Front. Struct. Civ. Eng. 2020, 14(2): 374–386
https://doi.org/10.1007/s11709-019-0600-0

RESEARCH ARTICLE

Prediction of bed load sediments using different artificial neural network models

Reza ASHEGHI, Seyed Abbas HOSSEINI*

Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
**Corresponding author. E-mail: abbas_hoseyni@srbiau.ac.ir*

© Higher Education Press and Springer-Verlag GmbH Germany, part of Springer Nature 2020

ABSTRACT Modeling and prediction of bed loads is an important but difficult issue in river engineering. The introduced empirical equations due to restricted applicability even in similar conditions provide different accuracies with each other and measured data. In this paper, three different artificial neural networks (ANNs) including multilayer percepterons, radial based function (RBF), and generalized feed forward neural network using five dominant parameters of bed load transport formulas for the Main Fork Red River in Idaho-USA were developed. The optimum models were found through 102 data sets of flow discharge, flow velocity, water surface slopes, flow depth, and mean grain size. The deficiency of empirical equations for this river by conducted comparison between measured and predicted values was approved where the ANN models presented more consistence and closer estimation to observed data. The coefficient of determination between measured and predicted values for empirical equations varied from 0.10 to 0.21 against the 0.93 to 0.98 in ANN models. The accuracy performance of all models was evaluated and interpreted using different statistical error criteria, analytical graphs and confusion matrixes. Although the ANN models predicted compatible outputs but the RBF with 79% correct classification rate corresponding to 0.191 network error was outperform than others.

KEYWORDS bed load prediction, artificial neural network, modeling, empirical equations

1 Introduction

Prediction of sediment transported loads due to nonlinear relationships and complicated interaction of involved parameters is a very difficult task. Moreover, the importance of accurate prediction and corresponding extracted information of sediment transport in providing more concise interpretation in different objectives such as morphological and sedimentological features [1,2], water engineering purposes [3–5], maintaining the channel geometry [6], and economic damage of sediment loads [7], as well as river basin management [8,9], has been approved.

Direct measurements or using the proposed empirical relations are the common methods in calculating the transported sediment loads. Due to possible unfeasible and uneconomical aspects in equipped all desired locations for long time direct measurement [4,10], the simple but accurate enough approaches for modeling of sediment loads is preferred. It is approved that the information on the channel, flow and sediment characteristics are essential elements on development of sediment load relations [11]. However, because of wide range of flow conditions as well as dependency of involved parameters to river and sediment characteristics the adopted equations are not unique [12–14]. Moreover, the sophisticated nonlinear relation between flow and sediment loads due to lack of accepted fundamental principles and increased demand of practical application as well as considering the effect of different auxiliary factors cannot be properly modeled by statistical and regression techniques [15–17]. Therefore, it can be expectable that the same formulation not only provide dissimilar scores of accuracy but also usually do not fit with the observed data. This reason implies why significant attempts for improvement of the existing equations should be made.

Due to approved difficulties in direct bed load measurements and involvement of several factors (e.g., change the

Article history: Received Nov 21, 2018; Accepted Mar 24, 2019

flow properties, modifying the bed geometry, friction, roughness, producing bed forms) integrating reliable bed load prediction into hydraulic calculations is a challenging task [18–20]. Furthermore, the associating of bed load measurements with slopes [21–23], make it not only more complicated and expensive than suspended load but also more interest in developing equations for the predictive models [16,24].

In such situations, by applying the soft computing approach and in particular artificial neural network (ANNs) techniques the complex nonlinear behavior can be captured [25,26]. The literatures show that the ANN techniques due to producing the results faster than most conventional methods have successfully been employed in sediment load predictions in hydrological sciences [14,27–37].

In this paper, different optimum multilayer percepterons (MLP), generalized feed forward neural network (GFFN), and radial based function (RBF) models subjected to various training algorithms and activation functions with the aim of bed load sediment prediction in the Main Fork Red River-Idaho, USA are introduced. Among the acquired data by United State Geological Survey (USGS) and United States Department of Agriculture (USDA), the five predominant parameters in empirical bed load transport formulas including flow discharge (Q), flow velocity (V), water surface slopes (S), flow depth (d), and mean grain size (D_{50}) were used. The results of developed models then were compared with previous proposed equations and evaluated using error criteria. The models then were ranked and the best accuracy in fitting with measured data was observed for RBF and GFFN, respectively.

2 Compiled datacenter of study area

The Main Fork Red River as part of the Spokane River Basin watershed is located in the Panhandle National Forest of northern Idaho (Fig. 1(a)). The records of stream flow for 1986 to 2000 and transported sediments for 1986 to 1999 including 200 and 136 measurements of transported bed and suspended loads has been extracted and compiled from USGS and USDA. The recorded data sets as well as information on channel profile, cross section and transported material can be found in URL link of US Forest service of USDA. The variation of river cross section at the measurement station for 3 years (1995–1998) is presented in Fig. 1(b). According to previous studies [2,4,14,19]; Q, V, S, d, and D_{50} due highest influence on bed load sediment transport were selected and analyzed. These factors can be categorized into hydrological, channel geometry, geomorphological and hydraulic characteristics. By sorting the time series of recorded data, a uniform processed data sets including 102 sets of five mentioned parameters were provided and statistically analyzed (Table 1). The wide range of observed variation in selected parameters can be referred to amount of precipitations in each water year during the measurements. The used data sets due to different units using $\frac{x-x_{\min}}{x_{\max}-x_{\min}}$ were normalized within the range of [0, 1] to provide dimensionless input data which are necessary to improve the learning speed and model stability. The created data sets was then randomized into 55%, 25%, and 20% to organize training, testing and validation of ANN-based models. Therefore training procedure implies that the models are run with

