

Sediment transport modeling in open channels using neuro-fuzzy and gene expression programming techniques

Katayoun Kargar, Mir Jafar Sadegh Safari, Mirali Mohammadi and Saeed Samadianfard

ABSTRACT

Deposition of sediment is a vital economical and technical problem for design of sewers, urban drainage, irrigation channels and, in general, rigid boundary channels. In order to confine continuous sediment deposition, rigid boundary channels are designed based on self-cleansing criteria. Recently, instead of using a single velocity value for design of the self-cleansing channels, more hydraulic parameters such as sediment, fluid, flow and channel characteristics are being utilized. In this study, two techniques of neuro-fuzzy (NF) and gene expression programming (GEP) are implemented for particle Froude number (Fr_p) estimation of the non-deposition condition of sediment transport in rigid boundary channels. The models are established based on laboratory experimental data with wide ranges of sediment and pipe sizes. The developed models' performances have been compared with empirical equations based on two statistical factors comprising the root mean square error (RMSE) and the concordance coefficient (CC). Besides, Taylor diagrams are used to test the resemblance between measured and calculated values. The outcomes disclose that NF4, as the precise NF model, performs better than the best GEP model (GEP1) and regression equations. As a conclusion, the obtained results proved the suitable accuracy and applicability of the NF method in Fr_p estimation.

Key words | gene expression programming, neuro-fuzzy, rigid boundary channel, sediment transport, self-cleansing, urban drainage

Katayoun Kargar
Mirali Mohammadi
Department of Civil Engineering,
Urmia University,
Urmia,
Iran

Mir Jafar Sadegh Safari (corresponding author)
Department of Civil Engineering,
Yaşar University,
Izmir,
Turkey
E-mail: jafar.safari@yasar.edu.tr

Saeed Samadianfard
Department of Water Engineering,
Faculty of Agriculture,
University of Tabriz,
Tabriz,
Iran

INTRODUCTION

Sediment transportation in open channels can be categorized into two classes of loose boundary and rigid boundary channels. Sediment transport in a rigid boundary channel has been the issue of studies in recent decades for the purpose of channel design. The presence of sediment has a significant effect on the hydraulic performance of open channels. It should be considered in designing sewers, drainage and irrigation systems. The sediment deposition in the channels alters the hydraulic resistance, distribution of wall shear stress, and velocity. On the other hand, it reduces the cross-section area of the channel and, accordingly, the hydraulic capacity of the channel decreases. Sedimentation spreads contamination through the urban area and has unwelcome environmental effects. To this end, sediment transport

modeling is a challenging task in hydraulic and urban hydrology engineering.

In order to avoid the aforementioned problems, channels are designed based on the self-cleansing concept. Self-cleansing satisfies a situation in which flow has ability to remove sediment from the bottom of the channel or prevent sediment deposition (Safari *et al.* 2018). To this extent, self-cleansing criteria could be categorized into two classes of 'moving bed particles' and 'non-deposition'. Each of the two categories can be broken down into sub-categories. Incipient motion and scouring are considered as the first class of self-cleansing model. In the incipient motion model, a minimum velocity or shear stress is required to start the sediment movement, while in scouring models, based on Camp criterion, the flow velocity is essential to clean the bed. The second group includes

'non-deposition without deposited bed' and 'non-deposition with deposited bed'; and incipient deposition criteria (Ab Ghani 1993; May 1993; Safari *et al.* 2018).

Non-deposition without deposited bed is a design criterion for designing small sewers, whereas non-deposition with deposited bed is used in the case of large sewers in which a small depth of existing deposited bed reduces the channel design bed slope. Non-deposition without deposited bed models are proposed for both suspended and bed loads, while incipient deposition and non-deposition with deposited bed models are developed merely for bed load. The non-deposition without deposited bed criterion is a protective method for channel design that keeps the channel bottom clean of sediment deposits. It was stated by Ab Ghani (1993) that self-cleansing velocity is reliant on sewer size, in which larger sewers need a higher self-cleansing velocity (Safari & Danandeh Mehr 2018). Nevertheless, large sewer design based on the non-deposition without deposited bed criterion produces inefficient design results in which the channel needs a steeper slope (Ab Ghani 1993; May 1993; Safari *et al.* 2017; Safari & Shirzad 2019).

The minimum design velocity or shear stress is necessary for the channels to have a clean bed in the non-deposition condition (Mayerle *et al.* 1991). Although minimum shear stress value (1–12.6 N/m²) is used in some cases, the minimum velocity (0.3–1 m/s) is frequently agreed in many countries. In the conventional method, many important factors such as the amount and type of sediment, and sewer size, are not considered. The self-cleansing models in the literature were frequently developed for bed load sediment transport, as bed load is close to the permanent deposition condition. Available models in the literature were mostly derived through applying the multiple non-linear regression technique. Although regression models have a fairly simple structure and they are able to work with partial input data, their computation ability is not as high as soft computing techniques.

