An efficient hybrid grey wolf optimization-based KELM approach for prediction of the discharge coefficient of submerged radial gates

Kiyoumars Roushangar1,2 · Saman Shahnazi1 · Arman Alirezazadeh Sadaghiani1

Accepted: 27 October 2022 / Published online: 15 November 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
Accurately prediction of discharge coefficient through radial gates is considered as a challenging hydraulic subject, particularly under highly submerged flow conditions. Incurring the advantages of kernel-depend extreme learning machine (KELM), this study offers a grey wolf optimization-based KELM (GWO-KELM) for effective prediction of discharge coefficient through submerged radial gates. Additionally, support vector machine (SVM) and Gaussian process regression (GPR) methods are also presented for comparative purposes. To build prediction models using GWO-KELM, GPR, and SVM, an extensive experimental database was established, consisting of 2125 data samples gathered by the US Bureau of Reclamation. From simulation results, it is observed that the proposed GWO-KELM approach with radial basis function and input parameters of the ratio of the downstream flow depth to the gate opening and submergence ratio provides the best performance with the correlation coefficient ($R$) of 0.983, the determination coefficient (DC) of 0.966 and the root-mean-squared error (RMSE) of 0.027. The obtained results showed that the proposed GWO-KELM with RBF kernel function gives better prediction accuracy than employed GPR and SVM approaches. Furthermore, the obtained results showed that the employed kernel-depend methods are capable of a statistically predicting the discharge coefficient under varied submergence conditions with satisfactory level of accuracy. Among them, proposed hybrid GWO-KELM method gave the most accurate results ($R = 0.873$, DC = 0.744, and RMSE = 0.035) for extremely highly submerged flow. Moreover, the results reflected that the employed kernel-depend methods give better predictions than the developed dimensionless formulas.

Keywords Radial gates · Discharge coefficient · Kernel-depend extreme learning machine · Grey wolf optimization · Gaussian process regression · Support vector machine

1 Introduction
Radial gates are known as the most common elements of water control structures in use today and can be found in most irrigation networks overall the world. They are
constructed by the use of a curved skin plate with a structural steel frame and play an essential role in facilitating water delivery, maintaining river environments, and controlling flooding. These types of gates are more cost-effective and simpler to utilize and maintenance in comparison to sluice gates (Clemmens et al. 2003; Shahrokhnia and Javan 2006). The performance accuracy of radial gates as check structures can vary according to the type of flow condition (i.e., free or submerged flow). Under a free flow condition, a free hydraulic jump is developed in the downstream part of the radial gate, and the flow is just affected by the upstream flow depth. On the other hand, with increasing tail-water depth, the jet emanating from a radial gate changes to partially submerged flow with an incomplete submerged jump and totally submerged flow with no jump. The discharge coefficient of gates constitutes as the basis for the determination of the opening, which is of considerable importance for accurate simulation and controlling of canal flow. Furthermore, accurate determination of discharge is crucial for informing water-saving policies and efficient operational management. Numerous studies have attempted to enhance discharge calibration of radial gates. An early study was carried out by Metzler (1948), who developed a series of maps to calculate the discharge coefficient of submerged radial gates. Toch (1955) and Buyalski (1983) applied the empirical relationship method to illustrate the discharge coefficient changes for both free and submerged flow conditions. Clemmens et al. (2003) acknowledged the inefficiency of empirical relationship approaches in calibrating submerged radial gates and developed the Energy–Momentum (E–M) approach to calibrate the discharge through a single radial gate. Using Buyalski’s experimental data, Wahl (2005) introduced relative gate opening to modify the correction factor. Shahrokhniya and Javan (2006) and Zahedani et al. (2012) derived equations for quantifying the submerged flow rate based on the downstream flow depth using dimensional analysis and nonlinear regression. Abdelhaleem (2016) conducted research on a discharge estimation of three submerged parallel radial gates and found that the method of dimensional analysis with the incomplete self-similarity concept excels any other calibration methods. The practical implementation of the aforementioned discharge calibration methods indicated inadequate performance under highly submerged flow (Guo et al. 2020). More recently, Guo et al. (2021) proposed new criteria in order to subdivided submerged flow into partially submerged and totally submerged flow. Using a new discharge calibration method, expressed as the identification method, they asserted that considering the classification method can satisfactorily enhance the discharge model accuracy for totally submerged flow.

Where classical methods are insufficient in providing persistent success due to the complexity and uncertainty of hydraulic of submerged radial gates, machine learning (ML) methods can be reliable tools that provide solutions for discharge estimation of radial gates. High efficiency and speed of machine learning methods in modeling sophisticated hydraulic problems have made them as one of the most widely used methods in predicting the discharge coefficient of different types of weirs. A promising application of artificial neural network (Bilhan et al. 2011; Salmasi et al. 2013; Parsaie and Haghiabi 2015; Karami et al. 2018), support vector machine (Azamathulla et al. 2016), and adaptive-network-based fuzzy inference systems (Shamshirband et al. 2016; Haghiabi et al. 2018) were reported in the literature. Moreover, more recently, deep learning techniques are employed to accurate discharge coefficient prediction of weirs (Chen et al. 2022). Generally, deep learning is a special kind of neural networks that improve modeling accuracy through increasing the number of layers (Dahou et al. 2022).

Kernel extreme learning machine (KELM) (Huang et al. 2011) is a single-layer forward neural network developed based on extreme learning machine (ELM) (Huang et al. 2006), which integrates the concept of kernel functions into ELM to enhance its generalization ability on prediction problems. Excellent prediction performance of KELM has encouraged researchers to employ this rigorous kernel-dependent method in predicting the discharge coefficient of piano key weirs (Roushangar et al. 2021a, b) and elliptical side orifices (Karbas, 2021).

On the other hand, a great diversity of meta-heuristic algorithms have been developed and utilized to solve the constrained optimization problems. These nature inspired algorithms such as genetic algorithm (GA), particle swarm optimization (PSO) and ant colony optimization (ACO) are attractive as they do not implement mathematical suppositions in the optimization problems and present better global search capability than conventional optimization algorithms. Combining the merits from different evolutionary algorithms has been also suggested to attain more efficiency from computational point of view. Outstanding performance of hybrid genetic algorithm and particle swarm optimization (Garg 2016; Mosleh et al. 2020), the fractional-order Caputo Manta-Ray Foraging Optimizer (Yousri et al. 2022), hybrid gravitation search algorithm and genetic algorithm (Garg 2015), modified whale optimization algorithm (Al-qaness et al. 2021), an effective hybrid method improved teaching–learning based on Harris hawks optimization (Kundu and Garg, 2021), and modified Aquila optimizer (Al-qaness et al. 2022) have been reported in the literature.