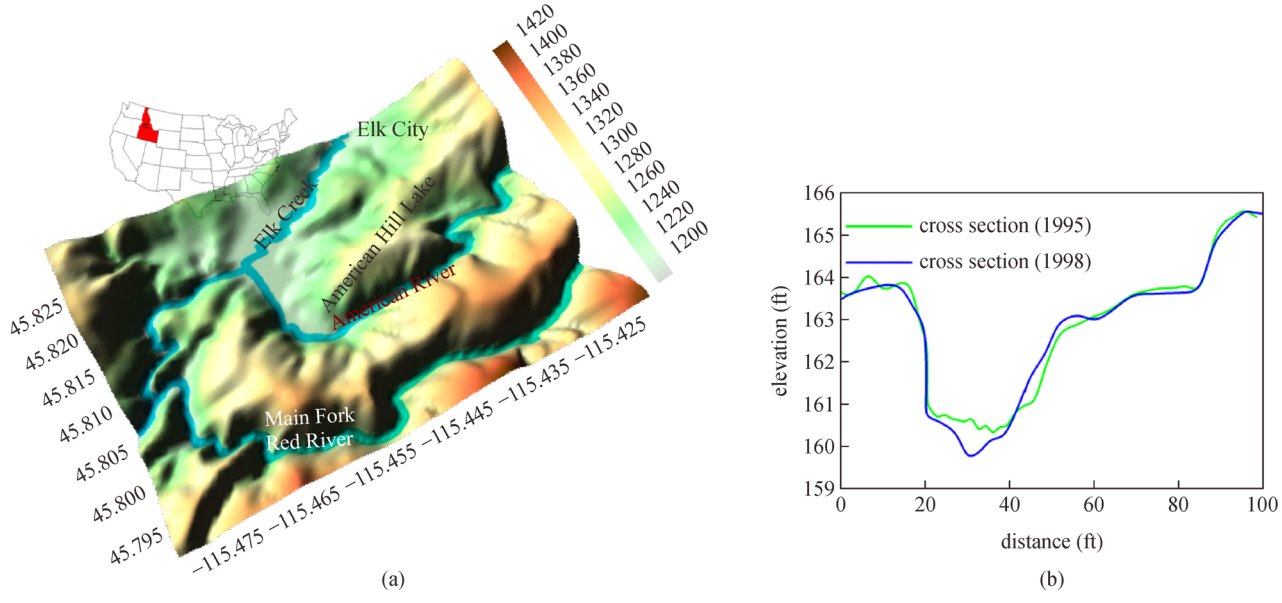

Fig. 1 Digital elevation map (DEM) of (a) studied area and (b) variation of cross section of river using USDA information for 3 years interval of the measurements.

Table 1 Statistical description of input parameters for prediction of sediment loads

variable	mean	standard error	standard deviation	variation range
Q (ft^3/s)	154.7	11.8	117.9	9.88–645.0
S (ft/ft)	0.002	0.0001	0.0011	0.0058–0.0003
V (ft/s)	2.851	0.095	0.949	0.56–5.01
d (ft)	1.428	0.049	0.490	0.34–3.13
D_{50} (mm)	8.852	0.929	9.38	0.40–45.50

Note: units of data: US measurement system.

similar data sets which provide more facilities in evaluating and comparison of the performance criteria.

3 ANN predictive models

ANNs as an information processing paradigm is one of the main tools used in machine learning and connectionist computing systems which aim to simulate brain structure and mimic the way that humans learn. Typically, the embedded neurons (processing units) in ANNs are arranged in a series of layers including input, hidden, and output. Input (x_i), weights ($w_{i,j}$), bias (b_i), activation function (f_{act}), and output ($O_{i,j}$) are the components of an artificial neuron. In human brain, the received inputs in each level of neurons provide insight and then the information gets passed on to the next, more senior level. This is precisely the mechanism that ANNs as fully connected and weighted network layers try to replicate. The input layer receives various forms of data which network aims to process or learn about. The data according to weighted connections then are passed through one or more hidden layers. The hidden layer is responsible to data processing and then transfer to output unit. The feed forward ANNs are those type of neural nets in which information travels in only one direction from input to output. The MLPs as typical feed forward networks are trained slowly but easy to use and can approximate any input/output map. The jth network output (net_j) contain a set of adaptive weight of $w_{i,j}$ and is expressed as:

$$net_j = \sum\nolimits_{x \in X} (O_i \cdot w_{i,j}) + b_i, \tag{1}$$

where b_i denotes the bias as a type of connection weight with a constant nonzero value which is set up into the all neurons in the back-propagation and transfer functions except for the input layer. The bias is much like a weight but with a constant input of 1, while the transfer function shifts the summed signals received from this neuron. The assigned activation state then by use of threshold value (θ_j) transforms the net_j from initial activation state $a_j(t-1)$ into new $a_j(t)$ by:

$$a_j(t) = f_{\text{act}}(net_j(t), a_j(t-1), \theta_j). \tag{2}$$

The output value O_j of the neuron j is then calculated from corresponding activation state a_j as:

$$f_{\text{out}}(a_j) = 0. \tag{3}$$

The error of each sample (E_{p}) and root mean square error ($RMSE$) between the input (x) and the actual output (y) for the kth output neuron is defined by:

$$E_{\text{p}} = \frac{1}{2} \sum\nolimits_{k \in O} (x_k - y_k)^2, \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{k \in O} (x_k - y_k)^2}{|O|}}. \tag{5}$$

The optimum weights then can be found using and updating procedure for $(n+1)$th pattern subjected to:

$$\Delta w_{i,k} = -\eta \frac{\partial Err(W)}{\partial w_{i,k}}, \tag{6}$$

$$w_{i,k}(n+1) = w_{i,k}(n) + \nabla w_{i,k}(n), \tag{7}$$

where η is the learning rate.

The GFFN due to ability in jumping over one or more layers showed more facilities in both selecting the optimum topology and increasing the computational potency [38–40]. Moreover, it was found that the in the same number of neurons the performance of GFFN is more efficient than MLPs [40]. The GFFN classifier uses a generalized shunting neuron (GSN) model which allows neurons to operate as adaptive nonlinear filters [38,40]. To produce the output in GFFN, all input is summed and passed through an activation function similar to a perceptron neuron (Eq. (8)).

$$x_j = \frac{b_j + f(\sum_i w_{ji} I_j + w_{jo})}{a_j + g(\sum_i c_{ji} I_i + c_{jo})}, \tag{8}$$

where x_j: output (activity) of the jth neuron; I_j and I_i: inputs to the ith and jth neurons; a_j: passive decay rate of the neuron (positive constant); w_{ji} and c_{ji}: connection weight from the ith inputs to the jth neuron; a_j and b_j: constant biases; g and f: activation functions.

The RBFs is a hybrid network with three layers (input layer, a hidden layer, and a linear output layer) in which uses nonlinear Gaussian activation transfer functions rather than the standard sigmoidal functions and tend to learn much faster than MLPs.

$$h_j(x) = \exp^{-\frac{(x-c_j)^{\mathrm{T}}(x-c_j)}{r^2}}, \quad (9)$$

where c_j denotes the center, r is the width and $(\cdot)^{\mathrm{T}}$ expresses the transpose of the argument.

Each input x_i at each hidden neuron j is weighted by w^h as:

$$S_j = [x_1 w^h_{1,j}, x_2 w^h_{2,j}, \ldots, x_n w^h_{n,j}, \ldots, x_N w^h_{N,j}], \quad (10)$$

where x_n is the nth input; $w^h{}_{n,j}$ is the input weight between input n and hidden neuron j. Accordingly the output of hidden neuron ($\phi_j(s_j)$) and the network output (O_m) is calculated by:

$$\phi_j(S_j) = \exp\left(\frac{||S_j - c_j||^2}{\sigma_j}\right), \quad (11)$$

$$O_m = \sum_{i=1}^{J} \phi_j(S_j) w^o_{j,m} + w^o_{0,m}, \quad (12)$$

where the activation function $\varphi_j(\cdot)$ for hidden neuron j is normally chosen as Gaussian function; c_j and σ_j denote the center and width of hidden neuron j, respectively. $w^o{}_{j,m}$ expresses the output weight between hidden neuron j and output neuron m and $w^o{}_{0,m}$ is the bias weight of output neuron m.

The approximating function expresses a sum of N radial basis functions associated with a different center x_i and appropriate weighted coefficients. The approximating function is linear and thus can be estimated using the matrix methods of linear least squares.

Therefore, producing optimum ANNs structure due to prevent the over-fitting problem and dependency to internal characteristics (e.g., training algorithm, number of neurons, learning rate, activation function, architecture, regularization) is a difficult but important task. Despite of different proposed relations for number of hidden neurons [41], no unique method for network configuration is accepted [42,43]. In this paper the optimum models and corresponding adjusted internal characteristics were found through an iterative procedure integrated with a developed code based on constructive technique (Fig. 2).

