



# Application of Gaussian Process Regression Model to Predict Discharge Coefficient of Gated Piano Key Weir

Masood Akbari<sup>1</sup> · Farzin Salmasi<sup>1</sup>  · Hadi Arvanaghi<sup>1</sup> · Masoud Karbasi<sup>2</sup> · Davood Farsadizadeh<sup>1</sup>

Received: 13 February 2019 / Accepted: 5 August 2019 /  
Published online: 21 August 2019  
© Springer Nature B.V. 2019

## Abstract

The Piano Key (PK) weir is a new type of long crested weirs. This study was involved the addition of a gate to PK weir inlet keys. It was conducted by the Department of Water Engineering, University of Tabriz, Iran to determine if the gate increased hydraulic performance. A Gated Piano Key (GPK) weir was constructed and tested for discharge ranges of between 10 and 130 l per second. To this end, 156 experimental tests were performed and the effective parameters on the GPK weir discharge coefficient ( $C_d$ ), such as gate dimensions ( $b$  and  $d$ ), gate insertion depth in the inlet key ( $H_{gate}$ ), the ratio of the inlet key width to the outlet key width ( $W_i/W_o$ ) and the head over the GPK weir crest ( $H$ ) were investigated. In addition, application of soft computing to estimate of  $C_d$  was carried out using MLP, GPR, SVM, GRNN, multiple linear and non-linear regressions methods using MATLAB 2018 software. This study suggests the relation for  $C_d$  with non-dimension parameters. The results of this study showed that  $H$ ,  $W_i/W_o$ ,  $H_{gate}$  and  $b$  and  $d$ , had the greatest effect on the GPK weir discharge coefficient, respectively. The GPR method was introduced as a new effective method for predicting discharge coefficient of weirs with  $RMSE = 0.011$ ,  $R^2 = 0.992$  and  $MAPE = 1.167\%$  and provided the best results when compared with other methods.

**Keywords** Gated piano key (GPK) weir · Experimental model · Discharge coefficient ( $C_d$ ) · Gaussian process regression (GPR) · Artificial intelligence

## 1 Introduction

The Piano Key (PK) weir is a long crested weir that improves the capacity of discharge through the constant upstream head by increasing the length of the crest. Due to the resulting high discharge capacity of these types of weir, reservoir capacity is increased and dams are

---

✉ Farzin Salmasi  
salmasi@tabrizu.ac.ir; ferzin.salmasi@gmail.com

stabilised against flooding. The PK weir has been developed by Lempérière, HydroCoop, France and Blanc, Hydraulic and Environmental Laboratory, University of Biskra, Algeria (Lempérière and Ouamane 2003). The first PK weir was built in 2006 on the Goulours dam in France, with the second being installed in 2008 on Saint-Marc dam in France (Laugier 2007; Laugier et al. 2009).

Lempérière (2011) suggested the head – discharge relation for the PK weirs. Kabiri-Samani and Javaheri (2012) presented the experimental relations of discharge coefficients for these types of weirs in two free and submerged forms. Anderson and Tullis (2012) compared the hydraulic performances of labyrinth weirs and rectangular PK weirs. They indicated that the reduced loss of inlet keys, led to better performance from the PK weirs. Anderson and Tullis (2012) showed that for the value of  $\frac{W_i}{W_o}=1.25-1.5$ , the PK weir had the highest efficiency. Machiels et al. (2013) studied the effect of parapet walls on the hydraulic performance of this type of weir and pointed out its positive effect due to the increased height of the weir. In addition to structural and hydraulic studies, the application of artificial intelligence techniques for estimating the weirs' discharge coefficients can be noted. Emiroglu et al. (2011) estimated the discharge coefficient of the triangular labyrinth side-weir on a straight channel using the Artificial Neural Network (ANN) model. Results of this model were more successful than the multiple non-linear regression model. By using the ANN model, Bilhan et al. (2011) estimated the discharge coefficient of triangular labyrinth side-weirs in curved channels. They suggested that the ANN model was more capable of estimating the discharge coefficient compared to the multiple non-linear regression model. Dursun et al. (2012) estimated the discharge coefficient of semi-elliptical side weirs by making use of Adaptive Neuro-Fuzzy Inference System (ANFIS). Their results indicated the success of this technique over Multiple Linear Regression (MLR) and Nonlinear Regression (NLR). Salmasi et al. (2012) estimated the discharge coefficient of compound broad-crested weirs by applying the Genetic Programming (GP) techniques and ANN. Results showed that the genetic programming technique is more capable than the ANN in terms of estimating the weir's discharge coefficient. Ebtehaj et al. (2015) estimated the discharge coefficient of triangular labyrinth side-weirs using the gene expression programming. By using the Support Vector Regression (SVR), Zaji et al. (2016) predicted the modified discharge coefficient of diagonal side weirs (in triangular form). Results obtained from their investigation showed that SVR-RBF outperforms the SVR-poly. Shamshirband et al. (2016) estimated the optimum discharge coefficient of side weirs using the ANFIS model. Results showed the ANFIS model with five inputs was more accurate than the ANFIS model with a single input. Parsaie (2016) investigated the discharge coefficient of side weirs using the experimental formulas, Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF). The MLP model yielded the best result. Haghiabi et al. (2018) estimated the discharge coefficient of triangular labyrinth side-weirs using the ANFIS. Results of this model were compared to MLP neural network's results. Results achieved from comparing the MLP to ANFIS indicated that both models function very appropriately, but the ANFIS structure is more accurate.

There have been many applications of the Gaussian Process Regression (GPR) method in water engineering sciences. For example, Pasolli et al. (2010) estimated the concentration of chlorophyll in subsurface waters by the GPR method using remote sensing data. Grbić et al. (2013) predicted the stream water temperature based on the GPR method. Their technique can be used as a basis for prediction tools for water resource managers. Once an online Bayes filtering was made on the global surface temperature data, Wang and Chaib-draa (2017) analysed the temperature using the GPR method. They suggested this technique was better than various types of Gaussian Process (GP) and was seen as an accurate and efficient system for analysis of global temperatures. Karbasi (2017) investigated the 10-year statistical data

(2000–2009) from Zanjan Synoptic Station, Iran, and predicted the multi-step ahead daily reference evapotranspiration by the method of Wavelet-GPR and GPR models. Results showed that the hybrid method of Wavelet-GPR was more capable and accurate over the GPR model in predicting the daily evapotranspiration.

A review of pertinent literature shows that previous studies regarding PK weirs have not investigated the effect of an additional gate on each of the inlet keys. This study is the first to examine such an effect. It also investigates other parameters of PK weirs, using the Gaussian Process Regression (GPR) method to predict coefficient of discharge ( $C_d$ ) and compare this result with Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), Generalized Regression Neural Networks (GRNN), multiple linear and nonlinear regression methods and with suggesting the relation for  $C_d$ . Thus, this study is distinguished from other studies in this area. In addition, the GPR method has not been used to investigate the discharge coefficient of weirs. The present study introduces this method for predicting discharge coefficient of weir.

## 2 Materials and Methods

### 2.1 Dimensional Analysis

Dimensional analysis aims at diagnosing the effective parameters in the studied phenomenon and determining the dimensionless ratios. Due to the geometrical complexity of Gated Piano Key weirs (GPK weirs), the discharge passing over them is subject to parameters of Eq. 1.

$$Q = f(H, H_{up}, H_{down}, b, d, H_{gate}, L, W, P, W_i, W_o, B_i, B_o, N, S_i, S_o, T_s, g, \mu, \sigma, \rho) \quad (1)$$

where  $Q$  is the discharge overflow from GPK weir,  $H$  is head on GPK weir,  $H_{up}$  is the upstream head,  $H_{down}$  is the downstream head,  $b$  is the length of rectangular gate,  $d$  is the width of rectangular gate,  $H_{gate}$  is the water head from the rectangular gate center to the crest of weir,  $L$  is the crest length [ $L = N(W_i + W_o + 2B)$ ],  $W$  is the total weir width or flume width,  $P$  is the total weir height,  $W_i$  is the inlet key width,  $W_o$  is the outlet key width,  $B_i$  is the downstream or inlet key overhang length,  $B_o$  is the upstream or outlet key overhang length,  $N$  is the weir cycle number,  $S_i$  is the inlet key slope,  $S_o$  is the outlet key slope,  $T_s$  is the weir wall thickness,  $g$  is the gravitational acceleration,  $\mu$  is dynamic viscosity,  $\rho$  is density and  $\sigma$  is surface tension. Figure 1 shows Gated Piano key weir geometric and hydraulic parameters.

