Engineering with Computers
DOI 10.1007/s00366-016-0446-1

ORIGINAL ARTICLE

Extreme learning machine assessment for estimating sediment transport in open channels

Isa Ebtehaj[1] · Hossein Bonakdari[1] · Shahaboddin Shamshirband[2]

Received: 31 December 2015 / Accepted: 14 March 2016

Abstract The minimum velocity required to prevent sediment deposition in open channels is examined in this study. The parameters affecting transport are first determined and then categorized into different dimensionless groups, including "movement," "transport," "sediment," "transport mode," and "flow resistance." Six different models are presented to identify the effect of each of these parameters. The feed-forward neural network (FFNN) is used to predict the densimetric Froude number (*Fr*) and the extreme learning machine (ELM) algorithm is utilized to train it. The results of this algorithm are compared with back propagation (BP), genetic programming (GP) and existing sediment transport equations. The results indicate that FFNN-ELM produced better results than FNN-BP, GP and existing sediment transport methods in both training (RMSE = 0.26 and MARE = 0.052) and testing (RMSE = 0.121 and MARE = 0.023). Moreover, the performance of FFNN-ELM is examined for different pipe diameters.

Keywords Bed load · Extreme learning machines (ELM) · Limit of deposition · Sediment transport · Storm water

Notations

A	Cross-sectional area of the flow
a_i	Learning factors of the hidden nodes (Eq. 6)
B_j	Bias of the jth neuron of the hidden layer (Eq. 8)
b_i	Learning factors of the hidden nodes (Eq. 6)
C_V	Volumetric sediment concentration
d	Median diameter of particle size
D	Pipe diameter
D_{gr}	Dimensionless particle number
Fr	Densimetric Froude number
$G(a_i, b_i, x)$	Is the output of the ith hidden node for input x (Eq. 6)
g	Gravitational acceleration
$g(.)$	Nonlinear activation piecewise continuous function (Eq. 8)
H_{ik}	Activation matrix of the jth neuron of the hidden layer for the kth training case (Eq. 8)
H'	is the Moore-Penrose inverse of Matrix H (Eq. 14)
R	Hydraulic radius
s	Specific gravity of sediment ($=\rho_s/\rho$)
S_0	Pipe slope
T	Axis representing the target values for the training cases (Eqs. 9, 14)
V	Flow velocity
V_t	Velocity required for the sediment's incipient motion (Eq. 2)
W_{ji}	Weight of the ith input neuron and the jth neuron of the hidden layer (Eq. 8)
X_{ik}	Input of the input neuron for the kth training case (Eq. 8)
y	Flow depth
β	Indicates the weight between the output layer neurons and hidden layer neurons (Eqs. 9, 14)
β_i	Weight between the ith hidden node and the output node (Eq. 6)
λ_c	Clear water friction factor

✉ Hossein Bonakdari
bonakdari@yahoo.com

[1] Department of Civil Engineering, Razi University, Kermanshah, Iran

[2] Department of Computer Science, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia

Published online: 29 March 2016

λ_s	Overall friction factor with sediment
ρ	Density of water
ρ_s	Density of sediment

1 Introduction

Among the many important issues that significantly affect the optimum design of open channel flow are problems including difficulties stemming from sediment deposition. Storm water often carries sediment along its path in terms of flow velocity and the path the flow takes. In case the input flow discharge is unable to transport sediment entering the channel with a fixed gradient, the sediment will deposit on the channel bed. Sediment deposition leads to reduced flow cross section area in addition to increased bed roughness and consequently, lower transport capacity. Conversely, maximum discharge occurs in flow passing through the channel, which washes away sediment deposited during low discharge condition. The channel must be designed in such manner that a balance is reached between the minimum and maximum velocities to prevent prolonged deposition of solid matter on the channel bed, keeping in view economic concerns. In case sediment remains on the channel bed undisturbed for a long time, it solidifies and gains cement properties.

The simplest way to prevent sediment deposition is to make use of minimum shear stress and velocity. However, this method does not consider specifications of the channel's hydraulic conditions and it presents different results under different conditions. It may consequently lead to underdesign, which causes sediment deposition on the channel bed, or overdesign, which leads to uneconomic planning [1]. A comprehensive review of minimum shear stress and velocity under different flow conditions was presented by Vongvisessomjai et al. [2]. Several researchers have studied the parameters influencing minimum velocity prediction and presented different equations to determine it [2–5]. Nalluri and Ab Ghani [4] presented an equation for sediment transport at the limit of deposition after conducting many experiments. They also suggested using the deposition bed state, which increases the effective bed width; in pipes with a diameter exceeding 1000 m. Ota and Nalluri [6] used physical concepts and presented a new model for examining sediment transport at the limit of deposition in pipes with a diameter above 500 mm. They also stated that the presented model has better economic justification than existing models. Banasiak [7] conducted a number of experiments on the movement of non-cohesive and semi-cohesive sediment and examined the behavior of sediment deposited in the channel besides its effect on the pipe's hydraulic performance. He showed that a transport regime occurs for Froude numbers higher than 0.5 accompanied by a washed-out bed form. Ebtehaj et al. [8] used three different data sets including a wide range of parameters affecting the minimum velocity and corrected the limit of deposition equation presented by Vongvisessomjai et al. [2]. They also presented an equation for pipes with larger diameters regarding the relative deposition depth.

Because of the complex mechanism in various geometrical and hydraulic conditions, soft computing methods have been widely applied in this field [9–18]. Numerous research works have also been conducted in the field of sediment deposition as one of the crucial topics of water engineering using the soft computing machine learning approach [19]; genetic programming [20]; gene expression programming [21, 22]; artificial neural networks [23]; adaptive neural fuzzy inference system [24, 25] and evolutionary algorithms [26].

Roushangar et al. [27] used gene expression programming (GEP) and adaptive neuro fuzzy inference system (ANFIS) to predict sediment transport in the Qotur River. According to their results, the method presented was more accurate than existing methods. Bravo et al. [28] presented a numerical model for the incipient motion of sediment transport based on the discrete element method (DEM). They used the proposed method to predict the border sheer stress field using fluid flow. Ebtehaj and Bonakdari [29] examined the parameters influencing sediment deposition at the limit of deposition. They defined the densimetric Froude number (Fr) parameter for the purpose of determining the minimum velocity required to prevent sediment deposition. They applied the genetic algorithm and imperialist competitive algorithms (GA and ICA) to predict this dimensionless parameter. It was found these algorithms predict Fr more accurately than existing sediment deposition equations. Kitsikoudis et al. [30] examined the performance of tree machine learning approaches (ANN, GP and ANFIS) to obtain an equation for sediment transport in sand-bed rivers. According to their results, the machine learning approach is more accurate that the other two existing methods.

Extreme learning machines (ELM) have been widely applied in engineering in recent years, such as in automatic human knee cartilage segmentation [31]; intrusion detection [32]; feature selection for nonlinear models [33]; the problem with noisy measurements [34]; hyperspectral image classification [35]; slope stability [36] and unconfined compressive strength of carbonate rocks [37]. So far, there are no studies devoted to assessing the proficiency of ELM in sediment transport analysis. Therefore, the advantage of the ELM approach helps in applying it to the estimation of sediment transport and Fr number in open channels.

The main objective of this article is to examine the ELM algorithm's performance in predicting Fr in open channels. Therefore, six different models are presented through examining the factors affecting sediment transport. Following that, the performance of this algorithm is compared

with the back propagation (BP) algorithm, genetic programming (GP) and existing sediment transport equations. Moreover, the performance of ELM is examined for different pipe diameters.

2 Sediment transport

2.1 Self-cleansing concept

Due to the generally poor performance of traditional methods in terms of accurate estimation of minimum velocity, there is a need for a certain equation to precisely predict the transfer of sediments in open channels. The presented equation should contain factors affecting this three-dimensional complex phenomenon. Therefore, several tests should be performed to obtain equations which can evaluate the ability to transfer the minimum concentration or lowest density of fine-grained suspended load, transport coarse particles as bed load and erode sediment particles from the deposited bed [38]. The test results indicate intense dependence of minimum velocity on the pipe diameter. Therefore, using pipes with large diameters in steep gradients seems to be uneconomical. Thus, a hypothesis for the minimum amount of settling was suggested. To prove this hypothesis, two fundamental criteria need to be verified: 1 considering the amount of minimum deposition which leads an increasing in the bed roughness but does not result in a decreasing of the bed transport capacity. 2 if the deposition depth can be sustainable, or will settling increase with time until the pipe becomes full.

