



# Evaluation of discharge coefficient of triangular side orifices by using regularized extreme learning machine

Rahim Gerami Moghadam<sup>1</sup> · Behrouz Yaghoubi<sup>1</sup> · Ahmad Rajabi<sup>1</sup> · Saeid Shabanlou<sup>1</sup> · Mohammad Ali Izadbakhsh<sup>1</sup>

Received: 16 September 2021 / Accepted: 29 March 2022 / Published online: 6 May 2022  
© The Author(s) 2022

## Abstract

The present paper attempts to reproduce the discharge coefficient (DC) of triangular side orifices by a new training approach entitled “Regularized Extreme Learning Machine (RELM).” To this end, all parameters influencing the DC of triangular side orifices are initially detected, and then six models are extended by them. For training the RELMs, about 70% of the laboratory measurements are implemented and the remaining (i.e., 30%) are utilized for testing them. In the next steps, the optimal hidden layer neurons number, the best activation function and the most accurate regularization parameter are chosen for the RELM model. As a result of a sensitivity analysis, we figure out that the most important RELM model simulates coefficient values with high exactness. The best RELM model estimates coefficients of discharge using all input factors. The efficiency of the best RELM model is compared with ELM, and it is demonstrated that the former has a lower error and better correlation with the experimental measurements. The error and uncertainty examinations are executed for the RELM and ELM models to indicate that RELM is noticeably stronger. At the final stage, an equation is proposed for computing this coefficient for triangular side orifices and a partial derivative sensitivity analysis is also carried out on it.

**Keywords** Discharge coefficient · Triangular side orifices · Partial derivative sensitivity analysis · Regularized extreme learning machine · Simulation

## Introduction

In general, side orifices are placed as a slot on main channel sidewalls to control the flow in irrigation systems, water treatment plants and flocculation units. Such orifices are utilized with various shapes such as square, rectangular, elliptical and triangular. The DC is probably the most important parameter for the design of side orifices. This parameter is affected by several variables which are hardly detected due to the existence of three-dimensional flows and high turbulence at the entrance of the side orifice. Therefore, many scientists have tried to understand the hydraulic behavior of such structures and their coefficient of discharge. For instance, the first researcher who evaluated the pattern of spatial variable flows around divert structures was Carballada (1979). Next, Ramamurthy et al.

(1986) and Ramamurthy et al. (1987) invented an analytical technique for the calculation of the flow rate of side orifices. They verified their findings with laboratory measurements. They proved that the proposed analytical model suitably approximated discharge data. Then, Hussein et al. (2010) through an experiment measured the discharge of circular side orifices in different geometric and hydraulic conditions. They conducted 215 experiments to conclude several regression formulae for computing the coefficient. Their proposed relations measured the coefficient via the channel width, the side orifice diameter and the Froude number. Hussein et al. (2011) through an experiment assessed the DC alongside the flow pattern of square side orifices. They executed 173 experiments and claimed that the DC extremely belongs to the Froude number, Reynolds number and geometric parameters. They came to the conclusion that the relations presented for circular side orifices could be also used for square orifices. Hussein et al. (2016) experimentally and analytically investigated the hydraulic behavior of circular side orifices in submerged and free conditions. By analyzing the findings of their

✉ Behrouz Yaghoubi  
byaghoubi@gmail.com

<sup>1</sup> Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran

experiment, they derived several relations to compute the coefficient and explained that the errors of these formulae for submerged and free conditions are 10 and 5%, separately. Moreover, Vatankhah and Mirnia (2018) through an experimental study measured coefficients of discharge of triangular side orifices for various geometric and hydraulic conditions. These authors introduced parameters influencing the DC and suggested some equations for computing the coefficient of discharge. Recently, artificial intelligence (AI) models and learning machines have been successfully implemented for simulating and predicting the DC of divert structures. Such models are accurate, reliable tools for the simulation of the discharge capacity and the determination of the impact level of different variables on the coefficient of discharge. Furthermore, AI techniques are very fast that researchers can save a considerable amount of time and money. For instance, the gene expression programming (GEP) approach has been successfully utilized for reproducing the DC of weirs (Ebtehaj et al. 2015a; Azimi et al. 2017a). Additionally, Khoshbin et al. (2016) and Azimi et al. (2017b) utilized hybrid neuro-fuzzy models to recreate the DC of weirs and side orifices. Via the group method of data handling technique, Ebtehaj et al. (2015b) forecasted the DC of square side orifices. Akhbari et al. (2017) managed to utilize radial base neural networks (RBNNs) for computing the DC of triangular-shaped weirs. In addition, Azimi et al. (2017c) employed the extreme learning machine (ELM) model to identify factors influencing the DC of weirs placed in trapezoidal canals. Moreover, Azimi et al. (2019) managed to utilize the support vector regression (SVM) to reproduce the DC of side weirs in rectangular shape. These authors proposed models for calculating the coefficient of discharge. They also detected factors impacting the discharge capacity of flow divert structures. Bagherifar et al. (2020) modeled the flow field within a circular flume along a rectangular side weir by means of a computational fluid dynamics. The authors showed that the specific energy along the side weir predicted by this model was roughly constant, and the energy drop along the side weir was negligible. This means that the average difference between the upstream and downstream specific energy was estimated as 2.1%.

As discussed, the ELM model has also been used for computing the coefficient of discharge. In general, ELM is one of the best learning machines with great ability to estimate and simulate different problems. This learning machine has acceptable exactness and high learning speed despite its drawbacks (Azimi and Shiri 2021a, 2021b; Azimi et al. 2021). For example, if we intend to gain the best outcomes, the input variables of ELM ought to be dimensionless. Moreover, this model does not include a regularization parameter, which negatively affects the performance of this model during the simulation process, and the phenomenon

of overfitting may occur for the ELM model. To address this restriction, Deng et al. (2009) suggested the regularized extreme learning machine (RELM) model.

Therefore, in this study, as the first time, the new RELM is adopted to recreate the DC of triangular side orifices situated inside rectangular canals. To this end, factors impacting the DC should be initially introduced. Then, the experimental measurements are classified into two categories including train and test. After that, the hidden layer neurons number (HLNN), the activation function and also the regularization parameter for the RELM model are chosen through a trial-and-error process. Then, six RELM models are developed and by executing a comprehensive sensitivity analysis, the best RELM model and the most influencing variable on the DC are identified and introduced. It ought to be noticed that the RELM performance is also contrasted with the ELM. For practical work, an equation is suggested for computing the DC of side weirs.

## Methodology

To train the single-layer feed-forward neural network, the extreme learning machine (ELM) is presented. This method is provided for the first time by Huang et al. (2006). Due to its high learning speed and reasonable generalizability compared to classical learning algorithms such as back-propagation, this algorithm has recently gained many water scientists' attention (Azimi et al. 2017c; Azimi and Shiri 2020a; Azimi and Shiri 2020b; Ebtehaj et al. 2019). Some of the advantages of the ELM method are: (i) low computational cost, (ii) low adjustable parameter, (iii) high generalization performance and scalability, and (iv) proficient approximation of the unknown functions. In modeling by the ELM algorithm, two matrices comprising input weights and the bias of hidden neurons are initialized in a random way; therefore, through the model training process only first matrix is approximated. In fact, a nonlinear problem becomes linear via in this method. Thus, the modeling speed of this approach is very fast. Prior to start modeling using the ELM, it is important to determine the type of the activation function (AF) and the HLNN. There are different AFs such as hyperbolic tangent, triangular basis, radial basis, hard limit, sigmoid, and sine. It is recommended to set the minimum HLNN equal to one and calculate the maximum by the following equation (Azimi and Shiri 2021c):

$$M_{\text{NHN}} \leq \frac{N}{\ln + 2} \quad (1)$$

where  $M_{\text{NHN}}$  represents the maximum HLNN,  $N$  denotes the number of samples considered for training the model and  $\ln$  denotes the number of input variables of the problem. In fact,

using this equation leads to the provision of a model with reasonable generalizability so that the model performance in dealing with data not involving model training is almost similar to the training mode. In fact, considering  $M_{NHN}$  more than the allowable value in the above relation leads to overfitting. In addition to the advantages described for ELM, this method has restrictions, which its weak performance in the presence of outliers is the most important drawback. In fact, due to the lack of heteroscedasticity observed in most real-world problems, the ELM performance in modeling data associated with outliers is significantly reduced. Therefore, in this study, this issue is addressed by providing the regularized ELM (RELM) method. First, the basic principles of ELM are presented, and then mathematical equations related to each developed method are presented.

### Original ELM

Figure 1 depicts the general form of the ELM. It is clear that this model consists of three layers: input, hidden and output. Moreover, the number of neurons existing in the input layer is the same as problem inputs, while this number for neurons of the output layer is equal to problem outputs as shown in the figure with one output. The connection of the hidden layer and the input layer is formed using input weights. Additionally, the output of the hidden layer is also computed by using the input weights and biases of the hidden layer so that by having the value of these two matrices that are determined randomly and using the desired activation function, a matrix named “output matrix of the hidden layer” is calculated. So, having this matrix and samples utilized

for model training, the output weight matrix is calculated through the process of solving a linear system.

