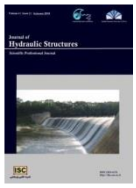
Journal of Hydraulic Structures
J. Hydraul. Struct., 2024; 10(1): 46-65
DOI: 10.22055/jhs.2024.45673.1278



OPEN ACCESS

Investigation of the Performance of Soft Computing Methods in the Hydraulic Evaluation of the Slot Fishway on the Inclined Drop

Ehsan Aminvash [1]
Farhoud Kalateh [2]
Rasoul Daneshfaraz[3]
John Abraham[4]

Abstract

The present research investigates K-Nearest Neighbor (K-NN) and Support Vector Machine (SVM) algorithms for evaluating hydraulic parameters of inclined drops equipped with fishway elements. 395 measurements were analyzed using these two algorithms with a focus on $\Delta E/E$ and y_d/H. To evaluate algorithm effectiveness, three accuracy measures: R^2, RMSE, and the KGE are used for a number of different scenarios. The SVM results show that in the first and second scenarios, respectively, model number 13 yielded RMSE=0.0156, R^2=0.973, and KGE=0.961 for the training mode and RMSE=0.0241, R^2=0.962, and KGE=0.952 for the testing mode. The K-NN algorithm with 75% of the total data employed for the initial sample size gave the best results with RMSE=0.0121, R^2=0.986, and KGE=0.975 for the first scenario and RMSE= 0.0123, R^2= 0.986, and KGE= 0.975 for the second scenario. Also, the results stated that the RBF kernel function is superior among the kernel functions of the SVM algorithm and the Cityblock distance measurement method among other methods of the K-NN algorithm. On the other hand, the K-NN method was superior to the SVM method.

Keywords: Fishway elements, Inclined drop, K-NN algorithm, SVM algorithm.

Received: 29 December 2023; Accepted: 18 March 2024

1. Introduction

One of the most essential things that must be paid attention to in the design of hydraulic structures is in terms of stability management. It is very important to keep the fish healthy and not lose them in the fishway structures. In the present research, we decided to investigate the flow on inclined drops by considering environmental considerations. The goals of the design are to

[1]Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran, Email: Ehsan.aminvash1994@gmail.com **(Corresponding author)**
[2] Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran.
[3] Faculty of Engineering, University of Maragheh, Maragheh, Iran.
[4] University of St. Thomas, School of Engineering, St Paul, USA.

simultaneously provide energy dissipation, protect downstream structures, increase aeration of the flow, improve water quality, and allow the migration of fish. Creating turbulence and mixing water and air is one of the most effective methods to increase energy losses downstream of an inclined drop. These structures are built for environmental and economic considerations. The design of these structures is such that it allows the fish to enter and exit easily without incurring injury and without unnecessary delay for spawning. The aim of the present research is to investigate the performance of neural network methods for evaluating the hydraulic parameters of inclined drops equipped with fishway elements.

Zhang et al. [1] designed a new multi-slot and pool-overflow combined fishway channel based on hydraulic property analysis and fish-passage test. Zhong et al. [2] numerically and experimentally evaluated the hydraulic parameters of the new semi-frustum in the pool-overflow fishway path and showed that the fish path with bed slope and inverse radius ratio, but with a longer pool distance, performs better. Zheng et al. [3] conducted a comparative study of the hydraulic characteristics of natural fish tracks and showed that a weak circulation zone is formed in the low-velocity region, which makes swimming easier for migratory fish. Baki and Azimi [4] studied the hydraulic design of slot and rock fishways in a review study. The results of their review and research showed that the long-term progress in the fishway lanes combines with vertical cracks and stone overflow, which finally helps the engineers in designing these instruments. Harris et al. [5] experimentally presented an innovation for tubular fishways and stated that the tubular fishways route has the potential to effectively pass fish upstream in new barriers of more than 5 inches with low construction and operating costs in the regime. Cea et al. [6] numerically investigated the average depth of turbulence for flow in vertical fishways. The results showed the importance of the turbulence model in numerical simulations of this nature.

Rodriguez et al. [7] evaluated the designs of fish passages with vertical slots and their effect on fish mobility. They showed that the amount of energy consumption per unit volume of the pool should not be more than 200 W/m^3 for large anadromous species and 150 W/m^3 for smaller species such as trout. Yagci [8] discussed fishways as eco-friendly structures. In addition to the speed and depth of the flow, which affects the ability of fish to swim, the intensity of turbulence is also important. The patterns of flow and turbulence can be controlled to provide suitable conditions for swimming.

Due to the interaction of the flow structure and the fishway elements, studies on specific laboratory conditions are generally limited in their applicability. Currently, artificial intelligence methods are taking on a larger role as investigation tools for these problems. Among the methods are: data mining methods that are used to model non-linear systems, Artificial Neural Networks (ANN), Gene Expression Programming (GEP), SVM, and K-NN as examples.

Arffin [9] used an ANN model and a linear regression model to predict sedimentation. Alp and Cigizoglu [10] compared the results of RBF and FFBP methods with linear regression and stated that the mentioned methods have provided better results. Using field and laboratory data, Goel and Pal [11] investigated the support vector machine for predicting scour depth and showed that changes in flow conditions, geometry, and bed materials have an effect on scour depth. Roshangar et al. [12] evaluated the efficiency of support vector machine in predicting hydraulic jump in contractions. The results showed that the relative energy consumption and downstream water depth varied with Froude number.

Sadeghfam et al. [13] studied the scrubbing of supercritical upstream of screens in a laboratory and used artificial intelligence methods (SFL, NF, and SVM methods). Daneshfaraz et al. [14] predicted hydraulic parameters of a drop by double screens using support vector machine. The results showed that the experimental hydraulic parameters are in good agreement with the output results from the support vector machine. Francisco Fuentes-Perez et al. [15] conducted to an autonomous obstruction detection system using hydraulic modeling and sensor networks in a step toward smart fishways with Paython. Results showed that the proposed system allows monitoring of the hydraulic performance of the fishway. Seifollahi et al. [16] estimated local scouring of cylindrical bridge piers using hybrid neural wavelets and artificial neural network methods. The results of their research showed that the neural wavelet method provided superior results than the ANN method. Daneshfaraz et al. [17, 18] have also conducted research in the field of vertical drops with screens and a blocky bed using the artificial network method.

Kalathil and Chandra [19] experimentally and numerically examined the hydraulic design of a step-pool fishway. The results showed that the velocity and the averaged velocity over the crest at 0.10 and 0.25 are considerably lower than at 0.25. Also, the maximum TKE and energy dissipation factors were within recommended limits for 0.10 and 0.25. Baharvand and Lashkar-Ara [20] investigated the hydrodynamic and biological characteristics of a modified meander C-type fishway to pass rainbow trout. The results showed that the TKE distribution graphs for different discharge and geometry scenarios through the longest path in the pools of the MMCF were investigated using the line probe technique for an average depth of 0.5d.

Mirkhorli et al. [21] examined hydraulic aspects of a rectangular labyrinth pool and weir fishway using FLOW-3D; they showed that a conventional weir has less energy dissipation than a rectangular labyrinth weir. Lower TKE and TI values were observed at the top of the labyrinth weir, at the corner of the wall downstream of the weir, and between the side walls of the weir and the channel wall. Baharvand et al. [22] conducted a comparative study using machine learning and regression-based approaches for predicting the hydraulic jump depth ratio. The results of their study show that the proposed model increases the correlation between observed and predicted values up to 0.987 and decreases the RMSE to 0.324. Cassan et al. [23] investigated the hydraulics of rock-ramp fishways with lateral slopes. The results of their research revealed that the method for assessing the stage-discharge relationship and maximum velocities is related to the lateral slope.

According to the research related to the prediction of hydraulic parameters in hydraulic structures, it was observed that the input parameters are very important for the evaluation of independent parameters. Therefore, in the present research, the prediction and evaluation of energy dissipation parameters and downstream depth of inclined structures equipped with fishway elements was performed using K-Nearest Neighbor (K-NN) and Support Vector Machine (SVM) methods. To the best knowledge of the authors, the nearest neighbor method (K-NN) has not been used in this type of study. Consequently, here, for the first time, the evaluation of the hydraulic parameters with K-NN and SVM methods have been performed.