If partly full flow passes through a channel loaded with sediment being transferred in the form of bed load and the flow velocity is not enough to prevent the formation from settling, then a sedimentary bed will be created. This problem increases the bed resistance because greater resistance depth reduces the velocity. It may be assumed that velocity reduction is the reason for reduced transport capacity as well as settling and obstruction. However, Ackers et al. indicated that the presence of substrate leads to higher sediment transport capacity [38, 39]. According to the description provided, the sediment transport in open channels should strike a balance between the settling amount and erosion rate in a specified time period, thus minimizing the combined costs of construction, operation and maintenance.

2.2 Sediment transport equations

Sediment transport equations at the limit of deposition can be divided into two groups: equations obtained from dimensional analysis and from semi-experimental equations. May et al. [38] solved sediment particle equations in the balanced state and used seven different data sets [39] to form the following equations:

$$C_V = 3.03 \times 10^{-2} \left(\frac{D^2}{A}\right) \left(\frac{d}{D}\right)^{0.6} \left(\frac{V^2}{g(s-1)D}\right)^{1.5} \left(1 - \frac{V_t}{V}\right)^4 \tag{1}$$

$$V_t = 0.125\sqrt{g(s-1)d}\left(\frac{y}{d}\right)^{0.47} \tag{2}$$

where C_V is volumetric sediment concentration, A is the cross-sectional flow area, D is the pipe diameter, s is the specific gravity of sediment ($=\rho_s/\rho$), d is the median diameter of particles, y is the flow depth, g is the gravitational acceleration, V is the flow velocity and V_t is the velocity required for incipient sediment motion (Eq. 2).

Azamathulla et al. [24] and Ebtehaj et al. [8] equations are amongst the sediment transport equations attained using dimensional analysis. Azamathulla et al. [24] used Ab Ghani [40] and Vongvisessomjai et al. [2] data sets and corrected the equation presented by Nalluri and Ab Ghani [4] as follows:

$$Fr = \frac{V}{\sqrt{g(s-1)d}} = 0.22 C_V^{0.16} D_{\mathrm{gr}}^{-0.14} \left(\frac{d}{R}\right)^{-0.29} \lambda_s^{-0.51} \tag{3}$$

where D_{gr} ($=d(g(\mathrm{s}-1)/\nu^2)^{1/3}$) is the dimensionless particle number, R is the hydraulic radius and λ_s is the overall sediment friction factor, which is calculated as follows:

$$\lambda_s = 0.851 \lambda_c^{0.86} C_V^{0.04} D_{gr}^{0.03} \tag{4}$$

where λ_c is the clear water friction factor of the channel. Ebtehaj et al. [8] used two different data sets [2, 40] and corrected the equation presented by Vongvisessomjai et al. [2] for bed load transport at the limit of deposition (Eq. 5). They also verified their equation results with Ota and Nalluri's [5] data set.

$$Fr = \frac{V}{\sqrt{g(s-1)d}} = 4.49 C_V^{0.21} \left(\frac{d}{R}\right)^{-0.54} \tag{5}$$

3 Materials and methods

3.1 Data collection

In this study, the data sets presented by Ab Ghani [40], Ota and Nalluri [5] and Vongvisessomjai et al. [2] are used. All data sets presented in this paper concern the limit of deposition state. Ab Ghani [40] conducted tests on 3 pipes, 20.5 m long and with different diameters, 0.154, 0.305 and 0.405 m. Smooth pipes were used in all tests, and the effect of roughness was studied on the pipe with 0.305 m diameter.

Ota and Nalluri [5] conducted limit of deposition tests at Chalmer University of Technology, Sweden, on 25 m long

Table 1 The ranges of all data sets

	Ab Ghani (1993)			Ota and Nalluri (1999)			Vongvisessomjai et al. (2010)		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
C_V	1.00E − 06	1.28E − 03	2.69E − 04	5.94E – 05	2.02E −05	3.48E − 05	9.00E − 05	4.00E − 06	3.36E − 05
d/R	6.33E − 03	2.46E − 01	7.21E − 02	9.86E – 02	9.45E − 03	3.85E − 02	2.53E − 02	7.14E − 03	1.49E – 02
d/D	1.51E − 03	3.70E − 02	1.34E − 02	2.49E – 02	3.16E − 03	1.16E − 02	4.30E − 03	1.33E − 03	2.62E – 03
D^2/A	8.31E − 01	1.06E + 01	4.21E + 00	3.51E + 00	9.12E − 01	1.83E + 00	1.61E + 01	3.41E + 00	5.58E + 00
λ_s	1.29E − 02	3.92E − 02	2.29E − 02	2.99E − 02	2.49E − 02	2.67E − 02	5.32E − 02	3.83E − 02	4.42E – 02
D_{gr}	5.40E + 00	9.75E + 01	4.44E + 01	1.42E + 02	1.80E + 01	6.60E + 01	1.09E + 01	5.06E + 00	7.92E + 00
R/D	1.08E − 01	3.02E − 01	2.06E − 01	3.82E − 01	2.24E − 01	3.09E − 01	2.80E − 01	8.00E − 02	1.85E – 01
Fr	1.52E + 00	1.13E + 01	3.85E + 00	6.32E + 00	1.86E + 00	3.40E + 00	8.98E + 00	2.84E + 00	5.62E + 00

pipes with 305 mm diameter. The slope of the pipes was 0.00315, roughness was 0.24 mm and Manning's coefficient of roughness was around 0.01.

Vongvisessomjai et al. [2] conducted experiments on two pipes with 100 and 150 mm diameter and 16 m length. These two sections were used to measure flow. One was positioned at a distance of 4.5 m upstream and the other at 5.5 m downstream, with 6 m separating these two points. In each section, the velocity at the flow surface, at intermediate depth and near the bed was measured and the average was deemed the average velocity. For the dual phase, water and air flow, Manning's coefficient of roughness (n) was equal to 0.0125. The ranges of all three data sets are presented in Table 1.

3.2 Extreme learning machines (ELM)

ELM as a learning algorithm tool was first introduced for single layer feed-forward neural network (SLFN) architecture [41, 42]. An ELM algorithm chooses the input weights randomly and determines the output weights of SLFN analytically. It has a more suitable general capability with faster learning speed and is also able to determine all network parameters analytically, thus preventing human intervention. Compared to other soft computing methods, additional advantages include ease of use, quick learning speed, higher performance and suitability for different nonlinear activation and kernel functions. The training process is faster than common epoch-based algorithms such as Levenberg–Marquardt, which depend on close-form in their training and do not include any sort of nonlinear optimization routine. A mathematical description of an SLFN function with n hidden nodes can be expressed as follows:

$$f_n(x) = \sum_{i=1}^{n} \beta_i G(a_i, b_i, x) \tag{6}$$

where β_i is the weight between the ith hidden node and the output node; a_i ($a_i \in R^n$) and b_i are the learning factors of the hidden nodes; and $G(a_i, b_i, x)$ is the output of the ith hidden node for input x. The activation function $g(x)$ (e.g., sigmoid and threshold) for an additive hidden node $G(a_i, b_i, x)$ is represented as:

$$G(a_i, b_i, x) = g(a_i.x + b_i) \tag{7}$$

Biases and weights are randomly allocated to the neurons of the input and hidden layers in ELM. Activating the hidden layer neurons for each training case in an ELM with "j" neurons in the hidden layer, "i" neurons in the input layer, and "k" training cases, are computed with the following equation:

$$H_{jk} = g\left(\sum \left(W_{ji} X_{ik}\right)\right) + B_j \tag{8}$$

where $g(.)$ could be any nonlinear activation piecewise continuous function; W_{ji} is the weight of the ith input neuron and the jth neuron of the hidden layer; B_j is the bias of the jth neuron of the hidden layer; X_{ik} is the input of the input neuron for the kth training case; and H_{ik} is the activation matrix of the jth neuron of the hidden layer for the kth training case. As such, the activation of all neurons in the hidden layer for the cases used in training are presented through this matrix, where j is a column and k is a row. Matrix H denotes the matrix of the neural network's output hidden layer. The weights between the output and hidden layers are used with the least-square fit application for the target values in the training mode in other outputs of the neurons in the hidden layer for each training case. The mathematical equation is presented as:

$$H\beta = T \tag{9}$$

$$\beta = \left(\beta_1, \ldots, \beta_j\right)_{j\times 1} \tag{10}$$

where β indicates the weight between the output layer neurons and hidden layer neurons; and T is the axis representing the target values for the training cases, expressed below:

$$T = (T_1, \ldots, T_k)_{k\times 1} \tag{11}$$

Consequently, the weights can be calculated as follows:

$$\beta = H'T \tag{12}$$

where

$$H(\tilde{a}, \tilde{b}, \tilde{x}) = \begin{bmatrix} G(a_1, b_1, x_1) \ldots G(a_L, b_L, x_L) \\ G(a_1, b_1, x_N) \ldots G(a_L, b_L, x_N) \end{bmatrix}_{N\times L} \tag{13}$$

with $\tilde{a} = a_1, \ldots, a_L;\ \tilde{b} = b_1, \ldots, b_L;\ \tilde{x} = x_1, \ldots x_L$

$$\beta = \begin{bmatrix} \beta_1^T \\ : \\ \beta_L^T \end{bmatrix}_{L\times m} \quad \text{and} \quad T = \begin{bmatrix} T_1^T \\ : \\ T_L^T \end{bmatrix}_{L\times m} \tag{14}$$

where β is the axis of the weight between the neurons in the hidden layer and output layers; H' is the Moore–Penrose inverse of Matrix H; and T is the axis between the weights of the training cases. Regarding the explanations provided, it can be stated that ELM training comprises two stages: (1) randomly allocating the biases and weights to the hidden layer neurons and calculating the hidden layer output from matrix H, and (2) calculating the output weights using the Moore–Penrose inverse of Matrix H and the target values for different training cases.

The training process for finding the inverse Moore–Penrose Matrix of the hidden layer is faster than common epoch-based algorithms such as Levenberg–Marquardt, which depend on close-form in their training and do not include any sort of nonlinear optimization routine. Therefore, the training period of the network is decreased significantly [42, 43]. ELM works with large numbers of random nonlinear input space projections, such that each neuron is connected to a single projection.

3.3 Methodology

Regarding studies conducted in the field of sediment transport at limit of deposition [2, 4, 38, 40], the fundamental parameters governing sediment transport in a pipe channel include flow depth (y) or hydraulic radius (R); velocity (V) or mean shear stress (τ_0); kinematic viscosity (ν); water density (ρ); median diameter of particle size (d); sediment density (ρ_s); volumetric sediment concentration (C_V); pipe diameter (D); overall sediment friction factor (λ_s); channel slope (S_0); and gravitational acceleration (g). Therefore, the functional relation of flow velocity is presented as follows:

$$V = f(g,\ d,\ s,\ C_V,\ D,\ y,\ R,\ A,\ \lambda_s) \tag{15}$$

The dimensionless parameters influencing sediment transport in open channels are the densimetric Froude number ($Fr = V/\sqrt{g(s-1)d}$), sediment volume concentration (C_V), dimensionless particle size ($D_{gr} = (d(g(s\text{-}1)/\nu^2)^{1/3})$), the average relative particle size (d/D), the ratio of average particle size to hydraulic radius (d/R), the square pipe diameter to the cross-sectional area of flow (D^2/A), the ratio of hydraulic radius to pipe diameter (R/D); and the overall sediment friction factor (λ_s) [23, 25, 26, 29]. By the nature of each dimensionless parameter mentioned above, Ebtehaj and Bonakdari [25, 26] classified these parameters in five groups: movement (Fr), transport (C_V), sediment (D_{gr}, d/D), transport mode (d/R, D^2/A, R/D) and flow resistance (λ_s). According to that, objective of this study is prediction of minimum velocity to prevent sedimentation; Fr parameter related to movement group is selected as target parameter (output of ELM). Therefore, the parameters from the other four groups are used to predict the target parameter (Fr). The reason for simultaneously using one dimensionless parameter from each group in predicting the Fr is that sensitivity analysis conducted in previous studies [23, 24] indicates the significance of using four groups in predicting the parameter related to the "movement (Fr)" group. As shown in Fig. 1, except for the "transport" (C_V) and "flow resistance" (λ_s) groups that only have one dimensionless parameter, the "sediment" (D_{gr} & d/D) and "transport mode" (d/R, D^2/A and R/D) groups have more than one dimensionless parameter. Therefore, to examine the performance of each presented parameter, the parameters of other groups are assumed constant and one group's dimensionless parameters are changed, eventually resulting in six different groups. It is also evident from the figure that only one parameter from each group has been used.

3.4 Performance evaluation index

Statistical indexes are often used to examine model performance accuracy. Considering the fact that using one single index is not a good criterion for evaluating model accuracy [44], relative and absolute indexes are used in this study. Absolute indexes use the measurement units of the target parameter and examine the prediction accuracy of a model while the relative indexes consider the target parameter to be dimensionless and solve this problem. The relative and absolute indexes applied in this study include the Root Mean Square Error (RMSE) and the Mean Absolute Relative Error (MARE), which are calculated as:

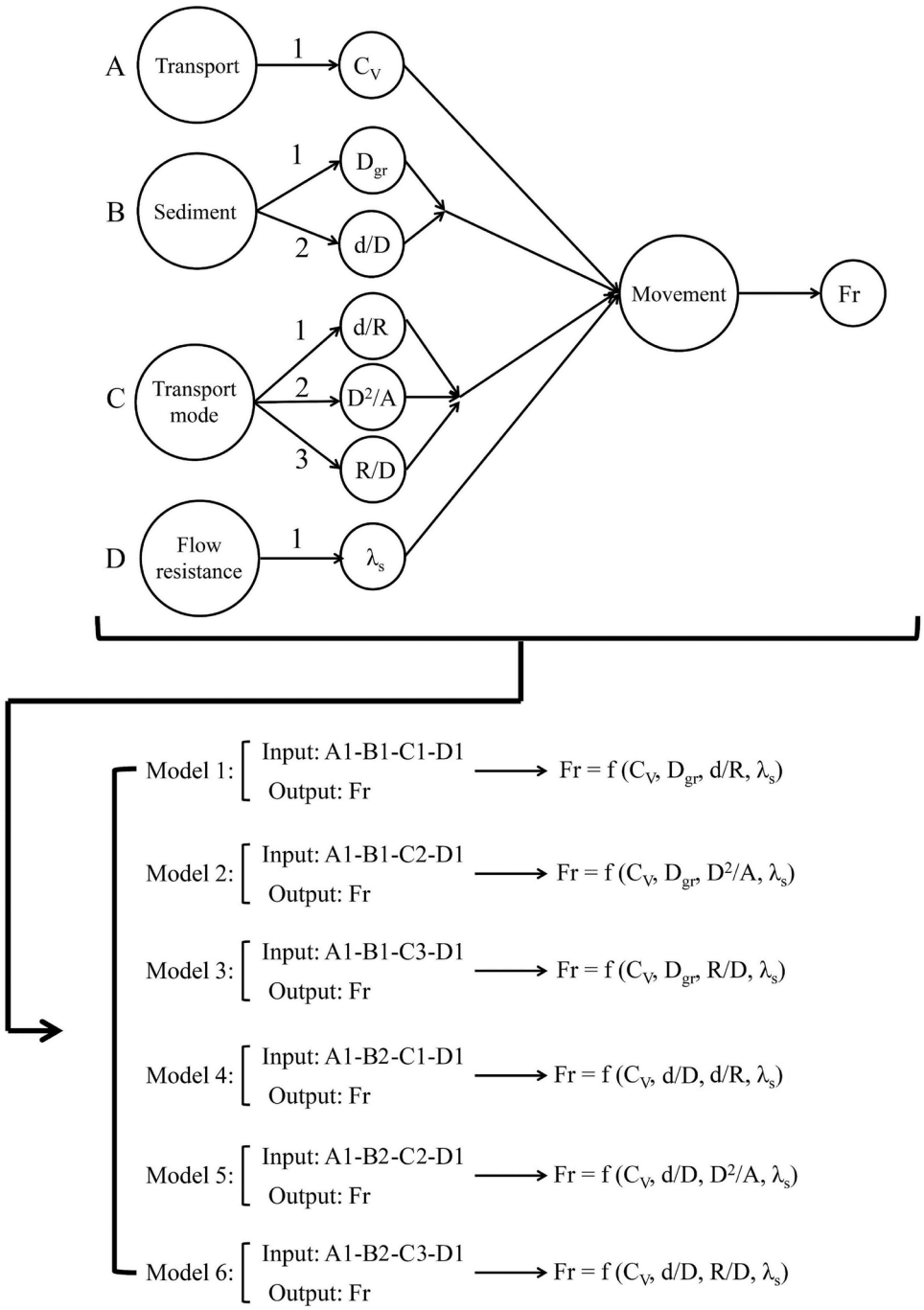

Fig. 1 Dimensionless groups and proposed models

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Fr_{\mathrm{Exp}} - Fr_{\mathrm{ELM}}\right)^2} \tag{16}$$

$$\mathrm{MARE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|Fr_{\mathrm{Exp}} - Fr_{\mathrm{ELM}}\right|}{Fr_{\mathrm{Exp}}} \tag{17}$$