With the rapid development of ML and meta-heuristic methods, a large number of intelligent integrated systems
have been proposed. Coupling meta-heuristic methods is an effective way to optimize the parameters of ML algorithms and achieve the best performance. In this context, Roushangar and Shahnazi (2019) used a combination of KELM and the PSO optimization algorithm in predicting bed load transport rate of gravel-bed rivers. Liao et al. (2019) applied a useful prediction method based on grey wolf optimization (GWO) algorithm and KELM, their hybrid model enjoyed greater accuracy when it came to predict the displacement of step like landslide. There are also a number of published studies that describe the significant role of GWO-KELM method in modeling various phenomena (Wang et al. 2017a, b; Luo 2019; Zhou 2021; Roushangar et al. 2022). A detailed review of mentioned studies confirmed the GWO-KELM as promising to serve as robust early warning tool with great performance for prediction of various phenomena. The KELM has achieved a lot benefits from the GWO strategy due to its unique global search ability. In other word, the most suitable parameter set of KELM have been determined with the aid of GWO algorithm. The proposed GWO-KELM model requires the least running time compared to other hybrid models. Moreover, GWO-KELM has more flexibility under different hydraulic conditions.

Prediction of the discharge coefficient is required as a basis for discharge measurements in irrigation channels and plays a fundamental role in evaluating the hydraulic efficiency of radial gates. A great deal of previous researches into the hydraulic of radial gates showed that the application of the ML approaches in radial gates has not been tried before. A review of the literature shows that the hybrid GWO-KELM method has a high capability in modeling various phenomena. Moreover, these studies confirm the complicated hydraulics of submerged radial gates, which necessitates the use of a robust model in predicting their discharge coefficient. In this regard, taking the advantages of hybrid GWO-KELM, for the first time, this method was used to predict the discharge coefficient of submerged radial gates. Apart from this, a high performance comparative study among two other kernel-depend methods, including Gaussian process regression (GPR) and support vector machine (SVM), has been applied to predict the discharge coefficient of submerged radial gates. Furthermore, it is justified to adopt the hybrid GWO-KELM as more efficacious substantial predicting model to predict the discharge coefficient of radial gates under varied submergence conditions with different submergence ratios. An extensive database including 2125 experimental data points gathered by the U.S. Bureau of Reclamation was used to feed the utilized kernel-depend models. Prediction accuracy of the employed kernel-depend techniques was assessed through various input combinations and under different submergence conditions.

### 2 Materials and methods

The experimental data provided by Buyalski (1983) were employed for prediction of the discharge coefficient of submerged radial gates. 2657 experiments were conducted based on three types of gate (Fig. 1). The gate lip seal was varied by modeling three commonly used designs: (1) the hard rubber bar; (2) the music note seal; and (3) without a seal, resulting in a sharp edge configuration. The sharp edge gate lip was achieved by simply removing the gate lip seal and filling in the screw holes with solder. Radial gates were established at three different trunnion height settings including 409, 461, and 511 mm. The radius of the radial gate for all the experiments was 702 mm, the gate width was 711 mm, and the downstream channel width was 762 mm. The floor of the laboratory flume was flat and at the same level over its entire length. The experiments were carried out under three different flow conditions, namely free, submerged, and jump (assumed to be the transition zone) flow (Fig. 2). The discharge coefficient of the submerged radial gate was obtained through the following equation:

\[
Q = C_d w b_g \sqrt{2g(y_1 - y_3)}
\]

where \(Q\) is the discharge \((\text{m}^3/\text{sec})\), \(w\) is the gate opening \((\text{m})\), \(b_g\) is the gate width \((\text{m})\), \(y_1\) and \(y_3\) are the upstream and downstream flow depth \((\text{m})\), respectively, \(y_2\) is the flow depth at vena contracta \((\text{m})\), \(x\) is the gate trunnion height above the invert \((\text{m})\), \(\theta\) is the gate leaf angle from horizontal \((^\circ)\) and \(g\) is the gravitational constant \((\text{m}/\text{s}^2)\). Here, with the removal of outliers, 2,125 experimental data samples under submerged flow conditions were prepared to feed the employed ML methods.

### 3 Support vector machine (SVM)

In this section, basic features of SVM are provided since this kernel-depend technique has been implemented in diverse areas of engineering. SVM can be formulated from a given dataset of \(\{(x_i, y_i)\}\) as:

\[
f(x) = W \varphi(x) + b
\]

where the vector \(W\) is called the weight factor and \(b\) stands as the bias. In the above equation, \(\varphi(x)\) represents the transfer function. Minimization procedure of risk function is utilized for estimation of the aforementioned coefficient as indicated below:

\[
R_{SVM}(C) = \frac{1}{2}||w||^2 + C \frac{1}{2} \sum_{i=1}^{n} L(x_i, y_i)
\]

where \(C\), \(\frac{1}{2} \sum_{i=1}^{n} L(x_i, y_i)\) stands as the empirical risk. The
The abovementioned coefficients can be determined through the minimization procedure of regularized risk function. This procedure is conducted by inserting slack variables $\xi_i$ and $\xi_i^+$ as representative of upper and lower excess deviation.

Minimize $R_{\text{SVM}}(w, \xi^+, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^+)$ \hspace{1cm} (4)

Subject to $\left\{ \begin{array}{l}
y_i - w \phi(x_i) + b_i \leq \xi_i + \xi_i^+ \\
w \phi(x_i) + b_i - y_i \leq \xi_i + \xi_i^+ \\
\xi_i, \xi_i^+ \geq 0, i = 1, \ldots, n
\end{array} \right.$

where $\frac{1}{2} \|w\|^2$ is the regularization term, $C$ denotes the cost factor, $\varepsilon$ is called as the loss function, and $n$ indicates the sample size.

