Water Supply

IWA PUBLISHING

Water Supply Vol 22 No 1, 44 doi: 10.2166/ws.2021.289

Uncertainty analysis regarding evaluating effective parameters on the hydraulic jump characteristics of different shape channels

Kiyoumars Roushangar[a,b,*], Farzin Homayounfar[c] and Roghayeh Ghasempour[a]

[a] Department of Water Resource Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran
[b] Center of Excellence in Hydroinformatics, University of Tabriz, Tabriz, Iran
[c] Department of Water Resource Engineering, Faculty of Civil and Environmental Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
*Corresponding author. E-mail: kroshangar@yahoo.com

FH, 0000-0002-3949-4656

ABSTRACT

The hydraulic jump phenomenon is a beneficial tool in open channels for dissipating the extra energy of the flow. The sequent depth ratio and hydraulic jump length critically contribute to designing hydraulic structures. In this research, the capability of the Support Vector Machine (SVM) and Gaussian Process Regression (GPR) as kernel-based approaches was evaluated to estimate the features of submerged and free hydraulic jumps in channels with rough elements and various shapes, followed by comparing the findings of the GPR and SVM models and the semi-empirical equations. The results represented the effect of the geometry (i.e., steps and roughness elements) of the applied appurtenances on hydraulic jump features in channels with appurtenances. Moreover, the findings confirmed the significance of the upstream Froude number in the estimating of sequent depth ratio in submerged and free hydraulic jumps. In addition, the immersion was the highest contributing variable regarding the submerged jump length on sloped smooth bed and horizontal channels. Based on the comparisons among kernel-based approaches and the semi-empirical equations, kernel-based models showed better performance than these equations. Finally, an uncertainty analysis was conducted to assess the dependability of the best applied model. The results revealed that the GRP model possesses an acceptable level of uncertainty in the modeling process.

Key words: GPR, negative step, sequent depth ratio, submerged hydraulic jump, trapezoidal channel

HIGHLIGHTS

- SVM and GPR methods were selected for determining influential elements regarding predicting hydraulic jump characteristics in various shape channels.
- Experimental datasets were considered to feed the applied models.
- The dependability of the best applied model was evaluated using uncertainty analysis.

INTRODUCTION

The hydraulic jump as an inherent phenomenon happens when the flow changes from supercritical to subcritical type (Corry *et al.* 1975). The water surface promptly increases during this transition, forming surface rollers and leading to intense mixing. Then, the air is entrained while dissipating a huge deal of energy. Hydraulic jump, as an effective phenomenon in rivers, open channels, and energy dissipater systems, is typically employed for kinetic energy dissipation downstream of several hydraulic structures such as gates, drops, spillways, and chutes. Thus, estimating the jump properties such as the sequent depth ratio and hydraulic jump length plays an essential role in designing hydraulic structures. A hydraulic jump can happen in submerged or free circumstances on sloped and horizontal beds. Accordingly, further examination of some basic jump features such as jump length and sequent depth ratio is necessary. For instance, the hydraulic jump length is considered one of the most vital variables in the design of stilling basins.

Nevertheless, it cannot be determined only by mathematical calculations. Thus, laboratory and experimental results require assessment. Numerous experimental works have focused on hydraulic jump formation regarding the estimating

of sequent depth ratio and length of the jump. In their study, Abbaspour *et al.* (2009) investigated hydraulic jumps on corrugated beds. Results demonstrated that the hydraulic jumps on corrugated beds have a smaller length and tail-water depth compared with those on smooth beds. Likewise, Gupta *et al.* (2013) studied free hydraulic jump length and energy losses in a horizontal prismatic channel. Based on their findings, the relative length and energy losses of the free jump increased with increasing the Froude number, although it reduced with increasing the Reynolds number. Similarly, Ead & Rajaratnam (2002) assessed the hydraulic jump on corrugated beds and indicated that a jump on corrugated beds had a length of half the jump over smooth beds. They further found that the hydraulic jumps had a smaller tail-water depth over corrugated beds compared with smooth beds. In a study, Ahmed *et al.* (2014) evaluated the submerged hydraulic jump over a triangle strip rough bed. It was concluded that the corrugated beds have permanently superior behavior compared with smooth beds. Also the length and tail-water depth of a jump over a corrugated bed were smaller in comparison with smooth beds. Carollo *et al.* (2007) assessed hydraulic jump features on a bed roughened by strictly packed bottom-cemented crushed gravel particles and revealed that boundary roughness decreases the sequent and the length of the hydraulic jump depth. Recently, soft computing methods such as Adaptive Neuro-Fuzzy Inference System (ANFIS), Artificial Neural Networks (ANN), Gene Expression Programming (GEP), Genetic Programming (GP), and Gaussian Process Regression (GPR) have been utilized in open channel hydraulics and water resources studies. In this regard, machine learning approaches can be evaluated for different purposes such as capacity prediction for the reservoir (Foddis *et al.* 2015), groundwater level fluctuation prediction (Gholami *et al.* 2015; Rajaee *et al.* 2019; Di Nunno & Granata 2020; Pandey *et al.* 2020), and roughness coefficient and sediment transport in sewer pipes (Roushangar & Ghasempour 2017; Roushangar *et al.* 2019), and tide level predictions (Granata & Di Nunno 2021). Other areas of focus are modeling side weir discharge (Azamathulla *et al.* 2016; Azimi *et al.* 2017; Granata *et al.* 2019), predicting daily flows and stage-discharge (Wu *et al.* 2009; Kumar *et al.* 2020; Malik *et al.* 2020), lake water level (Yaseen *et al.* 2020), water quality (Wang 2016), and energy dissipation in channels with rough beds (Roushangar & Ghasempour 2018), evaluating soft computing methods for predicting river bed load (Azamathulla *et al.* 2010; Chiang & Tsai 2011), and forecasting downstream river flows (Chen *et al.* 2015).

Based on literature reviews, no research has comprehensively focused on predicting hydraulic jump features in various shape channels by kernel-based methods. Due to the nonlinearity and uncertainties of the hydraulic jump phenomenon, the existing regression methods often do not show the desired accuracy. The existing semi-empirical equations rely on a limited database, untested model assumptions, and a general lack of field data, and do not show the same results under variable flow conditions. Consequently, the applications of many of these methods are limited to special cases of their development, and therefore do not show uniform results under different conditions. On the other hand, the correct estimation of hydraulic jump characteristics is of great importance in hydraulic engineering, as it directly affects the planning, design and management of hydraulic structures. However kernel-based approaches based on quadratic optimization of convex functions can easily switch from linear to nonlinear separation. This is realized by nonlinear mapping using so-called kernel functions. Therefore, the present study employed kernel-based approaches to predict the variable of interest. Among artificial models, kernel-based approaches such as SVM and GPR are relatively new and important methods based on the different kernel types which are based on statistical learning theory. Such models are capable of adapting themselves to predict any variable of interest via sufficient inputs. The training of these methods is fast, they have high accuracy, and the probability of occurrence of data overtraining in these methods is less. Also, the most important principle of SVM in comparison with other techniques (such as ANN, ANFIS, GP, and GEP) is the application of minimizing an upper bound to the generalization error instead of minimizing the training error. The current study aimed to expand former studies by representing the capability of kernel-based methods (i.e., SVM and GPR) in order to predict submerged and free hydraulic jump features in channels with various bed conditions and shapes (i.e., trapezoidal, rectangular, expanding, sloped, and with negative step channels with rough and smooth beds). Several models were prepared under various input combinations (based on hydraulic characteristics and geometry of the channels and appurtenances) in order to find the most appropriate input combination for modeling hydraulic jump characteristics in suddenly expanding channels. Then the accuracy of the best applied models was compared with the accuracy of several semi-empirical equations. In addition, the most essential parameters in the prediction procedure were determined by utilizing sensitivity analysis, followed by employing Monte Carlo uncertainty analysis (UA) to evaluate the dependability of the applied models.

