Estimation of scour depth around submerged weirs using self-adaptive extreme learning machine

M. Rashki Ghaleh Nou, M. Azhdary Moghaddam, M. Shafai Bajestan and H. Md. Azamathulla

ABSTRACT

In this study, the equilibrium scour depth downstream of the weir ($d_{s\text{-}a}$), the maximum scour depth downstream of the weir ($d_{s\text{-}max}$), the equilibrium scour depth upstream of the weir ($d_{us\text{-}a}$) and the maximum scour depth upstream of the weir ($d_{us\text{-}max}$) were simulated around the submerged weirs using the self-adaptive extreme learning machine (SAELM) model. In other words, the SAELM was utilized for the simulation of the scour depths around submerged weirs for the first time. In addition, Monte Carlo simulations (MCSs) were used to increase the accuracy of the artificial intelligence model. The results of modeling were validated using k-fold cross validation. At first, all effective parameters on the scour depth were determined and five distinct SAELM models were defined. Then, the optimal activation function of the SAELM model was obtained. By analyzing the results of modeling, the best models were identified to estimate $d_{s\text{-}a}/h_t$, $d_{s\text{-}max}/h_t$, $d_{us\text{-}a}/h_t$, and $d_{us\text{-}max}/h_t$, and the ratio of the average inflow velocity to the critical velocity (U_0/U_c) was determined as the most effective input parameter. In the following, the results of superior models were compared with the artificial neural network (ANN) and support vector machine (SVM). The results showed that SAELM models were more accurate. The uncertainty analysis was performed for these models, some of them were overestimated and others were underestimated. In addition, some equations were presented for equilibrium models for calculation of scour depth around the submerged weirs, which are used by environmental and hydraulic engineers without previous knowledge about the artificial intelligence models. Finally, a partial derivative sensitivity analysis (PDSA) was performed for all input parameters of the superior models.

Key words | K-fold cross validation, partial derivative sensitivity analysis (PDSA), scour, self-adaptive extreme learning machine (SAELM), submerged weirs

M. Rashki Ghaleh Nou (corresponding author)
University of Sistan and Baluchestan, Zahedan, Iran
E-mail: *mohammadrashki@pgs.usb.ac.ir*

M. Azhdary Moghaddam
Department of Civil Engineering, University of Sistan and Baluchestan, Zahedan, Iran

M. Shafai Bajestan
Department of Hydraulic Structures, Shahid Chamran University of Ahvaz, Iran

H. Md. Azamathulla
Department of Civil Engineering, University of the West Indies, St. Augustine, Trinidad

INTRODUCTION

In general, submerged weirs are widely used to raise the upstream water levels, determine the water depth for navigation, and prevent the erosion of the channel and river beds. In addition, the flow structure may cause the formation of sedimentary deposits or local scouring around the submerged weirs. Due to the importance of these hydraulic structures, numerous experimental, analytical and numerical studies have been carried out on the scouring pattern around the submerged weirs.

For example, Odgaard & Kennedy (1983) were among the first researchers to evaluate the flow pattern and scouring around the submerged weirs. By analyzing the results, they obtained a relationship to calculate the distance between submerged vanes. Subsequently, Odgaard &

doi: 10.2166/hydro.2019.070

Spoljaric (1986) investigated the effect of the submerged vanes on the flow pattern and bed topography in a laboratory rectangular flume. In addition, Odgaard & Wang (1991) presented an analytical method for design of submerged vanes in open channel flows. They verified their model results with a laboratory model and showed that the analytical model had an appropriate accuracy. Marelius & Sinha (1998) experimentally studied the effect of different inclination angles of submerged vanes on the flow pattern. They obtained the optimum angle of the submerged vanes. Gaudio *et al.* (2000) examined the long-term scouring around bed sills in different experimental conditions. By analyzing the results, they proposed many relationships for calculating the depth and length of the scour hole. Tan *et al.* (2005) studied the scouring pattern around the submerged vanes on a straight channel. They showed that the efficiency of submerged vanes depends on the impact angle of the flow to the plate, height, and length of the plates. They concluded that the optimum impact angle of the flow to the plates is 30 degrees and an optimum height of the submerged vanes is two to three times the height of the bed form. In addition, Ouyang (2009) evaluated the effects of the shape and size of submerged vanes on the scouring pattern in a rectangular channel. They provided an analytical model for calculating the changes in sedimentary beds and verified this model with experimental data. Bajestan & Azizi (2012) conducted a series of laboratory measurements on the scour depth around the submerged vanes. The analysis of the experimental results showed that cutting the leading edge of the submerged vanes reduced the amount of scouring around the submerged vanes. Guan *et al.* (2014) conducted an experimental study on the turbulence flow structure and shear stress in the submerged waters under clear water conditions. The experimental results showed that a large circulation and reattachment zone are formed downstream of the submerged weir. Guan *et al.* (2015) measured the scour values in upstream and downstream submerged weirs under live-bed conditions in a laboratory study. Using dimensional analysis and experimental results, they proposed the relationships for estimating scouring around the submerged weirs. Guan *et al.* (2016) evaluated experimentally the effect of sediment size, depth, flow rate and weir height on the scouring pattern around the submerged weirs. They provided several relationships to calculate the scour hole dimensions. Wang *et al.* (2018a) measured the local scouring between the circular bridge pier and submerged weir and concluded that with the application of the submerged weir downstream of the bridge pier, the scour depth is significantly reduced. Wang *et al.* (2018b) investigated the scour process around sloping submerged weirs in a laboratory study for two coarse-grained and fine-grained sediments. By analyzing the experimental results, a design approach was proposed for this structure. Wang *et al.* (2018c) investigated experimentally the effect of the slope of the submerged weir on the scouring pattern of coarse-grained and fine-grained sediments. They proposed a relationship as a function of medium and maximum scour.

In recent years, artificial intelligence and soft computing models have been widely used for modelling various phenomena. For example, Babovic (2000), Yu *et al.* (2004), Babovic (2009). Also, Bateni *et al.* (2007) estimated the time-dependent scour depth around the circular bridge piers using a neural network model. The results of their study showed that the most effective parameter on scour depth is the bridge pier diameter. In addition, Firat & Gungor (2009) estimated the scour around the bridge piers using GRNN and FFNN neural network algorithms. They showed that the GRNN model predicted the scour values more accurately. Zounemat-Kermani *et al.* (2009) modeled the scouring pattern around the bridge piers by the adaptive neuro-fuzzy inference system and the neural network. They showed that the FFBP-NN neural network model estimates the scour depth with great accuracy. Based on the results of sensitivity analysis, the dimensionless parameter of the bridge pier diameter and the ratio of the pier distance to the bridge pier diameter were the most effective parameters on the scouring pattern. Goel & Pal (2009) showed the potential of support vector machines in predicting the maximum scour depth on grade-control structures, from the available laboratory and field data. They compared the results of their study with some empirical relation and a feed forward back propagation neural network model. The outcome from the support vector machines-based modeling approach suggested a better performance in comparison to both the empirical relation and back propagation neural network approach with the

laboratory data. Muzzammil (2010) simulated the depth of the scour hole downstream of the abutments using adaptive neuro-fuzzy inference system (ANFIS) and neural network models. He found that the ANFIS model predicted the scour depth with more accuracy than neural network model. Najafzadeh *et al.* (2013) estimated the scour depth around the abutments using group method of data handling (GMDH) and support vector machines for clear-water conditions and lived-bed conditions. The results of the analysis showed that the GMDH model was more accurate than support vector machines. Etemad-Shahidi *et al.* (2015) simulated the scour depth around the bridges using an M5 model tree. They compared the results of the numerical model with empirical relationships, and found that the accuracy of the M5 model tree is great. Additionally, Sharafi *et al.* (2016) modeled the local scouring around the bridge bases using a support vector machine. Azimi *et al.* (2017a) estimated the scour around the bridge pile group with the combination of differential evolutionary algorithms and ANFIS. They presented ten combined models to identify the superior model and the most effective parameter for simulating scouring around the bridge pile group.

It should be noted that estimation of scouring around submerged weirs has not been simulated through artificial intelligence (AI) techniques yet. Additionally, classical AI models such as artificial neural networks (ANNs), ANFIS, support vector machine (SVM), gene expression programming (GEP) and GMDH have some disadvantages which a novel algorithm like extreme learning machine (ELM) overcomes. The ELM has one hidden layer and computational time is much less than other methods. Moreover, the approach provides a practical matrix to estimate the target function.

Therefore, for the first time, the scour depth upstream and downstream of the submerged weirs was simulated by the modern method of self-adaptive extreme learning machine in the present study. In order to achieve this goal, Monte Carlo simulations were used to increase the ability of the artificial intelligence model and the k-fold cross validation method to validate the results of the modeling. Initially, the parameters were determined and the optimal activation function was then introduced. After that, by analyzing the results of simulations, the best models and the most effective input parameters were identified. In addition, the results of SAELM models were compared with ANN and SVM models and uncertainty analysis was performed for these superior models. In the following, several equations were proposed for superior SAELM models and partial derivative sensitivity analysis (PDSA) for all input parameters.

MATERIALS AND METHODS

In this study, the self-adaptive extreme learning machine (SAELM) is employed to estimate scour depth around submerged weirs. The SAELM is a combination of ELM and a self-adaptive differential evolution algorithm (SADE) to improve the capability of the ELM in function approximation. Indeed, the SADE is combined with ELM to calculate the bias of hidden layer and input weight analytically that randomly assigned in convectional ELM. An overview of ELM, SADE and finally SAELM is presented in sections 2.1, 2.2 and 2.3, respectively.

Extreme learning machine (ELM)

Recently, a learning algorithm has been proposed for a single-layer feed-forward neural network (SLFFNN), known as the extreme learning machine (Huang *et al.* 2004). In this model, hidden node parameters (input weights, hidden neurons) are randomly selected and output weights are analytically determined using Moore-Penrose generalized inverse (MPGI) (Lei *et al.* 2015). The ELM method avoids many problems, such as stopping criteria, learning rates and learning courses based on learning methods.

Suppose that there are N arbitrary samples in the training phase as $(x_i, y_i)\ \varepsilon\ R^n \times R^m$ $(i = 1, 2, \ldots, n)$ with L hidden nodes. Therefore, the SLFFNN model with the transfer function of f(x) can be expressed as follows (Aghbashlo *et al.* 2016; Ding *et al.* 2016; Liu *et al.* 2016):

$$\sum_{i=1}^{L} \beta_i f_i(x_j) = \sum_{i=1}^{L} \beta_i f(a_i \cdot b_i \cdot x), \quad j = 1, 2, \ldots, N \tag{1}$$

in which, $a_i = [a_{i1}, a_{i2}, \ldots, a_{in}]^T$ is the input weight vector connected to the hidden layer node, b_i is the bias of the hidden

layer nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T$ is the output weight vector connected to the hidden layer node.

