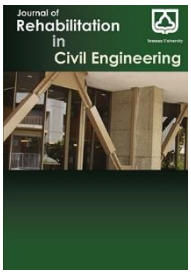
Journal of Rehabilitation in Civil Engineering



Journal homepage: http://civiljournal.semnan.ac.ir/

A laboratory investigation on the potential of computational intelligence approaches to estimate the discharge coefficient of piano key weir

E. Olyaie[1], M. Heydari[1*], H. Banejad[2], Kwok-Wing Chau[3]

1. Department of Water Engineering, College of Agriculture, Bu-Ali Sina University, Hamedan, Iran
2. Department of Water Engineering, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran
3. Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China
Corresponding author: *mheydari_ir@yahoo.com*

ARTICLE INFO

Article history:
Received: 03 December 2017
Accepted: 05 February 2018

Keywords:
Discharge coefficient,
ELM,
FFNN,
Piano key weir,
GEP,
LS.

ABSTRACT

The piano key weir (PKW) is a type of nonlinear control structure that can be used to increase unit discharge over linear overflow weir geometries, particularly when the weir footprint area is restricted To predict the outflow passing over a piano key weir, the discharge coefficient in the general equation of weir needs to be known. This paper presents the results of laboratory model testing of a piano key weir located on the straight open channel flume in the hydraulic laboratory of Bu-Ali Sina University. The discharge coefficient of piano key weir is estimated by using four computational intelligence approaches, namely, feed forward back-propagation neural network (FFBPN), an extension of genetic programming namely gene-expression programming (GEP), least square support vector machine (LSSVM) and extreme learning machine (ELM). For this purpose, 70 laboratory test results were used for determining discharge coefficient of piano key weir for a wide range of discharge values. Coefficient of determination (R^2), Nash-Sutcliffe efficiency coefficient (NS), root mean square error (RMSE), mean absolute relative error (MARE), scatter index (SI) and BIAS are used for measuring the models' performance. Overall performance of the models shows that, all the studied models are able to estimate discharge coefficient of piano key weir satisfactorily. Comparison of results showed that the ELM (R^2=0.997 and NS= 0.986) and LSSVM (RMSE=0.016 and MARE=0.027) models were able to produce better results than the other models investigated and could be employed successfully in modeling discharge coefficient from the available experimental data.

1. Introduction

Today's demand for more water reservoirs, magnitudes of probable maximum flood (PMF) events and the continuing need to improve dam safety has increased. So, the capacities of many existing spillways are currently inadequate and in need of upgrade or replacement. Typically, reservoir spillways use weirs as the flow control structure. In general, there are three techniques for raising the maximum weir discharge capacity of an uncontrolled weir spillway when limited by a maximum H or upstream pool elevation: (1) increase the width of the spillway (W), (2) lower the spillway crest elevation, or (3) increase L within the existing spillway footprint by replacing the existing linear weir with a nonlinear (labyrinth-type) weir. Increasing L of a linear weir, and consequently W, is often impractical due the dam geometry or economic constraints. Lowering the spillway crest elevation reduces base-flow reservoir storage volumes, reducing the amount of stored water available for the specific function of the reservoir (e.g., municipal, agriculture, commercial, and hydropower). The use of nonlinear weirs (labyrinth), however, where L > W, represents a viable and increasingly popular option for increasing discharge capacity while facilitating the sustainability of the existing spillway infrastructure (e.g., spillway channel). [1]. Labyrinth weir due to more effective length than conventional weirs, allows passing more discharge in narrow canals [2].

The piano key weir (PKW) is a modified labyrinth-type nonlinear weir designed specifically for applications with a limited footprint area. Due to increased flood discharges and strict specifications regarding the dam safety, many existing spillways require replacement or increased storage by optimizing their shape. So, there is a significant improvement in the use of piano key weir over other types of the weir.

Fig. 1 shows a longitudinal section and plan of a piano key weir for subcritical flow. As shown in Fig. 1, the PK weir has a rectangular nonlinear weir crest layout (in planform), with sloped floors in the inlet and outlet cycles referred to as keys. In the innovative developed PKWs, the sloped floors produce cantilevered apex overhangs, which help to increase the overall crest length (L) relative to a rectangular labyrinth weir with the same weir footprint [1,3]. An accurate estimation of the discharge coefficient (C_d) of weirs is a significant factor in designing various hydraulic structures.

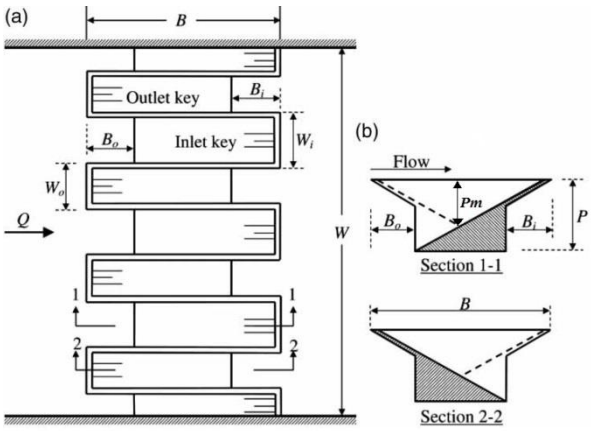


Fig. 1. PK weir parameters

Some useful empirical discharge equations for these weirs have been proposed. PKW may consist of various cross-sections depending on the flow requirements. There is a unique relationship between the unit discharge (the flow rate per unit width) and the upstream water depth measured relative to the weir crest, which is exploited for the purpose of flow measurement.

Recent works mainly focused on hydraulic behavior, flow conditions, and the discharge coefficient for different types of weirs. Kabiri-Samani and Javaheri carried out a set of laboratory experiments to investigate the effect of geometry on the discharge coefficient [4]. Head-discharge data and visual observations were collected for PKW by Anderson over a wide range of discharges and compare the appropriateness the recommended equations [1]. Ribeiro et al. (2012) as a technical note, reviewed the previous studies on the efficiency of planned and built PKW and the results were evaluated by comparing an actual PK-Weir's discharge to that theoretically obtained for a sharp-crested spillway with crest length equal to the width of the PK-Weir for a given hydraulic head.

Computational intelligence, a powerful tool in modeling and solving complex nonlinear problems in different fields such as hydraulic engineering [6-7, 28] and river engineering [8] has come to the attention of a number of researchers in recent years. By using intelligence techniques to predict the discharge coefficient of weirs, Bilhan et al. (2010) performed two different artificial neural network (ANN) techniques to lateral outflow over rectangular-side weirs located on a straight channel. They reported that the ANN technique could be employed successfully for the determination of discharge coefficient. According to their work, the feed-forward neural network (FFNN) model is found to be better than the radial basis neural network. Some of the other applied intelligence methods in predicting discharge coefficient of weirs are: prediction of discharge coefficient for trapezoidal labyrinth side weir using a neuro-fuzzy approach by Emiroglu and Kisi (2013); and also, Estimating Bed load sediment transport in a clean pipe by MLP network optimized with three different training algorithms, including variable learning rate (MLP-GDX), resilient back-propagation (MLP-RP) and Levenberg-Marquardt (MLP-LM) by Ebtehaj and Bonakdari (2016), Group Method of Data Handling (GMDH) was used for the purpose of predicting the discharge coefficient in a side weir by Ebtehaj et al. (2015), discharge coefficient of compound broad-crested weir by using genetic programming (GP) and artificial neural network (ANN) techniques by Salmasi et al. (2013); Gene expression programming (GEP) to predict the discharge coefficient in rectangular side weirs by Ebtehaj et al.(2015); 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 by Khoshbin et al, 2015, radial basis neural network and particle swarm optimization-based equations for predicting the discharge capacity of triangular labyrinth weirs by Zaji et al. (2015); prediction of lateral outflow over triangular labyrinth side weirs using two soft computing approaches , that is, the radial basis neural network and GEP by Kisi et al. (2012); Estimating discharge coefficient of semi-elliptical side weir using ANFIS by Dursun et al. (2012). Ebtehaj and Bonakdari (2017) used GEP model which properly predicted sediment transport in sewer. Also, Azimi et al, 2017 used a highly efficient GEP model for predicting the discharge coefficient in a side weir along a trapezoidal canal.

