Uncertainty assessment of kernel based approaches on scour depth modeling in downstream of ski-jump bucket spillways

Redvan Ghasemlounia and Seyed Mahdi Saghebian

ABSTRACT

From the hydraulic structures designer's point of view, the scour depth accurate estimation in downstream of spillways is so important. In this study, the scour depth was assessed downstream of ski-jump bucket spillways using two kernel based approaches namely Support Vector Machine (SVM) and Kernel Extreme Learning Machine (KELM). In the model developing process, two states were tested and the impacts of dimensional and non-dimensional parameters on model efficiency were assessed. The best applied model dependability was investigated via Monte Carlo uncertainty analysis. In addition, the model accuracy was compared with some available semi-theoretical formulas. It was observed that the applied models were more successful than available formulas. The sensitivity analysis results showed that q (unit discharge of spillway) variable in the state 1 and $q^2/[gY_t^3]$ (g is gravity acceleration and Y_t is tail water depth) variable in the state 2 were the most significant parameters in the modeling process. Comparison among applied methods and one other intelligence approach showed that KELM was more successful in predicting process. The obtained results from uncertainty analysis indicated that the KELM model had an allowable degree of uncertainty in the scour depth modeling.

Key words | KELM, scour depth, ski-jump bucket, uncertainty analysis

Redvan Ghasemlounia
Department of Civil Engineering,
Istanbul Gedik University,
Istanbul,
Turkey

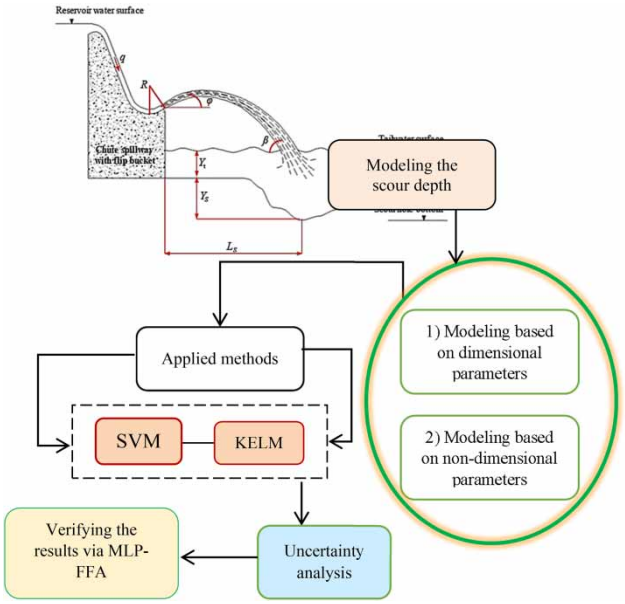
Seyed Mahdi Saghebian (corresponding author)
Department of Civil Engineering, Ahar Branch,
Islamic Azad University,
Ahar,
Iran
E-mail: *smsaghebian@iau-ahar.ac.ir*;
smsaghebian@gmail.com

HIGHLIGHTS

- The capability of two kernel based models (i.e. SVM, KELM) was investigated for scour depth assessing downstream of ski-jump bucket spillways.
- The capability of applied methods was compared with some available semi-empirical equations.
- The most important parameters were determined using sensitivity analysis.
- Monte Carlo uncertainty analysis was applied to investigate the dependability of the applied models.

doi: 10.2166/ws.2021.063

GRAPHICAL ABSTRACT



INTRODUCTION

Control structures are provided for dams to release flood water in excess of reservoir capacity. Ski-jump bucket spillways are one of the most commonly used structures in this regard. In ski-jumps, the whole jet of flow is thrown into the air using a bucket. Part of the energy in the jet is dissipated in the air because of friction and some other part is dissipated at the point of impact with the riverbed downstream as a result of excavating a large scour hole (Dargahi 2003). The scouring continues up to the point at which the rate of bed erosion is balanced by the rate of deposition of material brought back into the scour hole by the return flow (Chang *et al.* 2004). In the past decades, accurately prediction of scour dimensions has been of much interest among many investigators to prevent scour damages. The depth of scour is governed by various parameters, such as discharge intensity, height of fall, bucket radius, bucket lip angle, type and size of rock, degree of rock homogeneity, time, and mode of operation of spillway. In order to estimate the scour depth downstream ski-jump buckets, several empirical formulas have been developed (Damle *et al.* 1966; Martins 1975; Hoffmans 1998; Lopardo *et al.* 2002; Dargahi 2003).

Strelchuk (1969) investigated the scouring of the gravel beds at the downstream of ski-jump bucket spillways and showed that double and triple the flow discharge would increase the depth of the hole by 50 and 80 percent:

$$Ys = 3.69H_1\,(q^2/gH_1^3)^{0.3} \times (H_1/d_{50})^{0.1}\varphi^{0.36} \quad (1)$$

where Ys: scour depth, H_1: total head, q: unit discharge of spillway, φ: bucket lip angle, g: acceleration gravity, and d_{50} mean sediment size. Mason & Arumugam (1985) analyzed some of the existed relationships which had been proposed to estimate the maximum scour depth. They presented the following new relationship:

$$Ys = 3.27H_1^{0.05} \times q^{0.6} \times (Y_t^{0.15}/g^{0.3}d_{50}^{0.1}) \quad (2)$$

where Y_t is tail water depth. In 1974, Chee and Kung introduced the following relationship to estimate the maximum

scour depth of spillways:

$$Ys = 1.663q^{0.6H_1^{0.2}}/d_{50}^{0.1} \quad (3)$$

Chee & Padiyar (1969) provide the following relationship using three parameters of total head, unit discharge of spillway, and mean sediment size:

$$Ys = 2.12q^{0.67H_1 0.18}/d_{50}^{0.063} \quad (4)$$

The outcomes of conventional models are not general due to the scour depth complexity and uncertainty. Therefore, it is necessary to adopt or develop new methods for the accurate estimation of scour depth downstream of ski-jump bucket spillways. Over the past few decades several artificial intelligence (AI) methods [e.g., Artificial Neural Networks (ANNs), Neuro-Fuzzy models (NF), Genetic Programming (GP), Gene Expression Programming (GEP), Support Vector Machine (SVM), and Kernel Extreme Learning Machine (KELM)] have been developed and applied for assessing the complex hydraulic and hydrologic phenomena efficiency. Daily dewpoint temperature prediction (Al-Shammari *et al.* 2016), relative energy dissipation prediction (Saghebian 2019), longitudinal dispersion coefficients computing in natural streams (Azamathulla & Wu 2011), side weir discharge coefficient modeling (Azamathulla *et al.* 2017), monthly streamflow modeling (Zhu *et al.* 2018; Pandhiani *et al.* 2020), roughness coefficient modeling in sewer pipes (Roushangar *et al.* 2020), and forecasting long-term evapotranspiration rates (Ashrafzadeh *et al.* 2020) are some examples. In this study, the capability of two kernel based (KB) models (i.e. SVM, KELM) was investigated for scour depth assessing downstream of ski-jump bucket spillways. In this regard, two states were considered and, using dimensional parameters (sate 1) and non-dimensional parameters (state 2), several models were developed and tested. The capability of applied methods was compared with some available semi-empirical equations. In addition, the most important parameters were determined using sensitivity analysis. Also, Monte Carlo uncertainty analysis was applied to investigate the dependability of the applied models. In addition, the capability of SVM and KELM approaches was compared with Hybrid Multilayer Perceptron Firefly Algorithm (MLP-FFA) as new artificial intelligence approach.