4 Results of modeling process

The data were randomized as 55%, 25%, and 20% for training, testing and validation. In optimizing process not only different training algorithms and activation functions but also various arrangements of neurons in hidden layers were examined. To find the optimum model, seven training

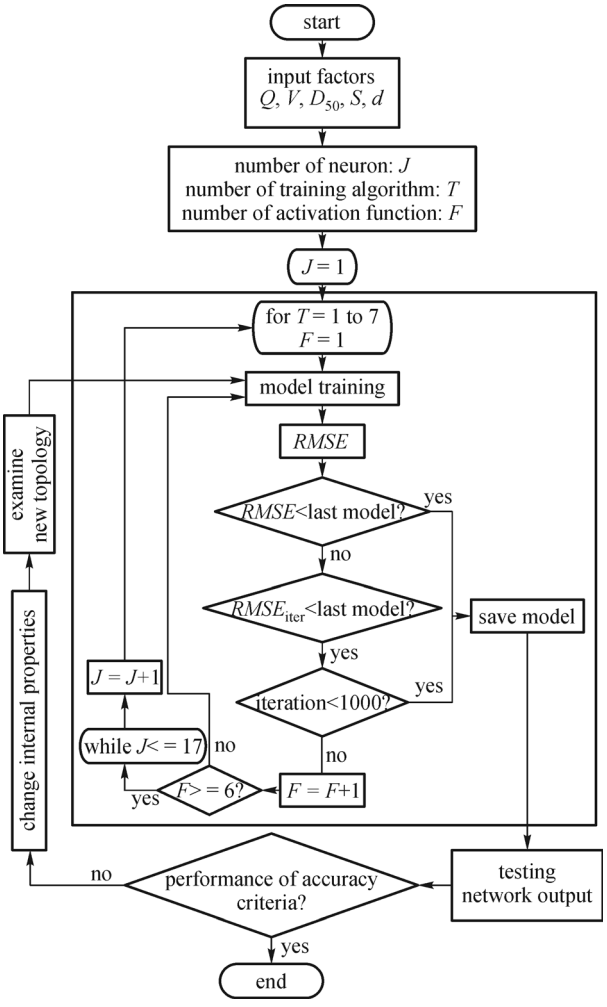

Fig. 2 The proposed block procedure to find the optimum architecture topology.

algorithms (quick propagation, QP; conjugate gradient descent, CGD; step, S; momentum, MO; quasi-Newton, QN; limited memory quasi-Newton, LMQN; Levenberg-Marquardt, LM) and six activation function (logistic, *Log*; hyperbolic tangent, *HyT*; linear, *Lin*; softmax axon, *SoA*; bias axon, *BiA*; squash, *Sq*) were examined. The sum of squares and cross-entropy were utilized as output error function, respectively.

To decrease the network variables, the learning rate of used algorithms and the domain of changes in step size for hidden layers was set 0.7 and [1.0–0.001], respectively. The process is terminated using two different termination criteria. The first priority is to satisfy the target root mean square error (*RMSE*) and if not achieved the number of iterations will use. The maximum iteration number in this study was set for 1000 which means that if cannot attain the desired *RMSE* then the minimum observed network error of each examined structure among 1000 repeated epochs will be considered. With this condition even the structures which couldn't show the target error but provide

lower *RMSE* than previous examined model were also captured. According to embedded loop in defined procedure (Fig. 2) numerous structures even with similar architecture but different internal characteristics were generated. As a description of executed effort (Fig. 2), the process with one neuron in hidden layer (J) is begun with one of the training algorithms (T) and activation functions (F). In the loop, T and F are fixed and only J is varying and while $t < 7$, the loop for 6 different F (*Log*, *Hyt*, *Lin*, *SoA*, *BiA*, *Sq*) is repeated. Therefore the process is repeated up to $F = 6$ and thus the condition $F \geqslant 6$ will break the internal iteration and goes to changing the J and the process with new number of neurons again is repeated. During this process, by achieving the minimum *RMSE*, the loop is terminated and the obtained number of neurons needs to be managed in appropriate layers to find the optimum network structure (e.g., topology characteristics, model regularization and new topology as presented in Fig. 2). The maximum number of neurons as a user option can be set by after neuron increment in which in this study using while command was managed for 17. Obviously, by changing the condition different number of neurons can be checked. In this step different arrangement of obtained neurons subjected to learning rate from 1 to 0.01 for each generated structure is examined and those topologies with the minimum observed *RMSE* as candidate of optimum models were considered. To have a similar condition in model evaluation, all models were trained subjected to initial randomized data sets.

Accordingly, the calculated minimum *RMSE* and the maximum network coefficient of determination (R^2) of each structure after 3 runs were considered. Structures refer to numerous generated topologies which have been obtained from different neurons arrangement in layers. Using described procedure more than 1200 different models even in similar structure but different internal characteristics were examined. Diverse training algorithms and activation functions under different step sizes for learning rate were used to avoid the over fitting problem and escape from local minima. For example replacing the gradient descent by the momentum optimizer with step size 0.001 minimizes the chances that it gets stuck in a local minimum. Furthermore, the results of 3 runs for each structure and investigation of error improvements were also monitored to be ensured that the examined models were not over fitted and trapped in local minima. The results of selected optimum models as well as example of tested network errors against the number of neurons for MLP, GFFN, and RBF subjected to different training algorithms are given in Table 2 and Fig. 3, respectively.

5 Discussion and validation

The confusion matrix is a specific table layout to analyze and visualize the performance of an algorithm in classification purposes of ANN models [44]. The rows (i) and columns (j) of this square matrix represent the used data sets in predicted and actual classes to display if the system is confusing two classes [45]. The value in the (i,j) position is the number of records for the target column in the ith category and corresponding network output in the jth category. The perfect classification would have 0 everywhere except on the diagonal entries. The confusion matrixes of all obtained optimum models with 10 rows and target value step of 1.04 were calculated. As an example, the results of RBF (Table 3) showed that the system have trouble in distinguishing the real values in the range of 1.08–2.11, 7.29–8.33, and 9.37–10.4, but can make proper distinction between other defined classes. The rows with 0 means that there is no input in this range and also the network perform appropriate prediction which there is no data to confuse with other categories. Similar interpretation for GFFN and MLP was also carried out. Then correct classification rate (*CCR*) as a qualitative characteristic in classification tasks (Eq. (13)) is employed to show the portion (in %) of data correctly recognized within the general data set [46]. The *CCR* and classification error (*CE*) were calculated for all optimized models in which for RBF can be found in Table 3. Furthermore, a comparison between the averages of *CCR* for all models were carried out and reflected in Table 4, respectively.

$$CCR = \frac{correctly\ predicted\ class}{total\ testing\ class} \longrightarrow$$

$$CE = 1 - CCR. \tag{13}$$

In the literatures, several empirical equations subjected to different parameters for determining the bed load have been proposed. In this paper the developed equations by [47–56], were investigated. The dimensionless bed load rate was calculated by the empirical equations in the spreadsheet software with the replacement of the flow and geometric characteristics of the river and the characteristics

Table 2 Results of implemented training algorithms to assess the optimized ANN based model