2. Methodology

2.1. Laboratory equipment and fishway elements

In order to check the performance of data mining methods for predicting relative energy losses and downstream depth parameters of the inclined drop; a total of 395 measurements are used. Experiments have been conducted in a hydraulic laboratory in a flume with length, width, and height of 5 meters, 30 cm, and 50 cm, respectively. The ranges of measured parameters are presented in Table 1. For fishway elements, glass plates with length, height, and thickness of 15, 10, and 6 cm respectively; with and without holes; and with angles of 45, 60, and 90 degrees to the horizon are used. The angle of the breakwater structure is 26.56 degrees with respect to the channel floor. The dimensions of the holes in the hollow elements are 5×5 cm. Figure 1 shows a model of the experimental channel and theequipment [24].

Table 1. Range of measured parameters and characteristics of fishway elements

Model Name	α (Degree of fishway elements)	Q (Lit/sec)	y_d (cm)	Fr_d (-)	Reynolds No.	Type of fishway element	
						with hole	without hole
Simple drop	-----		0.6~1.67	4.94~6.93		-----	-----
FW1	90		2.57~14.2	0.19~0.86			*
FW2	60	3.33~10	2.21~10	0.3~0.98	38986~582857		*
FW3	45		2.03~11.25	0.28~1.22			*
FW4	90		4.9~8.21	0.33~0.45		*	
FW5	60		4.46~8.73	0.37~0.42		*	
FW6	45		4.37~8.85	0.38~0.44		*	

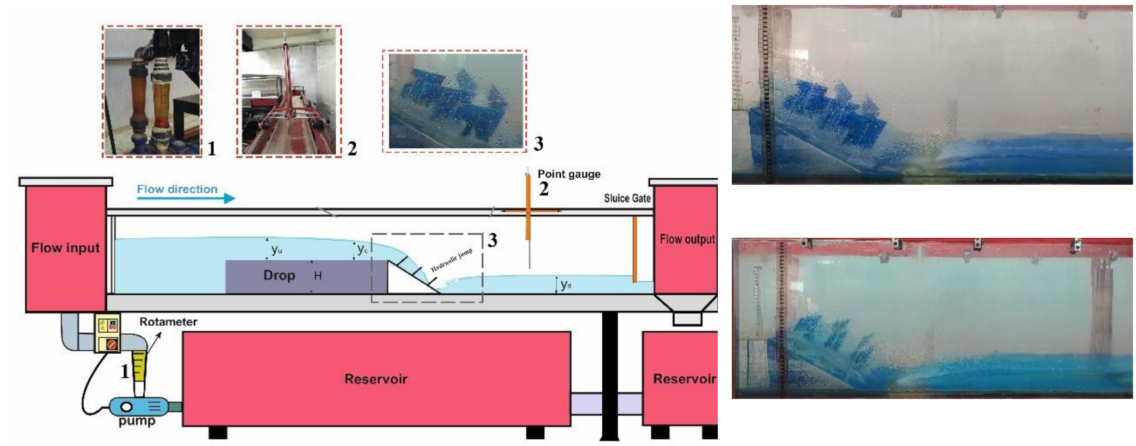


Figure 1. Experimental flume and equipment

2.2. Algorithm of SVM

SVM is a data-mining and machine-learning algorithm. It is divided into two modes (testing and training) and is applied for prediction and data classification. Machine learning requires input data; a portion of the data is used as training data and the remaining data is reserved for testing. The software calculates the dependent parameter based on independent parameter values. The SVM algorithm has various parameters, the most important of which is the kernel parameter or γ, the setting of which is important for the solution of the problem.

2.3. SVM theory

The SVM method entails classification and linear separation of data. If the data is linear, an attempt is made to separate them and maximum margins are calculated (shown graphically in Fig. 2); the size of the separating plate can be calculated from Eq. (1).

$$\text{Margin}=\frac{2}{\|w\|}=\frac{2}{w^{T}w} \tag{1}$$

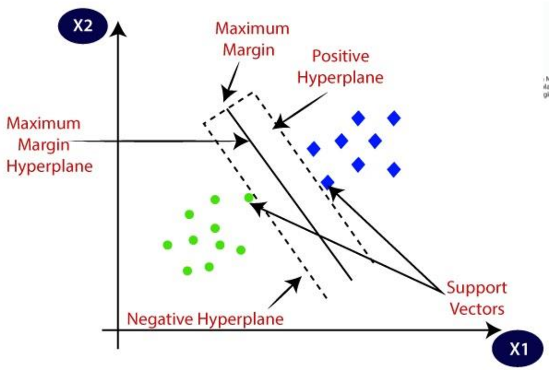


Figure 2. Data classification and support vector [17]

The best separating page is the page with the largest distance between two classes. In this case, the absolute value of *w* should be assigned the lowest value. The kernel function is defined as φ; it transfers data from the *x* space to the *z* space [17, 18].

$$\varphi : x \rightarrow z \qquad z=\varphi(x)$$

As a result, the separator equation is written as Eq. (2):

$$w^T z + b = f(x) = 0 \rightarrow w^T \varphi(x) + b = 0 \tag{2}$$

In Eq. (2), which is the relationship between the target variable and the input variables, φ(x) is the kernel, *f*(*x*) is the target function, *w* is the vector coefficient, and *b* is a constant value. The objective function, which is a convex function, has an optimal solution. Equation (3) is used to find the optimal page [12].

$$\min \frac{1}{2} \sum\nolimits_{i=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j y_i y_j \varphi(x_i)^T \varphi(x_j) - \sum_{i=1}^{n} \alpha_i \qquad 0 \le \alpha_i \le c \tag{3}$$

Instead of the expression $\varphi(x_j)\varphi(x_i)T$, a more general state can be defined as $k(x_i,x_j)$, which can be used to enter more complex spaces. The process is often called a kernel trick. As a result, Eq. (3) can be presented as Eqs. (4) and (5).

$$y = sign(\sum_{i=1}^{n} \alpha_i y_i k(x_i, x) + b) \tag{4}$$

$$b = \frac{1}{|s|} \sum_{i=1}^{n} [y_i - \sum_{j=1}^{n} \alpha_j y_j k(x_j, x)] \tag{5}$$

2.4. Algorithm of K-NN

The K-NN method is a non-stationary statistical method that is used for classification and regression. In both of the above cases, K-NN includes the closest training sample in the data space and its output varies depending on the type used in classification and regression. In the classification mode, it calculates the distance of a point marked by a label with the nearest points. Since the calculations of this algorithm are based on distance, data normalization can help improve its performance (Figure 3).

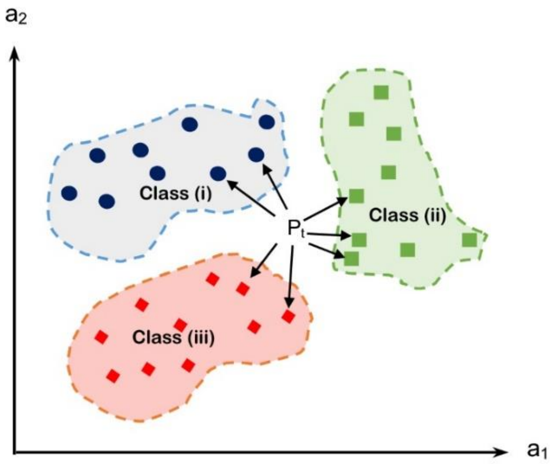


Figure 3. K-NN classification

The primary data are vectors in a multidimensional space, each of which contains a label called a category. The learning mode or training mode of the algorithm includes the storage of the feature vectors and the class label of the prototype. In the classification phase, K is a constant defined by the operator and the unlabeled vector is has the largest number in K-NN for that point. Distance measurements for continuous variables are usually a Euclidean distance. The K-NN algorithm includes the following steps:

Step 1: Load data
Step 2: Select K as the nearest neighbor number
Step 3: For each of the primary datapoints:

- The distance between the data in question and each of the primary data is calculated.
- Add the distance and index of the prototype.
- Sort the collection based on distance from small to large.
- Select K points of the first member of the sorted set.
- Output results.