MATERIALS AND METHODS

Used datasets

For determining the accuracy of the developed models laboratory investigations carried out at the Central Water and Power Research Station (CWPRS), India were used. Several experiments were performed considering different amounts for discharges and reservoir levels. The spillway gates were fully and partially open. The standing wave flume was used for hydraulic model discharge measurement. The sectional models were scaled to the range of 1:40–1:60, whereas comprehensive models had their scales varying from 1:50 to 1:100. A look at these observations revealed that additional measurements were necessary to make them more comprehensive; especially with respect to pattern of scour including width and distance of maximum scour depth from the spillway bucket lip (length). New hydraulic model studies were therefore conducted on three different bucket designs. The three hydraulic models simulated the dams across rivers Subarnarekha, Ranganadi, and Parbati Rivers in India.

The first dam was 52 m high and 720 m long. Its spillway consisted of 13 spans of 15 m wide each with crest at elevation 177 m. Radial gates of size 15 m16 m regulated the flow over this spillway. The design outflow flood was 26,150 m^3/s. This corresponded to a maximum water level at an elevation of 192.37 m. The ski-jump bucket with bucket radius of 25 m and lip angle of 32.5° was provided at the toe for energy dissipation. The second (Ranganadi River) dam was 60 m high, made up of concrete with a rockfill portion on its right side. It had an overflow spillway with seven spans of 10 m width and 12 m height. The spillway catered to a maximum outflow flood of 12,500 m^3/s. This corresponded to the maximum water level of 568.3 m and the full reservoir level of 567 m with the crest level of the spillway at 544 m. The ski-jump bucket modeled by a 1:60 scale model served as an energy dissipator at the toe of the spillway. It had a bucket radius of 18 m with 35° as the lip angle. The dam corresponding to the third spillway was 85 m high. It was

designed to pass a maximum discharge of 1,850 m^3/s at the full reservoir level of 2,198 m elevation. It had three spans, 6 m wide and 9 m high, separated by 6 m thick piers, and fitted with radial gates. An apron and a plunge pool along the downstream side fronted the bucket, which had a bucket radius of 28 m with the lip angle of 30°. This model based on Froude's law had a scale of 1:50. The downstream bed was made up of 2 mm diameter cohesionless sand particles. The riverbanks in this portion were assumed to be nonerodible and rigid. The various depths such as tail water depth, head over crest and other parameters were measured by using a point gauge having a graduation of 0.1 mm. The depth of scour was observed in a free formed plunge pool which was subsequently filled with sand having d50 size of 2 mm. Observations were made with four discharge passes (25, 50, 75, 100% of the maximum discharge) each with fully open as well as partially open gates. Experience showed that the equilibrium scour depth would be reached within this period, although the evolution of progressive scour depth is a function of time. In the end, 95 input-output pairs were compiled. Table 1 shows the range of data used in these experiments. The parameters mentioned in this table are Ys: scour depth, H_1: total head, Ls: scour length, q: unit discharge of spillway, φ: bucket lip angle, Yt: tail water depth, R: bucket radius, and d_{50} mean sediment size. Also, Ranganadi dam and the schematic view of spillway and scour hole notations are shown in Figure 1.

Kernel based methods

Kernel based approaches are new methods which are used for classification and regression purposes. Kernel based approaches are based on statistical learning theory initiated and can be used for modeling the complex and non-linear phenomenon. Two important KB approaches are KELM and SVM which work based on different kernal types such as linear, polynomial, radial basis function (RBF) and sigmoid functions in SVM and linear, polynomial, and RBF in KELM (Saghebian *et al.* 2020).

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

	Parameter							
Data range	**Q ($m^3/s/m$)**	**R (m)**	**Φ (rad)**	**H_1 (m)**	**Yt (m)**	**Ls (m)**	**Ys (m)**	**d_{50} (m)**
Minimum	0.0089	0.1	0.126	0.279	0.028	0.42	0.051	0.002
Maximum	0.204	0.61	0.78	1.79	0.265	2.24	0.55	0.008

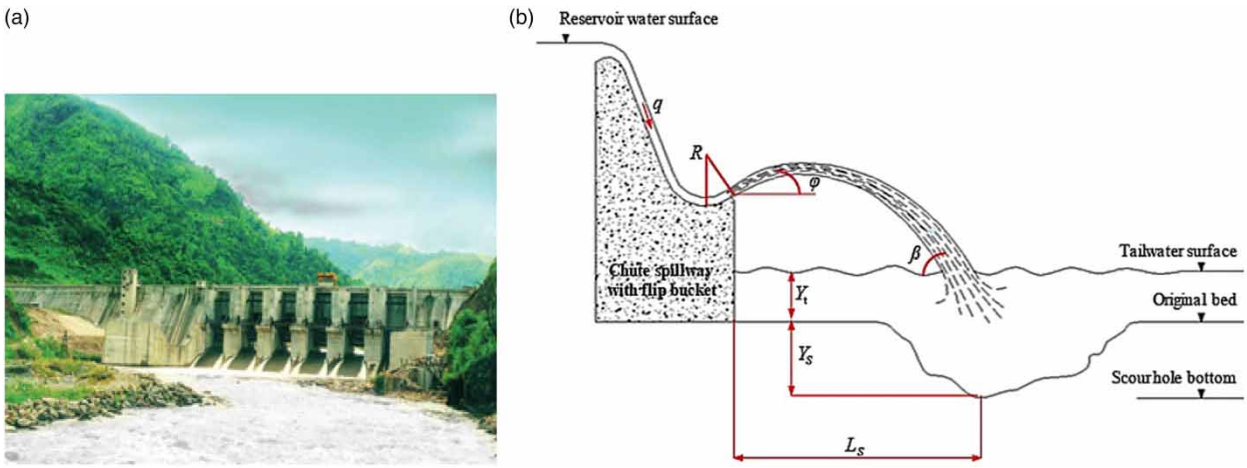


Figure 1 | Ranganadi dam (a) and spillway and scour hole notations (b).

