63, 381–396.

Bonakdar, L., and Etemad-Shahidi, A. (2011). "Predicting wave run-up on rubble-mound structures using M5 model tree." *Ocean Eng.*, 38, 111–118.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). "Classification and regression trees." Wadsworth Statistical Press, Belmont, CA.

Chen, X.Y., and Chau, K.W. (2016). "A hybrid double feedforward neural network for suspended sediment load estimation." *Water Resour. Manage.*, 30(7), 2179–2194.

Cigizoglu, H.K. (2004). "Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons." *Adv. Water Resour.*, 27, 185–195.

Cigizoglu, H.K., and Alp, M. (2006). "Generalized regression neural network in modelling river sediment yield." *Adv. Eng. Softw.*, 37, 63–68.

Cigizoglu, H.K., and Alp, M. (2007). "Suspended sediment load simulation by two artificial neural network methods using hydro meteorological data." *Environ. Model. Softw.*, 22, 2–13.

Cigizoglu, H.K., and Kisi, O. (2006). "Methods to improve the neural network performance in suspended sediment estimation." *J. Hydrol.*, 317, 221–238.

Daga, M., and Deo, M.C. (2009). "Alternative data-driven methods to estimate wind from waves by inverse modeling." *Nat. Hazards*, 49, 293–310.

Deo, M.C., Jha, A., and Chaphekar, A.S. (2001). "Neural networks for wave forecasting." *Ocean Eng.*, 28, 889–898.

Dogan, M., Yüksel, I., and Kisi, O. (2007). "Estimation of total sediment load concentration obtained by experimental study using artificial neural networks." *Environ. Fluid Mech.*, 7, 271–288.

Etemad-Shahidi, A., and Ghaemi, N. (2011). "Model tree approach for prediction of pile groups scour due to waves." *Ocean Eng.*, 38, 1522–1527.

Etemad-Shahidi, A., and Mahjoobi, J. (2009). "Comparison between M5′ model tree and neural networks for prediction of significant wave height in Lake Superior." *Ocean Eng.*, 36(15–16), 1175–1181.

Gharagheizi, F. (2007). "QSPR analysis for intrinsic viscosity of polymer solutions by means of GA-MLR and RBFNN." *Comput. Mater. Sci.*, 40, 159–167.

Haykin, S. (1999). *Neural networks: a comprehensive foundation*, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ.

Heng, S., and Suetsugi, T. (2014). "Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent instantaneous sediment yield prediction." *J. Hydrol.*, 512, 240–253.

Horowitz, A.J. (2003). "An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations." *Hydrol. Process*, 17(17), 3387–3409.

Huang, W., and Foo, S. (2002). "Neural network modeling of salinity variation in Apalachicola river." *Water Res.*, 36, 356–362.

Jain, S.K. (2001). "Development of integrated sediment rating curves using ANNs." *J. Hydraul. Eng.*, 127(1), 30–37.

Kisi, O. (2004). "Multi-layer perceptrons with Levenberge–Marquardt training algorithm for suspended sediment concentration prediction and estimation." *Hydrol. Sci. J.*, 49(6), 1025–1040.

Mahjoobi, J., and Etemad-Shahidi, A. (2008). "An alternative approach for the prediction of significant wave heights based on classification and regression trees." *Appl. Ocean Res.*, 30, 172–177.

Makarynskyy, O., Makarynska, D., Rayson, M., and Langtry, S. (2015). "Combining deterministic modelling with artificial neural networks for suspended sediment estimates." *Appl. Soft Comput.*, 35, 247–256.

Merritt, W.S., Letcher, R.A., and Jakeman, A.J. (2003). "A review of erosion and sediment transport models." *Environ. Model. Softw.*, 18(8–9), 761–799.

Nagy, H.M., Watanabe, K., and Hirano, M. (2002). "Prediction of sediment load concentration in rivers using artificial neural network model." *J. Hydraul. Eng.*, 128(6), 588–595.

Nourani, V., Alizadeh, F., and Roushangar, K. (2016). "Evaluation of a two-stage SVM and spatial statistics methods for modeling monthly river suspended sediment load." *Water Resour. Manage.*, 30(1), 393–407.

Quinlan, J.R. (1992). "Learning with continuous classes." *Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence*, World Scientific, Singapore, 343–348.

Reddy, M.J., and Ghimire, B.N.S. (2009). "Use of model tree and gene expression programming to predict the suspended sediment load in rivers." *J. Intell. Syst.*, 18(3), 211–227.

Sakhare, S., and Deo, M.C. (2009). "Derivation of wave spectrum using data driven methods." *Mar. Struct.*, 22, 594–609.

See, L., and Openshaw, S. (1999). "Applying soft computing approaches to river level forecasting." *Hydrol. Sci. J.*, 44, 763–778.

Shamaei, E., and Kaedi, M. (2016). "Suspended sediment concentration estimation by stacking the genetic programming and neuro-fuzzy predictions." *Appl. Soft Comput.*, 45, 187–196.

Shiau, J.T., and Chen, T.J. (2015). "Quantile regression-based probabilistic estimation scheme for daily and annual suspended sediment loads." *Water Resour. Manage.*, 29(8), 2805–2818.

Solomatine, D.P., and Dulal, K.N. (2003). "Model trees as an alternative to neural networks in rainfall – runoff modelling." *Hydrol. Sci. J.*, 48(3), 399–411.

Solomatine, D.P., and Yunpeng, X. (2004). "M5 model trees and neural networks: application to flood forecasting in the upper reach of the Huai River in China." *J. Hydrol. Eng.*, 9(6), 491–501.

Tayfur, G. (2002). "Artificial neural networks for sheet sediment transport." *Hydrol. Sci. J.*, 47(6), 879–892.

Wang, Y., and Witten, I.H. (1997). "Induction of model trees for predicting continuous lasses." *Proceedings of the Poster Papers of the European Conference on Machine Learning*, University of Economics, Faculty of Informatics and Statistics, Prague.

Wolfs, V., and Willems, P. (2014). "Development of discharge-stage curves affected by hysteresis using time varying models, model trees and neural networks." *Environ. Model. Softw.*, 55, 107–119.

Yitian, L., and Gu, R.R. (2003). "Modeling flow and sediment transport in a river system using an artificial neural network." *Environ. Manage.*, 31(1), 122–134.

Zounemat-Kermani, M., Kişi, Ö., Adamowski, J., and Ramezani-Charmahineh, A. (2016). "Evaluation of data driven models for river suspended sediment concentration modeling." *J. Hydrol.*, 535, 457–472.