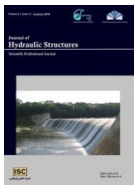
Journal of Hydraulic Structures
J. Hydraul. Struct., 2023; 9(1): 63-80
DOI: 10.22055/jhs.2023.43683.1251



OPEN ACCESS

Investigating Discharge Coefficient of Slide Gate-Sill Combination Using Expert Soft Computing Models

Yousef Hassanzadeh [1,2]
Hamidreza Abbaszadeh [1]

Abstract

The use of gate-sill combinations in recent years has been one of the new methods in increasing the hydraulic performance of gates, including the discharge coefficient (C_d). The present research aims to investigate the C_d of the gate with a sill in different dimensions in width and various positions relative to the gate using support vector machine (SVM) models, the K nearest neighbor (KNN) algorithm, and the artificial neural network (ANN) method using Statistica software. Out of 345 experimental data, 70% (241) were used for training and 30% (104) for testing. The best results are obtained when all dimensionless parameters (A_{total}/B^2, H_0/B, Z/B, ε/B, and X/B) are used. The results of different kernels showed that RBF kernel has better results in predicting C_d compared to Polynomial, Linear, and Sigmoid kernels. The results of the statistical indexes of R, KGE, RMSE, and Mean RE% for the RBF kernel in the test phase are 0.955, 0.90, 0.0192 and 1.82%, respectively. In the KNN model, Manhattan distance measure has favorable results compared to other Euclidean, Euclidean Squared, and Chebychev criteria. The results showed that the ANN method has the best performance compared to SVM and KNN models with values of 0.984, 0.976, 0.0098, and 1.15%, respectively.

Keywords: Discharge Coefficient (C_d), Gate-Sill, SVM, KNN, ANN.

Received: 06 May 2023; Accepted: 30 May 2023

1. Introduction

Slide gates are types of construction in which fluid passes below them. The most widely used of these gates are slide gates. Specifying the discharge and predicting the discharge coefficient (C_d) is significant. The control of water, the flow rate adjustment, and the flow passing the gate are done based on the opening and the estimation of the gate C_d. In recent decades, due to the lack of water resources, the need for optimal use of them has been felt more than ever. In controlling and distributing water in irrigation canals, utmost care should be taken to prevent water wastage.

[1]Department of Water Engineering, Center of Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran.
[2]Farazab Consulting Engineering Co., PMO, Tabriz, Iran. E-mail: yhassanzadeh@tabrizu.ac.ir **(Corresponding author)**

The slide gates should be chosen according to the situations of each zone; when the gate altitude goes beyond specific design criteria, more gates are needed in channels [1]. Applicability of these types of slide gates costs a lot. One of the solutions to solve this problem is to use the sill. Application of the sill increases the C_d of the gate.

In relation to investigating the (C_d) of the slide gate in normal state, many researches have been done [2-11]. In the context of the existence of a suppressed sill with the slide gates, several studies have been carried out, which can be referred to the experimental investigation of the use of polygonal and circular sills [12]. The results of Alhamid [12] showed that the C_d increases in the sill mode. Negm et al. [1] conducted an experimental investigation of the slide gate with sill under supercritical and subcritical flow conditions. They described the C_d as a function of the geometric parameters of the sill. Salmasi and Norouzi [13] investigated the various geometrical shapes of the sill on the C_d of the slide gate. They resulted that the circular one is an efficient geometrical shape. Karami et al. [14] investigated the effect of the sill parameters on the slide gates C_d using FLOW-3D. Their research indicated that the semicircular sill increases the C_d. Salmasi and Abraham [15] investigated the C_d of slide gates using polyhedral and non-polyhedral sills. Based on their results the circular sills have the high efficacy on C_d. Ghorbani et al. [16] analyzed the C_d of gates with suppressed sills using soft computing. The H_2O method has a fine efficiency in estimating the amount of C_d. Daneshfaraz et al. [17] investigated the application of sill with various width sizes on the slide gates C_d. Their results showed that the C_d in the upward location is more than the downward and below ones.