There are a number of advantages with the K-NN method. First, there is no presupposition about the data. Second, it has a simple algorithm but yields relatively high accuracy. Third, it is a multipurpose method (classification and regression). Among its disadvantages, since the algorithm stores all the previous data, it requires a lot of memory. The prediction stage with a large K may be slow. The flowchart of data evaluation steps from start to finish is shown in Figure 4.

CC BY

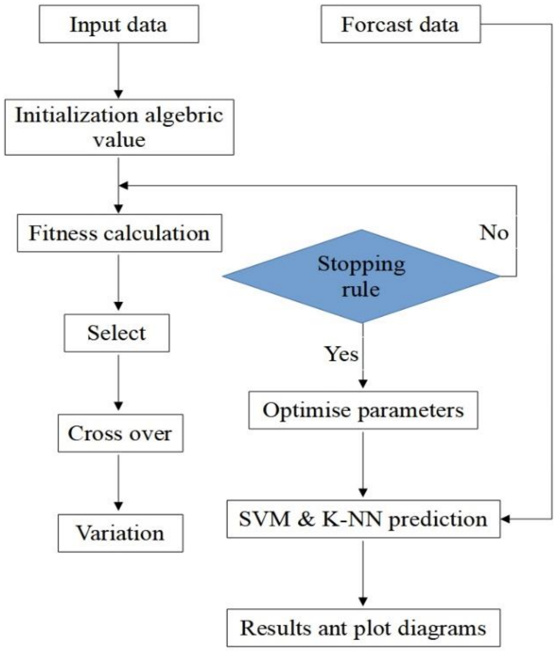


Figure 4. SVM and K-NN modelling and configuration flowchart

2.5. Dimensional analysis

Taking into account the geometric and hydraulic characteristics of the flow research models according to the research of [18], the effective parameters for the inclined drop along with the fishway elements are presented in the form of Eq. (6).

$$f_1(q, H, \theta, y_u, y_c, y_d, \mu, \rho, g, \alpha, a, b, x, L_{air}, L_j, E_u, E_d) = 0 \tag{6}$$

Dimensional analysis was performed by using the Buckingham-Π method, Eq. (7) was obtained.

$$f_2(\frac{y_u}{H}, \frac{y_c}{H}, \frac{y_d}{H}, \theta, \alpha, \frac{a}{H}, \frac{b}{H}, \frac{x}{H}, \frac{L_{air}}{H}, \frac{L_j}{H}, \frac{q}{H\sqrt{gH}} = Fr, \frac{q\rho}{\mu} = \mathrm{Re}, E_u, E_d) = 0 \tag{7}$$

Then we have:

$$\frac{\Delta \mathrm{E}}{E_u} = f_3(\frac{y_c}{H}, \frac{y_d}{H}, \theta, \alpha, \frac{a}{b}, \frac{x}{H}, \frac{L_{air}}{H}, \frac{L_j}{H}, \frac{q}{H\sqrt{gH}} = Fr_{0,d}, \frac{q\rho}{\mu} = \mathrm{Re}_u) \tag{8}$$

CC BY

Since the range of Reynolds number in this research is between 39000 and 58000, which represents the turbulent flow regime, the Reynolds number can be ignored. Also, parameters x/H and θ were omitted because they have fixed values. Thus, we have:

$$\frac{\Delta E}{E_u}, \frac{y_d}{H_d} = f_4\left(Fr_0, \frac{y_c}{H}, \alpha\right) \tag{9}$$

$$\frac{\Delta E}{E_u}, \frac{y_d}{H} = f_4\left(Fr_0, \frac{y_c}{H}, \frac{a}{b}, \alpha\right) \tag{10}$$

2.6. Evaluation parameters

Three statistical parameters are used to predict the relative energy consumption and the relative downstream depth using SVM and K-NN algorithms. RMSE, R^2, and the Kling-Gupta coefficient (KGE) are presented in the Eqs. (11) to (13), respectively. The index *E* is related to the experimental values, *P* is related to the evaluated values, and *n* is the numbers of data.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_E - S_P)^2} \qquad \text{if RMSE} \rightarrow 0\text{: best answer} \tag{11}$$

$$R^2 = \left[\frac{\left(n\sum S_E.S_P\right) - \left(\sum S_E\right).\left(\sum S_P\right)}{\sqrt{n\left(\sum S_E^2\right) - \left(\sum S_E\right)^2}.\sqrt{n\left(\sum S_P^2\right) - \left(\sum S_P\right)^2}}\right]^2 \qquad \text{if: R}^2 \rightarrow 1\text{: best answer} \tag{12}$$

$$KGE = 1 - \sqrt{(R-1)^2 + (\beta-1)^2 + (\gamma-1)^2}, \beta = \frac{\overline{Pre}}{\overline{Obs}}; \gamma = \frac{\sigma_{Pre}/\overline{Pre}}{\sigma_{Obs}/\overline{Obs}} \qquad \text{if: KGE} \rightarrow 1\text{: best answer} \tag{13}$$

3. Results and Discussion

In this study, two scenarios have been designated; the first scenario is the evaluation of relative energy dissipation and the second scenario is related to the evaluation of the relative downstream depth of the drop. In SVM, there are different kernels, among which RBF, linear and polynomial functions are used. For the K-NN algorithm, there are three methods to measure the distance between neighbors: Euclidean, Cityblock (Manhattan) and Chebychen. In order to estimate the relative energy dissipation from an inclined drop equipped with fishway elements, different data sets were allocated to training and testing. The results of the analysis are presented in Table 2 for both methods, 75% of the data was allocated to training and the remaining 25% for testing.

Table 2. Evaluation criteria of different percentages of training and testing in the best models in SVM

Scenario no.	Criteria evaluation	60-40%	65-35%	70-30%	75-25%	80-20%
1st scenario:	RMSE	0.0765	0.0739	0.0749	0.0241	0.0788
Energy	R^2	0.767	0.799	0.821	0.962	0.737
dissipation	KGE	0.796	0.782	0.864	0.952	0.765
2nd scenario:	RMSE	0.0523	0.0511	0.0518	0.022	0.044
Downstream	R^2	0.705	0.739	0.763	0.923	0.741
depth	KGE	0.743	0.741	0.751	0.941	0.758

In the first scenario ($\Delta E/E_0$) and the second scenario (y_d/H), 13 combinations have been used to evaluate the parameters using the K-NN and SVM methods. In these combinations, $\Delta E/E_0$ and y_d/H parameters are dependent parameters and Fr_0, y_c/H, and α are independent parameters.

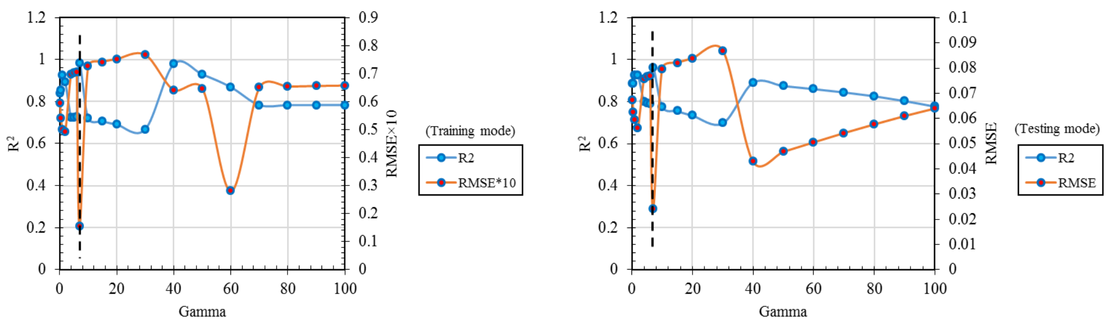
Table 3. Different input combinations in K-NN and SVM algorithms in the present study

Model no.	Input parameters	Model no.	Input parameters
First and second scenario: $\Delta E/E_0$ and y_d/H			
Model 1	α	Model 8	Fr_0, y_C/H
Model 2	a/b	Model 9	y_C/H, a/b, α
Model 3	Fr_0	Model 10	Fr_0, a/b, α
Model 4	y_C/H	Model 11	Fr_0, y_C/H, a/b
Model 5	y_C/H, α	Model 12	Fr_0, y_C/H, α
Model 6	Fr_0, α	Model 13	Fr_0, y_C/H, a/b, α
Model 7	Fr_0, a/b		

3.1. Results obtained from the SVM method

Three kernel functions, Radial Basic Function (RBF), Linear, and Polynomial, have been used in the SVM method for both scenarios. Each of the combinations presented in Table 3 is entered into the machine and the results of the best model are presented in Table 4. By analyzing the data, it was determined that the model that includes all the input parameters (model 13) has the best answer, the highest values of R^2 and KGE and the lowest value of RMSE. The value of RMSE motivates selection of this as the best model. Also, the RBF kernel function yields the best results among other kernel functions and the best model is related to the output of the RBF kernel function.