Since  $H_{up}$  and  $H_{down}$  are function of the  $H$ , they can be excluded. The number of cycles, slope of inlet and outlet keys can be seen as a proportion of other parameters, hence, they can be excluded, too. Using the dimensional analysis and considering the dimensionless numbers of Reynolds ( $R_e$ ), Weber ( $W_e$ ) and dimensionless ratios, Eq. 1 is reduced to Eq. 2.

$$\frac{Q}{g^{\frac{1}{2}}BH^{\frac{3}{2}}} = f\left(R_e, W_e, \frac{H}{P}, \frac{L}{B}, \frac{W_i}{W_o}, \frac{B_i}{L}, \frac{B_o}{L}, \frac{b}{P}, \frac{d}{P}, \frac{H_{gate}}{P}\right) \quad (2)$$

In accordance with Eq. (2), it can be said that  $C_d$  is subject to the geometrical parameters of GPK weirs and dimensionless numbers of Reynolds and Weber. The Reynolds number in channels is sufficiently large, hence the viscosity can be ignored (Henderson 1966). Therefore, the Reynolds and Weber numbers are deleted from Eq. 2. To investigate the geometrical parameters of GPK weir, Eq. 3 can be proposed.

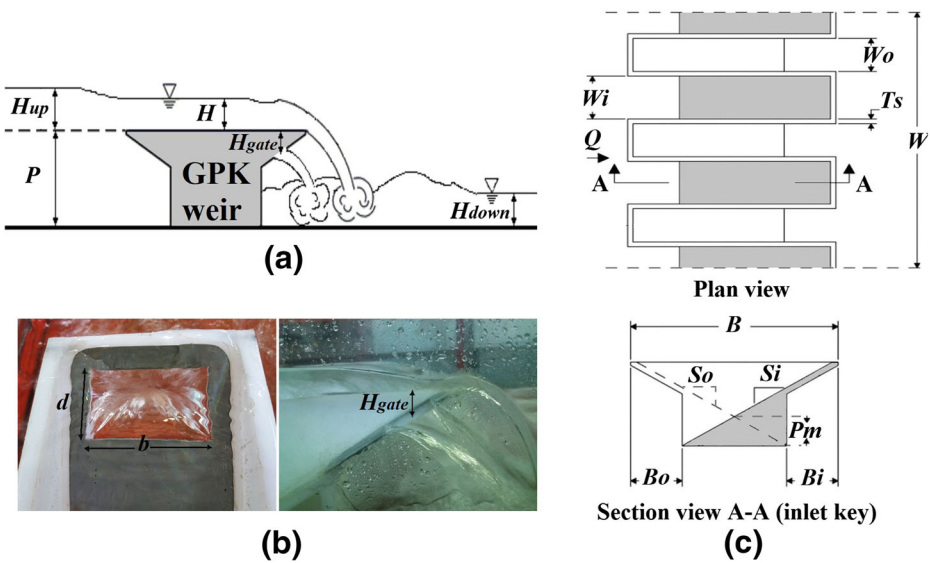


Fig. 1 (A) GPK weir hydraulic parameters, (B) The geometric parameters of gate in GPK weirs, (C) PK weir geometric parameters

$$C_d = f\left(\frac{H}{P}, \frac{L}{B}, \frac{W_i}{W_o}, \frac{B_i}{L}, \frac{B_o}{L}, \frac{b}{p}, \frac{d}{p}, \frac{H_{gate}}{P}\right) \tag{3}$$

### 2.2 Experimental Set up

These experiments were performed in the hydraulic laboratory, Department of Water Engineering, University of Tabriz, Iran. The experiments were conducted on a 10 m long and 0.933 m wide horizontal rectangular flume. The height of the first 3 m of the flume was 1 m with the rest of its length being 0.5 m in height. The floor of the flume was constructed from galvanized iron and the side walls from 10 mm thick glass. The GPK weir was installed on a 1 cm high ramp. The water circulation system included an underground tank equipped with a 100 hp. pump which provided a steady supply of water into a head tank. The upstream tank (head tank) provided water flow into the flume. The weir returned extra water from the head tank into the underground tank, thereby ensuring the discharge in the rotation system would be a fixed amount. The water flowed downstream through the flume, was collected in the collection tank, and then flowed into the underground tank to be pumped again into the head tank (Fig. 2).

The discharge rate was measured using an ultrasonic flow meter with 1% precision. Discharge measuring sensors were installed on a 10-in. pipe which supplied water to the flume. The discharge ranges were between 10 and 130 liters per second.

In order to reduce the turbulence of the input water flow from the pipe into the laboratory flume, the water travelled across a porous medium space created by sequential lattice plates placed along the first three meters of the laboratory flume. By placing the Styrofoam on the water surface oscillations were prevented and in this way the height of water was stabilized in the upstream of the GPK weir. The water depth was measured by

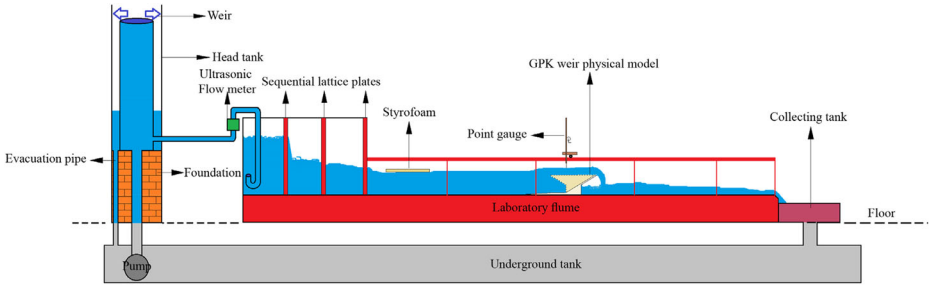


Fig. 2 The Schematic of laboratory flume and water circulation system

a point gauge moveable along the flume, with a precision of 0.1 mm. Due to the limitations imposed by surface tension forces the limitation of input discharge, the measured range of  $H$  is varied from 1.5–8 cm depending on the model type.

In this study, a four-cycle GPK weir was used. Figure 3 provides the tested physical models in this study. The variables in this study are: gate dimensions ( $b$  and  $d$ ), the different gate height ( $H_{gate}$ ), different ratios of the inlet key width to the outlet key width ( $W_i/W_o$ ), and the head over the GPK weir crest ( $H$ ). It can be noted that in addition of GPK weirs, un-gated PK weirs also were tested in this study.