The above equation can be rewritten as follows:

$$\sum_{i=1}^{L} \beta_i f_i(x_j) = H\beta \tag{2}$$

$$H = \begin{bmatrix} f(a_1 \cdot x_1 \cdot b_1) & \cdots & f(a_L \cdot x_1 \cdot b_L) \\ \vdots & \ddots & \vdots \\ f(a_1 \cdot x_N \cdot b_1) & \cdots & f(a_L \cdot x_N \cdot b_L) \end{bmatrix} \tag{3}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \tag{4}$$

$$T = \begin{bmatrix} y_1^T \\ \vdots \\ y_L^T \end{bmatrix}_{N \times m} \tag{5}$$

in which, H is the output matrix of the hidden layer, β is the output weight matrix, and T is called the label matrix.

The first step in this model is to determine the random values of input weights (α) and the bias of the hidden layer node (b) at the training phase. After determining these weights and biases, the output matrix of the hidden layer can be obtained from the input samples. Therefore, the SLFFNN training is converted into a least squared solution. After determining the input weights and the hidden layer biases, the output matrix of the hidden layer can be obtained from the training samples. Thus, SLFFNN training turns into the least square solution. By introducing the regularization theory for the ELM model, the objective function is expressed as follows:

$$\min L_{ELM} = \frac{1}{2}\|\beta\|^2 + \frac{c}{2}\|T - H\beta\|^2 \tag{6}$$

The least square solution of the above-mentioned equation is as follows:

$$V - cH^T(T - H\beta) = 0 \tag{7}$$

When the number of samples in the training phase are greater than the number of nodes in the hidden layer, one can write:

$$\beta = \left(\frac{1}{c} + H^T H\right)^{-1} H^T \tag{8}$$

When the number of samples in the training phase are less than the number of nodes in the pins, we have:

$$\beta = H^T \left(\frac{1}{c} + HH^T\right)^{-1} T \tag{9}$$

In the present study, the ELM non-linear activation functions include sigmoid (sig), sine (sin), hardlim, radial base (radbas) and triangular base (tribe).

Differential evolution (DE)

The differential evolution optimization algorithm (DE) is one of the most effective search-based methods (Price *et al.* 2005). For other evolutionary algorithms, this algorithm starts by creating an initial population. Then, by applying different operators, the newborn generation is formed and in the next stage (called the selection stage), the newborn generation with the parent generation is compared to evaluate the merit that is measured by the objective function. Then, the best members will enter as the next generation into the next step. This will continue to reach the desired results. In this section, the operation steps of this algorithm are respectively expressed, as follows:

Initial population

The number of variables in the problem is shown in Bot D algorithm. Each variable has one upper limit and one lower limit. An initial population of NP is randomly formed in D, as follows:

$$X_{io} = X_i \min + round(\delta_i \cdot (X_i \max - X_i \min)) \quad i = 1, 2, \ldots, NP \tag{10}$$

where, δ_i is a random number ranged between 0 and 1, X_i min and X_i max are respectively low and high limits of the variables in the problem and NP of the number of members.

Mutation and crossover

In this algorithm, five strategies can be used for the synthesis and production of nitrates (four strategies for mutations and one for crossover) defined as (Price *et al.* 2005):

$$Z_{i,G} = X_{r_1^i,G} + F.(X_{r_2^i,G} - X_{r_3^i,G}) \tag{11}$$

$$Z_{i,G} = X_{r_1^i,G} + F.(X_{\text{best},G} - X_{r_1^i,G}) + F.(X_{r_2^i,G} - X_{r_3^i,G}) + F.(X_{r_4^i,G} - X_{r_5^i,G}) \tag{12}$$

$$Z_{i,G} = X_{r_1^i,G} + F.(X_{r_2^i,G} - X_{r_3^i,G}) + F.(X_{r_4^i,G} - X_{r_5^i,G}) \tag{13}$$

$$Z_{i,G} = X_{i,G} + F.(X_{r_1^i,G} - X_{i,G}) + F.(X_{r_2^i,G} - X_{r_3^i,G}) \tag{14}$$

where, r_k^i are the integers attained randomly within the interval [1, 2, …, NP], F is the mutation factor, and X_{best} is the best member of the current population. Additionally, the crossover is defined as follows:

$$\text{if } u \le CR \text{ or } j = k \text{ then } Z_{i,j} = X_{r1,j} + CR(X_{r3,j} - X_{r2,j}) \text{ else } Z_{i,j} = X_{i,j} \tag{15}$$

where, j is the number of each variable from i^{th} member of the population and CR is a constant ranging between 0 and 1.

Selection

At this stage, the infants and parents are valued according to the objective function, and if the infant has a higher value than the parent, the parent will be replaced by an infant. Otherwise, the parent goes to the next stage along with the next generation.

$$Z_{i,g+1} = \arg\max\left(f(Z_i, g), f(Z_i, g+1)\right) \tag{16}$$

in which, the index g represents a generation, $Z_{i,g+1}$ is the population of the new generation (infants) and $Z_{i,g}$ of the previous generation population (parents). The function f is also the objective function.

Iteration

The sections 2.2.2 and 2.2.3 continue to reach the maximum iteration or convergence of the entire population.

Self-adaptive extreme learning machine (SAELM)

The method of the self-adaptive extreme learning machine is a hybrid method that uses the self-adaptive differential evolution algorithm (SADE) to increase the accuracy of ELM modeling (Cao *et al.* 2012). The SADE algorithm is used to estimate input weights and hidden node biases, and the output layer weights are determined according to the classical ELM method. First, an initial population vector, NP, is created using the SADE algorithm randomly in the first generation.

$$\theta_{k,G} = [a^T_{1,(k,G)}, \cdots, a^T_{L,(k,G)}, b_{1,(k,G)}, \cdots, b_{L,(k,G)}] \tag{17}$$

The output weight matrix is calculated using the following equation:

$$\beta_{k,G} = H^{-1}_{k,G} T \tag{18}$$

where, $H^{-1}_{k,G}$ is the generalized inverse of $H_{k,G}$. The $H_{k,G}$ is defined as follows:

$$H_{k,G} = \begin{bmatrix} g(a_{1,(k,G)}, b_{1,(k,G)}, x_1) & \cdots & g(a_{L,(k,G)}, b_{L,(k,G)}, x_1) \\ \vdots & \ddots & \vdots \\ g(a_{1,(k,G)}, b_{1,(k,G)}, x_N) & \cdots & g(a_{L,(k,G)}, b_{L,(k,G)}, x_N) \end{bmatrix} \tag{19}$$

The root mean squared error (RMSE) for each individual is calculated using the following equation:

$$RMSE_{k,G} = \sqrt{\frac{\sum_{i=1}^{N} \left| \sum_{j=1}^{L} \beta_j g(a_{j,(k,G)}, b_{j,(k,G)}, x_i) - t_i \right|}{m \times N}} \tag{20}$$

The population vector is maintained with the best value for RMSE in the first generation. In later generations, vectors of the parameters are evaluated as follows:

$$\theta_{k,G+1} = \begin{cases} u_{k,G+1} & \text{if } RMSE_{\theta_{k,G}} - RMSE_{\theta_{k,G+1}} > \varepsilon.RMSE_{\theta_{k,G}} \\ u_{k,G+1} & \text{if } |RMSE_{\theta_{k,G}} - RMSE_{\theta_{k,G+1}}| < \varepsilon.RMSE_{\theta_{k,G}} \\ & \text{and } |\beta_{u_{k,G+1}}| < |\beta_{\theta_{k,}}| \\ \theta_{k,G} & \text{else} \end{cases} \tag{21}$$

In the SADE algorithm used in the SAELM method, the trial vector for each objective vector is generated using the

strategies defined for the mutation. The selection of strategy for each generation is generated according to the probable process $P_{l,G}$. The parameter $P_{l,G}$ is the probability of the selection of l^{th} strategy in the G^{th} generation. In the developed model, the value of l can be in the range of 1 to 4. The parameter $P_{l,G}$ is updated when G is less than or equal to P (P is the number of vectors generated per population). Otherwise, if G is greater than P, $P_{l,G}$ is calculated from the following equation:

$$P_{l,G}=\frac{\frac{\sum_{g=G-P}^{G-1} ns_{l,g}}{\sum_{g=G-P}^{G-1} ns_{l,g}+\sum_{g=G-P}^{G-1} nf_{l,g}}+\varepsilon}{\sum_{l=1}^{4} S_{l,G}} \tag{22}$$

where, $nf_{l,g}$ is the number of trial vectors generated using l^{th} in the g^{th} generation that has successfully entered the next generation, $ns_{l,g}$ is the number of trial vectors generated using l^{th} in the g^{th} generation that is removed from the next generation, ε is a constant and positive number. The values of F and CR are randomly chosen for each objective vector using the normal distribution function. The production of trial vectors for future generations is determined by $\theta_{k,G+1}$ (Equation (20)). The evolutionary process continues until reaching the required amount of merit.

Artificial neural network

An artificial neural network is an idea for processing the information that is inspired by the biological nervous system and deals with information processing like the brain. This system has come from a large number of extraordinary processing elements, called neurons, working together to solve a problem (Vafakhah 2012). Artificial neural networks have the ability to learn, call, and generalize data learning patterns. These networks can be trained for a particular action by adjusting the element weights. The neural networks are usually trained to convert a particular input into a specific output (Riad *et al.* 2004). In order to select the optimal structure of the artificial neural network, the number of neurons in the intermediate layer is chosen with trial and error, so much of this number may lead to overfitting, and the low number of neurons may reduce the model accuracy for the data which have not been used in the training phase. For modelling, feedforward neural network with the backpropagation algorithm with the hyperbolic tangent function were used as a transfer function (Emiroglu *et al.* 2011).

$$tansig\ (x)=\frac{2}{1+e^{-2x}}-1 \tag{23}$$

Support vector machine (SVM)

The support vector machine (SVM) has been presented by Vapnik (1998). In a SVM regression model, it is necessary to estimate the functional dependency of the dependent variable of y on a set of independent variables of x. It is assumed that, as in other regression problems, the relation between dependent and independent variables is determined by a certain function, f, plus an additional amount of noise:

$$y=f(x)+noise \tag{24}$$

Therefore, the main issue is to find the form of the function f, which can correctly predict the new cases that SVM has not experienced so far. This function is accessible by training the SVM model on a dataset (called a training set), which includes a process for the continuous optimization of the error function. Based on the definition of this error function, two examples of SVM models are: (1) regression models of the type 1, known as v-SVM models and type II regression models, known as ε-SVM. In this study, the ε-SVM model is used for its wide application in the regression problems in recent years. For this model, the error function is defined as:

$$E=\frac{1}{2}W^TW+C\sum_{i=1}^{N}\zeta_i+C\sum_{i=1}^{N}\zeta_i^* \tag{25}$$

The error function is to be minimized according to the following constraints:

$$\begin{aligned} &W^T\phi(x)+b-y_i\le\varepsilon+\zeta_i^*\\ &y_i-W^T\phi(x)-b\le\varepsilon+\zeta_i\\ &\zeta_i,\zeta_i^*\ge 0 \qquad i=1,2,\ \ldots,N \end{aligned} \tag{26}$$

where, C is the capacity constant, W is the coefficient vector, WT is the transpose of the W, ζ_i and ζ_i^* are slack variables, b

is a constant, N is the training samples, and φ is a kernel function. The radial basis function (RBF) kernel is one of the best kernel functions in most of the regression problems in hydropower and hydrology sciences (Dibike *et al.* 2001; Noori *et al.* 2009a, 2009b). The RBF kernel function is defined as follows:

$$K(x_i, x) = \exp\left(-\gamma|x_i - x|^2\right) \tag{27}$$

According to the above-mentioned relations, the parameters C, ε, and γ must be defined before the start of modeling. The values of these parameters are calculated using trial and error.