ANN type	training algorithm	min *RMSE*	no. of neuron	corresponding topology	layer activation transfer function		R^2 for randomized data sets		
					hidden	output	train	test	validate
MLP	QN	0.289	9	5-5-4-1	*Log*	*HyT*	0.91	0.92	0.93
GFFN	MO	0. 201	9	5-4-5-1	*HyT*	*HyT*	0.94	0.96	0.97
RBF	LM	0.191	7	5-7-1	*Log*	*Log*	0.96	0.97	0.98

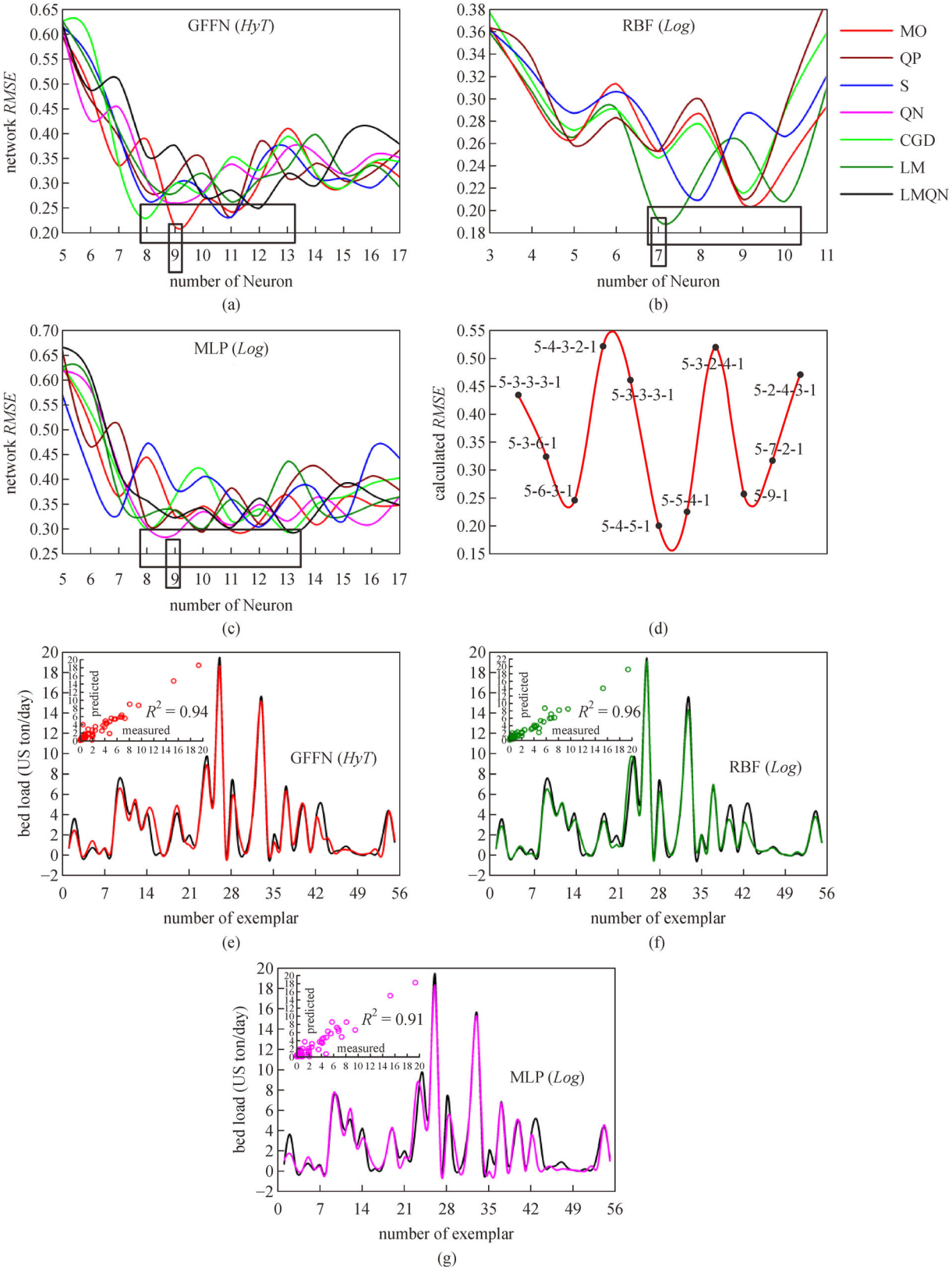

Fig. 3 Variation of calculated network *RMSE* for different training algorithms based on the number of neurons (the range of neurons for minimum observed error as well as used activation functions are given in rectangles and parentheses, respectively). (a) GFFN; (b) RBF; (c) MLP; (d) example of some tested structure to find the optimum GFFN model subjected to MO training algorithm and *HyT* activation function; (e) performance of optimized models corresponding to used learning rule in training stage in GFFN, (f) Performance of optimized models corresponding to used learning rule in training stage in RBF; (g) performance of optimized models corresponding to used learning rule in training stage in MLP.

Table 3 Confusion matrix of optimum RBF model

target output	network output RBF (validation data sets)										calculated results			
	0.04–1.08	1.08–2.11	2.11–3.15	3.15–4.19	4.19–5.22	5.22–6.26	6.26–7.29	7.29–8.33	8.33–9.37	9.37–10.4	total	true	false	*CR* (%)
0.04–1.08	9	0	0	0	0	0	0	0	0	0	9	9	0	100
1.08–2.11	0	4	1	0	0	0	0	0	0	0	5	3	2	75
2.11–3.15	0	0	0	0	0	0	0	0	0	0	0	0	0	–
3.15–4.19	0	0	0	0	0	0	0	0	0	0	0	0	0	–
4.19–5.22	0	0	0	0	1	0	0	0	0	0	1	1	0	100
5.22–6.26	0	0	0	0	0	0	0	0	0	0	0	0	0	–
6.26–7.29	0	0	0	0	0	0	0	0	0	0	0	0	0	–
7.29–8.33	0	0	0	0	0	0	1	1	0	0	2	1	1	50
8.33–9.37	0	0	0	0	0	0	0	0	0	0	0	0	0	–
9.37–10.4	0	0	0	0	0	0	0	0	1	1	2	1	1	50
note	9	3	2	0	1	0	1	1	1	1	19	16	3	84

target output	network output RBF (test data sets)										calculated results			
	0.01–1.89	1.89–3.77	3.77–5.65	5.65–7.53	7.53–9.41	9.41–11.29	11.29–13.17	13.17–15.05	15.05–16.93	>16.93	total	true	false	CR (%)
0.01–1.89	12	1	0	0	0	0	0	0	0	0	13	12	1	92.3
1.89–3.77	0	4	2	0	0	0	0	0	0	0	6	4	2	66.7
3.77–5.65	0	0	1	1	0	0	0	0	0	0	2	1	1	50
5.65–7.53	0	0	0	1	1	0	0	0	0	0	2	1	1	50
7.53–9.41	0	0	0	0	1	0	0	0	0	0	1	1	0	100
9.41–11.29	0	0	0	0	0	0	0	0	0	0	0	0	0	–
11.29–13.17	0	0	0	0	0	0	0	0	0	0	0	0	0	–
13.17–15.05	0	0	0	0	0	0	0	0	0	0	0	0	0	–
15.05–16.93	0	0	0	0	0	0	0	0	0	0	0	0	0	–
>16.93	0	0	0	0	0	0	0	0	0	1	1	1	0	100
note	12	5	3	2	2	0	0	0	0	1	25	20	5	80