Up to now, various regression relationships have been presented to estimate the C_d of the slide gate in normal state. In addition, a few studies have been done for the suppressed sill. The investigation of the research background showed that the use of non-suppressed sills in slide gates is a new issue, and except for a few studies conducted in the last year, no study has been conducted regarding non-suppressed sills. It should be mentioned that in the past research on the mentioned topic, the theoretical relationship for estimating the flow rate through the gates was presented for the first time, which was presented in the study of Daneshfaraz et al. [17-18]. Therefore, according to the uncertainty governing the problem in these studies, it seems necessary to conduct new research in the field of using intelligent models to predict the C_d of the slide gate with non-suppressed/suppressed sill with various positions relative to the gate. In addition, no study has been conducted on the application of soft computing models in this area. For this purpose, in the current research, using Artificial Neural Network method (ANN), Support Vector Machine model (SVM), and K Nearest Neighbor algorithms (KNN), the C_d of the slide gate with sill is evaluated based on 345 experimental data.

2. Materials and Methods

2.1. Experimental set up

In this research, the experimental data of Daneshfaraz et al. [17] has been used. All the experiments were conducted in a hydraulic laboratory of University of Maragheh in a channel with 5 m length, 0.30 m width, and 0.50 m depth (Figure 1-a). Experiments have been done in the discharge, and upstream depth ranges from 0.0025 to 0.0142 m^3/s and 0.05 to 0.44 m, respectively. Two pumps with a nominal capacity of 0.0075 m^3/s have been used. The sills were used in different widths of 0.025, 0.05, 0.075, 0.10, 0.15, 0.20, 0.25, and 0.30 m, thicknesses of 0.05 m, and heights of 0.03 m in various gate openings and positions. Here, 345 experimental data were analyzed to investigate the performance of SVM, KNN, and ANN models in predicting the C_d of the slide gate with a sill.