Three different sets of data with a total of 218 different data (Table 1) are used in this study to predict the *Fr* using FFNN-ELM, FFNN-BP and GP. 30 % of the total data were randomly selected to test the models and the remaining 70 % were used to train the models. The performance of each of these models was validated using the testing data after training the models.

3.5 Architecture of soft computing models

The parameters of the ELM and ANN modeling frameworks employed in the present study are presented in

Table 2 User-defined parameters for the (a) ELM model

(a) ELM model						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Hidden neurons	17	26	14	20	18	19
Activation function	Sigmoidal function	Sigmoidal function	Sigmoidal function	Sigmoidal function	Sigmoidal function	Sigmoidal function
Neurons—input	4	4	4	4	4	4
Neurons—output	1	1	1	1	1	1
Learning rule	ELM for SLFNs	ELM for SLFNs	ELM for SLFNs	ELM for SLFNs	ELM for SLFNs	ELM for SLFNs
Number of iteration	30,000	30,000	30,000	30,000	30,000	30,000

Table 3 User-defined parameters for the (b) ANN model

(b) ANN model	
Parameter	Value
Momentum constant	0.3
Learning rate	0.5
Maximum number of epochs	1000
Performance goal	10–7
Learning rule	Back propagation
Back propagation algorithm	Levenberg–Marquardt back propagation
Activation function	Sigmoidal function
Number of layers	3
Neurons—input	4
Neurons—hidden	19
Neurons—output	1
Number of iteration	1000
Activation function	Sigmoidal function

Table 4 User-defined parameters for the (c) GP model

(c) GP model	
GP	
Output	1
Population size	512
Function set	$+, -, \times, \div, \sqrt{}, x^2, ln, e^x, a^x$
Head size	5–9
Chromosomes	20–30
Number of genes	2–3
Mutation rate	91.46
Crossover rate	30.56
Inversion rate	108.53

Tables 2, 3, 4. The number of hidden neurons in the ELM models was determined through trial and error for both cases. The data was normalized prior to applying the process. In the present study, the 'fivefold' cross-validation method was employed to check the generalization ability of the model and to select the appropriate input vector. In the 'fivefold' cross-validation method, the data are segmented into five sub-samples. Of the five sub-samples, a single subsample is retained as validation data to test the model and the remaining four comprise the training data. The cross-validation process is then repeated five times with each of the five sub-samples used exactly once as validation data. This process can manage even badly mannered data.

4 Results and discussion

In Fig. 2, the performance of the feed-forward neural network (FFNN) is evaluated using two algorithms, extreme learning machine (ELM) and back propagation (BP), in the network training mode and genetic programming (GP). The results of predicting *Fr* using FFNN-ELM in models 1 (RMSE = 0.26, MARE = 0.052), 4 (RMSE = 0.295, MARE = 0.052), and 6 (RMSE = 0.309, MARE = 0.059), which consider *Fr* to be dependent on the $Fr = f(C_V, D_{gr}, d/R, \lambda_s)$, $Fr = f(C_V, d/D, d/R, \lambda_s)$ and $Fr = f(C_V, d/D, R/D, \lambda_s)$ parameters, respectively, are fairly accurate with the predicted values having less than 10 % relative error. FFNN-BP and GP also performed better in these models than the other models, but it was less accurate than FFNN-ELM; the *Fr* predicted by this algorithm had more than 10 % error in some cases, which was not observed with FFNN-ELM. It is evident that the index values in this figure are lower for FFNN-ELM than FFNN-BP and GP for all models, indicating that FFNN-ELM is more accurate than FFNN-BP and GP. When using the C_V, d/R, and λ_s parameters related to the transport, transport mode, and flow resistance groups, using each of the parameters presented for the sediment group (D_{gr}, d/D) (models 1 and 4) does not significantly affect the results or both indexes. Models 4 and 6 perform better in predicting the *Fr* in the training stage when using the D^2/A parameter from the "transport mode" group and when the other parameters are constant (C_V, d/D, λ_s) rather than when the two other parameters (R/D and d/R) from this group are used, which would lead to a significant decrease in the prediction

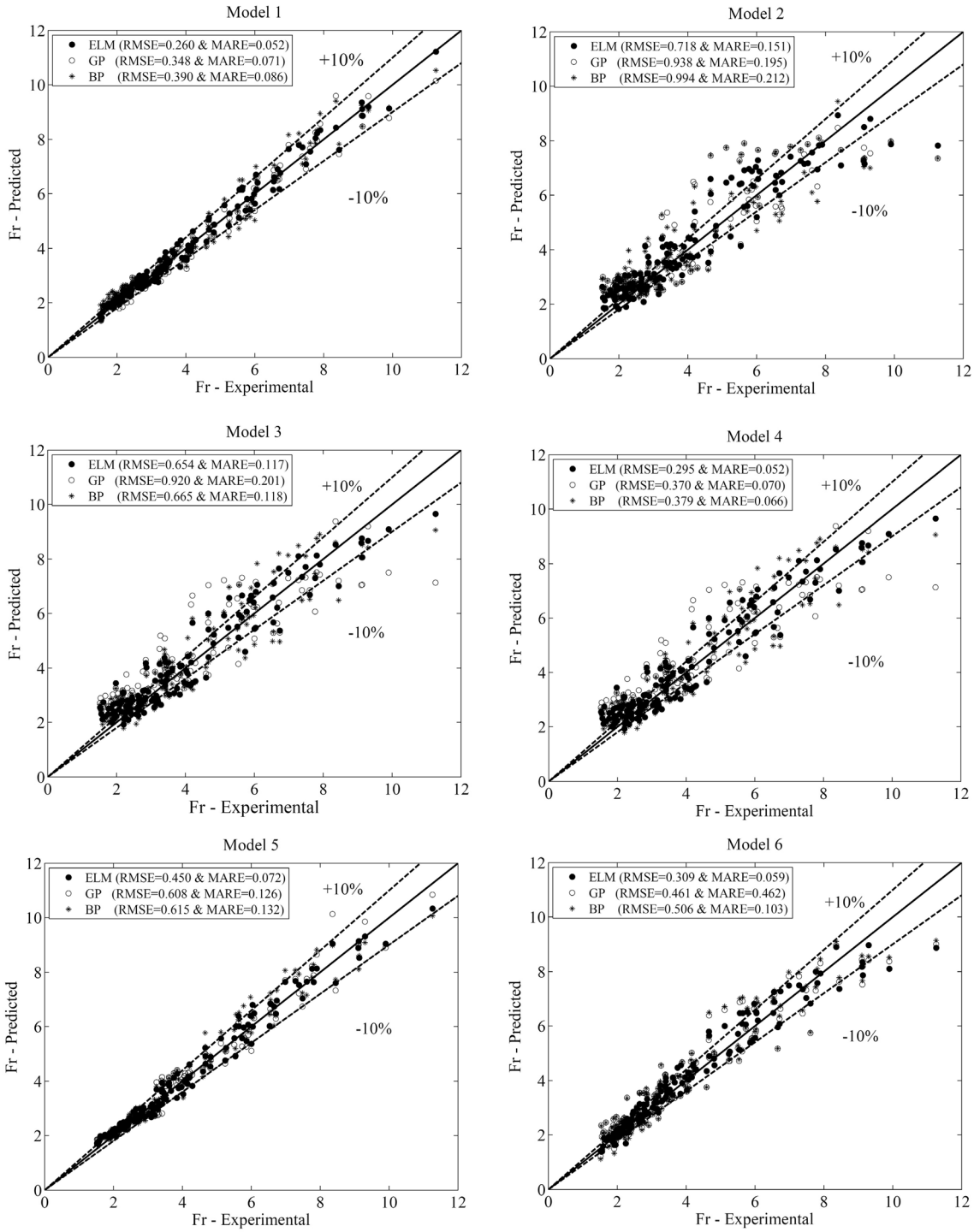

Fig. 2 Performance of FFNN-ELM, GP and FFNN-BP in estimating the *Fr* for all models (training)