Using Lagrange multiplier and optimality constraints in Eq. (2) thus yields a general form of function as indicated below:

$$f(x) = \sum_{i=1}^{n} (\beta_i - \beta_i^+) K(x_i, x_j) + b$$ \hspace{1cm} (5)

where

$$K(x_i, x_j) = \phi(x_i) \phi(x_j)$$ \hspace{1cm} (6)

The term $K(x_i, x_j)$ denotes the kernel function to achieve the inner products of two vector $x_i$ and $x_j$ in the feature space $\phi(x_i)$ and $\phi(x_j)$, respectively. The various kernel functions used in the present study are as follows:

- **Linear kernel**: $K(x_i, x_j) = \langle x_i, x_j \rangle$ \hspace{1cm} (7)

- **Polynomial kernel**: $K(x_i, x_j) = ((x_i, x_j) + c)^d$ \hspace{1cm} (8)

- **Radial Basis Function (RBF)**: $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\gamma^2}\right)$ \hspace{1cm} (9)

- **Sigmoid kernel**: $K(x_i, x_j) = \tanh\left(\pi (x_i, x_j) + C\right)$ \hspace{1cm} (10)

Furthermore, the generalization capability of the employed SVM method was assessed through the k-fold cross-validation techniques to prevent over-fitting over the testing phase. The datasets were divided into $k$ identical sub-samples (randomly selected), each of which was defined as a fold $(F_1, F_2, F_3, \ldots, F_k)$. Among subsamples $k$, one subsample was implemented for validation, and reminder were implemented for the training process of this model. The k-fold cross-validation approach was then repeated for $k$ times, each of the $k$ subsamples being utilized exactly once in the validation phase. The merit of this approach is the random repetition of subsamples in the test and training process for all observations, and each observation is just utilized one time for validation. In the present study, $k$ is considered to be 10. The separating mechanism of the k-fold cross-validation approach is illustrated in Fig. 3.
4 Gaussian process regression (GPR)

An infinite group of random variables is defined as a Gaussian process (GP) in which any of the finite subgroups contains a constant joint Gaussian distribution (Rasmussen and Nickisch, 2010). A GP is described through a mean and covariance (Eq. 9). Simply, the mean value is usually considered zero as the GP stands as a linear combination of multiple random variables with a high-dimensional joint normal distribution. The Gaussian process can be expressed as below:

\[ y \sim GP \left( m(x), k(x, x') \right) \]  

where \( m(x) \) and \( k(x, x') \) stands as the mean and the covariance functions, respectively. In the Gaussian process, the expected value of the function \( y^* \) at the input matrix point \( x \) can be obtained through the mean function of \( m(x) \) as expressed in Eq. (12):

\[ m(x) = E[f(X)] \]  

The covariance function of \( k(x, x') \) plays a role as a confidence level measure of the mean function as shown in Eq. (13). The covariance function captures any two arguments in such a way that it creates a positive definite matrix of \( K \).

\[ k(x, x') = E \left[ (f(x) - m(x))(f(x') - m(x')) \right] \]  

The covariance function determines particular features of the model, including stationarity, smoothness and periodicity. The primary and extensively used Gaussian process regression (GPR) consists of a mean value function zero and covariance function. There are different types of covariance functions (kernel functions) that can be utilized in a core tool of GPR as represented in Eqs. (14–19):

- **Exponential:**
  \[ k(x, x') = \sigma_f^2 \left[ -\frac{1}{2} \right] \]

- **Squared exponential:**
  \[ k(x, x') = \sigma_f^2 \left( 1 + \sqrt{3}r \right) \exp \left[ -\sqrt{3}r \right] \]

- **Matern 3/2:**
  \[ k(x, x') = \sigma_f^2 \left( 1 + \sqrt{5r} + \frac{5r}{3} \right) \exp \left[ -\sqrt{5r} \right] \]

- **Matern 5/2:**
  \[ k(x, x') = \sigma_f^2 \left( 1 + \frac{1}{2\sigma_f^2} \right) \left( x - x' \right)^T \left( x - x' \right) \]  

The joint Gaussian prior distribution on the observed value \( y \) can be analytically derived as:

\[ y^* | y \sim N(\overline{f}_{x^*}, \text{cov}(f_*)) \]

where \( \overline{f}_{x^*} \) represents the prediction mean, and it provides the...
best approximation for \( y \). The variance \( \text{cov}(f_\ast) \) is a sign of the prediction uncertainty. The mean value of the prediction, \( f_\ast \) in Eq. (21) is a linear combination of the target \( y \) while on the other hand, the variance \( \text{cov}(f_\ast) \) is not reliant on the target but only inputs:

\[
\left( \begin{array}{c}
y \\
f_\ast
\end{array} \right) \sim N \left( \begin{array}{c} m(X) \\ m(X_*) \end{array} \right) \left[ \begin{array}{ccc} K(X,X) + \sigma^2_n I & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{array} \right]^{-1} \left( \begin{array}{c} y - m(x) \\ y_* - m(x_*) \end{array} \right)
\]

(21)

where \( K(X,X_*) \) indicates the covariance matrix of the training set \( X \) and the test set \( X_* \). Furthermore, \( \text{cov}(f_\ast) \) represents the matrix of covariance between the test set itself. The marginal likelihood over \( f_\ast \) is denoted in Eq. (22) as:

\[
P(f_\ast|X,y,x_*) \sim N(\overline{f}_*, \text{cov}(f_\ast))
\]

(22)

where

\[
\overline{f}_* = E(f_\ast|X,y,x_*) = m(X_*) + K(X,X_*) \left( K(X,X) + \sigma^2_n I \right)^{-1} [y - m(x)]
\]

(23)

Log marginal likelihood is obtained by simplifying the integral expression of Eq. (22) using the logarithmic identifier:

\[
\log P(y|X,\theta) = -\frac{1}{2} y^T \left( K(X,X) \right)^{-1} y - \frac{1}{2} \log |K(X,X)| - \frac{1}{2} \log 2\pi
\]

(24)

where \( \theta \) is the component that sets the hyper-parameters. The minimum posterior hyper-parameter in the covariance function can be achieved through maximizing the marginal likelihood or minimizing the negative log marginal likelihood. Maximization of this marginal likelihood is accessible through taking derivatives over the parameters and utilizing gradient descent.

### 5 Kernel extreme learning machine (KELM)

The extreme learning machine (ELM) is known as a fast learning algorithm with better generalization capability. One of the weaknesses of ELM is the instability of this algorithm, which is caused by the stochastic adjustment of the initial weights. Kernel-based ELM as an extension of ELM can solve this problem through introducing kernel method and regularization theory (Wang et al. 2017a, b). Compared to ELM, kernel-depend ELM has more stability and generalizability, and it has been extensively applied regression fitting problems.