EXPERIMENTAL MODEL

In this study, the experimental data measured by Guan *et al.* (2014, 2016); Wang *et al.* (2018a, 2018b) were used for the verification of the results of a numerical model. Their experimental model includes a flume with length of 12 m, a depth of 0.38 m and a width of 0.44 m. They measured the parameters of the equilibrium scour depth downstream of the weir ($d_{s\text{-}a}$), the maximum scour depth downstream of the weir ($d_{s\text{-}max}$), the equilibrium scour depth upstream of the weir ($d_{us\text{-}a}$) and the maximum scour depth upstream of the weir ($d_{us\text{-}max}$). A schematic view of the experimental model and the scour process around the submerged weirs are depicted in Figure 1. The scour depth upstream and downstream of the submerged weirs (d_{us} and d_s) can be observed in this figure. In addition, the range of experimental data is presented in Table 1.

SCOUR AROUND THE SUBMERGED WEIRS

Wang *et al.* (2018b) considered the scour depth around the sloping submerged weirs (d_s) based on the sediment density (ρ_s), water density (ρ_w), kinematic viscosity of the fluid (ν), gravity acceleration (g), average inflow depth (h_0), tail-water depth (h_t), average flow velocity (U_o), average critical velocity of the inflow (U_c), average diameter of the sediment particles (d_{50}), the standard deviation of the particle sizes (σ_s), the width of the weir (b), the height of the weir (z) and the upstream slope of the weir (α).

$$ds = f(\rho_s, \rho, \nu, g, h_0, h_t, U_0, U_c, d_{50}, \sigma_s, b, z, \alpha) \tag{28}$$

They also considered a fixed value for the geometric characteristics of the experimental model and defined four dimensionless groups of the parameter. Equation (1) is re-written as follows:

$$ds/h_t = f\left(\frac{U_0}{U_c}, \frac{z}{h_t}, \frac{d_{50}}{h_t}, \frac{2\alpha}{\pi}\right) \tag{29}$$

Therefore, in the present study, the effects of dimensionless parameters of Equation (2) are considered as input parameters for the artificial intelligence models. Figure 2 shows the combination of input parameters for numerical models.

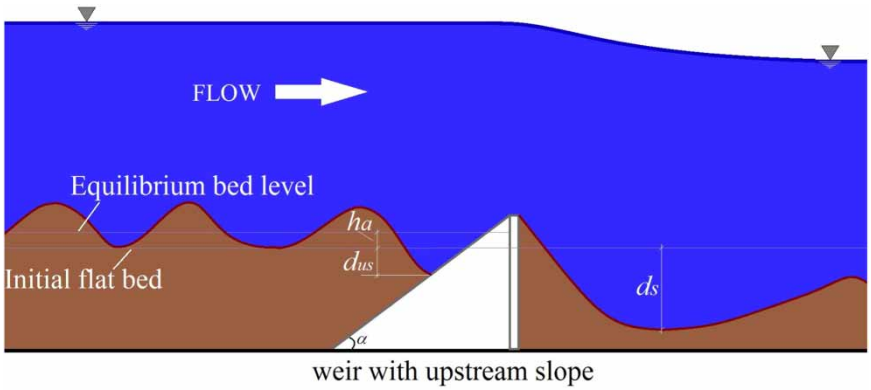

Figure 1 | Schematic view of the experimental model and scour process around the sloping submerged weir.

Table 1 | The range of experimental data used in this study

Parameter	Max	Min	Ave
U_o/U_c	3.696	0.722	2.209
z/h_t	0.333	0.133	0.215
d_{50}/h_t	0.007	0.002	0.005
$2\alpha/\pi$	1	0.167	0.835
$d_{s\text{-}a}/h_t$	1.507	0.113	0.620
$d_{s\text{-}max}/h_t$	1.673	0.120	0.773
$d_{us\text{-}a}/h_t$	0.740	0	0.354
$d_{us\text{-}max}/h_t$	1.333	0	1.333

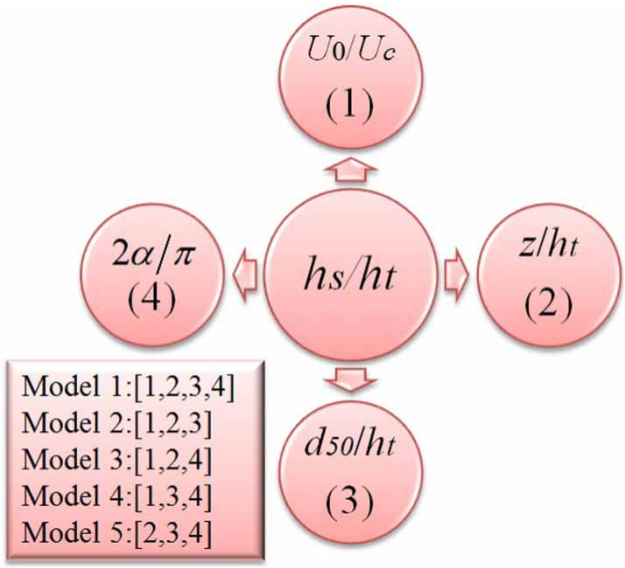

Figure 2 | The combination of input parameters for artificial intelligence models.

In this study, Monte Carlo simulations are used to assess the ability of extreme learning machine models. Monte Carlo simulations are a widespread classification of computational algorithms that use random sampling to calculate numerical results. In addition, the k-fold cross validation is used to evaluate the performance of these models. In a k-fold cross validation method, the main sample is randomly divided into k sub-samples of equal size. Among k sub-samples, one sub-sample is used for validation and the remaining are used to test this model. The k-fold cross validation process then repeats k times (equal to the number of layers), and each k sub-sample is exactly used one time as the validation data. The results obtained from the k-fold are averaged and presented as an estimate. The advantage of this method is the random repetition of sub-samples in the testing and training processes for all observations, and each observation is used exactly one time for the validation. In this study, the value of k is assumed to be equal to 4. In addition, the schematic view of the k-fold cross validation method and how to deal with the testing and training data are shown in Figure 3.

PERFORMANCE EVALUATION CRITERIA

In order to evaluate the accuracy of numerical models, the statistical indices of R^2, RMSE, MARE, (VAF) and scatter index (SI) was used as follows:

$$R^2 = \frac{\left(\frac{n\sum_{i=1}^{n} R_{(Predicted)i}R_{(Observed)i}}{-\sum_{i=1}^{n} R_{(Predicted)i}\sum_{i=1}^{n} R_{(Observed)i}}\right)^2}{\left(n\sum_{i=1}^{n}(R_{(Predicted)i})^2\sum_{i=1}^{n}(R_{(Predicted)i})^2\right)\left(n\sum_{i=1}^{n}(R_{(Observed)i})^2-\sum_{i=1}^{n}(R_{(Observed)i})^2\right)} \quad (30)$$

$$\mathrm{VAF} = \left(1 - \frac{\mathrm{var}(F_i - O_i)}{\mathrm{var}(F_i)}\right) \times 100 \quad (31)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(R_{(Predicted)i} - R_{(Observed)i})^2} \quad (32)$$

$$SI = \frac{RMSE}{(\overline{R})_{(Observed)}} \quad (33)$$

$$MAE = \frac{\sum_{i=1}^{n}|R_{(Predicted)i} - R_{(Observed)i}|}{n} \quad (34)$$

$$MARE = 100 \times \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|R_{(Predicted)i} - R_{(Observed)i}|}{R_{(Predicted)i}}\right) \quad (35)$$

in which, $(R)_{(observed)i}$, $(R)_{(predicted)i}$, $(\bar{R})_{(observed)i}$, and n are, respectively, experimental values, predicted results by numerical models, average experimental values, and the number of experimental data. The closeness of the MARE, RMSE, MAE, and SI to zero indicates the high accuracy of the numerical model. In addition, the closeness of the R2 index to 1 reveals high correlation of the numerical model. In general, the superior model has a larger VAF than other numerical models.

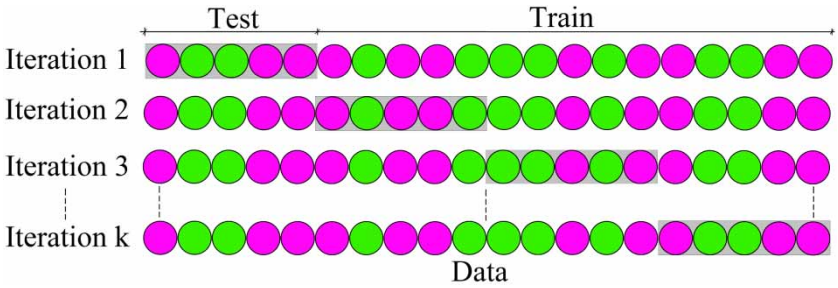

Figure 3 | Schematic view of the k-fold cross validation method and how to deal with testing and training data.

In this study, first, the optimal activation function is selected for all scouring values of $d_{s\text{-}a}/h_t$, $d_{s\text{-}max}/h_t$, $d_{us\text{-}a}/h_t$, and $d_{us\text{-}max}/h_t$. Then, the results of simulation by all SAELM models are evaluated. The modeling results are then compared with the artificial neural networks and the support vector machine models. In the next section, uncertainty analysis is performed on superior models. In addition, partial derivative sensitivity analysis is performed for superior models. Finally, for each superior model, an equation for estimation of $d_{s\text{-}a}/h_t$, $d_{s\text{-}max}/h_t$, $d_{us\text{-}a}/h_t$, and $d_{us\text{-}max}/h_t$ is developed.

RESULTS AND DISCUSSION

Activation function selection

In this section, an optimal activation function for simulating scouring around submerged weirs is investigated. As previously mentioned, the extreme learning machine has five activation functions entitled sigmoid, sin, hardlimit, tribas, and radbas. The comparison of experimental and simulated data along with scatter charts for these activation functions is shown in Figure 4. For example, for estimation of $d_{s\text{-}a}$, the activation function of sigmoid was selected as the most accurate function. The values of R^2, RMSE, and MAE for this activation function were calculated to be equal to 0.878, 0.095 and 0.068. In addition, the VAF and SI indices for estimation of $d_{s\text{-}a}$ by the activation function of sigmoid were 87.826 and 0.154, respectively. Also, to simulate the parameter $d_{s\text{-}max}$, the sigmoid function was more accurate than the other activation functions. For this activation function, the values of R^2 and SI were estimated to be equal to 0.862 and 0.136, respectively. It should be noted that for both parameters of $d_{us\text{-}a}/h_t$ and $d_{us\text{-}max}/h_t$, the sigmoid activation function was selected as the optimum activation function. For the parameter $d_{us\text{-}a}/h_t$, the RMSE and MARE indices of the sigmoid activation function were calculated to be 0.052 and 0.263, respectively. In addition, the values of VAF, SI and MAE for simulation of $d_{us\text{-}max}/h_t$ by sigmoid activation function were respectively 90.788, 0.175 and 0.083.