Consider a problem with one output and  $N$  samples for training the model  $\{(x_i, y_i)\}_{i=1}^N$  whose inputs are  $x_i \in R^n$  and the corresponding outputs are  $y_i \in R$ . If the HLNN is assumed to be  $L$  and  $g(x)$  is used as the AF, the mathematical form of the ELM method for creating a relationship between input and output parameters is explained as follows (Azimi and Shiri 2021c):

$$\sum_{i=1}^L \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = y_j, j = 1, 2, \dots, N \tag{2}$$

In this equation,  $\mathbf{x}_i$  denotes the input variables of the problem and  $y_i$  represents the output corresponding to inputs. Moreover, the parameter  $b_i$  represents the  $i$ th bias of the hidden layer. It should be stated that all hidden layer biases are randomly initialized. In addition to biases,  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$  which is known as the input weight is also initialized randomly. Also,  $g(\cdot)$  and  $\beta_i$  separately are the AF and the output weight vector. In addition,  $N$  and  $L$  are the numbers of training samples and hidden layer neurons, individually. It is worth mentioning that  $\mathbf{w}_i \cdot \mathbf{x}_j$  denotes the internal multiplication of  $\mathbf{w}_i$  and  $\mathbf{x}_j$ . Now, if Eq. (2) is rewritten as a matrix, the following relation is yielded (Azimi and Shiri 2021c):

$$\mathbf{H}\beta = \mathbf{y} \tag{3}$$

where

$$\mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_L, \mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_1, \dots, b_L, ) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \dots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \dots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L} \tag{4}$$

$$\mathbf{y} = [y_1, \dots, y_N]^T \tag{5}$$

$$\beta = [\beta_1, \dots, \beta_L]^T \tag{6}$$

The matrix presented for defining  $\mathbf{H}$  is known as the output matrix of the hidden layer and depends on the activation function, problem inputs and matrices of the input weight and biases of the hidden layer. Due to the fact that these two matrices are chosen randomly, this matrix is also obtained in the same way. Therefore, the value of this matrix is specified prior to the modeling start. In addition, we have the matrix  $\mathbf{y}$  comprising the corresponding values to input variables for all training samples. So, the only unknown matrix is the output weight ( $\beta$ ) which is computed through the training process. Equation (2) seems simple to solve; however, as the matrix  $\mathbf{H}$  is generally a non-square matrix, solving this relation comes with hardships. The simplest way to yield

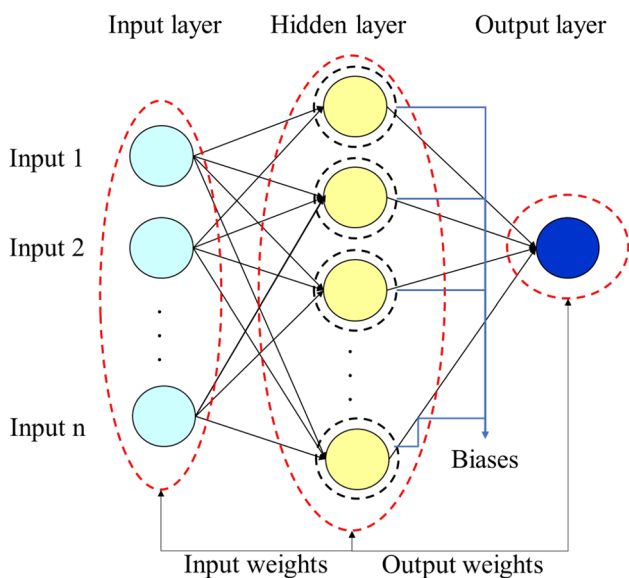


Fig. 1 Structure of ELM

the output weight matrix is to calculate the optimal value of least squares  $\hat{\beta}$  by minimizing the loss function as (Azimi and Shiri 2021c):

$$E_{ELM} = \min \|\mathbf{y} - \mathbf{H}\beta\| \tag{7}$$

The optimal response of the above formula is expressed in the following form:

$$\hat{\beta} = \mathbf{H}^+ \mathbf{y} \tag{8}$$

where  $\mathbf{H}^+$  denotes the Moore–Penrose generalized inverse of the matrix  $\mathbf{H}$  (Rao and Mitra 1971). Given the restrictions set for the maximum HLNN (Eq. 1), the HLNN is generally less than the number of training samples; thus,  $\hat{\beta}$  is computed by the following relationship:

$$\hat{\beta} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} \tag{9}$$

### RELM

Based on the structure defined for ELM, the only matrix determined via the training procedure is the output weight matrix. This matrix is computed in such a way that the training error has the least possible value. However, the minimization of the training error sometimes can decrease the generalizability of the developed model so that the proposed model might be weakened in the estimation of samples not involving in the model training. In the single-layer feed-forward neural network, the highest generalizability occurs once the norm of weights and training error are minimized simultaneously (Bartlett 1997). Thus, in this study, the regularization parameter ( $C$ ) is defined to increase the generalizability of the developed model. Due to the fact that the main difference between the developed method and the original ELM is in the use of regularization parameter, this technique is called regularization ELM (RELM). The regularization parameter is used to modify the loss function provided for ELM to minimize the norm of weights and the training blunder simultaneously. The modified function is expressed as:

$$E_{RELM} = \min_{\beta} C \|\mathbf{y} - \mathbf{H}\beta\|_2^2 + \|\beta\|_2^2 \tag{10}$$

The basic structure of the above equation is similar to Eq. (7), with the difference that the norm of weights (second term) is added to it and the regularization parameter is incorporated into the first term. The relation can be rewritten as follows:

$$E_{RELM} = \min_{\beta} C \|\mathbf{e}\|_2^2 + \|\beta\|_2^2 \text{ subjected to } \mathbf{y} - \mathbf{H}\beta = \mathbf{e} \tag{11}$$

By defining the corresponding Lagrangian as Eq. (12), the optimal response of Eq. (11) is presented as Eq. (13) (Tabak and Kuo 1971):

$$L(\beta, \mathbf{e}, \lambda) = C \|\mathbf{e}\|_2^2 + \|\beta\|_2^2 + \lambda^T (\mathbf{y} - \mathbf{H}\beta - \mathbf{e}) \tag{12}$$

$$\begin{cases} \frac{\partial L}{\partial \beta} = 0 & \Rightarrow 2\beta - \mathbf{H}^T \lambda = 0 \\ \frac{\partial L}{\partial \mathbf{e}} = 0 & \Rightarrow 2C\mathbf{e} - \lambda = 0 \\ \frac{\partial L}{\partial \lambda} = 0 & \Rightarrow \mathbf{y} - \mathbf{H}\beta - \mathbf{e} = 0 \end{cases} \tag{13}$$

where  $\lambda$  is the Lagrangian multiplier matrix,  $\mathbf{e} = [e_1, \dots, e_N]^T$  is the training error, and  $N$  represents the number of training samples. Based on the previous equation, output weights are yielded as

$$\hat{\beta} = \left( \mathbf{H}^T \mathbf{H} + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{H}^T \mathbf{y} \tag{14}$$

where  $\mathbf{I}$  is the unit matrix. It is obvious that neurons exist in the hidden layer are fewer than training samples (condition presented in Eq. (1)). However, if for any reason this condition is not met, the output weight is calculated in the following way:

$$\hat{\beta} = \mathbf{H}^T \left( \mathbf{H}\mathbf{H}^T + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{y} \tag{15}$$

The flowchart of the RELM method is presented in Fig. 2.

### Physical model

In this step, the experimental model utilized in this paper is introduced. In addition, the experimental measurements recorded by Vatankhah and Mirnia (2018) are employed to validate the outcomes of the numerical models. In the current study, the experimental measurements reported by Vatankhah and Mirnia (2018) were used since this dataset was quite new (2018) and comprehensive. Their model is composed of a rectangular-shaped canal equipped by a triangular side weir placed on the sidewall. The main channel is assumed to be 12 m in length, 0.25 m in width, and 0.5 m in height. To measure and control the flow in the rectangular channel, a tailgate is implemented at the end of the channel and a stilling tank is installed to mitigate the flow and avoid turbulence. Moreover, they employed side orifices with different shapes to examine various geometric conditions. All side orifices were made of Plexiglas sheets with a thickness of 0.01 m, while the crest thickness of triangular orifices was about 1 mm. Figure 3 illustrates a channel with side orifices in different shapes.