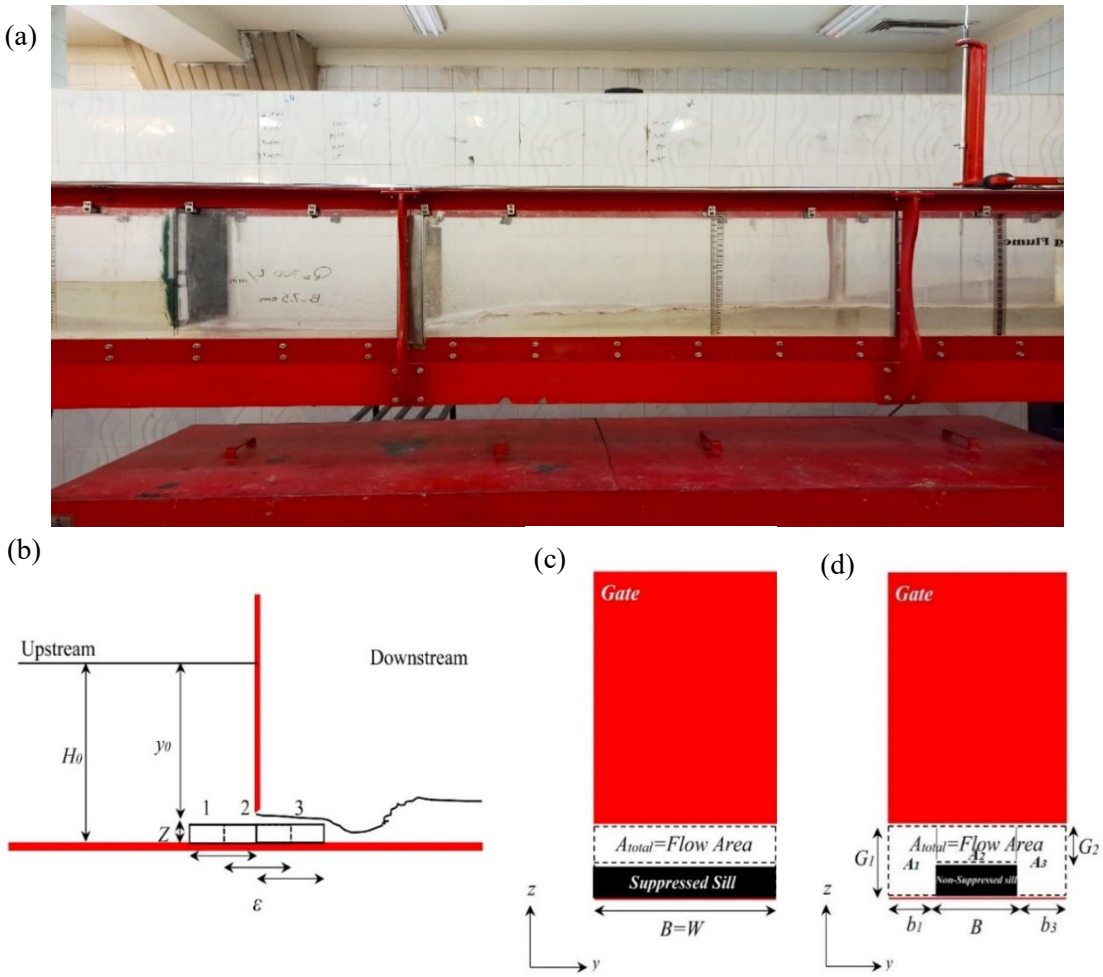


Figure 1. a) The experimental flume b) Slide gate and sills with different positions c) suppressed sill d) Non-suppressed sill [17]

2.2. Dimensional analysis

The discharge through the slide gate in non-sill state is computed as follows [3,5-8]:

$$Q = C_d W G \sqrt{2gH_0} \quad (1)$$

In Equation (1), Q represents the discharge, C_d is the coefficient of discharge, W is the width of channel, G is the opening, g is the gravitational acceleration, and H_0 is the upstream water level.

For the suppressed sill, the discharge is computed as relation (2) (Figure 1) [15-16]:

$$Q = C_d W G \sqrt{2g(H_0 - Z)} \quad (2)$$

In Equation (2), Z is the height of the sill. The discharge rate for the non-suppressed sill is calculated using relation (3) [17]:

$$Q = C_d\left(A_1\sqrt{2gH_0} + A_2\sqrt{2g(H_0 - Z)} + A_3\sqrt{2gH_0}\right) \quad (3)$$

In relation (3), A_1, A_3, and A_2 are the areas of the flow through the slide gates (Figure 1-c,d).

$$A_{total} = A_1 + A_2 + A_3 \quad (4)$$

In Equation (4), A_{total} is the flow total area. For the case with a sill, the most significant parameters that have a high effect on the C_d are as follows: [8]:

$$f_1 = (C_d . A_{total} . H_0 . B . Z . \varepsilon . W . X . \rho . g . \mu . \sigma) = 0 \quad (5)$$

In Equation (5), B is the width of the sill, ε is the thickness of the sill, X is the distance from the center to the center of the slide gate to the sill, ρ is the specific gravity, σ is the surface tension, and μ is dynamic viscosity. Using π-Buckingham method, the relation (6) can be presented:

$$f_2 \left(C_d . \frac{A_{total}}{B^2} . \frac{H_0}{B} . \frac{Z}{B} . \frac{\varepsilon}{B} . \frac{W}{B} . \frac{X}{B} . Re . We \right) = 0 \quad (6)$$

where, Re and We are the Reynolds and Weber numbers, respectively. According to the turbulent flow and $47222 \geq Re \geq 11111$, so the effect of the Re can be ignored [18-19]. When the fluid in the experiments is the same and the temperature is constant, Re and We are dependent on each other and change with the opening, then, the effect of We can be ignored [18,20-22]. In addition, the channel width parameter has assumed a constant value and is not the objective of this research, so the impact of this parameter was ignored. The examined parameters were presented in the form of Equation (7). The range of parameters obtained from the dimensional analysis is shown in Table (1).

$$C_d = f_3 \left(\frac{A_{total}}{B^2} . \frac{H_0}{B} . \frac{Z}{B} . \frac{\varepsilon}{B} . \frac{X}{B} \right) \quad (7)$$

Table 1. The range of parameters changes

no.	parameters	range of changes
1	A_{total}/B^2	$0.033 \leq A_{total}/B^2 \leq 18$
2	H_0/B	$0.296 \leq H_0/B \leq 8.6$
3	Z/B	$0.1 \leq Z/B \leq 1.2$
4	ε/B	$0.167 \leq \varepsilon/B \leq 2$
5	X/B	$-0.084 \leq X/B \leq 0.084$

The input parameters to Statistica 12 software were introduced to SVM, KNN, and ANN according to dimensional analysis in different cases (Table 2).