Therefore, on the basis of the results of all activation functions, the sigmoid activation function was identified as the optimal function, and this function is used to estimate the scouring values. The comparison of experimental and simulated data along with scatter charts for the activation function is shown in Figure 4.

Input combination selection

In this section, all SAELM models are evaluated in predicting the scour values around the submerged weirs. The most effective input parameter is also introduced. As mentioned in the previous sections, five distinct artificial intelligence models have been developed in this study. Comparison of the results of the statistical indices for various self-adaptive extreme training models can be seen in Figure 5. For example, for simulation of $d_{s\text{-}a}/h_t$, the SAELM1 model calculated the R2 and the SI equal to 0.878 and 0.154, respectively. For this model, the VAF, MARE, and RMSE indices were estimated as 87.826, 0.122 and 0.095, respectively. This model calculates the scour values for all input parameters of (U_o/U_c, z/h_t, d_{50}/h_t, $2\alpha/\pi$). In this study, in order to identify the most effective input parameter, four SAELM models were introduced with three input

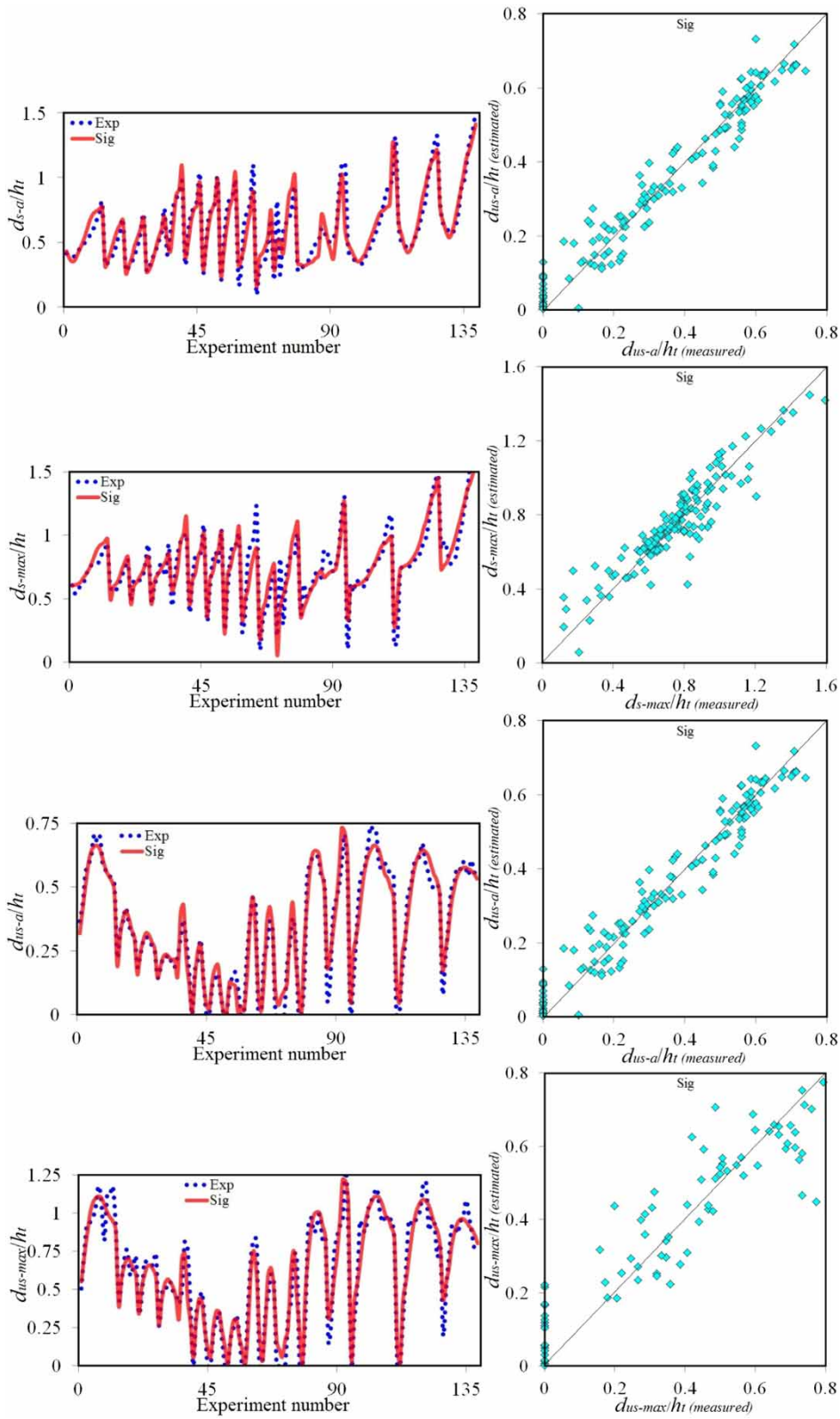

Figure 4 | Comparison of experimental scouring values with simulated values and scatter charts for various activation functions.

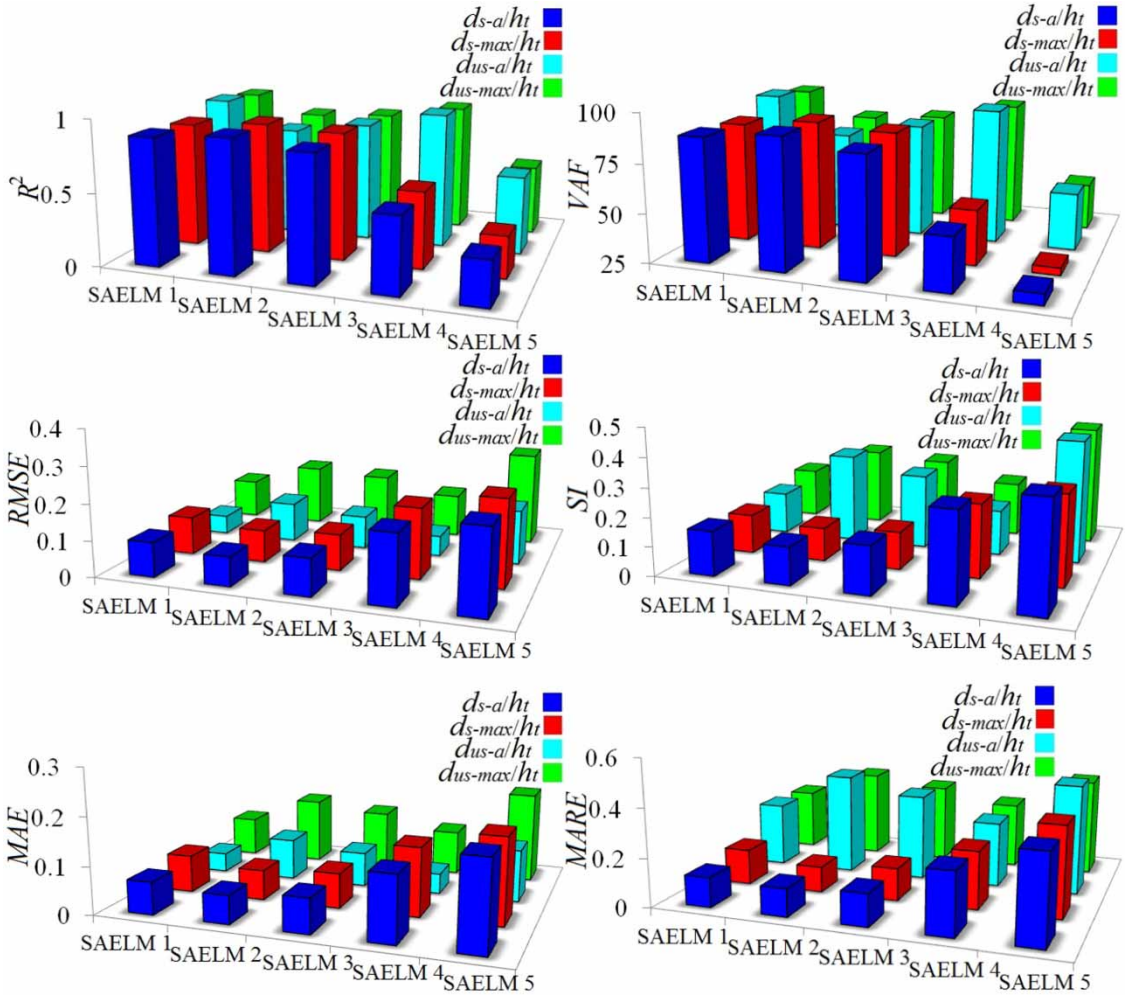

Figure 5 | Comparison of different statistical indices for SAELM models.

parameters, so that by removing each input parameter, the results of the model were extracted and error values were calculated. For example, for the SAELM2 model, the effect of the parameter $2\alpha/\pi$ is neglected. In other words, the model has simulated the scour values in terms of U_o/U_c, z/h_t, and d_{50}/h_t. Among all numerical models, the SAELM2 model has the highest accuracy in simulating scouring downstream of the submerged weirs. For this model, the values of SI and R^2 are equal to 0.130 and 0.913, respectively. In addition, the RMSE and MARE indices were 0.081 and 0.114, respectively. The SAELM3 model was a function of U_o/U_c, z/h_t, and $2\alpha/\pi$. For this model, the effect of the dimensionless parameter of d_{50}/h_t was removed and the values of R^2, RMSE, and VAF were calculated as 0.861, 0.081, and 91.286. For the SAELM3 model, the SI and MARE indices were equal to 0.164 and 0.132. The effect of parameter z/h_t was also ignored for the SAELM4 model. This means that the model was a function of U_o/U_c, d_{50}/h_t, and $2\alpha/\pi$. For the SAELM4 model, the R^2 was estimated equal to 0.519. For this model, the statistical indices of RMSE and VAF were also calculated to be 0.189 and 51.907, respectively. Additionally, the SAELM5 model estimates the values of the objective function in terms of z/h_t, d_{50}/h_t, and $2\alpha/\pi$. In other words, the effect of U_o/U_c has been eliminated for the simulation of scouring downstream of the submerged weirs. According to the results of modeling of the values of the objective function by artificial intelligence models, the SAELM5 model had the highest error value. For example, the values of VAF, MARE, and SI for this model were 30.487, 0.359, and 0.367, respectively.