Nowadays, soft computing techniques have been used in many engineering fields and applied for the purposes of design and management practices. Soft computing techniques have attracted the interest of many researchers for sediment transport modeling in sewer systems, such as Safari *et al.* (2016, 2017), Roushangar & Ghasempour (2017), Wan Mohtar *et al.* (2018), who used the artificial neural network, particle swarm optimization algorithm and evolutionary algorithm. Safari (2019) applied the decision tree, generalized regression neural network, and multivariate adaptive regression splines for modeling of sediment transport in open channel flow. As examples from the

application of gene expression programming (GEP) and adaptive neuro-fuzzy inference system (ANFIS), Ab Ghani & Azamathulla (2011) used GEP in the case of sediment transport in sewers. Azamathulla *et al.* (2012) found the functional relationships of sediment transport in sewers using the ANFIS technique and got acceptable results. Ebtehaj & Bonakdari (2014, 2017) used GEP and ANFIS techniques to predict sediment transport for sewer systems. It is reported that GEP and ANFIS provided better results in comparison with conventional regression methods. The aforementioned studies used limited ranges of experimental data in which, in most of them, only two data sets were used for modeling.

Calculation capability of models depends on the range of experimental data and the technique applied for the model development. Therefore, in the current research, the applicability of two soft computing techniques is evaluated for modeling sediment transport in the non-deposition without deposited bed condition in rigid boundary channels and their performances are compared with empirical equations. Four data sets with extensive ranges of channel size, sediment size and concentration are utilized to improve the precision of Neuro-Fuzzy (NF) and Gene Expression Programming (GEP), as two applicable soft computing techniques, in Fr_p estimation.

MATERIALS AND METHODS

Non-deposition without deposited bed

Self-cleansing models are developed based on experimental data and can be used to compute minimum velocity or sediment concentration for a non-deposition condition. Since the 1950s, studies on the self-cleansing concept in pipe flows were carried out by Craven (1953); however, the suggested model is too simple and seems unsuitable as a useful tool for channel design. The first efficient studies on sediment transport in rigid boundary channels were performed by Pedroli (1963), who accomplished experiments in rectangular channels and concluded that the sediment transport rate rises with increasing sediment size. May (1993) studied the non-deposition condition in circular pipe channels. Mayerle *et al.* (1991) studied bed load sediment transport in circular and rectangular channels and suggested

$$\frac{V}{\sqrt{gd(s-1)}} = 14.43 C_v^{0.18} D_{gr}^{-0.14} \left(\frac{d}{R}\right)^{-0.56} \lambda^{0.18} \quad (1)$$

for a circular cross-section, where V is the flow mean velocity, g acceleration due to gravity, d sediment median size,

s relative sediment density, C_v volumetric sediment concentration, R hydraulic radius, λ channel friction factor and D_{gr} dimensionless grain size parameters described by

$$D_{gr} = \left(\frac{(s-1)gd^5}{\nu^2} \right)^{\frac{1}{5}} \quad (2)$$

in which ν is kinematic viscosity of fluid. Ab Ghani (1993) considered the effect of pipe size in the non-deposition condition and noticed that higher velocity is required for large pipes. Ab Ghani (1993) examined the effect of pipe size and roughness and recommended a model based on experimental data, proposing

$$\frac{V}{\sqrt{gd(s-1)}} = 1.83C_v^{0.25}D_{gr}^{-0.10} \left(\frac{R}{d} \right)^{0.69} \lambda^{-0.04} \quad (3)$$

as a bed load self-cleansing model. It is found that the self-cleansing velocity increases when sediment concentration, channel size, and roughness increase. Vongvisessomjai et al. (2010) investigated the non-deposition condition in circular channels for both suspended and bed loads. The model for bed load according to the above-mentioned condition is expressed as follows

$$\frac{V}{\sqrt{gd(s-1)}} = 4.31C_v^{0.226} \left(\frac{d}{R} \right)^{-0.616} \quad (4)$$

Available models in the literature were developed applying non-linear regression analysis. Therefore, modeling sediment transport in open channels utilizing recent soft computing techniques seems to be quite helpful in providing powerful tools for particle Froude number estimation. To this end, as a primary application of the soft computing technique for sediment transport in sewer pipes, Ab Ghani & Azamathulla (2011) applied GEP for estimation of particle

Froude number and extracted a formula as

$$\frac{V}{\sqrt{gd(s-1)}} = \left(0.014 - \frac{\frac{-0.41}{R}}{\frac{d}{d}} \right) - \left(-0.411 - \frac{0.014}{\lambda} \right) + \frac{\lambda + \left(\frac{C_v - 5.91}{5.91} \right) \left(\frac{\lambda}{D_{gr}} \right)}{\lambda} + \lambda - \left(8.34\lambda D_{gr} \frac{R}{d} \sqrt{\lambda} \right) \quad (5)$$

Ab Ghani & Azamathulla (2011) demonstrated that the GEP-based equation provides better performance in comparison with conventional regression models.

Most of the studies on the application of soft computing techniques to sediment transport in rigid boundary channels used two or three data sets. To this end, this study utilizes four data sets, applying two powerful techniques of NF and GEP for model development.

Experimental data

In the current study, four data sets of May (1993), Ab Ghani (1993), Mayerle (1988), and Vongvisessomjai et al. (2010) are implemented for non-deposition without deposited bed channels (Table 1). The number of data taken from the above-mentioned studies is 375 observations. Mayerle (1988) carried out experiments in a circular channel of 152 mm diameter using four sediment sizes ranging from 0.5 to 8.74 mm. Mayerle (1988) studied rectangular channels with two different widths of 462 and 312 mm and five various sediment sizes ranging from 0.5 to 5.22 mm. May (1993) performed experiments in a 450 mm diameter circular channel using one sediment size of 0.73 mm.