Table 2. The cases defined in the present research

case no.	input parameters
1	$\frac{A_{total}}{B^2}.\frac{H_0}{B}$
2	$\frac{A_{total}}{B^2}.\frac{H_0}{B}.\frac{X}{B}$
3	$\frac{A_{total}}{B^2}.\frac{H_0}{B}.\frac{Z}{B}$
4	$\frac{A_{total}}{B^2}.\frac{H_0}{B}.\frac{\varepsilon}{B}$
5	$\frac{A_{total}}{B^2}.\frac{H_0}{B}.\frac{Z}{B}.\frac{\varepsilon}{B}$
6	$\frac{A_{total}}{B^2}.\frac{H_0}{B}.\frac{Z}{B}.\frac{\varepsilon}{B}.\frac{X}{B}$

2.3. Support Vector Machine model (SVM)

Vapnik [23] first used a Support Vector Machine model as a supervised learning model for classification and estimation. The SVM is an impressive learning machine that uses the principle of induction of structural error minimization and leads to a general optimal solution. Similar to other regression problems, it is assumed that the relation between the independent and dependent variables is determined by an algebraic function such as f(x) plus a disturbance value [24].

$$f(x) = W^T \emptyset(X) + b \quad (8)$$

$$Y = f(x) + noise \quad (9)$$

In Equation (8), W, b, and ø are the vector of factors, the characteristic constant of the regression function, and the kernel function, respectively. Here, the aim is to find a functional form for f(x). It is achieved by training the model with sample. The SVM function can be rewritten as Equation (10):

$$f(x) = \sum_{i=1}^{N} \bar{a}_i \emptyset(X_i)^T \emptyset(X) + b \quad (10)$$

The a_i parameter represents the average Lagrange factors. Calculating ø(X) in its characteristic space may be very complicated. To solve this issue, the usual procedure in SVM is to choose a kernel function. The choice of the kernel for SVM depends on the size of the training data and the dimensions of the feature vector. In other words, the kernel function should be selected that has the ability to learn the inputs of the problem. In practice, four types of Linear, Polynomial, Sigmoid, and Radial Basis Function (RBF) kernels are used.

$$K(X_i.X_j) = (X_i.X_j) \quad (11)$$

$$K(X_i.X_j) = \left(1 + (X_i.X_j)\right)^d \quad (12)$$

$$K(X_i.X_j) = tanh(-a(X_i.X_j) + C) \quad (13)$$

$$K(X_i.X_j) = exp(-\|X - X_i\|^2/\sigma^2) \quad (14)$$

where, C is an integer and positive, which determines the penalty when a model-training error occurs.

2.4. K Nearest Neighbor algorithm (KNN)

The K Nearest Neighbor algorithm is a common classification method and is based on distance measurement. KNN is also known as an instance-based model or a lazy learner; because it doesn't build an internal model and learn from training data to perform discriminatively; It keeps only the training samples that are used as knowledge for the prediction phase. For K regression problems, it finds the nearest neighbor and estimates the desired value by computing the mean value of the nearest neighbors. Distance functions such as Euclidean, Euclidean Squared, Manhattan, and Chebychev are used to determine the distance measure [25].

2.5. Artificial Neural Network method (ANN)

An Artificial Neural Network method generally consists of input, hidden and output layers. The neuron is the smallest information-processing unit that forms the basis of neural networks. A neural network is a set of neurons that, by being placed in various layers, forms a specific architecture based on the connections between neurons in various layers. In the neural network, each neuron acts independently and the overall behavior of the network is the result of the behavior of many neurons. It is possible to design a data structure that acts like a neuron-using computer programming [26]. By creating a network of these interconnected artificial neurons and creating a training algorithm for the network and applying these algorithms, it can be trained. Here, the Multi-Layer Perceptron (MLP) network was used.