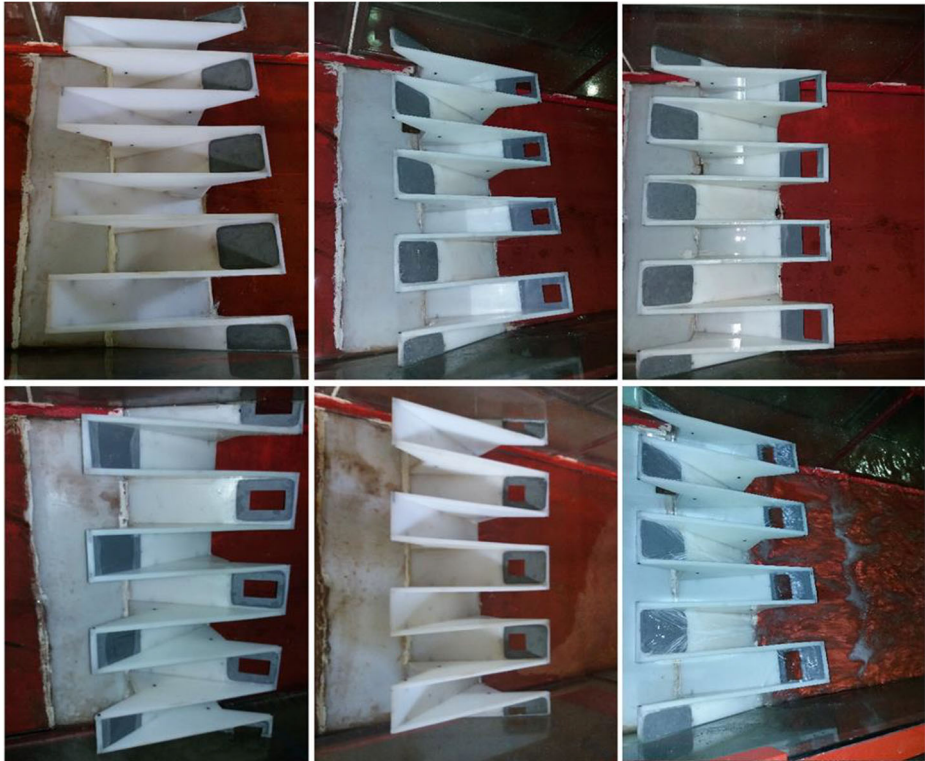


Fig. 3 Physical models of PK and GPK weirs with different geometric specifications

Tables 1 and 2 shows the PK and GPK weir geometric parameters in physical models. In Tables 1 and 2  $P_m$  is the weir wall height at the center of weir,  $B$  is the weir sidewall length and  $n$  is crest length to weir width ratio ( $n = L/W$ ). The other parameters were defined previously.

All physical models were built from polyethylene plates with 12 mm thickness. Gates frames were built from PVC with 1 mm thickness.

### 2.3 Experimental Data

As can be seen in Table 2, from 188 experiments for both PK and GPK weirs, 156 data set was relevant to the GPK weir. From this data, 70% and 30% were used as the training data and model testing data, respectively.

The statistical parameters of training and test data are given in Table 3. Training and test data were used for training process of models and evaluation of models accuracy, respectively.

### 2.4 Gaussian Process Regression (GPR)

The Gaussian Process is a suitable method to define the preferred distribution for the flexible models of regression and classification in which regression or class probability functions are not limited to the simple parametric forms.

One advantage of the Gaussian Process is the wide diversity of its covariance functions, which leads to functions with different degrees of smoothness, or various types of continuous structures, thereby allowing the scholar to choose appropriately from among them. These models can specify the distributions among the functions with one or more input variable. When this function defines the mean response in a regression model with Gaussian errors, the matrix calculations can be used for inference; this is feasible for those data sets with a sample of greater than one thousand. Gaussian processes are very important in statistical modeling, because they have normal characteristics (Neal 1997).

One can assume  $n$  observations in a desired dataset of  $Y = \{y_1, \dots, y_n\}$  as a single point sampled from the multiple Gaussian (It has  $n$  variables) distribution. Hence, data sets can be corresponded to a Gaussian process. The Gaussian process, therefore, is as simple as much as it is comprehensive. It is mostly assumed that the mean of correspondent the Gaussian process is zero everywhere. What connects one observation to another in such states is the covariance function,  $k(x, x)$ . Each observation  $y$  can be connected to a main function through the Gaussian noise model.

$$y = f(x) + N(0, \sigma_f^2) \tag{4}$$

**Table 1** The PK weir geometric parameters in physical models

$P$	$P_m$	$L$	$W$	$W_i$	$W_o$	$S_i$
0.197	0.129	4.744	0.933	0.116,0.093	0.093,0.116	0.006
$S_o$	$B$	$B_i$	$B_o$	$T_s$	$N$	$n$
0.006	0.489	0.121	0.121	0.012	4	5

**Table 2** Different geometric parameters of the GPK weir in physical model

<i>d</i>	<i>b</i>	<i>H<sub>gate</sub></i>	<i>w<sub>f</sub>/w<sub>o</sub></i>	<i>d</i>	<i>b</i>	<i>H<sub>gate</sub></i>	<i>w<sub>f</sub>/w<sub>o</sub></i>
0*	0	0	0.8	0	0	0	1.25
0.05	0.08	0.0272	0.8	0.05	0.08	0.0272	1.25
0.05	0.08	0.0348	0.8	0.05	0.08	0.0348	1.25
0.05	0.05	0.0201	0.8	0.05	0.05	0.0201	1.25
0.05	0.05	0.0442	0.8	0.05	0.05	0.0442	1.25
0.08	0.05	0.0201	0.8	0.08	0.05	0.0201	1.25
0.08	0.05	0.0442	0.8	0.08	0.05	0.0442	1.25

\* Note: zero values denote PK weir without gate

Where,  $N(0, \sigma_f^2)$  is the noise of normal distribution function with mean 0 and variance  $\sigma_f^2$ . Regression, in fact, means looking for  $f(x)$ . To simplify the next step, a new way for noise combination in  $k(x, x)$  is used by writing the following term:

$$k(x, x') = \sigma_f^2 \exp\left[\frac{-(x-x')^2}{2\tau^2}\right] + \sigma_n^2 \delta(x, x) \tag{5}$$

where,  $\delta(x, x)$  is the Kronecker delta function. Therefore,  $n$  observations of  $y$  is taken into account; the goal is to predict  $y^*$ .

The predicted values of observations are the same in accordance with Eq. 4, but variances vary upon due to the observational noise process. To prepare the GPR for covariance function, Eq. 5 is calculated among all possible combinations of these points and findings are summarized in three matrices:

$$k = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{bmatrix} \tag{6}$$

$$K^* = [k(X^*, X_1)k(X^*, X_2)\dots k(X^*, X_n)], k_{ij} = k(x_i, x_j) \tag{7}$$

It should also be noted that the diagonal elements,  $K$ , are in the form of  $\sigma_f^2 + \sigma_n^2$ . When the  $x$  receives a large domain, the non-diagonal elements approaches zero (Ebden 2015).

**Table 3** Statistic parameters of training phase and test dataset

Variable	Data range		Mean		Standard deviation	
	Training phase	Testing phase	Training phase	Testing phase	Training phase	Testing phase
<i>H/P</i>	0.051–0.391	0.083–0.376	0.233	0.211	0.090	0.075
<i>H<sub>gate</sub>/P</i>	0.097–0.215	0.097–0.215	0.151	0.156	0.051	0.047
<i>w<sub>f</sub>/w<sub>i</sub></i>	0.8–1.25	0.8–1.25	1.035	1.049	0.225	0.224
<i>b/P</i>	0.242–0.388	0.242–0.388	0.293	0.285	0.069	0.066
<i>d/P</i>	0.242–0.388	0.242–0.388	0.282	0.301	0.065	0.071
<i>C<sub>d</sub></i>	0.398–0.9	0.405–0.885	0.589	0.596	0.121	0.114

## 2.5 Multilayer Perceptron Model

The Multilayer Perceptron Neural networks (MLP-NN) are the feed-forward networks including one or more hidden layers. One way to train the widely used MLP-NN is the back propagation training rule, which is based on the error correction learning rule (moving in the negative direction of momentary slope subject to the performance (error) function which reduces the model error)(Dibike and Solomatine 2001; Negnevitsky 2005).

The back propagation rule is composed of two main paths. On the first path, the inlet vector is applied to the multilayer network, the effects of which are propagated through the middle layers to the output layers. The outlet vector formed on the output layer makes the real response of MLP-NN. On the second path, known as the backward path, the parameters of MLP-NN are modified and regulated. Such regulation is done in accordance with the error correction rule. Similarly, the weights of neurons in the middle layers change in such a way that the error value between the output of neural network and real output is minimized (Demuth et al. 2014; Goh 1995; Rafiq et al. 2001). Once the artificial neural network is developed, data is usually divided into two training and test periods. There is no exact rule for determining the minimum size of training and test sets. As per such suggestions, approximately 70% of total data is sufficient for network training and the remaining 30% is used for network testing (Baum and Haussler 1989; Zare et al. 2012). In this study, structure and architecture of ANN was optimum as  $1 \times 11 \times 1$  on the basis of trial and error conducted on the number of neurons in the middle layer. This is a three-layer network where 1, 11 and 1 indicate the number of neurons in input, hidden and output layers, respectively. The activation functions of this network are hyperbolic tangent sigmoid (tansig) and linear (purelin) functions, respectively, for hidden and output layers. The learning algorithm is the error back-propagation algorithm on the basis of Levenberg–Marquardt optimization method.