accuracy of both algorithms. Models 2 [$Fr = f(C_V, d/D, D^2/A, \lambda_s)$] and 3 [$Fr = f(C_V, d/D, R/D, \lambda_s)$] present weaker *Fr* prediction results in network training for both FFNN-ELM, FFNN-BP and GP methods compared with model 1. Therefore, the best *Fr* prediction performance is obtained when the *d*/*R* parameter is used as the representative of the "transport mode" group and the C_V, *d*/*D*, λ_s parameters are used from the remaining dimensionless groups.

The *Fr* predictions by FFNN-ELM model 1 (RMSE = 0.121, MARE = 0.023), model 4 (RMSE = 0.141, MARE = 0.025), and model 6 (RMSE = 0.152, MARE = 0.025) were acceptable, similar to the model training

results. In fact, in addition to C_V and λ_s which are constant in all models, directly utilizing the hydraulic radius (R) is necessary to attain accurate prediction. This variable takes into account the flow depth and pipe diameter, whereas using other parameters in the "transport mode" group (D^2/A) leads to lower accuracy (model 2). FFNN-BP (RMSE = 0.31, MARE = 0.058) and GP (RMSE = 0.327, MARE = 0.07) also achieved good results in these models; even for model 1, all predicted values had less than 10 % relative error, signifying better results than the model's training mode (BP: RMSE = 0.39, MARE = 0.086; GP: RMSE = 0.348, MARE = 0.071). *Fr* prediction did not present acceptable results with model 2 and with both algorithms. Both predicted the *Fr* with more than 10 % error and in some cases it even exceeded 20 %. FFNN-ELM was relatively more accurate than FFNN-BP and GP in this model. In addition, according to the statistical index, Model 2 presented almost the weakest results, whereby the results of the predictions made by the three methods (FFNN-ELM; RMSE = 0.625, MARE = 0.123; FFNN-BP; RMSE = 1.073, MARE = 0.138; GP; RMSE = 1.089, MARE = 0.172) were less accurate than the other models. Model 3 [$Fr = f(C_V, D_{gr}, R/D, \lambda_s)$] made predictions with relative error above 10 % in both presented methods. This model made most predictions with underestimated values while both methods presented in this study made under- and overestimated predictions in network training. This model, however, has nearly equal statistical indexes for both methods in network training and testing modes. The only difference between model 5 and model 4, which predicted the *Fr* fairly accurately, is that model 5 used the D^2/A parameter instead of R/D (from the "transport mode" group). This model presents the *Fr* parameter in the form of underestimated and overestimated values with a relatively great relative error when using the FFNN-BP (RMSE = 0.783, MARE = 0.123) method, resulting in either sediment deposition in the channel bed or uneconomical planning (respectively). The FFNN-ELM (RMSE = 0.487, MARE = 0.071) method predicts *Fr* more accurately, with most predictions having a relative error less than 10 %.

In general, the effect of particle size in the "sediment" group as the D_{gr} parameter led to better results than d/D. This is because pipe diameter is considered a hydraulic parameter, so models 1, 2 and 3 were more accurate than 4, 5 and 6, respectively. In the "transport mode" group the use of hydraulic radius with median diameter of particles (d/R) is better than using D^2/A and R/D. The worst results when considering C_V, D_{gr} and λ_s as effective parameters of transport, sediment and flow resistance were achieved when D^2/A was used as a parameter related to the "transport mode" group. The reason is that this input combination does not contain flow depth, whereas models 1 and 3 contain flow depth as a hydraulic parameter.

Figure 4 shows the performance of FFNN-ELM in different pipe channels (100, 150, 154, 225, 305 and 450 mm) for model 1. The *Fr* prediction results were examined for $D = 305$ mm in two rough beds ($k = 0.53$ and 1.34 mm) as well as a smooth bed. It is clear from the figure that the *Fr* values related to different diameters have a relatively good range. FFNN-ELM is fairly accurate for all pipe diameters, so much so that the predicted *Fr* values are quite consistent with the actual values, with a relative error less than 10 %, indicating reliable predictions. The minimum MARE index value is $D = 100$ mm, which is equal to 0.019, and $D = 225$ mm has the minimum RMSE value in predicting the *Fr* using the FFNN-ELM and input data presented for model 1 (equal to 0.11). For $D = 450$ mm (RMSE = 0.471, MARE = 0.059), the *Fr* prediction is relatively less accurate than for the other diameters. However, according to the figure, the maximum relative error presented by this model was 10 % and the mean prediction relative error was approximately 5 % for this diameter, which indicates that *Fr* predictions are desirable when the FFNN-ELM method is used.

Table 5 serves to quantitatively examine *Fr* prediction accuracy when using FFNN-ELM. It can be seen that except for $D = 225$ mm where model 4 has values below the statistical indexes and for $D = 305$ mm ($k = 1.34$ mm) where model 6 experiences the same, model 1 yields better results in other states. In the two states where model 1 does not give the best results, the values obtained for the RMSE and MARE indexes are slightly different from the superior models. FFNN-ELM performs best at D = 100 mm and $D = 150$ mm with a relative error of approximately 2 %. FFNN-ELM performs better for larger diameters as well, with a relative error less than 6 % for $D = 305$ mm ($k = 0.53$ mm), which presents the weakest results. Therefore, it can be stated that FFNN-ELM predicts *Fr* fairly accurately for all diameters.

In selecting the optimal amongst models presented in this study in terms of testing and training mode results whose quantitative statistical index values are given in Figs. 2 and 3, and the results in Table 5 relate to different diameters, model 1 appears more accurate for all states (Fig. 4).

Figure 5 compares the performance of FFNN-ELM with the existing sediment transport equation at limit of deposition. The predictions made using FFNN-ELM are very consistent with the actual values. In nearly all cases, prediction was achieved with small relative error (RMSE = 0.023 and MARE = 0.121). Ebtehaj et al.'s [8] equation is fairly accurate for small *Fr* (RMSE = 0.076 and MARE = 0.556); however, as the *Fr* increases the equation's accuracy drops, but it clearly makes more accurate predictions regarding *Fr* than the two other equations [24, 38]. May et al.'s [38] equation predicts differently for different cases. It presents *Fr* with relatively large error and underestimates and overestimates (RMSE = 0.094 and MARE = 0.741), which leads to sediment deposition on the channel bed or uneconomical design, respectively. Azamathulla et al.'s [24] equation considers the parameters affecting *Fr* prediction similar to