MATERIALS AND METHODS

Datasets

In this research, two hydraulic jump types were considered, including the free jump in trapezoidal, horizontal rectangular, expanding, rectangular with slope, and with negative step channels with rough and smooth beds and the submerged jump in horizontal and sloped rectangular channels with smooth beds. Bhutto (1987) performed hydraulic jump experiments on a rectangular smooth bed in the hydraulic laboratory of Mehran University of Engineering and Technology, Jamshoro. Accordingly, Bhutto's measurements yielded smooth-bed datasets. The experiments were conducted in a channel 16 m long. The cross-section of the channel was rectangular, 0.3 m wide and 0.4 m high. The channel included steel and glass-sided floors (70, 82, and 65 data for the submerged jump–horizontal, free jump–horizontal, and submerged jump–sloped smooth beds, respectively). The free-type hydraulic-jump datasets on the rectangular smooth bed channel with slope were achieved from the experiments conducted by Li (1995), including an adjustable channel (with length of 2.43 m, width of 0.15 m, and depth of 0.305 m) on different channel bed slopes in the range of 1°–5° (88 data). Hughes & Flack (1984) performed hydraulic jump experiments over a rectangular rough bed by artificially roughened test beds in a horizontal rectangular flume with a width of 0.305 m, and smooth Plexiglas side walls and length of 2.13 m. They utilized two kinds of roughness elements, including some parallel square bars perpendicular to the flow direction, along with strictly packed gravel particles cemented to the base (168 data). Additionally, free jump experimental data provided by Bremen (1990) and Evcimen (2012) were applied to predict the objectives. Bremen's experiments (1990) were anticipated for smooth (178 data) and with negative step (129 data) sudden diverging basins. Accordingly, 17 m^3 upstream basins were supplied by two 0.30-m-diameter conduits (with a maximum total discharge of 0.375 m^3/s) during the experiments. Further, a horizontal and prismatic rectangular channel (with a width of 0.5 m and length of 10.8 m) was linked to the basin. Similarly, Evcimen (2012) performed some experiments at the hydraulic laboratory of the Middle East Technical University. The trapezoidal shape channels were used during the experiments and the impacts of prismatic roughness on the characteristics of hydraulic jump were evaluated (107 data). In Figure 1 a hydraulic jump in various shape channels is illustrated.

Kernel-based techniques

Kernel-based methods such as Support Vector Machine (SVM) and Gaussian Process Regression (GPR) are considered as relatively innovative and significant techniques in terms of various kernel types and statistical learning theory. These models can adapt themselves for predicting any parameter of interest by adequate inputs. Furthermore, they can model non-linear decision boundaries, and numerous kernels exist in this regard. These methods are also objectively strong against overfitting, particularly in high-dimensional spaces. Nevertheless, proper selection of the kernel kind is the most essential step in GPR and SVM methods because of its direct effect on classification precision and training. In this study the SVM and GRP methods were implemented using the MATLAB computer programming language.

GPR

In GPR models, it is assumed that the nearby observations require the conveying of data regarding each other. Gaussian processes are a way of specifying a prior directly over function space. This is a natural generalization of the Gaussian distribution, whose mean and covariance are a vector and matrix, respectively. Although the Gaussian distribution is over vectors, its procedure is considered over functions. Accordingly, generalization requires no validation process considering the former knowledge on functional dependencies and data. Moreover, GP regression models can recognize the predictive distribution equivalent to test inputs (Rasmussen & Williams 2006). A GP is referred to as a combination of random variables for which any finite number possesses a mutual multivariate Gaussian distribution. For instance, $\chi \times \gamma$ denote input and output domains, respectively, and n pairs (x_i, y_i) are independently drawn and identically distributed from these domains. As regards regression, it is assumed that $y \subseteq \Re$. Then, a GP on χ is determined by the covariance function of $k{:}\chi \times \chi \rightarrow \Re$ and the mean function of $\mu{:}\chi \rightarrow \Re$.

In GP regression, it is primarily assumed that y is obtained by $y = f(x) + \xi$, in which $\xi \sim N(0,\ \sigma^2)$. The symbol $\sim$ in statistics indicates the sampling. It is noteworthy that a related random variable $f(x)$ exists as the value of the stochastic function f at that location for every input x in the GP regression. In the current study, it is presumed that the observational error ξ is typically independent and identically distributed with a mean value of zero ($\mu(x) = 0$), a variance of σ^2, and $f(x)$ resulting from the Gaussian process χ specified by k. In other words, $Y = (y_1, \ldots, y_n) \sim N(0,\ K + \sigma^2 I)$ where $K_{ij} = k(x_i, x_j)$ and I represent

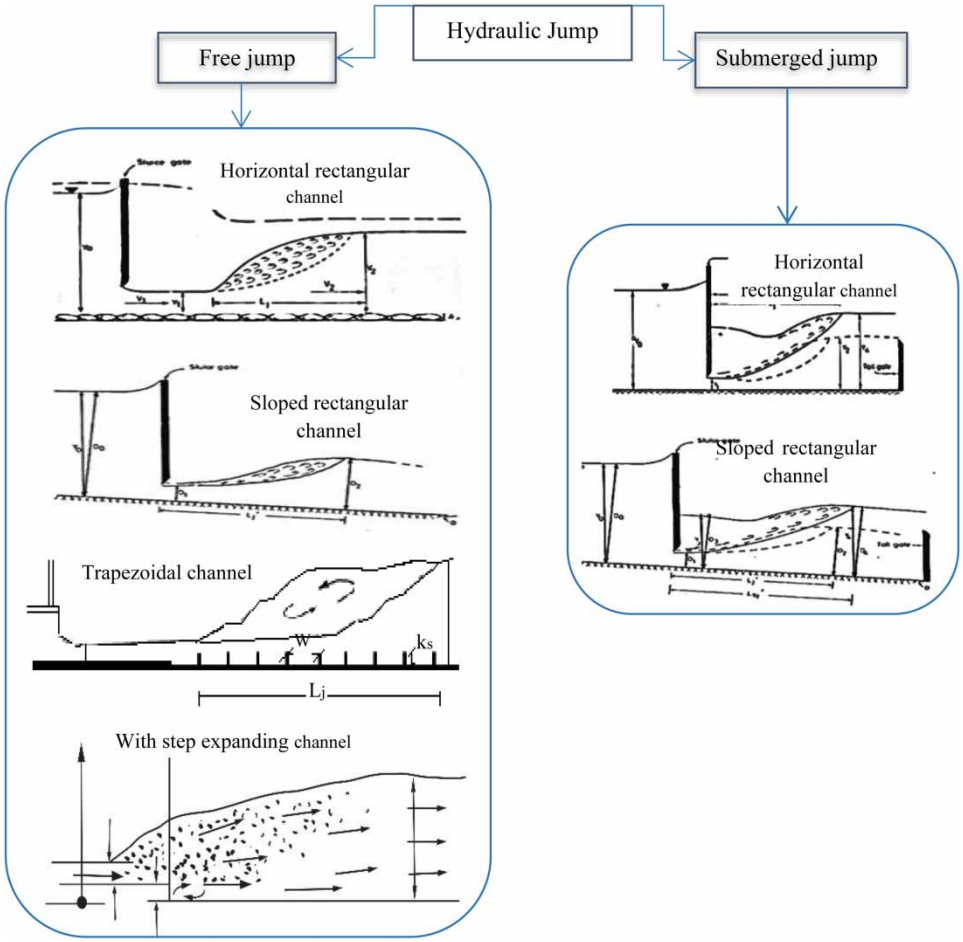

Figure 1 | Forms of hydraulic jump in different shape channels.

the identity matrix. Considering that $Y/X \sim N\,(0,\ K + \sigma^2 I)$ is normal, it is the conditional distribution of test labels based on the training and test data of $p(Y_*/Y, X, X_*)$. Thus, one possesses $Y_*/Y,\ X,\ X_* \sim N(\mu,\ \Sigma)$ where:

$$\mu = K(X_*, X)(K(X, X) + \sigma^2 I)^{-1} Y \tag{1}$$

$$\Sigma\ =\ K(X_*,\ X_*)\ -\ \sigma^2\, I\ -\ K(X_*, X)\,(K(X,\ X)\ +\ \sigma^2\, I))^{-1}\, K(X,\ X_*) \tag{2}$$

Given the existence of n_* test data and n training data, $K(X, X_*)$ denotes the $n \times n_*$ matrix of covariances assessed for all pairs of test data and training sets, which is similarly true for the other values of $K(X, X)$, $K(X_*, X)$, and $K(X_*, X_*)$, where X and Y are the training data vectors and also y_i are the training data labels while X_* demonstrates the testing data vector. A specified covariance function is needed for making a positive semi-definite covariance matrix K where $K_{ij} = k(x_i, x_j)$. The applied kernel function term in SVM equals the utilized covariance functions in GP regression. By the values of the kernel k and level of noise σ^2, Equations (1) and (2) would be sufficient for inference. An appropriate covariance function should be selected along with its parameters during the training procedure of GP regression models. Regarding GP regression with a fixed Gaussian noise value, a GP model can be trained by Bayesian inference, namely, the maximization of the marginal likelihood, leading to the maximization of the negative log-posterior:

$$p(\sigma^2,\ k)\ =\ \frac{1}{2}\, y^T\,(K\ +\ \sigma^2\, I)^{\,-1}\, y\ \ +\ \frac{1}{2}\log|K\ +\ \sigma^2\, I|\ -\ \log p(\sigma^2)\ -\ \log p(k) \tag{3}$$

To determine hyperparameters, the partial derivative of Equation (3) can be obtained based on σ^2 and k, and the gradient descent should be considered to get minimization. Due to higher effects on regression model accuracy, calculating the optimal value of the capacity constant (C), the size of the error-intensive zone (ε) in SVM and Gaussian noise in GPR are necessary. The optimal values for setting parameters of each kernel-based model should be obtained following the trial-and-error procedure.