Table 4 Comparison of *CCR* and classification error of optimized models for validation and test data sets

model	*CCR* (%)		*CE* (%)	
	test	validate	test	validate
RBF	80	84	20	16
GFFN	80	79	20	21
MLP	65	68	35	32

of the river bed materials. This replacement in empirical relationships in spreadsheet software can easily be done. The performance of these equations can be compared with ANN models using validation data sets which randomly selected and didn't previously have been fed to network in training and testing stages [42]. Due to specified conditions in development of each equation as well as inherent approved limitation of implemented statistical techniques not only high scattering in produced outputs can be observed but also some of them even couldn't provide estimation for all fed validation sets (Table 5). Therefore the comparisons between those equations without misestimated outputs were conducted and reflected in Fig. 4. The imperfect predictability of empirical equations showed that in similar condition the accuracy and fitting to actual data are not uniform and thus appropriate results cannot be expected. As presented in Table 5 the coefficient of determination (R^2) of estimated bed loads using empirical predictors vary from 0.08 [48] to 0.21 [56], whereas in ANN models more consistency and closer values from 0.93 in MLP to 0.98 for RBF model were obtained. It was observed that proposed equation by Ref. [54] provide under estimate values whereas other formulas exhibited over estimate in predicted values.

Despite of approved applicability of statistical error criteria in evaluating the performance of predictive models, the associated interpretation in assessment of low-volume data across multiple items should be considered [40,43]. Here, the accuracy performance of models were evaluated suing the mean absolute percentage error (*MAPE*), variance account for (*VAF*), *RMSE*, mean squared deviation (*MSD*), mean absolute deviation (*MAD*), R^2, calculated residuals (*CR*) and absolute error (*AE*) criteria. *MAPE* as the most common used index of forecast error represents the accuracy using the average of the unsigned percentage error. *MAD* is the average distance between each data point and the mean. This statistic gives an idea about the variability in a data set for size of the error in units and calculates the average of the unsigned errors. *MSD* as a nonnegative index measures the quality and accuracy of a predictor in fitted values and also express a risk function, corresponding to the expected values of the squared error loss. *AE* as an indicator of physical error and the uncertainty in a measurement expresses the quality of model and defines the deviation between predicted and measured values. *CR* is the differences between the measured and predicted values. In performance analyzing, higher values of *VAF* and R^2 as well as smaller and lower values of *MAPE*, *CR*, *AE*, *MAD*, *MSD*, and *RMSE* is of interested. According to reflected results in Table 6 and Fig. 5, in the studied area the equation proposed by

Table 5 Some of the tested bed load empirical equations in this study

researcher(s)	equation	misestimated data	R^2
Nielsen [47]	$\varphi_b = [12(\theta-0.05)(\theta^{0.5})]$	0	0.11
Ackers and White [48]	$\varphi_b = [\theta^{1.25}Log(16.5\theta/S)]$	0	0.10
Wong and Parker [49]	$\varphi_b = 3.97(\theta-0.0495)^{1.5}$	7/19	–
Wilson [50]	$\varphi_b = 12(\theta-\theta_c)^{1.5}$	7/19	–
Paintal [51]	$\varphi_b = 6.56\times10^{18}\theta^{16}$	0	0.14
Madsen [52]	$\varphi_b = [k(\theta-\theta_c)(\theta^{0.5}-0.7\theta_c^{0.5})]$	0	0.11
Meyer-Peiter and Muler [53]	$\varphi_b = 8(f\theta-\theta_c)^{1.5}$	12/19	–
Rottner [54]	$\varphi_b = \left\{\left(\frac{V}{(g(G_S-1)D_{50})^{0.5}}\right)\left[0.667\left(\frac{D_{50}}{d}\right)^{0.67}+0.14\right]-0.778\left(\frac{D_{50}}{d}\right)^{0.67}\right\}^3$	0	0.13
Van-Rijn [55]	$\varphi_b = \left[\left(\frac{0.053}{D_{50}\left(\frac{g(G_S-1)}{\theta^2}\right)^{\left(\frac{1}{3}\right)}}\right)\left(\frac{\theta}{\theta_c}-1\right)^{2.1}\right]$	12/19	–
Kalinske [56]	$q_b = \sqrt{gdS}D_{50}f\left(\frac{\tau_C}{\tau}\right)$	0	0.21

Note: $\theta = \frac{RS}{(G_S-1)}$; $\varphi_b = \frac{q_b}{[gD_{50}^3(G_S-1)]^{\left(\frac{1}{2}\right)}}$.

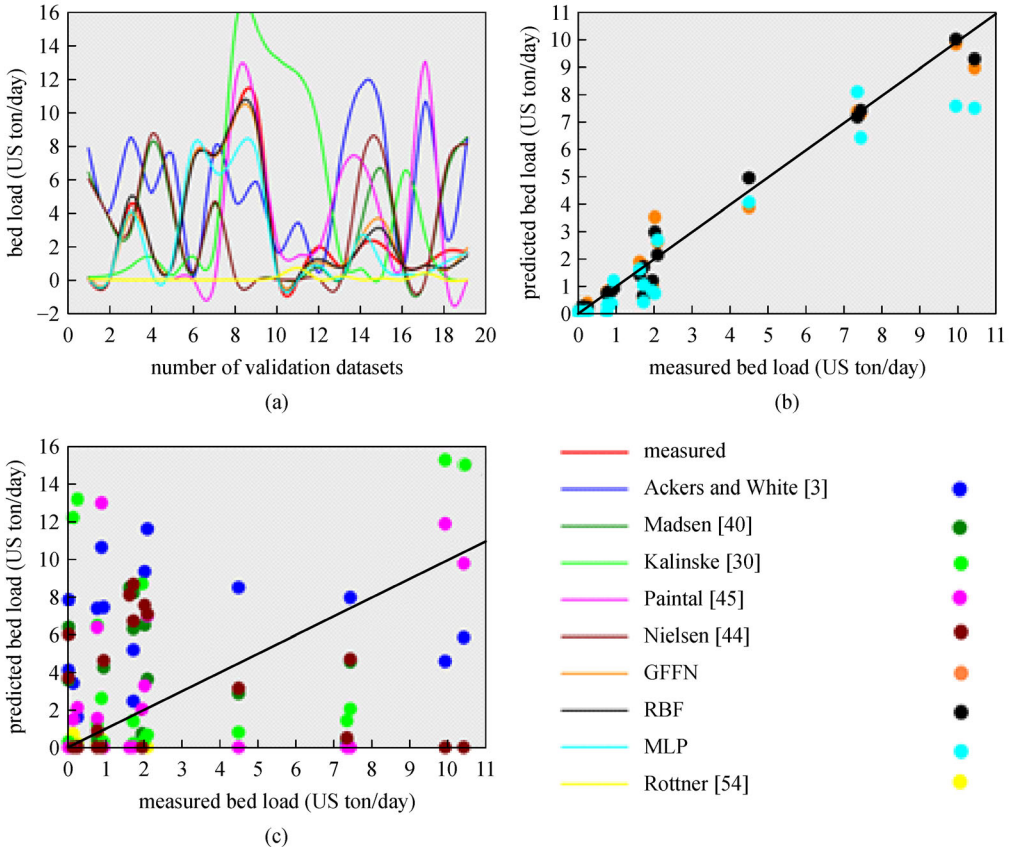

Fig. 4 (a) Comparing the predicted bed load values using ANN and empirical models; (b) scattering of predicted bed loads using ANN models regarding 1:1 line; (c) predicted bed loads using empirical equations.