Table 5 FFNN-ELM evaluation (all models) for different diameters

Diameter	Statistical Index	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$D = 100$ mm	RMSE	**0.164**	0.677	0.669	0.243	0.731	0.234
Smooth bed	MARE	**0.019**	0.093	0.088	0.027	0.095	0.030
$D = 150$ mm	RMSE	**0.122**	0.769	0.759	0.173	0.560	0.205
Smooth bed	MARE	**0.020**	0.091	0.095	0.023	0.082	0.026
$D = 154$ mm	RMSE	**0.248**	0.850	0.761	0.237	0.285	0.256
Smooth bed	MARE	**0.053**	0.256	0.209	0.063	0.070	0.057
$D = 225$ mm	RMSE	0.110	0.519	0.596	**0.106**	0.258	0.116
Smooth bed	MARE	**0.024**	0.110	0.134	**0.024**	0.053	0.028
$D = 305$ mm	RMSE	**0.179**	0.659	0.623	0.274	0.343	0.218
Smooth bed	MARE	**0.043**	0.187	0.192	0.066	0.061	**0.043**
$D = 450$ mm	RMSE	**0.471**	1.375	0.821	0.516	0.991	0.658
Smooth bed	MARE	**0.059**	0.151	0.115	0.063	0.105	0.075
$D = 305$ mm	RMSE	**0.246**	0.418	0.487	0.263	0.329	0.262
Rough bed ($k = 0.53$ mm)	MARE	**0.059**	0.104	0.129	0.060	0.079	0.061
$D = 305$ mm	RMSE	0.163	0.427	0.438	0.147	0.161	**0.156**
Rough bed ($k = 1.34$ mm)	MARE	0.049	0.113	0.114	0.046	0.052	**0.045**
All	RMSE	**0.233**	0.691	0.624	0.257	0.462	0.266
	MARE	**0.044**	0.136	0.141	0.048	0.071	**0.044**

Bold values show the minimum value for the RMSE and MARE ceritaria

the parameters considered for model 1 of this study, but does not predict well. Most predictions are underestimated (RMSE = 0.784 and MARE = 0.139), which leads to sediment deposition on the channel bed. Therefore, using May et al. [38] and Azamathulla et al.'s [24] equations for the purpose of predicting Fr are not considered reliable and it is advisable to avoid using them in design.

Since FFNN-ELM (a) had the best results, Fr can be calculated with the following equation:

$$Fr = \left[\frac{1}{(1+\exp\left(\mathrm{In}W \times \mathrm{In}V + \mathrm{BHI}\right))}\right]^{T} \times \mathrm{Out}W \tag{18}$$

where

$$\mathrm{In}V = \begin{bmatrix} C_V \\ D_{gr} \\ d/R \\ \lambda_s \end{bmatrix} \quad \mathrm{BHI} = \begin{bmatrix} 0.34 \\ 0.85 \\ 0.51 \\ 0.77 \\ 0.51 \\ 0.19 \\ 0.01 \\ 0.98 \\ 0.40 \\ 0.87 \\ 0.18 \\ 0.65 \\ 0.26 \\ 0.99 \\ 0.07 \\ 0.51 \\ 0.89 \end{bmatrix} \quad \mathrm{Out}W = \begin{bmatrix} 9.27 \\ 0.38 \\ 199.40 \\ 2.25 \\ 222.33 \\ -46.98 \\ -127.30 \\ 22.47 \\ -5.84 \\ -10.20 \\ 17.61 \\ -115.02 \\ -5.59 \\ -28.38 \\ 22.03 \\ 9.06 \\ -152.45 \end{bmatrix} \quad \mathrm{In}W = \begin{bmatrix} 0.08 & -0.61 & -0.26 & 0.49 \\ -0.87 & 0.99 & -0.28 & 0.86 \\ 0.00 & 0.60 & 0.23 & -0.60 \\ -0.14 & 0.77 & 0.11 & 0.90 \\ 0.59 & -0.43 & 0.83 & 0.08 \\ 0.38 & -0.90 & -0.35 & -0.84 \\ 0.51 & 0.47 & 0.54 & -0.17 \\ 0.49 & -0.44 & -0.77 & 0.47 \\ -0.39 & 0.14 & -0.78 & -0.97 \\ -0.60 & -0.81 & 0.64 & -0.45 \\ -0.98 & -0.50 & 0.29 & 0.44 \\ 0.11 & 0.46 & 0.29 & -0.33 \\ 0.65 & 0.00 & -0.92 & 0.49 \\ -0.72 & 0.48 & 0.06 & 0.29 \\ 0.90 & 0.91 & 0.36 & -0.32 \\ 0.68 & 0.43 & -0.26 & 0.44 \\ 0.45 & -0.28 & 0.97 & 0.10 \end{bmatrix}$$

5 Conclusions

Regarding the importance of studying sediment transport in open channels, the minimum velocity required to prevent sediment deposition was estimated in this study using extreme learning machines (ELM) and feed-forward neural networks (FFNN) (FFNN-ELM). The results obtained from this research are briefly explained below:

- Among the six different presented models, models 1 and 4 produced better results than the other models, such that the mean relative error in predicting the Fr was

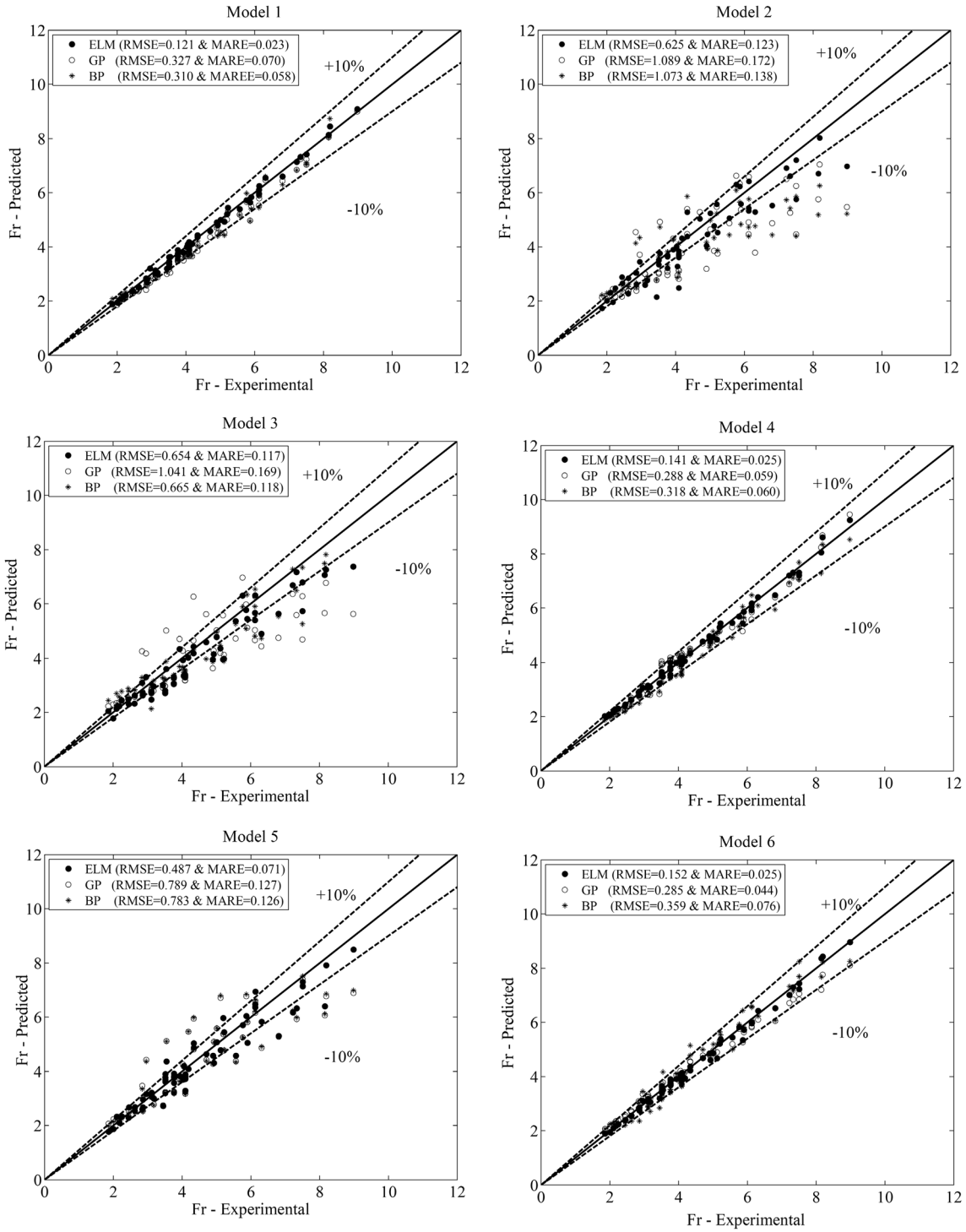

Fig. 3 Performance of FFNN-ELM, GP and FFNN-BP in estimating *Fr* for all models (testing)

