

New stochastic modeling strategy on the prediction enhancement of pier scour depth in cohesive bed materials

Ahmad Sharafati, Ali Tafarajnoruz and Zaher Mundher Yaseen

ABSTRACT

Scouring around the piers, especially in cohesive bed materials, is a fully stochastic phenomenon and a reliable prediction of scour depth is still a challenging concern for bridge designers. This study introduces a new stochastic model based on the integration of Group Method of Data Handling (GMDH) and Generalized Likelihood Uncertainty Estimation (GLUE) to predict scour depth around piers in cohesive soils. The GLUE approach is developed to estimate the related parameters whereas the GMDH model is used for the prediction target. To assess the adequacy of the GMDH-GLUE model, the conventional GMDH and genetic programming (GP) models are also developed for evaluation. Several statistical performance indicators are computed over both the training and testing phases for the prediction accuracy validation. Based on the attained numerical indicators, the proposed GMDH-GLUE model revealed better predictability performance of pier scour depth against the benchmark models as well as several gathered literature studies. To provide an informative comparison among the proposed techniques (i.e. GMDH-GLUE, GMDH, and GP models), an improvement index (*IM*) is employed. Results indicated that the GMDH-GLUE model achieved $IM_{train} = 6\%$ and $IM_{test} = 3\%$, demonstrating satisfying performance improvement in comparison with the previously proposed GMDH model.

Key words | cohesive bed materials, GLUE model, pier scour depth prediction, stochastic model

Ahmad Sharafati
Department of Civil Engineering, Science and
Research Branch,
Islamic Azad University,
Tehran,
Iran

Ali Tafarajnoruz
Dipartimento di Ingegneria Civile,
Università della Calabria,
Cubo 42B, Rende,
Italy

Zaher Mundher Yaseen (corresponding author)
Sustainable Developments in Civil Engineering
Research Group, Faculty of Civil Engineering,
Ton Duc Thang University,
Ho Chi Minh City,
Vietnam
E-mail: yaseen@tdtu.edu.vn

INTRODUCTION

Since the 1950s, several field surveys have clarified that scour development around piers is the main cause of bridge collapses across the globe and, due to the importance of this issue, a large amount of research has been performed to study the pier scour phenomenon (Breusers *et al.* 1977; Chiew & Melville 1987; Ferraro *et al.* 2013). Briefly, construction of a bridge on a river alters the flow pattern, resulting in the development of two main types of vortices (i.e. horseshoe and wake vortices). Horseshoe vortices extend like a necklace at the base of the pier front, while the development and shedding of the wake vortices occur downstream of the pier. Both experimental and numerical studies on the turbulent flow field around the piers confirmed that the shape,

dimensions and location of these vortices may significantly vary during time (Dargahi 1990; Kirkil *et al.* 2008). The scouring process initiates when such stochastic phenomena interact with bed sediment. It is worthwhile noting that the sediment transport is also stochastic and can be estimated through some statistical approaches (Dodaro *et al.* 2014, 2016). Therefore, it is quite reasonable to consider the scouring process entirely as a complex stochastic phenomenon.

Complicated scour development is usually experienced when the bridge piers are found in cohesive soils. For granular non-cohesive soil particles, the submerged density of the soil particles and the gravity forces are the major resistance agents to the sediment motion, while in cohesive

sediments the physico-chemical characteristics play a critical role. In fact, scouring occurs when the fluid shear on the erodible bed exceeds the physico-chemical forces between the cohesive bed particles and the submerged unit weight of the sediment particles (Rambabu *et al.* 2003). Therefore, scour depth around a pier in cohesive bed materials is associated with higher uncertainty than in non-cohesive ones (Chang *et al.* 2004; Firat & Gungor 2009).

The use of semi-empirical regression-based equations is the most common method applied for scour estimation at bridge piers. In these equations, the ultimate pier scour depth is defined as a function of flow properties and sediment characteristics (Breusers *et al.* 1977; Melville 1997; Arneson *et al.* 2012). In spite of extensive laboratory studies, owing to the existence of a complex flow field around the pier foundation and unknown influencing parameters on scour development, the regression-based equations have not always provided promising predictions (Gaudio *et al.* 2010, 2013; Tafarojnoruz 2012). This is due to the high stochasticity phenomena of the scouring problem. Recently, the implementation of artificial intelligence (AI) methodologies has been considered as a reliable alternative to the existing conventional equations to solve complicated problems when a simple equation may not represent the whole complexity of a phenomenon (Najafzadeh *et al.* 2017, 2018; Sharafati *et al.* 2019a, 2019c). Regarding the pier scour depth issue, researchers have already reported several AI models such as artificial neural networks (ANNs) (Toth & Brandimarte 2011), genetic programming (GP) (Azamathulla *et al.* 2010), linear genetic programming (LGP) (Güven *et al.* 2009), group method of data handling (GMDH) (Najafzadeh & Barani 2011) and model tree (MT) (Etemad-Shahidi & Ghaemi 2011; Etemad-Shahidi *et al.* 2015) in which better modeling methodologies are demonstrated in comparison with the empirical equations. However, several limitations still exist, particularly the stochastic intrinsic problem. Hence, some prediction inaccuracies relating to the stochastic issues still exist and hydraulic scientists need to tackle these problems to improve predictions.

It is noteworthy that none of the above approaches take into account the stochastic behavior of the scouring process. In fact, researchers assume scour development is a deterministic phenomenon. On the contrary, owing to the random nature and complexity of the scouring process, expected uncertainties may lead to an unavoidable risk in pier scour

estimation during foundation design (Yanmaz 2002). Therefore, the use of a probabilistic framework to evaluate the likelihood of achieving different scour depths is essential to estimate failure probabilities due to excessive scour depth (Johnson & Dock 1998). Variable uncertainties that may originate due to the difficulties of an accurate measurement should be predicted, aimed at quantifying the risk of pier failure (Yanmaz 2001).

Quantifying the uncertainty of model parameters can be accomplished through several approaches. Among them, generalized likelihood uncertainty estimation (GLUE), developed in the 1990s (Beven & Binley 1992), has been adopted by several researchers for uncertainty analysis in hydraulics and water resource engineering. For example, it was recently successfully used to assess sediment yield in a watershed (Ayele *et al.* 2017), analysis of the groundwater transport time (Zell *et al.* 2018) and increasing the accuracy level in the prediction of pipeline scour depth (Sharafati *et al.* 2018). In this method, it is not necessary to maximize or minimize the objective function, while the information concerning various parameter sets is derived from the likelihood measure indices. In the following sections, this technique is described in detail and evaluated for the prediction of scour depth around a pier in cohesive bed materials.

This investigation aims to reanalyze and revise the coefficients and exponents of recently developed scour depth equations of a cylinder founded in cohesive soils taking into account the stochastic methodology GLUE. The selected equations, as described in the following section, were proposed by Najafzadeh *et al.* (2013) based on the the GMDH network. Moreover, in this study, a new equation is extracted utilizing the GP approach and its performance and the accuracy of the novel stochastic equation in estimation of the scour depth is evaluated using common statistical indices during the training and testing phases.

METHODS

Influencing parameters and scour depth prediction formulations

In addition to the effective parameters on scouring phenomena in non-cohesive bed sediment, scour depth evolution in

cohesive sediment particles may be also affected by the initial water content of soil, clay percentage, and bed shear strength (Briaud *et al.* 1999; Ansari *et al.* 2002; Rambabu *et al.* 2003; Debnath & Chaudhuri 2010). The parameters within the following equation are commonly considered as effective parameters on maximum scour depth, d_s , around a cylinder founded in a cohesive bed sediment:

$$d_s = g_1(\rho, \mu, d_{50}, \gamma, S, U, D, C_p, IWC, g) \quad (1)$$

where d_s denotes maximum scour depth, g_1 stands for an unknown function, and the dependent variables $\rho, \mu, d_{50}, \gamma, S, U, D, C_p, IWC, g$ are mass density of water, water dynamic viscosity, median diameter of soil particles, flow depth, bed shear strength, mean flow velocity, pier diameter, clay percentage, initial water content, and acceleration due to gravity, respectively. After dimensional analysis, the following dimensionless groups are extracted for prediction of non-dimensional maximum scour depth d_s/D :

$$\frac{d_s}{D} = g_2\left(Re_p, \frac{\gamma}{D}, Fr_p, \frac{D}{d_{50}}, IWC, C_p, \frac{S}{\rho U^2}\right) \quad (2)$$

where g_2 represents an unknown function, $Re_p = \frac{\rho U D}{\mu}$, $Fr_p = \frac{U}{\sqrt{gD}}$, and $\frac{S}{\rho U^2}$ are, respectively, pier Reynolds number, pier Froude number and non-dimensional bed shear strength. In line with Najafzadeh *et al.* (2013), Fr_p and $\frac{\gamma}{D}$ were replaced with the flow Froude number, $Fr = \frac{U}{\sqrt{g\gamma}}$. Moreover, the influence of pier Reynolds number on scour evolution is negligible assuming that the flow around the pier is turbulent and in general it is insignificant in the prediction of ultimate pier scour depth (Tafarojnoruz *et al.* 2010). Concerning the selected data utilized in the present investigation, the magnitude of non-dimensional sediment size, $\frac{D}{d_{50}}$, is not significant (Chiew & Melville 1987). Therefore, Equation (2) may be written as expressed in Equation (3) and adopted for analysis of the

present study. g_3 is also an unknown function:

$$\frac{d_s}{D} = g_3\left(Fr, IWC, C_p, \frac{S}{\rho U^2}\right) \quad (3)$$

In this study, it was attempted to improve the traditional GMDH method in the estimation of ultimate scour depth at a pier founded in cohesive bed materials using a stochastic approach. In this way, a new hybrid model named GMDH-GLUE is introduced as a novel approach. Further, the accuracy of the newly developed model was checked by comparing it with the GP model. In this section, first the GMDH and GP methods and their formulas are described then the GLUE approach, as well as the GMDH-GLUE model, is presented.