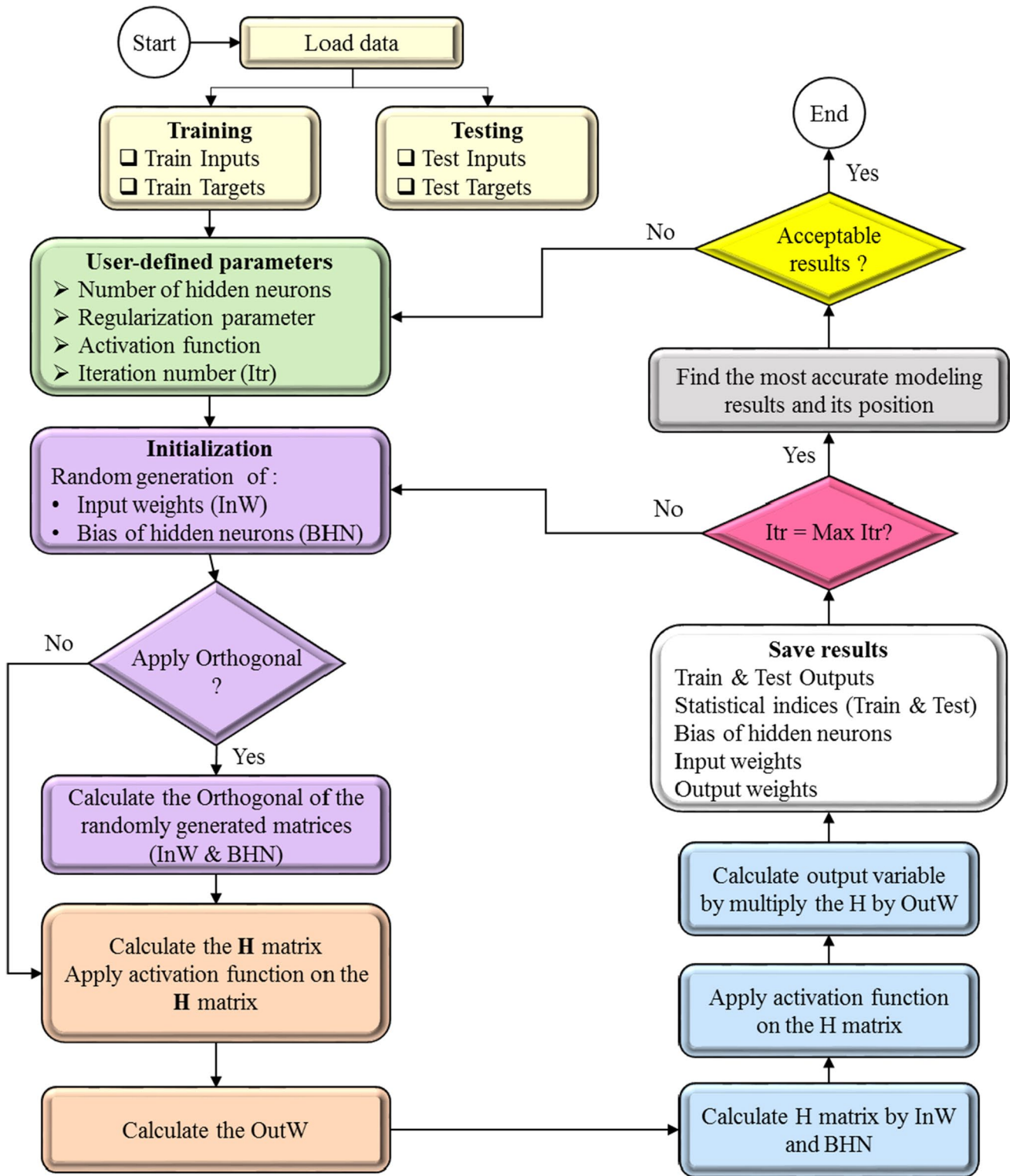


Fig. 2 Flowchart of RELM

### DC of triangular side orifice

Vatankhah and Mirnia (2018) demonstrated that the DC of triangular side orifices ( $C_d$ ) can be interpreted as a function

of the triangular orifice length ( $L$ ), side orifice height ( $H$ ), orifice crest height ( $W$ ), main channel width ( $B$ ), flow depth upstream the triangular side orifice ( $y_1$ ), flow velocity inside the main channel upstream the triangular side orifice ( $V_1$ ),

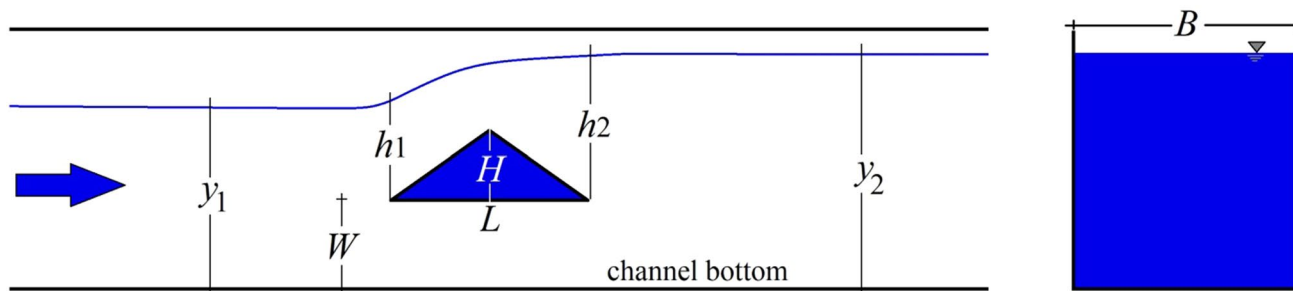


Fig. 3 Layout of experimental model used in study

gravitational acceleration ( $g$ ), water surface tension ( $\sigma$ ), water viscosity ( $\mu$ ), and water density ( $\rho$ ):

$$C_d = f_1(L, H, W, B, y_1, V_1, g, \sigma, \mu, \rho) \tag{16}$$

They proved that Eq. (1) is expressed as a function of the following dimensionless groups:

$$C_d = f_2\left(\frac{W}{H}, \frac{B}{L}, \frac{B}{H}, \frac{y_1}{H}, F_1 = \frac{V_1}{\sqrt{gy_1}}, Re = \frac{\rho V_1 L}{\mu}, We = \frac{\rho V_1^2 L}{\sigma}\right) \tag{17}$$

Vatankhah and Mirnia (2018) asserted that the Reynolds number ( $Re = \frac{\rho V_1 L}{\mu}$ ) and the Weber number ( $We = \frac{\rho V_1^2 L}{\sigma}$ ) have no noticeable impact on the DC of triangular side orifices. Thus, Eq. (17) is expressed as

$$C_d = f_2\left(\frac{W}{H}, \frac{B}{L}, \frac{B}{H}, \frac{y_1}{H}, F_1\right) \tag{18}$$

Therefore, the parameters introduced in the above equation are used for the development of the RELM models. In Fig. 4, the RELM 1 to RELM 6 models alongside the combinations of the input parameters are shown.

**Goodness of fit**

To test the precision of the presented RELM models, different statistical indices such as the correlation coefficient ( $R$ ), variance accounted for (VAF), root-mean-square error (RMSE), BIAS, mean absolute relative error (MARE), and Nash–Sutcliffe efficiency (NSC) are utilized (Azimi and Shiri 2021b).

$$R = \frac{\sum_{i=1}^n (F_i - \bar{F})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (F_i - \bar{F})^2 \sum_{i=1}^n (O_i - \bar{O})^2}} \tag{19}$$

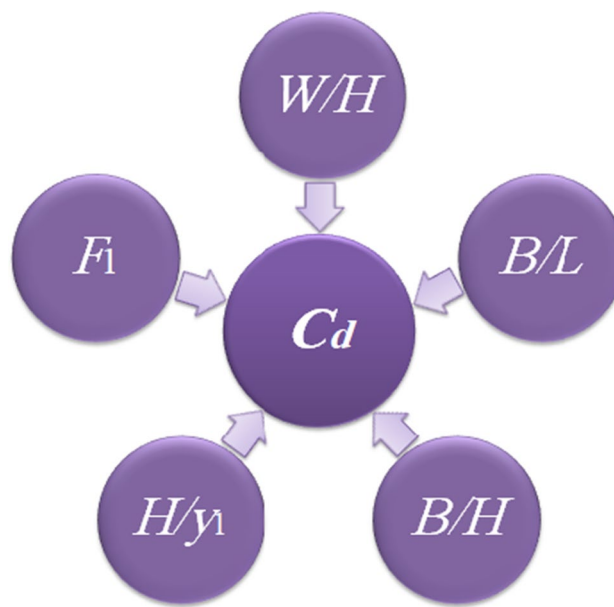


Fig. 4 RELM 1 to RELM 6 models and combinations of different input variables

$$VAF = \left(1 - \frac{\text{var}(F_i - O_i)}{\text{var}(F_i)}\right) \times 100 \tag{20}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_i - O_i)^2} \tag{21}$$

$$BIAS = \frac{1}{n} \sum_{i=1}^n |F_i - O_i| \tag{22}$$

$$MARE = \frac{1}{n} \sum_{i=1}^n \left| \frac{F_i - O_i}{O_i} \right| \tag{23}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - F_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \tag{24}$$

where  $O_i$  denotes the observed data,  $F_i$  represents the data forecasted via numerical models,  $\bar{O}$  represents the average of observed data, and  $n$  denotes the number of observed data. In the current study, the dataset was divided into two major subsamples, the training and testing datasets. In the first round, 50% of the data were utilized for testing and 50% of the rest for testing. In the next round, 60% of the data were applied for training, while 40% of the remaining were considered for testing. This procedure was repeated for other ratios comprising 70% for training vs. 30% for testing, and finally 80% for training against 20% for testing. Analyses showed that the machine learning models had

better performance when 70% of data were employed for the training and 30% of the rest for testing.