SVM

SVM as an intelligence method is utilized in information and dataset categorization. Such a method, which was first introduced by Vapnik (1995), is recognized as structural risk minimization, minimizing an upper bound on the expected risk contrary to the traditional empirical risk that reduces the error on the training data. The SVM technique relies on the optimal hyperplane concept separating the samples of two classes while taking into account the widest gap between the two groups. Support Vector Regression (SVR) is an expanded version of SVM regression that aims to identify a function with the most deviation from the actual target vectors for all considered training data and the flatness as possible (Smola 1996). Vapnik (1995) represented the kernel function concept for non-linear SVR. It should be noted that proper selection of the kernel kind is the most significant step in the SVM owing to its direct effect on classification precision and training. Generally, various kinds of kernel functions exist, including linear, polynomial, and sigmoid types, along with radial basis function (RBF).

Model evaluation criteria

Equations (4)–(6) present the applied statistical measurements for assessing the behavior of various models, namely the Nash–Sutcliffe (DC), Correlation Coefficient (R), and Root Mean Square Error (RMSE). The linear dependence between observations and corresponding simulated values is provided by the R value (Legates & McCabe 1999), in other words, R is selected as the degree of collinearity criterion of level prediction. Its values range between -1 and $+1$. The optimal value of R is equal to 1. RMSE demonstrates the standard deviation of the differences between modeled and observed values (Chen & Chau 2016). The smaller the RMSE value, the better the model, viz., the more precise the predictions (Ratner 2009). DC shows the relative assessment of the model performance in dimensionless measures (Nash & Sutcliffe 1970), and it exhibits the relative magnitude of the residual variance compared with the observed data variance. The optimal value of DC is 1 (Tikhamarine *et al.* 2019).

$$DC = 1 - \frac{\sum_{i=1}^{N} (l_o - l_p)^2}{\sum_{i=1}^{N} (l_o - \overline{l_p})^2} \tag{4}$$

$$R = \frac{\sum_{i=1}^{N} (l_o - \overline{l_o}) \times (l_p - \overline{l_p})}{\sqrt{\sum_{i=1}^{N} (l_o - \overline{l_o})^2 \times (l_p - \overline{l_p})^2}} \tag{5}$$

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(l_o - l_p)^2}{N}} \tag{6}$$

where l_p, l_o, $\overline{l_p}$, $\overline{l_o}$, are the predicted, observed, mean predicted, and mean observed amounts, respectively. Also, N shows the data samples' number. It should be noted that in this study all used variables (input and output parameters) were scaled between 0 and 1 in order to eliminate the input and output variables' dimensions. Data normalizing eliminates the influence of variables with different absolute magnitudes and increases the training speed. The used normalization only rescales the data to another range which after modeling can be de-normalized to real values.

Simulation and model development

In this study, the GPR and SVM models' capabilities were examined for measuring hydraulic jump characteristics such as jump length and sequent depth ratio. Several datasets from previous studies were utilized to estimate hydraulic jump features (Figure 2). The intended datasets were classified into testing and training data types comprising 30% and 70% of samples. It

should be noted that the order of the datasets was selected in a way such that the training dataset contains a representative sample of all the behavior in the data in order to obtain a model with higher accuracy. One method for finding a good training set which can give good accuracy both in training and testing sets is an instance exchange (Bolat & Yildirim 2004). The process starts with a randomly selected training set. After the initial training process, the test data is applied to the network. A randomly selected true classified instance in the training set (I1) is thrown into the test set, a wrongly classified instance in the test set (I2) is put into the training set and the network re-trained. The process is repeated until reaching the maximum training and test accuracy. On the other hand, selecting appropriate variables as inputs for the models is essential through the modeling procedure. Several hydraulic jump parameters were considered in the present study, including flow features such as upstream flow depth, tail-water depth, immersion factor, and bed specifications (e.g., bed slope and roughness). Moreover, the Froude number is considered as a major parameter extensively utilized in a wide variety of hydraulic studies. Different arrangements of such variables were studied for the considered kinds of the jump, followed by evaluating the effects of these variables on features of hydraulic jumps. Tables 1–3 present the recommended input arrangements. The applied parameters for modeling include Y_1, Y_2, $\tan\theta$, Fr_1, Y_4, w, and k_s, indicating the initial depth of hydraulic jump, the second depth of hydraulic jump, the slope of the channel's bed, upstream Froude number, tail-water depth, rough element distance, and rough element height, along with S and S', which denote the immersion factor in free and submerged states, respectively.

In this research, in order to determine the best performance of the SVM and GPR models and select the best kernel functions, several developed models were tested via the SVM and GPR considering various kernels. Also, the optimal value of capacity constant (C) and the size of the error-intensive zone (ε) in these methods are required due to their high impact on the accuracy of the mentioned regression approaches. The coefficient C is a positive constant that influences a trade-off between the approximation error and the regression and must be selected by the user, and ε has an effect on the smoothness of the SVM and GPR responses, so both the complexity and the generaliz-ation capability of the network depend on its value. If epsilon is larger than the range of the target values, we cannot expect a good result. If epsilon is 0, we can expect overfitting. Also, the variable parameter used with a kernel function considerably affects the flexibility of the function. These parameters should be selected by the user. In the current study, optimization of these parameters has been done by a systematic grid search of the parameters using cross-validation on the training set of each considered state. For selecting the optimum parameters and assessing average developed model performance, the RMSE was used based on some well-known studies. Also, designing SVM- and GP-based regression methods involves using the kernel function concept. In this study, to determine the best structure of GPR and SVM models and choose the best kernel function, the depth ratio variable in the sloped smooth bed channel with the free jump was estimated by utilizing different kernels and using the model with the input parameters of $\tan\theta$ and Fr_1. Figure 3 displays the findings of the statistical parameters of various kernels for this model. Based on the findings, both the SVM and GPR models resulted in better prediction accuracy using the kernel function

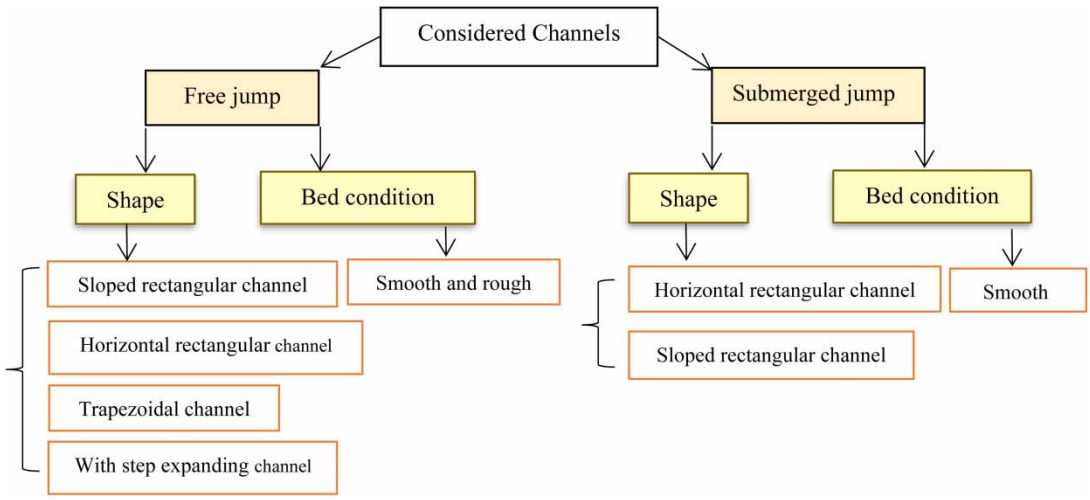

Figure 2 | Various forms of applied channels.

of RBF compared with other kernels. Thus, the RBF kernel was utilized as a core instrument of GPR and SVM applied for the remaining models.

RESULTS AND DISCUSSION

The results obtained for the sequent depth ratio modeling

Various models were established for evaluating hydraulic jump features in channels with various shapes based on the geometry and flow condition of rough elements and the applied channels. The expanded models were examined with the GPR and SVM models in order to perform depth ratio and jump length estimation in such channels. The best model was identified for each dataset. The model with the closest RMSE to 0 and the correlation coefficient (R) and determination correlation (DC) with closest value to 1 was selected as the best one. Table 1 and Figure 4 represent the findings of the SVM and GPR models. According to the results, the established models for the case of the sloped smooth bed channel performed more positively among all kinds of channels in the state of a free hydraulic jump compared with the other case. In most practical uses, it is desired and occasionally essential to create the hydraulic jump on sloped beds since this can present effective design information. According to Table 1, it can be seen that the combination of tanθ and Fr_1 parameters was the best model. This model predicted the depth ratio more accurately in the case of the channel with a sloped smooth bed (DC = 0.952, R = 0.974, and RMSE = 0.625). Nonetheless, analytical studies reveal that including the gravitational forces over the flow is essential for predicting the free jump characteristics over the sloped bed channels. As expected, adding the bed slope variable to input combinations enhanced the sequent depth ratio and the jump length prediction accuracy. Furthermore, according to the findings, utilizing Y_1/K_s as an input parameter increased the modeling performance for channels with the rough bed. Likewise, the expanding channel with no appurtenances resulted in better prediction compared with the expanding channel with steps.