Description of GMDH based scouring model

The GMDH technique is a system identification methodology able to model and predict the behavior of a complicated system according to certain input–output data pairs (Ivakhnenko & Ivakhnenko 1995). Its principle was developed on the basis of heuristic self-organizing taking into account some certain operations such as seeding, rearing, crossbreeding, and selection and rejection of seeds attributable to the input model parameters, and model definition (Ivakhnenko & Ivakhnenko 2000; Amanifard *et al.* 2008; Najafzadeh & Tafarojnoruz 2016). The capability of this technique in scour depth prediction has been extensively assessed. For instance, the GMDH network was successfully developed to model basin sediment yield (Garg 2015), scour depth downstream of grade control structures (Najafzadeh 2015), stable channel design (Shaghghi *et al.* 2017), discharge coefficient of cylindrical weirs (Parsaie *et al.* 2018), scouring rate under pipelines due to waves (Najafzadeh & Saberi-Movahed 2018). All these investigations and other similar studies prove the merits of the GMDH technique in the estimation of scour depth around river and marine structures.

The relationship between the output and input model parameters in GMDH methodology was built based upon a function of the following nonlinear Volterra functional series which is known as Kolmogorov-Gabor polynomial

(Farlow 1981; Ivakhnenko & Müller 1997):

$$\hat{y} = a_0 + \sum_{i=1}^n a_i z_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} z_i z_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_{ijk} z_i z_j z_k + \dots \tag{4}$$

where x stands for the input variables, a denotes the weighting coefficients and \hat{y} represents the model output. Having employed the quadratic polynomial terms, the GMDH network has the following structure (Najafzadeh et al. 2014):

$$\hat{y} = a_0 + \sum_{i=1}^n a_i z_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} z_i z_j \tag{5}$$

In the training step, the weighting coefficients of GMDH network are estimated using the standard orthogonal least square (OLS) method so that the difference between observed output (y) and model output (Equation (5)) is minimized as follows (Najafzadeh & Saberi-Movahed 2018):

$$a = (Z^T Z)^{-1} Z^T Y \tag{6}$$

$$Z = \begin{bmatrix} 1 & z_{1p} & z_{1q} & z_{1p}z_{1q} & z_{1p}^2 & z_{1q}^2 \\ 1 & z_{2p} & z_{2q} & z_{2p}z_{2q} & z_{2p}^2 & z_{2q}^2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 1 & z_{mp} & z_{mq} & z_{mp}z_{mq} & z_{mp}^2 & z_{mq}^2 \end{bmatrix} \quad p, q \in \{1, 2, \dots, n\} \tag{7}$$

$$Y = \{y_1, y_2, \dots, y_m\}^T \tag{8}$$

where m and n denote the number of observed data and input parameters, respectively.

Regarding the prediction of pier scour depth, combinations of iterative and evolutionary algorithms by means of GMDH technique may offer more reliable predictions than those resulting from traditional equations and some other soft computing techniques (Najafzadeh & Barani 2011; Najafzadeh & Azamathulla 2013). Concerning the pier scour estimation at cohesive bed soils, Najafzadeh et al. (2013) suggested the quadratic polynomial GMDH based formulas to forecast scour depth around a pier in

cohesive soils as follows:

$$\hat{y}_1 = 0.3278 + 0.0022IWC - 0.585Fr + 0.1309IWC.Fr - 48 \times 10^{-5}IWC^2 - 1.35Fr^2 \tag{9}$$

$$\hat{y}_2 = -0.2806 + 0.00288 \frac{S}{\rho U^2} + 5.34Fr - 0.0232 \frac{S.Fr}{\rho U^2} - 7.62 \times 10^{-7} \left(\frac{S}{\rho U^2} \right)^2 - 2.42133Fr^2 \tag{10}$$

$$\hat{y}_3 = 1.657 - 0.0573C_p + 2.304Fr - 0.0675C_p.Fr + 5.86 \times 10^{-4}C_p^2 + 7.834Fr^2 \tag{11}$$

$$\hat{y}_4 = 0.1303 - 0.128\hat{y}_1 + 0.507\hat{y}_3 - 0.5687\hat{y}_1\hat{y}_3 + 0.5865\hat{y}_1^2 + 0.4168\hat{y}_3^2 \tag{12}$$

$$\hat{y}_5 = 0.0489 - 0.369\hat{y}_2 + 1.138\hat{y}_3 - 0.6505\hat{y}_2\hat{y}_3 + 0.5937\hat{y}_2^2 + 0.2156\hat{y}_3^2 \tag{13}$$

$$\frac{d_s}{D} = 0.0208 + 2.9068\hat{y}_4 - 1.957\hat{y}_5 - 0.9366\hat{y}_4.\hat{y}_5 - 0.445\hat{y}_4^2 + 1.4092\hat{y}_5^2 \tag{14}$$

From Equations (9)–(14), it is revealed that Najafzadeh et al. (2013) employed four non-dimensional parameters (e.g. IWC , Fr , $\frac{S}{\rho U^2}$ and C_p) to predict non-dimensional scour depth $\left(\frac{d_s}{D}\right)$.

Description of GMDH-GLUE based scouring model

Beven & Binley (1992) proposed the GLUE methodology that is a stochastic approach to quantify the variability of a model output. Having employed the GLUE approach, it is possible to extract the best parameters set among many random combinations of them. Therefore, this method generates many sets of parameters using their probability distribution functions. Afterwards, all generated sets are divided into behavioral and non-behavioral sets using a specified value of the behavioral threshold. Then, the re-scaled likelihood weights of the i th set of parameters (l_w^i) would be determined as follows (Wang et al. 2006):

$$\begin{cases} l_w^i = \frac{l^i}{\sum_{k=1}^N l^k} \\ l^i = \exp\left(-\frac{RMSE^i}{RMSE^{min}}\right) \end{cases} \tag{15}$$

where l^i is the value of the i th likelihood measure of behavioral and $RMSE^{min}$ is the minimum value of $RMSE$ (herein the root mean square error parameter was considered as a performance measure) among the total number of behavioral simulations (N). Ultimately, having assigned the uniform distribution as the prior probability distribution of stochastic parameters, it is possible to determine the posterior distribution of them using Equation (15) (Freni et al. 2009).

In the current investigation, the parameters of the GLUE-GMDH model are modified by estimating the weighting coefficients of the GMDH network using the GLUE approach. In this way, the parametric model (random model) was developed as Equations (16)–(21) while all weighting coefficients of GMDH (36 coefficients) were considered as random variables:

$$\hat{y}_1 = a_1 + a_2 IWC + a_3 Fr + a_4 IWC \cdot Fr + a_5 IWC^2 + a_6 Fr^2 \quad (16)$$

$$\hat{y}_2 = b_1 + b_2 \frac{S}{\rho U^2} + b_3 Fr + b_4 \frac{S \cdot Fr}{\rho U^2} + b_5 \left(\frac{S}{\rho U^2} \right)^2 + b_6 Fr^2 \quad (17)$$

$$\hat{y}_3 = c_1 + c_2 C_p + c_3 Fr + c_4 C_p Fr + c_5 C_p^2 + c_6 Fr^2 \quad (18)$$

$$\hat{y}_4 = d_1 + d_2 \hat{y}_1 + d_3 \hat{y}_3 + d_4 \hat{y}_1 \hat{y}_3 + d_5 \hat{y}_1^2 + d_6 \hat{y}_3^2 \quad (19)$$

$$\hat{y}_5 = e_1 + e_2 \hat{y}_2 + e_3 \hat{y}_3 + e_4 \hat{y}_2 \hat{y}_3 + e_5 \hat{y}_2^2 + e_6 \hat{y}_3^2 \quad (20)$$

$$\frac{d_s}{D} = f_1 + f_2 \hat{y}_4 + f_3 \hat{y}_5 + f_4 \hat{y}_4 \cdot \hat{y}_5 + f_5 \hat{y}_4^2 + f_6 \hat{y}_5^2 \quad (21)$$

where $a_1 \sim a_6$, $b_1 \sim b_6$, $c_1 \sim c_6$, $d_1 \sim d_6$, $e_1 \sim e_6$, $f_1 \sim f_6$ are the stochastic coefficients.