## Results and discussion

### Neurons of hidden layer

First, the HLNN is examined. It should be reminded that by increasing the HLNN, the exactness of RELM increases; however, the computational time significantly increases. In Fig. 5, the changes in the HLNN of the RELM model versus the computed statistical indices are depicted. In the current study, one hidden layer neuron is considered for the model in the beginning. For the model with one neuron in its hidden layer,  $R$ , MARE, and BIAS, respectively, are yielded to be 0.953, 0.011, and 0.004. Next, the HLNN increases to 12. As shown, the model with two neurons within the hidden layer demonstrates the lowest accuracy and the poorest

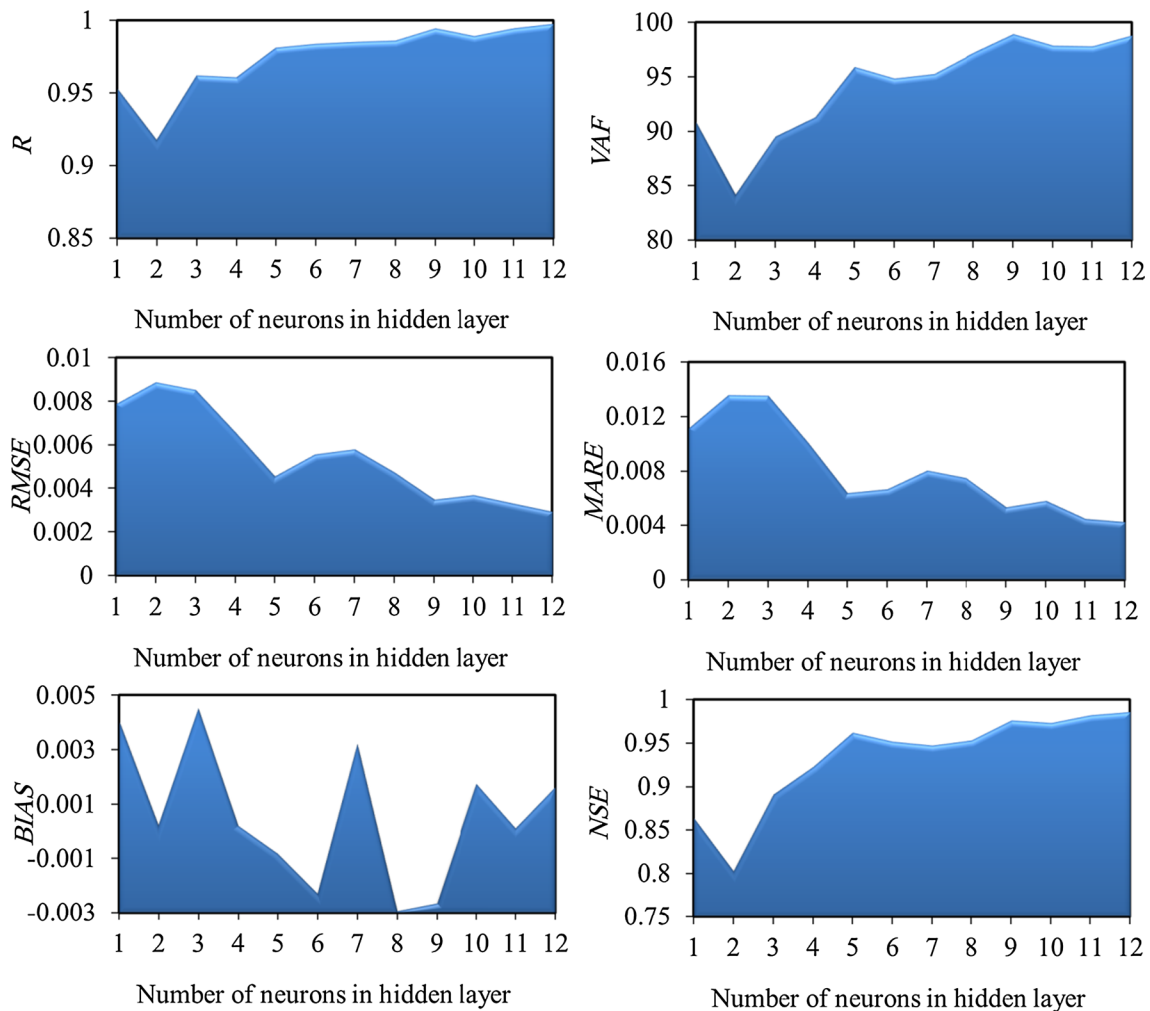


Fig. 5 Changes of HLNN versus different statistical indices

performance. Moreover, for this model, RMSE, BIAS, and NSE are computed to be 0.009, 0.0002, and 0.803, separately. Regarding the results yielded from the sensitivity analysis executed on hidden layer neurons, the outcomes of the models with more than 9 neurons within their hidden layer have no significant changes. Thus, this value is chosen for the HLNN. It implies that for the RELM model, HLNN=9, NSC, VAF, and R, respectively, are estimated to be 0.976, 98.930, and 0.995.

Thus, in the present study, 9 neurons are incorporated within the RELM hidden layer. The scatter plot of RELM with 9 hidden layer neurons is shown in Fig. 6. As displayed, the model displays a good correlation with the observed measurements and predicts coefficients of discharge with the highest exactness and the least error.

### Activation function

The RELM model has six AFs including triangle basis (tribas), tangent hyperbolic (tanh), radial basis (radbas), hard limit (hardlim), sinusoidal (sin), and sigmoid (sig) whose efficiencies are studied in this part. In Fig. 7, the results of such activation functions are depicted. Among all activation functions, hardlim showed the weakest performance. In addition,  $R$ , MARE, and BIAS for this function are equal to 0.589, 0.039, and 0.017, separately. By contrast, sig has the highest accuracy among these functions so that MARE, RMSE, and VAF are computed for this function to be 0.007, 0.005, and 97.178, respectively. In general, the function sig produces better values in the range between zero and one using the curve provided by the function. Thus, this function acts better than step functions.

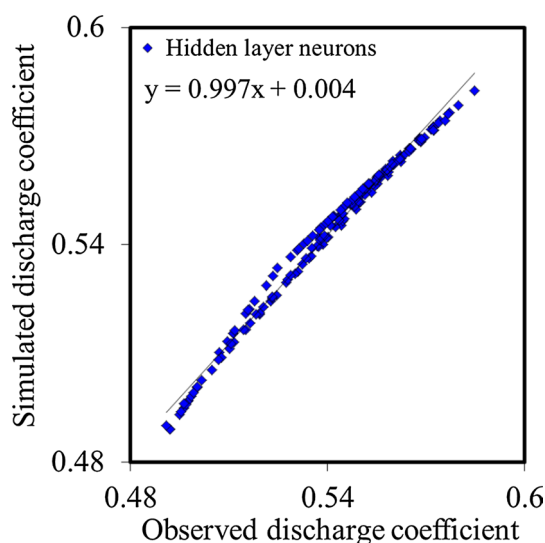


Fig. 6 Scatter plot for RELM with HLNN = 9

As discussed, the function sig owns the highest exactness and correlation with the observed measurements. The results of the coefficients of discharge recreated by sig are shown in Fig. 8. In the next steps, the activation function sig is employed for forecasting the DCs of triangular side orifices by the RELM model.

### Regularization parameter (C)

One of the most important features of RELM is the implementation of the regularization parameter ( $C$ ). The proper selection of the regularization parameter value noticeably enhances the RELM efficiency. For the regularization parameter of the RELM, six different values are chosen. Once  $C=1$ , RELM has lower accuracy so that VAF, NSC, and  $R$ , respectively, are achieved to be 64.937, 0.012, and 0.843. By contrast, in the situation where  $C=0.0001$ , the RELM demonstrates its best performance in terms of accuracy and correlation. It means that for the RELM with  $C=0.0001$ , MARE and  $R$  are yielded 0.005 and 0.995, separately. The statistical indices estimated for regularization parameters are displayed in Fig. 9.

In the developed machine learning model, the optimized value of regularization parameter ( $C$ ) was tuned as 0.0001 because the RELM demonstrated a good performance to predict the target parameter.

As shown, once  $C=0.0001$ , the RELM model displays its best performance. Moreover, for this model, BIAS, VAF, and NSE for this model are yielded to be -0.003, 98.930, and 0.976, individually. The outcomes of the coefficients of discharge for the model with  $C=0.0001$  are shown in Fig. 10.

### RELM models

In this step, the introduced RELM models are examined. The comparison of the performances of the RELM 1 to RELM 6 models is displayed in Fig. 11. It ought to be noticed that the RELM1 model estimates the coefficients of discharge of triangular side orifices using all input parameters. It means that the factors  $W/H, B/L, B/H, H/y_1, F1$  are chosen as the inputs. Compared to the other RELM models, RELM 1 shows the lowest error and the highest correlation. For instance, for RELM1,  $R$ , MARE, and NSE, respectively, are calculated to be 0.995, 0.005, and 0.976. Next, the impact of each input is removed and RELM 2 to RELM 6 are developed. For RELM 2, the impacts of the Froude number are not taken into accounts and this model simulates the DC of triangular side orifices using the input factors such as  $W/H, B/L, B/H, H/y_1$ . It should be considered that RELM2 shows the worst performance among all RELM models. For example, RMSE, MARE, and NSC for RELM2 are yielded to be 0.014, 0.021, and 0.224, separately.