The results of the SVM method show that for the first and second scenarios, combination number 13 with inputs Fr_0, y_C/H, a/b and α has the lowest error (RMSE=0.0156), the highest correlation coefficient (R^2=0.984), and the highest Kling-Gupta coefficient (KGE= 0.961) for the first-scenario training mode. The corresponding values are RMSE= 0.0241, R^2= 0.962 and KGE= 0.952 for the first scenario testing mode. For the second scenario, the training mode has values of RMSE= 0.0212, R^2= 0.973, and KGE = 0.916; values from the testing mode are: RMSE= 0.0291, R^2= 0.971, and KGE= 0.939 and with an evaluation percentage of 75-25%. Values γ=7 and γ=10 are used, respectively for the first and second scenario (Figure 5).



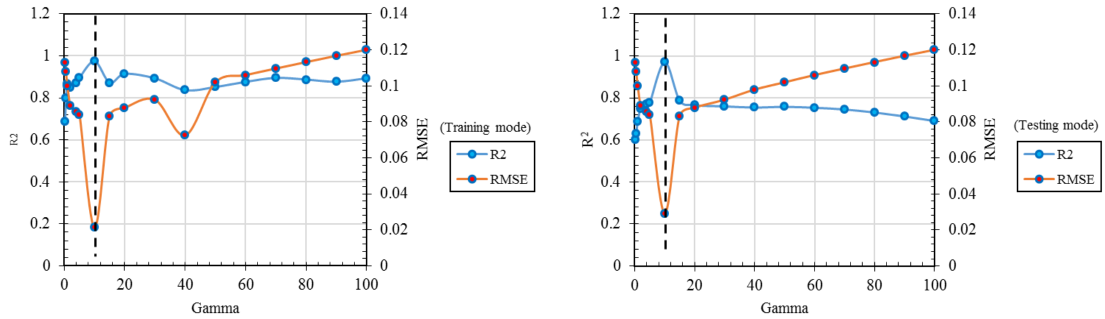
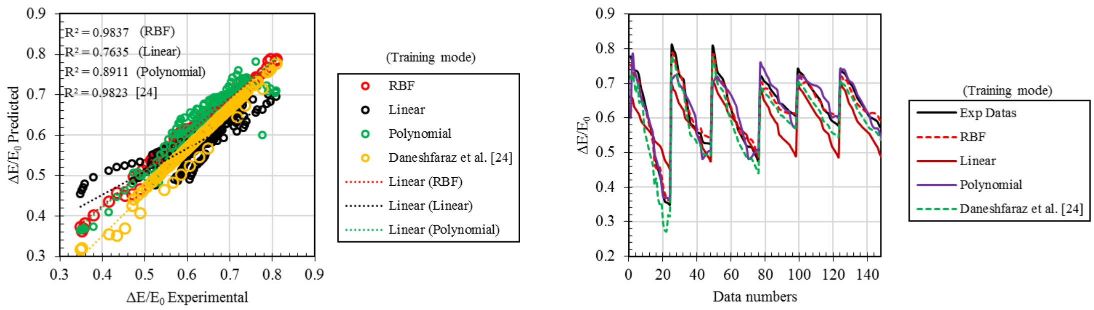


Figure 5. R^2 and RMSE Vs. γvalues for the SVM algorithm for RBF: a, b) First scenario c, d) Second scenario

Table 4. Evaluation parameters of the best combination for scenario 1 and 2 for all SVM kernels

Kernel type	Training mode			Testing mode				
	Best combination: Combination of number 13 (in table 3)							
	RMSE	R^2	KGE	RMSE	R^2	KGE	γ	Scenario
RBF	0.0156	0.984	0.961	0.0241	0.962	0.952	7	First: $\Delta E/E_0$
Linear	0.0676	0.873	0.852	0.08	0.836	0.842	---	
Polynomial	0.0888	0.956	0.946	0.0938	0.945	0.955	1	
RBF	0.0212	0.973	0.916	0.0291	0.971	0.0939	10	Second: y_d/H
Linear	0.0689	0.818	0.871	0.0817	0.761	0.859	---	
Polynomial	0.109	0.581	0.549	0.112	0.551	0.485	10	

Figures (6a, c, e, g) show the distribution curves of the predicted data for the best combination of hydraulic parameters (combination no. 13) for the first and second scenarios for both the training and testing modes. These data have less dispersion than other kernel functions and are placed on a bisector. Also, Figures (6-b, d, f, h) compare the experimental and predicted data of the superior model and show a good correlation in these two scenarios for the distribution of the RBF kernel with respect to the linear and polynomial functions. On the other hand, Comparing the data of this research with the data of research [24] for different methods of SVM algorithm shows that they match very well. So that the data of research [24] with the present research has R^2 parameter equal to 0.9823 and 0.9752 for the training and testing mode, respectively in the parameter of relative energy dissipation, and for the relative depth of downstream parameter R^2 is equal to 0.9958 and 0.982 for the training and testing mode, respectively.



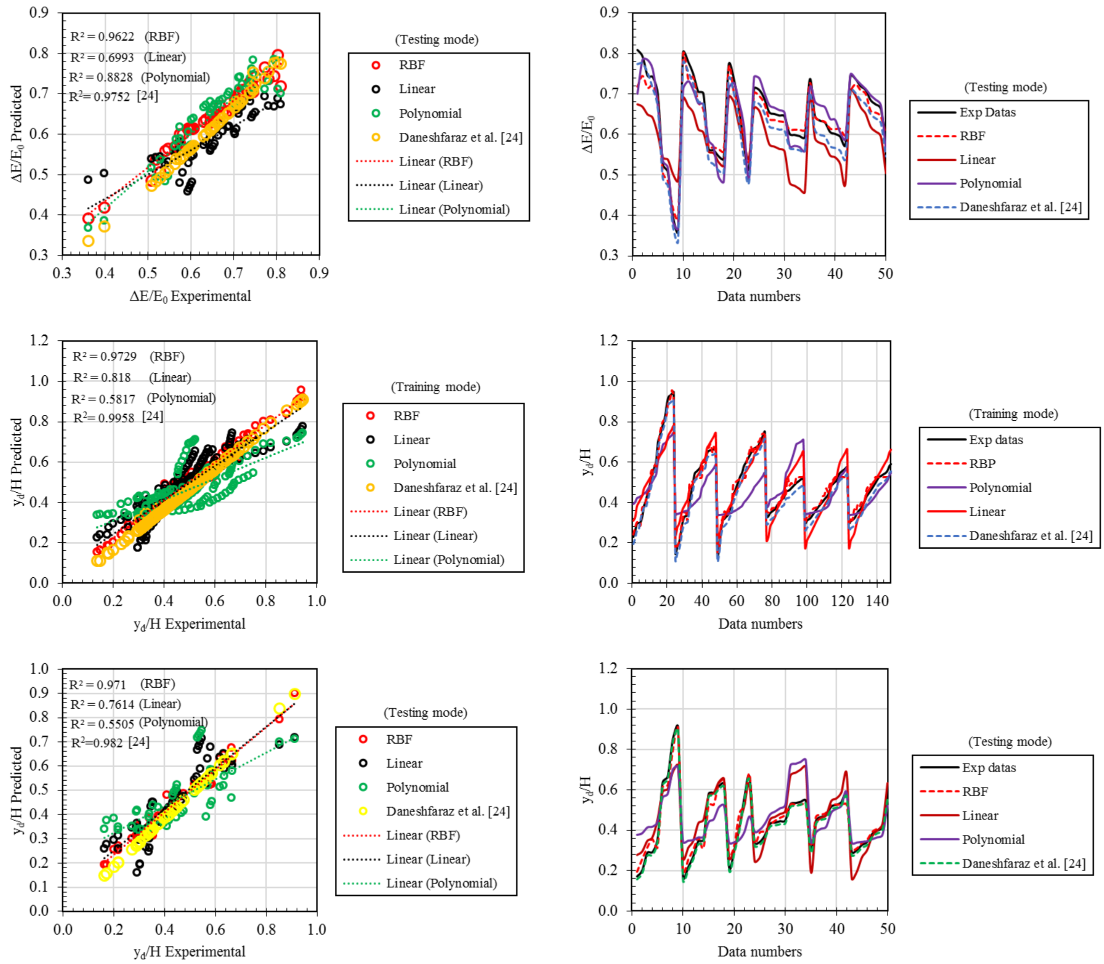


Figure 6. a, c, e, g) Distribution curve of experimental and predicted data in the training and testing modes for the best model (a, c: first scenario; e, g: second scenario)
b, d, f, h) Comparison of experimental and predicted data in the training and testing modes for the best model in the SVM algorithm (b, d: first scenario; f, h: second scenario)