Table 1 | Modeling results of the testing datasets for predicting the sequent depth ratio

Channel type	Model		R	DC	RMSE	Channel type	Model		R	DC	RMSE
Free hydraulic jump											
Horizontal smooth channel	**Fr_1**	**SVM**	**0.978**	**0.934**	**1.121**	Expanding channel with smooth bed	Fr_1	SVM	0.945	0.882	0.364
		GPR	**0.981**	**0.940**	**1.071**			GPR	0.953	0.887	0.314
	Fr_1, Y_1	SVM	0.955	0.900	1.171		**Fr_1, Y_1/B**	**SVM**	**0.955**	**0.915**	**0.299**
		GPR	0.958	0.905	1.121			**GPR**	**0.963**	**0.920**	**0.249**
Horizontal rough bed channel	Fr_1	SVM	0.954	0.741	0.681	Expanding channel with step	Fr_1	SVM	0.893	0.744	0.488
		GPR	0.964	0.745	0.631			GPR	0.900	0.755	0.438
	Y_2/k_s, Y_1/k_s	SVM	0.667	0.451	1.114		**Fr_1, Y_1/B**	**SVM**	**0.959**	**0.911**	**0.324**
		GPR	0.674	0.454	1.064			**GPR**	**0.967**	**0.920**	**0.305**
	Fr_1, Y_1/k_s	**SVM**	**0.987**	**0.942**	**0.324**		Fr_1, Y_1/k_s	SVM	0.931	0.864	0.345
		GPR	**0.997**	**0.948**	**0.274**			GPR	0.938	0.873	0.295
Sloped smooth bed channel	Fr_1	SVM	0.946	0.929	0.619	Trapezoidal channel with rough bed	Fr_1	SVM	0.958	0.913	0.306
		GPR	0.955	0.935	0.609			GPR	0.966	0.922	0.256
	Fr_1, Y_1	SVM	0.958	0.935	0.604		Fr_1, Y_1/k_s	SVM	0.964	0.915	0.298
		GPR	0.968	0.941	0.587			GPR	0.972	0.922	0.248
	Fr_1, tan θ	**SVM**	**0.974**	**0.952**	**0.498**		**Fr_1, w/k_s**	**SVM**	**0.972**	**0.949**	**0.286**
		GPR	**0.982**	**0.958**	**0.435**			**GPR**	**0.980**	**0.955**	**0.236**
Submerged hydraulic jump											
Horizontal smooth channel	Fr_1	SVM	0.969	0.956	0.319	Sloped smooth bed channel	Fr_1	SVM	0.740	0.475	1.049
		GPR	0.972	0.962	0.303			GPR	0.746	0.478	0.997
	Y_1, Fr_1	SVM	0.960	0.948	0.396		S', Fr_1	SVM	0.745	0.455	1.097
		GPR	0.968	0.954	0.376			GPR	0.745	0.459	1.050
	S, Fr_1	**SVM**	**0.997**	**0.993**	**0.012**		**Fr_1, tan θ**	**SVM**	**0.998**	**0.987**	**0.119**
		GPR	**0.997**	**0.996**	**0.010**			**GPR**	**0.998**	**0.990**	**0.113**
	S, Y_1, Fr_1	SVM	0.978	0.963	0.301		S', Fr_1, tan θ	SVM	0.977	0.956	0.167
		GPR	0.978	0.966	0.256			GPR	0.985	0.962	0.159

Table 2 | Comparing the results of the various models of the testing datasets for modeling jump length (free hydraulic jump)

Channel type	Model		*R*	DC	RMSE	Channel type	Model		*R*	DC	RMSE
Free hydraulic jump											
Horizontal smooth channel	$\boldsymbol{Fr_1}$	**SVM**	**0.992**	**0.979**	**5.050**	Expanding channel with smooth bed	Fr_1	SVM	0.855	0.855	0.420
		GPR	**0.996**	**0.985**	**4.899**			GPR	0.858	0.858	0.424
	Y_2/Y_1	SVM	0.965	0.933	6.180		$\boldsymbol{Fr_1, (Y_2-Y_1)/Y_1}$	**SVM**	**0.898**	**0.898**	**0.305**
		GPR	0.969	0.939	5.995			**GPR**	**0.901**	**0.901**	**0.301**
	Fr_1, Y_2/Y_1	SVM	0.989	0.971	5.320		Fr_1, Y_2/Y_1	SVM	0.865	0.874	0.373
		GPR	0.992	0.977	5.160			GPR	0.871	0.888	0.366
Horizontal rough bed channel	Fr_1	SVM	0.638	0.403	6.714	Expanding channel with step	Fr_1	SVM	0.884	0.884	0.369
		GPR	0.640	0.406	6.513			GPR	0.887	0.887	0.367
	Fr_1, Y_1/k_s	SVM	0.935	0.876	2.730		$\boldsymbol{Fr_1, (Y_2-Y_1)/Y_1}$	**SVM**	**0.912**	**0.909**	**0.327**
		GPR	0.938	0.881	2.648			**GPR**	**0.915**	**0.914**	**0.323**
	$Fr_1, \frac{Y_2}{k_s}-\frac{Y_1}{k_s}$	SVM	0.947	0.899	2.430		Fr_1, Y_2/Y_1	SVM	0.910	0.910	0.325
		GPR	0.951	0.904	2.357			GPR	0.913	0.913	0.330
	$\boldsymbol{Y_2/k_s, Y_1/k_s}$	**SVM**	**0.969**	**0.939**	**1.990**		Fr_1, Y_1/k_s	SVM	0.930	0.847	0.312
		GPR	**0.972**	**0.945**	**1.930**			GPR	0.933	0.852	0.303
Sloped smooth bed channel	Fr_1	SVM	0.967	0.935	1.130	Trapezoidal channel with rough bed	Fr_1	SVM	0.898	0.804	0.410
		GPR	0.970	0.940	1.096			GPR	0.901	0.809	0.406
	Fr_1, $\tan\theta$	SVM	0.978	0.957	0.913		Fr_1, $(Y_2-Y_1)/Y_1$	SVM	0.909	0.811	0.394
		GPR	0.981	0.962	0.886			GPR	0.912	0.816	0.382
	$\boldsymbol{Fr_1, Y_1, \tan\theta}$	**SVM**	**0.995**	**0.997**	**0.889**		Fr_1, Y_1/k_s	SVM	0.927	0.857	0.350
		GPR	**0.998**	**0.998**	**0.862**			GPR	0.930	0.862	0.340
	Fr_1, $\tan\theta$, Y_2/Y_1	SVM	0.972	0.938	1.012		$\boldsymbol{Fr_1, w/k_s}$	**SVM**	**0.935**	**0.858**	**0.346**
		GPR	0.975	0.944	0.982			**GPR**	**0.938**	**0.863**	**0.335**

Note: DC, determination coefficient; RMSE, root mean square error; SVM, support vector machine; GPR, Gaussian process regression.

Table 3 | Statistical parameters of the testing datasets for modeling the jump length (the submerged hydraulic jump)

Channel type	Model		*R*	DC	RMSE	Channel type	Model		*R*	DC	RMSE
Submerged hydraulic jump											
Horizontal smooth channel	Fr_1	SVM	0.891	0.845	7.438	Sloped smooth channel	S'	SVM	0.959	0.908	0.382
		GPR	0.894	0.850	7.066			GPR	0.963	0.913	0.371
	S	SVM	0.932	0.864	5.776		Fr_1	SVM	0.255	0.230	1.360
		GPR	0.935	0.869	5.603			GPR	0.256	0.232	1.319
	Y_4/Y_1	SVM	0.957	0.922	4.748		$\boldsymbol{S', \tan\theta}$	**SVM**	**0.983**	**0.958**	**0.223**
		GPR	0.960	0.928	4.606			**GPR**	**0.986**	**0.964**	**0.216**
	S, Fr_1	SVM	0.963	0.935	3.750		Fr_1, $\tan\theta$	SVM	0.394	0.311	1.277
		GPR	0.967	0.940	3.638			GPR	0.395	0.313	1.239
	S, Y_4/Y_1	SVM	0.973	0.955	3.528		S', Fr_1, & $\tan\theta$	SVM	0.949	0.893	0.412
		GPR	0.976	0.961	3.422			GPR	0.952	0.898	0.400
	$\boldsymbol{S, Y_4/Y_1, Fr_1}$	**SVM**	**0.994**	**0.985**	**3.159**		S', Fr_1, & $\tan\theta$	SVM	0.970	0.945	0.254
		GPR	**0.998**	**0.991**	**3.064**			GPR	0.973	0.951	0.246
							S', Fr_1, $\tan\theta$, & Y_4/Y_1	SVM	0.950	0.918	0.321
								GPR	0.953	0.924	0.311