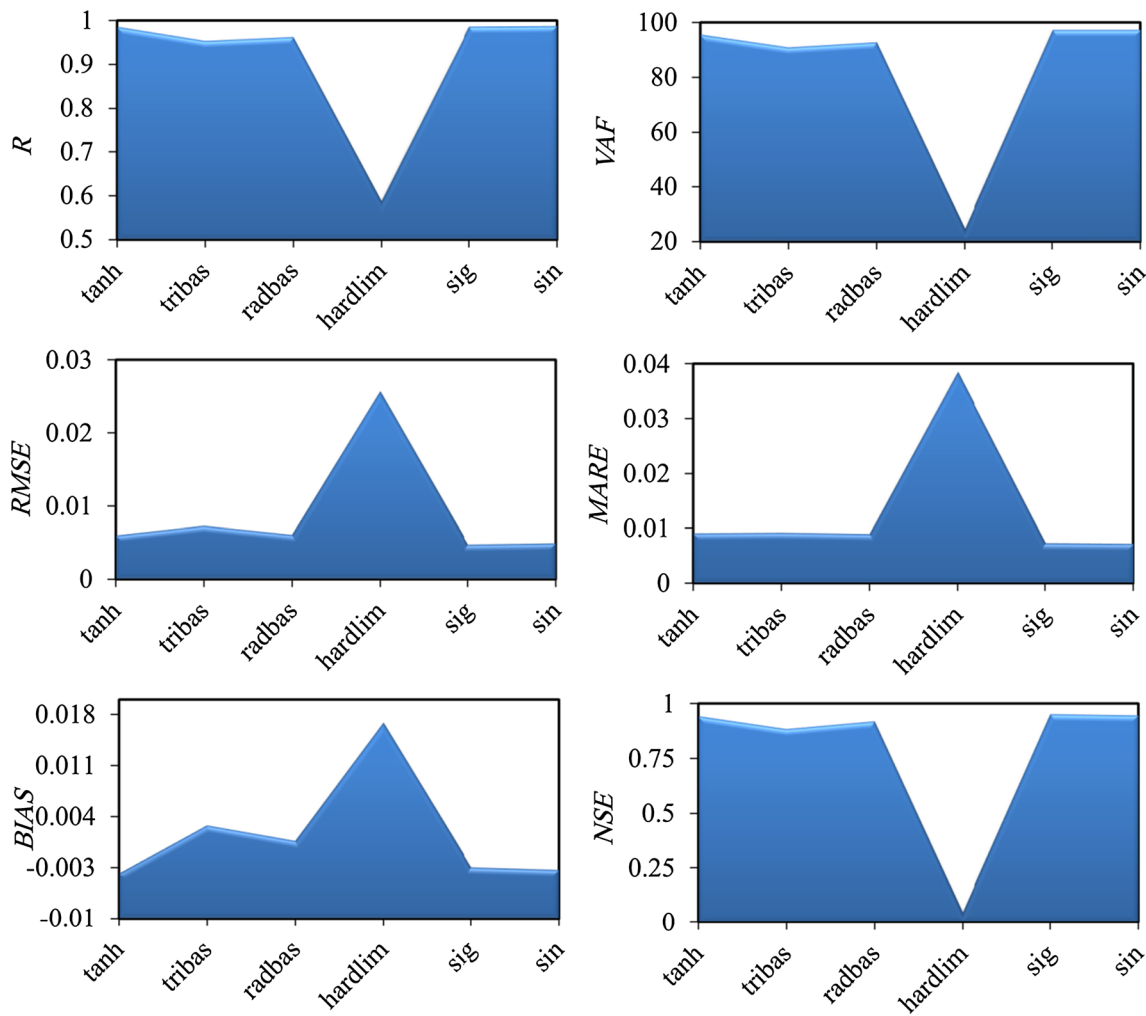


Fig. 7 Different statistical indices for RELM AFs

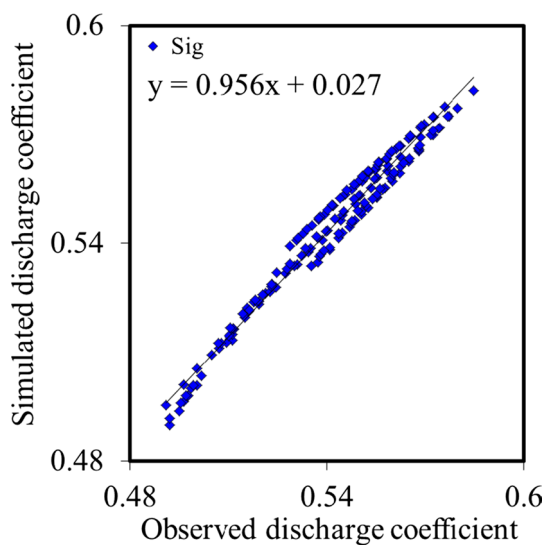


Fig. 8 Scatter plot of activation function sig

For RELM 3, the impact of the parameter  $H/y_1$  is neglected and the model computes the target function in terms of  $W/H, B/L, B/H, F1$ . For RELM3, RMSE, BIAS, and  $R$ , respectively, are equal to 0.005,  $-0.001$ , and 0.980. Moreover, for RELM4, NSC, VAF, and MARE are achieved to be 0.862, 86.317, and 0.017, respectively. It ought to be noticed that RELM4 forecasts the DCs via parameters such as  $W/H, B/L, H/y_1, F1$  and eliminates the impact of the dimensionless parameter  $B/H$ . It ought to be noticed that for RELM5 the impacts of the input factor  $B/L$  are neglected and other variables such as  $W/H, B/H, H/y_1, F1$  are selected as the inputs of this model. Also, NSC, MARE, and BIAS for RELM5 individually are equal to 0.108, 0.067, and 0.037. Furthermore, for RELM6, VAF, RMSE, and  $R$ , respectively, are achieved to be 85.049, 0.010, and 0.923. It ought to be noticed that that RELM6 forecasts the coefficients of discharge by means of the factors  $B/L, B/H, H/y_1, F1$  and eliminates the impact of  $W/H$ .

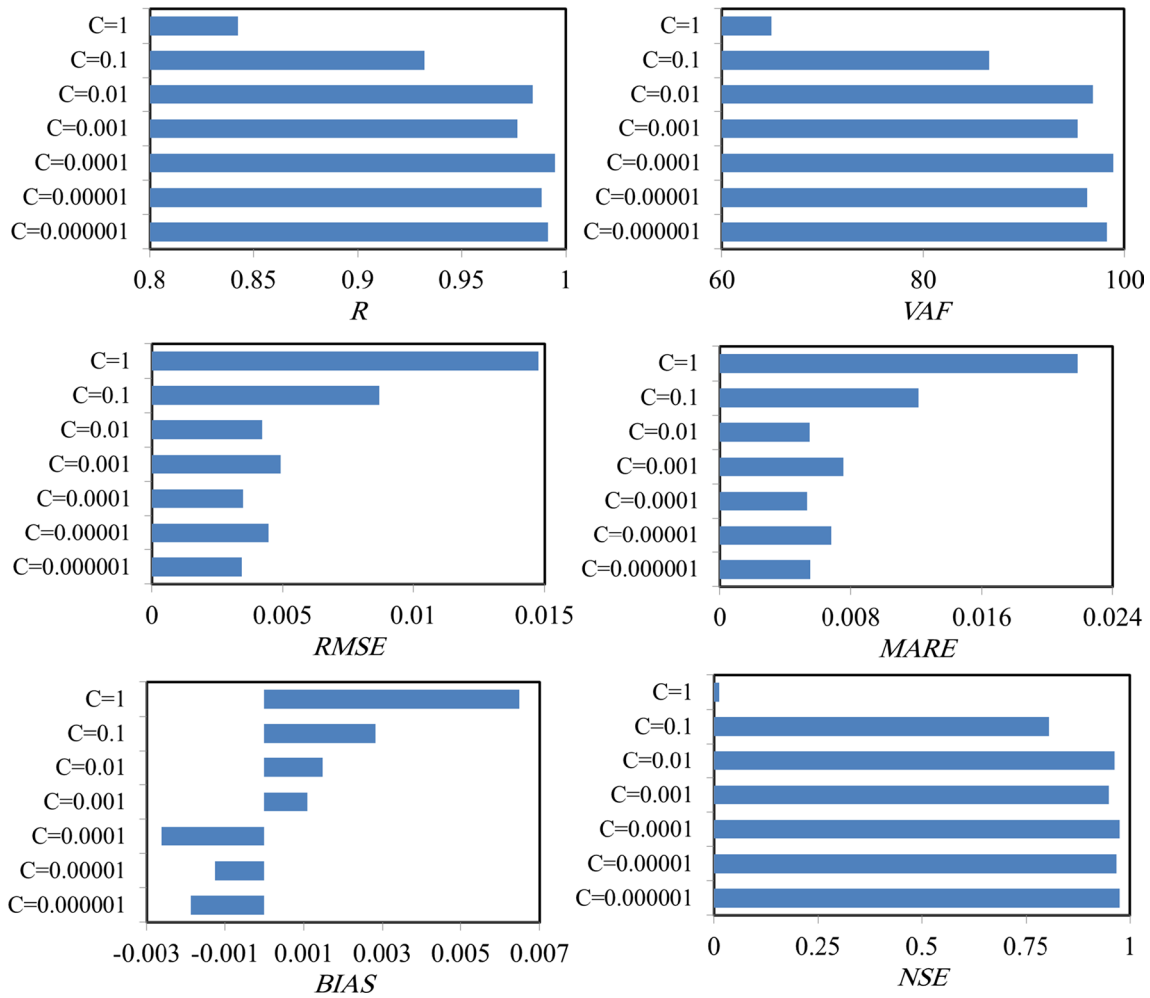


Fig. 9 Different statistical indices estimated for different regularization parameters