Kernel Extreme Learning Machine (KELM)

Extreme Learning Machine (ELM) is a Single Layer Feed Forward Neural Network (SLFFNN) preparing method initially introduced by Huang *et al.* (2006). SLFFNN is a straight framework where information weights linked to hidden neurons and hidden layer biases are haphazardly chosen, while the weights among the hidden nodes are resolved logically. This strategy likewise has preferred execution and adapts progressively over the bygone era learning methods (Huang *et al.* 2006), as everyone does not like traditional techniques that involve numerous variables to setup, to demonstrate a complex issue this technique does not need much human involvement to accomplish ideal parameters. The standard single-layer neural system with N random information (a_i,b_i) (where $a_i = [a_{i1}, a_{i2}, \ldots, a_{in}]^T \in R^n$, $b_i = [b_{i1}, b_{i2}, \ldots, b_{im}]^T \in R^m$), M hidden neurons, and the active function f(a) is shown as follows:

$$\sum_{i=1}^{M} \varphi_i f(w_i a_j + c_i) = O_j \quad j = 1, 2, \ldots, N \tag{5}$$

where $w_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^T$ is the weight vector that joins the input layer to the hidden layer, $\varphi_i = [\varphi_{i1}, \varphi_{i2}, \ldots, \varphi_{im}]^T$ is the weight vector that joins the hidden layer to the target layer. c_i shows the hidden neuron biases. The general SLFFNN network aim is to minimize the difference between the predicted (O_j) and target (t_j) values which can be expressed as below:

$$\sum_{i=1}^{M} \varphi_i f(w_i \alpha_j + c_i) = T_j \quad j = 1, 2, \ldots, N \tag{6}$$

Equation (6) can be summarized as:

$$H\varphi = T \tag{7}$$

where

$$H(w_1, \ldots, w_M, c_1, \ldots, c_M, a_1, \ldots, a_N) = \begin{bmatrix} f(w_1 a_1 + c_1 \ldots f(w_M a_M + c_M) \\ \ldots \\ f(w_1 a_N + c_1 \ldots f(w_M a_M + c_M) \end{bmatrix} \tag{8}$$

$$T = \begin{bmatrix} T_1^T \\ . \\ . \\ T_M^T \end{bmatrix}, \varphi = \begin{bmatrix} \varphi_1^T \\ . \\ . \\ \varphi_M^T \end{bmatrix} \tag{9}$$

Matrix T is identified as the target matrix of the hidden layers of the neural network. H is considered as output matrix of neural network. Huang *et al.* (2012) also introduced kernel functions in the design for ELM. Now some kernel functions are used in the design of ELM such as linear, radial basis, normalized polynomial, and polynomial kernel functions. Kernel function based ELM design is named KELM. For more details about KELM, readers and researchers are referred to Huang *et al.* (2012).

Support Vector Machine (SVM)

Support Vector Machines as structural risk minimization (SRM) methods minimize an upper boundary on the expected risk (Vapnik 1995). According to Ji *et al.* (2017), this approach is applied for information categorization and dataset classification and regression (see Figure 2). The SVMs are based on the concept of the optimal hyper plane that separates samples of two classes by considering the widest gap between two classes. Support Vector Regression (SVR) is an extension of SVM regression. In the use of SVMs for regression aims, we tried to obtain a function with the most deviation from the actual target vectors for all given training data. For the nonlinear SVR the kernel function concept is used (see Vapnik 1995 for more details).

Hybrid Multilayer Perceptron Firefly Algorithm (MLP-FFA)

The MLP-FFA method is the integration of Multilayer Perceptron (MLP) and Firefly Algorithm (FFA) models. The

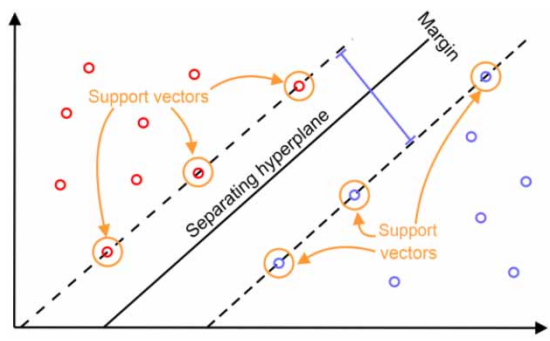


Figure 2 | Graphical presentation of the SVM classification.

MLP is a class of feed-forward artificial neural networks (Ghorbani *et al.* 2018). The FFA as a swarm intelligence optimization method is conceptually based on the fireflies motion. This method is applied in strategic searching for optimal parameters of the MLP model. The MLP-FFA modeling flowchart is shown in Figure 3. Formulating the objective function and modified the light intensity is the main aim of MLP-FFA model designing. Equations (10)–(12) can be applied to calculate the light intensity $I(r)$, attractiveness η, and the cartesian distance between any two i and j fireflies (Ghorbani *et al.* 2018):

$$I(r) = I_O \exp(-\gamma \times r^2) \tag{10}$$

$$\eta(r) = \eta_O \exp(-\gamma \times r^2) \tag{11}$$

$$r_{ij} = \|x_i + x_j\| = \sqrt{\sum_{K=1}^{d} (x_{i,k} - x_{j,k})} \tag{12}$$

where $x_{i,k}$, γ, d, $I(r)$, I_o, respectively, are the *kth* component of the spatial coordinate *xi* of the *ith* firefly, the coefficient of light absorption, the given problem dimensionality; the light intensity at distance r, and initial light intensity from a firefly. Also, $\eta(r)$ and η_O are the attractiveness η at a distance r and $r = 0$, respectively. The next firefly i coordinate is expressed as below:

$$x_i^{(i+1)} = x_i + \Delta x_i \tag{13}$$

$$\Delta x_i = \eta_o e^{-\gamma \times r^2}(x_j - x_i) + \chi \varepsilon_i \tag{14}$$

In Equation (14), the attraction effect is shown using the first term, and randomization is shown using the second term. χ is the randomization coefficient. The χ value varies from 0 to 1 (for the current study $\chi = 0.5$). ε_i is the random number vector. This parameter is obtained from a Gaussian distribution (ε_i is 0.96 in this study).