The algorithm model of kernel-depend ELM can be denoted as follows:

\[
f(x) = h(x)\beta
\]

(25)

where \( h(x) \) is the hidden layer output matrix, and \( \beta \) is the connection weight between the hidden layer and the output layer. In the next step, the regularization coefficient of \( C \) is introduced for improving the stability and generalization capability of the model. Then, the least squares solution of the output weight can be achieved based on the theory of generalized inverse matrix:

\[
\hat{\beta} = H^T \left( \frac{1}{C} + HH^T \right)^{-1} y
\]

(26)

where \( H \) is the output matrix of the hidden layer and \( y \) is desired output vector. Therefore, the output function of the KELM can be expressed as:

\[
f(x) = \left( H^T \left( \frac{1}{C} + HH^T \right)^{-1} y \right) h(x)
\]

(27)

Due to the unknown nature of feature mapping function \( h(x) \), the kernel function should be embedded into the structure of KELM. In order to determine the kernel matrix \( \Omega_{\text{KELM}} \), the element of the matrix is:

\[
\Omega_{\text{KELM}}(i,j) = h(x_i)h(x_j) = K(x_i,x_j)
\]

(28)

And then, the network output of KELM can be written as:

\[
f(x) = h(x)H^T \left( \frac{1}{C} + HH^T \right)^{-1} y = \left[ K(x_i,x_1) \ldots K(x_i,x_N) \right] \left( \frac{1}{C} + \Omega_{\text{KELM}} \right)^{-1} y
\]

(29)

where \( K(x_i,x_j) \) stands as the kernel function. To select the appropriate kernel function, several most frequently used kernel functions are as follows:

- **Linear kernel**:
  \[
  K(x_i,x_j) = (x_i,x_j)
  \]
  (30)

- **Polynomial kernel**:
  \[
  K(x_i,x_j) = (\langle x_i,x_j \rangle + C)^d
  \]
  (31)

- **Radial Basis Function (RBF)**:
  \[
  K(x_i,x_j) = \exp \left( -\frac{\|x_i - x_j\|^2}{2\gamma^2} \right)
  \]
  (32)

- **Wavelet kernel**:
  \[
  K(x_i,x_j) = \cos \left( \omega \frac{\|x_i - x_j\|}{\beta} \right) \exp \left( -\frac{\|x_i - x_j\|^2}{\beta} \right)
  \]
  (33)
6 Grey Wolf Optimization (GWO)

GWO is a newly introduced intelligent algorithm inspired by the hunting behavior and social hierarchy of grey wolves developed by Mirjalili et al. (2014). Considering the social rank of wolves, the pack is divided into four levels from alpha (α) to omega (ω), as depicted in Fig. 4. The α wolf plays a leadership role as the group tracks his/her commands.

In the mathematical model of GWO, α, β, and δ are considered as the best solutions among the whole wolf pack and remaining solutions belong to ω wolves. In the mathematical model of GWO, the first three solutions are classified as α, β, δ, and the remaining solutions have relevance to ω wolves. The hierarchy in each iteration is updated according to the three best solutions to implement this mechanism. The location update of grey wolves is illustrated in Fig. 5.

The concept of GWO algorithm is based on searching, encircling, hunting, and attacking the prey. The first attempt before hunting process is encircling the prey by the pack of grey wolves. The encircling model of wolves is as follows:

\[
\vec{X}(t+1) = \vec{X}_p(t) - A \cdot D
\]

where \(\vec{X}\) stands as the position vector of wolf, \(\vec{X}_p(t)\) is the prey position, \(t\) and \(t+1\) represents the current and next locations and \(A\) is the convergence factor, and \(D\) represents the distance between the prey and the wolf which is determined as follows:

\[
D = |C \cdot \vec{X}_p(t) - \vec{X}_t|
\]

\[
\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a}
\]

\[
\vec{C} = 2\vec{r}_2
\]

where \(\vec{r}_1\) and \(\vec{r}_2\) are random vectors in \([0, 1]\). The α wolf is usually guided the hunt and is closer to the prey position. The β and δ wolves might also contribute in hunting. For mathematically simulation of hunting behavior, the α, β, and δ wolves as the first, second and third best candidate solutions are supposed to have more information about the possible position of prey. Therefore, ω wolves should be obligated to renovate their positions as specified by the position of the best search agents as follows:

\[
\vec{X}(t+1) = \frac{1}{3} \vec{X}_1 + \frac{1}{3} \vec{X}_2 + \frac{1}{3} \vec{X}_3
\]

where \(\vec{X}_1\), \(\vec{X}_2\) and \(\vec{X}_3\) are the position of the grey wolf α, β, and δ, respectively:

\[
\vec{X}_1 = \vec{X}_α(t) - A_α \cdot \vec{D}_α
\]

\[
\vec{X}_2 = \vec{X}_β(t) - A_β \cdot \vec{D}_β
\]

\[
\vec{X}_3 = \vec{X}_δ(t) - A_δ \cdot \vec{D}_δ
\]

where

\[
\vec{D}_α = |C_1 \cdot \vec{X}_α(t) - \vec{X}|
\]

\[
\vec{D}_β = |C_2 \cdot \vec{X}_β(t) - \vec{X}|
\]

\[
\vec{D}_δ = |C_3 \cdot \vec{X}_δ(t) - \vec{X}|
\]

Fig. 5 Position updating procedure of ω wolves based on the positions of α, β and δ wolves

Fig. 4 The social hierarchy of grey wolves
position of the grey wolf in the pack can be obtained after iteration as expressed in Eq. (45):

$$\overline{X}(t + 1) = \frac{(\overline{X}_1 + \overline{X}_2 + \overline{X}_3)}{3}$$ (45)

7 Modeling configurations

In this paper, propose hybrid GWO-KELM technique was implemented in the MATLAB® environment. Optimal adjustment of constant parameters $\rho$ and especially the kernel parameters might considerably affect the proposed modeling process. The optimum values of the above-mentioned effective parameters were obtained using the GWO algorithm. Therefore, GWO was embedded into the employed KELM method to enhance the modeling capability of discharge coefficient of the submerged radial gate. The GWO algorithm does not have any affinity to stick in local optimum points in the complex multimodal optimization problem. This nature inspired algorithm provides more varied search of the solution space for solving complex optimization problems. Improved optimum solutions with reduced computation burden can be accessed in GWO in comparison with the existing stochastic search methods (Roushangar et al. 2021a, b). The GWO is superior to these methods because (i) The GWO has better capability for information sharing and benefits of improved conveying mechanism; (ii) it utilizes random function and considers three candidate solutions for getting the best performance and converges quickly by making jump from local minima towards global minima. The main advantage of GWO algorithm over most of the widely known meta-heuristic algorithms is that the GWO algorithm operation requires no specific input parameters. Additionally, it is straightforward and free from computational complexity. In order to divide the total data into training and testing sections, K-fold cross validation approach is applied by considering $K = 10$ for random selection of data division. The data set containing 2125 data samples were partitioned into two sub-sets. The first separation as 75% of total data was associated with training set while the remaining 25% of data were used for testing goals. It is noteworthy that, the validity of a predicting model is significantly attributed to the range of data used in building the model. Generally, using abnormal data may lead to undesirable results, so in this study all input variables are scaled in the range of 0.1–0.9. This leads to uniformity of data values for the network and increases the speed of training and model capability (Dawson and Wilby, 1998). The following equation was used to normalize the utilized data in this study:

$$x_n = 0.1 + 0.9 \times \left( \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \right)$$ (46)

where $x_n$, $x$, $x_{\text{max}}$, $x_{\text{min}}$, respectively, are: the normalized value of variable $i$, the original value, the maximum and minimum of variable $i$.