Recently, support vector machines (SVMs) are a popular learning technique for classification, regression, and other learning task. Unlike traditional artificial neural network technique, the quadratic

programming (QP) with linear limitations is formulated in SVM problems [19- 21]. Also, simplifying the optimization processes of SVMs can be performed through a modification version of SVM namely Least Square Support Vector Machine (LSSVM) [22, 23]. Researchers successfully applied LSSVM for solving different problems in engineering [4, 16, 24]. Also, in recent years, there has been increasing interest in extreme learning machine (ELM), which is an extraordinary learning approaches for single hidden layer feed forward neural network (SHLFN) [25, 26]. In ELM, the weights and biases of input layers are assigned randomly and the output weights are determined analytically. ELM produces high generalization performance at very high speed [27]. There are different applications of ELM in the literature [11, 18, 30].

To the best knowledge of the authors, there is not any published study indicating the input–output mapping capability of computational intelligence techniques in modeling the discharge coefficient of PKW. In addition, there is no report presenting the using of LSSVM and ELM in modeling discharge coefficient of the any type of weirs. Thus, the present study is focused on the construction of different computational intelligence approaches, such as FFBPN, GEP, LSSVM and ELM, to predict the discharge coefficient using a measured data set. The obtained results are finally compared to each other. For this purpose, we put forward a black-box FFBPN structure as reference models for C_d prediction. Then we applied GEP, LSSVM and ELM to model our reference scenarios. These methods offer advantages over conventional modeling, including the ability to handle large amounts of noisy data from dynamic and nonlinear systems, especially, when the underlying physical relationships are not fully understood. Ultimately, we discussed the accuracy of these techniques via the comparison of their performances. It is relevant to note that the models investigated in the present study are normally applied within deterministic frameworks in professional practices, which encouraged the practice of comparing actual with predicted values. Therefore, the paper presents a comparative study on new generation computational intelligence approaches as superior alternative to the linear and nonlinear regression models for predicting the discharge coefficient of PKW located on a straight open channel.

This paper is organized as follows. The next section provides compendious description of the experimental setup of the study. Brief introduction and model development of these soft computing methods are also described before discussing the results and making concluding remarks. In the “Applications and Results” section, the performances of various models developed are demonstrated by forecasting discharge coefficient of PK weir. Finally, the last section, the “Conclusion” section, provides the concluding remarks of the present study.

2. Experimental setup

This research was carried out in the hydraulic laboratory of water engineering department, faculty of agriculture, Bu-Ali Sina University, Hamadan, Iran. The experiments were carried out in a 0.83-m-wide by 0.5-m-deep by 10-m-long rectangular flume. The main channel consisted of a smooth, horizontal, well-painted steel bed and glass lateral walls. The flume was equipped with a rolling point gauge (readability 0.01-cm) instrumentation carriage, which was used to measure water surface and crest elevations at

various locations upstream of the PK weir after the water level had been allowed to stabilize for a minimum of 5 minutes. All weir data were collected for flows ranging from 5 to 90 l/s. Flow rates were measured using calibrated ultrasonic flow meter. The channel was fed by a centrifuge pump delivering flow in an upstream stilling basin. The slope of the channel was considered zero in all the experiments. To dissipate large eddies, grid walls and wave suppressors were laid upstream along the channel. Water was pumped from the main tank to the flume. Experiments were done for subcritical flow, stable flow conditions and, free overflow conditions. An elevation view of the flume is presented in Fig. 2. A total of 70 experiments were performed for various hydraulic conditions to calculate C_d.

Fig. 2. Overview of weir test setup

The weirs walls were fabricated using 6-mm-thick plexiglass acrylic sheeting. The weirs were assembled with acrylic glue, and the crest level was machined using a computer numerical controlled (CNC) milling machine. Models were designed with N = 4, and featured with a flat top crest. Weirs were sealed with silicon and other sealants. The weirs were installed in the flume on a short, adjustable base and leveled (±0.5mm) using surveying equipment.

An ultrasonic flowmeter with high accuracy was used to measure the discharge of flow in the laboratory flume. An ultrasonic flow meter is a type of flow meter that measures the velocity of a fluid with ultrasound to calculate volume flow. Using ultrasonic transducers, the flow meter can measure the average velocity along the path of an emitted beam of ultrasound, by averaging the difference in measured transit time between the pulses of ultrasound propagating into and against the direction of the flow or by measuring the frequency shift from the Doppler Effect. The flowmeter got calibrated at the beginning of this study. In order to change and control the amount of discharge, a power inverter was used. Invertor changes the input voltage of pump's engine which leads to a change in engine's revolution and finally changes the discharge of flow.

Point gage with accuracy of ±1 mm was used to measure the head of water in the flume.

An overview of the test facility with the PKW installed is presented in Fig. 3. Table 1 presents the description of the weir configurations tested and their corresponding prefixes.

Table 1. Test Weir Dimensions

Test weir dimensions			
P	*24 cm*	*B*	*48 cm*
L	*462.16 cm*	*Bb*	*20 cm*
W	*83 cm*	*Bi*	*14 cm*
Wo	*9.77 cm*	*Bo*	*14 cm*
Wi	*9.77*	*Ts*	*0.6 cm*

In this table, important geometric parameters for PK weir design (most of which are shown in Fig. 1) include the total weir height (P), crest centerline length (L), inlet and outlet key width (W_i and W_o), PK weir spillway width (W), footprint length (B_b), upstream (outlet key) and downstream (inlet key) apex cantilever lengths (B_o and Bi, respectively), inlet or outlet key length (B), weir wall thickness (T_s) and number of keys (N).

3. Methodology

3.1. Feed forward back-propagation neural networks

Neural networks are highly flexible modelling tool with the ability to learn the mapping between input and output without knowing a prior relationship between them; therefore, it can be applied to model such random and complex process. In this paper, a feed-forward back-propagation neural network (FFBPN) is used to model the process. The architecture of a FFBPN refers to its framework as well as its interconnection approaches. The framework is often specified by the number of layers and the number of nodes per layer. Back propagation uses a set of pairs of input and output values (called patterns). An input pattern is fed into the network to produce an output, which is then compared with the actual output pattern. The knowledge is presented by the interconnection weights between layers which are adjusted during learning stage. The output from any neuron at any layer is calculated by the following formula:

$$Y_j = f\sum_{i=1}^{n} w_{ij}x_i + \Theta_j \quad (3)$$

where, Y_j=final output from j^{th} neuron, f=activation function (here logsimoidal transfer function has been used), n=number of neurons in the previous layer, w_{ij}=synaptic weight between i^{th} and j^{th} neuron, x_i=output from i^{th} neuron, Θ_j=bias at j^{th} neuron. The typical performance function used for training feedforward neural networks is the mean sum of squares (MSE) of the network errors:

$$MSE = \frac{1}{N}\sum_{j=1}^{N}\left(T_j - Y_j\right)^2 \quad (4)$$

where, T_j=target output from j^{th} neuron. The weights are updated by gradient descent method which is as follows:

$$\left[w_{ij}\right]_{p+1} = \left[w_{ij}\right]_p + \left[\Delta w_{ij}\right]_{p+1} \quad (5)$$

where, p=iteration number

$$\left[w_{ij}\right]_{p+1} = -\eta\frac{\partial E_j}{\partial w_{ij}} + \alpha\left[\Delta w_{ij}\right]_p \quad (6)$$

where, η=learning rate, α=momentum coefficient.