Performance criteria

In this study, two criteria were used to assess the applied models capability including Determination Coefficient (DC) and Root Mean Square Errors (RMSE). These

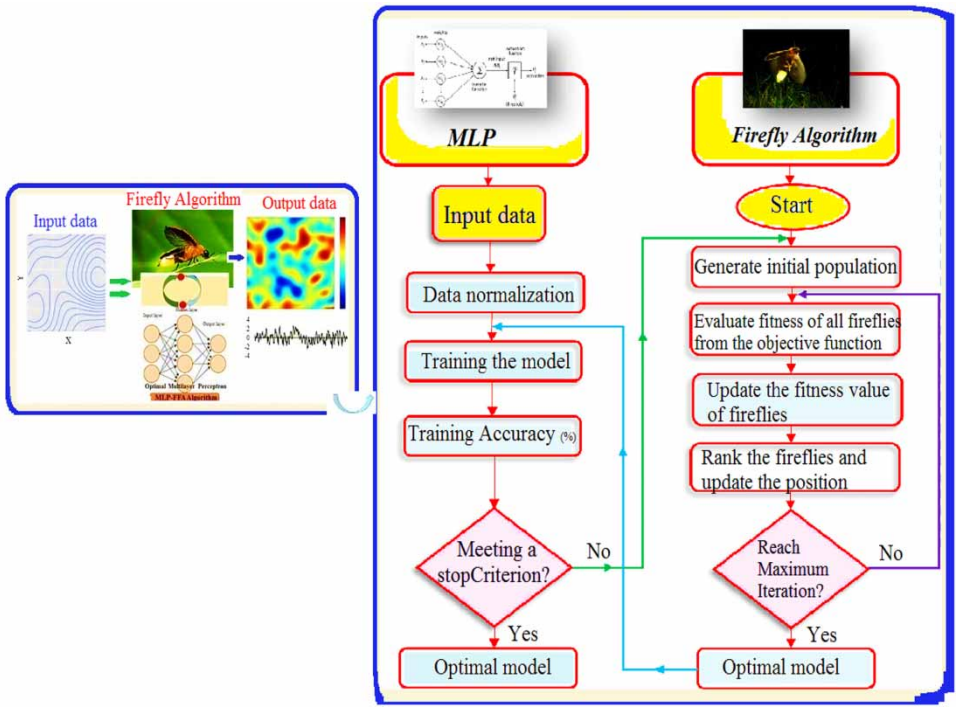


Figure 3 | A schematic view of the MLP-FFA method.

statistical criteria are formulated as:

$$DC = 1 - \frac{\sum_{i=1}^{N}(l_o - l_p)^2}{\sum_{i=1}^{N}(l_o - \overline{l_p})^2}, R = \frac{\sum_{i=1}^{N}(l_o - \overline{l_o}) \times (l_p - \overline{l_p})}{\sqrt{\sum_{i=1}^{N}(l_o - \overline{l_o})^2 \times (l_p - \overline{l_p})^2}},$$

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{(l_o - l_p)^2}{ND}} \quad (15)$$

where l_o, l_p, $\overline{l_o}$, $\overline{l_p}$, and ND, respectively, show the observed, predicted, mean observed, mean predicted values and data number. It should be noted that in this study all input variables were scaled between 0 and 1 in order to eliminate the input and output variables dimensions.

Uncertainty analysis

The aim of a model uncertainty analysis is to determine the statistical characteristics of the outputs of that model as a function of the uncertainty of the input parameters (Noori *et al.* 2015). Uncertainty is a factor associated with the estimation result which determines the estimation values range. Its value indicates the level of confidence in which the actual measured value falls, within the specified range (Noori *et al.* 2015). In the current research, the Monte Carlo method proposed by Abbaspour *et al.* (2007) was used to evaluate the uncertainty of the AI models in SPEIs series modeling. To verify model results uncertainties, 95% confidence interval (95PPU) and bandwidth factor (d-factor) which is the average distance between the upper (XU) and lower (XL) uncertainty bands should be used (Noori *et al.* 2015). In this regard, the considered model should be developed many times (1000 in this research), and the empirical cumulative distribution probability of the models should be calculated.

The appropriate confidence limits are mostly from measured data within the width of 95PPU and have a reasonable average width (d-factor → 0) (Abbaspour *et al.* 2007). For evaluating the mean width of the confidence band Abbaspour *et al.* (2007) suggested the below equation:

$$\text{d} - \text{factore} = \frac{\overline{dx}}{\sigma x} \quad (16)$$

where σ_x and $\overline{dx}$ are the observed data standard deviation and the confidence band's average width, respectively. The percentage of the data within the confidence band of 95% is calculated as:

$$\text{Bracketed by 95PPU} = \frac{1}{k} Cont(j|X_L < X_{reg} < X_U \quad (17)$$

where 95PPU shows 95% predicted uncertainty; k shows the number of observed data and X_{reg} shows the current registered data.

Models development

Input variables

Based on the experimental studies (Mason 1984; Lopardo *et al.* 2002) the important variables which affect the depth of scour can be a function of the parameters:

$$Y_s = f(q, H_1, R, \varphi, Y_t, d_{50}, g, \rho_s, \rho_w) \quad (18)$$

which g: gravity acceleration, ρ_w: density of water, ρ_s: density of sediment particles. Using dimensional analysis (13) can be expressed as following:

$$\frac{Y_s}{Y_t} = f\left(\frac{q^2}{gY_t^3}, \frac{d_{50}}{Y_t}, \frac{R}{Y_t}, \varphi, \frac{H_1}{Y_t}, \frac{\rho_s}{\rho_w}\right) \quad (19)$$

the ratio of ρ_s/ρ_w is constant and can be eliminated from the modeling process. Two states of modeling were considered. In the first state, dimensional parameters and in the second state, non-dimensional parameters were applied. Table 2 shows the developed models in this study. It should be noted that 75% of the whole dataset were used for training the models, and 25% data were used for testing the models. In addition, the applied models were run via coding in MATLAB software.

RESULTS AND DISCUSSION

Selecting the appropriate kernel types of applied approaches

Each artificial intelligence method has its own settings and parameters and, to achieve the desired results, the optimized

Table 2 | Developed models for predicting E_L/E_1

State 1			State 2		
Model	**Input variable(s)**	**Input variable**	**Model**	**Input variable(s)**	**Input variable**
D(I)	d_{50}, q, H_1, R, φ, Y_t	Y_s	N(I)	$\frac{q^2}{gY_t^3}$	Y_s/Y_t
D(II)	q, H_1, R, φ, d_{50}	Y_s	N(II)	$\frac{q^2}{gY_t^3}$, $\frac{d_{50}}{Y_t}$	Y_s/Y_t
D(III)	H_1, R, φ, d_{50}	Y_s	N(III)	$\frac{q^2}{gY_t^3}$, $\frac{d_{50}}{Y_t}$, $\frac{H_1}{Y_t}$	Y_s/Y_t
D(V)	q, R, φ, d_{50}	Y_s	N(IV)	$\frac{q^2}{gY_t^3}$, $\frac{d_{50}}{Y_t}$, $\frac{R}{Y_t}$, φ, $\frac{H_1}{Y_t}$	Y_s/Y_t
			N(V)	$\frac{q^2}{gY_t^3}$, $\frac{d_{50}}{Y_t}$, φ	Y_s/Y_t
			N(VI)	$\frac{q^2}{gY_t^3}$, φ, $\frac{H_1}{Y_t}$	Y_s/Y_t

amount of these parameters should be determined. In SVM and KELM, designing the selection of appropriate type of kernel function is important. In this section of the paper, SVM and KELM methods were evaluated using model D(I) in order to select the best kernel functions of each model. Figure 4 indicates the percentage error of statistical RMSE parameter for model D(I) with different kernel functions. From the results, it was found that, among all kernel functions, the RBF yielded more accurate outcomes. Therefore, RBF was selected as a core tool of SVM and KELM methods and used in the rest of the models.