To simplify the proposed model, the number of stochastic coefficients is reduced using sensitivity analysis. To this end, the impact of each coefficient on the model performance is measured. Based on the results of sensitivity analysis, it was found that among all considered stochastic coefficients, nine stochastic coefficients (e.g. $a_4 \sim a_5$, $b_2 \sim b_5$, c_2 and $c_4 \sim c_5$) exert a significant effect on maximum scour depth. Therefore, these variables were estimated using the GLUE approach and the other ones were considered by means of the Najafzadeh et al. (2013) formulas. In total, the proposed GMDH-GLUE has nine influencing parameters. To estimate the prior probability

Table 1 | Prior distribution parameters of stochastic coefficients

Scouring formula coefficients	PDF	PDF parameters	
		(Lower limit)	(Upper limit)
a_4	Uniform	0.113826	0.150535
a_5		-55.2×10^{-5}	-41.7×10^{-5}
b_2		$2.50E \times 10^{-3}$	3.31×10^{-3}
b_3		4.643478	6.141000
b_4		-0.026680	-0.020170
b_5		-8.76×10^{-7}	-6.63×10^{-7}
c_2		-0.065900	-0.049830
c_4		-0.077630	-0.058700
c_5		5.10×10^{-4}	6.74×10^{-4}

distribution function (uniform distribution) of final stochastic variable, the cross-validation method was applied to the employed dataset. The parameters of the extracted prior probability distribution functions are presented in Table 1. To estimate the stochastic coefficients by the GLUE approach, the posterior probability distribution functions of stochastic coefficients were extracted using Equation (15) and methodology, which is presented in Figure 1.

Description of GP based scouring model

Genetic Programming (GP) presents a functional expression which consists of some logical and mathematical formulas (Koza 1994). In this way, GP generates many equations (fitting equations) and selects the best one using performance criteria such as $RMSE$ (root mean square error) and MAE (mean absolute error) (Barmapalexis et al. 2018). The development of each GP based formulas evolves in five steps as (Najafzadeh & Saberi-Movahed 2018): (a) specifying terminals set (e.g. input variables, constants), (2) defining operators and symbolic functions (e.g. $+$, $-$, \times , \div , $\sqrt{\quad}$, \dots), (3) measuring the fitness of individuals in the population using goodness of fit criterion (e.g. $RMSE$), (4) controlling the run using specified parameters (e.g. population size), (5) specifying the termination criterion (e.g. maximum number of generations). In this study, a new scouring equation (Equation (22)) was developed using genetic programming and compared with other developed stochastic scouring models. In this way, the

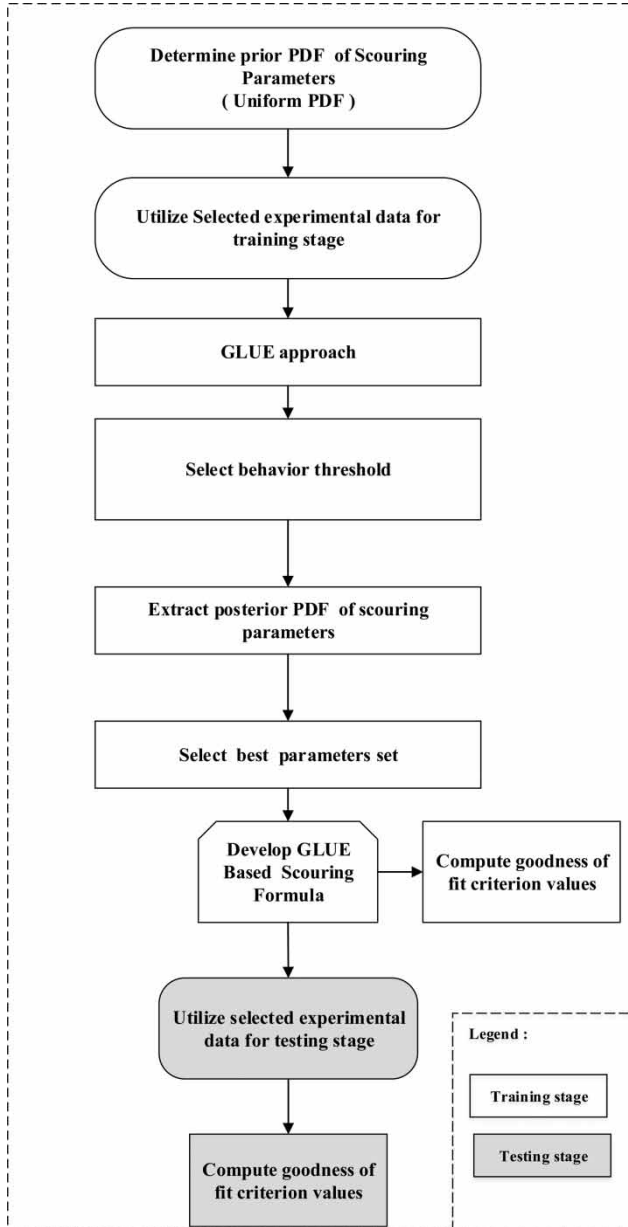


Figure 1 | The GMDH-GLUE method for predicting pier scour depth in cohesive bed materials.

arithmetic operators { + , - , × , / } and also basic functions { exp , log , power , √ } were utilized for developing the GP model (see Figure 2). Moreover, the variables of the proposed GP model are shown in Table 2. Note that a comprehensive analysis by Najafzadeh et al. (2013) demonstrates that the performance of Equations (14)–(21) is only acceptable for certain ranges of the dimensionless groups. The recommended ranges are $IWC = 36\text{--}43\%$, $C_p = 35\text{--}100\%$

and $Fr = 0.07\text{--}0.44$, adopted herein for the application of Equation (22):

$$\frac{d_s}{D} = \left(1.72 \left(\frac{-0.957Fr}{\left(\frac{0.558C_p}{-0.897IWC} \right)} + \exp\left(\frac{-0.09S}{\rho U^2}\right) + \exp(-0.093C_p) \right) \right) + 0.107 \quad (22)$$

DATASETS DESCRIPTION

In this study, a total number of 95 experimental datasets were collected from three open source researches (Rambabu et al. 2003; Debnath & Chaudhuri 2010; Najafzadeh & Barani 2011). For analysis and deriving new equations, 70% of the collected data (i.e. 67 data sets) were randomly selected for training the proposed predictive model and the remaining 30% (i.e. 28 data sets) were utilized for the testing phase. A frequency diagram of the employed dataset and statistics of them in both training and testing phases are shown in Figure 3 and Table 3, respectively. Furthermore, Table 4 presents the range of dimensionless groups of the datasets.

DESCRIPTION OF EMPLOYED PERFORMANCE INDICES

The adequacy of the predictive equation sets in training and testing phases are measured using performance indices such as root mean square error (RMSE), mean absolute error (MAE) and correlation coefficient (CC) as in the following formulas (Yaseen et al. 2018a, 2018b; Abdulelah Al-Sudani et al. 2019; Sharafati et al. 2019b):

$$RMSE = \sqrt{\frac{1}{N_T} \sum_{j=1}^{N_T} \left(\left(\frac{d_s}{D} \right)_{Obs,j} - \left(\frac{d_s}{D} \right)_{Sim,j} \right)^2} \quad (23)$$

$$MAE = \frac{1}{N_T} \sum_{j=1}^{N_T} \left| \left(\frac{d_s}{D} \right)_{Obs,j} - \left(\frac{d_s}{D} \right)_{Sim,j} \right| \quad (24)$$

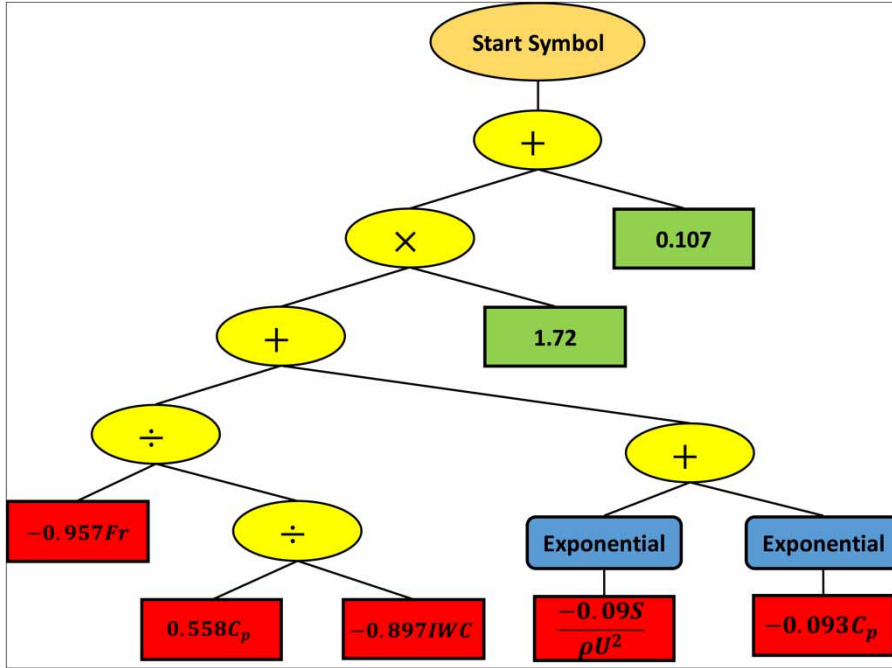


Figure 2 | Tree structure of the developed GP model to predict the pier scour depth in cohesive bed materials.