Given the results of the established model regarding the trapezoidal channel with the rough bed, higher accuracy was presented by the model with parameters Fr_1 and w/k_s. As regards channels with a step or rough bed, it was found that Y_1/k_s and w/k_s increased the efficiency of the models and the impacts of w/k_s in further enhancing model accuracy compared with Y_1/k_s. Furthermore, favorable accuracy was found for the model, including only the input parameter of Fr_1. Therefore, the applied methods are capable of successfully predicting the sequent depth ratio by utilizing only upstream flow features as

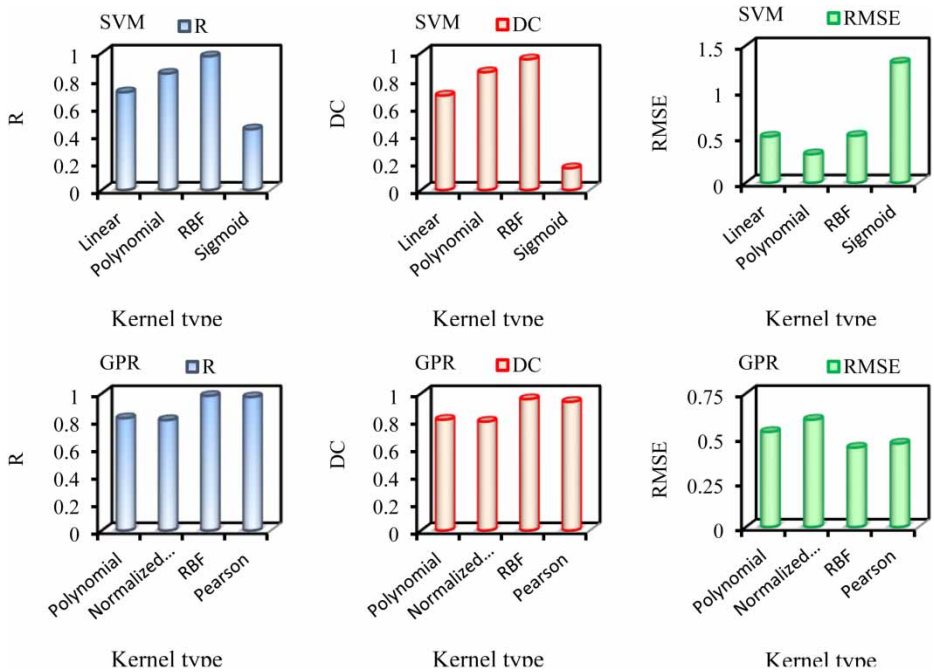

Figure 3 | Statistics parameters via SVM and GPR kernel functions for the testing set regarding the depth ratio of the sloped smooth bed channel. *Note*: DC, determination coefficient; RMSE, root mean square error; SVM, support vector machine; GPR, Gaussian process regression.

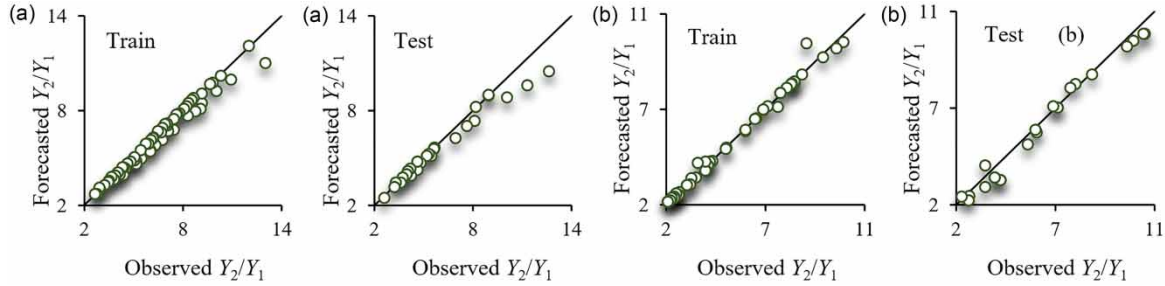

Figure 4 | Prediction of GPR against the perceived depth ratio of hydraulic jump: (a) free jump over the sloped smooth bed and (b) submerged jump over the horizontal smooth bed.

input data. The results of GPR models are negligibly accurate compared with those of the SVM model. Figure 4 depicts the verification between estimated and measured values of the test series for the best models of each state. In the form of a submerged hydraulic jump, the model with parameters Fr_1 and S in the case of the smooth horizontal channel functioned more successfully in comparison with the other models. Based on the comparison of the obtained results for free and submerged states, models for the submerged state led to better outcomes.

The results obtained for the length of hydraulic jump modeling

The statistical criteria obtained for the developed models for jump length in channels with a varied shape are listed in Table 2. Regarding channels with appurtenances, the horizontal rough bed channel resulted in better estimation accuracy compared with the expanding channel with step and trapezoidal channel and a rough bed, the maximum R and DC, and the lowest RMSE values. Based on the obtained data, the model with parameters Y_2/k_s and Y_1/k_s demonstrated a higher rate of accuracy. In addition, adding $(Y_2 - Y_1)/Y_1$, Y_1/k_s, w/k_s, and Y_2/Y_1 as input parameters increased the model's capability, confirming the impact of the geometry (i.e., roughness and step elements) of the applied appurtenances on the jump length in channels with various appurtenances. As regards the channels without appurtenances, the model including Fr_1, Y_1, and tan θ as input

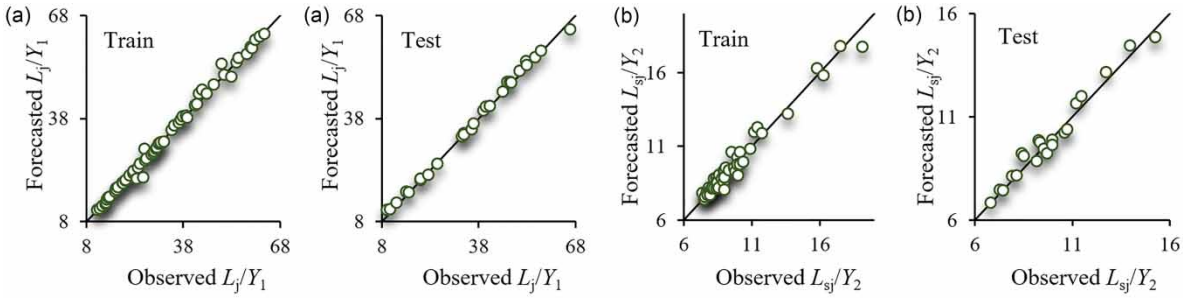

Figure 5 | Prediction of the best GPR model versus the obtained hydraulic jump length: (a) free jump on the sloped smooth bed and (b) submerged jump on the horizontal smooth bed.

parameters and using the sloped smooth bed channel gave optimal results. It was observed that in the jump length prediction in the expanding channel with a negative step the model with input parameters of Fr_1, S/Y_1 led to more accurate predictions, while for the case of the asymmetric channel the model L(II) with parameters of Fr_1, $(Y_2 - Y_1)/Y_1$ represented higher efficiency. Based on the results it could be inferred that for modeling the length of hydraulic jump in basins with a negative step, using parameters of Y_2/Y_1 and Y_1/k_s increased the accuracy of the models. Parameter Y_1/k_s confirms the importance of the relative height of the applied step in the hydraulic jump characteristic predicting process in channels with a negative step.

Table 3 provides validation results for predicting the submerged jump length over the horizontal channels with smooth beds. Investigation of the results confirms that the input combination with parameters S, Y_4/Y_1, and Fr_1 for length prediction led to superior performance regarding other combinations. Therefore, the immersion factor and the upstream Froude number were the most critical variables in modeling the submerged jump features over a horizontal channel with a smooth bed. Moreover, the addition of Y_2/Y_1 to the input parameters increased the model accuracy. Contrary to former cases, the objective function was L_{sj}/Y_2 regarding modeling the submerged jump length in the sloped channels with smooth beds. Generally, the literature review indicated that submerged jump on the sloped channel is nearly untouched, and restricted empirical equations exist for estimating the length of such jumps. Statistical parameters confirmed that the application of the upstream Froude number as an input parameter led to poor predictions in jump modeling. Accordingly, the slope of the bed and the immersion factor (S' and tan θ) played a crucial role in modeling the submerged jump length over the channels with sloped beds. Moreover, investigating input parameters to estimate the submerged jump sequent depth ratio in these channels confirmed that adding the slope of the bed (tan θ) to the preliminary Froude number enhances the findings (Table 3). Figure 5 shows the experimental values against the length values of the simulated jump with regard to the GPR best models in the channel with a varied shape for the test series.