2.6. Statistical indexes

Here, the following statistical indexes were used to investigate the effectiveness of the used methods to predict the C_d.

$$RE\% = \frac{C_{d_{Obs}} - C_{d_{cal}}}{(C_d)_{Obs}} \times 100 \tag{15}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(C_{d_{Obs}} - C_{d_{Cal}})_i^2}{n}} \tag{16}$$

$$KGE = 1 - \sqrt{(R-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$$

$$\beta = \frac{\overline{C_{d_{Cal}}}}{\overline{C_{d\ Obs}}} . \gamma = \frac{CV_{Cal}}{CV_{Obs}} = \frac{\sigma_{Cal}/\overline{C_{d\ Cal}}}{\sigma_{Cal}/\overline{C_{d\ Obs}}}$$

$$R = \frac{[\sum_{i=1}^{n}(C_{d\ Obs\ i} - \overline{C_{d\ Obs}}) \times (C_{d\ Cal\ i} - \overline{C_{d\ Cal}})]}{\sum_{i=1}^{n}(C_{d\ Obs\ i} - \overline{C_{d\ Obs}}) \sum_{i=1}^{n}(C_{d\ Cal\ i} - \overline{C_{d\ Cal}})}$$

$$\begin{aligned} &0{,}7 < KGE < 1 \quad \text{very good} \\ &0{,}6 < KGE < 0{,}7 \quad \text{good} \\ &0{,}5 < KGE \leq 0{,}6 \quad \text{satisfactory} \\ &0{,}4 < KGE \leq 0{,}5 \quad \text{acceptable} \\ &KGE \leq 0{,}4 \quad \text{unsatisfactory} \end{aligned} \tag{17}$$

Here, RE%, RMSE and KGE are the percentage Relative Error, Root Mean Square Error and Kling Gupta Efficiency, respectively. *Obs*, *Cal* and *n* represent observational, computational, and the total of data, respectively. In relation (17), R is the correlation coefficient, β is the calculated average data relative to the observed average data, and γ is the calculated standard deviation (SD) relative to the observed standard deviation [27].

3. Results and Discussion

Different dimensionless parameters were considered as inputs of various models. The C_d of the slide gate with the sill was considered as the output and target characteristic, and the possibility of using modern data mining methods in the estimation of the C_d was tried to be evaluated. To predict

the C_d by data mining methods, in general, 70% of the data were chosen for the training phase and rest of the data were chosen for the test phase. In the present research, based on dimensional analysis, the C_d depends on various parameters. These parameters were checked according to Table (2) in different cases. The results of statistical indexes of support vector machine for these cases are shown in Table (3). The best case was chosen in such a way that it has appropriate R, RMSE, Mean RE, and KGE statistical indexes compared to the experimental results. According to Table (3), case no.6 has favorable results compared to other cases, so that for this case in the training phase R=0.960, RMSE=0.0162, Mean RE=0.0168, and KGE=0.939. In addition, for the test phase, the values of these indexes are 0.955, 0.0182, 0.0192, and 0.9, respectively. Case no.6 represents the superior case in the SVM section. Figure (2) shows the data graph of the experimental and predicted values obtained from the SVM model for the superior case. The results indicate that the trend of the changes obtained from the experimental results is the same as predicted. According to Figure (2-c, d), it can be seen that for the superior case, in the training and testing phase, a wide range of data are in the ±3% error band. So that in the training and test phase, more than 90% and 81% of the data are in the ±3% error band, respectively, which indicates the high accuracy of the solution when selecting all the effective input parameters to the SVM model.