Additionally, the SAELM2 model was introduced as the superior model for estimation of $d_{s\text{-}max}/h_t$. For this model, the values of R^2, VAF and SI were calculated equal to 0.901, 90.115, and 0.115, respectively. Also, for simulation of $d_{us\text{-}a}/h_t$ and $d_{us\text{-}max}/h_t$, the SAELM1 model was identified as the superior model. The MAE was calculated for modeling of $d_{us\text{-}a}/h_t$ and $d_{us\text{-}max}/h_t$ by the SAELM1 model, respectively equal to 0.040 and 0.083.

It should be noted that for all objective functions including $d_{s\text{-}a}/h_t$, $d_{s\text{-}max}/h_t$, $d_{us\text{-}a}/h_t$ and $d_{us\text{-}max}/h_t$, with removing the parameter U_o/U_c, the accuracy of the modeling was significantly decreased. Therefore, the ratio of U_o/U_c was identified as the most effective input parameter.

The comparison of observed and simulated values of scouring along with the scatter diagram for the superior SAELM models are shown in Figure 6.

Comparison of superior models with ANN and SVM

In the following, the results of the superior models for estimation of $d_{s\text{-}a}/h_t$, $d_{s\text{-}max}/h_t$, $d_{us\text{-}a}/h_t$ and $d_{us\text{-}max}/h_t$ are compared with the results of ANN and SVM models (Figure 7). Based on the results of modeling, for modeling all objective functions, SAELM models were identified as superior models. For example, the values of R^2 and SI were calculated for modeling the parameter of $d_{s\text{-}a}/h_t$ by ANN model as 0.863 and 0.163, respectively. In addition, the values of MARE and RMSE for estimation of $d_{s\text{-}a}/h_t$ by SVM model were respectively 0.114 and 0.089. In contrast, for estimation of $d_{s\text{-}max}/h_t$ by the ANN and SVM models, the value of VAF was calculated as 83.006 and 83.890, respectively. However, the R^2 index for modeling $d_{us\text{-}a}/h_t$ using ANN and SVM models was calculated to be 0.861 and 0.898, respectively. For modeling of $d_{us\text{-}max}/h_t$ by ANN model, the SI and MARE indices were respectively 0.246 and 0.341. Additionally, for modelling $d_{us\text{-}max}/h_t$ by the SVM model, the values of VAF, MAE, and RMSE were estimated to be 84.523, 0.113, and 0.141, respectively.

Uncertainty analysis

In this section, the uncertainty analysis of superior models is implemented. The uncertainty analysis is used to estimate the predicted error by numerical models in which the predicted error by the numerical model (e_j) is calculated as the difference between the predicted (P_j) and observed (T_j) values ($e_j = P_j - T_j$). In addition, the average value of the predicted error is also obtained as:

$$\bar{e} = \sum_{j=1}^{n} e_j \tag{36}$$

Additionally, the standard deviation of the predicted error is defined as:

$$S_e = \sqrt{\sum_{j=1}^{n} (e_j - \bar{e})^2 / n - 1} \tag{37}$$

The negative values of $\bar{e}$ represent the underestimated function of the numerical model. In contrast, positive values reveal the overestimated function of the numerical model. It should be noted that using the parameters $\bar{e}$ and S_e a confidence band is produced around the predicted errors by the Wilson score method without continuity correction. In the following, using $\pm 1.64 S_e$ may lead to a confidence band of 95%. Based on the results of uncertainty analysis, the SAELM2 model, which was used to estimate the parameters $d_{s\text{-}a}$ and $d_{s\text{-}max}$, had an overestimated function. On the other hand, the SAELM1 model, which was used for the prediction of $d_{us\text{-}a}$ and $d_{us\text{-}max}$ was underestimated. Also, the width of the uncertainty band used to estimate the parameters of $d_{s\text{-}a}$ and $d_{s\text{-}max}$ by the SAELM2 model was -0.014 and -0.015, respectively. Additionally, the 95% prediction error interval of the SAELM2 model for simulations of $d_{us\text{-}max}$ was between -0.024 and 0.013. Table 2 shows the results of uncertainty analysis for the superior SAELM models.

According to the results of numerical models, the SAELM2, SAELM2, SAELM2 and SAELM2 models respectively estimated the values of $d_{s\text{-}a}$, $d_{s\text{-}max}$, $d_{us\text{-}a}$, and $d_{us\text{-}max}$ with higher accuracy than the other models. Therefore, an equation is developed for each model.

The development of SAELM-based equations for local scouring around the submerged weirs is as follows:

$$\frac{d_S}{h_t} = \left[\frac{1}{(1 + \exp(InW \times InV + BHN))}\right]^T \times OutW \tag{38}$$

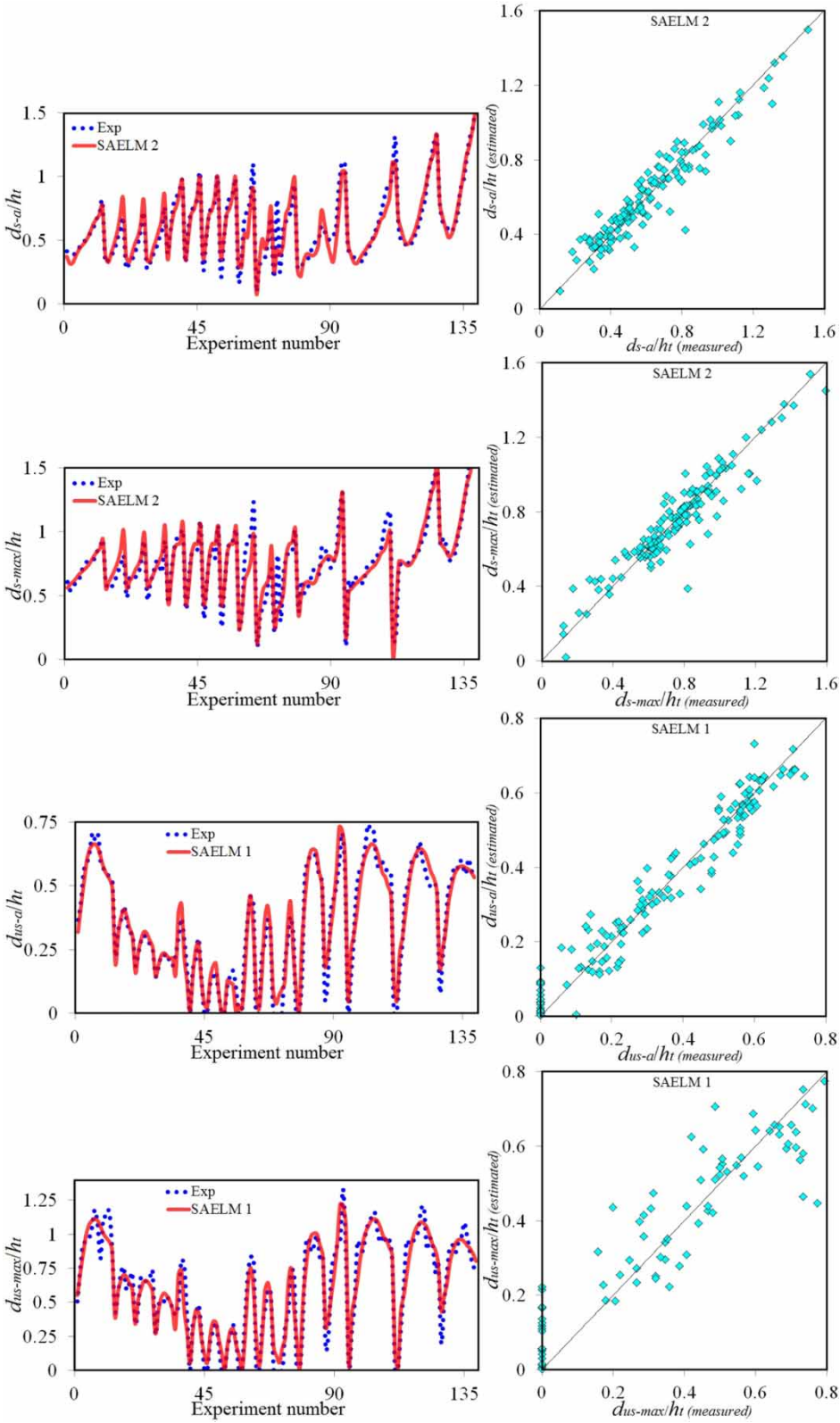

Figure 6 | Comparison of observed and simulated values of scouring and scatter diagram for the superior SAELM models.

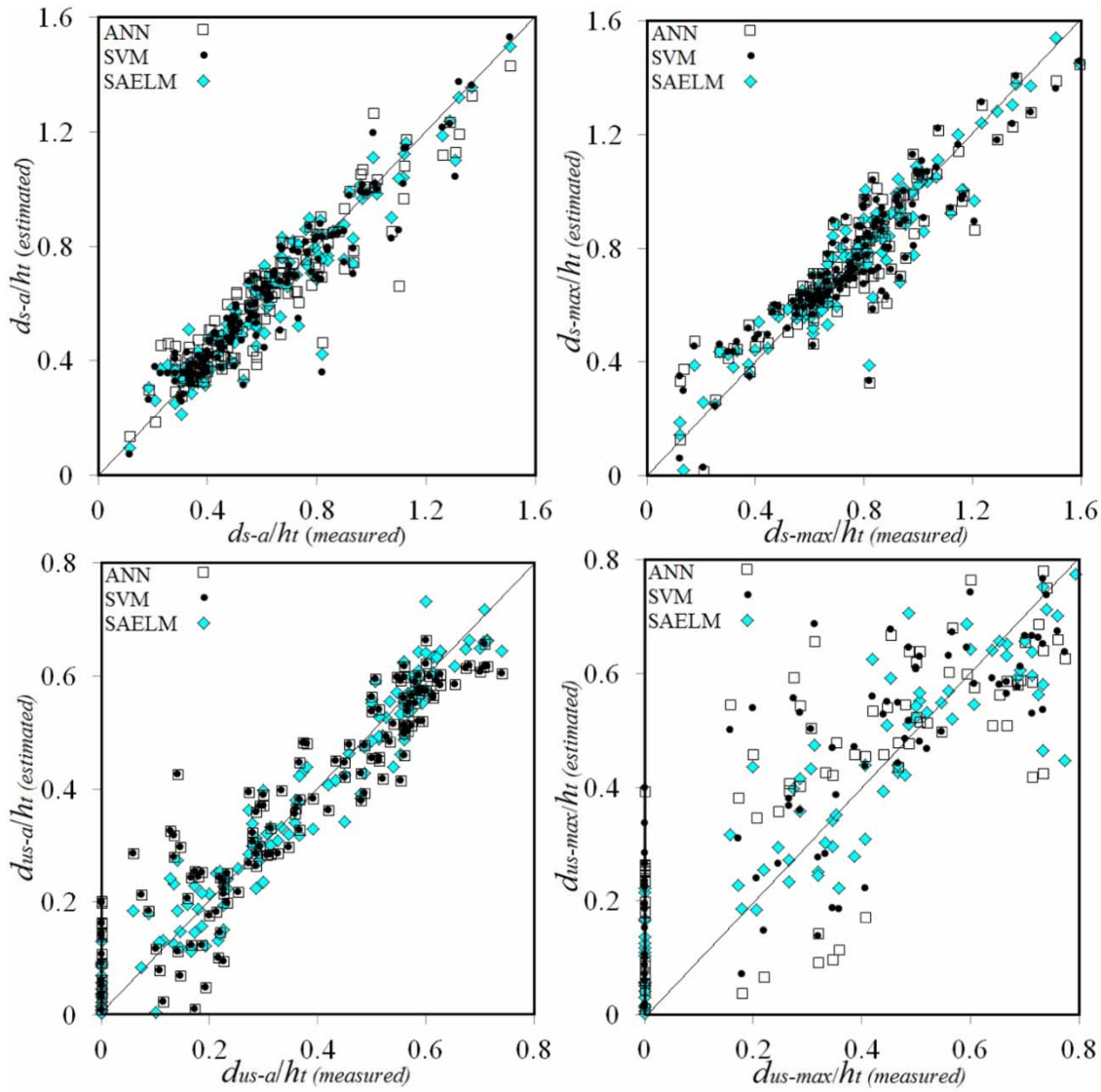

Figure 7 | Comparison of the results of superior models of ANN and SVM.