Developed models base on dimensional parameters

In the first state, several dimensional parameters were used for model development. These models were tested via SVM and KELM methods. The obtained results are shown in Table 3 and Figure 5. According to the performance criteria results, it could be deduced that the model *D(I)*, with a combination of q, H_1, R, φ, Y_t, and d_{50} as inputs, performed more successfully. Considering the four developed models results, it was found that using the combination of Y_t and q parameters, the models efficiency increased. Also, it was observed that the effect of the q variable on increasing the accuracy of the models was more than the H_1 variable. Based on Table 3 results, the efficiency of the KELM model in predicting the downstream scour depth of ski-jump bucket spillways was higher than for the SVM approach.

Developed models base on non-dimensional parameters

In the second state, several non-dimensional parameters were applied for model development. The results obtained for this state are indicated in Table 4 and Figure 6. The results showed that the model *N(IV)* with a combination of $\frac{q^2}{gY_t^3}$, $\frac{d_{50}}{Y_t}$, $\frac{R}{Y_t}$, φ, $\frac{H_1}{Y_t}$ parameters had the superior performance. However, the model *N(V)* with parameters $q^2/[gY_t^3]$, d_{50}/Y_t, and φ showed approximately the same results, therefore, this model with only three inputs could be selected as the optimum model. Based on the results, it could be seen that the use of $q^2/[gY_t^3]$, d_{50}/Y_t, R/Y_t, and φ parameters as inputs increased the applied method capabilities in the scour depth modeling.

Sensitivity analysis

In this section, the most effective variables in predicting the scours depth of ski-jump bucket spillways was assessed using sensitivity analysis. In this regard, for each state, the superior models of KELM method were selected and re-run by removing each input variable. Table 5 and Figure 7 show the sensitivity analysis results. Based on Table 5, it could be induced that q in the state 1 and $q^2/[gY_t^3]$ in the state 2 were the most important variables in estimating the scour depth downstream of ski-jump bucket spillways.

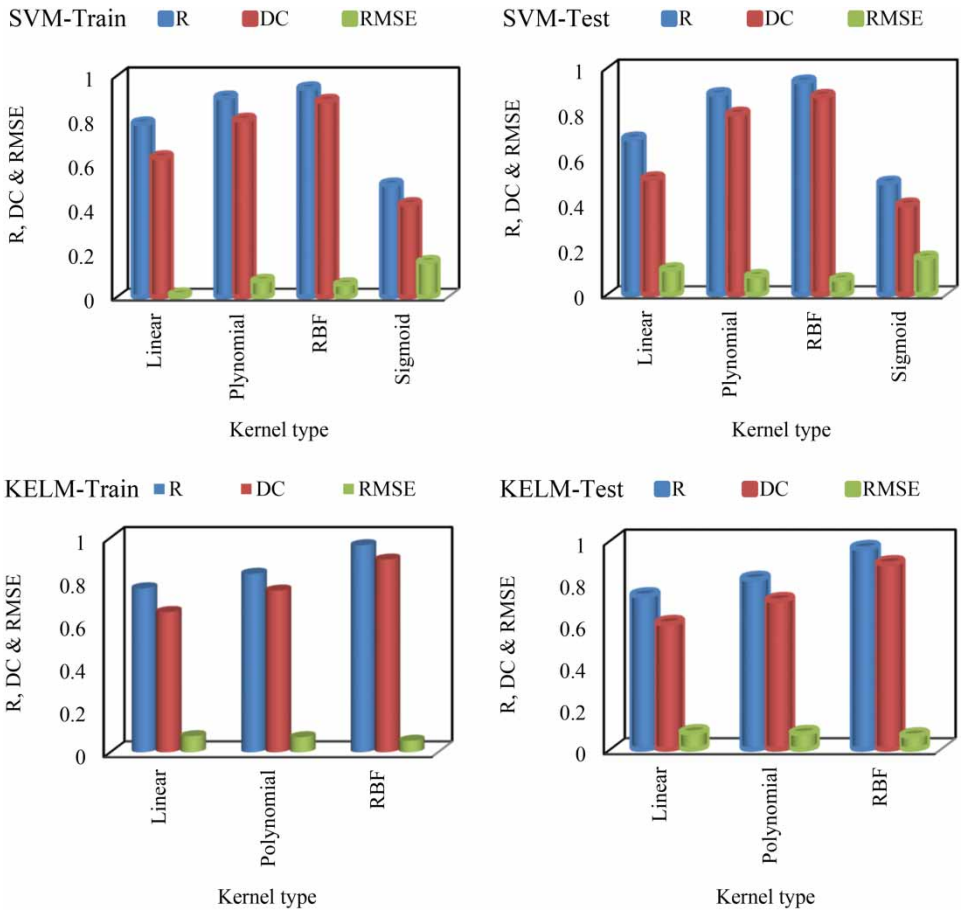


Figure 4 | Statistical parameters of SVM and KELM models with different kernel functions for the model D(I).

Table 3 | Statistical parameters of the SVM and KELM models for the state 1 modeling

		Performance criteria					
		Traina			Testa		
Method	Model	R	DC	RMSE	R	DC	RMSE
SVM	**D(I)**	**0.938**	**0.886**	**0.057**	**0.939**	**0.876**	**0.067**
	D(II)	0.919	0.842	0.061	0.918	0.838	0.077
	D(III)	0.732	0.662	0.098	0.538	0.481	0.165
	D(V)	0.942	0.835	0.058	0.912	0.806	0.085
KELM	**D(I)**	**0.966**	**0.899**	**0.054**	**0.962**	**0.889**	**0.064**
	D(II)	0.942	0.855	0.064	0.941	0.851	0.073
	D(III)	0.750	0.672	0.093	0.551	0.501	0.157
	D(V)	0.953	0.848	0.055	0.935	0.818	0.080

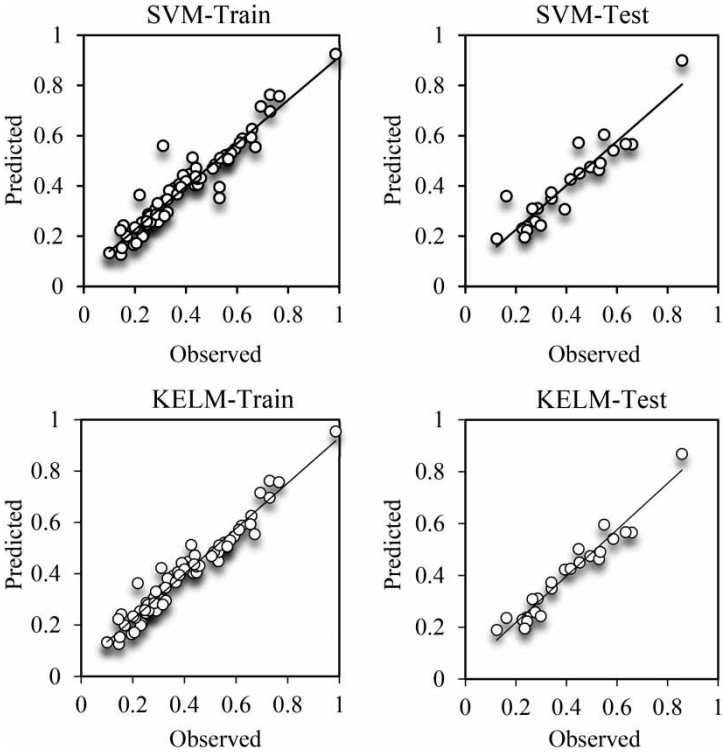


Figure 5 | Comparison of the observed and predicted scour depth for superior models of the state 1.