Sensitivity analysis

The effects of various parameters on hydraulic jump features were assessed by a sensitivity analysis using the RMSE parameter. In this regard, the superior model was rerun via the GPR model, and the importance of each variable was evaluated by omitting the variable. Based on the RMSE values presented in Figure 6, it can be seen that the omitting of Fr_1 led to a significant reduction in modeling accuracy. Therefore, this parameter is the most effective variable in predicting the free hydraulic jump length over horizontal smooth bed channels, smooth and with negative step expanding channels, and trapezoidal channels with rough and sloped smooth beds. The secondary depth ratio normalized by the roughness diameter (Y_2/k_s) strongly affects the models' accuracy for predicting the length of a free hydraulic jump over horizontal rough beds. The immersion factor statistically is significant for predicting the length of the submerged hydraulic jumps over the sloped and horizontal smooth bed channels. As shown in Figure 6, Fr_1 was the essential factor for predicting the sequent depth ratio in all types of jumps, although it may not be the first leading parameter regarding modeling the hydraulic jump lengths.

Comparison of kernel-based approaches and classical equations

Various semi-empirical and analytical equations have been presented for forecasting hydraulic jump features having simple and complex structures. In this regard, empirical equations mentioning free hydraulic jumps are the most famous and valuable equations. The basis of the equations is obtained using the momentum principle regarding the sequent depth ratio. Nonetheless, using experimental equations, some references calculated several parts of the equation in some types of

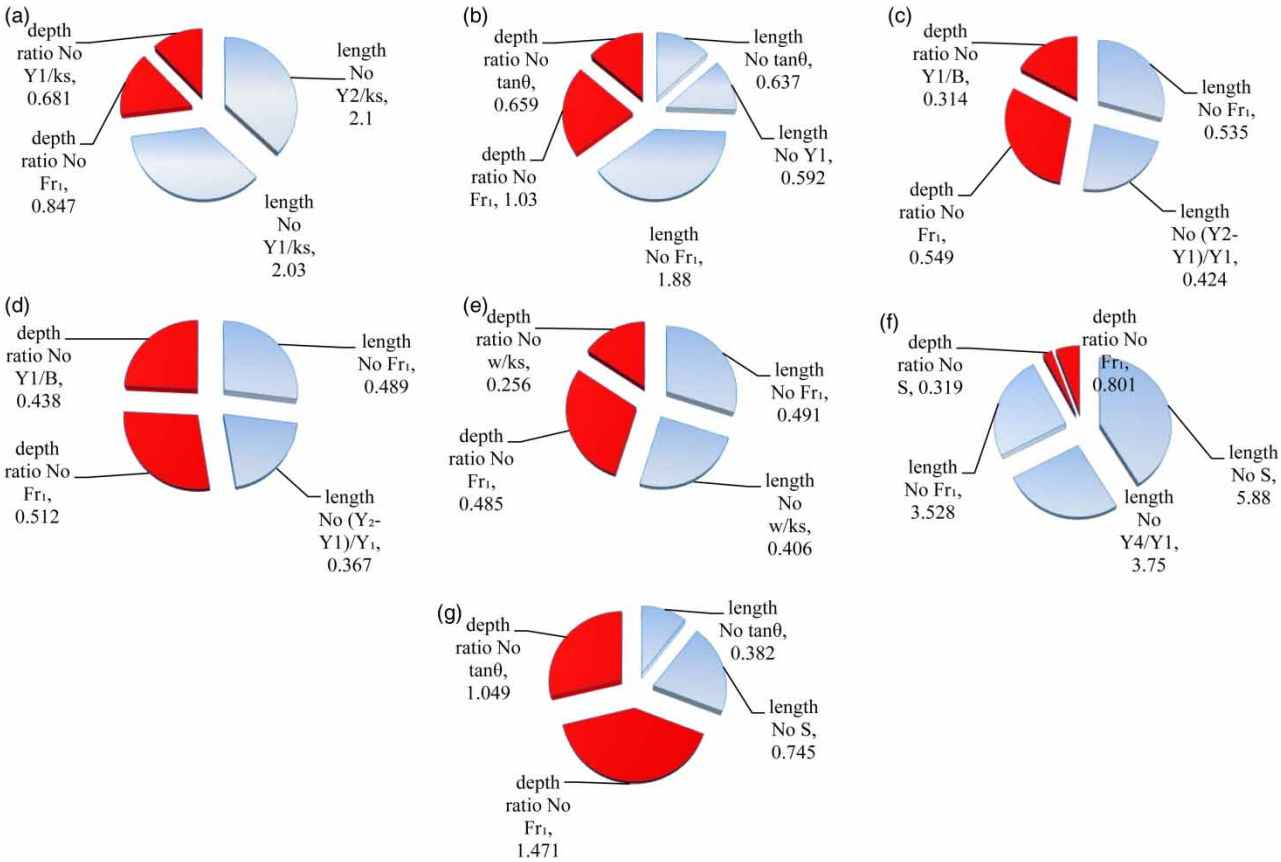

Figure 6 | The relative implication of every input parameter of the best GPR models: (a) free jump on the horizontal rough bed, (b) free jump on sloped smooth bed, (c) free jump on expanding smooth bed, (d) free jump on expanding channel with step, (e) free jump on the trapezoidal rough bed, (f) submerged jump on the horizontal smooth bed, and (g) submerged jump on sloped smooth bed. *Note*: GPR, Gaussian process regression.

hydraulic jumps. In addition, some of the studies focused on submerged hydraulic jumps. The stated equations for such a jump are valid for estimating jump features. In the present study, the accuracy of the best-introduced models and several available semi-theoretical formulae in the literature were compared to investigate the performance of the employed approaches. Table 4 summarizes the obtained data from the comparison. Based on three evaluated criteria (i.e., R, DC, & RMSE), the estimated values by the GPR and SVM models gave more precise results compared with equations. The results further revealed that the formulae used for computing the sequent depth ratio presented a rational fit to the experimental data. The length of the equations of the hydraulic jump represented no desired consistency between the approximated values and the observed sets in comparison with the sequent depth ratio formulae. It is worth noting that the present equations are based on untested model assumptions and a restricted database. Moreover, they lack field data and do not yield the same findings under variable flow conditions. Based on the results, it can be seen that the use of the Fr_1 parameter in the semi-empirical equations led to an increment in the equations' efficiency; however, the obtained results from the equations differ from each other and their application is limited to special cases of their development. Nonetheless, regardless of the uncertainty and complexity of the hydraulic jump phenomenon, the obtained data show the workability of GPR and SVM as kernel-based machine learning approaches for modeling the characteristics of free and submerged hydraulic jumps.

The usefulness of the applied methods

The most important characteristics of hydraulic jump, length, and sequent depth ratio play a key role in the design of hydraulic structures. However, the results of the semi-empirical equations which have been developed to estimate the length and sequent depth ratio of hydraulic jumps are not general and acceptable due to the uncertainty of the phenomenon. Therefore, artificial intelligence models, especially kernel-based approaches, can be used to improve the modeling efficiency. In the

Table 4 | Comparison of statistical parameters between the best GPR and SVM models and formulae

Channel type		Researcher	Equation	*R*	DC	RMSE
Length of jump	Free hydraulic jump on the horizontal smooth bed	Smetana (1934)	$L_j = 6 \times (Y_2 - Y_1)$	0.735	0.408	10.98
		Bakhmeteff & Matzke (1936)	$L_j = 5 \times (Y_2 - Y_1)$	0.733	0.411	10.08
			SVM	0.992	0.979	5.050
			GPR	0.996	0.985	4.899
	Free hydraulic jump on the horizontal rough bed	Ead & Rajaratnam (2002)	$\frac{L''_j}{Y_1} = 1.74Fr_1 + 3.62$	0.805	0.408	6.48
			SVM	0.969	0.939	1.990
			GPR	0.972	0.945	1.930
	Free hydraulic jump on the sloped smooth bed	Henderson (1966)	$\frac{L'_j}{Y_2} = 6.1 + 4\tan\theta$	0.962	0.499	4.851
			SVM	0.995	0.997	0.889
			GPR	0.998	0.998	0.862
	Free hydraulic jump on the expanding channel	Hager (1985)	$Lj = \left\{1 + \left(1 - \frac{1}{\sqrt{B}}\right) \times [1 - th(1.9X_1)]\right\} Lj^*$	0.712	0.311	0.733
			$Lj^* = Y_1 \times 220 \times th\left(\frac{Fr_1 - 1}{22}\right)$			
			SVM	0.898	0.898	0.305
			GPR	0.901	0.901	0.301
	Submerged hydraulic jump on the horizontal smooth bed	Wóycicki (1931)	$L_{sj} = H_{sj} \times \left[6.0 - 0.05 \times \frac{Y_4}{Y_1}\right]$	0.512	0.221	55.912
		Govinda Rao & Rajaratnam (1963)	$L_{sj} = Y_2 \times [\,4.9 \times S + 6.10\,]$	0.617	0.223	51.68
			SVM	0.994	0.985	3.159
			GPR	0.998	0.991	3.064
Sequent depth ratio	Free hydraulic jump on the horizontal smooth bed	Bélanger (1849)	$\frac{Y_2}{Y_1} = \frac{1}{2}\left(\sqrt{1 + 8Fr_1^2} - 1\right)$	0.828	0.618	6.12
			SVM	0.978	0.934	1.121
			GPR	0.981	0.940	1.071
	Free hydraulic jump on the horizontal rough bed	Carollo *et al.* (2007)	$\frac{Y_2}{Y_1} = \frac{1}{2}\left(\sqrt{1 + 8(1-\beta)Fr_1^2} - 1\right)$	0.801	0.699	0.619
			$\left[\beta = 0.42 \times \frac{K_s}{h_1}\right]$			
			SVM	0.987	0.942	0.324
			GPR	0.997	0.948	0.274
	Free hydraulic jump over the smooth expanding channel	Herbrand (1973)	$Y = Fr_1\sqrt{\frac{2}{B}} - \frac{1}{2B}$	0.905	0.654	0.597
			SVM	0.955	0.915	0.299
			GPR	0.963	0.920	0.249
	Free hydraulic jump on the sloped smooth bed	Rajaratnam (1966)	$\frac{Y_2}{Y_1} = \frac{1}{2}\left(\sqrt{1 + 8G^2} - 1\right)$	0.993	0.698	1.686
			$[G^2 = \Gamma^2 \times Fr_1^2], \quad \Gamma = 10^{0.027\theta}$			
			SVM	0.974	0.952	0.513
			GPR	0.982	0.958	0.498