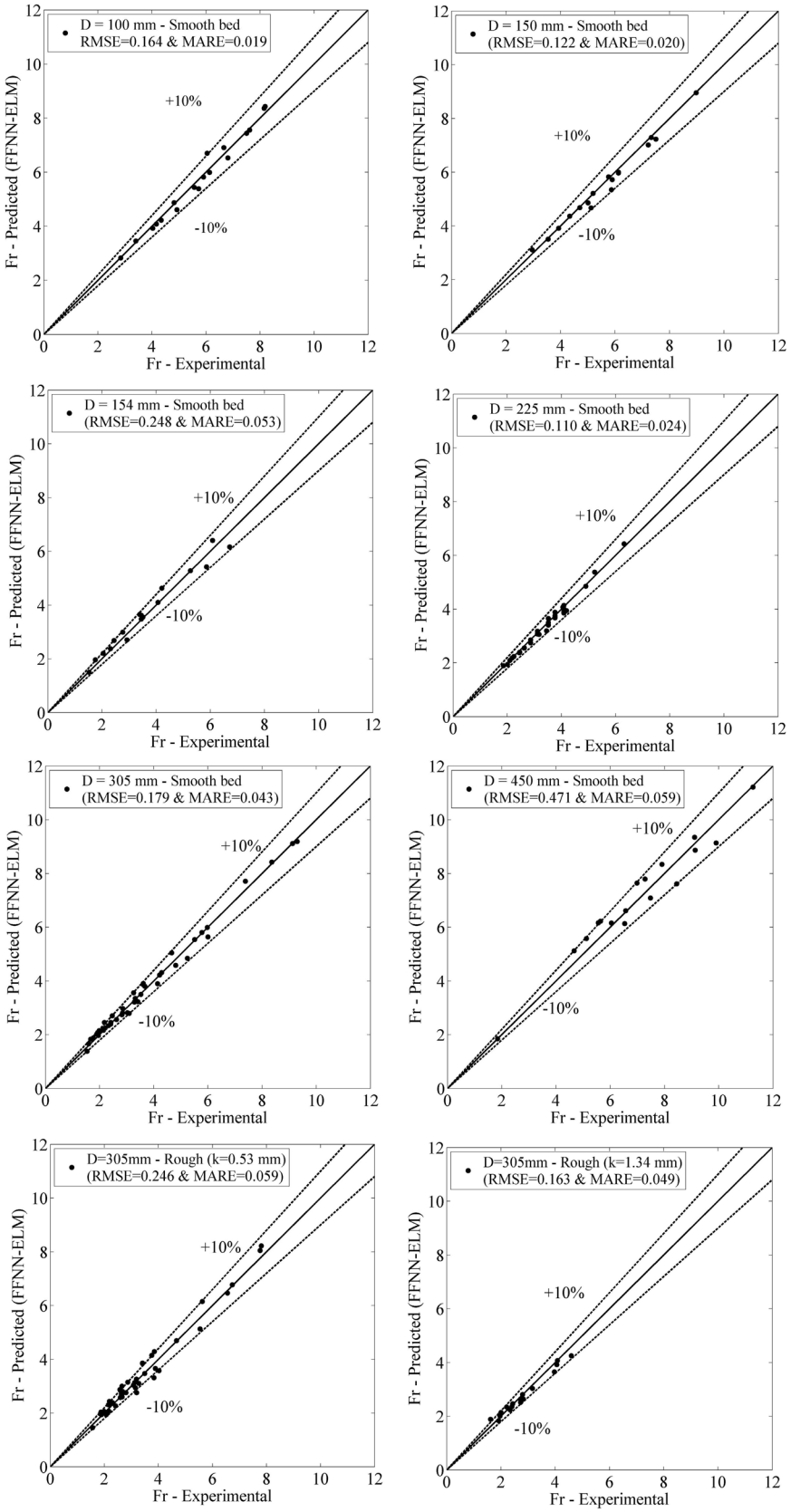

Fig. 4 FFNN-ELM performance in estimating *Fr* for model 1 with different diameters

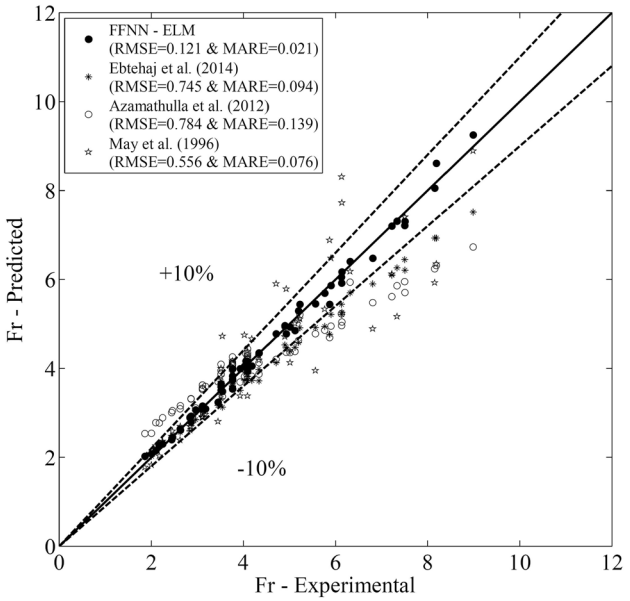

Fig. 5 Performance of FFNN-ELM versus existing equations for estimating *Fr*

approximately 5 % (MARE = 0.052). However, the best model is fairly accurate regarding all statistical indexes. Therefore, model 1 performs best according to Tables 2, 3, 4 and in all methods used in this study (FFNN-ELM, FFNN-BP and GP).

- Comparing the three techniques in this study indicates that FFNN-ELM outperformed the FFNN-BP and GP method in both training (RMSE = 0.26 and MARE = 0.052) and testing (RMSE = 0.121 and MARE = 0.023) modes. In addition, all three algorithms presented similar results for all models, whereby model 1 [$Fr = f(C_V, D_{gr}, d/R, \lambda_s)$] performed best and model 2 [$Fr = f(C_V, D_{gr}, D^2/A, \lambda_s)$] performed the worst. It could be stated that if the d/R parameters from the "transport mode" group and the D_{gr} parameter from the "sediment" group are simultaneously used with the parameters from the "transport" and "flow resistance" groups, the outcome will be optimal.
- Comparing the performance of the FFNN-ELM method with existing equations indicates that this method (RMSE = 0.121 and MARE = 0.023) predicted the *Fr* more accurately than the best existing method, which is Ebtehaj et al.'s [8] method (RMSE = 0.556 and MARE = 0.076).
- Examining the performance of FFNN-ELM for different diameters additionally indicates that FFNN-ELM produced good results under different conditions. A change in pipe diameter does not significantly affect FFNN-ELM performance in predicting *Fr*, but it presents good results in all states. Regarding the explanations given, it can be concluded that FFNN-ELM performs well in predicting the *Fr* and can be used as an alternative method in predicting the *Fr*.