Table 4 | Statistical parameters of the SVM and KELM models for the state 2 modeling

		Performance criteria					
		Train			Test		
Method	Model	R	DC	RMSE	R	DC	RMSE
SVM	N(I)	0.697	0.508	0.121	0.515	0.315	0.168
	N(II)	0.833	0.756	0.088	0.609	0.471	0.155
	N(III)	0.841	0.721	0.094	0.513	0.344	0.163
	N(IV)	**0.888**	**0.818**	**0.078**	**0.812**	**0.769**	**0.110**
	N(V)	0.882	0.807	0.079	0.816	0.765	0.112
	N(VI)	0.861	0.780	0.084	0.767	0.693	0.134
KELM	N(I)	0.710	0.515	0.115	0.525	0.319	0.163
	N(II)	0.848	0.766	0.084	0.621	0.477	0.147
	N(III)	0.856	0.730	0.089	0.523	0.328	0.161
	N(IV)	**0.904**	**0.829**	**0.073**	**0.830**	**0.791**	**0.107**
	N(V)	0.898	0.818	0.075	0.832	0.778	0.109
	N(VI)	0.876	0.790	0.080	0.782	0.702	0.127

Validation of proposed best SVM models using ANFIS and ANN

The experimental data were used to evaluate the performance of proposed best applied models compared with other data driven models. In this regard, for each state (i.e. modeling base on dimensional and non-dimensional parameters) the superior model was run using the MLP-FFA model and the results were compared with SVM and KELM. Table 6 shows the results of this comparison.

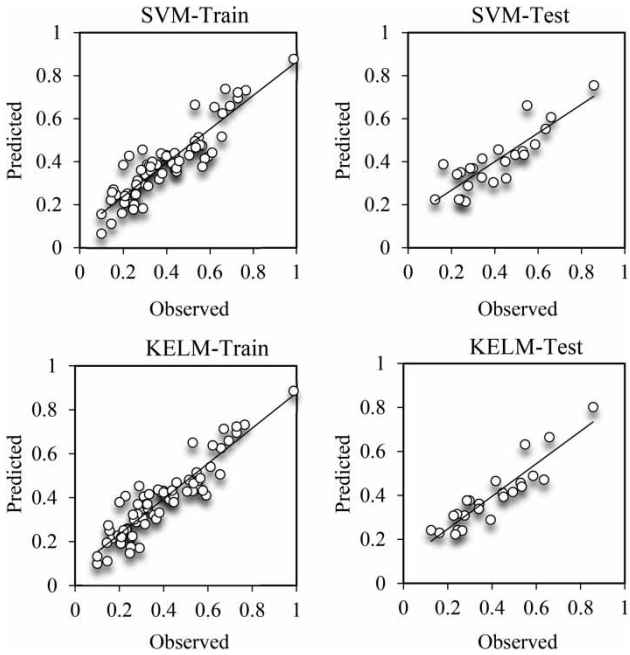


Figure 6 | Comparison of the observed and predicted scour depth for superior models of the state 2.

As seen from Table 5, the MLP-FFA model led to desired accuracy and the efficiency of this model was more than the SVM. However, the KELM model slightly yielded better results in comparison with SVM and MLP-FFA models.

Investigating the capability of several available scour depth prediction equations

The capability of several existing scour depth equations was evaluated and the results were compared with KB models. The RMSE criterion was used as an indication of the accuracy of the equations. The results are shown on Figure 8. Based on the results, among all used formulas, Chee & Kung's (1974) equation led to reasonable accuracy. The correlation between models and observed values was rather good for the Y_s smaller values. However, Figure 8 showed that in general, the applied formulas were not so successful in scour depth modeling, while the SVM and KELM results were close to the observed data. These models had the lowest RMSE and this proved that the KB methods are efficient approaches in scour depth modeling of ski-jump bucket spillways.

Table 5 | Relative significance of each of input parameters of the best models for each state

		Performance criteria for test series		
Model	**Eliminated variable**	**R**	**DC**	**RMSE**
State 1 D(I)	D(I)	0.962	0.889	0.064
	d_{50}	0.915	0.838	0.077
	q	**0.724**	**0.601**	**0.119**
	H_1	0.928	0.871	0.069
	R	0.901	0.832	0.078
	φ	0.834	0.802	0.082
	Y_t	0.902	0.833	0.078
State 2 N(V)	N(V)	0.832	0.744	0.109
	$\mathbf{q^2/[gY_t^3]}$	**0.598**	**0.328**	**0.168**
	φ	0.607	0.466	0.159
	d_{50}/ Y_t	0.609	0.471	0.155

The results of uncertainty analysis

In this part of the study the uncertainty analysis was done in order to find the uncertainty of the superior KELM

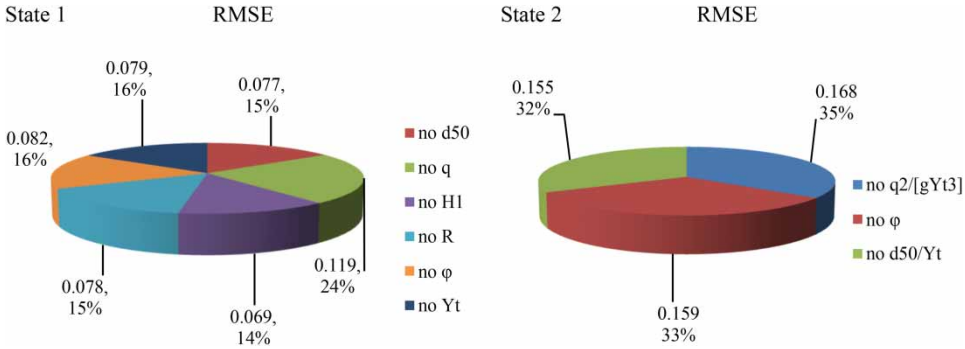


Figure 7 | Comparison of statistical parameters obtained from sensitivity analysis.