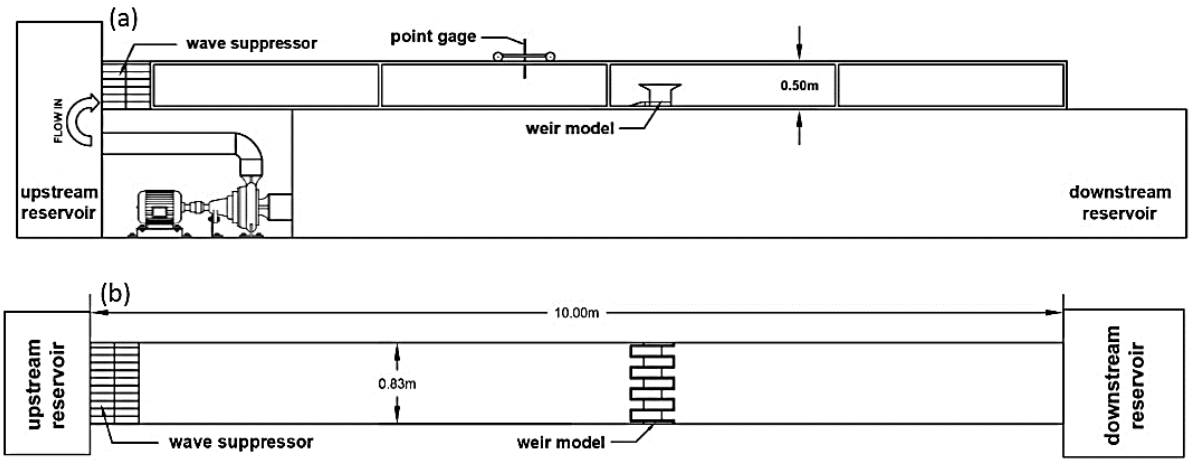


Fig. 3. Overview of Test flume (a) side view and (b) plan view

3.2. Gene-expression programming

Gene-expression programming (GEP) is an extension to genetic programming (GP) in which straightforward or direct chromosomes are encoded to the entities, after that changed into an expression parse tree totally isolating the genotype & phenotype which makes GEP much quicker (100-10,000 times) than the GP [5, 31]. For instance, an expression tree of an arithmetical expression (Eq. (7)) is shown in Fig. 4.

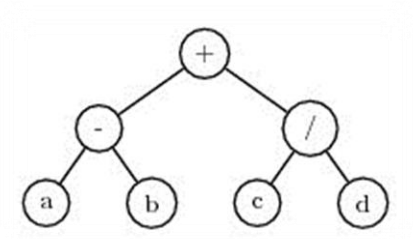


Fig. 4. Expression tree of Eq. (5). [32]

$$(a/b) + (c * d) \quad (7)$$

In GEP, there are numerous qualities in a chromosome & a few subprograms are encoded with every quality. In this way, any project can be encoded for productive advancement of the arrangements by the novel structures of the qualities in the GEP design. The whole achievable locale of the issue is utilized by the novel structure of the qualities in the GEP to have all around sorted out hereditary administrators searching for the arrangements. In GEP, more perplexing investigative & innovative projects can be comprehended with the assistance of straight chromosomes and expression trees (ET). Each direct chromosome is controlled hereditarily, i.e. replication, mutation, recombination and transposition [33]. They are made out of qualities basically involved the head & tail part. The tail length (t_l) is an element of head length (h_l) and number of contentions of the capacity (n) and communicated as the accompanying mathematical statement:

$$t_1 = h_1(n-1) + 1 \quad (8)$$

The flow chart of GEP is presented in Fig. 5 [34]. The fitness value of ETs is the selection criteria for the existence in the procedure. For reaching the optimal result, trees having the poorest fitness value are terminated. After terminating the poorest fitness valued trees, left over population embraces persisting trees depending on the recognized selection technique.

In this study, the discharge coefficient was selected as the output to develop the GEP model. Four basic arithmetic operators (+, -, *, /) and some basic mathematical functions (sin, cos, tan, log, power) were utilized. A large number of generations (5,000) were tested. First, the maximum size of each program was specified as 256, starting with 64 instructions for the initial program. The functional set and operational parameters used in GEP modelling during this study are listed in Table 2.

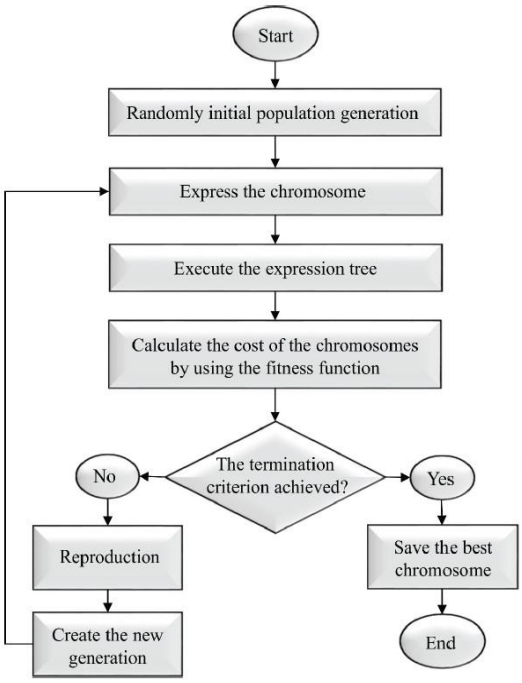


Fig. 5. Flowchart of GEP algorithm [12].

3.3. Least Square support vector machine

Least square support vector machine (LSSVM) models are used to approximate the nonlinear relationship between input variables and output variables with certain accuracy [13, 35]. The LSSVM model is to use the least square linear system as the loss function, and the inequality constraints are revised as the equality constraints in the LSSVM model.

The given training sample is $S\ \{(x_i, y_i)|\ i= 1,2,3,...,m\}$, m is the number of samples, the set $\{x_i\}\in R$ represents the input vector, $y\in\{-1,1\}$ indicates the corresponding desired output vector, the input data is mapped into the high dimensional feature space by using nonlinear mapping function ϕ. Then the existing optimal classification hyperplane must meet the following conditions:

Table 2. Parameters of the optimized GEP model

Parameter	Description of parameter	Setting of parameter
P_1	Function set	+,–, *, /
P_2	Population size	250
P_3	Mutation rate	0.034
P_4	Inversion rate	0.1
P_5	One-point recombination rate	0.3
P_6	Two-point recombination rate	0.3
P_7	Gene recombination rate	0.1
P_8	Gene transportation rate	0.1
P_9	Linking Function	Addition
P_{10}	Program size	Initial64, maximum 256

$$\begin{cases} \omega^T x_i + b \geq 1, & y_i = 1 \\ \omega^T x_i + b \leq -1, & y_i = -1 \end{cases} \quad (9)$$

where ω is Omega vector of superplane, *b* is offset quantity. Then the classification decision function is described as follow:

$$f(x_i) = sgn(\omega^T x_i + b) \quad (10)$$

The classification model of LSSVM is described by the optimization function $\min_{\omega,\xi,b} J(\omega, \xi_i)$:

$$\min_{\omega,\xi,b} J(\omega, \xi_i) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\gamma \sum_{i=1}^{m} \xi_i^2 \quad (11)$$

$$s.t. \;\; y_i[\omega^T \phi(x_i) + b] = 1 - \xi_i, i = 1,2,3,\dots,m \quad (12)$$

where ξ_i is slack variable, *b* is offset, ω is support vector, $\xi=(\xi_1,\xi_2,\dots,\xi_m)$,γ is classification parameter to balance the fitness error and model complexity.