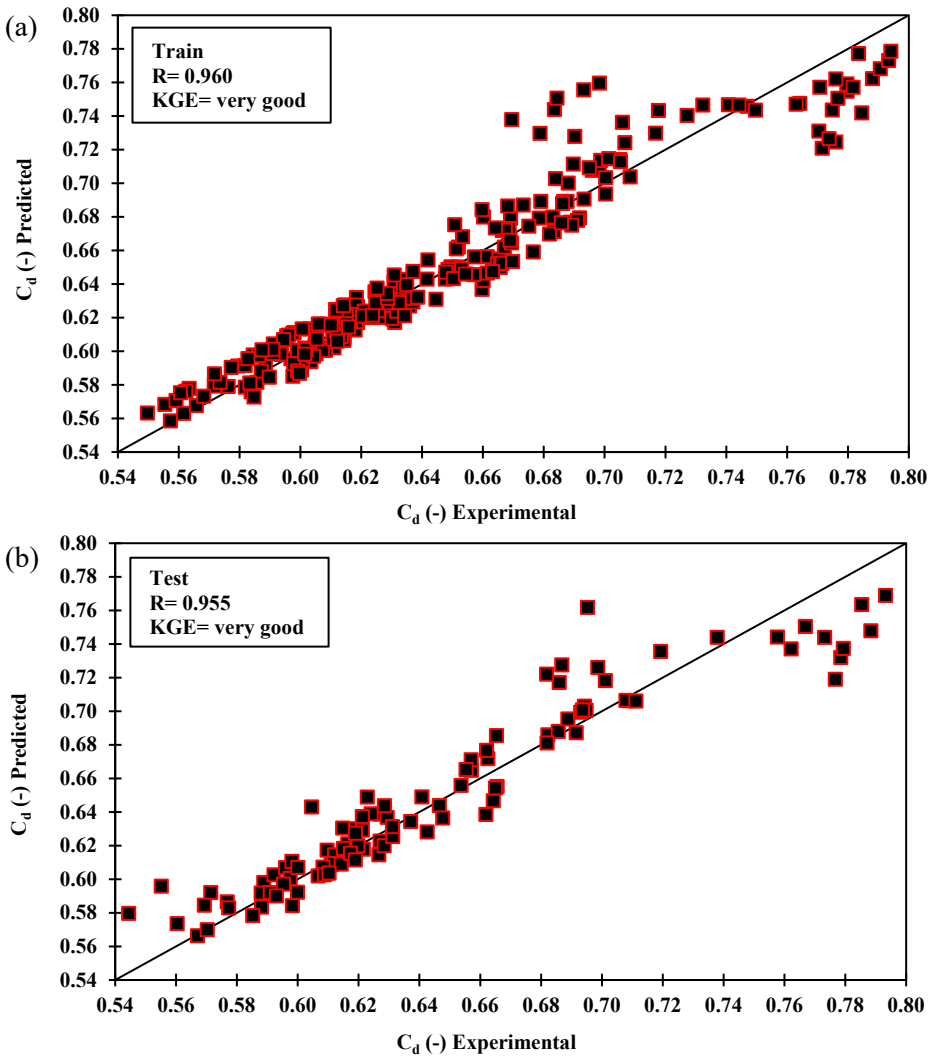
Table 3. The results of C_d prediction in different input parameters

cases no.	Train				Test			
	R (-)	RMSE (-)	KGE (-)	Mean RE (-)	R (-)	RMSE (-)	KGE (-)	Mean RE (-)
1	0.871	0.0295	0.743	0.0268	0.855	0.0326	0.695	0.0295
2	0.880	0.0292	0.722	0.0263	0.862	0.0325	0.672	0.0287
3	0.951	0.0182	0.894	0.0177	0.944	0.0207	0.840	0.0202
4	0.951	0.0181	0.916	0.0183	0.944	0.0204	0.866	0.0207
5	0.955	0.0173	0.940	0.0176	0.951	0.0188	0.901	0.0199
6	0.960	0.0162	0.939	0.0168	0.955	0.0182	0.9	0.0192

According to Table (4), among the Linear, Polynomial, RBF, and Sigmoid kernels, the RBF was selected as the best kernel for the SVM according to the results of its statistical indexes.

Table 4. The statistical indexes of various kernels in the SVM model

statistical index (test phase)	Kernels			
	Linear	Polynomial	RBF	Sigmoid
R	0.810	0.653	0.955	0.337
KGE	0.7	0.556	0.9	-2.073
RMSE	0.0356	0.0496	0.0182	0.2433
Mean RE	0.0406	0.0547	0.0192	0.2210