## 2.6 Support Vector Machines (SVM) Regression Model

Support Vector Machines (SVM) include two categories: 1. Support vector classifier; 2. Support vector regression. The SVM is a supervised learning technique which was introduced by Vapnik (1995) and is based on statistical learning theory (SLT)(Vapnik 2013). In some cases, complicated and nonlinear structures are needed for data separation. In this case, the main data is mapped and rearranged in a new space by the SVM through applying a set of mathematical functions, known as kernel. A SVM algorithm looks for a hyperplane with maximum margin. From the geometrical perspective, margin is calculated by the distance between the hyperplane and the closest training samples. The shortest distance from a hyperplane to sample with label +1 is equal to the shortest distance from that hyperplane to the samples with label -1. In fact, the margin is calculated by doubling this distance. A separating hyperplane can be defined as below:

$$W.X + b = 0 \quad (8)$$

where,  $W = \{w_1, \dots, w_n\}$  is a vector in which the number of existing elements is equal to the attributes and is a constant. In 2D space where the data sets are defined with two attributes and one class label, Eq. 8 is rewritten as Eq. 9, assuming the  $w_0 = b$ :



$$w_0 + w_1x_1 + w_2x_2 = 0 \quad (9)$$

Accordingly, the samples (points) located on the space above this hyperplane complete the inequality (10) and the ones below this hyperplane complete the inequality (11).

$$w_0 + w_1x_1 + w_2x_2 > 0 \quad (10)$$

$$w_0 + w_1x_1 + w_2x_2 < 0 \quad (11)$$

Adjusting the  $W$  and  $b$ , we have:

$$w_0 + w_1x_1 + w_2x_2 \geq 1 \text{ if } y_i = +1 \quad (12)$$

$$w_0 + w_1x_1 + w_2x_2 \leq -1 \text{ if } y_i = -1 \quad (13)$$

This means that each sample located over or on the hyperplane  $H_1$  belongs to the class +1 and the one below and on the hyperplane  $H_2$  belongs to the class -1. Those samples which are exactly on the hyperplane  $H_1$  and  $H_2$ , are called “support vectors”(Burges 1998).

### 2.6.1 Kernel Functions

A common way to solve the nonlinear problems is to use the kernel functions; these functions are defined based on the inner product of given data. Designing the GPR techniques includes application of kernel function concept. In fact, with a nonlinear transformation from inner space to the attribute space with more dimensions (even infinite), problems can be separated linearly. By transforming the samples from the input space into the attribute space, the nonlinear separator will become linear.

$$w^T \phi(x_j) + b = 0 \quad (14)$$

Primary and dual problems create the attribute space problem. With this difference: that instead of  $(x_i^T \cdot x_j)$  the value of  $x_j = \phi^T(x_i) \phi(x_j)$  and  $k(x_i, x_j)$  is used, where,  $k(x_i, x_j)$  is the kernel function for linearization of nonlinear problems. Among the most important kernels, linear, polynomial, normalized polynomial, Radial Basis Function (RBF) kernel and Pearson kernel function can be mentioned.

Simple polynomial kernel function:

$$k(x_i, x_j) = (x_i^T x_j + 1)^p \quad (15)$$

Normalized polynomial kernel function:

$$k(x_i, x_j) = \frac{(x_i^T x_j + 1)^p}{\sqrt{(x_i^T x_i) + (x_j^T x_j)}} \quad (16)$$

Radial basis function:

$$k(x_i, x_j) = \exp(-Y|x_i - x_j|^2) \quad (17)$$

Pearson kernel function:

$$k(x_i, x_j) = \frac{1}{\left[1 + \left(2\sqrt{x_i, x_j^2} \sqrt{2^{\frac{1}{w}} - \frac{1}{\emptyset}}\right)^2\right]^w} \quad (18)$$

These kernel functions have individual parameters in their structures, known as the hyper-parameters  $\lambda$ . The radial basis function, for instance, has Gamma hyper-parameter and Pearson function has the Sigma and Omega parameters. In addition, to select the optimum kernel function in applying the kernel function based methods, it is very important to determine the optimum hyper-parameter related to each function. The kernel based modeling techniques needs to create the appropriate parameters defined by the user, since the accuracy of this regression model is highly dependent on selection of these parameters. In addition to selecting the specific kernel parameters, GPR needs to specify the optimum of Gaussian noise level. To select the parameters chosen by the user (i.e.  $C$ ,  $\gamma$ ,  $\sigma$ ,  $\omega$ ,  $\varepsilon$  and Gaussian noise), numerous techniques have been proposed; manual method (trial and error), network search method, Genetic Algorithm (GA) and particle swarm optimization. In the present study, the trial and error is used for selecting the parameters chosen by the user. The optimums of various parameters defined by the user are selected in such a way that they minimize the root-mean-square error (RMSE) and maximize the correlation coefficient (Pal and Deswal 2010).

## 2.7 Generalized Regression Neural Networks Model

The Generalized Regression Neural Networks (GRNN) are a set of Radial basis function networks together with a linear layer (Chen et al. 1991). GRNN model is based on a statistically standard technique, known as kernel regression (Cigizoglu and Alp 2006; Li et al. 2013). These networks, include 4 layers: input layer, pattern layer, summation layer and output layer (Cigizoglu and Alp 2006). The input layer receives the information. The number of neurons is equal to the input vector dimension. Then, the inlet neuron from the input layer transfers the data to the pattern layer which has nonlinear transformation from the input space to the pattern space. Neurons in the pattern layer (which are also called the “pattern neurons”) can keep the relationship between the input neuron and the proper response from the pattern layer and also the number of neurons equal to the input variables. The summation layer includes a simple summation and a weighted summation. The former calculates the output arithmetic sum from the pattern layer and the connection weight is 1. The latter calculates the output weighting sum from the pattern layer. Once the summated neurons are transferred to the output layer, output of the GRNN model can be measured. In these networks, the number of neurons in the output layer is equal to the output vector dimension (Li et al. 2013). Unlike the back propagation training algorithm, GRNNs do not need frequent training processes (Chen et al. 1991; Cigizoglu and Alp 2006). These networks take a radius into account around each observational data where each input data within that radius makes that data involved in its estimation for a new input value. The effect of the radius of GRNN used in this study was measured based on trial and error.

## 2.8 Statistical Criteria

To investigate the accuracy of models suggested in this study, the following statistical measures have been used:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{N}} \quad (19)$$

$$NRMSE = \frac{RMSE}{\bar{O}_i} \times 100 \quad (20)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{O_i - P_i}{O_i} \right| \times 100 \quad (21)$$

$$R^2 = \frac{\left( \sum_{i=1}^N (O_i - \bar{O}_i) (P_i - \bar{P}_i) \right)^2}{\sum_{i=1}^N (O_i - \bar{O}_i)^2 \sum_{i=1}^N (P_i - \bar{P}_i)^2} \quad (22)$$

In these equations,  $\bar{O}_i$  is the mean of observations,  $\bar{P}_i$  is the mean of predictions and  $N$  is the total number of data. Root Mean Square Error (RMSE) is the difference between the predicted and observed data. Mean Absolute Percentage Error (MAPE) shows the accuracy of model prediction.  $R^2$  describes the connection between the predicted and observed data.

Normalized Root Mean Square Error (NRMSE) is used for comparing the models with various measures. Model performance with regard to NRMSE is defined as below (Mihoub et al. 2016):

Excellent if:  $NRMSE < 10\%$ .