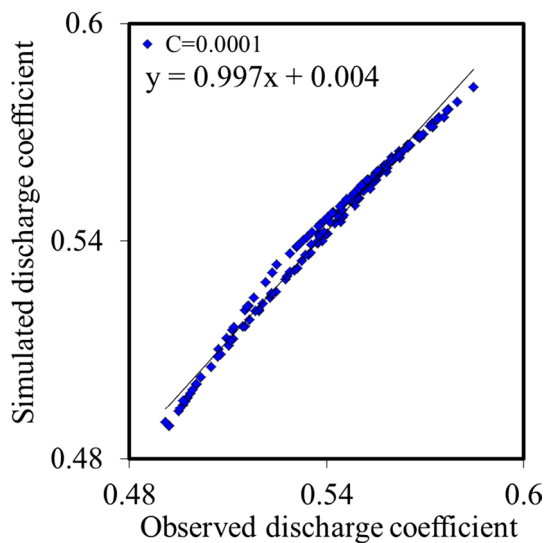


Fig. 10 Scatter plot for regularization parameter equal to 0.0001

The RELM 1 model was the best machine learning model in this study because RELM 1 had the highest level of precision and correlation in comparison with other machine learning models.

As discussed, RELM 1 has the best performance among all RELMs. The comparison of the DCs predicted by RELM1 with the observed data is depicted in Fig. 12. RELM 1 displays the lowest blunder and the highest correlation. After RELM 1, other models such as RELM 3, RELM 4, RELM 5, and RELM 6, respectively, are chosen as the second, third, fourth, fifth, and sixth best models. By contrast, RELM 2 is identified as the sixth or the worst AI model of the present study in estimating the coefficient of discharge. According to the outcomes, the Froude number is considered as the most impacting input variable, because by eliminating this variable, the modeling accuracy is significantly declined. Also, other parameters such as  $W/H$ ,  $B/L$ ,  $B/H$ , and  $H/y_1$  are considered as the second, third, fourth, and fifth influencing variables, separately.

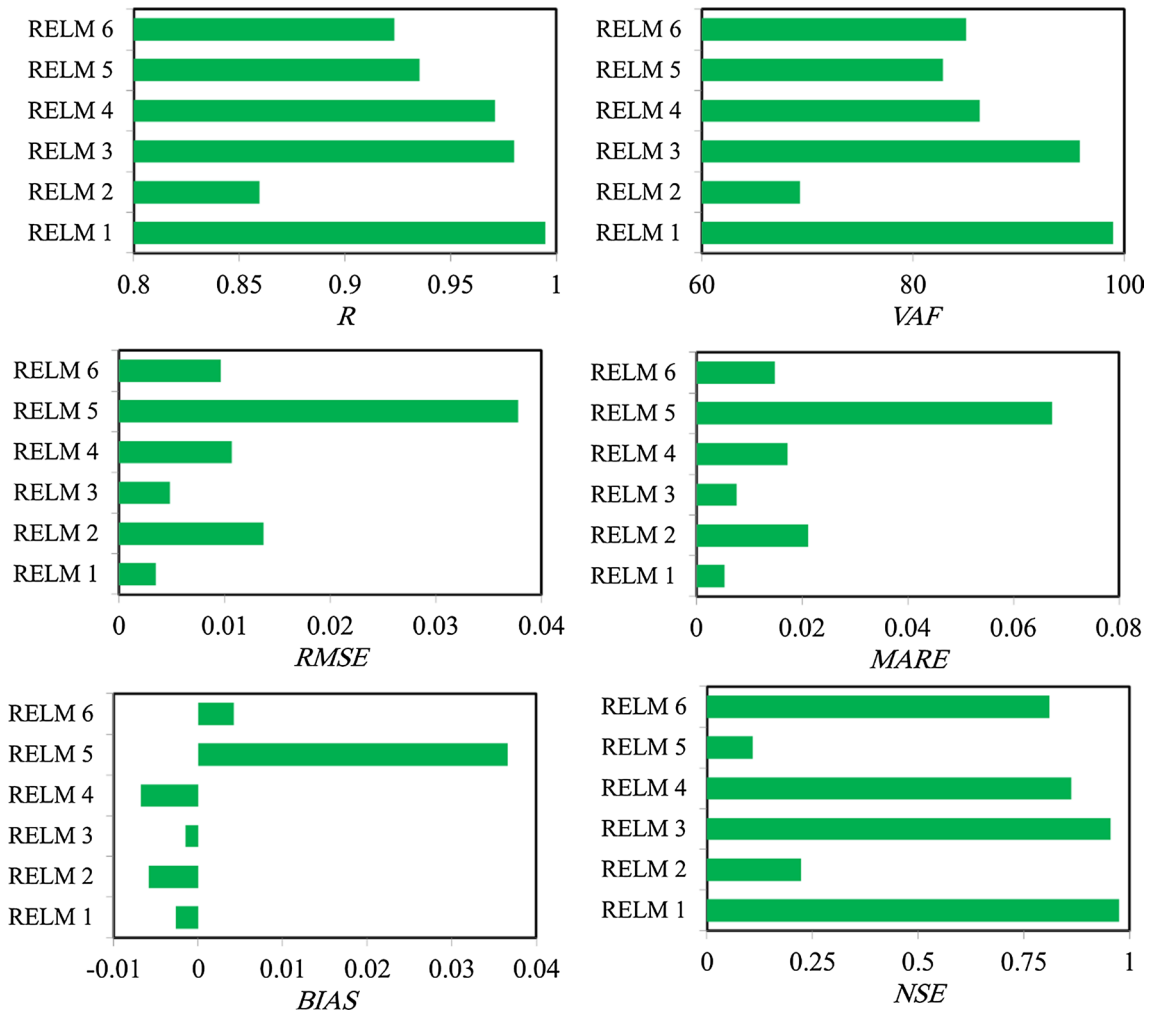


Fig. 11 Values of statistical indices for RELM1 to RELM6

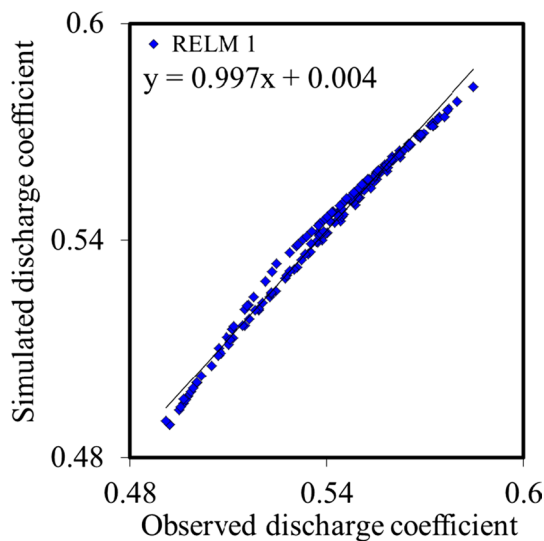


Fig. 12 Scatter plot for RELM1

### Comparison between RELM and ELM

This section attempts to compare the performances of the RELMs and ELMs in estimating the DC of triangular side orifices. In Table 1, the statistical indices computed for these two AI models are listed. According to the simulation results, RELM showed higher efficiency than ELM. As an example, VAF, NSE, and BIAS for ELM separately are obtained to be 89.019, 0.885, and 0.006. However, other indices such as RMSE, *R*, and MARE for ELM are estimated to be 0.010, 0.982, and 0.016, individually.

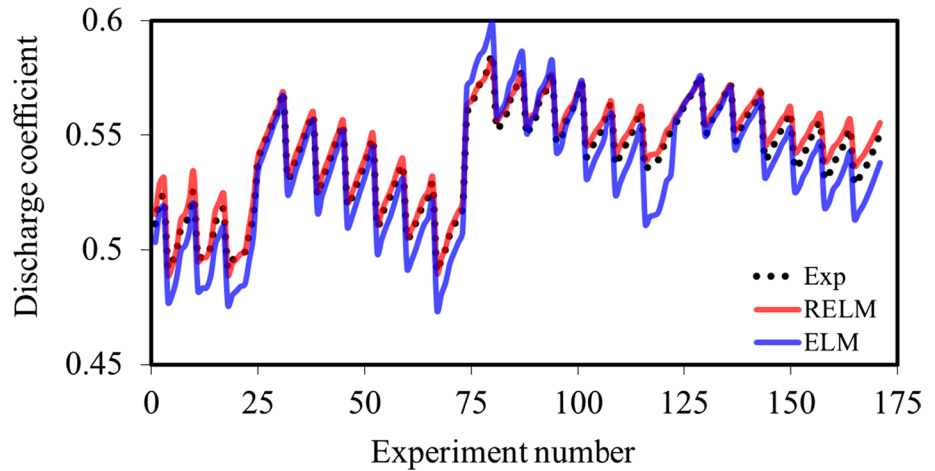
Therefore, the RELM model has higher correlation and less error compared to the ELM model. The comparison of the DCs reproduced by RELM and ELM with the experimental measurements is illustrated in Fig. 13. In the following, further examination is done on the performances of the RELMs and ELMs.