Table 2 | The results of uncertainty analysis for superior SAELM models

Target parameter	Models	Number of samples	Mean prediction error	Standard Division	S_e	Width of uncertainty band	95% prediction error interval
d_{s-a}	SAELM 2	139	1.271E-09	0.081	0.007	−0.014	−0.014 and 0.014
$d_{s-\max}$	SAELM 2	139	1.050E-09	0.089	0.008	−0.015	−0.015 and 0.015
d_{us-a}	SAELM 1	139	−0.004	0.052	0.004	−0.009	−0.013 and 0.005
$d_{us-\max}$	SAELM 1	139	−0.005	0.109	0.009	−0.019	−0.024 and 0.013

where, InW is the matrix of input weights, InV is the matrix of the input variables, BHN is the matrix of bias hidden neurons, and OutW is the matrix of output weights. The optimal values of these matrices are as follows.

For estimation of $d_{s\text{-}a}$ using the SAELM2 model:

$$InV = \begin{bmatrix} U_0/U_c \\ z/h_t \\ d_{50}/h_t \end{bmatrix} \quad BHN = \begin{bmatrix} 0.429 \\ 0.416 \\ 0.088 \\ 0.308 \\ 0.769 \\ 0.657 \\ 0.673 \\ 0.699 \\ 0.133 \\ 0.991 \\ 0.648 \\ 0.018 \\ 0.196 \\ 0.812 \\ 0.081 \\ 0.447 \\ 0.118 \\ 0.882 \\ 0.746 \\ 0.336 \end{bmatrix} \quad InW = \begin{bmatrix} -0.755 & 0.703 & -0.906 \\ -0.755 & -0.684 & -0.136 \\ 0.484 & -0.762 & -0.443 \\ -0.916 & -0.700 & 0.233 \\ 0.187 & 0.449 & -0.267 \\ 0.785 & -0.893 & 0.372 \\ 0.358 & 0.891 & -0.415 \\ 0.027 & -0.922 & 0.813 \\ 0.424 & -0.362 & -0.546 \\ -0.734 & -0.139 & 0.822 \\ -0.510 & -0.712 & 0.270 \\ -0.233 & -0.305 & 0.773 \\ -0.760 & -0.743 & 0.075 \\ 0.171 & 0.069 & 0.505 \\ 0.839 & 0.601 & 0.266 \\ 0.365 & -0.833 & -0.660 \\ 0.758 & 0.584 & -0.364 \\ 0.769 & 0.913 & 0.439 \\ -0.626 & -0.101 & 0.314 \\ 0.191 & -0.971 & 0.797 \end{bmatrix} \quad OutW = \begin{bmatrix} -5182.117 \\ 41461.834 \\ -25410.525 \\ -9397.317 \\ 132049.116 \\ 2101.598 \\ -70098.007 \\ 8945.292 \\ 37064.258 \\ -1796.013 \\ -20649.296 \\ 45128.990 \\ -77697.496 \\ -49552.988 \\ -31058.601 \\ 6096.474 \\ -34604.612 \\ 27892.949 \\ 9202.837 \\ -432.649 \end{bmatrix} \tag{39}$$

For estimation of $d_{s\text{-}max}$ using the SAELM2 model:

$$InV = \begin{bmatrix} U_0/U_c \\ z/h_t \\ d_{50}/h_t \end{bmatrix} \quad BHN = \begin{bmatrix} 0.100 \\ -0.688 \\ -0.104 \\ -0.855 \\ -0.664 \\ 0.709 \\ 0.247 \\ -0.959 \\ -0.633 \\ -0.200 \\ -0.744 \\ -0.369 \\ -0.242 \\ -0.231 \\ 0.532 \\ -0.263 \\ -0.963 \\ 0.934 \\ -0.524 \\ -0.560 \end{bmatrix} \quad InW = \begin{bmatrix} 0.687 & -0.632 & -0.228 \\ -0.757 & -0.603 & 0.389 \\ -0.178 & 0.168 & -0.656 \\ -0.167 & 0.144 & 0.344 \\ -0.819 & -0.583 & 0.174 \\ 0.658 & -0.212 & -0.033 \\ 0.399 & -0.788 & 0.223 \\ -0.057 & 0.309 & 0.139 \\ -0.235 & -0.507 & 0.770 \\ 0.663 & 0.903 & -0.802 \\ 0.322 & -0.089 & -0.160 \\ 0.125 & 0.690 & 0.266 \\ -0.909 & 0.132 & 0.382 \\ -0.429 & -0.589 & -0.018 \\ -0.033 & -0.379 & 0.377 \\ 0.390 & 0.0946 & -0.982 \\ -0.510 & -0.437 & 0.999 \\ 0.274 & 0.163 & 0.999 \\ 0.847 & -0.049 & 0.355 \\ 0.644 & 0.658 & 0.139 \end{bmatrix} \quad OutW = \begin{bmatrix} 3077.385 \\ -140328.103 \\ 6517.276 \\ -145387.976 \\ 64905.481 \\ 37818.822 \\ -3174.665 \\ 227343.025 \\ -9995.323 \\ -18301.876 \\ 53705.396 \\ 12972.207 \\ 13117.018 \\ -57252.748 \\ 115142.823 \\ -30088.598 \\ 135487.140 \\ -161550.463 \\ -1039.080 \\ 14655.244 \end{bmatrix} \tag{40}$$

For estimation of $d_{us\text{-}a}$ using the SAELM1 model:

$$InV = \begin{bmatrix} U_0/U_c \\ z/h_t \\ d_{50}/h_t \\ 2\alpha/\pi \end{bmatrix} \quad BHN = \begin{bmatrix} -0.052 \\ -0.931 \\ 0.424 \\ 0.855 \\ 0.286 \\ -0.694 \\ -0.276 \\ 0.548 \\ 0.303 \\ 0.609 \\ -0.237 \\ 0.723 \\ -0.181 \\ 0.646 \\ -0.368 \\ 0.705 \\ 0.431 \end{bmatrix} \quad InW = \begin{bmatrix} 0.369 & 0.626 & -0.276 & 0.927 \\ -0.275 & -0.707 & -0.194 & -0.709 \\ -0.571 & -0.826 & 0.261 & -0.461 \\ -0.887 & -0.643 & -0.341 & -0.029 \\ -0.775 & 0.262 & 0.566 & -0.347 \\ 0.861 & -0.088 & -0.472 & 0.209 \\ -0.168 & 0.356 & -0.702 & -0.145 \\ -0.178 & 0.807 & -0.830 & -0.306 \\ -0.976 & -0.886 & -0.507 & -0.560 \\ 0.707 & -0.053 & -0.546 & 0.086 \\ 0.938 & -0.157 & 0.876 & -0.670 \\ -0.808 & -0.628 & -0.812 & 0.538 \\ 0.626 & 0.598 & -0.0715 & -0.025 \\ -0.533 & 0.044 & 0.006 & -0.664 \\ -0.595 & 0.989 & -0.274 & -0.337 \\ 0.764 & -0.180 & 0.763 & 0.182 \\ -0.121 & -0.194 & 0.193 & 0.124 \end{bmatrix} \quad OutW = \begin{bmatrix} 239.348 \\ 447.778 \\ -354.972 \\ -197.708 \\ 116.239 \\ -27.718 \\ -1759.703 \\ 716.108 \\ 51.371 \\ 1089.362 \\ -13.691 \\ -27.427 \\ -688.459 \\ 252.449 \\ 71.504 \\ -652.324 \\ 347.021 \end{bmatrix} \tag{41}$$

For estimation of $d_{us\text{-}max}$ using the SAELM1 model:

$$InV = \begin{bmatrix} U_0/U_c \\ z/h_t \\ d_{50}/h_t \\ 2\alpha/\pi \end{bmatrix} \quad BHN = \begin{bmatrix} -0.260 \\ -0.488 \\ 0.325 \\ -0.825 \\ -0.371 \\ -0.999 \\ -0.714 \\ 0.733 \\ -0.986 \\ -0.798 \\ 0.034 \\ 0.092 \\ -0.731 \\ -0.077 \\ 0.638 \\ 0.271 \\ -0.231 \end{bmatrix} \quad InW = \begin{bmatrix} -0.299 & -0.671 & 0.991 & 0.391 \\ 0.005 & -0.371 & -0.233 & 0.676 \\ 0.901 & -0.869 & -0.646 & -0.031 \\ -0.945 & 0.277 & 0.986 & -0.030 \\ 0.048 & 0.173 & -0.513 & -0.940 \\ 0.454 & -0.318 & 0.212 & -0.247 \\ 0.581 & 0.445 & 0.583 & 0.090 \\ -0.360 & -0.371 & 0.451 & 0.216 \\ 0.432 & -0.021 & -0.385 & -0.207 \\ -0.034 & -0.670 & 0.031 & -0.765 \\ -0.442 & -0.019 & -0.856 & 0.030 \\ 0.606 & -0.077 & 0.224 & 0.296 \\ 0.315 & 0.222 & 0.921 & -0.878 \\ -0.315 & 0.219 & 0.730 & 0.413 \\ -0.791 & -0.336 & -0.747 & 0.663 \\ -0.619 & -0.751 & 0.770 & 0.985 \\ -0.610 & -0.966 & -0.326 & 0.064 \end{bmatrix} \quad OutW = \begin{bmatrix} -641.892 \\ 287.941 \\ 233.214 \\ -90.424 \\ 445.982 \\ 2409.836 \\ 533.799 \\ -625.087 \\ -3043.478 \\ -620.429 \\ 844.119 \\ -289.120 \\ -34.748 \\ 153.656 \\ -101.050 \\ 152.261 \\ 305.846 \end{bmatrix} \tag{42}$$

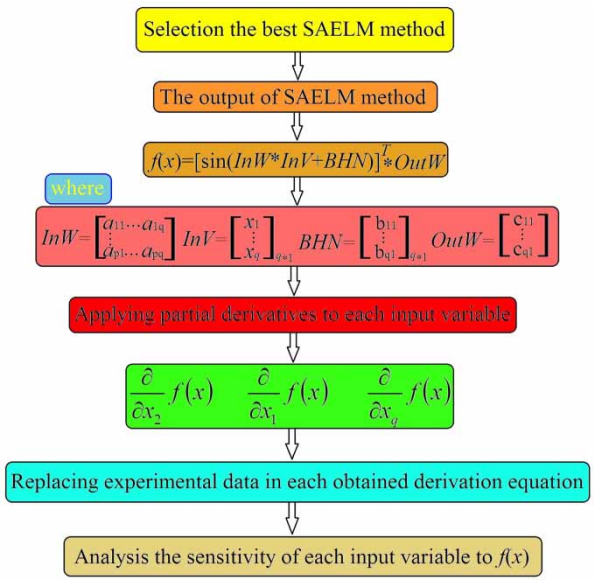

Figure 8 | Sensitivity analysis flowchart used in this study.