Good if:  $10\% < NRMSE < 20\%$ .

Fair if:  $20\% < NRMSE < 30\%$ .

Poor if:  $NRMSE > 30\%$ .

## 3 Results and Discussion

In this study, the results of artificial intelligence models, MLP, GPR, GRNN, SVM, together with two linear and nonlinear regression models, were investigated and assessed in terms of estimating the GPK weir discharge coefficient. The results of the various models are given below.

### 3.1 Multi-Layer Perceptron (MLP) Neural Network

A three-layer artificial neural network model (with a hidden layer) was used in this study. The input and output activation functions were selected as sigmoid and linear. Due to higher accuracy and speed, the Levenberg-Marquardt algorithm was used for network training

(Lourakis and Argyros 2004; Ngia and Sjoberg 2000). The most important part in modeling using the MLP neural network was to determine the number of optimal neurons in the hidden layer. For this purpose, a trial and error method was applied. The number of neurons in the hidden layer was changed from 1 to 10 and the best number of neurons was chosen using the RMSE for the test data of model. Table 4 shows the modeling results by making use of the MLP neural network. From these results it can be determined that the MLP neural network with two neurons in the hidden layer has yielded the best outcome with  $RMSE = 0.024$  and  $R^2 = 0.961$ . Based on Table 4, it can be seen that as the number of neurons increases, the model accuracy decreases.

### 3.2 Gaussian Process Regression (GPR) Model

Table 5 shows the results of the GPR model in estimating the GPK weir discharge coefficient. To investigate the effect of different kernels on the model accuracy, five types of kernels, i.e., Squared Exponential, Exponential, Matern 3/2, Matern 5/2 and Rational quadratic, were evaluated. The GPR model with the Squared Exponential kernel has yielded the best results concerning  $RMSE = 0.011$ ,  $R^2 = 0.992$  and  $MAPE = 1.167\%$ . Comparing different kinds of kernels (Squared Exponential, Exponential and etc.) indicated that application of different kernels had minor impact on model performance and varied from  $RMSE = 0.018$  for the Rational quadratic kernel to  $RMSE = 0.015$  for the Exponential kernel. Results revealed that the GPR model was not sensitive to the kernel changes. Figure 4 shows the Scatter plots of observed and predicted values of  $C_d$  for GPR model with different kernels. Closeness of the data to line 1:1 shows the appropriate accuracy of models in estimating the GPK weir discharge coefficient.

### 3.3 Support Vector Machine (SVM) Model

Table 6 shows the modeling results obtained from the SVM model. In this study, RBF, Linear and polynomial kernels were used for modeling. To find the optimum kernel parameters ( $\epsilon$  and  $\gamma$ ), the Bayesian optimization algorithm in MATLAB was used. The SVM model with the RBF kernel gave the best result with  $RMSE = 0.015$ ,  $R^2 = 0.982$  and  $MAPE = 1.961\%$ . Our results were consistent with ones obtained by Zaji et al. (2016). Comparing two linear and

**Table 4** Results of MLP model with different number of neuron at hidden layer

Model	Neurons	Training phase				Testing phase			
		RMSE	R <sup>2</sup>	MAPE%	NRMSE%	RMSE	R <sup>2</sup>	MAPE%	NRMSE%
MLP	1	0.033	0.970	4.800	5.677	0.031	0.934	4.568	5.202
<b>MLP</b>	<b>2</b>	<b>0.028</b>	<b>0.979</b>	<b>3.953</b>	<b>4.752</b>	<b>0.024</b>	<b>0.961</b>	<b>3.441</b>	<b>4.031</b>
MLP	3	0.009	0.998	0.919	1.601	0.080	0.833	4.362	13.334
MLP	4	0.009	0.998	0.881	1.502	0.055	0.889	3.266	9.202
MLP	5	0.008	0.998	0.834	1.419	0.112	0.756	5.664	18.805
MLP	6	0.016	0.994	1.660	2.655	0.030	0.937	2.945	4.947
MLP	7	0.008	0.999	0.767	1.282	0.032	0.941	2.276	5.436
MLP	8	0.009	0.998	0.816	1.451	0.041	0.921	2.667	6.865
MLP	9	0.010	0.998	0.903	1.651	0.300	0.391	10.890	50.221
MLP	10	0.008	0.998	0.768	1.406	0.082	0.818	4.664	13.746

Note: Bold numbers refer to the best models

**Table 5** Results of GPR model with different kernel types

Model	Kernel	Training phase				Testing phase			
		RMSE	R <sup>2</sup>	MAPE%	NRMSE%	RMSE	R <sup>2</sup>	MAPE%	NRMSE%
<b>GPR</b>	<b>SqExponential</b>	<b>0.009</b>	<b>0.998</b>	<b>0.787</b>	<b>1.558</b>	<b>0.011</b>	<b>0.992</b>	<b>1.167</b>	<b>1.805</b>
GPR	Exponential	0.000	1.000	0.006	0.014	0.015	0.983	1.558	2.496
GPR	Matern 3/2	0.001	1.000	0.123	0.187	0.015	0.984	1.538	2.448
GPR	Matern 5/2	0.004	1.000	0.408	0.648	0.017	0.979	1.753	2.849
GPR	Rational quad	0.005	0.999	0.533	0.888	0.018	0.976	1.894	3.065

Note: Bold numbers refer to the best models

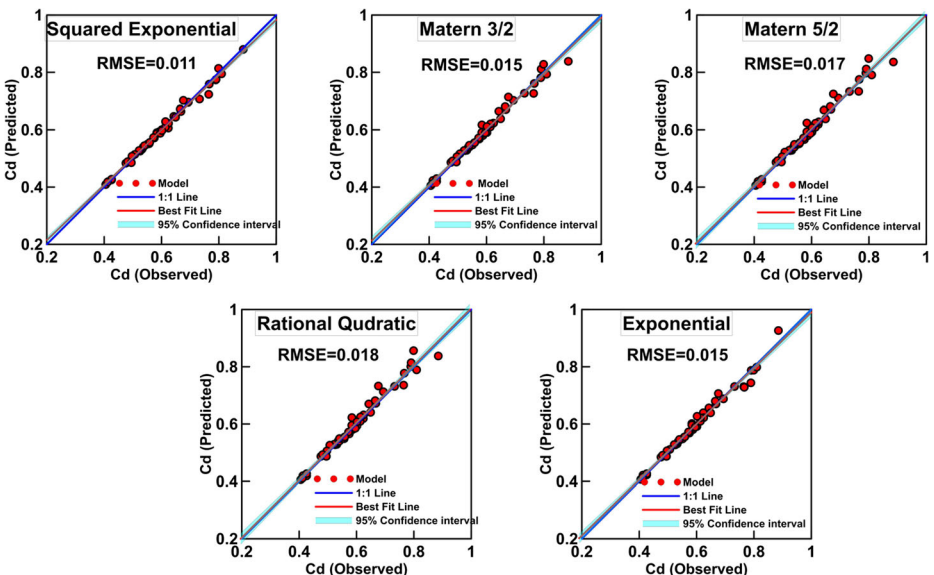
polynomial kernels indicates the superiority of the SVM model with the linear kernel (RMSE = 0.042) over the polynomial model (RMSE = 0.057). Figure 5 shows Scatter plots of observed and predicted values of  $C_d$  for SVM model with different kernels.

### 3.4 Generalized Regression Neural Networks (GRNN) Model

Table 7 shows the modeling results of GRNN in estimating the GPK weir discharge coefficient. The optimal model was obtained by changing the Spread parameter in GRNN. In view of Table 7, the Spread parameter equal 0.001 yielded the best results for with RMSE = 0.073, R<sup>2</sup> = 0.651 with further reduction of this parameter having no significant impact on results.

### 3.5 Multiple Linear Regression

Analysis of multiple linear regression was conducted on the experimental data. The Dimensionless eq. (22) was obtained for estimating the discharge coefficient of the GPK weir. Analyses were made in MATLAB.