Note: DC, determination coefficient; RMSE, root mean square error; SVM, support vector machine; GPR, Gaussian process regression; L'_j, length of the free jump on the sloped smooth bed channel; L_{sj}, length of the submerged jump over the horizontal smooth bed channel; Y_1, initial depth of the hydraulic jump; Y_2, the second depth of the hydraulic jump; Fr_1, upstream Froude number; $\tan\theta$, slope of the bed of the channel; Y_4, tail-water depth; H_{sj}, height of the jump; S, immersion factor; Γ, positive coefficient, which was obtained by Rajaratnam's proposed equation.

present investigation, the SVM and GPR modeling approaches were applied because of their high learning ability and information processing potentiality, which make them suitable for complex nonlinear modeling without prior knowledge about the input–output relationships, which is difficult to handle with the statistical approach. The kernel-based approaches have better generalization ability, less susceptibility to noise and outliers than the regression models, and can handle incomplete data. As can be seen from the obtained results, these methods led to more accurate outcomes compared with the semi-empirical equations.

The results of UA

The UA determined the uncertainty of the best GPR model. In the current study, the Monte Carlo UA method was utilized as well. In the UA method, two elements are applied to measure robustness and evaluate model uncertainty. The first one is the percentage of the investigated outputs within the range of 95PPU, and the next one represents the average distance between the lower (X_L) and upper (X_U) uncertainty bands (Noori *et al.* 2015). Accordingly, the intended model needs to be run several times (1,000 times in this study), and the experimental cumulative distribution probability of the models should be determined as well. The lower and upper bands are regarded as the probabilities of 97.5% and 2.5% for the cumulative distribution, respectively. Two important indices should be taken into account at the appropriate confidence level. First, the 95PPU band brackets most observations. In addition, the average distance between the upper and lower parts of the 95PPU (*d*-factor) should be smaller than the standard deviation of the observed data (Abbaspour *et al.* 2007). The indicated indices were used for estimating input uncertainties. According to Abbaspour *et al.* (2007), the average width of the confidence interval band can be computed by Equation (7):

$$d\text{-factor} = \frac{\overline{\mathrm{dx}}}{\sigma \mathrm{x}} \tag{7}$$

where $\overline{dx}$ and σx denote the observed average width of the confidence band and data standard deviation, respectively. The 95PPU can be determined as follows:

$$\text{Bracketed by 95PPU} = \frac{1}{k}\mathrm{Cont}(j|X_L < X_{reg} < X_U \tag{8}$$

where 95PPU, k, and X_{reg} represent the predicted uncertainty of 95%, the observed data number, and the currently registered data, respectively. Table 5 presents the obtained results for the UA. According to the *d*-factor values and 95% PPU, the predicted and observed values were within the 95% PPU band in most cases. Additionally, the findings revealed that the rate of *d*-factors for training and testing datasets was lower compared with the standard deviation of the observed data. Therefore, modeling the hydraulic jump characteristics via the GPR model resulted in an allowable uncertainty level.

Table 5 | Uncertainty indices for the GRP model

			Performance criteria				Performance criteria	
Channel type		**Model**	**95%PPU**	***d*-factor**	**Channel type**	**Model**	**95%PPU**	***d*-factor**
	Free hydraulic jump				Submerged hydraulic jump			
Sequent depth ratio	Horizontal smooth bed	Fr_1;	80.05%	1.321	Horizontal smooth bed	Fr_1, Y_1/B	78.12%	1.373
	Horizontal rough bed	Fr_1, Y_1/k_s	81.12%	1.160	Sloped smooth bed	Fr_1, tan θ	75.24%	1.424
	Sloped smooth bed	Fr_1, tan θ	88.15%	1.082				
	Expanding smooth bed	Fr_1, Y_1/B	76.23%	1.113				
	Expanding channel with step	Fr_1, Y_1/B	79.18%	1.055				
	Trapezoidal channel with rough bed	Fr_1, w/k_s	84.10%	1.081				
Length of jump	Horizontal smooth bed	Fr_1	73.18%	2.880	Horizontal smooth bed	S, Y_4/Y_1, Fr_1	64.81%	4.310
	Horizontal rough bed	Y_2/k_s, Y_1/k_s	71.14%	3.923	Sloped smooth bed	S', tan θ	60.15%	4.651
	Sloped smooth bed	Fr_1,Y_1, tan θ	79.55%	3.121				
	Expanding smooth bed	Fr_1, $(Y_2-Y_1)/Y_1$	68.17%	2.155				
	Expanding channel with step	Fr_1, $(Y_2-Y_1)/Y_1$	70.05%	2.283				
	Trapezoidal channel with rough bed	Fr_1, w/k_s	66.28%	4.050				

Note: GRP, Gaussian process regression; PPU, prediction uncertainty.

CONCLUSIONS

In the present study, novel data-mining approaches (i.e. GPR and SVM) were used in order to predict the sequent depth ratio and the length in various kinds of channels and hydraulic jumps in trapezoidal, rectangular, expanding, sloped, and with negative step channels with smooth and rough beds. The applied meteorological data-sets for establishing the models were found from former valid studies. The results from the presented GPR and SVM methods and the obtained values from other semi-empirical equations were compared, and it was demonstrated that the SVM- and GPR-based models overwhelm the employed semi-empirical equations regarding predicting the sequent depth ratio and the length of the jump in submerged and free jumps. Thus, some notable conclusions were drawn as follows.

In the channels with smooth horizontal beds, it was observed that the upstream Froude number increased the modeling accuracy and led to better prediction of free hydraulic jump length. Regarding free hydraulic jumps over horizontal rough bed channels, using both the ratio of primary and secondary depths to roughness diameter (Y_1/k_s and Y_2/k_s) plays a key role in predicting the length of this jump. Regarding the free hydraulic jump over sloped smooth bed channels, the addition of the bed slope to the upstream Froude number and the initial depth results in an enhancement in the findings. Moreover, for channels with a step or rough bed, parameters w/k_s and Y_1/k_s incremented model efficiency, revealing the effect of the geometry of the applied appurtenances (i.e., roughness elements and step) over hydraulic jump features in channels with numerous appurtenances. Based on the obtained results, the developed models for the sloped smooth bed channel were more positive between all types of channels in the state of a free hydraulic jump than other cases.

Concerning the submerged jumps on channels with sloped and horizontal smooth beds, the immersion factor was considered as the most influential variable. It had a bold contribution to the submerged jump length predictions.

In general, it was confirmed that the upstream Froude number is the most contributing factor regarding predicting the sequent depth ratio in all kinds of hydraulic jumps. It was further found that the results of the GPR models are slightly more precise compared with the SVM models. On the other hand, it should be noted that although GPR and SVM can develop an extremely precise model for predicting the objective function, they cannot establish explicit rules. Moreover, the dependability of the superior applied models was assessed through UA, and it was revealed that the GPR model possessed an acceptable uncertainty degree in the modeling of hydraulic jump characteristics in channels with shapes. It should, however, be noted that the AI models are data-driven models, and thus, they are data sensitive. The prediction of data using SVM and GPR depends on the availability of laboratory data. Therefore, different types of channels to be used (as the ones in the present study) may affect the fitting and prediction accuracy of applied models, which should be further discussed in the future. Also, it is suggested to use other methods of artificial intelligence and compare the obtained results with the present research.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

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. *Journal of Hydrology* **333** (2–4), 413–430.

Abbaspour, A., Hosseinzadeh Dalir, A., Farsadizadeh, D. & Sadraddini, A. A. 2009 Effect of sinusoidal corrugated bed on hydraulic jump characteristics. *Journal of Hydro-Environment Research* **3**, 109–117.

Ahmed, H. M. A., El Gendy, M., Mirdan, A. M. H., Mohamed Ali, A. A. & Abdel Haleem, F. S. F. 2014 Effect of corrugated beds on characteristics of submerged hydraulic jump. *Ain Shams Engineering Journal* **5** (4), 1033–1042.