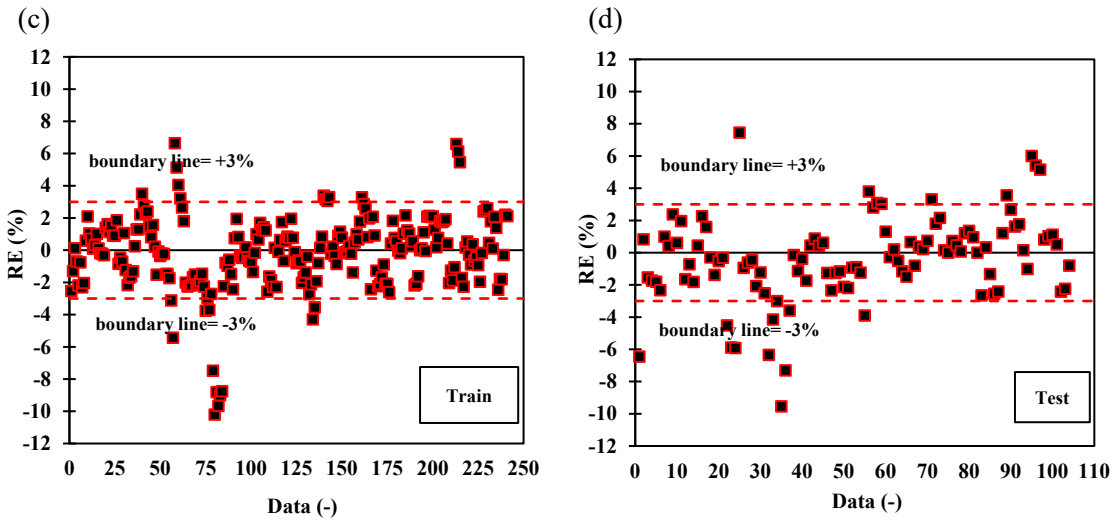
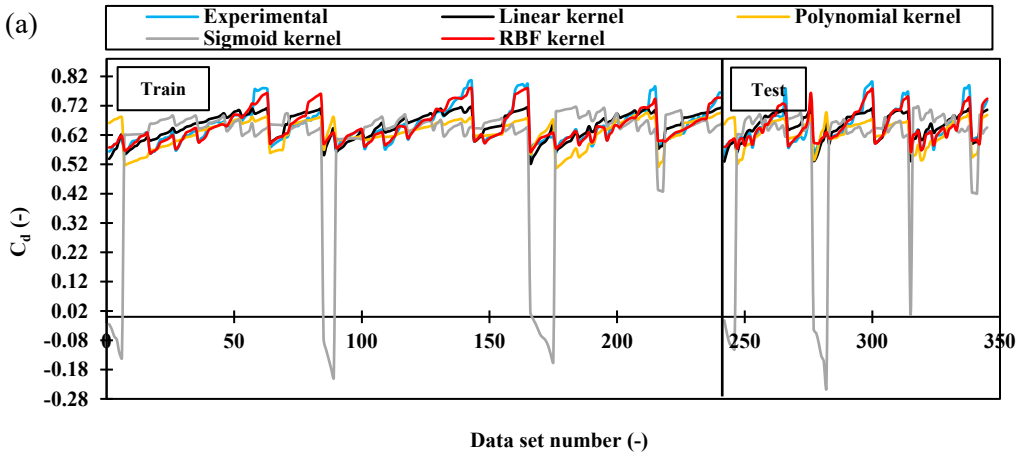
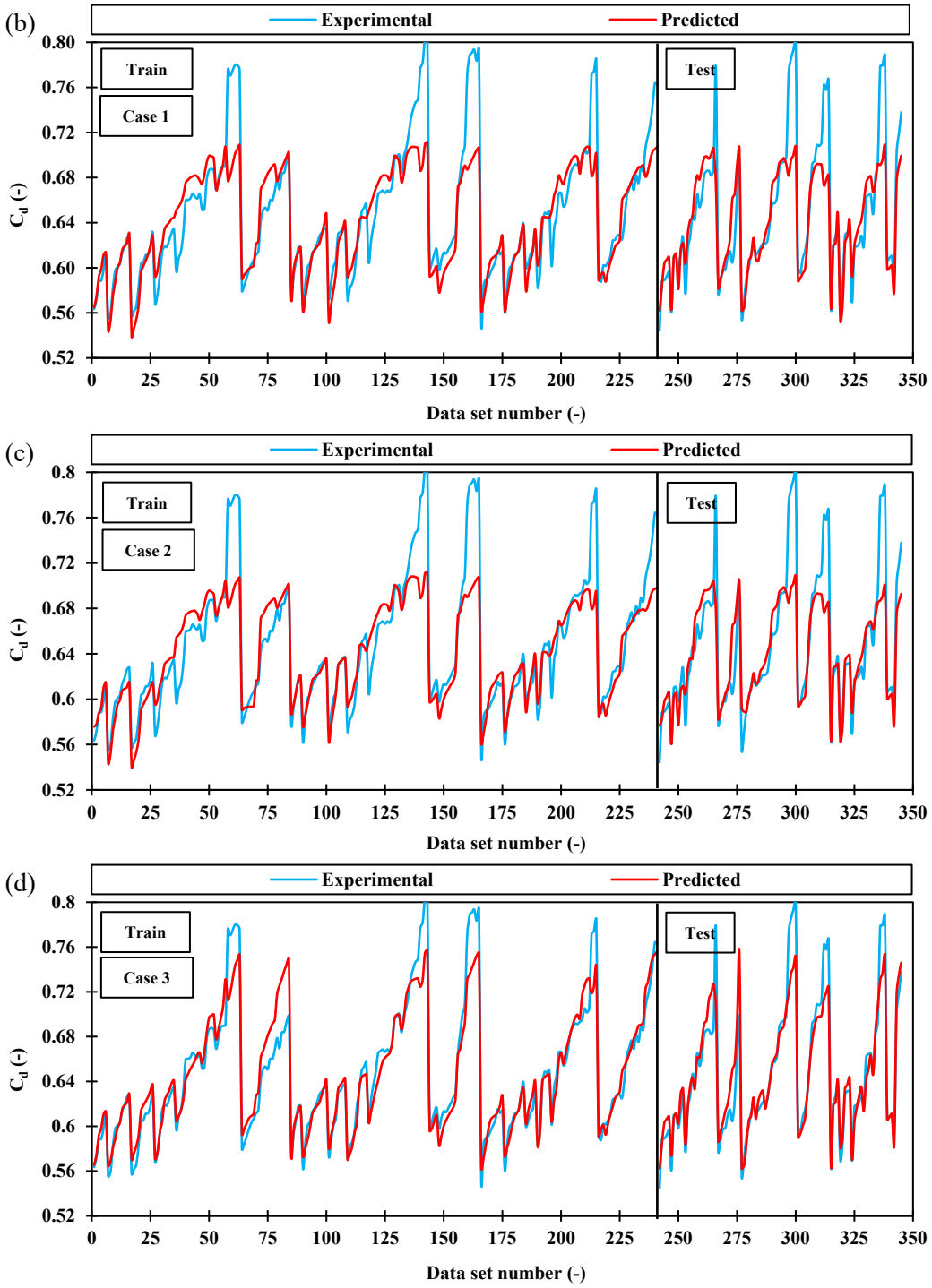