Ref. [56] was ranked as the most appropriate predictors whereas Ref. [48], represented the most improper result.

The applicability of sensitivity analysis methods in determining the influence of input parameters on predicted outputs has been approved [57–59]. Using sensitivity analysis methods not only simplified but robust calibrated model from large number of parameters can be utilized but also important connections between observations and model output as well as evaluating the impacts of the uncertainties in the output can be determined [58–63]. In partial derivative algorithm (PaD) as one of the most known ANN-based sensitivity analysis technique [64,65], the contribution of inputs is analyzed using Jacobian matrix of the partial derivatives of outputs (Eq. (14)) with respect to inputs [66].

$$\text{contribution of } i\text{th variable} = \frac{SSD_i}{\sum_i SSD_i},$$

$$SSD_i = \sum_p \left(\frac{\partial o_k^p}{\partial x_i^p}\right)^2, \tag{14}$$

where $O_k{}^p$ and $x_i{}^p$ are output and input values for pattern P the SSD_i is sum of the squares of the partial derivatives, respectively.

Therefore, PaD is able to assess the sensitivity of the output against slight changes in inputs. The results of calculated influence of fed inputs on predicted output using PaD method is reflected in Fig. 6.

It can be seen that the Q and V as tow of the main hydraulic parameters are ranked as the most effective parameters whereas the d which also can be categorized into hydraulic factors is placed in lowest rank. However the values of Q and V are approximately similar and their differences are not significant. The contribution of D_{50} as one of the main important sediment characteristics in predicted bed loads is ranked in the third position. The influence of S which is a geomorphological related factor is scored in fourth place but both D_{50} and S are recognized with moderate influences more than d.

6 Conclusions

ANN models are able to reveal hidden laws of natural phenomena such as sediment transport process. In the current study, different MLP, GFFN, and RBF models for the purpose of bed load prediction in the Main Fork Red River in Idaho, USA were successfully developed and examined. The models were optimized based on the

Table 6 Results of statistical criteria to evaluate the performance of used models

Model	*MAPE*	*RMSE*	*MAD*	*MSD*	*VAF*	R^2
MLP	6.98	0.289	0.93	0.127	81.56	0.93
GFFN	5.23	0.201	0.80	0.025	96.49	0.97
RBF	5.27	0.191	0.77	0.018	95.77	0.98
Nielsen [47]	18.08	7.673	1.75	0.463	26.38	0.11
Ackers and White [48]	20.82	8.517	2.09	1.05	23.16	0.10
Paintal [51]	14.07	4.291	1.73	0.365	35.22	0.14
Madsen [52]	17.20	6.424	1.73	0.533	27.06	0.11
Rottner [54]	15.92	4.348	1.72	0.517	30.88	0.13
Kalinske [56]	11.38	3.562	1.45	0.316	54.24	0.21

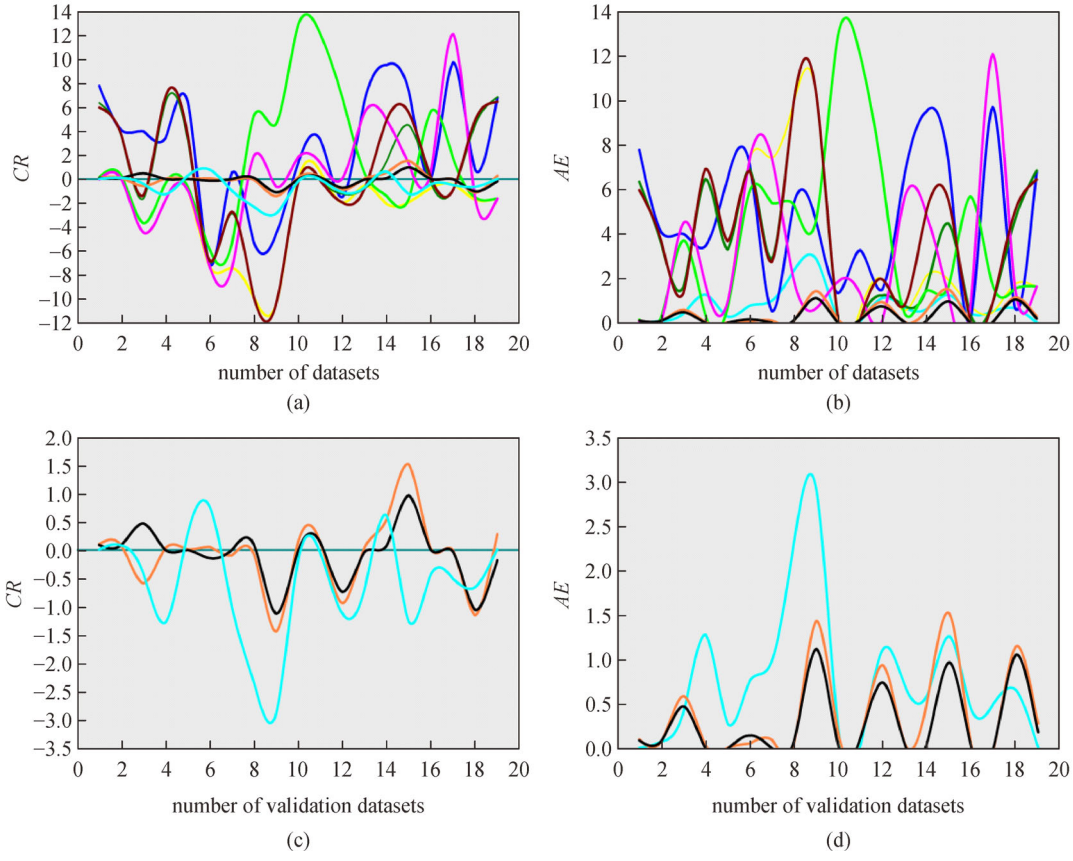

Fig. 5 (a) Comparison of CR for ANN and empirical models; (b) comparison of AE for ANN and empirical models; (c) variation of CR values based on the used data sets in ANN models; (d) variation of AE values based on the used data sets in ANN models (The used colors are similar to those defined in Fig. 4).

variability of internal characteristics through an organized iterative procedure using 102 sets of five dominant implemented hydraulic variables and sediment data (Q, V, S, d, and D_{50}). The calculated network *RMSE* and R^2 revealed that the optimized RBF (*RMSE* = 0.191) due to more compatible results between measured and predicted bed load transport slightly outperforms than GFFN. It was also observed that in similar number of neurons the GFFN (*RMSE* = 0.201) reflect much more robust results than MLP (*RMSE* = 0.289). The summary of predicted results for ANN models using validation and test data sets were presented in separated confusion matrixes which gave insight into the errors by classifier and showed the *CCR* probability. The corresponding level of *CCR* for RBF, GFFN, and MLP was 79%, 74% and 67%, respectively. Furthermore the predictability of several empirical bed load equations respect to ANN models for the study area was analyzed. In compare to ANN models which their