Table 6 | Statistical parameters of the SVM, KELM, and MLP-FFA models for the superior models

		Performance criteria					
		Train			Test		
Method	Model	R	DC	RMSE	R	DC	RMSE
Modeling base on dimensional parameters							
SVM	D(I)	0.938	0.886	0.057	0.939	0.876	0.067
KELM	D(I)	0.966	0.899	0.054	0.962	0.889	0.064
MLP-FFA	D(I)	0.955	0.891	0.055	0.954	0.883	0.066
Modeling base on non-dimensional parameters							
SVM	N(IV)	0.888	0.818	0.078	0.812	0.769	0.110
KELM	N(IV)	0.904	0.829	0.073	0.830	0.791	0.107
MLP-FFA	N(IV)	0.911	0.823	0.075	0.828	0.790	0.108

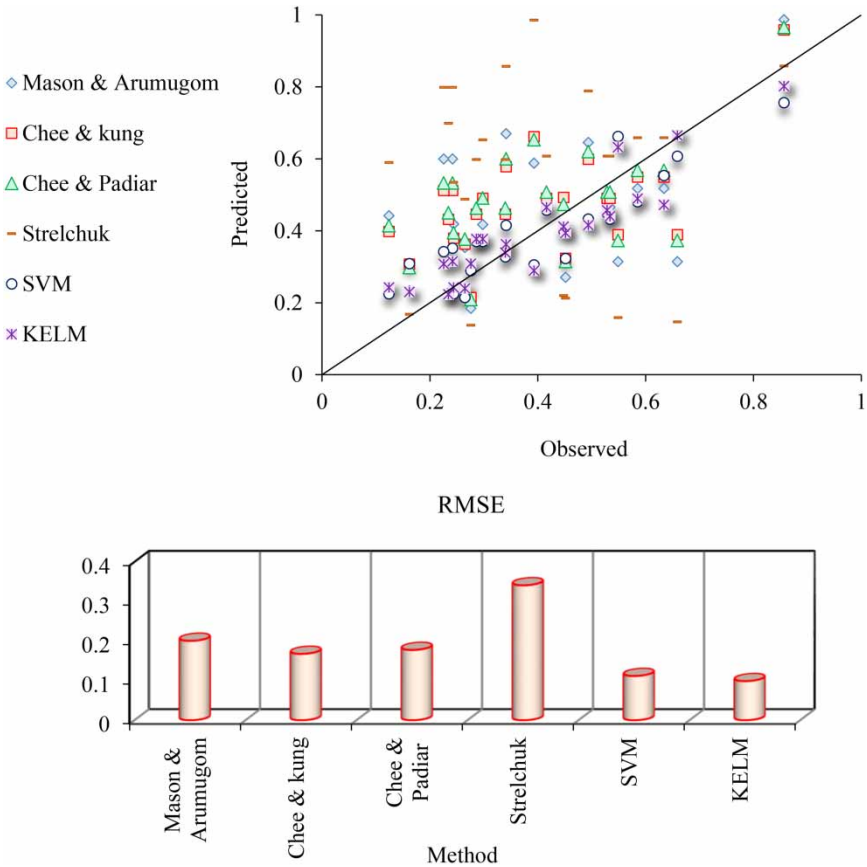


Figure 8 | Comparison of prediction from proposed equations and the best SVM and KELM models.

model. In the proper confidence level, two important indices should be considered. First: the 95PPU band brackets most of the observations, and second: the d-Factor is smaller than the standard deviation of the observed data. The two mentioned indices were applied to account for input uncertainties. The obtained results for the uncertainty

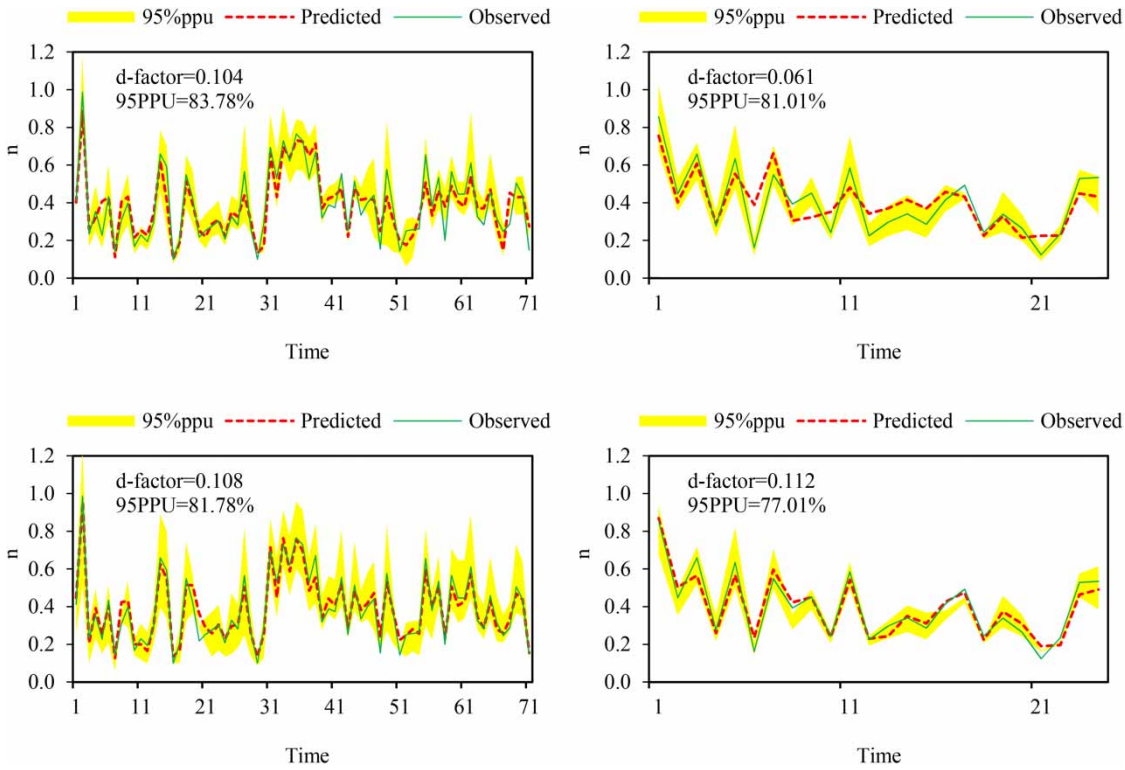


Figure 9 | Uncertainty analysis for the best model of the KELM method.

analysis are shown in Figure 9. Based on the values obtained for the d-Factor and 95%PPU, it could be indicated that for both considered states the observed and predicted values were within the 95PPU band in most of the cases. Also, it was found that the amount of d-Factors for train and test datasets was smaller than the standard deviation of the observed data. Therefore, it could be induced that scour depth modeling downstream of the ski-jump bucket spillway via KELM model led to an allowable degree of uncertainty.