3.2. K-NN Results

The K-NN pattern method is a machine learning method that is classified as a non-stationary statistical method used for regression and classification. The distance criterion for continuous variables in this algorithm is usually the Euclidean distance. Euclidean distance is the distance between two points which is obtained by using the Pythagorean theory. In the K-NN algorithm, there are three methods for distance measurement, which are Euclidean, Cityblock (Manhattan) and Chebychev methods. As with the earlier SVM results, statistical metrics for K-NN are listed in Table 3 are entered in order and the results of the best models are provided in Table 5. It was determined that combination 13, which includes all the input parameters, has the best response and has thus been selected as the best model. The working method of this algorithm is such that after selecting the distance measurement methods, a percentage is selected for the initial sample size, the number of neighbors is determined, and the correct setting of this value for the number of neighbors helps to improve the results as much as possible.

Table 5. Results of evaluation parameters of the best combination for scenario 1 and 2 for K-NN algorithm

Distance measure	RMSE	R^2	KGE	Number of Nearest Neighbors	Scenario	Size of prototype sample
	Best combination: Combination of number 13 (in table 3)					
Euclidean	0.0123	0.984	0.966			
Cityblock	0.0121	0.986	0.975	12	First: $\Delta E/E_0$	75%
Chebychev	0.0157	0.980	0.968			
Euclidean	0.0225	0.985	0.956			
Cityblock	0.0123	0.986	0.967	12	Second: y_d/H	75%
Chebychev	0.0157	0.980	0.973			

Up to a total of 100 neighbors were examined. It is clear in Figure 7 that for all distance measurement methods, 12 neighbors have been chosen as the optimal number of neighbors. Figures (7a and b) shows the changes of RMSE and R^2 against the number of nearest neighbors for the first scenario and Figures (7c and d) show the parameters for the second scenario. The expression indicates that 12 neighbors is optimal.

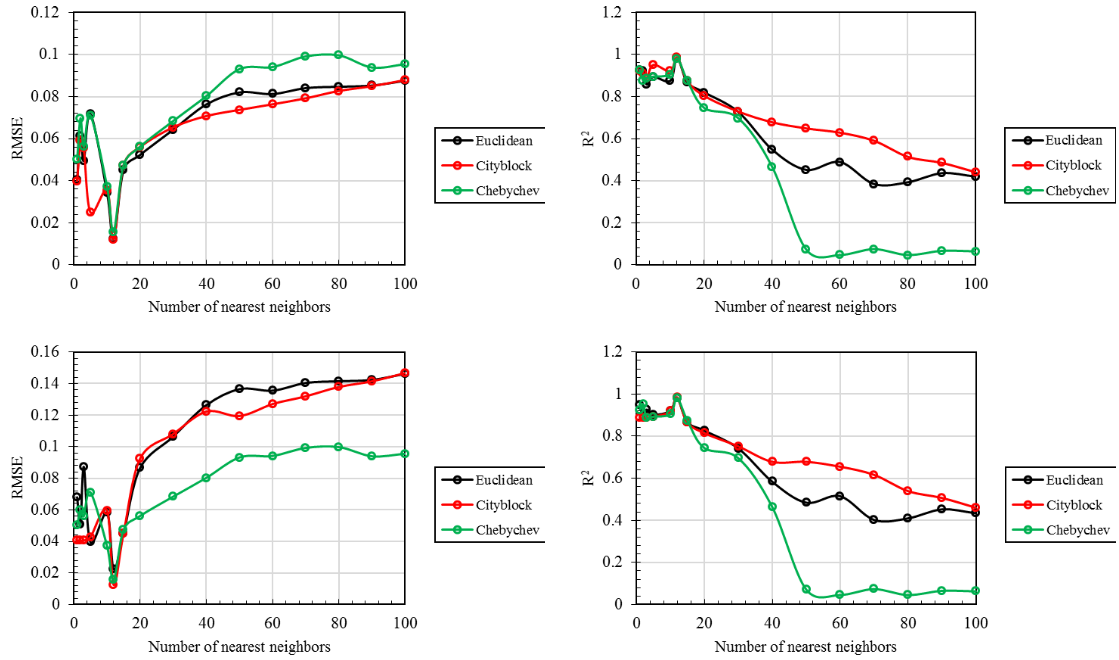


Figure 7. R^2 and RMSE Vs. number of nearest neighbors in K-NN algorithm: a, b) First scenario c, d) Second scenario

The results of the K-NN method shows that in the first and second scenarios, combination number 13 with inputs Fr_0, y_C/H, a/b and α yields RMSE = 0.0123, the highest R^2 (R^2= 0.984), and the highest Kling-Gupta coefficient (KGE= 0.966) for the Euclidean method; RMSE= 0.0121, R^2= 0.986 and KGE= 0.975 for the Cityblock method; and RMSE= 0.0157, R^2= 0.980 and KGE=

0.968 for the Chebychev method in the first scenario. For the second scenario, RMSE= 0.0225, R^2= 0.985 and KGE= 0.956 for the Euclidean method, RMSE= 0.0123, R^2 = 0.985 and KGE = 0.967 for the Cityblock method, and RMSE= 0.0157, R^2= 0.980 and KGE= 0.973 for the Chebychev method.

Figures (8a and c) show the distribution curves of the experimentally predicted data of the best combination of hydraulic parameters (combination no. 13) in the first and second scenarios using 75% of the data. The data related to this combination in the K-NN algorithm and in all three methods for distance measurement have less dispersion than the SVM kernel functions and all these methods are placed on the same median. Also, Figures (8-b and d) compare the experimental and predicted data of the superior model and show a good correlation in these two scenarios for Euclidean, Cityblock and Chebychev methods. It can be seen in these figures that the K-NN algorithm provides better results compared to those obtained from the SVM algorithm. On the other hand, with a more detailed investigation, it can be seen that by increasing the number of nearest neighbors, unfavorable results are obtained.

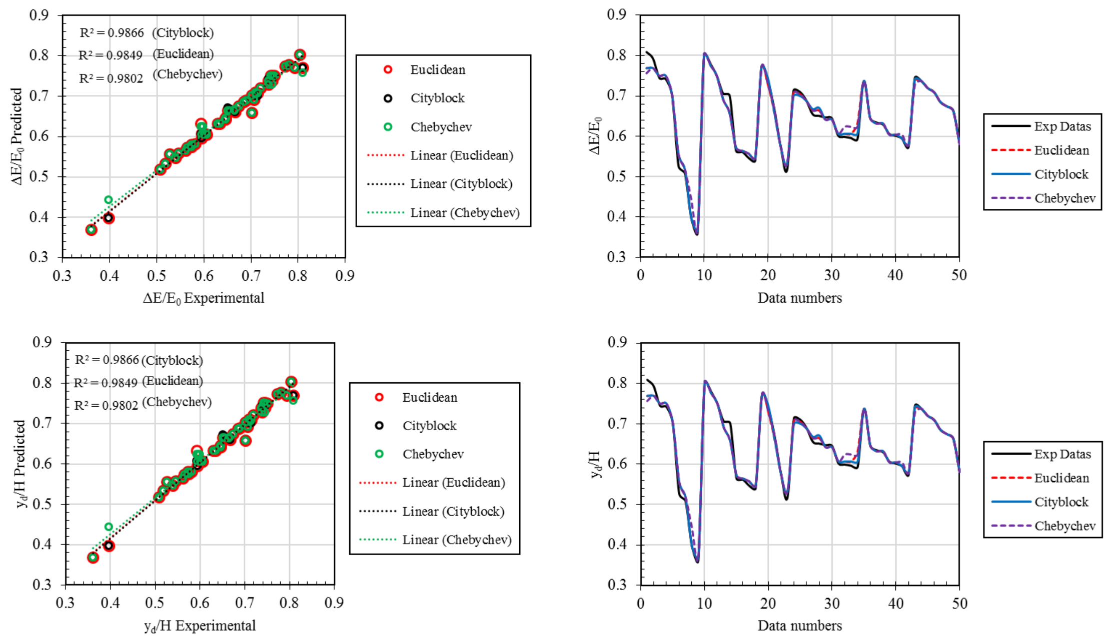


Figure 8. a, c) Distribution curves of experimental and predicted results for the best model in the K-NN algorithm (a: first scenario; c: second scenario)
(b, d, f, h) Comparison of experimental and predicted data for the best model in the K-NN algorithm (b: first scenario; d: second scenario)