Table 1 | Ranges of experimental data

	<i>D</i> or <i>W</i> (mm)	<i>d</i> (mm)	λ	<i>Y</i> (mm)	<i>V</i> (m/s)	C_v (ppm)
Mayerle (1988)	<i>D</i> = 152	0.50–8.74	0.016–0.034	28–122	0.37–1.10	20–1,275
	<i>W</i> = 311–462	0.50–5.22	0.011–0.025	31–111	0.41–1.04	14–1,568
May (1993)	<i>D</i> = 450	0.73	0.014–0.018	222–338	0.50–1.22	2–38
Ab Ghani (1993)	<i>D</i> = 154–450	0.46–8.30	0.013–0.048	24–342	0.24–1.21	4–1,450
Vongvisessomjai et al. (2010)	<i>D</i> = 100–150	0.20–0.43	0.034–0.053	30–60	0.24–0.63	4–90

D: circular channel diameter; *W*: rectangular channel bed width; *d*: sediment median size; λ : channel friction factor; *Y*: flow depth; *V*: flow mean velocity; C_v : sediment volumetric concentration.

Ab Ghani (1993) conducted experiments in circular channels having diameters of 154, 305, and 450 mm using seven different sediments with sizes ranging from 0.46 to 8.3 mm. In the experiments of Vongvisessomjai *et al.* (2010), pipes with diameters of 100–150 mm with sediment sizes ranging from 0.2 to 0.43 mm were adapted.

Model development

Several parameters such as fluid, flow, channel and sediment features are important for modeling sediment transport in open channels. For this purpose, flow velocity (V), hydraulic radius (R), acceleration due to gravity (g), fluid kinematic viscosity (ν) and density (ρ), sediment density (ρ_s), sediment median size (d), volumetric sediment concentration (C_v) and channel friction factor (λ) are considered for model development. Reviewing the models available in the literature, they can be stated as

$$\frac{V}{\sqrt{gd(s-1)}} = f\left(C_v, D_{gr}, \frac{d}{R}, \lambda\right) \quad (6)$$

The left hand side of Equation (6) is known as the particle Froude number (Fr_p). Due to Fr_p consisting of flow mean velocity as the desired variable for channel design, in most of the studies in the literature Fr_p is selected as the dependent parameter. Therefore, Fr_p is considered as the output of the model and the parameters given on the right side of Equation (6) are considered as independent parameters and the input to the model.

The model development procedure comprises training and testing stages. To this end, data were divided into two parts. Among 375 experimental data, 300 data were randomly selected and used for training and the remaining 75 experiments were used for testing the models. The training part discovers probable models among the independent and dependent variables though, in the testing step, the overview capability of the models is estimated with the unused data set.

Neuro-fuzzy

Jang (1993) introduced the adaptive neuro-fuzzy inference system for the first time. ANFIS is more prevalent in hydrological programs and it incorporates the privileges of both ANN and the fuzzy inference system. Despite many problems for ANN and ANFIS in dealing with non-constant data, both of them are widely used to predict hydrological

variables. NF systems utilize the learning algorithm of neural networks to adapt their rule-base parameters. Most fuzzy inference systems are divided into three types according to 'if-then' rules: the Mamdani system, Sugeno system and Tsukamoto system. In an organized and systematic method, ANFIS creates unknown fuzzy rules by using a set of input and output data. ANFIS uses different layers and node functions for learning and regulating fuzzy systems. The hybrid learning algorithm contains front and rear passes to reduce calculated errors and training parts. By calculating the minimum square error, the assumed parameters are fixed while the corresponding parameters are updated. Later, in the rear passes, the corresponding parameters are fixed, and the assumed parameters are updated by using the gradient descent algorithm (Khayyam *et al.* 2012).

Gene expression programming (GEP)

Ferreira (2001a, 2001b) introduced GEP as an evolutionary algorithm. It includes features of Genetic Programming and Genetic Algorithm. GEP computer programs are encoded in linear chromosomes. Chromosomes with dissimilar sizes and forms in GEP can code in a simple graph (Ferreira 2001a, 2001b). Like other evolutionary methods, in GEP the procedure starts by randomization of early population chromosomes. Then, each single chromosome of the early population is estimated using a fitness function and its value. Diverse fitness functions can be used in the GEP model, such as mean squared error (MSE), root mean square error ($RMSE$), relative standard error (RSE), and root relative squared error $RRSE$ (Ferreira 2001a). The best chromosomes are more likely to be transported to the subsequent generation. After choosing the best chromosomes, these acts are repeated by genetic operators with some changes. In order to modify the GEP, genetic operators such as transposition, mutations, recombination and inversion are used. GEP is a complete genotype/phenotype system, with a genotype that is completely isolated from the phenotype, while in the GP genotype and phenotype in a simple replicator system are combined together (Ferreira 2001a, 2001b). In the GEP method, the formation of genetic diversity becomes too simple because genetic operators work at the chromosome level. Additionally, GEP includes an exclusive and multigenic nature that lets more intricate programs consisting of numerous sub-programs be completed. Consequently, GEP exceeds the old GP system. More details about the structure and application of GEP and GP can be found in Danandeh Mehr *et al.* (2018).