The optimization problem is transformed into its dual space. Lagrange function is introduced to solve it. The corresponding optimization problem of LSSVM model with Lagrange function is described as follow:

$$L(\omega, b, \xi, \alpha) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\gamma \sum_{i=1}^{m} \xi_i^2 - \sum_{k=1}^{m} \alpha_i\{y_i[\omega^T \phi(x_k) + b] - 1 + \xi_i\} \quad (13)$$

where α_i is the Lagrange multiplier, and $\alpha_i \geq = 0 (i=1,2,3,\dots,m)$. Then the classification decision function is described as follow:

$$f(x_i) = sgn(\sum_{i=1}^{m} \alpha_i y_i K(x, x_i) + b) \quad (14)$$

3.4. Extreme learning machine

Extreme learning machine (ELM) is the modified version of single-hidden layer feed forward networks (SLFN) (26, 36) Due to the random determination of the input weights and hidden biases, ELM requires numerous hidden neurons. In practice, the number of hidden neurons should be larger than the number of the variables in dataset, since the useless neurons from the hidden layer will be pruned automatically. Fig.6 shows the topology of a single hidden layer feed forward neural network based on ELM using the activation function, $g(x)= sig(w_i. x_i+b)$.

In SLFN, the relationship between input (x) and output (t) is given below:

$$\sum_{i=1}^{L} \beta_i G(a_i, x_j, b_j) = t_j \quad j = 1, \dots, N \quad (15)$$

where *N* is the number of samples, *L* is the number of hidden nodes, β_i is the output weight and (a_i, b_i) is the *i*th parameter of the *i*th hidden node. In this study, L_x and L_y are used as inputs of the ELM. The output of ELM is d. So, x = [L_x, L_y] and $t = [d]$. Eq. (15) can be written in Eq. (16):

$$H\beta = T \quad (16)$$

where

$$H = \begin{bmatrix} G(a_1, x_1, b_1) & \dots & G(a_L, x_L, b_L) \\ \vdots & \dots & \vdots \\ G(a_1, x_N, b_1) & \dots & G(a_L, x_N, b_L) \end{bmatrix}_{N\times L}, \quad \beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L\times m}, \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_L^T \end{bmatrix}_{N\times m} \quad (17)$$

The value of β is determined from Eq. (18):

$$\beta = H^{-1}T \quad (18)$$

where H^{I} is the Moore-Penrose generalized inverse [35].

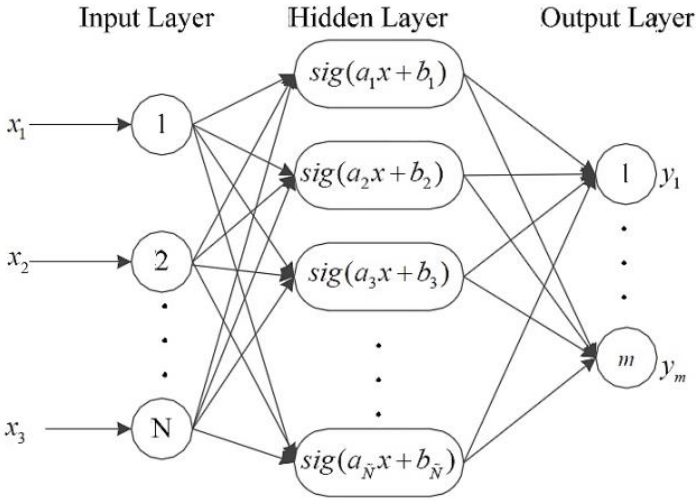


Fig 6. The topology single hidden layer feed forward neural network using ELM. [37]

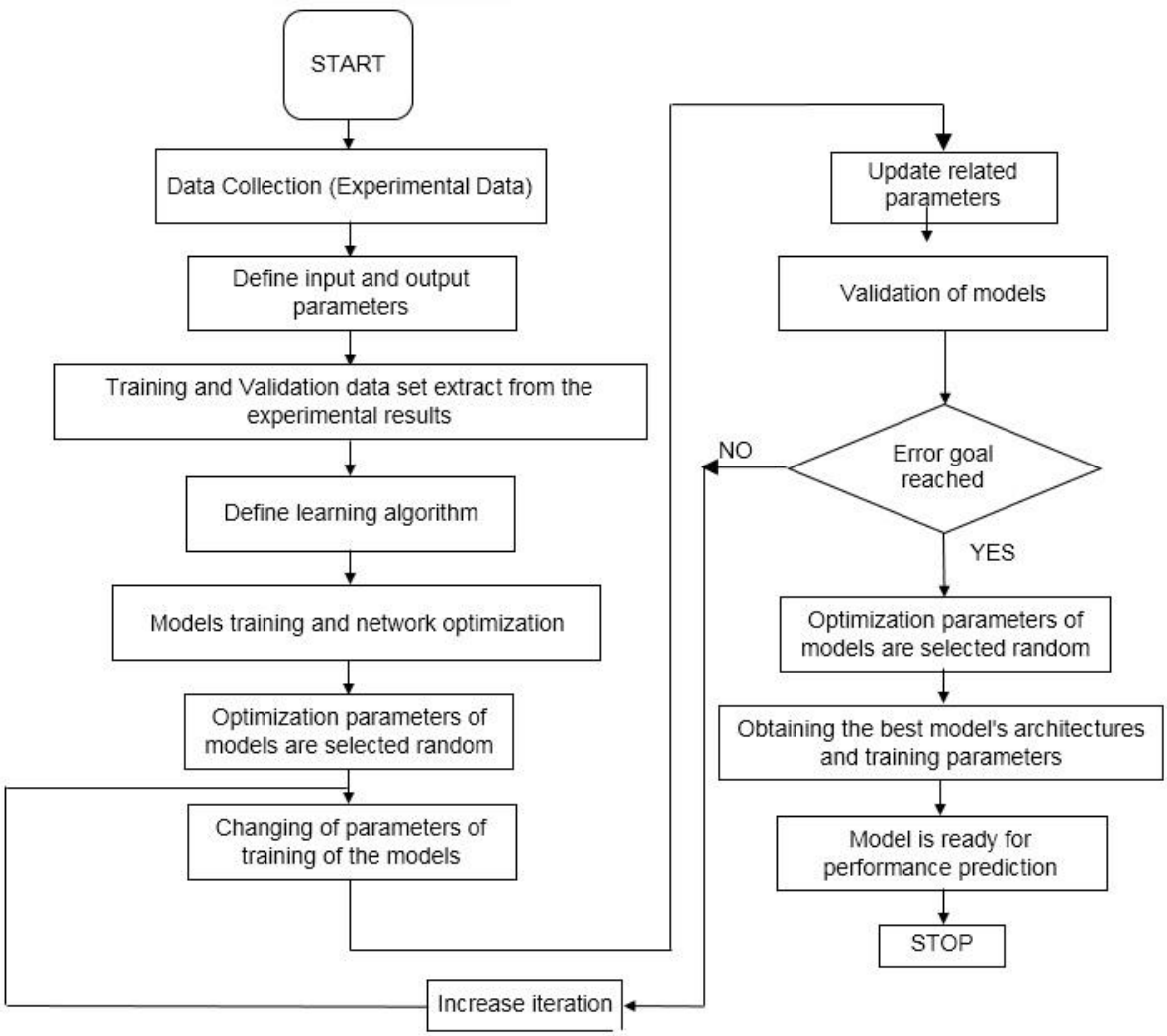


Fig 7. The flow chart of overall research framework

3.5. Model verification

The results of the comparison between equation and those proposed models (FFBPN, GEP, LSSVM, and ELM) in this study are presented herein, in terms of criteria for coefficient of determination (R^2), The Nash–Sutcliffe model efficiency coefficient (NS), root mean square error (RMSE), mean absolute relative error (MAPE), scatter index (SI) and BIAS, as defined below:

$$R^2 = \left[\sum_{i=1}^{n}(x_i-\bar{x})\,(y_i-\bar{y})/\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}\right]^2 \quad (19)$$

$$NS = 1-\frac{\sum_{i=1}^{n}(x_i-y_i)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad (20)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-y_i)^2} \quad (21)$$

$$MARE = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i-y_i|}{x_i} \quad (22)$$

$$SI = \frac{RMSE}{\bar{x}} \quad (23)$$

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(x_i-y_i) \quad (24)$$

where y_i and x_i are the modeled and calculated C_d values, respectively, and $\bar{y}$ and $\bar{x}$ are the mean modeled and calculated C_d values, respectively. The afore-mentioned indexes represent the estimated values as prediction error average but provide no information on the prediction error distribution of the presented models. The RMSE index only indicates a model's ability to predict a value away from the mean. Therefore, in order to test the effectiveness of the model developed, it is important to test the model using some other performance evaluation criteria such as threshold statistics [38, 39]. The threshold statistics (TS) not only give the performance index in terms of predicting C_d but also the distribution of the prediction errors.