Table 2 | Parameters of developed GP model

Parameter	Description of parameter	Value of parameter
P1	Maximum symbolic expression tree depth	5
P2	Maximum symbolic expression tree length	150
P3	Population size	1,000
P4	Maximum generations	50
P5	Mutation probability	15%
P6	Internal crossover point probability	90%

where $\left(\frac{d_s}{D}\right)_{Obs,j}$ and $\left(\frac{d_s}{D}\right)_{Sim,j}$ represent the j th observed and predicted normalized scour depth magnitudes, $\overline{\left(\frac{d_s}{D}\right)_{Obs}}$ and $\overline{\left(\frac{d_s}{D}\right)_{Sim}}$ are the observed and predicted mean values of the normalized scour depth. N_T is the number of datasets.

In order to determine the proficiency of the introduced models (e.g. GMDH-GLUE and GP) to a traditional GMDH model (base model), an improvement index (IM) is computed in both training and testing phases (Sharafati et al. 2018). It should be noted that the training performance metrics have been used to evaluate the performance of a predictive model for memorizing the presented information by training dataset, while the testing performance metrics measure the model adequacy for predicting target values using the attained framework during the training phase. Owing to the different concept between training and testing phases, the IM values were computed separately in training (IM_{train}) and testing (IM_{test}) phases as follows:

$$CC = \frac{\sum_{j=1}^{N_T} \left(\left(\frac{d_s}{D} \right)_{Obs,j} - \overline{\left(\frac{d_s}{D} \right)_{Obs}} \right) \left(\left(\frac{d_s}{D} \right)_{Sim,j} - \overline{\left(\frac{d_s}{D} \right)_{Sim}} \right)}{\sqrt{\sum_{j=1}^{N_T} \left(\left(\frac{d_s}{D} \right)_{Obs,j} - \overline{\left(\frac{d_s}{D} \right)_{Obs}} \right)^2} \sqrt{\sum_{j=1}^{N_T} \left(\left(\frac{d_s}{D} \right)_{Sim,j} - \overline{\left(\frac{d_s}{D} \right)_{Sim}} \right)^2}} \quad (25)$$

$$IM_{train} = \frac{(IM_{train}^{RMSE} + IM_{train}^{MAE} + IM_{train}^{CC})}{3} \quad (26)$$

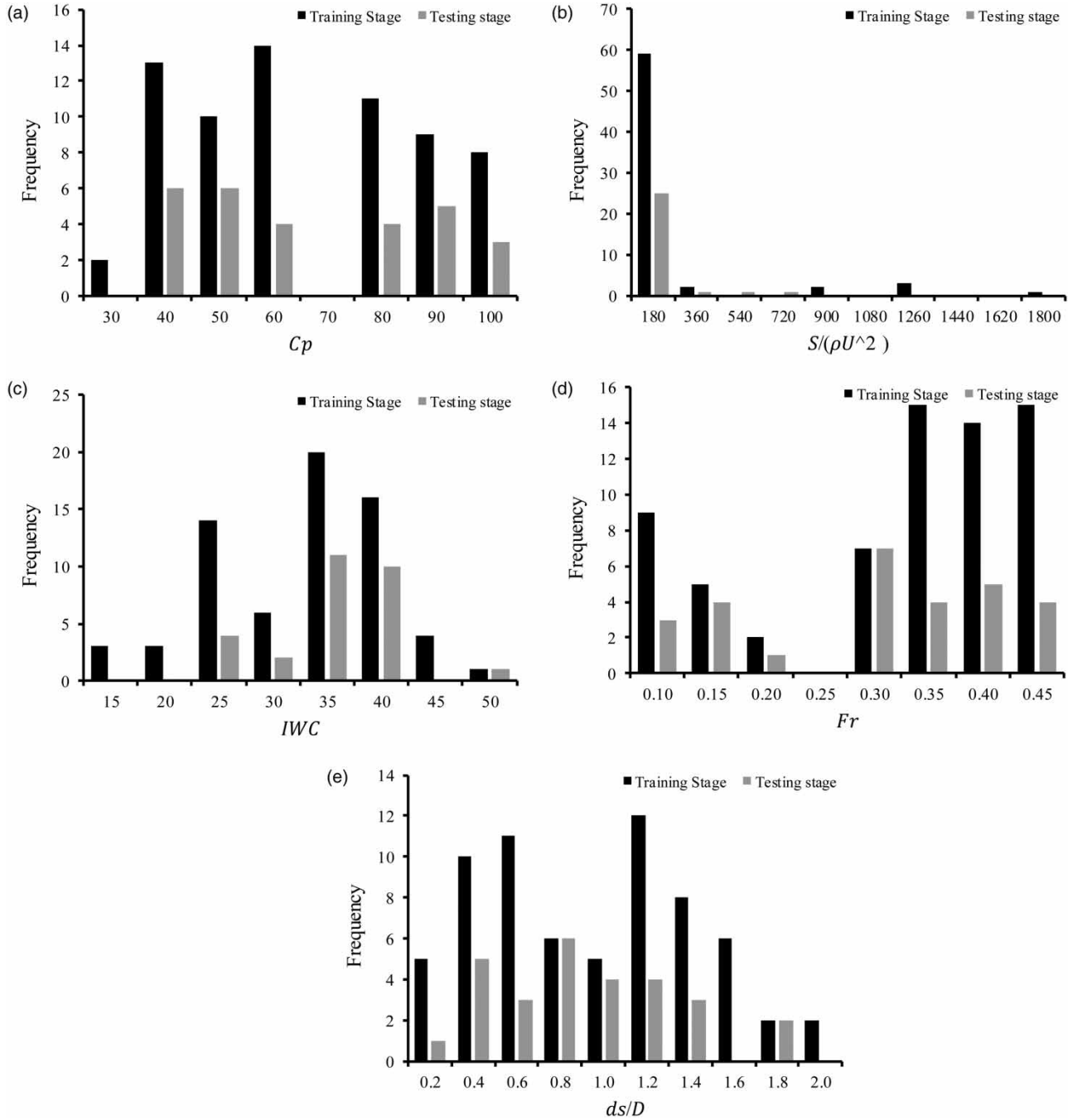


Figure 3 | Frequency of the employed data test in training and testing stages: (a) C_p , (b) $\frac{S}{\rho U^2}$, (c) IWC , (d) Fr , (e) $\frac{d_s}{D}$.

Table 3 | Statistics of employed dataset in training and testing stages

Statistic	Training stage					Testing stage				
	$\frac{d_s}{D}$	C_p	$\frac{S}{\rho U^2}$	IWC	Fr	$\frac{d_s}{D}$	C_p	$\frac{S}{\rho U^2}$	IWC	Fr
Max	1.91	100	1790.65	45.84	0.44	1.71	100	633.04	45.92	0.43
Min	0.15	20	9.23	10.70	0.07	0.17	34	11.61	20.20	0.08
Ave	0.86	59.21	131.89	30.38	0.30	0.79	59.79	76.68	32.88	0.28
S.D*	0.48	22.86	333.57	8.15	0.12	0.40	22.55	142.15	5.81	0.11
C.V ⁺	0.56	0.39	2.53	0.27	0.40	0.51	0.38	1.85	0.18	0.40

Note: *S.D: Standard deviation. + C.V: Coefficient of variation.

Table 4 | Range of dimensionless groups used in present study and Najafzadeh et al. (2013)

Parameter	Range
C_p (%)	20–100
$\frac{S}{\rho U^2}$ (-)	9–1791
IWC (%)	11–46
$Fr = \frac{U}{\sqrt{gD}}$ (-)	0.07–0.44

$$IM_{test} = \frac{(IM_{test}^{RMSE} + IM_{test}^{MAE} + IM_{test}^{CC})}{3} \quad (27)$$

$$IM_{train/test}^{RMSE} = \frac{(RMSE_{train/test}^{base\ model} - RMSE_{train/test}^{Model})}{RMSE_{train/test}^{base\ model}} \times 100 \quad (28)$$

$$IM_{train/test}^{MAE} = \frac{(MAE_{train/test}^{base\ model} - MAE_{train/test}^{Model})}{MAE_{train/test}^{base\ model}} \times 100 \quad (29)$$

$$IM_{train/test}^{CC} = \frac{(CC_{train/test}^{Model} - CC_{train/test}^{base\ model})}{CC_{train/test}^{base\ model}} \times 100 \quad (30)$$

where $RMSE_{train/test}^{base\ model}$, $CC_{train/test}^{base\ model}$ and $MAE_{train/test}^{base\ model}$ are, respectively, the computed RMSE, CC and MAE of the base model (Najafzadeh et al. 2013) while $RMSE_{train/test}^{Model}$, $CC_{train/test}^{Model}$ and $MAE_{train/test}^{Model}$ are the calculated indices of the proposed models (e.g. GMDH-GLUE and GP) in training or testing phases.