Performance evaluation

Validation of the model is essential for the credibility of the established model. For this purpose, the performance of developed models is evaluated in terms of several criteria; the RMSE and concordance coefficient (CC), Taylor diagrams and scatter plots. RMSE shows contradistinction among the calculated and measured Fr_p by

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Fr_p^m - Fr_p^c)^2}{n}} \quad (7)$$

in which n is the number of data and Fr_p^m and Fr_p^c are, respectively, the measured and calculated particle Froude numbers. CC is the concordance among the measured and calculated values and it is more acceptable when it approaches 1 and can be expressed as

$$CC = \frac{2r\sigma_m\sigma_c}{\sigma_m^2 + \sigma_c^2 + (\bar{Fr}_p^m - \bar{Fr}_p^c)^2} \quad (8)$$

where σ_m and σ_c are standard deviation, \bar{Fr}_p^m and \bar{Fr}_p^c are the average of the measured and calculated Fr_p , respectively, and r is the correlation coefficient. Furthermore, Taylor diagrams (Taylor 2001) were used to check the accuracy of the mentioned models in Fr_p estimation. It is remarkable that Taylor suggested a diagram in which measured parameters and some characteristics of the model are summed up, coincidentally. Surprisingly, Taylor diagrams utilize several points on a polar plot for comparing the accuracy of measured and calculated values. In these diagrams, the correlation coefficient and normalized standard deviation are represented by azimuth angle and radial distances from the base point, respectively (Taylor 2001).

RESULT AND DISCUSSION

Four scenarios shown in Table 2 are considered for estimation of Fr_p in this study. Due to the fact that d/R

Table 2 | Input and output parameter combinations of studied models

Model number	Input parameters	Output parameter
1	$C_v, D_{gr}, \frac{d}{R}, \lambda$	Fr_p
2	$C_v, \frac{d}{R}, \lambda$	Fr_p
3	$C_v, \frac{d}{R}$	Fr_p
4	$C_v, D_{gr}, \frac{d}{R}$	Fr_p be completed

and C_v are the most important independent parameters in sediment transport in sewers; they are used in all input combinations. Moreover, according to the alteration of the sediment size and density, as well as velocity and Fr_p changes dependent upon numerous situations such as particle properties, turbulence intensity, and sediment concentration, these variables are therefore considered in input combinations.

Application of neuro-fuzzy

Different NF designs were applied using code developed in MATLAB and the best model constructions are specified. There are several types of membership functions, like shape-shaped, trapezoids, Gaussian, and sigmoid. Through minimizing the target function, the optimal parameters of the membership functions (MFs) are determined based on RMSE between calculated and measured values. In the current computations, different numbers and various types of MFs are examined for different input parameters. Additionally, to reduce the complexity of the NF models, lower MF numbers are selected. Moreover, it should be noticed that the optimum structure of NF models is obtained after performing numerous trials for different input combinations (Table 3).

After the trial-error procedure, the result obtained from Table 3 showed that the optimum NF1 model implemented (3, 3, 3, 3) gauss2 MFs for the inputs of $C_v, D_{gr}, d/R$ and λ with 100,000 generations. Additionally, (2, 2, 2) and (4, 4, 4) gauss MFs are selected as the optimum structures of NF2 and NF3 with the same number of generations as NF1. Finally, (4, 4) dsig MFs with 100,000 generations and using input parameters of NF4 produced more accurate predictions of measured values comparing with other structures.

Application of gene expression programming

There are some significant stages in the arrangement of the GEP model; at the first step, choosing a fitness function is essential. Therefore, in this study, the RRSE function is used. The next stage is choosing the collection of terminals and functions to create the chromosomes. In this study, after performing different functions, 14 functions were designated, comprising (+, -, ×, /) and (Ln, Sin, Exp, Cos, Power, Sqrt, Tan, Arcsin, Arccos, Arctan). In order to achieve an uncomplicated and sensible GEP model, the functions were selected based on their coherence to the quiddity of the problem.

In this step, first, one gene and two head sizes were used. After that, by increasing the number of genes, the accuracy and strength of the model were remarkably enhanced and

Table 3 | The RMSE values of NF models with different activation functions and MFs in the layers

No. of model	Activation function	No. of MFs in the layers	RMSE	No. of model	Activation function	No. of MFs in the layers	RMSE
1	(3 3 3 3)	gauss2	1.285	3	(2 2)	gauss2	1.340
	(4 4 4 4)	gauss2	1.740		(3 3)	gauss2	1.600
	(2 2 2 2)	gauss2	1.536		(4 4)	gauss2	1.550
	(2 2 2 2)	gauss	1.644		(2 2)	gauss	1.530
	(3 3 3 3)	gauss	1.308		(3 3)	gauss	1.500
	(4 4 4 4)	gauss	1.736		(4 4)	gauss	1.434
	(2 2 2 2)	dsig	1.439		(2 2)	dsig	1.366
	(3 3 3 3)	dsig	1.790		(3 3)	dsig	1.330
	(4 4 4 4)	dsig	10.000		(4 4)	dsig	1.323
	(2 2 2 2)	gbell	1.610		(2 2)	gbell	1.614
	(3 3 3 3)	gbell	1.858		(3 3)	gbell	1.469
	(4 4 4 4)	gbell	2.310		(4 4)	gbell	1.454
	(2 2 2 2)	psig	1.390		(2 2)	psig	1.366
	(3 3 3 3)	psig	1.880		(3 3)	psig	1.333
	(4 4 4 4)	psig	2.190		(4 4)	psig	1.323
2	(2 2 2)	gauss2	1.480	4	(2 2 2)	gauss2	1.640
	(3 3 3)	gauss2	1.486		(3 3 3)	gauss2	1.190
	(4 4 4)	gauss2	1.540		(4 4 4)	gauss2	1.274
	(2 2 2)	gauss	1.349		(2 2 2)	gauss	1.360
	(3 3 3)	gauss	1.600		(3 3 3)	gauss	1.374
	(4 4 4)	gauss	1.660		(4 4 4)	gauss	1.038
	(2 2 2)	dsig	1.467		(2 2 2)	dsig	1.297
	(3 3 3)	dsig	1.387		(3 3 3)	dsig	1.389
	(4 4 4)	dsig	1.428		(4 4 4)	dsig	2.830
	(2 2 2)	gbell	2.100		(2 2 2)	gbell	1.470
	(3 3 3)	gbell	1.860		(3 3 3)	gbell	1.304
	(4 4 4)	gbell	2.030		(4 4 4)	gbell	1.150
	(2 2 2)	psig	1.467		(2 2 2)	psig	1.297
	(3 3 3)	psig	1.467		(3 3 3)	psig	1.200
	(4 4 4)	psig	1.375		(4 4 4)	psig	2.830