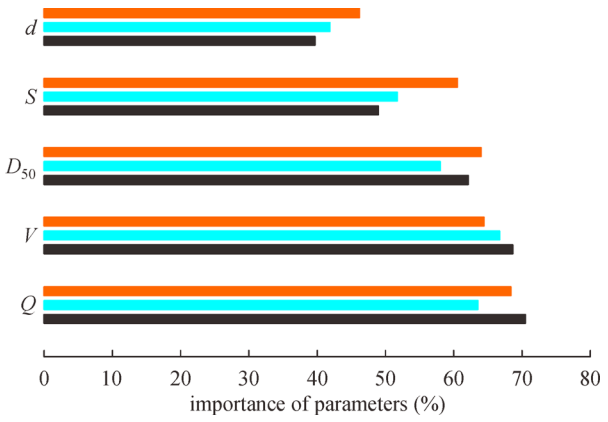

Fig. 6 Sensitivity analysis of ANN models to identify the contribution of input parameters on predicted bed loads (the used colors are similar to Fig. 4).

corresponding R^2 values varied from 0.93 to 0.98, the empirical equations using similar fed data provided diverse predicted scatters with R^2 values between 0.10 and 0.21. Comparison between the empirical formulas also demonstrated an overall unsatisfactory predictions of the bed loads for a wide range of hydrodynamic and sediment conditions. This issue can be considered as one of the main shortcomings of the empirical equations that in a same condition exhibit different accuracy scores and thus the performance of such relations were found comparatively insufficient. The accuracy performance of the ANN models and empirical equations were then assessed and ranked using *MAPE*, *RMSE*, *MAD*, *VAF*, R^2, *CR*, and *AE* error criteria. It was observed that RBF due to satisfied more criteria (*RMSE*, 0.191; *MAD*, 0.77; *MSD*, 0.018; R^2, 0.98) than GFFN (*MAPE*, 5.23; *VAF*, 96.49) demonstrated more accurate performance. Moreover, the comparative analytical graphs of variations of *CR* and *AE* showed lower values in ANN models than empirical equations in which the lowest variation was tracked in RBF. The *Q*, *V*, and *d* using the PaD method were recognized as the most and least effective factors on predicted bed loads.

Acknowledgements The authors greatly expressed their appreciate to Dr. Abbas Abbaszadeh Shahri for his expert advice and encouragement through this study.

References

1. Francalanci F, Solari L, Toffolon M. Local high-slope effects on sediment transport and fluvial bed form dynamics. Water Resources Research, 2009, 45: W05426
2. Camenen B, Larson M. A general formula for noncohesive suspended sediment transport. Journal of Coastal Research, 2008, 24(3): 615–627
3. Helton J C, Johnson J D, Sallaberry C J, Storlie C B. Survey of sampling based methods for uncertainty and sensitivity analysis. Reliability Engineering & System Safety, 2006, 91(10–11): 1175–1209
4. Pektas A O, Dogan E. Prediction of bed load via suspended sediment load using soft computing methods. Geofizika, 2015, 32(1): 27–46
5. Ma H, Nittrouer J A, Naito K, Fu X, Zhang Y, Moodie A J, Wang Y, Wu B, Parker G. The exceptional sediment load of fine-grained dispersal systems: Example of the Yellow River, China. Science Advances, 2017, 3(5): e1603114
6. Goodwin P. Analytical solutions for estimating effective discharge. Journal of Hydraulic Engineering, 2004, 130(8): 729–738
7. Lovejoy S B, Lee J G, Randhir T O, Engel B A. Research needs for water quality management in the 21st century: A spatial decision support system. Journal of Soil and Conservation, 1997, 52(1): 18–22
8. Cigizoglu H K. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. Advances in Water Resources, 2004, 27(2): 185–195
9. Kisi O, Shiri J. River suspended sediment estimation by climatic variables implication: Comparative study among soft computing techniques. Computers & Geosciences, 2012, 43: 73–82
10. Heng S, Suetsugi T. Using artificial neural network to estimate sediment load in ungauged catchments of the Tonle Sap River Basin, Cambodia. Journal of Water Resource and Protection, 2013, 5(2): 111–123
11. Yang C T, Wan S. Comparisons of selected bed-material load formulas. Journal of Hydraulic Engineering, 1991, 117(8): 973–989
12. Yang S Q. Formula for sediment transport in rivers, estuaries, and coastal waters. Journal of Hydraulic Engineering, 2005, 131(11): 968–979
13. Khorram S, Ergil M. Most influential parameters for the bed-load sediment flux equations used in alluvial rivers. Journal of the American Water Resources Association, 2010, 46(6): 1065–1090
14. Sirdari Z Z, Ab Ghani A, Sirdari N Z. Bedload transport predictions based on field measurement data by combination of artificial neural network and genetic programming. Pollution, 2015, 1(1): 85–94
15. Kişi Ö. River suspended sediment concentration modeling using a neural differential evolution approach. Journal of Hydrology (Amsterdam), 2010, 389(1–2): 227–235
16. Cao M, Alkayem N F, Pan L, Novák D. Advanced Methods in Neural Networks-based Sensitivity Analysis with Their Applications in Civil Engineering. INTECH Press, 2016
17. Melesse M A, Ahmad S, McClain M E, Wang X, Lim Y H. Suspended sediment load prediction of river systems: An artificial neural network approach. Agricultural Water Management, 2011, 98(5): 855–866
18. Ham D G, Church M. Bed material transport estimated from channel morphodynamics: Chilliwack River, British Columbia. Earth Surface Processes and Landforms, 2000, 25(10): 1123–1142
19. Song T, Chiew Y M, Chin C O. Effects of bed-load movement on flow friction factor. Journal of Hydraulic Engineering, 1998, 124(2): 165–175
20. Recking A. An experimental study of grain sorting effects on bedload. Dissertation for the Doctoral Degree. Lyon: Institut National Des Sciences Appliquess De Lyon, 2006
21. Parker G. Lateral bed load transport on side slopes. Journal of the