CONCLUSION

Scour depth accurate prediction downstream of spillways is an important issue due to its impact on such structures performances. In this study, the efficiency of two KB methods was investigated in scour depth modeling of ski-jump bucket spillways. In the model developement process, based on dimensional (state 1) and non-dimensional (state 2) parameters, two states were tested. The results showed that in the state 1 model with input parameters q, H_1, R, φ, Y_t, and d_{50} and in the state 2 model with input parameters of $q^2/[gY_t^3]$, d_{50}/Y_t, R/Y_t, H_1/Y_t, and φ performed more successful. It was observed that using Y_t, q, $q^2/[gY_t^3]$, $d_{50}/\ Y_t^3$, $R/\ Y_t$, and φ as inputs enhanced the model efficiency. Also, comparison among applied kernel based models and some available equations revealed that the SVM and KELM models were more reliable in modeling the scour depth downstream of spillways. The results of sensitivity analysis indicated that the impact of variable q in the state 1 and variable $q^2/[gY_t^3]$ in the state 2 on obtaining a model with higher accuracy were more than other used parameters. A comparison was performed between the SVM and KELM results and the MLP-FFA method. The results showed that KELM is more accurate than the other two intelligence approaches. Also, the superior applied models dependability was investigated using uncertainty analysis. The results showed that the KELM model had an allowable degree of uncertainty in scour depth modeling downstream of the ski-jump bucket spillway.

DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories at http://cwprs.gov.in/.

REFERENCES

Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J. & Srinivasan, R. 2007 Modelling hydrology and water quality in the prealpine/alpine Thur watershed using SWAT. *J. Hydrol.* **333** (2), 413–430.

Al-Shammari, E. T., Mohammadi, K., Keivani, A., Ab Hamid, S. H., Akib, S., Shamshirband, S. & Petković, D. 2016 Prediction of daily dewpoint temperature using a model combining the support vector machine with firefly algorithm. *J. Irrig. Drain. Eng.* **142** (5), 04016013.

Ashrafzadeh, A., Kişi, O., Aghelpour, P., Biazar, S. M. & Masouleh, M. A. 2020 Comparative study of time series models, support vector machines, and GMDH in forecasting long-term evapotranspiration rates in northern Iran. *J. Irrig. Drain. Eng.* **146** (6), 04020010.

Azamathulla, H. M. & Wu, F. C. 2011 Support vector machine approach for longitudinal dispersion coefficients in natural streams. *Appl. Soft Comput.* **11**, 2902–2905.

Azamathulla, H. M., Haghiabi, A. H. & Parsaie, A. 2017 Prediction of side weir discharge coefficient by support vector machine technique. *Water Supply* **16** (4), 1002–1016.

Chang, W. Y., Lai, J. S. & Yen, C. L. 2004 Evolution of scour depth at circular bridge piers. *J. Hydraul. Eng.* **130** (9), 905–913.

Chee, S. P. & Kung, T. 1974 Piletas de derrubio autoformadas. In *6th Latin American Congress of the International Association for Hydraulic Reasearch*, Bogata, Columbia.

Chee, S. P. & Padiyar, P. V. 1969 Erosion at the base of flip buckets. *Eng. J.* **52** (111), 22–24.

Damle, P. M., Venkatraman, C. P. & Desai, S. C. 1966 Evaluation of scour below skijump buckets of spillways. *CWPRS Golden Jubilee Symposia*, Poona, India, **1**, 154–163.

Dargahi, B. 2003 Scour development downstream of a spillway. *J. Hydraul. Res.* **41** (4), 417–426.

Ghorbani, M. A., Deo, R. C., Karimi, V., Yaseen, Z. M. & Terzi, O. 2018 Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey. *Stoch. Environ. Res. Risk Assess.* **32** (6), 1683–1697.

Hoffmans, G. J. C. M. 1998 Jet scour in equilibrium phase. *J. Hydraul. Eng.* **124** (4), 430–437.

Huang, G. B., Zhu, Q. Y. & Siew, C. K. 2006 Extreme learning machine: theory and applications. *Neurocomputing* **70** (1–3), 489–501.

Huang, G. B., Zhou, H., Ding, X. & Zhang, R. 2012 Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **42** (2), 513–529.

Ji, J., Zhang, C., Gui, Y., Lü, Q. & Kodikara, J. 2017 New observations on the application of LS-SVM in slope system reliability analysis. *J. Comput. Civ. Eng.* **31** (2), 06016002.

Lopardo, R. A., Lopardo, M. C. & Casado, J. M. 2002 Local rock scour downstream large dams. In *International Workshop on Rock Scour due to High Velocity Jets*, Lausanne, Switzerland, pp. 55–58.

Martins, R. B. F. 1975 Scouring of rocky river beds by free jet spillways. *Int. Water Power and Dam Construction, England* **27** (5), 152–153.

Mason, P. J. 1984 Erosion of plunge pools downstream of dams due to the action of free trajectory jets. *Proc. Inst. Civ. Eng.* **76**(5), 523–537.

Mason, P. J. & Arumugam, K. 1985 Free jet scour below dams and flip buckets. *J. Hydraul. Eng.* **1119** (2), 220–235.

Noori, R., Deng, Z., Kiaghadi, A. & Kachoosangi, F. T. 2015 How reliable are ann, anfis, and svm techniques for predicting longitudinal dispersion coefficient in natural rivers? *J. Hydraul. Eng.* **142** (1), 04015039.

Pandhiani, S. M., Sihag, P., Shabri, A. B., Singh, B. & Pham, Q. B. 2020 Time-series prediction of streamflows of Malaysian rivers using data-driven techniques. *J. Irrig. Drain. Eng.* **146** (7), 04020013.

Roushangar, K., Ghasempour, R. & Biukaghazadeh, S. 2020 Evaluation of the parameters affecting the roughness coefficient of sewer pipes with rigid and loose boundary conditions via kernel based approaches. *Int. J. Sediment Res.* **35** (2), 171–179.

Saghebian, S. M. 2019 Predicting the relative energy dissipation of hydraulic jump in rough and smooth bed compound channels using SVM. *Water Supply.* **19** (4), 1110–1119.

Saghebian, S. M., Roushangar, K., Ozgur Kirca, V. S. & Ghasempour, R. 2020 Modeling total resistance and form resistance of movable bed channels via experimental data and a kernel-based approach. *J. Hydroinf.* **22** (3), 528–540.

Strelchuk, D. L. 1969 *Scour at the Base of Spillway Buckets*. M.S Thesis, University of Windsor. Ontario.

Vapnik, V. 1995 *The Nature of Statistical Learning Theory*. Data Mining and Knowledge Discovery, pp. 1–47.

Zhu, S., Luo, X., Xu, Z. & Ye, L. 2018 Seasonal streamflow forecasts using mixture-kernel GPR and advanced methods of input variable selection. *Hydrol Res.* **50** (1), 200–214.

First received 18 November 2020; accepted in revised form 24 February 2021. Available online 10 March 2021