The threshold statistic for a level of *x%* is a measure of the consistency in forecasting errors from a particular model. The threshold statistics are represented as TS_x and expressed as a percentage. This criterion can be expressed for different levels of absolute relative error from the model. It is computed for x% level (TL) as

$$TS_x = \frac{Y_x}{n}\times 100 \quad (25)$$

where Y_x is the number of computed C_d (out of n total computed) for which the absolute relative error is less than x% from the model.

4. Applications and results

The FFBPN, GEP, LSSVM and ELM models were compared based on their performance in training and validation sets. For implementation of the AI techniques, among the total 70 data sets, 48 sets (of about 70% of total) were randomly selected as "training data", while the remaining 22 sets (of about 30% of total) were employed for "validation data". To predict discharge coefficient of the piano key weir, three input parameters, that is, the dimensionless upstream head H/P, the dimensionless parameter H/L, and discharge (Q) were considered in the study.

FFBPN and GEP includes all parameters and LSSVM and ELM are used to study the effects of not using each of these parameters on predicting the target in order to analyze sensitivity.

For the same basis of comparison, the same training and validation sets, respectively, were used for all the above models developed, whilst the six quantitative

standard statistical performance evaluation measures were employed to evaluate the performances of various models developed. The experiments were restricted to subcritical flow in the main and side channels. The range of various parameters used in the present study is given in Table 3 and 4.The results are summarized in Tables 5 and 6. It is apparent that all of the performances of these models are almost similar during training as well as validation. In order to get an effective evaluation of the FFBPN, GEP, LSSVM and ELM models performance, the best model structures, have been used to compare the models. From the best fit model, it was found that the difference between the values of the statistical indices of the training and validation set did not vary substantially. It was observed that all four models generally gave low values of the *RMSE*, *MARE* and *SI* as well as high *NS* and R^2, the performances of the FFBPN, GEP, LSSVM and ELM models performance in the C_d estimating were satisfactory. The present research is conducted accounting to the flow chart shown in Fig. 7.

Table 3. Range of variables studied

H/P	*0.027-0.304*
H/L	*0.002-0.015*
Q (lit/sec)	*5.2-90*

Table 4. Statistical parameters of training and testing sets

	H/P		*H/L*		*Q*	
Parameter	Training set	Validation set	Training set	Validation set	Training set	Validation set
Maximum	*0.295*	*0.304*	*0.015*	*0.011*	*90*	*80*
Minimum	*0.030*	*0.027*	*0.002*	*0.002*	*11.1*	*5.2*
Average	*0.159*	*0.148*	*0.008*	*0.010*	*43.80*	*41.33*

Table 5. Quantitative study of the presented models' results in comparison with the experimental results using verification criteria (training period).

Models	R^2	*NS*	*RMSE*	*MARE*	*SI*	*BIAS*
FFBPN	0.910	0.905	0.029	0.063	0.070	-0.0008
GEP	0.957	0.942	0.017	0.041	0.043	-0.0005
LSSVM	0.983	0.971	0.008	0.016	0.024	-0.0003
ELM	0.999	0.988	0.003	0.005	0.010	-0.0001

Table 6. Quantitative study of the presented models' results in comparison with the experimental results using verification criteria (validation period).

Models	R^2	*NS*	*RMSE*	*MARE*	*SI*	*BIAS*
FFBPN	0.883	0.820	0.038	0.072	0.084	-0.0093
GEP	0.939	0.927	0.026	0.054	0.058	-0.0006
LSSVM	0.978	0.969	0.015	0.027	0.033	-0.0004
ELM	0.997	0.986	0.016	0.029	0.011	-0.0002

The optimal architecture of the FFBPN models and its parameter variation were determined based on the minimum value of the mean squared error (*MSE*) of the training and testing sets. For this model, the *logsig* and *purelin* functions were respectively found to be optimal activation functions for the hidden and output layers. In the models, the number of iterations was 10,000 and the optimal number of neurons in the hidden layer was found to be 1. With the increase in number of neurons, the networks yielded several local minimum values with different *MSE* values for the training set. Selection of an appropriate number of nodes in the hidden layer is a very important aspect as a larger number of them may result in over-fitting, while a smaller number of nodes may not capture the information adequately. Fletcher and Goss (1993) suggested that the appropriate number of nodes in a hidden layer ranges from $(2n^{1/2}+m)$ to $(2n+1)$, where n is the number of input nodes and m is the number of output nodes. Subsequently, the best scenario was developed using various architectures of FFBPN.

A model can be claimed to produce a perfect estimation if the *NS* criterion is equal to 1. Normally, a model can be considered as accurate if the NS criterion is greater than 0.8 [36]. It can be observed from Tables 5 and 6 that the NS values for various applied data driven methods in this study are over 0.8. This indicated that they had good performance during both training and validation and these models achieved acceptable results. It also showed that the ELM model had the smallest value of the *RMSE* as well as higher value of R^2 in the training as well as validation period, so, it was selected as the best-fit model for predicting the C_d in this study. Also, the NS values for the ELM model predict of the C_d value were higher than those for the other models, which indicates that the overall quality of estimation of the ELM model is better than the FFBPN, GEP and LSSVM models according to NS. Compared among different models from the *MARE* and *SI* viewpoints, the ELM model performed a bit better than the other. For analyzing the results during training phase, it can be observed that the ELM model outperformed all other models. But in the validation phase, the LSSVM model obtained the best *RMSE* and *MSRE* statistics of 0.015 and 0.027, respectively; while the ELM model obtained the best R^2, *NS*, *SI* and *BIAS* statistics of 0.997, 0.986, 0.011 and -0.0001. Thus, in the validation phase, as seen in Tables 5 and 6, the values with the ELM model prediction were able to produce a good, near forecast, as compared to those with other models, whilst it can be concluded that the LSSVM model obtained the best relative error between the observed and modeled C_d. Furthermore, as can be seen from Tables 5 and 6 that the virtues or defect degrees of forecasting accuracy are different in terms of different evaluation measures during the training phase and the validation phase. ELM model is able to obtain the better forecasting accuracy in terms of different evaluation measures not only during the training phase but also during the validation phase.

It appears that while assessing the performance of any model for its applicability in predicting C_d, it is not only important to evaluate the average prediction error but also the distribution of prediction errors. The statistical performance evaluation criteria employed so far in this study are global statistics and do not provide any information on the distribution of errors. Therefore, in order to test the robustness of the model developed, it is important to test

the model using some other performance evaluation criteria such as threshold statistics. The distribution of errors is presented in Fig. 8, which gives a clear indication of better performance by the ELM model. It can be also observed from Fig. 8 that all applied models perform almost similarly in terms of the distribution of the errors. According to this figure, nearly 92% of the discharge coefficients estimated using ELM have a relative error lower than 2%. However, according to Fig. 8, for FFBPN, GEM and LSSVM models, almost 58%, 75% and 98% of the estimated amounts have less than 8% error, respectively, and this is approximately 100% for ELM. Based on the given explanations so far, it can be said that ELM is fairly accurate in estimating C_d.