The flowchart of the suggested hybrid model is depicted in Fig. 6.

8 Performance evaluation criteria

The correlation coefficient ($R$), the determination coefficient ($\text{DC}$), and the root-mean-square error ($\text{RMSE}$) were employed as statistical indices for evaluating the models. $R$ was applied to show linear correlation between simulated and observed values. $\text{DC}$ is a commonly used index in water resources modeling which has a range of (-∞, 1]. A value of unity demonstrates a perfect match between the simulated and observed values. Finally, $\text{RMSE}$ shows the model’s estimation error.

$$R = \frac{\sum_{i=1}^{N} (X_i - \overline{X}) \times (Y_i - \overline{Y})}{\sum_{i=1}^{N} (X_i - \overline{X})^2 \times (Y_i - \overline{Y})}$$ (47)

$$\text{DC} = 1 - \frac{\sum_{i=1}^{N} (X_i - Y_i)^2}{\sum_{i=1}^{N} (X_i - \overline{X})^2}$$ (48)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (X_i - Y_i)^2}{N}}$$ (49)

where $X_i$ and $Y_i$ denote the observed and simulated values of the discharge coefficient, respectively, and $N$ represents the number of data, also:

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$$ (50)

And

$$\overline{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i$$ (51)

9 Input parameters

Design of the optimum input matrix of the ML models that can give the best simulation of the target parameter is an important step with application of these models. The flow motion through a submerged radial gate can be expressed by the following functional relationship:
In which \( Q \) = flow discharge, \( y_1 \) = upstream flow depth, \( y_3 \) = downstream flow depth, \( L \) = channel width, \( w \) = gate opening, \( \mu \) = water viscosity and \( \rho \) = water density.

Considering \( w \), \( g \) and \( \mu \) as independent variables, the \( \Pi \)-theorem provides the following five dimensionless groups:

\[
F(Q, y_1, y_3, L, w, g, \mu, \rho) = 0
\]  

(52)

\[
\Pi_1 = \frac{Q}{w^2 g^2}
\]

(53a)

\[
\Pi_2 = \frac{y_1}{w}
\]

(53b)

\[
\Pi_3 = \frac{y_3}{w}
\]

(53c)
\[ \Pi_4 = \frac{B}{w} \]  
(53d)

\[ \Pi_5 = \frac{w^3 g \rho}{\mu} \]  
(53e)

Rearranging the dimensionless group yields:

\[ \frac{\Pi_1 \Pi_3}{\Pi_4} = \frac{\rho Q}{\mu L} = Re \]  
(54)

In which \( Re \) = Reynolds number, and:

\[ \Pi_2 - \Pi_3 = \frac{y_1 - y_3}{w} \]  
(55)

Accordingly, the functional relationship, Eq. (51), can be rearranged in the following form:

\[ F\left(\frac{y_1}{w}, \frac{y_3}{w}, Re, \frac{y_1 - y_3}{w}\right) = 0 \]  
(56)

On the other hand, an experimental study by Toch (1955) demonstrated influential effect of \( w/R \), \( y_3/R \), \( y_1/R \), and \( a/R \) on the discharge coefficient of submerged radial gates. Neglecting the influence of the Reynolds number and rearranging the dimensionless groups yields:

\[ F\left(\frac{y_1 - y_3}{w}, \frac{y_1}{w}, \frac{y_3}{w}, \frac{w}{R}, \frac{y_3}{R}, \frac{a}{R}\right) = 0 \]  
(57)

Various combinations of the aforementioned parameters were prepared, and the best ones were selected through and trial and error procedure as listed in Table 1. Statistical information of input and output parameters in terms of the maximum (max), minimum (min), mean and standard deviation (S.t.d.) value is represented in Table 2.

### 10 Results and discussions

Table 3 provides the statistical indicators of the employed kernel-depend models during the training and testing phases. For the implementation of GWO-KELM, the population (pack size) was set to 300 and iteration was applied as 50. Taking into consideration the statistical indices, all the employed kernel-depend models generated the most accurate results with the model (4), where \( y_1-y_3/w \) and \( y_3/w \) were utilized as the input parameters. The comparison results of the employed kernel-depend methods revealed that the hybridized GWO-KELM model the (4) as input combination outperformed the SVM, and GPR standalone models with the lowest RMSE (0.027) and highest R (0.983) and DC (0.966) values. The results presented in Table 3 also showed that the submergence ratio \( (y_1-y_3)/w \) as a common parameter in all developed input models has an essential role in the modeling process, and is the most influential parameter in predicting the discharge coefficient of submerged radial gates. The results are consistent with the findings of Ansar (2001), Ferro (2001), and Shahrokhnia and Javan (2006) that the discharge of a submerged radial gate can be considered as a function of the differential flow depth. Comparing the model (2) and (3) demonstrated that the addition of \( y_3/w \) instead of \( y_1/w \) has a positive effect on the prediction accuracy of the employed kernel-depend methods, and it can be declared that the flow through the radial gate is more affected by downstream flow depth. Similarly, using the ratio of the downstream flow depth to the gate opening \( (y_3/w) \) significantly increases the modeling accuracy. The scatter plots of predicted discharge coefficient \( (C_d) \) vs. observed values for proposed GWO-KELM models are depicted in Fig. 7.

In order to provide more information about the performance of employed kernel-depend approaches, the distribution of relative errors and their histograms are plotted for the best input combination (Fig. 8). The relative errors of the proposed GWO-KELM met the characteristic of normal distribution (mean = 0.03). The relative errors were restricted to a limited range between \((-0.5 \) and \(+1.5\)) for GPR, \((-1 \) and \(+2.5\)) for SVM, and \((-0.8 \) and \(+1.6\)) for GWO-KELM. Assessing the histogram of errors reveals that the distribution of the errors was symmetrical and the errors were accumulated around zero for the GWO-KELM, which indicated better performance of GWO-KELM approach in the prediction of \( C_d \) compared to the GPR, and SVM methods.