Partial derivative sensitivity analysis (PDSA)

In the following, partial derivative sensitivity analysis (PDSA) is considered for superior models and calibration of input parameters. In other words, the partial derivative sensitivity analysis (PDSA) is one of the most important methods for identifying the pattern of changes in input parameters (Azimi *et al.* 2017b). In general, the positive values of PDSA indicate the increase of the objective function; in contrast, the negative values mean a decrease in the output value. The sensitivity analysis flowchart used in this study is depicted in Figure 8. As shown in this figure, the relative derivative of f(x) is used for each input variable. It should be noted that each input variable is a matrix k × l, in which k is the number of samples.

The results of partial derivative sensitivity analysis for the input parameters of the SAELM2 and SAELM1 superior models that simulate the values of $d_{s\text{-}a}$, $d_{s\text{-}max}$, $d_{us\text{-}a}$, and $d_{us\text{-}max}$ are shown in Figures 9–12. For example, for $d_{s\text{-}a}$ and $d_{s\text{-}max}$, by increasing the ratio of U_o/U_c, the value of PDSA increases. In contrast, for the parameters $d_{us\text{-}a}$ and $d_{us\text{-}max}$, the sensitivity analysis decreased when increasing the ratio of U_o/U_c. In addition, in the simulation of the parameter $d_{us\text{-}max}$ by the SAELM1 model, the sensitivity analysis decreased with increasing z/h_t.

CONCLUSION

In this paper, the scour depth upstream and downstream of submerged weirs was simulated by the self-adaptive extreme learning machine (SAELM) model for the first time. In addition, for validation of SAELM model results, k-fold cross validation method was used, the value of k in this study was assumed to be 4. Using non-dimensional input parameters, five different SAELM models (SAELM 1 to SAELM 5) were defined. Also, the sigmoid activation function was introduced as the optimal activation function. The determination coefficient (R^2) and the scatter index (SI) for the activation function were calculated equal to 0.944 and 0.147, respectively. Then, by analyzing the results of modeling, SAELM 1 and SAELM 2 models were introduced for estimation of scour depth upstream and downstream of the submerged weirs. The models had a reasonable performance for simulating the

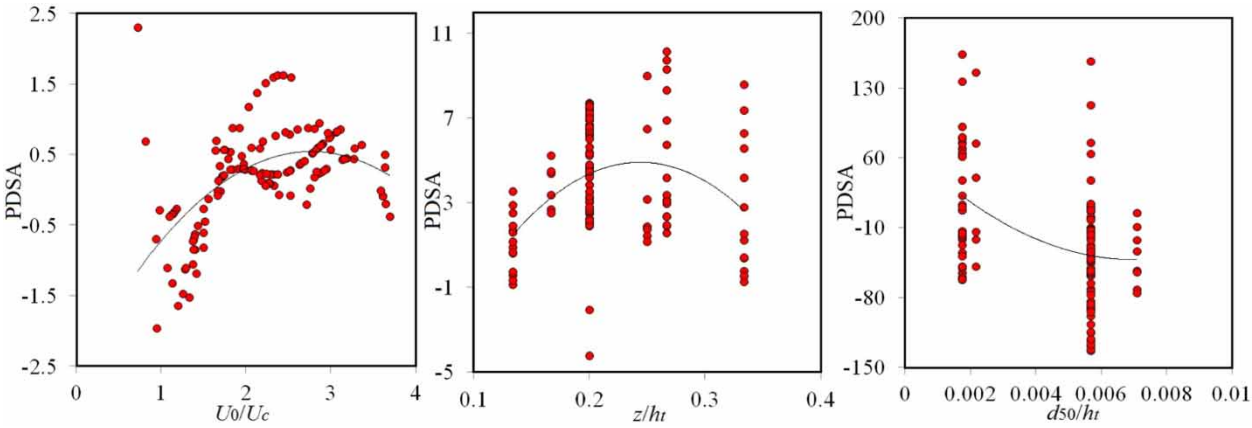

Figure 9 | The results of PDSA for simulation of $d_{s\text{-}a}$ by the SAELM2 model.

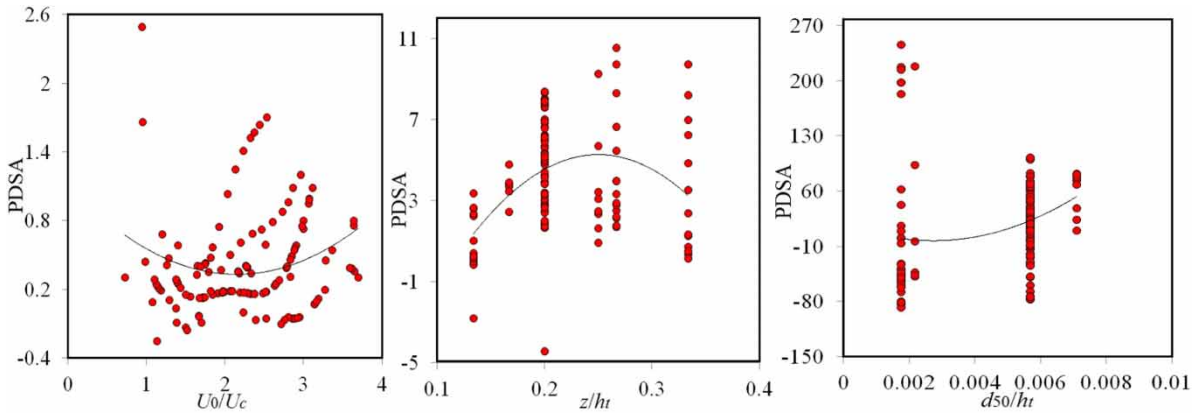

Figure 10 | The results of PDSA for simulation of $d_{s\text{-}max}$ by the SAELM2 model.

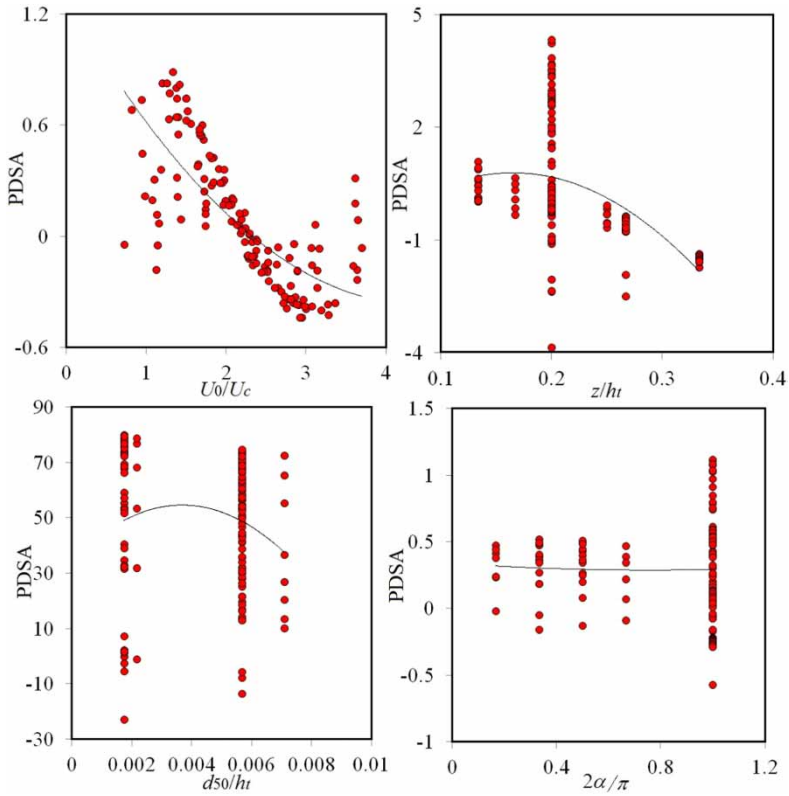

Figure 11 | The results of PDSA for simulation of $d_{us\text{-}a}$ by the SAELM1 model.

scour depth, for instance, the RMSE and MARE for the best model were respectively estimated as 0.052 and 0.040, respectively. In addition, the dimensionless parameter of U_o/U_c was detected as the most effective input parameter. Also, it was shown that the SAELM model had better performance than classical artificial intelligence techniques. Finally, some equations were presented for calculating the scour depth around the submerged weirs and partial derivative sensitivity analysis (PDSA) was performed for all input parameters of these equations. All in all, SAELM was identified as a reliable method to simulate scour depth in the vicinity of submerged weirs. The model had acceptable accuracy and presented matrices can be utilized by engineers and scholars without knowledge of soft computing.