APPLICATION RESULTS AND DISCUSSION

The main objective of this study is to make an enhancement on the published research by (Najafzadeh et al. 2013), in which the scour depth around a pier in cohesive soil is predicted by the GMDH technique. Due to the high stochasticity of the targeted variable $\frac{d_s}{D}$ in response to various random variables (e.g.; Fr , IWC , C_p , $\frac{S}{\rho U^2}$), the authors were inspired to adopt a reliable stochastic modeling strategy (i.e. GLUE) integrated with GMDH as a tuning procedure. In addition to the proposed GMDH-GLUE, an outstanding evolutionary computing GP model was developed for validation purposes.

Development of GMDH-GLUE model for estimating of scour depth

To develop a GMDH-GLUE model for estimating the scour depth around a cylinder founded in cohesive sediments, the significant stochastic parameters (e.g. $a_4 \sim a_5$, $b_2 \sim b_5$, c_2 and $c_4 \sim c_5$), as shown in Equations (16)–(18), were employed as basic stochastic variables. To estimate the stochastic parameters, uniform probability distributions were employed as the prior distribution of the mentioned parameters (Table 1). Further, the GLUE method was employed to extract the posterior distributions of the variables by performing 10,000 simulations (Figure 4).

From Figure 4 it is revealed that the b_3 , c_4 and c_5 (coefficients of Fr and C_p) have positive skewness. Therefore, the performance of scour depth prediction is increased using

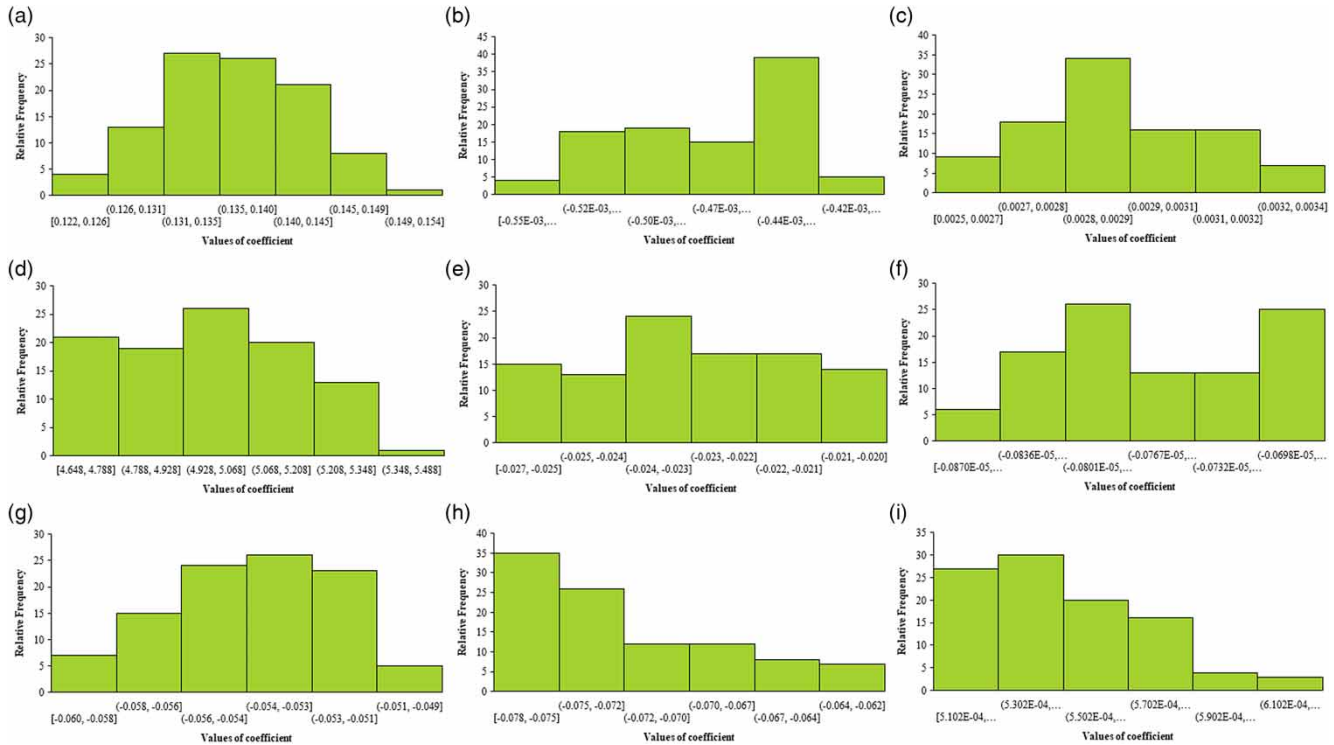


Figure 4 | Posterior distribution of the employed stochastic coefficients in GMDH-GLUE model: (a) a_4 , (b) a_5 , (c) b_2 , (d) b_3 , (e) b_4 , (f) b_5 , (g) c_2 , (h) c_4 , (i) c_5 .

the lower value of these parameters. Also, the negative skewness values of a_5, b_b (coefficients of IWC and $\frac{S}{\rho U^2}$) implies that the accuracy of scour depth prediction can be increased using large values of these parameters resulting in increasing the accuracy of scouring prediction.

Having generated the posterior distributions of the stochastic parameters, the best parameters set was selected among preformed simulations by the GLUE approach. The GMDH-GLUE model produced several stochastic formulations based on the influential nine hyper-parameters toward the scouring depth, expressed as follows:

$$\hat{y}_1 = 0.3278 + 0.0022IWC - 0.585Fr + 0.141IWC.Fr - 47.2 \times 10^{-5}IWC^2 - 1.35Fr^2 \quad (34)$$

$$\hat{y}_2 = -0.2806 + 0.00291 \frac{S}{\rho U^2} + 5.157Fr - 0.0232 \frac{S.Fr}{\rho U^2} - 6.86 \times 10^{-7} \left(\frac{S}{\rho U^2} \right)^2 - 2.42133Fr^2 \quad (35)$$

$$\hat{y}_3 = 1.657 - 0.0539C_p + 2.3047Fr - 0.0766C_p.Fr + 5.62 \times 10^{-4}C_p^2 + 7.834Fr^2 \quad (36)$$

As noted before, the use of Equations (34)–(36), as well as Equations (9)–(14) is limited to $IWC = 36\text{--}43\%$, $C_p = 35\text{--}100\%$ and $Fr = 0.07\text{--}0.44$.

Assessment of GMDH-GLUE performance

The performance of the proposed GMDH-GLUE model was validated using an uncertainty potential limit prior to the modeling phase during the training learning process. The criterion of 95PPU and P-factor were examined as reported in Figure 4. In accordance to the established research by Abbaspour et al. (2007), the computed value of 95PPU is considered using a minimal percentage of 50 for the modeled observations in the 95PPU bound, which is an acceptable range. On the basis of the GMDH-GLUE model results over the training phase (Figure 5), nine observations out of 67 were located out of the 95PPU bound with P-factor magnitude (87%). In general, GMDH-GLUE exhibited a reliable modeling strategy during the training phase.

One of the very common graphical examinations usually performed for prediction evaluation is the scatterplot (see

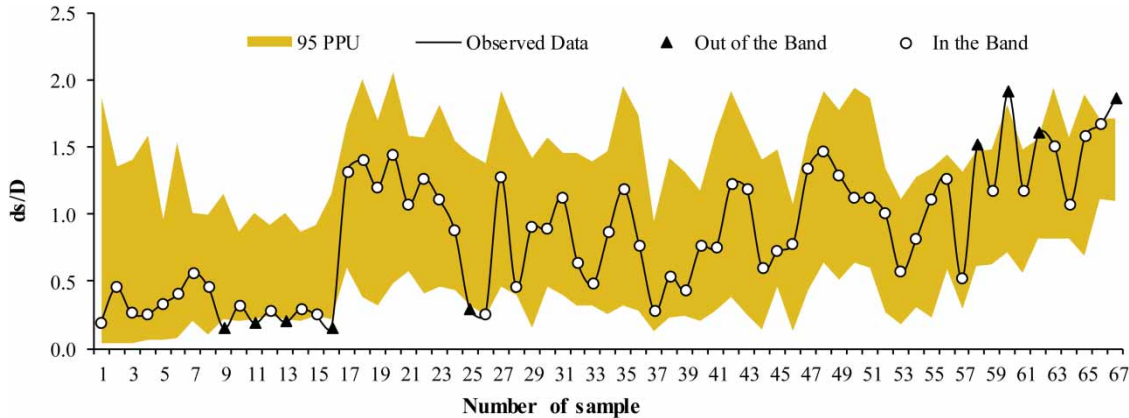


Figure 5 | Comparison between the generated 95PPU band of GMDH-GLUE scouring model and the observed training data.

Figure 6). The figure reports an informative presentation of the variance between the observed scouring data and the predicted values. Figure 6(a) and 6(b) generated the scatter-plots over the training and testing phases. The figures indicate the square correlation coefficient R^2 values where the variance magnitudes are scattered. The R^2 value of the GMDH-GLUE model denoted a superior result followed by the evolutionary GP and GMDH-BP models, with quantitative values of approximately $R^2 \approx 0.84$, 0.80 and 0.79, in that order, over the training phase, whereas the testing phase GMDH-GLUE, GP and GMDH-BP models attained $R^2 \approx 0.82$, 0.62 and 0.81, respectively. Markedly, the GMDH-GLUE model displayed the closest variance towards the ideal fit line 45° .