**Table 1 Selected input combinations for prediction of discharge coefficient of submerged radial gates**

| Model | Input parameters |
|-------|------------------|
| (1)   | \( \frac{y_1-y_3}{w} \) |
| (2)   | \( \frac{y_1}{w}, \frac{y_3}{w} \) |
| (3)   | \( \frac{y_1-y_3}{w}, \frac{y_3}{w} \) |
| (4)   | \( \frac{y_1-y_3}{w}, \frac{y_3}{w} \) |

**Table 2 The statistical details of the input and output parameters**

| Parameter | Min | Max | Mean | SD  |
|-----------|-----|-----|------|-----|
| Input parameters | \( \frac{y_1-y_3}{w} \) | -3.71 | 14.88 | 0.928 | 1.44 |
| | \( \frac{y_1}{w} \) | 0.078 | 2.17 | 0.693 | 0.234 |
| | \( \frac{y_3}{w} \) | 0.054 | 1.02 | 0.585 | 0.227 |
| | \( \frac{a}{R} \) | 0.686 | 17.77 | 4.03 | 3.04 |
| Output parameter | \( C_d \) | 0.025 | 0.705 | 0.313 | 0.151 |
can be classified into two major groups: global and local. Global kernel functions, such as linear and polynomial kernel functions, show high generalization performance but a weak learning ability and are affected by samples far from each other. Local kernel functions, such as RBF and wavelet kernel functions, have a strong learning ability but low generalization performance and are affected by samples close to each other. Here, to find an appropriate kernel function, different kernel functions were utilized in the structure of the employed kernel-depend techniques. For the implementation of the embedded SVM method, the best values of kernel parameters were achieved through a trial-and-error procedure. Additionally, optimization process of the cost factor (C) and the loss function (ε) was conducted by a grid search in the C and ε parameter space using the cross-validation technique. In this grid search a normal range of parameters settings are examined. First, the optimal values of C and ε were obtained in the intervals of

| Method   | Model | Train | Test |
|----------|-------|-------|------|
|          |       | R     | DC   | RMSE | R     | DC   | RMSE |
| GWO-KELM | (1)   | 0.802 | 0.644 | 0.090 | 0.783 | 0.613 | 0.092 |
|          | (2)   | 0.840 | 0.706 | 0.082 | 0.814 | 0.662 | 0.086 |
|          | (3)   | 0.870 | 0.758 | 0.074 | 0.841 | 0.707 | 0.080 |
|          | (4)   | 0.988 | 0.994 | 0.016 | 0.983 | 0.966 | 0.027 |
| GPR      | (1)   | 0.814 | 0.663 | 0.088 | 0.783 | 0.613 | 0.092 |
|          | (2)   | 0.860 | 0.744 | 0.077 | 0.810 | 0.656 | 0.087 |
|          | (3)   | 0.860 | 0.738 | 0.077 | 0.838 | 0.702 | 0.081 |
|          | (4)   | 0.994 | 0.988 | 0.016 | 0.973 | 0.956 | 0.030 |
| SVM      | (1)   | 0.790 | 0.620 | 0.093 | 0.773 | 0.595 | 0.094 |
|          | (2)   | 0.806 | 0.648 | 0.090 | 0.792 | 0.627 | 0.090 |
|          | (3)   | 0.841 | 0.705 | 0.082 | 0.827 | 0.683 | 0.083 |
|          | (4)   | 0.974 | 0.949 | 0.034 | 0.972 | 0.944 | 0.035 |

Fig. 7 Scatter plot of observed versus predicted $C_d$ for different developed models (test series)
The result demonstrated that SVM-RBF showed better prediction accuracy \((R = 0.972, \text{DC} = 0.944, \text{RMSE} = 0.035)\) as reported in the literature (Azamathulla et al. 2016; Le et al. 2020; Roushangar and Shahnazi 2020a, b). It was followed by linear kernel function \((R = 0.761, \text{DC} = 0.558, \text{RMSE} = 0.099)\) and polynomial kernel \((R = 0.645, \text{DC} = 0.398 \text{ and RMSE} = 0.115)\). In modeling the present phenomenon, the linear kernel function as the simplest kernel outperformed the polynomial kernel with the shortest processing time. The results indicated the poor performance of the SVM model using the sigmoid kernel function \((R = 0.127, \text{DC} = -0.00043 \text{ and RMSE} = 0.132)\). In the case of GPR, for tuning the associated hyper-parameters, a standard gradient descent optimizer was employed with maximizing the log marginal likelihood. Table 4 exhibits the performance of different kernel functions in the structure of employed GPR model. In consideration of the DC values, small variation can be observed throughout the employed kernel functions. The GPR performance ranges between 0.957 (for Squared Exponential kernel) and 0.960 (for Matern 3/2 kernel).

In the SVM, model performance is substantially affected by the RBF kernel parameter \((\gamma)\), which may cause under-fitting and over-fitting problems (Roushangar and Shahnazi, 2020a, b and 2021). Figure 9 illustrates the statistical parameter of DC via gamma values of the SVM model. From the figure, it can be observed that the accuracy of the SVM model shows various behavior with changes of gamma values for different developed models. Fluctuations with the introduction of the model (1) as an input to SVM are more evident. The best-fitting gamma values for the best input combination (model 4) are obtained when \(\gamma \geq 150\). The SVM model with small values of \(\gamma\) has tends to overfit, or memorize of data. The ideal value of the RBF kernel parameter was obtained \(\gamma = 250\) for the model (4) from the analysis performed. The optimal hyper-parameters of SVM for the model (4) were achieved as \((C = 1, \varepsilon = 0.001)\).

Since the gate trunnion pin height is an effective geometrical characteristic of radial gates on flow rate, it may be beneficial to check out the prediction capability of employed kernel-depend methods for different trunnion pin heights. For this purpose, developed input combinations were re-run for three different trunnion pin-heights, including 409, 461, and 511 mm and the obtained results are presented in Table 5. It can be noticed a slightly better performance of employed SVM, GPR, and GWO-KELM techniques in predicting discharge coefficient of submerged radial gates with the gate trunnion heights of \(\alpha = 409\) mm and \(\alpha = 511\) mm. For radial gates with the gate trunnion heights of \(\alpha = 409\) mm, introducing \(y_1/R\) and \(y_3/R\) (models 2 and 3) led to a decrease in overall modeling performance. It was more evident in the case of SVM, where modeling accuracy was reduced by 27% and 22% in terms of DC, respectively.
Table 4  The results of GPR method with different kernel functions (fed with model III)

| Kernel type          | Performance criteria | Train | Test |
|----------------------|----------------------|-------|------|
|                      |                      | R     | DC   | RMSE | R   | DC   | RMSE |
| Exponential          |                      | 0.988 | 0.994 | 0.016 | 0.973 | 0.956 | 0.030 |
| Squared Exponential  |                      | 0.979 | 0.960 | 0.030 | 0.978 | 0.957 | 0.030 |
| Matern 3/2           |                      | 0.983 | 0.966 | 0.027 | 0.980 | 0.960 | 0.029 |
| Matern 5/2           |                      | 0.981 | 0.962 | 0.029 | 0.979 | 0.958 | 0.030 |
| Rational Quadratic   |                      | 0.981 | 0.962 | 0.029 | 0.979 | 0.958 | 0.030 |