**Fig. 4** Scatter plots of observed and predicted  $C_d$  for GPR model with different kernels

**Table 6** Results of SVM models with different kernels and optimized parameters

Model	Kernel	Training phase				Testing phase					
		$\epsilon$	$\gamma$	RMSE	R <sup>2</sup>	MAPE%	NRMSE%	RMSE	R <sup>2</sup>	MAPE%	NRMSE%
SVM	<b>RBF</b>	<b>0.015</b>	<b>0.273</b>	<b>0.012</b>	<b>0.996</b>	<b>1.926</b>	<b>2.109</b>	<b>0.015</b>	<b>0.982</b>	<b>1.961</b>	<b>2.566</b>
SVM	Line	0.091	0.001	0.117	0.708	6.517	19.846	0.042	0.879	4.692	6.983
SVM	Poly	0.002	0.472	0.096	0.820	3.752	16.328	0.057	0.779	4.634	9.591

Note: Bold numbers refer to the best models

$$C_d = 0.672 - 1.618 \left(\frac{H}{P}\right) + 0.23 \left(\frac{b}{P}\right) + 0.056 \left(\frac{d}{P}\right) + 0.534 \left(\frac{H_{gate}}{P}\right) + 0.125 \left(\frac{w_i}{w_0}\right) \quad (23)$$

### 3.6 Multiple Nonlinear Regression

To evaluate the nonlinear regression in estimation of the discharge coefficient of the GPK weir, an equation within the general form (23) was evaluated:

$$C_d = a' \left(\frac{H}{P}\right)^{b'} \left(\frac{b}{P}\right)^{c'} \left(\frac{d}{P}\right)^{d'} \left(\frac{H_{gate}}{P}\right)^{e'} \left(\frac{w_i}{w_0}\right)^{f'} \quad (24)$$

Coefficients of  $a', b', c', d', e', f'$  were determined using MATLAB:

$$C_d = 0.375 \left(\frac{H}{P}\right)^{-0.53} \left(\frac{b}{P}\right)^{0.096} \left(\frac{d}{P}\right)^{0.012} \left(\frac{H_{gate}}{P}\right)^{0.148} \left(\frac{w_i}{w_0}\right)^{0.276} \quad (25)$$

### 3.7 Comparing the Techniques of Artificial Intelligence and Regression Analysis

Table 8 compares the best model built in each method by artificial intelligence with linear and nonlinear regression analyses. In Table 8, the statistical parameters of RMSE, R<sup>2</sup>, MAPE and NRMSE were used for training and testing datasets. As seen in Table 8, three artificial intelligence models, GPR, SVM and MLP, outperformed the regression models and only the GRNN model did not give acceptable results. Comparison between the GPR, SVM and MLP models indicated that the GRP model yielded the best results with RMSE = 0.011, R<sup>2</sup> = 0.992 and MAPE = 1.167%, followed by the SVM model with RBF kernel with RMSE = 0.015, R<sup>2</sup> = 0.982 and MAPE = 1.961%. The MLP model with two neurons in the hidden layer is ranked as the third with RMSE = 0.024, R<sup>2</sup> = 0.961 and MAPE = 3.441%. The GRNN model with

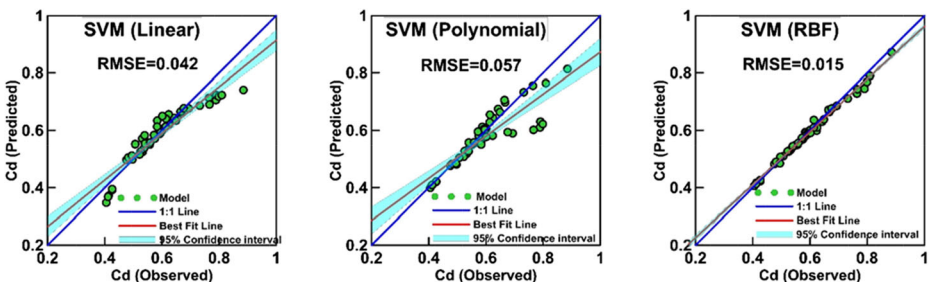


Fig. 5 Scatter plots of observed and predicted  $C_d$  for SVM model with different kernels

**Table 7** Results of GRNN model with different spread parameter values

Model	Spread	Training phase				Testing phase			
		RMSE	R <sup>2</sup>	MAPE%	NRMSE%	RMSE	R <sup>2</sup>	MAPE%	NRMSE%
GRNN	1	0.190	0.401	20.103	32.181	0.111	0.154	14.608	18.651
GRNN	0.5	0.177	0.405	18.031	30.043	0.106	0.152	14.027	17.811
GRNN	0.1	0.097	0.842	6.413	16.411	0.078	0.547	8.048	13.090
GRNN	0.01	0.008	0.998	0.359	1.413	0.073	0.650	5.589	12.262
<b>GRNN</b>	<b>0.001</b>	<b>0.000</b>	<b>1.000</b>	<b>0.001</b>	<b>0.002</b>	<b>0.073</b>	<b>0.651</b>	<b>5.757</b>	<b>12.304</b>
<b>GRNN</b>	<b>0.0001</b>	<b>0.000</b>	<b>1.000</b>	<b>0.001</b>	<b>0.002</b>	<b>0.073</b>	<b>0.651</b>	<b>5.757</b>	<b>12.304</b>

Note: Bold numbers refer to the best models

RMSE = 0.073 was ranked sixth and yielded weaker results than the linear regression. Comparison between the two regression models indicated the superiority of nonlinear regression model over the linear one. With RMSE = 0.035, R<sup>2</sup> = 0.961 and MAPE = 4.562%, nonlinear regression gave acceptable results for estimating the discharge coefficient of GPK weir. In the linear regression model, values of RMSE = 0.061, R<sup>2</sup> = 0.865 and MAPE = 9.491% were obtained. Figure 6 shows the Scatter plots of observed and predicted values of C<sub>d</sub> for all of the six models compared in Table 8. As can be seen in Fig. 6, the data for GPR model is closer to line 1:1.

Figure 7 shows the points related to the observational and estimated data together with a confidence interval (CI) 95% for the GPR model. With regard to this figure, most of the observed data points are within the CI 95%.

### 3.8 Sensitivity Analysis

In order to evaluate the impact of input parameters on estimated discharge coefficient of the GPK weir, the sensitivity analysis was conducted using the GRP model (given its less error). In this analysis, each time a parameter was omitted from the model inputs, the model was implemented and accuracy was evaluated. The deleted parameter with the highest effect on decreased model accuracy and increased model error was evaluated as the most important parameter. In view of the analysis results, seen in Table 9, deleting the parameter H/P had the highest effect on accuracy and caused the RMSE to increase to 0.128. The error percentage achieved MAPE = 16.621%. Parameter w<sub>o</sub>/w<sub>i</sub> was the second most effective parameter on the discharge coefficient of the GPK weir. With its deletion RMSE increased to 0.079.

Comparison of the parameters related with gate ( $\frac{H_{gate}}{P}, \frac{b}{P}, \frac{d}{P}$ ) showed that  $\frac{H_{gate}}{P}$  (RMSE = 0.032) affected more than two parameters of  $\frac{b}{P}$  (RMSE = 0.029) and  $\frac{d}{P}$  (RMSE = 0.027). Given the statistical measures, two parameters of  $\frac{b}{P}$  and  $\frac{d}{P}$  had the same effect on model accuracy.