Using the intercept between the correlation coefficient, standard division and the root mean square metrics, the Taylor diagram was generated for more constructive validation of the developed stochastic predictive model. Figure 7(a) and 7(b) shows the Taylor diagram visualization over the training and testing phases of the modeling. In both figures the GMDH-GLUE model evidenced a closer coordinate to the observed benchmark scouring depth value, while in the testing phase the GP model was less accurate in the prediction of the scour depth. Indeed, this is normal due to the lower magnitude achieved for the correlation coefficient as reported in the scatterplot presentation.

A boxplot is generated for the modeling evaluation of the scour depth prediction where the degree of spread in the prediction data is reported in the form of different quartiles (25, 50, 75 and the interquartile range, IQR) (see

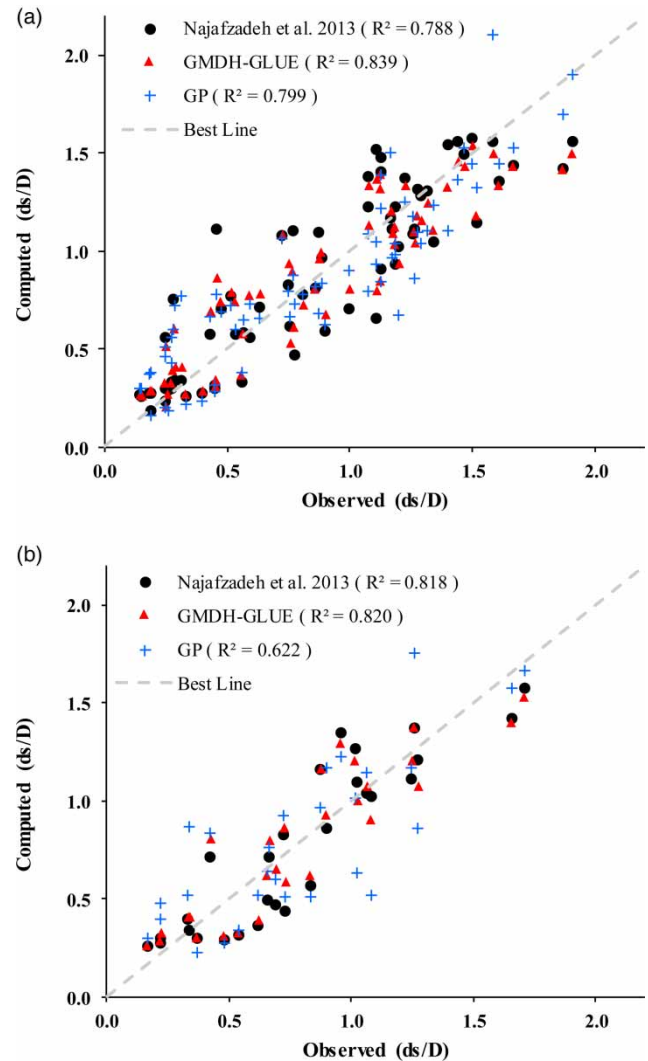


Figure 6 | Scatter plot of the computed versus the observed scour depth: (a) training phase, (b) testing phase.

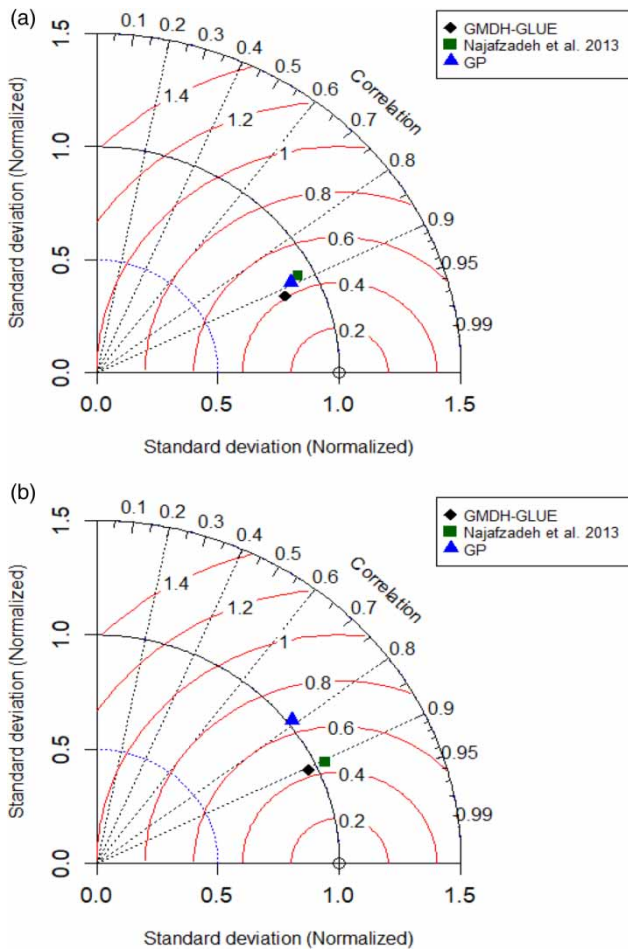


Figure 7 | Normalized Taylor diagrams of the predicted and the observed scour depth: (a) training phase, (b) testing phase.

Figure 8). Based on the numerical magnitudes of the lower (Q25%), median (Q50%) and upper (Q75%) quartiles, the GMDH-GLUE model offered the best performance in comparison with the GMDH-BP and GP models. However, in terms of Q25% the GP model provided more reliable predictions in some cases. These are usual outcomes where statistical models exhibit different behavior from a certain case to another one when the nature of the modeled problem may significantly change. In general, the IQR is an indicator for the data variability proving the capability of the GMDH-GLUE and GP models over the GMDH-BP model.

In numerical presentation (see Table 5), through the absolute error metrics evaluation (RMSE and MAE), the GMDH-GLUE model demonstrated better prediction

results over the classical GMDH modelled using a back propagation learning algorithm. On the contrary, the reported quantitative results clarify that the GP model performed adversely. The statistical indices of the GMDH-GLUE model in terms of RMSE and MAE metrics have been improved by 3.23 and 12.11% in the training stage, while those indices were enhanced by 4.02 and 4.76% over the testing phase. It is clear from the statistical presentation of the RMSE, and MAE metrics, that it is not adequate to explain the superiority of the model's accuracy. Hence, another comprehensive metric, i.e. IM = improvement index is computed for the validation of the proposed method. A satisfactory improvement is achieved using the GMDH-GLUE model over the training ($IM_{train} = 6\%$) and testing ($IM_{test} = 3\%$) phases, respectively.

To demonstrate the superiority of the new GMDH-GLUE model, it is validated against the other studies in this context. Here, the RMSE metric was utilized as a determination factor for the evaluation. Several studies were surveyed and their results are tabulated in Table 6. Based on the reported numerical values of the RMSE, the proposed GMDH-GLUE model demonstrated good enhancement over the literature studies. GMDH-GLUE provided prediction enhancement of 70.6% against the ANN model (Kaya 2010), 30.8% against GMDH-BP (Najafzadeh & Barani 2011), 50.5% against the support vector regression (SVR) model (Pal et al. 2011), 4.76% against GMDH (Najafzadeh et al. 2013), 24.7% against the gene expression programming (GEP) model (Muzzammil et al. 2015), and 24.6% against the evolutionary radial basis function neural network (ERBFNN) model (Cheng et al. 2015). Apparently, the proposed GMDH-GLUE model revealed an acceptable prediction accuracy improvement over the literature studies and that demonstrated its potential for comprehending the actual relationship between the influencing parameters and the maximum scour depth.

CONCLUSIONS AND FUTURE RESEARCH DIRECTION

Scouring phenomena is one of the vital problematic issues in riverine and marine structures and hydraulic engineers need to take into account optimal effective parameters in