Fig. 9  Variation Statistics parameter of DC vs. RBF kernel parameter (γ) to find SVM optimums
The performance of the employed kernel-depend methods was compared with those of the traditional models based on the dimensional analysis. To achieve this aim, the formulas of Shahrokhniya and Javan (2005), (2006) and Zahedani et al (2012) were used as depicted in Table 6. It should be noted that the selected dimensionless formulas are developed based on experimental data and under submerged flow conditions.

The results illustrated in Fig. 10 reflect that the employed kernel-depend methods give better predictions than the selected dimensionless formulas. As is shown in this figure, Shahrokhniya and Javan’s (2005) approach underestimates the discharge coefficient of submerged radial gates with statistical indices of $R = 0.942$, $DC = -4.29$, and $RMSE = 0.348$. By contrast, Shahrokhniya and Javan’s (2006) approach overestimates the discharge coefficient with statistical indices of $R = 0.969$, $DC = -1.46$, and $RMSE = 0.237$. Existing equations and related constant parameters are developed within a specific range of flow conditions, and this inefficiency can be due to the

### Table 5

| Method  | Trunnion pin-heights | Model | Performance criteria |
|---------|-----------------------|-------|-----------------------|
|         |                       |       | Train | Test               |
|         |                       |       | $R$    | $DC$ | $RMSE$ | $R$   | $DC$ | $RMSE$ |
| GWO-KELM| 409                   | (1)   | 0.805 | 0.643 | 0.076 | 0.801 | 0.640 | 0.079 |
|         |                       | (2)   | 0.867 | 0.752 | 0.065 | 0.813 | 0.652 | 0.075 |
|         |                       | (3)   | 0.854 | 0.729 | 0.068 | 0.785 | 0.614 | 0.079 |
|         |                       | (4)   | 0.982 | 0.965 | 0.024 | 0.982 | 0.965 | 0.023 |
|         | 461                   | (1)   | 0.812 | 0.659 | 0.093 | 0.812 | 0.659 | 0.097 |
|         |                       | (2)   | 0.843 | 0.711 | 0.085 | 0.832 | 0.692 | 0.092 |
|         |                       | (3)   | 0.865 | 0.748 | 0.079 | 0.856 | 0.731 | 0.086 |
|         |                       | (4)   | 0.979 | 0.959 | 0.032 | 0.966 | 0.933 | 0.043 |
|         | 511                   | (1)   | 0.830 | 0.688 | 0.080 | 0.819 | 0.668 | 0.080 |
|         |                       | (2)   | 0.846 | 0.715 | 0.076 | 0.838 | 0.701 | 0.076 |
|         |                       | (3)   | 0.909 | 0.826 | 0.060 | 0.846 | 0.706 | 0.075 |
|         |                       | (4)   | 0.991 | 0.983 | 0.018 | 0.986 | 0.971 | 0.023 |
| GPR     | 409                   | (1)   | 0.838 | 0.699 | 0.072 | 0.835 | 0.693 | 0.070 |
|         |                       | (2)   | 0.894 | 0.792 | 0.060 | 0.804 | 0.640 | 0.076 |
|         |                       | (3)   | 0.878 | 0.767 | 0.063 | 0.804 | 0.641 | 0.076 |
|         |                       | (4)   | 0.999 | 0.998 | 0.005 | 0.989 | 0.978 | 0.018 |
|         | 461                   | (1)   | 0.822 | 0.675 | 0.090 | 0.822 | 0.676 | 0.095 |
|         |                       | (2)   | 0.870 | 0.756 | 0.078 | 0.835 | 0.696 | 0.092 |
|         |                       | (3)   | 0.882 | 0.777 | 0.075 | 0.854 | 0.729 | 0.087 |
|         |                       | (4)   | 0.994 | 0.988 | 0.016 | 0.973 | 0.947 | 0.038 |
|         | 511                   | (1)   | 0.864 | 0.745 | 0.072 | 0.833 | 0.693 | 0.077 |
|         |                       | (2)   | 0.884 | 0.778 | 0.067 | 0.824 | 0.677 | 0.078 |
|         |                       | (3)   | 0.916 | 0.836 | 0.058 | 0.846 | 0.711 | 0.074 |
|         |                       | (4)   | 0.999 | 0.998 | 0.005 | 0.989 | 0.978 | 0.020 |
| SVM     | 409                   | (1)   | 0.848 | 0.715 | 0.070 | 0.822 | 0.673 | 0.073 |
|         |                       | (2)   | 0.819 | 0.670 | 0.075 | 0.741 | 0.532 | 0.087 |
|         |                       | (3)   | 0.827 | 0.677 | 0.075 | 0.755 | 0.551 | 0.085 |
|         |                       | (4)   | 0.981 | 0.962 | 0.025 | 0.979 | 0.957 | 0.026 |
|         | 461                   | (1)   | 0.810 | 0.656 | 0.093 | 0.805 | 0.647 | 0.099 |
|         |                       | (2)   | 0.841 | 0.701 | 0.087 | 0.823 | 0.669 | 0.096 |
|         |                       | (3)   | 0.847 | 0.714 | 0.085 | 0.830 | 0.681 | 0.094 |
|         |                       | (4)   | 0.974 | 0.948 | 0.036 | 0.962 | 0.925 | 0.045 |
|         | 511                   | (1)   | 0.828 | 0.686 | 0.080 | 0.819 | 0.671 | 0.079 |
|         |                       | (2)   | 0.846 | 0.714 | 0.077 | 0.804 | 0.637 | 0.083 |
|         |                       | (3)   | 0.864 | 0.744 | 0.072 | 0.843 | 0.709 | 0.075 |
|         |                       | (4)   | 0.987 | 0.975 | 0.022 | 0.983 | 0.963 | 0.026 |
method’s range of validity. However, the proposed formulas of Shahrokhniya and Javan (2005 and 2006) enjoyed better correlation with the observed data as they incorporate the gate leaf angle (θ) that plays an effective role on the discharge coefficient. Moreover, the dimensionless formula developed by Zahedani et al. (2012) presented an extremely poor performance on prediction of discharge coefficient with statistical indices of $R = -0.233$, DC = -4.29, and RMSE = 0.348. Generally, proposed traditional methods are not flexible enough to present uniform results under different conditions, and their applicability is limited to the specific cases of their development.