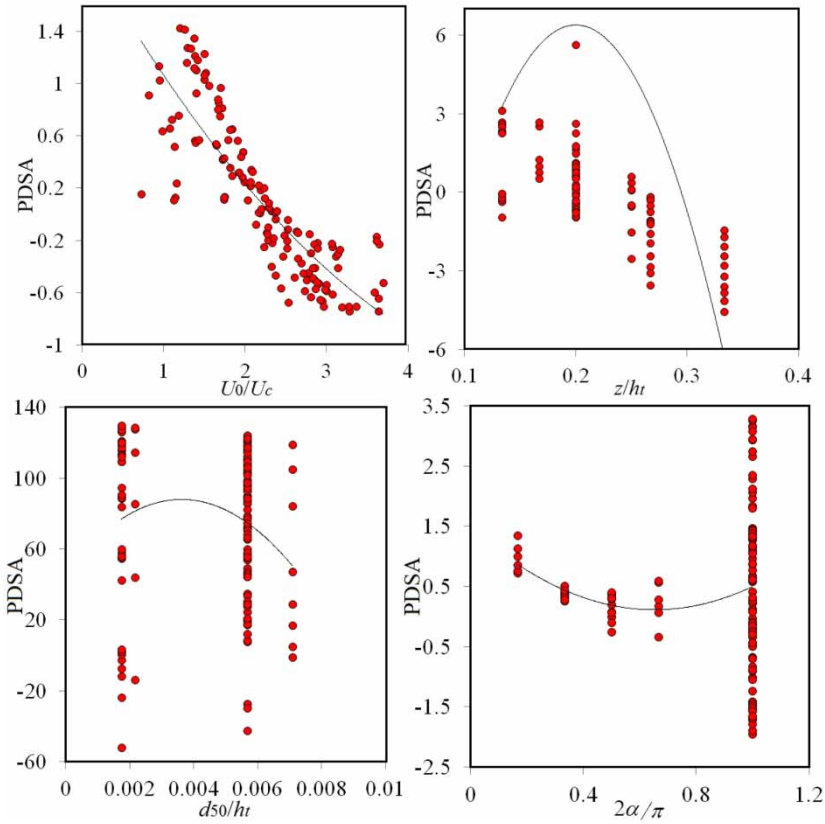

Figure 12 | The results of PDSA for simulation of $d_{us\text{-}max}$ by SAELM1 model.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interest: The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

Aghbashlo, M., Shamshirband, S., Tabatabaei, M., Lip Yee, P. & Nabavi Larimi, Y. 2016 The use of ELM-WT (extreme learning machine with wavelet transform algorithm) to predict exergetic performance of a DI diesel engine running on diesel/biodiesel blends containing polymer waste. *Energy* **94**, 443–456. doi:10.1016/j.energy.2015.11.008.

Azimi, H., Bonakdari, H., Ebtehaj, I. & Ashraf Talesh, S. H. 2017a Evolutionary Pareto optimization of an ANFIS network for modeling scour at pile groups in clear water condition. *Fuzzy Sets Syst.* **319**, 50–69. doi:10.1016/j.fss.2016.10.010.

Azimi, H., Bonakdari, H. & Ebtehaj, I. 2017b A highly efficient gene expression programming model for predicting the discharge coefficient in a side weir along a trapezoidal canal. *Irrig. Drain.* **66**, 655–666. doi:10.1002/ird.2127.

Babovic, V. 2000 Data mining and knowledge discovery in sediment transport. *Comput.-Aided Civ. Infrastruct. Eng.* **15**, 383–389.

Babovic, V. 2009 Introducing knowledge into learning based on genetic programming. *J. Hydroinform.* **11**, 181–193.

Bajestan, M. S. & Azizi, R. 2012 Experimental Investigation of Scour Depth at the Edge of Different Submerged Vane Shapes. In: *World Environmental and Water Resources Congress 2012: Crossing Boundaries*, pp. 1376–1385.

Bateni, S. M., Jeng, D.-S. & Melville, B. W. 2007 Bayesian neural networks for prediction of equilibrium and time-dependent scour depth around bridge piers. *Adv. Eng. Softw.* **38**, 102–111. doi:10.1016/j.advengsoft.2006.08.004.

Cao, J., Lin, Z. & Huang, G.-B. 2012 Self-adaptive evolutionary extreme learning machine. *Neural Process. Lett.* **36**, 285–305. doi:10.1007/s11063-012-9236-y.

Dibike, Y. B., Velickov, S., Solomatine, D. & Abbott, M. B. 2001 Model induction with support vector machines: introduction and applications. *J. Comput. Civ. Eng.* **15**, 208–216. doi:10.1061/(ASCE)0887-3801(2001)15:3(208).

Ding, S., Zhang, J., Xu, X. & Zhang, Y. 2016 A wavelet extreme learning machine. *Neural Comput. Appl.* **27**, 1033–1040. doi:10.1007/s00521-015-1918-8.

Emiroglu, M. E., Bilhan, O. & Kisi, O. 2011 Neural networks for estimation of discharge capacity of triangular labyrinth side-

weir located on a straight channel. *Expert Syst. Appl.* **38**, 867–874. doi:10.1016/j.eswa.2010.07.058.

Etemad-Shahidi, A., Bonakdar, L. & Jeng, D.-S. 2015 Estimation of scour depth around circular piers: applications of model tree. *J. Hydroinformatics* **17**, 226–238.

Firat, M. & Gungor, M. 2009 Generalized regression neural networks and feed forward neural networks for prediction of scour depth around bridge piers. *Adv. Eng. Softw.* **40**, 731–737. doi:10.1016/j.advengsoft.2008.12.001.

Gaudio, R., Marion, A. & Bovolin, V. 2000 Morphological effects of bed sills in degrading rivers. *J. Hydraul. Res.* **38**, 89–96. doi:10.1080/00221680009498344.

Goel, A. & Pal, M. 2009 Application of support vector machines in scour prediction on grade-control structures. *Eng. Appl. Artif. Intell.* **22**, 216–233. https://doi:10.1016/j.engappai.2008.05.008.

Guan, D., Melville, B. W. & Friedrich, H. 2014 Flow patterns and turbulence structures in a scour hole downstream of a submerged weir. *J. Hydraul. Eng.* **140**, 68–76. doi:10.1061/(ASCE)HY.1943-7900.0000803.

Guan, D., Melville, B. W. & Friedrich, H. 2015 Live-bed scour at submerged weirs. *J. Hydraul. Eng.* **141**, 04014071. doi:10.1061/(ASCE)HY.1943-7900.0000954.

Guan, D., Melville, B. & Friedrich, H. 2016 Local scour at submerged weirs in sand-bed channels. *J. Hydraul. Res.* **54**, 172–184. doi:10.1080/00221686.2015.1132275.

Huang, G.-B., Zhu, Q.-Y. & Siew, C.-K. 2004 Extreme learning machine: a new learning scheme of feedforward neural networks. In: *2004 IEEE Int. Jt. Conf. Neural Networks (IEEE Cat No04CH37541)*. IEEE, pp. 985–990

Lei, Y., Zhao, D. & Cai, H. 2015 Prediction of length-of-day using extreme learning machine. *Geod. Geodyn.* **6**, 151–159. doi:10.1016/j.geog.2014.12.007.

Liu, Q., Yin, J., Leung, V. C. M., Zhai, J. H., Cai, Z. H. & Lin, J. 2016 Applying a new localized generalization error model to design neural networks trained with extreme learning machine. *Neural Comput. Appl.* **27**, 59–66. doi:10.1007/s00521-014-1549-5.

Marelius, F. & Sinha, S. K. 1998 Experimental investigation of flow past submerged vanes. *J. Hydraul. Eng.* **124**, 542–545. doi:10.1061/(ASCE)0733-9429(1998)124:5(542).

Muzzammil, M. 2010 ANFIS approach to the scour depth prediction at a bridge abutment. *J. Hydroinformatics* **12**, 474–485.

Najafzadeh, M., Barani, G.-A. & Kermani, M. R. H. 2013 Abutment scour in clear-water and live-bed conditions by GMDH network. *Water Sci. Technol.* **67**, 1121–1128.

Noori, R., Abdoli, M. A., Ghasrodashti, A. A. & Jalili Ghazizade, M. 2009a Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: a case study of Mashhad. *Environ. Prog. Sustain. Energy* **28**, 249–258. doi:10.1002/ep.10317.

Noori, R., Karbassi, A., Farokhnia, A. & Dehghani, M. 2009b Predicting the longitudinal dispersion coefficient using support vector machine and adaptive neuro-fuzzy inference system techniques. *Environ. Eng. Sci.* **26**, 1503–1510. doi:10.1089/ees.2008.0360.

Odgaard, A. J. & Kennedy, J. F. 1983 River-bend bank protection by submerged vanes. *J. Hydraul. Eng.* **109**, 1161–1173. doi:10.1061/(ASCE)0733-9429(1983)109:8(1161).

Odgaard, A. J. & Spoljaric, A. 1986 Sediment control by submerged vanes. *J. Hydraul. Eng.* **112**, 1164–1180. doi:10.1061/(ASCE)0733-9429(1986)112:12(1164).

Odgaard, A. J. & Wang, Y. 1991 Sediment management with submerged vanes. II: applications. *J. Hydraul. Eng.* **117**, 284–302. doi:10.1061/(ASCE)0733-9429(1991)117:3(284).

Ouyang, H.-T. 2009 Investigation on the dimensions and shape of a submerged vane for sediment management in alluvial channels. *J. Hydraul. Eng.* **135**, 209–217. doi:10.1061/(ASCE)0733-9429(2009)135:3(209).

Price, K., Storn, R. M. & Lampinen, J. A. 2005 *Differential Evolution: A Practical Approach to Global Optimization*. doi:10.1007/3-540-31306-0.

Riad, S., Mania, J., Bouchaou, L. & Najjar, Y. 2004 Predicting catchment flow in a semi-arid region via an artificial neural network technique. *Hydrol. Process* **18**, 2387–2393. doi:10.1002/hyp.1469.

Sharafi, H., Ebtehaj, I., Bonakdari, H. & Zaji, A. H. 2016 Design of a support vector machine with different kernel functions to predict scour depth around bridge piers. *Nat. Hazards* **84**, 2145–2162. doi:10.1007/s11069-016-2540-5.

Tan, S.-K., Yu, G., Lim, S.-Y. & Ong, M.-C. 2005 Flow structure and sediment motion around submerged vanes in open channel. *J. Waterw. Port, Coastal, Ocean Eng.* **131**, 132–136. doi:10.1061/(ASCE)0733-950X(2005)131:3(132).

Vafakhah, M. 2012 Application of artificial neural networks and adaptive neuro-fuzzy inference system models to short-term streamflow forecasting. *Can. J. Civ. Eng.* **39**, 402–414. doi:10.1139/l2012-011.

Vapnik, V. N. 1998 *Statistical Learning Theory*. Wiley.

Wang, L., Melville, B. W., Whittaker, C. N. & Guan, D. 2018a Effects of a downstream submerged weir on local scour at bridge piers. *J. Hydro-Environment. Res.* **20**, 101–109. doi:10.1016/j.jher.2018.06.001.

Wang, L., Melville, B. W., Guan, D. & Whittaker, C. N. 2018b Local scour at downstream sloped submerged weirs. *J. Hydraul. Eng.* **144**, 04018044. doi:10.1061/(ASCE)HY.1943-7900.0001492.

Wang, L., Melville, B. W. & Guan, D. 2018c Effects of upstream weir slope on local scour at submerged weirs. *J. Hydraul. Eng.* **144**, 04018002. doi:10.1061/(ASCE)HY.1943-7900.0001431.

Yu, X., Liong, S. & Babovic, V. 2004 EC-SVM approach for real-time hydrologic forecasting. *J. Hydroinform.* **6**, 209–223.

Zounemat-Kermani, M., Beheshti, A.-A., Ataie-Ashtiani, B. & Sabbagh-Yazdi, S.-R. 2009 Estimation of current-induced scour depth around pile groups using neural network and adaptive neuro-fuzzy inference system. *Appl. Soft. Comput.* **9**, 746–755. doi:10.1016/j.asoc.2008.09.006.

First received 28 March 2019; accepted in revised form 24 July 2019. Available online 13 September 2019