3 genes were selected. After some trials, 8 is considered as the head size and, according to the trials, a head size of more than eight did not ameliorate the performance of the model. Therefore, the number of chromosomes was designated as 30, head size as 8 and number of genes as 3. After completing the architecture of the chromosome, the genetic operators used in this study are mutations, inversion, transposition, and recombination, and their rates are shown in Table 4. Moreover, addition is selected as the linking function. The simulation of the model begins after the parameters are determined. As a result, GEP1 gives a

formula that can be used in sediment transport modeling as

$$\frac{V}{\sqrt{gd(s-1)}} = \frac{3.05C_v^{0.16}}{\text{Arctan}\left(\text{Arctan}\left(\sqrt{\frac{d}{R}}\right)\right) + \text{Arctan}(3.41 - \text{Ln}(D_{gr})) + \text{Arctan}\left(\text{Tan}\left(\left(8.37 - 7.99\lambda + \frac{d}{R}\lambda\right)^2\right)\right) + \text{Ln}\left(\left(\left(\frac{d}{R}\right)^3\right)^{2\lambda}\right)} \quad (9)$$

Table 4 | Parameters of the optimized GEP model

Description of parameter	Setting of parameter
Mutation rate	0.044
Inversion rate	0.1
One point and two point recombination rate	0.3
Gene recombination and transportation rate	0.1

Comparison of models

Developed models in this study applying NF and GEP techniques are compared with Mayerle *et al.* (1991),

Ab Ghani (1993) and Vongvisessomjai *et al.* (2010) conventional regression models as well as a GEP-based model of Ab Ghani & Azamathulla (2011).

Table 5 shows the comparison of the models based on *RMSE* and *CC* values on the test data set. The choice of an appropriate combination of input parameters is an important point in the modeling. The influences of various formations of parameters are evaluated and results are shown in Table 5. It is seen that model of GEP1 with *RMSE* of 1.245 and *CC* of 0.837 has the best performance among other GEP models and NF4 with *RMSE* of 1.039 and *CC* of 0.878 provides more accurate predictions of Fr_p in comparison with other NF models. Among the conventional regression equations, Ab Ghani (1993) has the best result with *RMSE* of 1.460 and *CC* of 0.749. The Ab Ghani & Azamathulla (2011) GEP-based equation outperforms Mayerle *et al.* (1991) and Vongvisessomjai *et al.* (2010) conventional regression models in terms of *RMSE* and *CC*. The Ab Ghani (1993) and Ab Ghani & Azamathulla (2011) models can compete with models developed in the study in terms of accuracy for particle Froude number estimation. It has to be emphasized that for both the GEP-based equations of Ab Ghani & Azamathulla (2011) and the model developed in this study (Equation (9)), real data (not normalized data) should be used. On the other hand, the sediment volumetric concentration (C_v) is in ppm (parts per million) in Ab Ghani & Azamathulla (2011) while, similar to the conventional regression models in the literature, the real value of C_v should be used in application of Equation (9).

Evaluation of models in terms of goodness-of-fit with scatter plots of the calculated and measured Fr_p in non-deposition condition with different input parameters

are shown in Figure 1. It is observed that Fr_p calculated by GEP1 and NF4 models matches appropriately with the measured Fr_p for the non-deposition condition. It is imperative to note that although NF4 and GEP1 slightly overestimated the Fr_p values, their results are close to the best-fit line in comparison with other models. Additionally, even though models of NF1, GEP2, GEP3, GEP4 overestimate Fr_p in the non-deposition condition, a few data remain under the bisector line. Also, a slight underestimation is seen in the case of Fr_p values, which is more noticeable for NF2 and NF3 models. The Ab Ghani (1993) and Ab Ghani & Azamathulla (2011) models slightly underestimate while the Vongvisessomjai *et al.* (2010) and Mayerle *et al.* (1991) models significantly overestimate Fr_p with a wide scatter. The slight overestimation of GEP1 and NF4 models can be linked to the ranges of experimental data used in this study. In fact, the credibility of a sediment transport model depends on the ranges of experimental data used in the modeling procedure. Behind the higher credibility of the developed model using wide ranges of experimental data, it may cause an adverse effect on the accuracy of the model and slight overestimation or underestimation of the developed model. Figure 2 shows a three-dimensional bar graph of the *RMSE* and *CC* created by NF and GEP models. The results for the most and least precisely predicted models are presented. As can be seen in Figure 2, NF4 has the lowest *RMSE* and highest *CC* compared to the other NF models. In addition, GEP1 has the best performance compared to the other GEP models. This indicates that the NF and GEP models are able to achieve an accurate performance in predicting Fr_p .