Generally, kernel-depend techniques allow high dimension genomic data, fit nonlinear relations between outcomes and genomic data and find flexible ways to include structured information and computational sophistication. Among them, KELM can obtain comparative and more excellent property with faster training speed and much easier implementation in application. The GWO strategy conveys advantages into KELM due to its unique global search ability to acquire the most suitable parameter set of KELM. The set of mentioned reasons increases the efficiency of the employed hybrid method compared to the dimensionless formulas and under different hydraulic conditions.

A review of previous studies shows that the prediction performance of existing calibration methods is poor for submerged flows (Guo et al. 2020). In order to assess the capability of proposed kernel-depend approaches in the quantification of discharge coefficient under varied submergence conditions, the submergence ratio, defined as $S_r = \frac{y_1 - y_3}{w}$ was considered as a representative of subdivision criteria for submerged flow through a radial gate. Different intervals of mentioned parameters were evaluated through the large number of trial and error processes, and the generalization ability of proposed kernel-depend approaches with selected best input combinations were analyzed for each interval. The obtained results are depicted in Fig. 11, which demonstrating an obvious rising trend in prediction accuracy of kernel-depend techniques. These findings confirmed that the performance of GWO-KELM, SVM and, GPR techniques with the specified input parameters inclines to be more robust with decreasing levels of submergence. According to the obtained results, proposed hybrid GWO-KELM method gave the most accurate results ($R = 0.873$, DC = 0.744, and RMSE = 0.035) for extremely highly submerged flow with $S_r < 0.05$ (where $S_r = y_1 - y_3/w$).

### 11 Conclusions

The present paper is an attempt to implement a new kernel-depend extreme learning machine grey wolf optimization algorithm (GWO-KELM), Gaussian process regression (GPR), and support vector machine (SVM) for modeling discharge coefficient of submerged radial gates. The optimization capability of GWO that naturally inspired by the

---

**Table 6** Selected dimensionless formulas for prediction of discharge coefficient

| Name | Formula |
|------|---------|
| Shahrokhniya and Javan (2006) | $C_d = 0.62\left(\frac{\theta}{\pi}\right)^{-0.06}\left(\frac{y_1 - y_3}{y_3}\right)^{0.37}$ |
| Shahrokhniya and Javan (2005) | $C_d = 0.53\left(\frac{\theta}{\pi}\right)^{0.87}\left(\frac{y_1 - y_3}{w}\right)0.33$ |
| Zahedani et al. (2012) | $C_d = 0.8814\left(y_1 - y_3\right)^{0.0246}w^{-0.0273}R^{-0.0511}$ |

**Fig. 10** The obtained results of dimensionless formulas  
(a) Shahrokhniya and Javan (2005)  
(b) Shahrokhniya and Javan (2006)  
(c) Zahedani et al. (2012)
swarm evolutionary feature is utilized to tune the KELM model for the implemented application. The obtained results showed that the proposed GWO-KELM gives better prediction accuracy than employed GPR and SVM approaches. Among the developed input combinations, the model having input parameters of the ratio of the downstream flow depth to the gate opening \((y_3/w)\) and submergence ratio \((y_1-y_3/w)\) provides the best performance. The submergence ratio as a common parameter in all developed input models is the most influential parameter in predicting the discharge coefficient of submerged radial gates. The introduced influential parameters can be used in the development of new relationships to estimate the discharge coefficient of submerged radial gates. The introduced influential parameters can be used in the development of new relationships to estimate the discharge coefficient of submerged radial gates. The results also showed that despite other calibration methods, the employed kernel-depend approaches have high flexibility in predicting the discharge coefficient under different submergence conditions. It was showed that the prediction performance of employed kernel-depend methods tends to be more robust with decreasing levels of submergence. It is clear that there are variations in the prediction process of discharge coefficient under varied submergence Therefore, the exact state of flow condition should be determined before any modeling attempts are made. The outcomes of this work will help engineers to make some contribution towards the practical application of AI methods in hydraulic science. The results of this study indicate that by knowing the hydraulic conditions of the radial gate, a relative understanding of the accuracy of discharge coefficient estimation can be achieved. It is noteworthy to mention that the problems related to the lack of field data related to radial gates are among the limitations of the present study. However, the proposed kernel-depend techniques are data-driven, so further studies should be conducted using data ranges beyond this study and field data to confirm the high capabilities of the proposed prediction tools to model the discharge coefficient of submerged radial gates.

**Funding** No funding was received for conducting this study.

**Availability of data and materials** The data and materials that support the findings of this study are available on request from the corresponding author.

**Declarations**

**Conflict of interest** The authors declare that they have no competing interests.

**Ethical approval** Not applicable, because this article does not contain any studies with human or animal subjects.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**References**

Abdelhaleem FSF (2016) Discharge estimation for submerged parallel radial gates. Flow Meas Instrum 52:240–245

Al-qaness MA, Ewees AA, Abd Elaziz M (2021) Modified whale optimization algorithm for solving unrelated parallel machine scheduling problems. Soft Comput 25(14):9545–9557

Al-qaness, M. A., Ewees, A. A., Fan, H., AlRassas, A. M., & Abd Elaziz, M. (2022). Modified aquila optimizer for forecasting oil production. Geo-spatial Information Science, 1–17.

Ansar M, Ferro V (2001) Simultaneous Flow over and under a Gate. J Irrig Drain Eng 127(5):325–328

Azamathulla HM, Haghiahi AH, Parsaie A (2016) Prediction of side weir discharge coefficient by support vector machine technique. Water Sci Technol Water Supply 16(4):1002–1016

Bilhan O, Emiroglu ME, Kisi O (2011) Use of artificial neural networks for prediction of discharge coefficient of triangular
An efficient hybrid grey wolf optimization-based KELM approach for prediction of the discharge...
Zahedani MR, Keshavarzi A, Javan M, Shahrokhnia MA (2012) New equation for estimation of radial gate discharge. In Proceedings of the institution of civil engineers-water management (vol 165, No 5, pp 253–263). Thomas Telford Ltd.

Zhou F, Han J, Yang X (2021) Multivariate hierarchical multiscale fluctuation dispersion entropy: applications to fault diagnosis of rotating machinery. Appl Acoust 182:108271

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.