Now, an error examination is performed for the RELMs and ELMs. Figure 14 depicts the error analysis results for

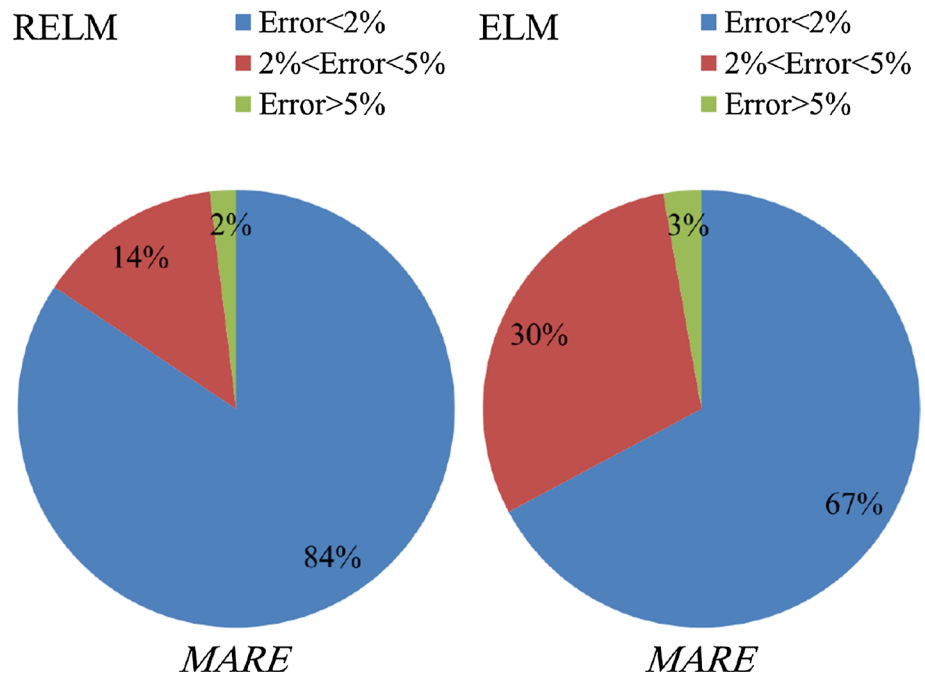
**Table 1** Comparison of RELMs and ELMs

Model	<i>R</i>	VAF	RMSE	MARE	BIAS	NSE
RELM	0.995	98.930	0.003	0.005	-0.003	0.976
ELM	0.982	89.019	0.010	0.016	0.006	0.885

**Fig. 13** Comparison of coefficients of discharge simulated by RELM and ELM



**Fig. 14** Error analysis outcomes for RELMs and ELMs



these two models. Based on the corresponding results, the vast majority of the outcomes achieved by the RELM model have an error of less than 2% (i.e., 84% of the RELM results), while this percentage for ELM is equal to 67%. Moreover, about a third of the ELM model outcomes display an error between 2 and 5%; however, this percentage is calculated to be approximately 14% for the RELM model. Approximately 2 and 3% of the coefficients of discharge

recreated using the RELM and ELM models have an error of more than 5%, respectively.

The uncertainty analysis (UA) is one of the best tools for specifying the efficiency of numerical models. Using this analysis, error values can be interpreted properly (Azimi and Shiri 2020a). Table 2 lists the UA results for RELM and ELM. According to Table 2, the RELM model demonstrates an underestimated performance for the simulation of the DCs of triangular side orifices, while the ELM

**Table 2** Uncertainty analysis results for RELMs and ELMs

Model	Mean prediction error	Standard deviation of prediction error	95% prediction error interval	Uncertainty bandwidth
ELM	0.0063	0.0074	(0.0052, 0.0074)	±0.0011
RELM	-0.0026	0.0023	(-0.00297, 0.00227)	±0.00035

demonstrates an overestimated performance for predicting the coefficients. It should be noted that the uncertainty bandwidth of the RELM model is much narrower than the uncertainty band of the ELM model (approximately three times narrower). In other words, the uncertainty band values for the ELMs and RELMs, respectively, are yielded to be ±0.0011 and ±0.00035.

**The superior RELM model**

According to the findings of the present research, RELM1 is specified as the most successful model for simulating the DC of triangular side orifices. It ought to be noticed that RELM1 reproduces the DCs in terms of all input variables. The general expression of the formula obtained from the RELM is as follows:

$$C_d = \left[ \frac{1}{(1 + \exp(\text{InW} \times \text{InV} + \text{BHN}))} \right]^T \times \text{OutW} \quad (25)$$

where

- InW = matrices of input weights.
- InV = input variables.
- BHN = bias of hidden neurons.
- OutW = matrices of output weights.
- The matrices of RELM1 are presented as

$$\text{InV} = \begin{bmatrix} W/H \\ B/L \\ B/H \\ H/y_1 \\ F_1 \end{bmatrix} \quad (26)$$

$$\text{BHN} = \begin{bmatrix} -0.3548 \\ 0.3036 \\ 0.1627 \\ -0.2202 \\ -0.4865 \\ -0.4437 \\ 0.2869 \\ 0.0362 \\ -0.4354 \end{bmatrix} \quad \text{InW} = \begin{bmatrix} 0.0918 & -0.5653 & 0.2762 & -0.2999 & -0.0988 \\ 0.1848 & 0.2208 & 0.4795 & -0.0270 & 0.3259 \\ -0.1488 & -0.0752 & 0.3683 & -0.4585 & 0.0602 \\ 0.3847 & 0.1675 & -0.3718 & 0.2252 & 0.2741 \\ -0.1717 & 0.6127 & -0.0885 & -0.4988 & -0.2484 \\ 0.5988 & 0.3706 & 0.3086 & -0.0805 & -0.1549 \\ 0.0367 & 0.0290 & 0.4558 & 0.3487 & 0.3001 \\ -0.0582 & 0.0500 & 0.2722 & 0.4665 & -0.7552 \\ 0.6280 & -0.2858 & -0.1863 & -0.2316 & -0.2429 \end{bmatrix} \quad \text{OutW} = \begin{bmatrix} 3.8312 \\ 14.3517 \\ -8.0767 \\ -2.7308 \\ 1.5018 \\ -3.3035 \\ -6.8588 \\ 0.6647 \\ 0.1774 \end{bmatrix} \quad (27)$$

In this step, a partial derivative sensitivity analysis (PDSA) is done for the best model. The PSDA is a useful tool for evaluating the effectiveness of each input variable of the target function value, which in this research is the DC side orifices with triangular shape (Azimi and Shiri 2020b). Moreover, the positive indication of the PSDA implies increasing impacts of the input parameters on the coefficient of discharge, while the negative sign uncovers that the impact is diminishing. In other words, the relative derivative of each input variable is computed relative to the target function (Azimi et al. 2017a). The PSDA findings for the best model (i.e., RELM 1) are shown in Fig. 15. As discussed, RELM 1 predicts the coefficients of discharge via the inputs  $W/H, B/L, B/H, H/y_1, F_1$ . Based on the findings, all the PSDA values for the parameters  $W/H, B/L$ , and  $F_1$  are negative. Moreover, for the other input parameters such as  $B/H, H/y_1$  part of the PSDA results are positive and the remaining are yielded negatively.

**Conclusion**

Basically, the DC is the most effective variable for the design of a side orifice. In the present paper, as the first time, a new training approach entitled “Regularized Extreme Learning Machine (RELM)” was utilized for reproducing the DC of triangular-shaped side orifices. To do this, six different RELM models including RELM 1 to RELM 6 were developed using the variables impacting triangular side orifices. After that, the laboratory measurements were classified into two categories including training (70%) and testing (30%). The most highlighted outcomes of this paper are as follows:

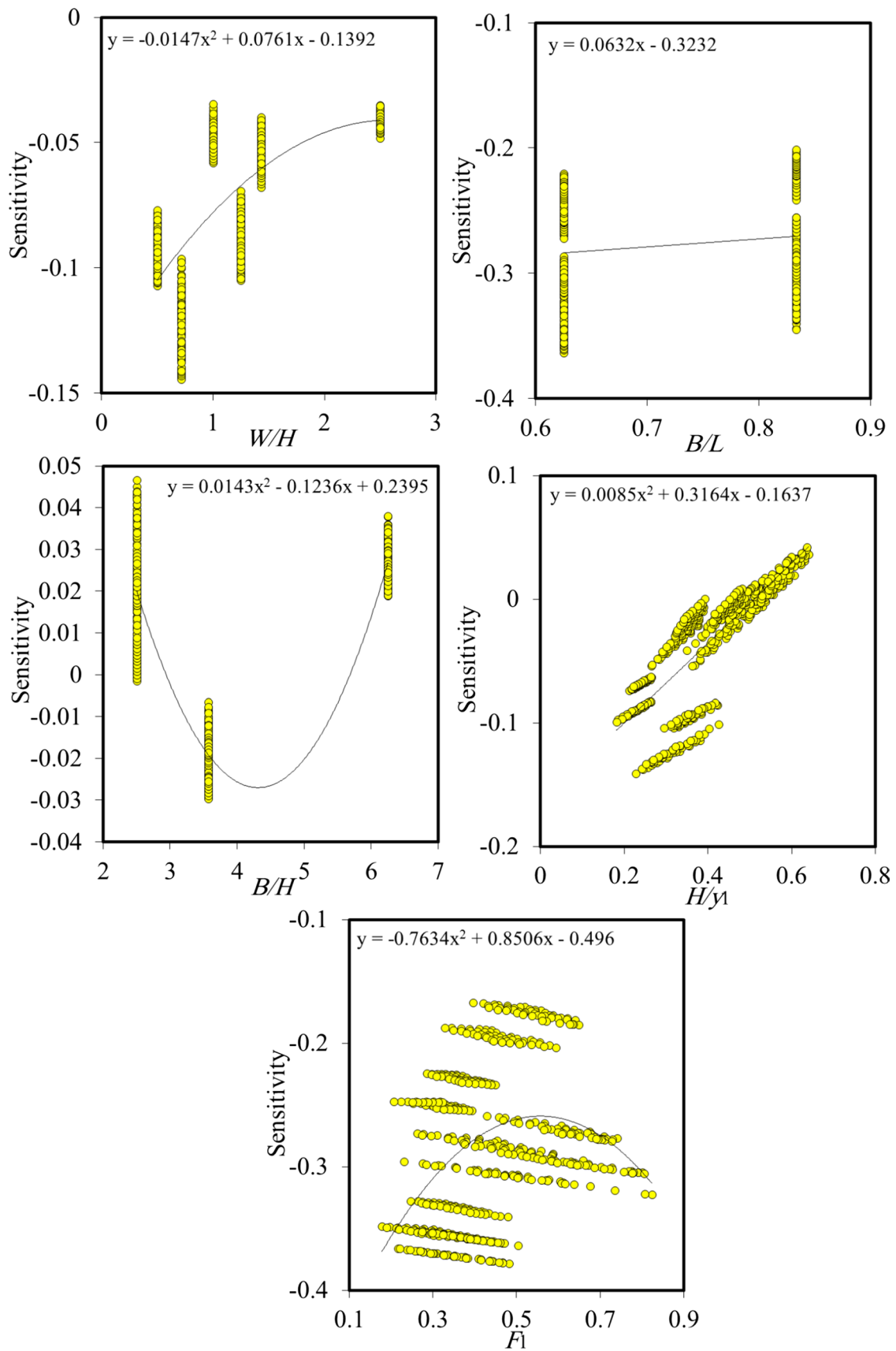


Fig. 15 PDSA results for RELM1

- The most optimal HLNN was chosen to be 9 through a trial-and-error procedure. RMSE, MARE, and VAF for the RELM HLNN=9 hidden layer were obtained to be 0.003, 0.005, and 98.930, individually.
- The function sigmoid was taken into account as the best activation function of the RELM model for recreating the coefficients of discharge. For this activation function,  $R$  and NSE, respectively, were approximated to be 0.986 and 0.953.
- Seven different values were considered for the regularization parameter ( $C$ ). It was observed that when  $C = 0.0001$ , RELM demonstrated its best performance. In addition, the estimations of MARE and BIAS for the regularization parameter were yielded to be 0.005 and -0.003, separately.
- Among all RELM models, RELM1 as the best one predicted the DCs of triangular side orifices via the inputs  $W/H, B/L, B/H, H/y_1, F1$ . This model displayed an acceptable performance, and  $R$  and RMSE for it were computed to be 0.995 and 0.003, individually.
- The RELM findings were contrasted with the ELM to emphasize the superiority of the former. For instance, about 86% of the DC recreated by RELM had an error of less than 2%. However, this percentage for ELM was about 67%.
- The RELM model estimated the coefficients of discharge with an underestimated performance, while the ELM performance was assessed overestimated.
- A RELM-based formula was ultimately put forward for the estimation of the DC of triangular side orifices. Furthermore, a partial derivative sensitivity analysis was conducted on the proposed formula.

For the future works, other machine learning methods and optimization algorithms can be used to simulate the discharge coefficient of side weirs.

**Funding** All authors declare they have no funding.

## Declarations

**Conflict of interest** All authors declare they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Akhbari A, Zaji AH, Azimi H, Vafaeifard M (2017) Predicting the discharge coefficient of triangular plan form weirs using radian basis function and M5' methods. *Appl Res Water Wastewater* 4(1):281–289
- Azimi H, Shiri H (2020a) Dimensionless groups of parameters governing the ice-seabed interaction process. *J Offshore Mech Arct Eng* 142(5):051601–051613
- Azimi H, Shiri H (2020b) Ice-seabed interaction analysis in sand using a gene expression programming-based approach. *App Ocean Res* 98:102120–102212
- Azimi H, Shiri H (2021a) Sensitivity analysis of parameters influencing the ice-seabed interaction in sand by using extreme learning machine. *Nat Hazards* 106(3):2307–2335
- Azimi H, Shiri H (2021) Evaluation of ice-seabed interaction mechanism in sand by using self-adaptive evolutionary extreme learning machine. *Ocean Eng* 239:109795
- Azimi H, Shiri H (2021c) Assessment of ice-seabed interaction process in clay using extreme learning machine. *Int J Offshore Polar Eng* 31(04):411–420
- Azimi H, Bonakdari H, Ebtehaj I (2017a) A highly efficient gene expression programming model for predicting the discharge coefficient in a side weir along a trapezoidal canal. *Irrig Drain* 66(4):655–666
- Azimi H, Shabanlou S, Ebtehaj I, Bonakdari H, Kardar S (2017b) Combination of computational fluid dynamics, adaptive neuro-fuzzy inference system, and genetic algorithm for predicting discharge coefficient of rectangular side orifices. *Irrig Drain Eng* 143(7):04017015
- Azimi H, Bonakdari H, Ebtehaj I (2017c) Sensitivity analysis of the factors affecting the discharge capacity of side weirs in trapezoidal channels using extreme learning machines. *Flow Meas Instrum* 54:216–223
- Azimi H, Bonakdari H, Ebtehaj I (2019) Design of radial basis function-based support vector regression in predicting the discharge coefficient of a side weir in a trapezoidal channel. *App Water Sci* 9(4):78
- Azimi H, Shiri H, Malta ER (2021) A non-tuned machine learning method to simulate ice-seabed interaction process in clay. *J Pipeline Sci Eng* 1(3):1–18
- Bagherifar M, Emdadi A, Azimi H, Sanahmadi B, Shabanlou S (2020) Numerical evaluation of turbulent flow in a circular conduit along a side weir. *App Water Sci* 10(1):1–9
- Bartlett PL (1997) For valid generalization the size of the weights is more important than the size of the network. In: *Advances in Neural Information Processing Systems*, pp 134–140
- Carballada BL (1979) Some characteristics of lateral flows. *Concordia Univ, Montreal*
- Deng W, Zheng Q, Chen L (2009) Regularized extreme learning machine. In: *2009 IEEE Symposium on Computational Intelligence and Data Mining*. pp 389–395.
- Ebtehaj I, Bonakdari H, Zaji AH, Azimi H, Sharifi A (2015a) Gene expression programming to predict the discharge coefficient in rectangular side weirs. *Appl Soft Comput* 35:618–628
- Ebtehaj I, Bonakdari H, Khoshbin F, Azimi H (2015b) Pareto genetic design of group method of data handling type neural network for prediction discharge coefficient in rectangular side orifices. *Flow Meas Instrum* 41:67–74

- Ebtehaj I, Bonakdari H, Gharabaghi B (2019) Closure to “An integrated framework of extreme learning machines for predicting scour at pile groups in clear water condition” by: Ebtehaj I, Bonakdari H, Moradi F, Gharabaghi B Sheikh Khozani z. *Coastal Eng* 147:135–137
- Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. *Neurocomputing* 70(1–3):489–501
- Hussain A, Ahmad Z, Ojha CSP (2016) Flow through lateral circular orifice under free and submerged flow conditions. *Flow Meas Instrum* 52:57–66
- Hussein A, Ahmad Z, Asawa GL (2010) Discharge characteristics of sharp-crested circular side orifices in open channels. *Flow Meas Instrum* 21(3):418–424
- Hussein A, Ahmad Z, Asawa GL (2011) Flow through sharp-crested rectangular side orifices under free flow condition in open channels. *Agric Water Manag* 98:1536–1544
- Khoshbin F, Bonakdari H, Ashraf Talesh SH, Ebtehaj I, Zaji AH, Azimi H (2016) Adaptive neuro-fuzzy inference system multi-objective optimization using the genetic algorithm/singular value decomposition method for modelling the discharge coefficient in rectangular sharp-crested side weirs. *Eng Optim* 48(6):933–948
- Ramamurthy AS, Udoyara ST, Serraf S (1986) Rectangular lateral orifices in open channel. *J Environ Eng* 135(5):292–298
- Ramamurthy AS, Udoyara ST, Rao MVJ (1987) Weir orifice units for uniform flow distribution. *J Environ Eng* 113(1):155–166
- Rao CR, Mitra SK (1971) Generalized inverse of matrices and its applications. John Wiley & Sons. Inc., New York
- Tabak D, Kuo BC (1971) Optimal control by mathematical programming. Prentice-Hall, Englewood Cliffs
- Vatankhah AR, Mirnia SH (2018) Predicting discharge coefficient of triangular side orifice under free flow conditions. *J Irrig Drain* 144(10):04018030

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.