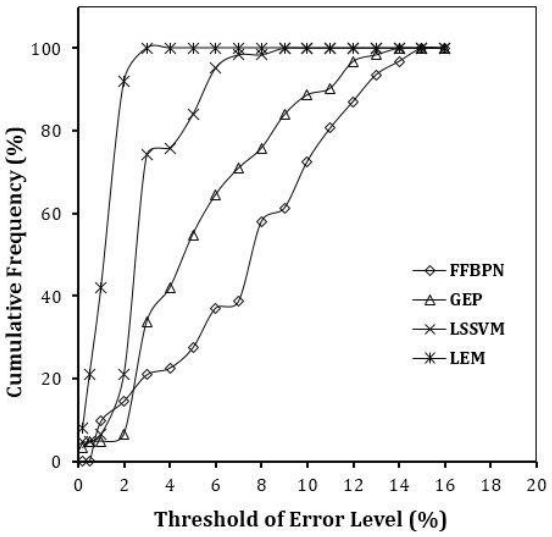


Fig. 8. Distribution of forecast error across different error thresholds for all the models.

Fig. 9 displays the estimation discharge coefficients as presented by the 4 models in this study against the experimental results. This figure demonstrates that LSSVM and ELM estimated the results fairly accurately, with all the discharge coefficient amounts estimated by this model having a relative error below 10%. It was obviously seen from these figures that the ELM estimates were closer to the corresponding observed flow values than those of the other models. As seen from the fit line equations (assume that the equation is y = ax + b) in the scatter plots that *a* and *b* coefficients for the ELM model are, respectively, closer to the 1 and 0 with a higher R^2 value of 0.997 than other models for C_d.

Overall, LSSVM and GEP models gave good prediction performance and were successfully applied to establish the forecasting models that could provide accurate and reliable C_d prediction. The results suggested that the ELM model was superior to the other in this forecasting. The reason for a better prediction accuracy of ELM model than other models primarily lied in the shortcoming of the models, e.g. slowly learning speed, over-fitting, curse of dimensionality and convergence to local minimum. Conversely, ELM model was based on the empirical risk minimization principle, which could attack the problem in theory.

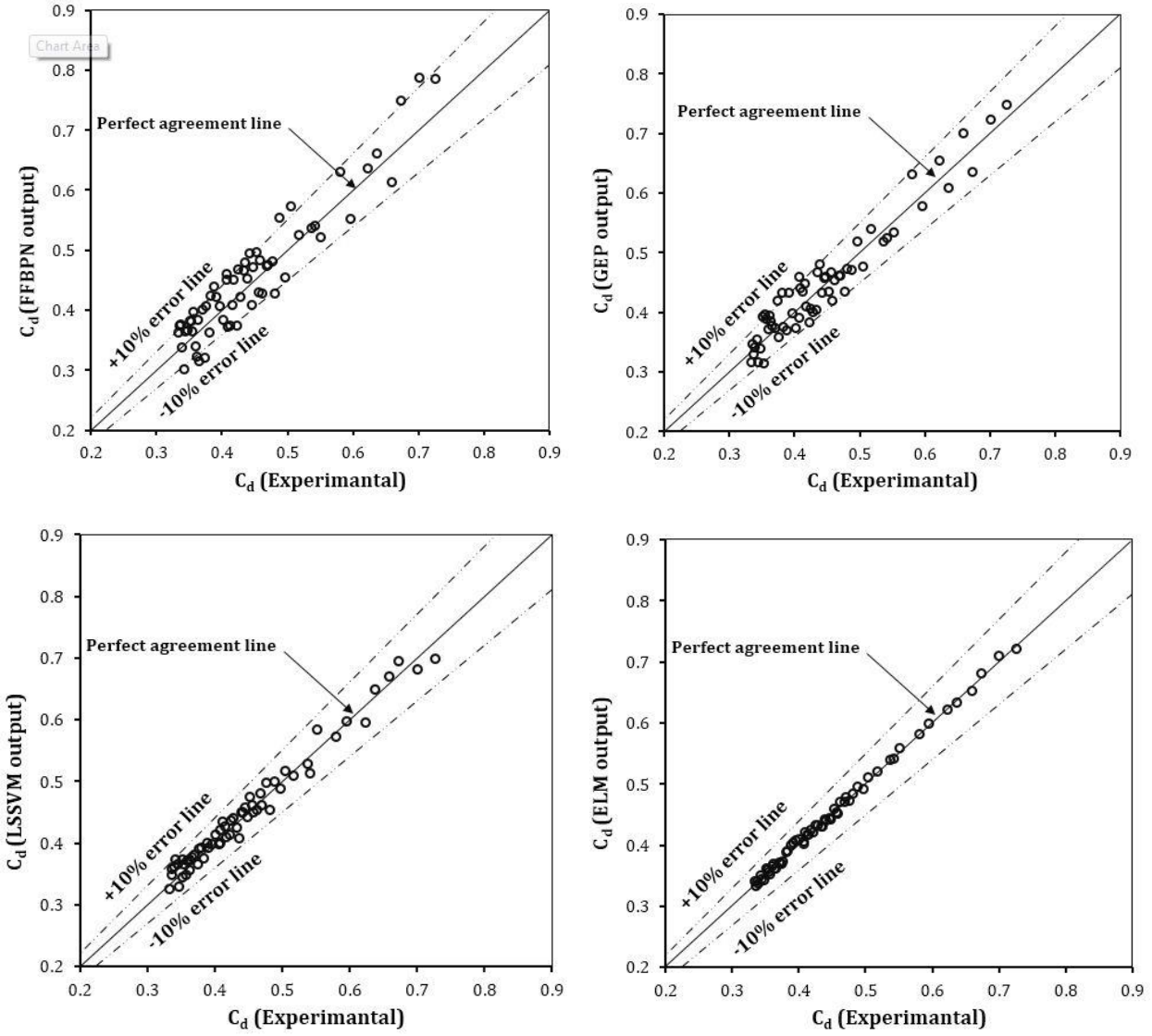


Fig. 9. Comparing the estimated results of the four models to the laboratory results

The distributions of forecast errors by all models for C_d were presented in Fig. 10 from which it was clearly evident that the ELM performed better than the other models. It could be observed from Fig. 10, for all developed models no specific clustering was observed.

The data in Fig. 11 showed that the ELM was extremely closer to the experimental C_d values than other approaches used in this study. So it was evident that ELM and LSSVM consistently performed better than GEP and FFBPN.

C_d data for each models, presented in Fig. 12, were normalized relative to modeled results to experimental data at a dimensionless upstream head, H_T/P in performance. As shown in Fig. 12, for ELM model, C_d (ELM)/C_d (experimental) was nearest to 1.0 and the results of C_d prediction demonstrated the effectiveness and efficiency of the ELM and LSSVM models.