Figure 9 shows the violin plot of the basic distribution of the data. From this figure, it can be concluded that the best SVM match of the evaluated data with the experimental data is related to the RBF kernel function, but the best match of the data occurs with the K-NN algorithm. Among the distance measurement methods in the K-NN method, all three methods have more accuracy and better results and compared to each other. Compared to the SVM algorithm, the results of the K-NN method are in good agreement with the experimental data.

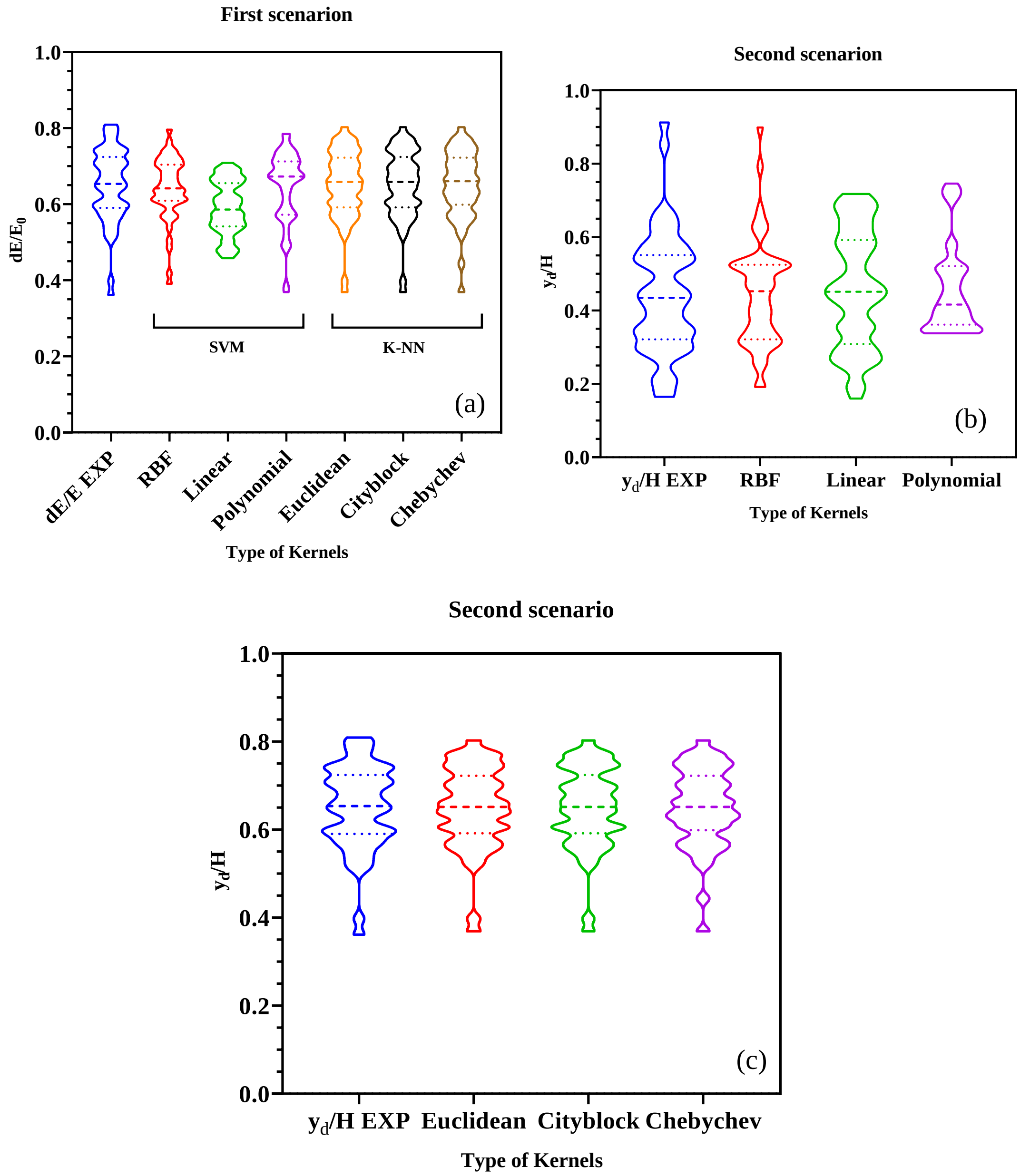


Figure 9. Violin plot for predicted and experimental data: a) First scenario, b) Second scenario, in SVM c) Second scenario in K-NN

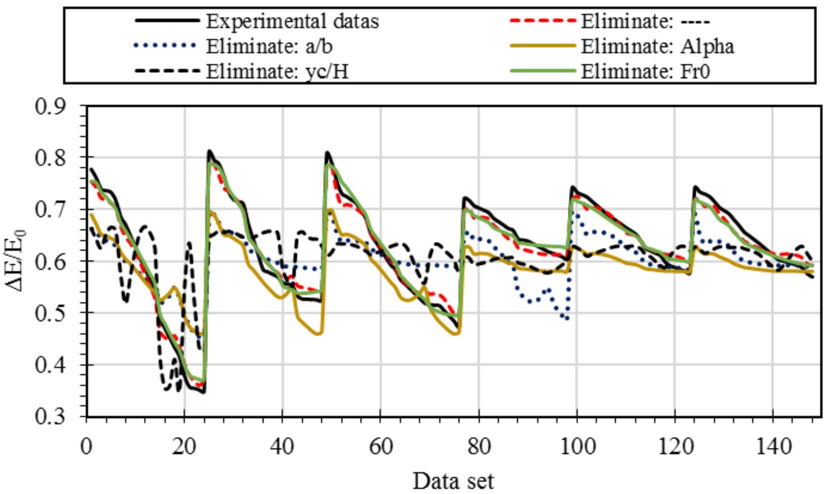
3.3. Sensitivity Analysis

Sensitivity analysis examines the important of the input variables of a statistical model at affecting the output results. By removing the input parameters one by one in both scenarios, and by using dimensional analysis, the parameter that had the most impact in predicting the relative energy dissipation and the relative depth of the downstream can be identified. Results of this sensitivity analysis are listed in Table 6.

÷

Input parameters	Eliminate parameter	Training		Testing	
		R^2	RMSE	R^2	RMSE
		First scenario ($\Delta E/E_0$)			
Fr_0, y_C/H, a/b, α	----	0.984	0.0156	0.962	0.0241
Fr_0, y_C/H, α	a/b	0.656	0.0676	0.568	0.0806
Fr_0, y_C/H, a/b	α	0.846	0.0673	0.892	0.0739
Fr_0, a/b, α	y_C/H	0.281	0. 1039	0.0963	0.1024
y_C/H, a/b, α	Fr_0	0.985	0.0138	0.966	0.02
		Second scenario (y_d/H)			
Fr_0, y_C/H, a/b, α	----	0.0156	0.984	0.0241	0.962
Fr_0, y_C/H, α	a/b	0.073	0.802	0.0832	0.605
Fr_0, y_C/H, a/b	α	0.0502	0.905	0.0478	0.912
Fr_0, a/b, α	y_C/H	0.138	0.267	0.154	0.127
y_C/H, a/b, α	Fr_0	0.0302	0.966	0.0328	0.959

Based on the sensitivity analysis carried out in the present research, it was found that for both the first and second scenarios, the independent parameter of the relative critical depth (y_c/H) has the greatest effect on the relative energy dissipation and the relative depth of the downstream according to Table 5 and Figure 10. According to Figure 10, in both scenarios, it is clear that the models in which the relative critical depth parameter is removed from the input parameters have a very bad match with the experimental data.