**Table 8** Comparison of best models

Model	Training phase				Testing phase			
	RMSE	R <sup>2</sup>	MAPE%	NRMSE%	RMSE	R <sup>2</sup>	MAPE%	NRMSE%
LM	0.104	0.717	10.267	17.566	0.061	0.865	9.491	10.283
NLM	0.038	0.963	4.336	6.440	0.035	0.961	4.562	5.882
<b>GPR</b>	<b>0.009</b>	<b>0.998</b>	<b>0.787</b>	<b>1.558</b>	<b>0.011</b>	<b>0.992</b>	<b>1.167</b>	<b>1.805</b>
SVM	0.012	0.996	1.926	2.109	0.015	0.982	1.961	2.566
MLP	0.028	0.979	3.953	4.752	0.024	0.961	3.441	4.031
GRNN	0.000	1.000	0.001	0.002	0.073	0.651	5.757	12.304

Note: Bold numbers refer to the best models

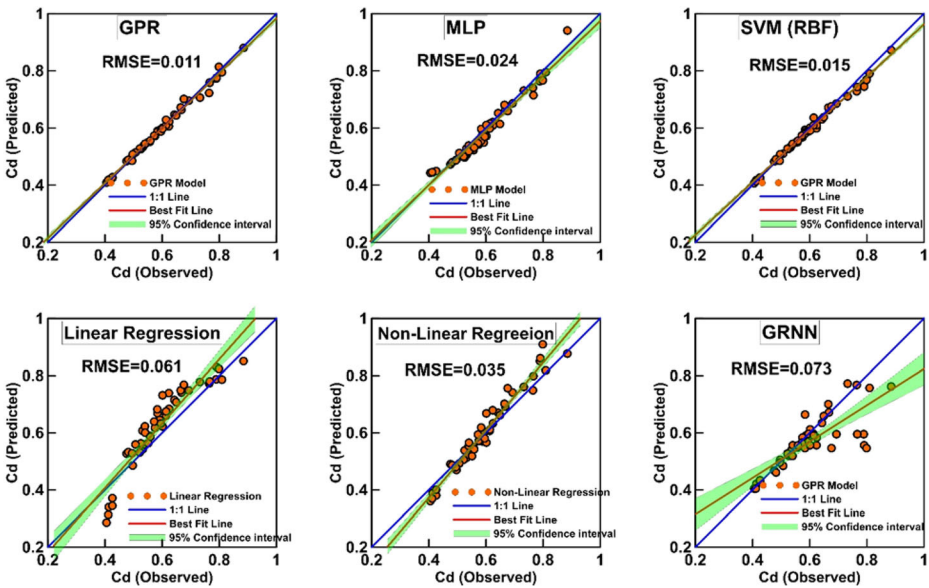


Fig. 6 Scatter plots of observed and predicted values of  $C_d$  for the six methods

### 4 Conclusion

Designing weirs with high discharge capacity is important in order to control flood flows entering the rivers and reservoirs of dams and maintaining safe levels. The GPK weir was proposed as a new idea in order to increase the capacity of weir discharge and improve its performance. In this study, effective parameters on the GPK weir discharge were investigated using soft computing methods MLP, GPR, SVM, GRNN, linear and non-linear regressions.

For this purpose, data from 156 laboratory tests was collected and processed in MATLAB software 2018. The results of this study showed that  $H$ ,  $W_f/W_o$ ,  $H_{gate}$ ,  $b$  and  $d$ , (defined in fig. 1),

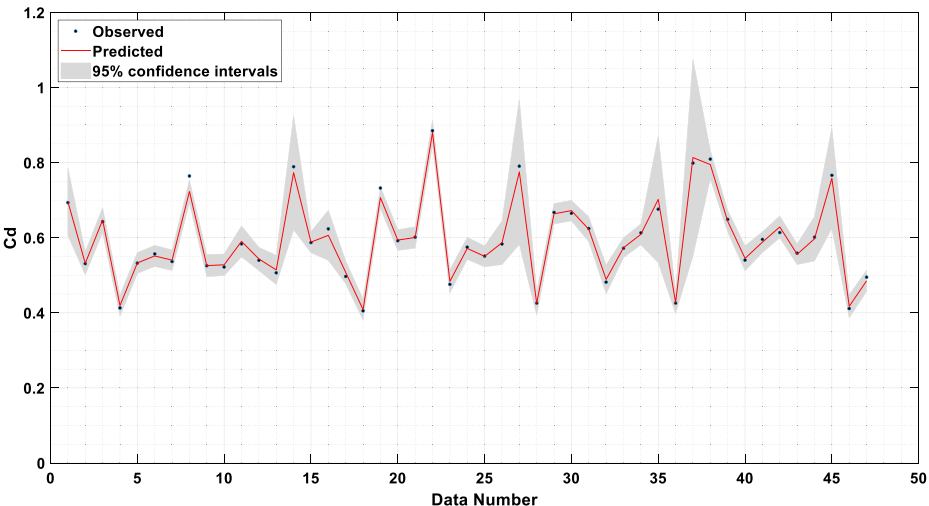


Fig. 7 Point and interval predictions of  $C_d$  by GPR model



**Table 9** Sensitivity analysis of input variables

Model	Omitted variable	Rank	Training phase				Testing phase			
			RMSE	R <sup>2</sup>	MAPE%	NRMSE%	RMSE	R <sup>2</sup>	MAPE%	NRMSE%
GPR	$H/P$	1	0.165	0.280	16.848	21.518	0.128	0.083	16.621	27.997
GPR	$H_{gate}/P$	3	0.027	0.981	1.894	5.312	0.032	0.933	2.766	4.574
GPR	$w_i/w_o$	2	0.033	0.971	4.743	13.215	0.079	0.787	8.163	5.629
GPR	$b/P$	4	0.022	0.987	1.405	4.783	0.029	0.943	2.370	3.716
GPR	$d/P$	5	0.008	0.998	0.783	4.452	0.027	0.952	1.851	1.416

had the greatest effect on the GPK weir discharge coefficient, respectively. Three artificial intelligence models, GPR, SVM and MLP, outperformed the regression models and only the GRNN model did not give acceptable results. Comparison between the GPR, SVM and MLP models indicated that the GRP model yielded the best results, followed by the SVM model with RBF kernel. The MLP model with two neurons in the hidden layer was ranked third. The GRNN model was ranked sixth and yielded weaker results than the linear regression. Comparison between the two regression models indicated the superiority of the nonlinear regression model over the linear one. According to the results of artificial intelligence techniques in this study, the GPR method was found to be a new and accurate method for predicting the discharge coefficient of weirs.

## Compliance with Ethical Standards

**Conflict of Interest** None.

**Notation** The following symbols are used in this paper:  $b$ : Length of rectangular gate;  $B$ : Weir sidewall length;  $d$ : Width of rectangular gate;  $B_d$ : Downstream or inlet key overhang length;  $B_o$ : Upstream or outlet key overhang length;  $C_d$ : Dimensionless discharge coefficient;  $g$ : Gravitational acceleration; GPK weir: Gated piano key weir;  $H$ : Head over the crest;  $H_{down}$ : Downstream head;  $H_{gate}$ : Water head from the rectangular gate center to the crest of weir;  $H_{up}$ : Upstream head;  $L$ : Crest length; [ $L = N(W_i + W_o + 2B)$ ].  $n$ : Crest length to weir width ratio; ( $n = L/W$ ).  $N$ : Weir cycle number;  $P$ : Total weir height;  $P_m$ : Weir wall height at the center of weir; PK weir: Piano key weir;  $Q$ : Discharge;  $S_i$ : Inlet key slope;  $S_o$ : Outlet key slope;  $T_s$ : Weir wall thickness;  $W$ : Total weir width or flume width;  $W_i$ : Inlet key width;  $W_o$ : Outlet key width;  $\mu$ : Water dynamic viscosity;  $\rho$ : Water density and  $\sigma$ : Water surface tension.