Furthermore, Taylor charts were used to examine standard deviation and correlation values among calculated and measured Fr_p for the NF and GEP models with different input parameters. Taylor diagrams for the above-mentioned models are shown in Figure 3. The length of the space from the reference point (a circular point) to each point is defined as centered *RMSE* (Taylor 2001). Therefore, the most accurate model has the minimum distance between the circular point and its correspondent point. According to Figure 3, GEP1 (a rectangular point) and NF4 (a triangular point) with several input combinations offered the best estimates of Fr_p in the non-deposition condition.

Table 5 | Presentation of the NF and GEP models for the Fr_p based on *RMSE* and *CC*

Method		Statistical parameter	
		<i>RMSE</i>	<i>CC</i>
GEP	GEP1	1.25	0.84
	GEP2	1.39	0.81
	GEP3	1.52	0.74
	GEP4	1.33	0.81
	Ab Ghani & Azamathulla (2011)	1.73	0.63
NF	NF1	1.29	0.82
	NF2	1.35	0.71
	NF3	1.32	0.72
	NF4	1.04	0.88
Regression equations	Ab Ghani (1993)	1.46	0.75
	Mayerle <i>et al.</i> (1991)	3.72	0.52
	Vongvisessomjai <i>et al.</i> (2010)	2.29	0.68

CONCLUSIONS

The applicability of NF and GEP to sediment transport modeling as self-cleansing design criteria are examined based on

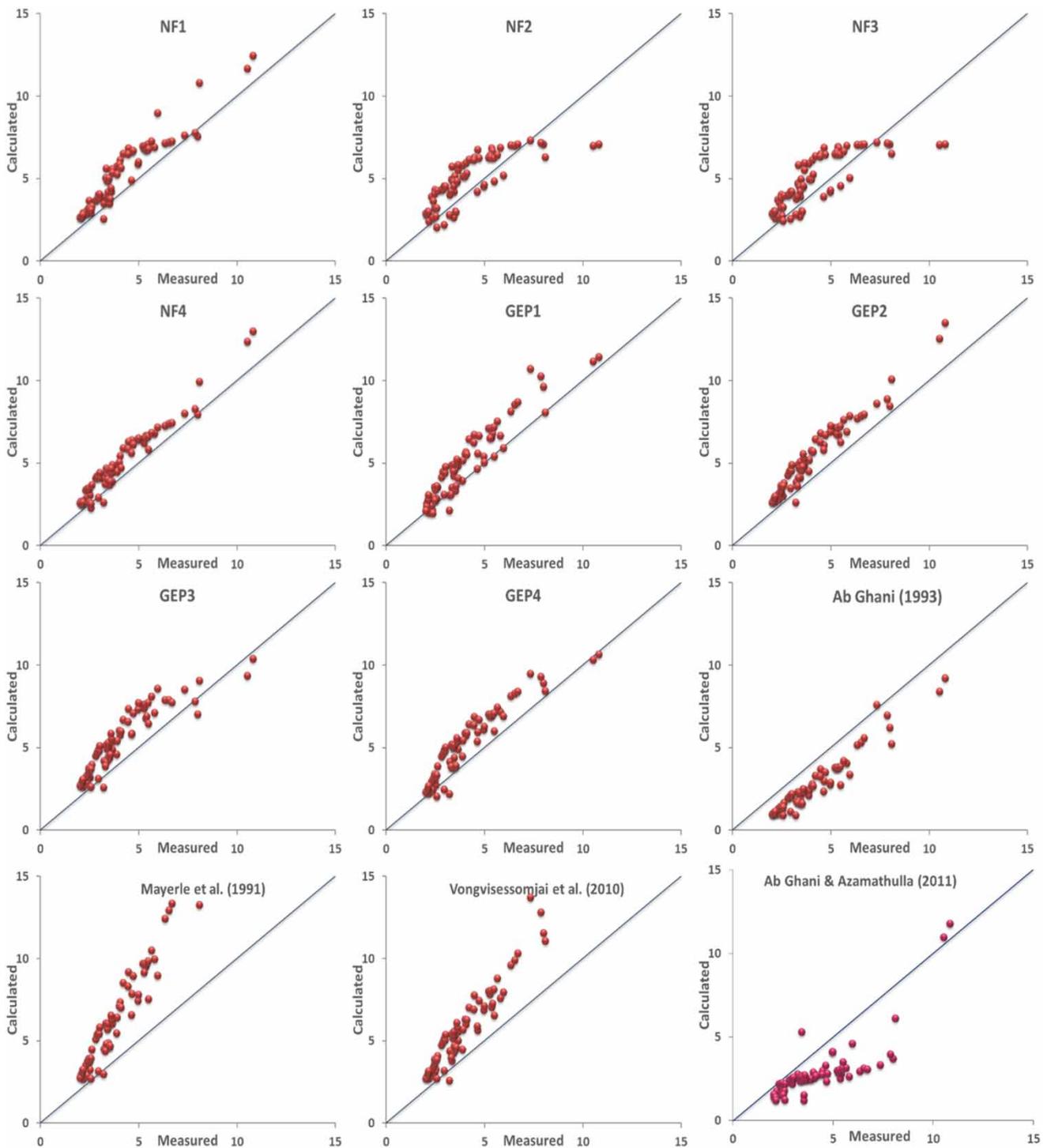


Figure 1 | Comparison of models in terms of measured and calculated Fr_p .