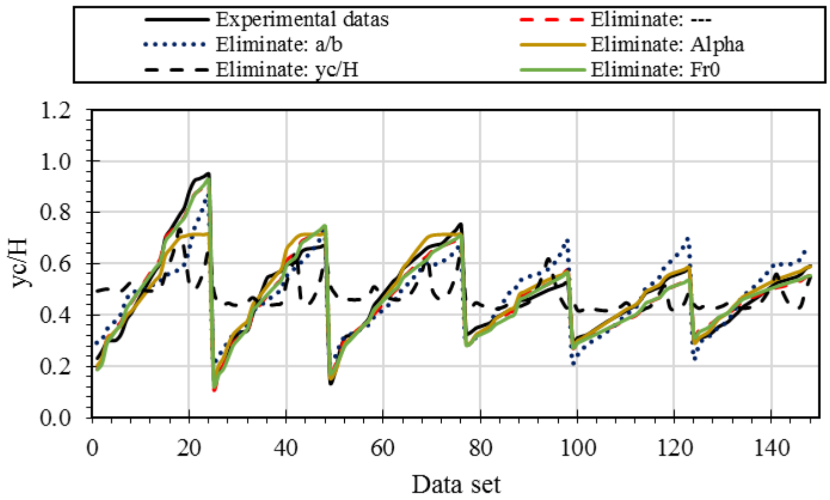
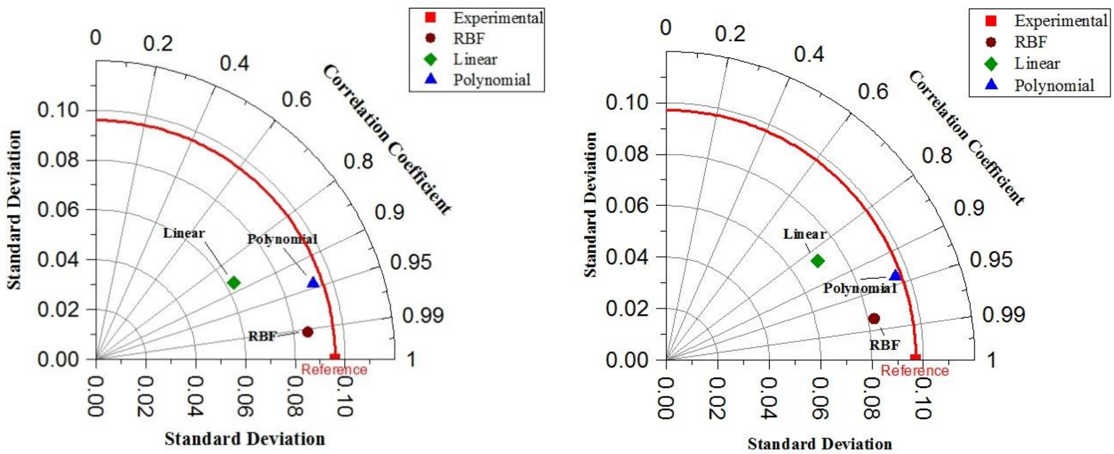


Figure 10. Comparison of laboratory data with results evaluated in sensitivity analysis: a) Relative energy dissipation, b) Relative downstream depth

In order to analyze and evaluate the models used in this research, according to Figure 11, Taylor diagram was used. As shown in Figure 11, one of the advantages of the Taylor diagram is that it uses two statistical indicators, correlation coefficient and standard deviation. Based on the Taylor diagram, the closer the predicted value is to the observed value in terms of correlation coefficient and standard deviation, the better the assessment and prediction has been made. The Taylor diagram in Figure 11 shows that the performance of the K-NN algorithm in all the investigated models, which includes models Eucliden, Chebichev, and Cityblock, has the highest performance and efficiency compared to the SVM model. Because the predicted standard deviation has the closest distance to the standard deviation of the observed data and shows the highest correlation coefficient. Based on all evaluation criteria, it can be stated that the models investigated in the present research have a good performance in estimating the parameters of relative energy dissipation and relative depth downstream. But among the reviewed models, the K-NN algorithm is more accurate than the SVM algorithm. Of course, this point should also be kept in mind that in the SVM algorithm, the linear model is the most consistent in observational data with prediction.



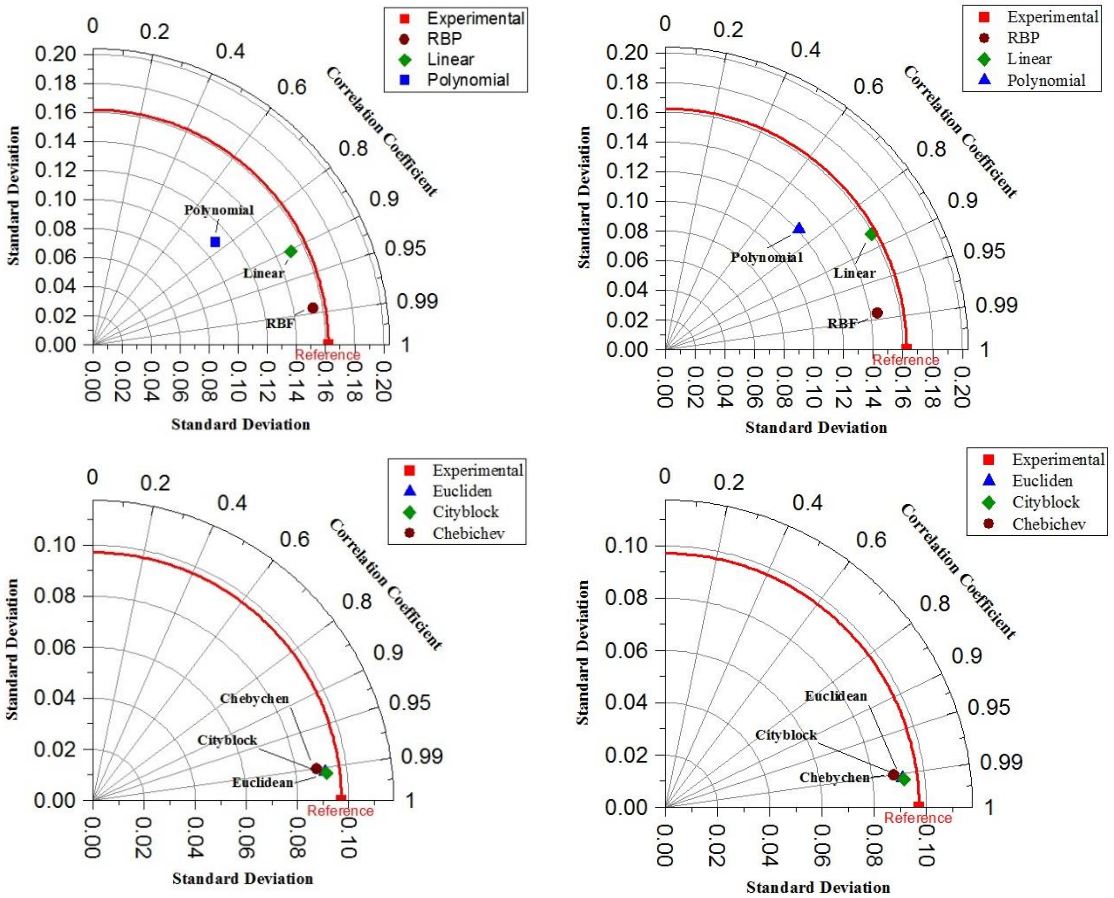


Figure 11. Taylor diagram in the investigated models of present research

4. Conclusion

In the present study, the parameters of relative energy dissipation and relative downstream depth in an inclined drop equipped with fishway elements have been evaluated using K-NN and SVM algorithms. A total of 395 data points with the same experimental conditions were used. It was shown that the prediction of these two parameters in both scenarios, a combination that included all independent input parameters, provided the best combination. In both scenarios, combination number 13 with independent input parameters Fr_0, y_C/H, a/b and α was superior and its results are as follows:

In the SVM algorithm, the Radial Basic Function (RBF) kernel type provides the best result among the kernel functions investigated in this research. In the K-NN algorithm, the Euclidean, Cityblock and Chebychev distance measurement methods have provided results that are in good agreement with the experimental data. The first scenario ($\Delta E/E_0$), in the SVM algorithm results were quantified with R^2, RMSE and KGE evaluation values equal to 0.984, 0.0156 and 0.961 respectively for the training mode and 0.962, 0.0248 and 0.952 for testing mode.