Azamathulla, H. M., Ghani, A. A., Chang, C. K., Hasan, Z. A. & Zakaria, N. A. 2010 Machine learning approach to predict sediment load – a case study. *Clean Soil, Air, Water* **38** (10), 969–976.

Azamathulla, H. M., Haghiabi, A. H. & Parsaie, A. 2016 Prediction of side weir discharge coefficient by support vector machine technique. *Water Supply* **16** (4), 1002–1016.

Azimi, H., Bonakdari, H. & Ebtehaj, I. 2017 Sensitivity analysis of the factors affecting the discharge capacity of side weirs in trapezoidal channels using extreme learning machines. *Flow Measurement and Instrumentation* **54**, 216–223.

Bakhmeteff, B. A. & Matzke, A. E. 1936 The hydraulic jump in terms of dynamic similarity. *Transactions of the American Society of Civil Engineers* **101**, 630–647.

Bélanger, J. B. 1849 *Notes sur le Cours d'Hydraulique (Notes on Hydraulic Engineering)*, Session 1849–1850. École nationale des ponts et chaussées, Paris, France (in French).

Bhutto, H. B. G. 1987 *Hydraulic Jump Control and Energy Dissipation*. PhD thesis, Mehran University of Engineering and Technology, Jamshoro, Pakistan.

Bolat, B. & Yildirim, T. 2004 A data selection method for probabilistic neural networks. *Journal of Electrical and Electronics Engineering* **4** (2), 1137–1140.

Bremen, R. 1990 *Expanding Stilling Basin*. Laboratoire de Constructions Hydrauliques, EPFL, Lausanne, Switzerland.

Carollo, F. G., Ferro, V. & Pampalone, V. 2007 Hydraulic jumps on rough beds. *Journal of Hydraulic Engineering* **133** (9), 989–999.

Chen, X. Y. & Chau, K. W. 2016 A hybrid double feedforward neural network for suspended sediment load estimation. *Water Resources Management* **30** (7), 2179–2194.

Chen, X. Y., Chau, K. W. & Busari, A. O. 2015 A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. *Engineering Applications of Artificial Intelligence* **46**, 258–268.

Chiang, J. & Tsai, Y. 2011 Suspended sediment load estimate using support vector machines in Kaoping river basin. In: *2011 International Conference on Consumer Electronics, Communications and Networks (CECNet),* IEEE, Piscataway, NJ, USA, *pp. 1750–1753.*

Corry, M. L., Thompson, P. L., Watts, F. J., Richards, D. L., Jones, J. S. & Bradley, J. N. 1975 *Hydraulic Design of Energy Dissipators for Culverts and Channels.* Hydraulic Engineering Circular No. 14, US Department of Transportation, Federal Highway Administration, Washington, DC, USA.

Di Nunno, F. & Granata, F. 2020 Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network. *Environmental Research* **190**, 110062.

Ead, S. A. & Rajaratnam, N. 2002 Hydraulic jumps on corrugated beds. *Journal of Hydraulic Engineering* **128** (7), 656–663.

Evcimen, T. U. 2012 *Effect of Prismatic Roughness on Hydraulic Jump in Trapezoidal Channels*. PhD thesis, Middle East Technical University, Ankara, Turkey.

Foddis, M. L., Ackerer, P., Montisci, A. & Uras, G. 2015 ANN-based approach for the estimation of aquifer pollutant source behaviour. *Water Supply* **15** (6), 1285–1294.

Gholami, V., Chau, K. W., Fadaee, F., Torkaman, J. & Gaffari, A. 2015 Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers. *Journal of Hydrology* **529** (3), 1060–1069.

Govinda Rao, N. S. & Rajaratnam, N. 1963 The submerged hydraulic jump. *Journal of the Hydraulics Division* **89** (1), 139–162.

Granata, F. & Di Nunno, F. 2021 Artificial Intelligence models for prediction of the tide level in Venice. *Stochastic Environmental Research and Risk Assessment*. doi:10.1007/s00477-021-02018-9.

Granata, F., Di Nunno, F., Gargano, R. & de Marinis, G. 2019 Equivalent discharge coefficient of side weirs in circular channel – a lazy machine learning approach. *Water* **11** (11), 2406.

Gupta, S. K., Mehta, R. C. & Dwivedi, V. K. 2013 Modeling of relative length and relative energy loss of free hydraulic jump in horizontal prismatic channel. *Procedia Engineering* **51**, 529–537.

Hager, W. H. 1985 Hydraulic jump in non-prismatic rectangular channels. *Journal of Hydraulic Research* **23** (1), 21–35.

Henderson, F. M. 1966 *Open Channel Flow*. MacMillan, New York, USA.

Herbrand, K. 1973 The spatial hydraulic jump. *Journal of Hydraulic Research* **11** (3), 205–218.

Hughes, W. C. & Flack, J. E. 1984 Hydraulic jump properties over a rough bed. *Journal of Hydraulic Engineering* **110** (12), 1755–1771.

Kumar, M., Kumari, A., Kushwaha, D. P., Kumar, P., Malik, A., Ali, R. & Kuriqi, A. 2020 Estimation of daily stage–discharge relationship by using data-driven techniques of a perennial river, India. *Sustainability* **12** (19), 7877.

Legates, D. R. & McCabe Jr., G. J. 1999 Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research* **35** (1), 233–241.

Li, C. F. 1995 *Determining the Location of Hydraulic Jump by Model Test and HEC-2 Flow Routing*. MSc thesis, Ohio University, Athens, OH, USA.

Malik, A., Tikhamarine, Y., Souag-Gamane, D., Kisi, O. & Pham, Q. B. 2020 Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction. *Stochastic Environmental Research and Risk Assessment* **34** (11), 1755–1773.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10** (3), 282–290.

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? *Journal of Hydraulic Engineering* **142** (1), 04015039.

Pandey, K., Kumar, S., Malik, A. & Kuriqi, A. 2020 Artificial neural network optimized with a genetic algorithm for seasonal groundwater table depth prediction in Uttar Pradesh, India. *Sustainability* **12** (21), 8932.

Rajaee, T., Ebrahimi, H. & Nourani, V. 2019 A review of the artificial intelligence methods in groundwater level modeling. *Journal of Hydrology* **572**, 336–351.

Rajaratnam, N. 1966 The hydraulic jump in sloping channels. *Water and Energy International* **23** (2), 137–149.

Rasmussen, C. E. & Williams, C. K. I. 2006 *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA, USA.

Ratner, B. 2009 The correlation coefficient: its values range between +1/−1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing* **17** (2), 139–142.

Roushangar, K. & Ghasempour, R. 2017 Prediction of non-cohesive sediment transport in circular channels in deposition and limit of deposition states using SVM. *Water Supply* **17** (2), 537–551.

Roushangar, K. & Ghasempour, R. 2018 Evaluation of the impact of channel geometry and rough elements arrangement in hydraulic jump energy dissipation via SVM. *Journal of Hydroinformatics* **21** (1), 92–103.

Roushangar, K., Ghasempour, R. & Biukaghazadeh, S. 2019 Evaluation of the parameters affecting the roughness coefficient of sewer pipes with rigid and loose boundary conditions via kernel based approaches. *International Journal of Sediment Research* **35** (2), 171–179.

Smetana, J. 1934 *Experimental Study of the Submerged or Expanded Hydraulic Jump. Report Translated from the Russian by A.D. Kalal*. Bureau of Reclamation, Denver, CO, USA.

Smola, A. J. 1996 *Regression Estimation with Support Vector Learning Machines*. Master's thesis, Technische Universität München, Munich, Germany.

Tikhamarine, Y., Malik, A., Kumar, A., Souag-Gamane, D. & Kisi, O. 2019 Estimation of monthly reference evapotranspiration using novel hybrid machine learning approaches. *Hydrological Sciences Journal* **64** (15), 1824–1842.

Vapnik, V. N. 1995 *The Nature of Statistical Learning Theory*, 2nd edn. Springer-Verlag, New York, USA, pp. 1–47.

Wang, D. 2016 Research on raw water quality assessment oriented to drinking water treatment based on the SVM model. *Water Supply* **16** (3), 746–755.

Wóycicki, K. 1931 *Wassersprung, Deckwalze und Ausfluss unter einer Schütze (The Hydraulic Jump, Its Top Roll and Discharge through a Sluice Gate*). PhD thesis, Polnische Akademie der Wissenschaften, Warsaw, Poland.

Wu, C. L., Chau, K. W. & Li, Y. S. 2009 Methods to improve neural network performance in daily flows prediction. *Journal of Hydrology* **372** (1–4), 80–93.

Yaseen, Z. M., Naghshara, S., Salih, S. Q., Kim, S., Malik, A. & Ghorbani, M. A. 2020 Lake water level modeling using newly developed hybrid data intelligence model. *Theoretical and Applied Climatology* **141**, 1285–1300.

First received 7 April 2021; accepted in revised form 17 August 2021. Available online 31 August 2021