Fr_p estimation in the non-deposition without deposited bed condition in rigid boundary channels. In the present study, for model development, different characteristics of channel, sediment, fluid and flow were used. Through the modeling

process, various input parameters were considered to find the best result for Fr_p estimation. In this study, Fr_p , d/R and C_v were selected as the most important parameters for improving the accuracy of the model. Thus, two mentioned

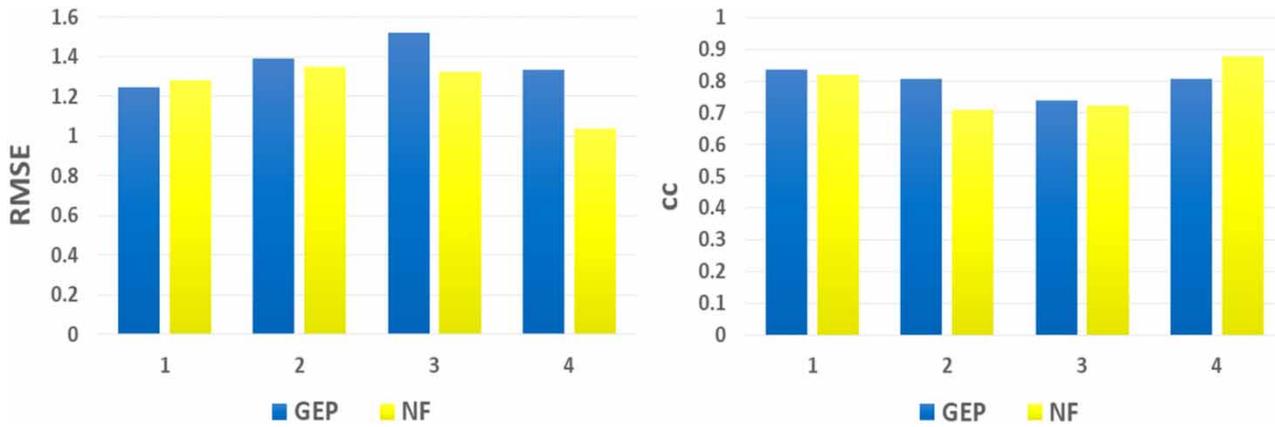


Figure 2 | Two-dimensional bar graph of the $RMSE$ and CC values for different input combinations.

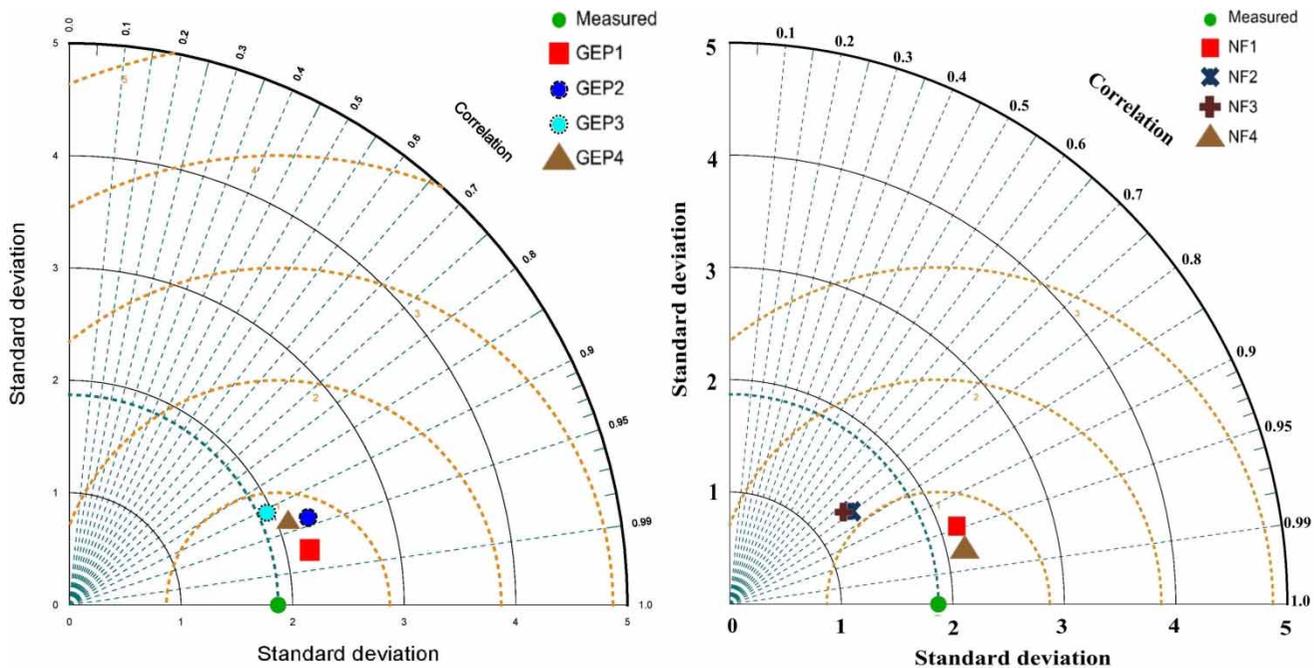


Figure 3 | Taylor diagrams of calculated Fr_p values in test period.

parameters are utilized in all input combinations. The obtained results indicate that NF4 with $RMSE$ of 1.039 and CC of 0.878, and GEP1 with $RMSE$ of 1.245 and CC of 0.837 have more suitable accuracy compared with other models in Fr_p predictions. Moreover, Ab Ghani (1993), with $RMSE$ of 1.460 and CC of 0.749 provided more accurate predictions of Fr_p in comparison with other studied regression equations. In other words, the statistical analyses indicate that the performances of NF and GEP models are superior to the regression equations. Moreover, the most accurate NF model (NF4) created slightly more precise predictions than the best GEP model (GEP1). Conclusively, it is

inferred that the implemented soft computing models can successfully be applied for Fr_p estimation for self-cleansing channel design.