## References

- Anderson R, Tullis B (2012) Piano key weir hydraulics and labyrinth weir comparison. *J Irrig Drain Eng* 139: 246–253. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000530](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000530)
- Baum EB, Haussler D (1989) What size net gives valid generalization? In: *Advances in neural information processing systems*, pp 81–90
- Bilhan O, Emiroglu ME, Kisi O (2011) Use of artificial neural networks for prediction of discharge coefficient of triangular labyrinth side weir in curved channels. *Adv Eng Softw* 42:208–214. <https://doi.org/10.1016/j.advengsoft.2011.02.006>
- Burges CJC (1998) A tutorial on support vector machines for pattern recognition. *Data Min Knowl Disc* 2:121–167. <https://doi.org/10.1023/a:1009715923555>
- Chen S, Cowan CF, Grant PM (1991) Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Trans Neural Netw* 2:302–309. <https://doi.org/10.1109/72.80341>
- Cigizoglu HK, Alp M (2006) Generalized regression neural network in modelling river sediment yield. *Adv Eng Softw* 37:63–68. <https://doi.org/10.1016/j.advengsoft.2005.05.002>
- Demuth HB, Beale MH, De Jess O, Hagan MT (2014) *Neural network design*. Martin Hagan

- Dibike YB, Solomatine DP (2001) River flow forecasting using artificial neural networks. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* 26:1–7. [https://doi.org/10.1016/s1464-1909\(01\)85005-x](https://doi.org/10.1016/s1464-1909(01)85005-x)
- Dursun OF, Kaya N, Firat M (2012) Estimating discharge coefficient of semi-elliptical side weir using ANFIS. *J Hydrol* 426–427:55–62. <https://doi.org/10.1016/j.jhydrol.2012.01.010>
- Ebden M (2015) Gaussian processes: A quick introduction. arXiv preprint arXiv:150502965
- Ebtehaj I, Bonakdari H, Zaji AH, Azimi H, Sharifi A (2015) Gene expression programming to predict the discharge coefficient in rectangular side weirs. *Appl Soft Comput* 35:618–628. <https://doi.org/10.1016/j.asoc.2015.07.003>
- Emiroglu ME, Bilhan O, Kisi O (2011) Neural networks for estimation of discharge capacity of triangular labyrinth side-weir located on a straight channel. *Expert Syst Appl* 38:867–874. <https://doi.org/10.1016/j.eswa.2010.07.058>
- Goh AT (1995) Back-propagation neural networks for modeling complex systems. *Artif Intell Eng* 9:143–151. [https://doi.org/10.1016/0954-1810\(94\)00011-s](https://doi.org/10.1016/0954-1810(94)00011-s)
- Grbić R, Kurtagić D, Slišković D (2013) Stream water temperature prediction based on Gaussian process regression. *Expert Syst Appl* 40:7407–7414. <https://doi.org/10.1016/j.eswa.2013.06.077>
- Haghiabi AH, Parsaie A, Ememgholizadeh S (2018) Prediction of discharge coefficient of triangular labyrinth weirs using adaptive neuro fuzzy inference system. *Alexandria Engineering Journal* 57:1773–1782. <https://doi.org/10.1016/j.aej.2017.05.005>
- Henderson FM (1966) Open channel flow
- Kabiri-Samani A, Javaheri A (2012) Discharge coefficients for free and submerged flow over piano key weirs. *J Hydraul Res* 50:114–120. <https://doi.org/10.1080/00221686.2011.647888>
- Karbasi M (2017) Forecasting of multi-step ahead reference evapotranspiration using wavelet- Gaussian process regression model. *Water Resour Manag* 32:1035–1052. <https://doi.org/10.1007/s11269-017-1853-9>
- Laugier F (2007) Design and construction of the first piano key weir spillway at Goulours dam. *International journal on hydropower and dams* 14:94
- Laugier F, Lochu A, Gille C, Leite Ribeiro M, Boillat J-L (2009) Design and construction of a labyrinth PKW spillway at saint-Marc dam, France. *International journal on hydropower and dams* 16:100–107
- Lempérière F (2011) New labyrinth weirs triple the spillways discharge. *Water and Energy International* 68:77–78
- Lempérière F, Ouamane A (2003) The piano keys weir: a new cost-effective solution for spillways. *International Journal on Hydropower and Dams* 10:144–149
- Li H-Z, Guo S, Li C-J, Sun J-Q (2013) A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowl-Based Syst* 37:378–387. <https://doi.org/10.1016/j.knosys.2012.08.015>
- Lourakis M, Argyros A (2004) The design and implementation of a generic sparse bundle adjustment software package based on the levenberg-marquardt algorithm. Technical Report 340, Institute of Computer Science-FORTH, Heraklion, Crete, Greece
- Machiels O, Ericum S, Archambeau P, Dewals B, Piroton M (2013) Parapet Wall effect on piano key weir efficiency. *J Irrig Drain Eng* 139:506–511. [https://doi.org/10.1061/\(asce\)ir.1943-4774.0000566](https://doi.org/10.1061/(asce)ir.1943-4774.0000566)
- Mihoub R, Chabou N, Guermoui M (2016) Modeling soil temperature based on Gaussian process regression in a semi-arid-climate, case study Ghardaia, Algeria. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources* 2:397–403. <https://doi.org/10.1007/s40948-016-0033-3>
- Neal RM (1997) Monte Carlo implementation of Gaussian process models for Bayesian regression and classification. arXiv preprint physics/9701026
- Negnevitsky M (2005) Artificial intelligence: a guide to intelligent systems. Pearson Education
- Ngia LS, Sjoberg J (2000) Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm. *IEEE Trans Signal Process* 48:1915–1927. <https://doi.org/10.1109/78.847778>
- Pal M, Deswal S (2010) Modelling pile capacity using Gaussian process regression. *Comput Geotech* 37:942–947. <https://doi.org/10.1016/j.compgeo.2010.07.012>
- Parsaie A (2016) Predictive modeling the side weir discharge coefficient using neural network. *Modeling Earth Systems and Environment* 2. <https://doi.org/10.1007/s40808-016-0123-9>
- Pasolli L, Melgani F, Blanzieri E (2010) Gaussian process regression for estimating chlorophyll concentration in subsurface waters from remote sensing data. *IEEE Geosci Remote Sens Lett* 7:464–468. <https://doi.org/10.1109/LGRS.2009.2039191>
- Rafiq M, Bugmann G, Easterbrook D (2001) Neural network design for engineering applications. *Comput Struct* 79:1541–1552. [https://doi.org/10.1016/S0045-7949\(01\)00039-6](https://doi.org/10.1016/S0045-7949(01)00039-6)
- Salmasi F, Yıldırım G, Masoodi A, Parsamehr P (2012) Predicting discharge coefficient of compound broad-crested weir by using genetic programming (GP) and artificial neural network (ANN) techniques. *Arab J Geosci* 6:2709–2717. <https://doi.org/10.1007/s12517-012-0540-7>

- Shamshirband S, Bonakdari H, Zaji AH, Petkovic D, Motamedi S (2016) Improved side weir discharge coefficient modeling by adaptive neuro-fuzzy methodology. *KSCE J Civ Eng* 20:2999–3005. <https://doi.org/10.1007/s12205-016-1723-7>
- Vapnik V (2013) *The nature of statistical learning theory*. Springer science & business media
- Wang Y, Chaib-draa B (2017) An online Bayesian filtering framework for Gaussian process regression: application to global surface temperature analysis. *Expert Syst Appl* 67:285–295. <https://doi.org/10.1016/j.eswa.2016.09.018>
- Zaji AH, Bonakdari H, Shamshirband S (2016) Support vector regression for modified oblique side weirs discharge coefficient prediction. *Flow Meas Instrum* 51:1–7. <https://doi.org/10.1016/j.flowmeasinst.2016.08.006>
- Zare M, Pourghasemi HR, Vafakhah M, Pradhan B (2012) Landslide susceptibility mapping at Vaz watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arab J Geosci* 6:2873–2888. <https://doi.org/10.1007/s12517-012-0610-x>

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

## Affiliations

Masood Akbari<sup>1</sup> • Farzin Salmasi<sup>1</sup> • Hadi Arvanaghi<sup>1</sup> • Masoud Karbasi<sup>2</sup> • Davood Farsadizadeh<sup>1</sup>

Masood Akbari  
masood.akbari@gmail.com

Hadi Arvanaghi  
arvanaghi@tabrizu.ac.ir

Masoud Karbasi  
m.karbasi@znu.ac.ir

Davood Farsadizadeh  
farsadi@tabrizu.ac.ir

<sup>1</sup> Department of Water Engineering, University of Tabriz, P.O. Box: 5166616471, Tabriz, Iran

<sup>2</sup> Department of Water Engineering, University of Zanjan, Zanjan, Iran