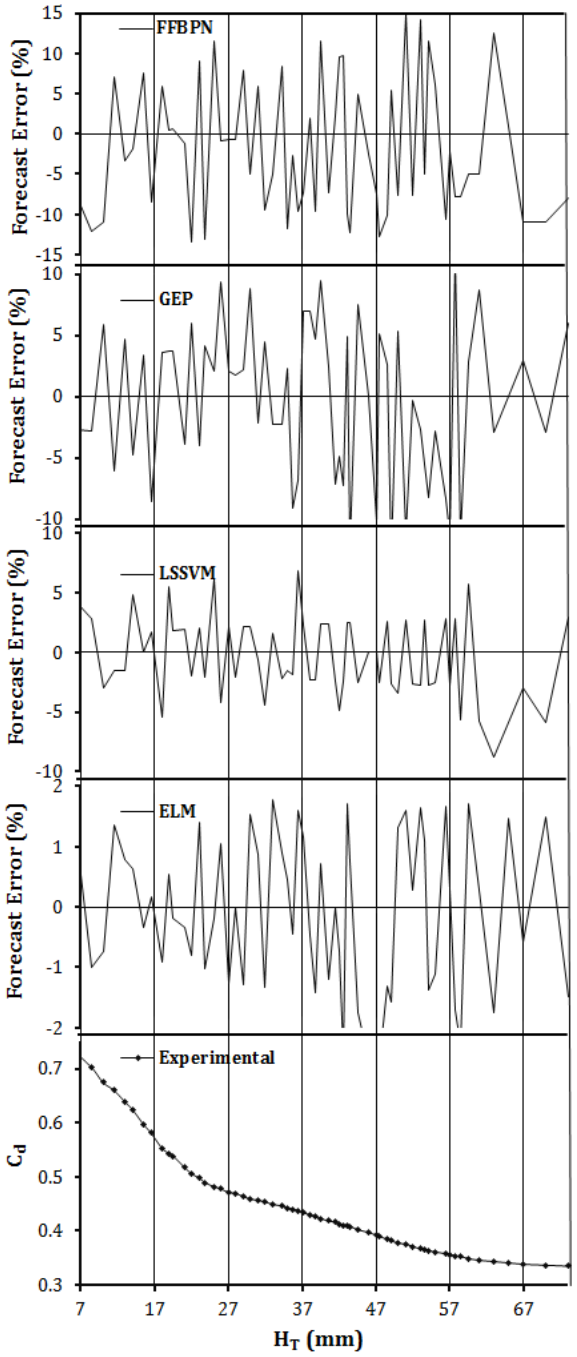


Fig. 10. Distribution of forecast error across the full range of discharge coefficient.

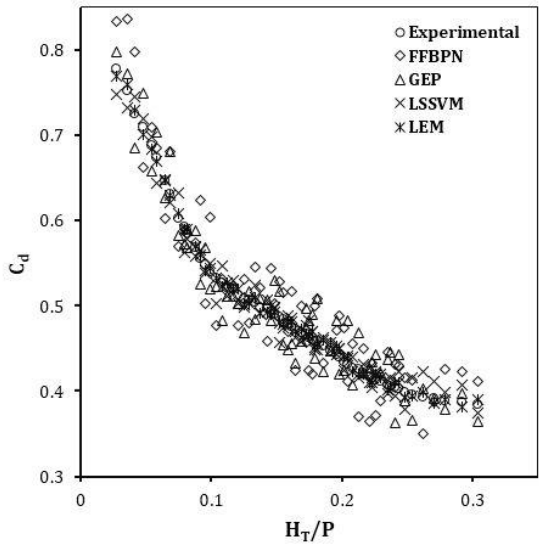


Fig. 11. FFBPN, GEP, LSSVM and ELM modeled C_d versus HT/P data

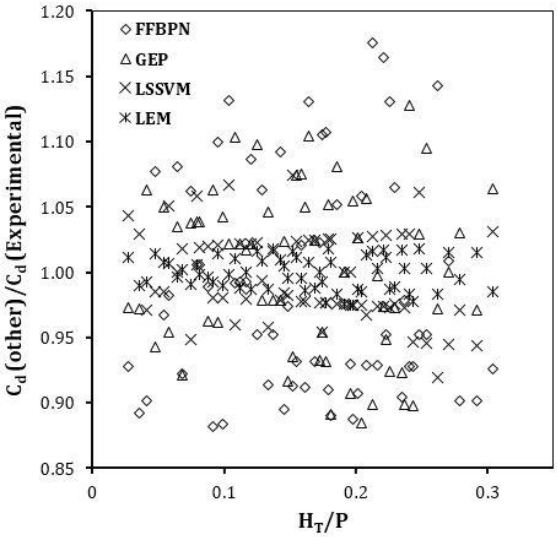


Fig. 12. FFBPN, GEP, LSSVM and ELM C_d versus HT/P data normalized to experimental data

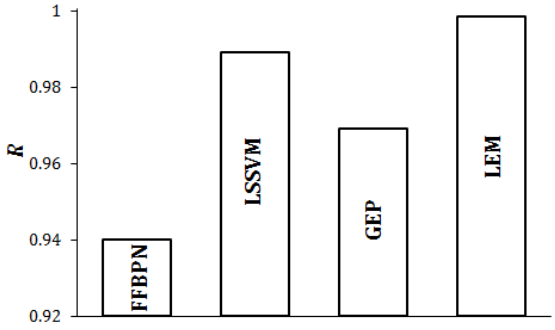


Fig. 13. Bar chart of R values of the different models.

The LSSVM model's performance provided encouraging output. Since four models had been developed to predict the discharge coefficient of piano key weir, a comparative study had been carried out between the adopted GPR, ELM and LSSVM models (Fig. 13).

It can be concluded that all five presented models provide fairly good results for prediction of discharge coefficient of PKW. The one of the difference between the estimations of different models is their relative error. Accordingly, considering the given inputs, the fact that compared to other models the greatest relative error by LEM was the lowest for both sets of data (training and validation), and the RMSE and NS indexes presented for this model are good, it can be argued that compared to other models, LEM and LSSVM can serve as a replacement method. Regarding the explanations for the different models presented in this study, ELM is more accurate than the rest. Therefore, it can be stated that LEM, which produced good results with both different datasets, can be used confidently to estimate the discharge coefficient of PK weirs. It is also worth mentioning that LSSVM and GEP has high precision in discharge coefficient estimation and may be used as an alternative to existing methods.

5. Conclusion

In the present study, various data driven techniques (e.g., FFBPN, GEP, LSSVM and ELM), were compared to estimate discharge coefficient of a piano key weir. To achieve this objective, an experimental data set in the hydraulic laboratory of Bu-Ali Sina University were employed to develop various models investigated in this study. The methods utilized the statistical properties of the data series. The obtained results indicated that computational intelligence methods were powerful tools to model C_d and could give good estimation performance. Therefore, the results of the study were highly encouraging and suggested that LSSVM and ELM approaches were promising in modeling C_d, and this might provide valuable reference for researchers and engineers who applied the methods for modeling long-term hydraulic parameters estimating. As the next step, comparing the results of the models, it was seen that the values of R^2 and NS of ELM models were higher than those of other models. Moreover, the RMSE and MSRE values of LSSVM models were lower than those of FBBPN, GEP and ELM models. Therefore, ELM and LSSVM models could improve the accuracy over the other applied models. Overall, the analysis presented in this study showed that the ELM method was superior to the FFBPN, GEP and LSSVM in forecasting C_d of piano key weir. In general, the implementation of all intelligence models in the present study illustrated the flexibility of C_d modeling. It is hoped that future research efforts will focus in these directions, i.e. more efficient approach for training multi-layer perceptron of ANN model, improve the prediction accuracy, especially for the higher values of C_d, by combining or improving model parameters, fine-tuning of algorithms for selecting more appropriate parameters of GP evolution, saving computing time or more efficient optimization algorithms in searching optimal parameters of SVM model etc., to improve the accuracy of the forecasting models in terms of different evaluation measures for better planning, design, operation, and management of various engineering systems.

Acknowledgments

The authors would like to express their appreciation to the anonymous reviewers for their valuable comments and suggestions to improve the manuscript

References

[1] Anderson, R.M. (2011). "Piano key weir head discharge relationships." M.S. thesis, Utah State University, Logan, UT.

[2] Emami, S., Arvanaghi, H., Parsa, J. (2018). "Numerical Investigation of Geometric Parameters Effect of the Labyrinth Weir on the Discharge Coefficient" J. Rehab. Civil Eng. Vol. 6, pp. 1-9.

[3] Ahmadi, M.A., Pouladi, B., Javvi, Y., Alfkhani, S., Soleimani, R. (2015). "Connectionist technique estimates H2S solubility in ionic liquids through a low parameter approach." J Supercrit Fluids, Vol. 97, pp. 81–87.