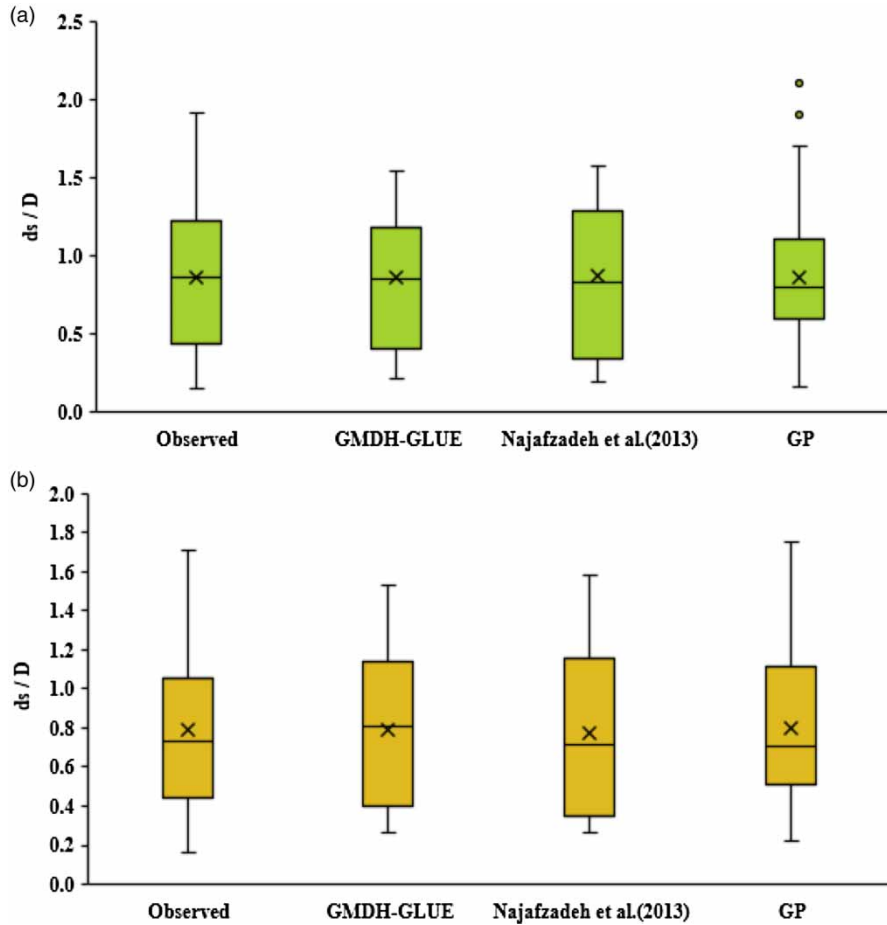


Figure 8 | Boxplot of the observed scour depth against the predicted ones: (a) training phase, (b) testing phase.

Table 5 | Values of performance measures and improvement indices of the predictive models

Equation		Train			IM _{Train}	Test			IM _{Test}
		RMSE	MAE	CC		RMSE	MAE	CC	
Najafzadeh et al. (2013)		0.221	0.170	0.888	–	0.182	0.150	0.905	–
GMDH-GLUE		0.195	0.164	0.916	–	0.173	0.144	0.906	–
GP		0.214	0.176	0.894	–	0.266	0.216	0.789	–
Improvement (%)	GMDH-GLUE	12.11	3.23	3.16	6	4.76	4.02	0.11	3
	GP	3.31	–3.55 ^a	0.71	0.2	–46.27	–44.11	–12.80	–34

^aNote: the negative value means the GP model has lower performance compared to the GMDH model.

the design of those structures, ensuring adequacy and accuracy of the prediction results. The current study provided a new integrated stochastic model based on GMDH-GLUE for predicting scour depth around piers using effective variables. In fact, the use of the standard GMDH technique

provides deterministic predictions based on a complex relationship between the influencing parameters and the output variable. In the present study, the prediction was improved, taking into account the stochastic behavior of the data that is generally neglected when only the GMDH

Table 6 | The validation of the current research modeling approach against the established researches over the literature for depth scouring around piers using the root mean square error metric

Literature research	Current research (GMDH-GLUE)	Kaya (2010) (ANN)	Najafzadeh & Barani (2011) (GMDH-BP)	Pal et al. (2011) (SVR)	Najafzadeh et al. (2013) (GMDH)	Muzzammil et al. (2015) (GEP)	Cheng et al. (2015) (ERBFNN)
RMSE (testing phase)	0.173	0.59	0.25	0.35	0.182	0.23	0.23

approach is considered. This investigation proves the acceptable adequacy of the selected stochastic technique, GLUE, to extract the stochastic characteristics of the experimental data. The new concept, GMDH-GLUE, denotes a combined model whose predictions are based on both the stochastic and deterministic properties of the input and output variables.

To validate the selected methodology, the proposed stochastic model was validated against the previously published GMDH model (Najafzadeh et al. 2013) in addition to one reliable evolutionary computing model called genetic programming. The modeling results were authenticated using numerous statistical metrics and graphical evaluation indicators. The findings of this study evidenced the feasibility and accuracy of the adopted GMDH-GLUE model in comparison to the benchmark models. The statistical metric *IM* substantiated a satisfactory accuracy improvement of 6 and 3%, respectively, over the train and test modeling phases against the original GMDH model. Overall, the explored stochastic predictive model captured the internal nonlinear mathematical behavior of the influencing parameters and the output variables. Although satisfactory prediction accuracy was achieved utilizing the selected methodology, the modeling strategy would be enhanced through the incorporation of a non-linear mask approach to abstract the mostly correlated input attributes using the potential of the recently explored nature-inspired optimization algorithms.

ACKNOWLEDGEMENTS

The authors would like to extend their gratitude and appreciation to Prof. Amir Etemad Shahidi for offering the employed data set and his constructive comments for developing this research. We would also like to express

our profound gratitude to the respected editor Prof. Dimitri Solomatine and two unknown reviewers for sharing their ideas and admirable comments.

REFERENCES

- Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J. & Zobrist, J. 2007 *Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT*. *J. Hydrol.* **333**, 413–430.
- Abdulah Al-Sudani, Z., Salih, S. Q., Sharafati, A. & Yaseen, Z. M. 2019 *Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation*. *J. Hydrol.* **573**, 1–12.
- Amanifard, N., Nariman-Zadeh, N., Farahani, M. H. & Khalkhali, A. 2008 *Modelling of multiple short-length-scale stall cells in an axial compressor using evolved GMDH neural networks*. *Energy Convers. Manage.* **49**, 2588–2594.
- Ansari, S. A., Kothiyari, U. C. & Ranga Raju, K. G. 2002 *Influence of cohesion on scour around bridge piers*. *J. Hydraul. Res.* **40**, 717–729.
- Arneson, L. A., Zevenbergen, L. W., Lagasse, P. & Clopper, P. E. 2012 *Evaluating Scour at Bridges*. Federal Highway Administration, USA.
- Ayele, G., Teshale, E., Yu, B., Rutherford, I. & Jeong, J. 2017 *Streamflow and sediment yield prediction for watershed prioritization in the Upper Blue Nile River Basin, Ethiopia*. *Water* **9**, 782.
- Azamathulla, H. M., Ghani, A. A., Zakaria, N. A. & Guven, A. 2010 *Genetic programming to predict bridge pier scour*. *J. Hydraul. Eng.* **136**, 165–169.
- Barmapalexis, P., Grypioti, A., Eleftheriadis, G. K. & Fatouros, D. G. 2018 *Development of a new aprepitant liquisolid formulation with the aid of artificial neural networks and genetic programming*. *AAPS Pharm. Sci. Tech.* **19**, 741–752.
- Beven, K. & Binley, A. 1992 *The future of distributed models: model calibration and uncertainty prediction*. *Hydrol. Processes* **6**, 279–298.
- Beusers, H. N. C., Nicolle, G. & Shen, H. W. 1977 *Local scour around cylindrical piers*. *J. Hydraul. Res.* **15**, 211–252.
- Briaud, J.-L., Ting, F. C. K., Chen, H. C., Gudavalli, R., Perugu, S. & Wei, G. 1999 *SRICOS: prediction of scour rate in cohesive soils at bridge piers*. *J. Geotech. Geoenviron. Eng.* **125**, 237–246.