The second scenario (y_d/H), in the SVM algorithm resulted in R^2, RMSE and KGE values equal to 0.973, 0.0212, and 0.916, respectively, for training mode and 0.971 and 0.0291 and 0.931 for the testing mode and RBF kernel function. In the first and second scenarios with the

K-NN algorithm, the evaluation values in all three distance measurement methods are in the range of $0.98 \leq R^2 \leq 0.986$, $0.0123 \leq$ RMSE ≤ 0.0225 and $0.956 \leq$ KGE ≤ 0.973. By examining the results, it was found that the K-NN algorithm provides better results compared to the SVM algorithm.

The sensitivity analysis showed that the relative critical depth parameter has the greatest effect in predicting energy dissipation and downstream relative depth. In the first scenario, by removing the desired parameter, the evaluation parameters RMSE and R^2 are respectively equal to 0.1039 and 0.281 in the training mode and 0.1024 and 0.0963 in the testing mode. For the second scenario these quantities are, respectively, equal to 0.267 and 0.138 in the training mode and 0.127 and 0.154 in the testing mode.

Notation

q: The flow rate per unit width
H: The height of the drop
θ: The angle of the drop with the horizon
y_u: The upstream depth of the drop
y_c: The critical depth
y_d: The downstream depth of the drop
μ: Dynamic viscosity
ρ: Density
g: The acceleration of gravity
α: The placement angle of the fishway elements,
a and b: The length and height of the opening of the elements,
x: The distance between the elements,
L_{air}: The length of the aeration,
L_j: The length of the hydraulic jump,
E_u: The upstream energy of the drop,
E_d: The specific energy downstream of the drop.

References

1. Zhang D, Qu Y, Shi X, Liu Y, Jiang, C, (2024). Design of a Novel Multislot and Pool‑Weir Combined Fishway Based on Hydraulic Properties Analysis and Fish-Passage Experiments. Journal of Hydraulic Engineering, 150(3), 04024004.
2. Zhong Z, Ruan T, Hu Y, Liu J, Liu B, Xu, W, (2021). Experimental and numerical assessment of hydraulic characteristic of a new semi-frustum weir in the pool-weir fishway. Ecological Engineering, 170, 106362.
3. Zheng T, Niu Z, Sun S, Shi J, Liu H, Li G, (2020). Comparative Study on the Hydraulic Characteristics of Nature-Like Fishways. Water, 12(4), 955.
4. Baki A.B.M, Azimi A.H, (2021). Hydraulics and design of fishways II: vertical-slot and rock-weir fishways. Journal of Ecohydraulics, 1-13.
5. Harris J.H, Peirson W.L, Mefford B, Kingsford R.T, Felder S, (2020). Laboratory testing of an innovative tube fishway concept. Journal of Ecohydraulics, 5(1), 84-93.
6. Cea L, Pena L, Puertas J, Vázquez-Cendón ME, Peña E, (2007). Application of several depth-averaged turbulence models to simulate flow in vertical slot fishways. Journal of Hydraulic Engineering, 133(2), 160-172.

7. Rodríguez TT, Agudo JP, Mosquera LP, González EP, (2006). Evaluating vertical-slot fishway designs in terms of fish swimming capabilities. Ecol. Eng, 27(1), 37–48.
8. Yagci O, (2010). Hydraulic aspects of pool-weir fishways as ecologically friendly structure, Elsevier, Ecological Engineering, 36(1), 36-46.
9. Ariffin J, Ghani AA, Zakaria NA, Yahya AS, (2004). Sediment prediction using ANN and regression approach. 1st International Conference on Managing Rivers in the 21st Century.
10. Alp M, Cigizoglu K, (2007). Suspended sediment load simulated by two artificial neural network methods using hydrometeorological data. Environmental Modeling and Software, 22(1), 2-13.
11. Goel A, Pal M, (2009). Application of support vector machines in scour prediction on grade-control structures. Engineering Applications of Artificial Intelligence, 22(2), 216-223.
12. Roushangar K, Valizadeh R, Ghasempour R, (2017). Estimation of hydraulic jumps characteristics of channels with sudden diverging side walls via SVM. Water Science and Technology, 76(7), 1614-1628.
13. Sadeghfam S, Daneshfaraz R, Khatibi R, Minaei O, (2019). Experimental studies on scour of supercritical flow jets in upstream of screens and modelling scouring dimensions using artificial intelligence to combine multiple models (AIMM). Journal of Hydroinformatics 21(5), 893-907.
14. Daneshfaraz R, Bagherzadeh M, Esmaeeli R, Norouzi R, Abraham J, (2021a). Study of the performance of support vector machine for predicting vertical drop hydraulic parameters in the presence of dual horizontal screens. Water supply 21(1), 217-231.
15. Fuentes-Pérez J.F, García-Vega A, Bravo-Córdoba F.J, Sanz-Ronda, F.J, (2021). A step to smart fishways: an autonomous obstruction detection system using hydraulic modeling and sensor networks. Sensors, 21(20), 6909.
16. Seifollahi M, Lotfollahi-Yaghin MA, Kalateh F, Daneshfaraz R, Abbasi S, Abraham J, (2021). Estimation of the local scour from a cylindrical bridge Pier using a compilation wavelet model and artificial neural network. Journal of Hydraulic Structures, 7(3), 1-22.
17. Daneshfaraz R, Aminvash E, Mirzaei R, Abraham J, (2021b). Predicting the energy dissipation of a rough sudden expansion rectangular stilling basins using the SVM algorithm. Journal of Applied Research in Water and Wastewater 8(2), 98-106.
18. Daneshfaraz R, Aminvash E, Ghaderi A, Abraham J, Bagherzadeh M, (2021c). SVM performance for predicting the effect of horizontal screen diameters on the hydraulic parameters of a vertical drop. Applied Sciences 11(9), 4238.
19. Kalathil S.T, Chandra V, (2023). Experimental and numerical investigation on the hydraulic design criteria for a step-pool nature-like fishway. Progress in Physical Geography: Earth and Environment, 47(6), 831-851.
20. Baharvand S, Lashkar-Ara B, (2023). Hydrodynamic and Biological Assessment of Modified Meander C-type Fishway to pass Rainbow Trout (Oncorhynchus mykiss) Fish Species. Scientia Iranica.
21. Mirkhorli P, Ghaderi A, Alizadeh Sanami F, Mohammadi M, Kuriqi A, Kisi O, (2024). An Investigation on Hydraulic Aspects of Rectangular Labyrinth Pool and Weir Fishway Using FLOW-3D. Arabian Journal for Science and Engineering, 1-27.
22. Baharvand S, Jozaghi A, Fatahi-Alkouhi R, Karimzadeh S, Nasiri R, Lashkar-Ara B, (2021). Comparative study on the machine learning and regression-based approaches to predict the hydraulic jump sequent depth ratio. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 45(4), 2719-2732.

23. Cassan L, Miranda F.C, Laurens P, Courret D, (2023). Hydraulic of rock-ramp fishway with lateral slope. Environmental Fluid Mechanics, 23, 1-18.
24. Daneshfaraz R, Aminvash E, Bagherzadeh M, Ghaderi A, Kuriqi A, Najibi A, Ricardo A.M, (2021d). Laboratory investigation of hydraulic parameters on inclined drop equipped with fishway elements. Symmetry, 13(9),1643.

© 2024 by the authors. Licensee SCU, Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0 license) (http://creativecommons.org/licenses/by/4.0/).