Hydraulics Division, 1984, 110(2): 197–199
22. Sekine M, Parker G. Bedload transport on transverse slopes. Journal of Hydraulic Engineering, 1992, 118(4): 513–535
23. Damgaard J S, Whitehouse R J S, Soulsby R S. Bed load sediment transport on steep longitudinal slopes. Journal of Hydraulic Engineering, 1997, 123(12): 1130–1138
24. Yadav S M, Samtani B K. Bed load equation evaluation based on alluvial river data. India. KSCE Journal of Civil Engineering, 2008, 12(6): 427–433
25. Hamdia K M, Arafa M, Alqedra M. Structural damage assessment criteria for reinforced concrete buildings by using a Fuzzy Analytic Hierarchy process. Underground Space, 2018, 3(3): 243–249
26. Hamdia K H, Lahmer T, Nguyen-Thoi T, Rabczuk T. Predicting the fracture toughness of PNCs: A stochastic approach based on ANN and ANFIS. Computational Materials Science, 2015, 102: 304–313
27. Bouzeria H, Ghenim A N, Khanchoul K. Using artificial neural network (ANN) for prediction of sediment loads, application to the Mellah catchment, northeast Algeria. Journal of Water and Land Development, 2017, 33(1): 47–55
28. Cigizoglu H K. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. Advances in Water Resources, 2004, 27(2): 185–195
29. Chang C K, Azamathulla H M, Zakaria N A, Ghani A A. Appraisal of soft computing techniques in prediction of total bed material load in tropical rivers. Journal of Earth System Science, 2012, 121(1): 125–133
30. Kakaei Lafdani E K, Moghaddam Nia A M, Ahmadi A. Daily suspended sediment load prediction using artificial neural networks and support vector machines. Journal of Hydrology (Amsterdam), 2013, 478: 50–62
31. Khanchoul K, Tourki M, Le Bissonnais Y. Assessment of the artificial neural networks to geomorphic modelling of sediment yield for ungauged catchments, Algeria. Journal of Urban and Environmental Engineering, 2015, 8(2): 175–185
32. Kisi Ö. Constructing neural network sediment estimation models using a data-driven algorithm. Mathematics and Computers in Simulation, 2008, 79(1): 94–103
33. Kumar B. Neural network prediction of bed material load transport. Hydrological Sciences Journal, 2012, 57(5): 956–966
34. Mustafa M R, Isa M H. Comparative study of MLP and RBF neural networks for estimation of suspended sediments in Pari River, Perak. Research Journal of Applied Sciences, Engineering and Technology, 2014, 7(18): 3837–3841
35. Sasal M, Kashyap S, Rennie C D, Nistor I. Artificial neural network for bedload estimation in alluvial rivers. Journal of Hydraulic Research, 2009, 47(2): 223–232
36. Bhattacharya B, Price R K, Solomatine D P. Machine learning approach to modeling sediment transport. Journal of Hydraulic Engineering, 2007, 133(4): 440–450
37. Kisi O, Dailr A H, Çimen M, Shiri J. Suspended sediment modeling using genetic programming and soft computing techniques. Journal of Hydrology (Amsterdam), 2012, 450–451: 48–58
38. Arulampalam G, Bouzerdoum A. A generalized feed forward neural network architecture for classification and regression. Neural Networks, 2003, 16(5–6): 561–568
39. Worden K, Wong C X, Parlitz U, Hornstein A, Engster D, Tjahjowidodo T, Al-Bender A, Rizos D D, Fassois S D. Identification of pre-sliding and sliding friction dynamics: Grey box and black-box models. Mechanical Systems and Signal Processing, 2007, 21(1): 514–534
40. Abbaszadeh Shahri A, Larsson S, Johansson F. CPT-SPT correlations using artificial neural network approach-A case study in Sweden. Electronic Journal of Geotechnical Engineering, 2015, 20: 13439–13460
41. Sheela K G, Deepa S N. Review on methods to fix number of hidden neurons in neural networks. Mathematical Problems in Engineering, 2013, 2013: 425740
42. Ghaderi A, Abbaszadeh Shahri A, Larsson S. An artificial neural network based model to predict spatial soil type distribution using piezo cone penetration test data (CPTu). Bulletin of Engineering Geology and the Environment, 2018, 78: 4579–4588
43. Abbaszadeh Shahri A. An optimized artificial neural network structure to predict clay sensitivity in a high landslide prone area using piezo cone penetration test (CPTu) data—A case study in southwest of Sweden. Geotechnical and Geological Engineering, 2016, 34(2): 745–758
44. Stehman S V. Selecting and interpreting measures of thematic classification accuracy. Remote Sensing of Environment, 1997, 62(1): 77–89
45. Powers D M W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. Journal of Machine Learning Technologies, 2011, 2(1): 37–63
46. Dreižienė L, Dučinskas K, Paulionienė L. Correct classification rates in multi-category discriminant analysis of spatial Gaussian data. Open Journal of Statistics, 2015, 5(1): 21–26
47. Nielsen P. Combined convection-diffusion modelling of sediment entrainment. In: Proceedings of the 23rd International Conference on Coastal Engineering, Venice. New York: ASCE, 1992, 1–14
48. Ackers P, White W R. Sediment transport: New approach and analysis. Journal of the Hydraulics Division, 1973, 99(HY11): 2041–2060
49. Wong M, Parker G. Re analysis and correction of bedload relation of Meyer-Peter and Muller using their own database. Journal of Hydraulic Engineering, 2006, 132(11): 1159–1168
50. Wilson K C. Bedload transport at high shear stresses. Journal of the Hydraulics Division, 1966, 92(6): 49–59
51. Paintal A. Concept of critical shear stress in loose boundary open channels. Journal of Hydraulic Research, 1971, 9(1): 91–113
52. Madsen O S. Mechanics of cohesionless sediment transport in coastal waters. In: Kraus N C, Gingrich K J, Krieel D L, eds. Coastal sediment. New York: ASCE, 1991, 15–27
53. Meyer-Peter E, Muller R. Formulas for bed load transport. In: Proceedings of the Second Meeting of International Association for Hydraulic Research. Stockholm: IAHR, 1948, 39–64
54. Rottner J. A formula for bed load transportation. Houille Blanche, 1959, 14(3): 285–307
55. Van Rijn L C. Sediment transport, part I: Bed load transport. Journal of Hydraulic Engineering, 1984, 110(10): 1431–1456
56. Kalinske A A. Movement of sediment as bed-load in rivers. Transactions American Geophysical Union, 1947, 28(4): 615–620
57. Cacuci D G, Ionescu-Bujor M, Navon M. Sensitivity and Uncertainty Analysis: Applications to large-Scale Systems. Boca

Raton, FL: Chapman & Hall, 2005
58. Saltelli A. Sensitivity analysis for importance assessment. Risk Analysis, 2002, 22(3): 579–590
59. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. Global Sensitivity Analysis: The Primer. Chichester: John Wiley & Sons, 2008
60. Bahremand A, De Smedt F. Distributed hydrological modeling and sensitivity analysis in Torysa watershed, Slovakia. Water Resources Management, 2008, 22(3): 393–408
61. Wang W, Jones P, Partridge D. Assessing the impact of input features in a feedforward neural network. Neural Computing & Applications, 2000, 9(2): 101–112
62. Hamdia K H, Silani M, Zhuang X, He P, Rabczuk T. Stochastic analysis of the fracture toughness of polymeric nanoparticle composites using polynomial chaos expansions. International Journal of Fracture, 2017, 206(2): 215–227
63. Vu-Bac N, Lahmer T, Zhuang X, Nguyen-Thoi T, Rabczuk T. A software framework for probabilistic sensitivity analysis for computationally expensive models. Advances in Engineering Software, 2016, 100: 19–31
64. Tchaban T, Taylor M J, Griffin A. Establishing impacts of the inputs in a feedforward neural network. Neural Computing & Applications, 1998, 7(4): 309–317
65. Dimopoulos Y, Bourret P, Lek S. Use of some sensitivity criteria for choosing networks with good generalization ability. Neural Processing Letters, 1995, 2(6): 1–4
66. Chiang W Y K, Zhang D, Zhou L. Predicting and explaining patronage behavior toward web and traditional stores using neural networks: A comparative analysis with logistic regression. Decision Support Systems, 2006, 41(2): 514–531