- Chang, W.-Y., Lai, J.-S. & Yen, C.-L. 2004 Evolution of scour depth at circular bridge piers. *J. Hydraul. Eng.* **130**, 905–913.
- Cheng, M.-Y., Cao, M.-T. & Wu, Y.-W. 2015 Predicting equilibrium scour depth at bridge piers using evolutionary radial basis function neural network. *J. Comput. Civ. Eng.* **5**, 4014070.
- Chiew, Y. M. & Melville, B. W. 1987 Local scour around bridge piers. *J. Hydraul. Res.* **25**, 15–26.
- Dargahi, B. 1990 Controlling mechanism of local scour. *J. Hydraul. Eng.* **116**, 1197–1214.
- Debnath, K. & Chaudhuri, S. 2010 Laboratory experiments on local scour around cylinder for clay and clay-sand mixed beds. *Eng. Geol.* **111**, 51–61.
- Dodaro, G., Tafarojnoruz, A., Calomino, F., Gaudio, R., Stefanucci, F., Adduce, C. & Sciortino, G. 2014 An experimental and numerical study on the spatial and temporal evolution of a scour hole downstream of a rigid bed. In: *Proceedings of the International Conference on Fluvial Hydraulics, RIVER FLOW 2014*. Taylor and Francis Group plc, Lausanne, Switzerland, pp. 1415–1422.
- Dodaro, G., Tafarojnoruz, A., Sciortino, G., Adduce, C., Calomino, F. & Roberto, G. 2016 Modified Einstein sediment transport method to simulate the local scour evolution downstream of a rigid bed. *J. Hydraul. Eng.* **142**, 4016041.
- Etemad-Shahidi, A. & Ghaemi, N. 2011 Model tree approach for prediction of pile groups scour due to waves. *Ocean Eng.* **38**, 1522–1527.
- Etemad-Shahidi, A., Bonakdar, L. & Jeng, D.-S. 2015 Estimation of scour depth around circular piers: applications of model tree. *J. Hydroinf.* **17**, 226–238.
- Farlow, S. J. 1981 The GMDH algorithm of Ivakhnenko. *Am. Stat.* **35**, 210–215.
- Ferraro, D., Tafarojnoruz, A., Gaudio, R. & Cardoso, A. H. 2013 Effects of pile cap thickness on the maximum scour depth at a complex pier. *J. Hydraul. Eng.* **139**, 482–491.
- Firat, M. & Gungor, M. 2009 Generalized regression neural networks and feed forward neural networks for prediction of scour depth around bridge piers. *Adv. Eng. Software* **40**, 731–737.
- Freni, G., Mannina, G. & Viviani, G. 2009 Urban runoff modelling uncertainty: comparison among Bayesian and pseudo-Bayesian methods. *Environ. Model. Software* **24**, 1100–1111.
- Garg, V. 2015 Inductive group method of data handling neural network approach to model basin sediment yield. *J. Hydrol. Eng.* **20**, C6014002.
- Gaudio, R., Grimaldi, C., Tafarojnoruz, A. & Calomino, F. 2010 Proceedings 1st European IAHR congress. In: *Comparison of Formulae for the Prediction of Scour Depth at Piers*. Heriot-Watt University Edinburgh, UK, p. 6.
- Gaudio, R., Tafarojnoruz, A. & De Bartolo, S. 2013 Sensitivity analysis of bridge pier scour depth predictive formulae. *J. Hydroinf.* **15**, 939–951.
- Güven, A., Azamathulla, H. M. & Zakaria, N. A. 2009 Linear genetic programming for prediction of circular pile scour. *Ocean Eng.* **36**, 985–991.
- Ivakhnenko, A. G. & Ivakhnenko, G. A. 1995 The review of problems solvable by algorithms of the group method of data handling (GMDH). *Pattern Recognit. Image Anal.* **5**, 527–535.
- Ivakhnenko, A. G. & Ivakhnenko, G. A. 2000 Problems of further development of the group method of data handling algorithms. Part I. *Pattern Recognit. Image Anal.* **10**, 187–194.
- Ivakhnenko, A. G. & Müller, J. A. 1997 Recent developments of self-organising modeling in prediction and analysis of stock market. *Microelectron. Reliab.* **37**, 1053–1072.
- Johnson, P. A. & Dock, D. A. 1998 Probabilistic bridge scour estimates. *J. Hydraul. Eng.* **124**, 750–754.
- Kaya, A. 2010 Artificial neural network study of observed pattern of scour depth around bridge piers. *Comput. Geotech.* **37**, 413–418.
- Kirkil, G., Constantinescu, S. G. & Ettema, R. 2008 Coherent structures in the flow field around a circular cylinder with scour hole. *J. Hydraul. Eng.* **134**, 572–587.
- Koza, J. R. 1994 Genetic programming as a means for programming computers by natural selection. *Stat. Comput.* **4**, 87–112.
- Melville, B. W. 1997 Pier and abutment scour: integrated approach. *J. Hydraul. Eng.* **123**, 125–136.
- Muzzammil, M., Alama, J. & Danish, M. 2015 Scour prediction at bridge piers in cohesive bed using gene expression programming. *Aquat. Procedia.* **4**, 789–796.
- Najafzadeh, M. 2015 Neuro-fuzzy GMDH based particle swarm optimization for prediction of scour depth at downstream of grade control structures. *Eng. Sci. Technol. Int. J.* **18**, 42–51.
- Najafzadeh, M. & Azamathulla, H. M. 2013 Group method of data handling to predict scour depth around bridge piers. *Neural Comput. Appl.* **23**, 2107–2112.
- Najafzadeh, M. & Barani, G. A. 2011 Comparison of group method of data handling based genetic programming and back propagation systems to predict scour depth around bridge piers. *Sci. Iran* **18**, 1207–1215.
- Najafzadeh, M. & Saberi-Movahed, F. 2018 GMDH-GEP to predict free span expansion rates below pipelines under waves. *Mar. Georesour. Geotechnol.* **37**, 375–392.
- Najafzadeh, M. & Tafarojnoruz, A. 2016 Evaluation of neuro-fuzzy GMDH-based particle swarm optimization to predict longitudinal dispersion coefficient in rivers. *Environ. Earth Sci.* **75**, 157.
- Najafzadeh, M., Barani, G. A. & Azamathulla, H. M. 2013 GMDH to predict scour depth around a pier in cohesive soils. *Appl. Ocean Res.* **40**, 35–41.
- Najafzadeh, M., Barani, G.-A. & Hessami Kermani, M. R. 2014 Estimation of pipeline scour due to waves by GMDH. *J. Pipeline Syst. Eng. Pract.* **5**, 06014002.
- Najafzadeh, M., Tafarojnoruz, A. & Lim, S. Y. 2017 Prediction of local scour depth downstream of sluice gates using data-driven models. *ISH J. Hydraul. Eng.* **23**, 195–202.
- Najafzadeh, M., Rezaie-Balf, M. & Tafarojnoruz, A. 2018 Prediction of riprap stone size under overtopping flow using data-driven models. *Int. J. River Basin Manage.* **16**, 505–512.
- Pal, M., Singh, N. K. & Tiwari, N. K. 2011 Support vector regression based modeling of pier scour using field data. *Eng. Appl. Artif. Intell.* **24**, 911–916.

- Parsaie, A., Azamathulla, H. M. & Haghiabi, A. H. 2018 Prediction of discharge coefficient of cylindrical weir-gate using GMDH-PSO. *ISH J. Hydraul. Eng.* **24**, 116–123.
- Rambabu, M., Narasimha Rao, S. & Sundar, V. 2003 Current-induced scour around a vertical pile in cohesive soil. *Ocean. Eng.* **30**, 893–920.
- Shaghghi, S., Bonakdari, H., Gholami, A., Ebtehaj, I. & Zeinolabedini, M. 2017 Comparative analysis of GMDH neural network based on genetic algorithm and particle swarm optimization in stable channel design. *Appl. Math. Comput.* **313**, 271–286.
- Sharafati, A., Yasa, R. & Azamathulla, H. M. 2018 Assessment of stochastic approaches in prediction of wave-induced pipeline scour depth. *J. Pipeline Syst. Eng. Pract.* **9**, 4018024.
- Sharafati, A., Haghbin, M., Motta, D. & Yaseen, Z. M. 2019a The application of soft computing models and empirical formulations for hydraulic structure scouring depth simulation: a comprehensive review, assessment and possible future research direction. *Arch. Comput. Methods Eng* 1–25. doi:10.1007/s11831-019-09382-4.
- Sharafati, A., Khosravi, K., Khosravinia, P., Ahmed, K., Salman, S. A., Mundher, Z. & Shamsuddin, Y. 2019b The potential of novel data mining models for global solar radiation prediction. *Int. J. Environ. Sci. Technol.* **16**, 7147–7164.
- Sharafati, A., Tafarajnoruz, A., Shourian, M. & Yaseen, Z. M. 2019c Simulation of the depth scouring downstream sluice gate: the validation of newly developed data-intelligent models. *J. Hydro. Environ. Res.* doi.org/10.1016/j.jher.2019.11.002.
- Tafarajnoruz, A. 2012 Discussion of ‘Genetic programming to predict bridge pier scour’. *J. Hydraul. Eng.* **138**, 669–671.
- Tafarajnoruz, A., Gaudio, R., Grimaldi, C. & Calomino, F. 2010 Required conditions to achieve the maximum local scour depth at a circular pier. In: *Proceedings of the XXXII Italian Conference of Hydraulics and Hydraulic Constructions*, Palermo, Italy.
- Toth, E. & Brandimarte, L. 2011 Prediction of local scour depth at bridge piers under clear-water and live-bed conditions: comparison of literature formulae and artificial neural networks. *J. Hydroinf.* **13**, 812–824.
- Wang, X., Frankenberger, J. R. & Klavivko, E. J. 2006 Uncertainties in DRAINMOD predictions of subsurface drain flow for an Indiana silt loam using the GLUE methodology. *Hydrol. Processes Int. J.* **20**, 3069–3084.
- Yanmaz, A. M. 2001 Uncertainty of local scouring parameters around bridge piers. *Turk. J. Eng. Environ. Sci.* **25**, 127–137.
- Yanmaz, A. M. 2002 Dynamic reliability in bridge pier scouring. *Turk. J. Eng. Environ. Sci.* **26**, 367–375.
- Yaseen, Z. M., Awadh, S. M., Sharafati, A. & Shahid, S. 2018a Complementary data-intelligence model for river flow simulation. *J. Hydrol.* **567**, 180–190.
- Yaseen, Z. M., Ehteram, M., Sharafati, A., Shahid, S., Al-Ansari, N. & El-Shafie, A. 2018b The integration of nature-inspired algorithms with Least Square Support Vector regression models: application to modeling river dissolved oxygen concentration. *Water (Switzerland)* **10**, 1124.
- Zell, W. O., Culver, T. B. & Sanford, W. E. 2018 Prediction uncertainty and data worth assessment for groundwater transport times in an agricultural catchment. *J. Hydrol.* **561**, 1019–1036.

First received 24 February 2019; accepted in revised form 9 December 2019. Available online 3 February 2020