References

1. Bonakdari H, Ebtehaj I. (2014) Study of sediment transport using soft computing technique. In: 7th International Conference on Fluvial Hydraulics, RIVER FLOW 2014, Lausanne, Switzerland, 933–940. doi:10.1201/b17133-126
2. Vongvisessomjai N, Tingsanchali T, Babel MS (2010) Non-deposition design criteria for sewers with part-full flow. Urban Water J 7(1):61–77. doi:10.1080/15730620903242824
3. Bonakdari H, Ebtehaj I. (2014) Verification of equation for non-deposition sediment transport in flood water canals. In: 7th International Conference on Fluvial Hydraulics, RIVER FLOW 2014, Lausanne, Switzerland, 1527–1533. doi:10.1201/b17133-203
4. Nalluri C, Ab Ghani A (1996) Design options for self-Cleansing storm sewers. Water Sci Technol 33(9):215–220. doi:10.1016/0273-1223(96)00389-7
5. Ota JJ, Nalluri C (1999) Graded sediment transport at limit deposition in clean pipe channel. In: 28th International Association for Hydro-Environment Engineering and Research, Graz, Austria
6. Ota JJ, Nalluri C (2003) Urban storm sewer design: approach in consideration of sediments. J Hydraul Eng 129(4):291–297. doi:10.1061/(ASCE)0733-9429(2003)129:4(291)
7. Banasiak R (2008) Hydraulic performance of sewer pipes with deposited sediments. Water Sci Technol 57(11):1743–1748. doi:10.2166/wst.2008.287
8. Ebtehaj I, Bonakdari H, Sharifi A (2014) Design criteria for sediment transport in sewers based on self-cleansing concept. J Zhejiang Univ Sci-A 15(11):914–924. doi:10.1631/jzus.A1300135
9. Azmathullah HMd, Deo MC, Deolalikar PB (2005) Neural networks for estimation of scour downstream of a ski-jump bucket. J Hydraul Eng 131(10):898–908. doi:10.1061/(ASCE)0733-9429(2005)131:10(898)
10. Azmathullah HMd, Deo MC, Deolalikar PB (2006) Estimation of scour below spillways using neural networks. J Hydraul Res 44(1):61–69. doi:10.1080/00221686.2006.9521661
11. Azmathullah HMd, Deo MC, Deolalikar PB (2008) Alternative neural networks to estimate the scour below spillways. Adv Eng Softw 39(8):689–698. doi:10.1016/j.advengsoft.2007.07.004
12. Zaji AH, Bonakdari H (2014) Performance evaluation of two different neural network and particle swarm optimization methods for prediction of discharge capacity of modified triangular side weirs. Flow Meas Instrum 40:149–156. doi:10.1016/j.flowmeasinst.2014.10.002
13. Esmaeili M, Osanloo M, Rashidinejad F, Bazzazi AA, Taji M (2014) Multiple regression, ANN and ANFIS models for prediction of backbreak in the open pit blasting. Eng Comput 30(4):549–558. doi:10.1007/s00366-012-0298-2
14. Ebtehaj I, Bonakdari H, Khoshbin F, Azimi H (2015) Pareto genetic design of GMDH-type neural network for predict discharge coefficient in rectangular side orifices. Flow Meas Instrum 41:67–74. doi:10.1016/j.flowmeasinst.2014.10.016
15. Faradonbeh RS, Monjezi M, Armaghani DJ (2015) Genetic programing and non-linear multiple regression techniques to predict backbreak in blasting operation. Eng Comput. doi:10.1007/s00366-015-0404-3
16. Armaghani DJ, Mohamad ET, Hajihassani M, Yagiz S, Motaghedi H (2015) Application of several non-linear prediction

tools for estimating uniaxial compressive strength of granitic rocks and comparison of their performances. Eng Comput. doi:10.1007/s00366-015-0410-5
17. Armaghani DJ, Hasanipanah M, Mohamad ET (2015) A combination of the ICA-ANN model to predict air-overpressure resulting from blasting. Eng Comput. doi:10.1007/s00366-015-0408-z
18. Zahiri A, Dehghani AA, Azamathulla HMd (2015) Application of Gene-Expression programming in hydraulic engineering. In: Handbook of Genetic Programming Applications (pp 71–97). Springer International Publishing. doi:0.1007/978-3-319-20883-1_4
19. Bhattacharya B, Price R, Solomatine D (2007) Machine Learning Approach to Modeling Sediment Transport. J Hydraul Eng 133(4):440–450. doi:10.1061/(ASCE)0733-9429(2007)133:4(440)
20. Aytek A, Kisi O (2008) A genetic programming approach to suspended sediment modeling. J Hydrol 351:288–298. doi:10.1016/j.jhydrol.2007.12.005
21. Ab Ghani A, Azamathulla HMd (2010) Gene-expression programming for sediment transport in sewer pipe systems. J Pipeline Syst Eng Pract 2(3):102–106. doi:10.1061/(ASCE)PS.1949-1204.0000076
22. Ebtehaj I, Bonakdari H, Zaji AH, Azimi H, Sharifi A (2015) Gene expression programming to predict the discharge coefficient in rectangular side weirs. Appl Soft Comput 35:618–628. doi:10.1016/j.asoc.2015.07.003
23. Ebtehaj I, Bonakdari H (2013) Evaluation of sediment transport in sewer using artificial neural network. Eng Appl Comput Fluid Mech 7(3):382–392. doi:10.1080/19942060.2013.11015479
24. Azamathulla HMd, Ab Ghani A, Fei SY (2012) ANFIS—based approach for predicting sediment transport in clean sewer. Appl Soft Comput 12(3):1227–1230. doi:10.1016/j.asoc.2011.12.003
25. Ebtehaj I, Bonakdari H (2014) Performance evaluation of adaptive neural fuzzy inference system for sediment transport in sewers. Water Resour Manage 28(13):4765–4779. doi:10.1007/s11269-014-0774-0
26. Ebtehaj I, Bonakdari H (2015) Assessment of evolutionary algorithms in predicting non-deposition sediment transport. Urban Water J. doi:10.1080/1573062X.2014.994003
27. Roushangar K, Mehrabani FV, Shiri J (2014) Modeling river total bed material load discharge using artificial intelligence approaches (based on conceptual inputs). J Hydrol 514(6):114–122. doi:10.1016/j.jhydrol.2014.03.065
28. Bravo R, Ortiz P, Pérez-Aparicio JL (2014) Incipient sediment transport for non-cohesive landforms by the discrete element method (DEM). Appl Math Model 38(4):1326–1337. doi:10.1016/j.apm.2013.08.010
29. Ebtehaj I, Bonakdari H (2014) Comparison of genetic algorithm and imperialist competitive algorithms in predicting bed load transport in clean pipe. Water Sci Technol 70(10):1695–1701. doi:10.2166/wst.2014.434
30. Kitsikoudis V, Sidiropoulos E, Hrissanthou V (2014) Assessment of sediment transport approaches for sand-bed rivers by means of machine learning. Hydrolog Sci J. doi:10.1080/02626667.2014.909599
31. Zhang K, Lu W (2011) Automatic human knee cartilage segmentation from multi-contrast MR images using extreme learning machines and discriminative random fields. Machine learning in medical imaging. Springer, Berlin Heidelberg, pp 335–343
32. Cheng C, Tay WP, Huang GB (2012) Extreme learning machines for intrusion detection. Neural networks (IJCNN), the 2012 international joint conference on. Brisbane, Australia, IEEE, pp 1–8
33. Benoit F, Van Heeswijk M, Miche Y, Verleysen M, Lendasse A (2013) Feature selection for nonlinear models with extreme learning machines. Neurocomputing 102(15):111–124. doi:10.1016/j.neucom.2011.12.055
34. Lu X, Long Y, Zou H, Yu C, Xie L (2014) Robust extreme learning machine for regression problems with its application to wifi based indoor positioning system. In: Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop on, IEEE, 1–6
35. Duan W, Li S, Fang L (2014) Spectral–spatial hyperspectral image classification using superpixel and extreme learning machines. Pattern Recognition. Springer, Berlin Heidelberg, pp 159–167
36. Liu Z, Shao J, Xu W, Chen H, Zhang Y (2014) An extreme learning machine approach for slope stability evaluation and prediction. Nat Hazards 73(2):787–804
37. Liu Z, Shao J, Xu W, Wu Q (2014) Indirect estimation of unconfined compressive strength of carbonate rocks using extreme learning machine. Acta Geotech. doi:10.1007/s11440-014-0316-1
38. May RWP, Ackers JC, Butler D (1996) Development of design methodology for self-cleansing sewers. Water Sci Technol 33(9):195–205. doi:10.1016/0273-1223(96)00387-3
39. Ackers JC, Butler D, May RWP (1996) Design of sewers to control sediment problems. Report No. 141 CIRIA, Construction Industry Research and Information Association, London, UK
40. Ab Ghani A (1993). Sediment Transport in Sewers, Ph.D. Thesis, University of Newcastle Upon Tyne, UK
41. Annema AJ, Hoen K, Wallinga H (1994) Precision requirements for single-layer feedforward neural networks. Fourth international conference on microelectronics for neural networks and fuzzy systems. Italy, Turin, pp 145–151
42. Huang GB, Zhu QY, Siew CK (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. Proceedings of International Joint Conference on neural networks, Budapest, Hungary
43. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. Neurocomputing 70:489–501. doi:10.1016/j.neucom.2005.12.126
44. Sudheer KP, Jain SK (2003) Radial basis function neural networks for modeling stage discharge relationship. J. Hydrolog Eng 8(3):161–164. doi:10.1061/(ASCE)1084-0699(2003)8:3(161)