[4] Kabiri-Samani, A., Javaheri, A. (2012). "Discharge coefficients for free and submerged flow over Piano Key weirs." J. Hydraulic Res. Vol. 50, pp. 114-120.

[5] Ribeiro, M.L., Bieri, M., Boillat, J.L., Schleiss, A.J., Singhal, G., Sharma, N. (2012). "Discharge Capacity of Piano Key Weirs." J Hydraul Eng, Vol.138, pp.199-203.

[6] Du, S., Zhang, J., Deng, Z., Li, J. (2014). "A novel deformation prediction model for mine slope surface using meteorological factors based on kernel extreme learning machine." Int. J. Eng. Res. Africa, Vol. 12, pp. 67-81.

[7] Dursun, O.F., Kaya, N., Firat, M. (2012). "Estimating discharge coefficient of semi-elliptical side weir using ANFIS." J. Hydrol. pp. 426-427, pp. 55–62.

[8] Baghalian, S., Bonakdari, H., Nazari, F., Fazli, M. (2012). "Closed-form solution for flow field in curved channels in comparison with experimental and numerical analyses and Artificial Neural Network." Eng. Appl. Comput. Fluid Mech. Vol. 6, pp. 514-526.

[9] Bilhan, O., Emiroglu, M.E., Kisi, O. (2010). "Application of Two different neural network techniques to lateral outflow over rectangular side weirs located on a straight channel." Adv. Eng. Softw. Vol. 41, Issue 6, pp. 831-837.

[10] Emiroglu, M.E., Kisi, O. (2013). "Prediction of discharge coefficient for trapezoidal labyrinth side weir using a neuro-fuzzy approach." Water Resour. Manage, Vol. 27, pp. 1473-1488.

[11] Ebtehaj, I., Bonakdari, H. (2016). "Bed load sediment transport estimation in a clean pipe using multilayer perceptron with different training algorithms." KSCE J. Civil Eng. Vol. 20, pp. 581-589.

[12] Ebtehaj, I., Bonakdari, H., Zaji, A.H., Azimi, H., Khoshbin, F. (2015). "GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs." Eng. Sci. Technol. Intern. J. Vol. 4, pp. 746-757.

[13] Salmasi, F., Yıldırım, G., Masoodi, A., Parsamehr, P. (2013). "Predicting discharge coefficient of compound broad-crested weir by using genetic programming (GP) and artificial neural network (ANN) techniques." Arab. J. Geosci, Vol 6, pp. 2709-2717.

[14] Khoshbin, F., Bonakdari, H., Ashraf, Talesh, S.H., Ebtehaj, I., Zaji, A.H., 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. Vol. 48, pp.933-948.

[15] Zaji, A.H., Bonakdari, H., Shamshirband, S., Qasem, S.N. (2015). "Potential of particle swarm optimization based radial basis function network to predict the discharge coefficient of a modified triangular side weir." Vol. 45, pp. 404-407.

[16] Kisi, O., Emiroglu, M.E., Bilhan, O., Guven, A. (2012). "Prediction of lateral outflow over triangular labyrinth side

weirs under subcritical conditions using soft computing approaches." Expert Syst Appl, Vol 39, Issue 3, pp. 3454-3460.

[17] Ebtehaj, I., Bonakdari, H. (2017). "No-deposition Sediment Transport in Sewers Using of Gene Expression Programming". Soft Computing in Civil Engineering, Vol 1(1): pp. 29-53. doi: 10.22115/scce.2017.46845

[18] Azimi, H., Bonakdari, H., Ebtehaj, I. (2017). "A Highly Efficient Gene Expression Programming Model for Predicting the Discharge Coefficient in a Side Weir along a Trapezoidal Canal." Irrig. Drain. Vol. 66(4), pp.655–666.

[19] Huang, G.B., Zhu, Q.Y., Siew, C.K. (2004). "Extreme learning machine: a new learning scheme of feed forward neural networks." Neural Networks, Vol. 2, pp. 985-990.

[20] Huang, G.B., Zhu, Q.Y., Siew, C.K. (2006). "Extreme learning machine: theory and applications." Neurocomputing, Vol. 70, Issue 1, pp. 489–501. http://dx.doi.org/ 10.1016/j.neucom.2005.12.126.

[21] Ebtehaj, I., Bonakdari, H., Shamshirband, S. (2016). "Extreme learning machine assessment for estimating sediment transport in open channels." Engineering with Computers. Vol. 32, pp.691-704.

[22] Huang, G.B., Chen, L., (2007). "Convex incremental extreme learning machine." Neurocomputing, Vol. 70, pp. 3056-3062.

[23] Jain, A., Ormsbee, L.E. (2002) "Evaluation of short-term water demand forecast modeling techniques: conventional methods versus AI." J Am Water Works Assoc Vol. 94, pp. 64-72.

[24] Jain, A., Varshney, A.K., Joshi, U.C. (2001). "Short-term water demand forecast modeling at IIT Kanpur using artificial neural networks." Water Resour Manag, Vol. 15, pp. 299-321.

[25] Laugier, F. (2007). "Design and construction of the first piano key weir (PKW) spillway at the Goulours dam." The Int J Hydropower and Dams, Vol. 5, pp.94-101.

[26] Laugier, F., Lochu, A., Gille, C., Leite, Ribeiro. M., Boillat, J.L. (2009). "Design and construction of a labyrinth PKW spillway at St-Marc dam, France." The Int. J. Hydropower and Dams, Vol 5, pp.100-107.

[27] Huang, G.B., Wang, D.H., Lan, Y. (2011). "Extreme learning machines: a survey." Int J Machine Learn Cybernet, Vol. 2, pp. 107-122.

[28] Naderpour, H., Rafiean, A. H., Fakharian, P. (2018). "Compressive strength prediction of environmentally friendly concrete using artificial neural networks". Journal of Building Engineering. Vol. 16, pp. 213-219, doi: 10.1016/j.jobe.2018.01.007

[29] Negnevitsky, M. (2005) "Artificial Intelligence: A Guide to Intelligent Systems", Pearson Education.

[30] Rajurkar, M.P., Kothyarib, U.C., Chaube, U.C. (2004). "Modeling of the daily rainfall-runoff relationship with artificial neural network." J. Hydrol. Vol. Vol. 285, pp. 96-113.

[31] Sadri, S., Burn, D.H. (2012). "Nonparametric methods for drought severity estimation at ungauged sites." Water Resour Res 48(12): W12505.

[32] Day, P., Das, A.K. (2016), "A utilization of GEP (gene expression programming) metamodel and PSO (particle swarm optimization) tool to predict and optimize the forced convection around a cylinder." Energy. Vol.95, pp.447-458.

[33] Ferreira, C. (2001). "Gene expression programming: a new adaptive algorithm for solving problems." Complex Syst, Vol. 13, pp. 87-129.

[34] Ferreira, C. (2002) "Gene expression programming in problem solving, in soft computing and industry." Springer. Pp. 635-653.

[35] Serre, D. (2002). Matrices: "Theory and Applications." Springer, Berlin.

[36] Shu, C., Ouarda, T.B.M.J. (2008) "Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system." J Hydrol, Vol. 349, pp. 31-43.

[37] Ertugrul, O.F. (2016). "Forecasting electricity load by a novel recurrent extreme learning machines approach."

Elec. Power Energy Syst. Vol. 78 pp. 429-435.

[38] Suykens, J. (2001). "Support vector machines: a nonlinear modeling and control perspective." Europ J Control, Vol. 7, pp. 311-327.

[39] Toro, C.H.F., Meire, S.G., Gálvez, J.F., Fdez-Riverola, F. (2013). "A hybrid artificial intelligence model for river flow forecasting." Appl Soft Comput, Vol.13, pp. 3449-3458.

[40] Fletcher, D., Goss, E. (1993). "Forecasting with neural networks: an application using bankruptcy data." Inform Manag, Vol. 24, pp. 159-167.