REFERENCES

- Ab Ghani, A. 1993 *Sediment Transport in Sewers*. PhD thesis, University of Newcastle upon Tyne, UK.
- Ab Ghani, A. & Azamathulla, H. M. 2011 *Gene-expression programming for sediment transport in sewer pipe systems*. *Journal of Pipeline Systems Engineering and Practice* 2 (3), 102–106.

- Azamathulla, H. M., Ab Ghani, A. & Fei, S. Y. 2012 ANFIS-based approach for predicting sediment transport in clean sewer. *Applied Soft Computing* **12** (3), 1227–1230.
- Craven, J. P. 1953 The transportation of sand in pipes – full pipe flow. In: *Proc., 5th Hydraulics Conf, Ames, IA*. Iowa State University.
- Danandeh Mehr, A., Nourani, V., Kahya, E., Hrnjica, B., Sattar, A. M. A. & Yaseen, Z. M. 2018 Genetic programming in water resources engineering: a state-of-the-art review. *Journal of Hydrology* **566**, 643–667.
- Ebtehaj, I. & Bonakdari, H. 2014 Performance evaluation of adaptive neural fuzzy inference system for sediment transport in sewers. *Water Resources Management* **28** (13), 4765–4779.
- Ebtehaj, I. & Bonakdari, H. 2017 No-deposition sediment transport in sewers using of gene expression programming. *Soft Computing in Civil Engineering* **1** (1), 29–53.
- Ferreira, C. 2001a Gene expression programming in problem solving. In: *6th Online World Conference on Soft Computing in Industrial Applications* (invited tutorial).
- Ferreira, C. 2001b Gene expression programming: a new adaptive algorithm for solving problems. *Complex Systems* **13** (2), 87–129.
- Jang, J.-S. R. 1993 ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics* **23** (3), 665–685.
- Khayyam, H., Nahavandi, S. & Davis, S. 2012 Adaptive cruise control look-ahead system for energy management of vehicles. *Expert Systems with Applications* **39** (3), 3874–3885.
- May, R. W. P. 1993 *Sediment Transport in Pipes and Sewers with Deposited Beds*. Technical Report. Hydraulic Research Ltd, Wallingford, UK. Report SR 320.
- Mayerle, R. 1988 Sediment transport in rigid boundary channels. PhD thesis, University of Newcastle upon Tyne, Newcastle upon Tyne, Northumbria, UK.
- Mayerle, R., Nalluri, C. & Novak, P. 1991 Sediment transport in rigid bed conveyances. *Journal of Hydraulic Research* **29** (4), 475–495.
- Pedroli, R. 1963 *Bed Load Transportation in Channels with Fixed and 1290 52 Smooth Inverts*. PhD thesis, Scuola Politecnica Federale.
- Roushangar, K. & Ghasempour, R. 2017 Estimation of bedload discharge in sewer pipes with different boundary conditions using an evolutionary algorithm. *International Journal of Sediment Research* **32** (4), 564–574.
- Safari, M. J. S. 2019 Decision tree (DT), generalized regression neural network (GR) and multivariate adaptive regression splines (MARS) models for sediment transport in sewer pipes. *Water Science and Technology* **79** (6), 1113–1122.
- Safari, M. J. S. & Danandeh Mehr, A. 2018 Multigene genetic programming for sediment transport modeling in sewers for conditions of non-deposition with a bed deposit. *International Journal of Sediment Research* **33** (3), 262–270.
- Safari, M. J. S. & Shirzad, A. 2019 Self-cleansing design of sewers: definition of the optimum deposited bed thickness. *Water Environment Research* **91** (5), 407–416.
- Safari, M. J. S., Aksoy, H. & Mohammadi, M. 2016 Artificial neural network and regression models for flow velocity at sediment incipient deposition. *Journal of Hydrology* **541**, 1420–1429.
- Safari, M. J. S., Shirzad, A. & Mohammadi, M. 2017 Sediment transport modeling in deposited bed sewers: unified form of May's equations using the particle swarm optimization algorithm. *Water Science and Technology* **76** (4), 992–1000.
- Safari, M. J. S., Mohammadi, M. & Ab Ghani, A. 2018 Experimental studies of self-cleansing drainage system design: a review. *Journal of Pipeline Systems Engineering and Practice* **9** (4), 04018017.
- Taylor, K. E. 2001 Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres* **106**, 7183–7192.
- Vongvisessomjai, N., Tingsanchali, T. & Babel, M. S. 2010 Non-deposition design criteria for sewers with part-full flow. *Urban Water Journal* **7** (1), 61–77.
- Wan Mohtar, W. H. M., Afan, H., El-Shafie, A., Bong, C. H. J. & Ab Ghani, A. 2018 Influence of bed deposit in the prediction of incipient sediment motion in sewers using artificial neural networks. *Urban Water Journal* **15** (4), 296–302.

First received 17 April 2019; accepted in revised form 1 July 2019. Available online 5 July 2019