Figure 2. Experimental C_d against predicted a) training phase b) test phase, scatter percentage of relative error of data c) train phase d) test phas

In order to examine the different cases (Table 2) more precisely, the experimental and predicted C_d values for the number of data related to the training and test phases are shown in Figure (3). The experimental and predicted C_d values for different kernels are shown in Figure (3-a). As can be seen, the RBF kernel has high accuracy compared to other kernels and has predicted the C_d with high accuracy. According to Figure (3), among the cases that case no.6 was introduced as the superior model. Case no.5 also has results close to the experimental results. The results of statistical indexes for case no.5 in the training phase are R=0.955, RMSE=0.0173, KGE=940, and Mean RE=0.0176. The results of these indexes for the test phase are 0.951, 0.0188, 0.901, and 0.0199, respectively. In cases no.1 and 2, this difference is greater. The reason for this issue can be attributed to the non-use of other dimensionless parameters such as X/B, Z/B, and ε/B. Although the X parameter does not play much role: the results of cases no.5 and 6 are close to each other. The reason for this problem can be the provision of the role of this parameter by other parameters. Although the position of the sill has an effective role on increasing the C_d, its highest value is related to the upstream tangent position.




(b)
Experimental
Predicted
Train
Case 1
Test
C_d (-)
Data set number (-)
(c)
Experimental
Predicted
Train
Case 2
Test
C_d (-)
Data set number (-)
(d)
Experimental
Predicted
Train
Case 3
Test
C_d (-)
Data set number (-)

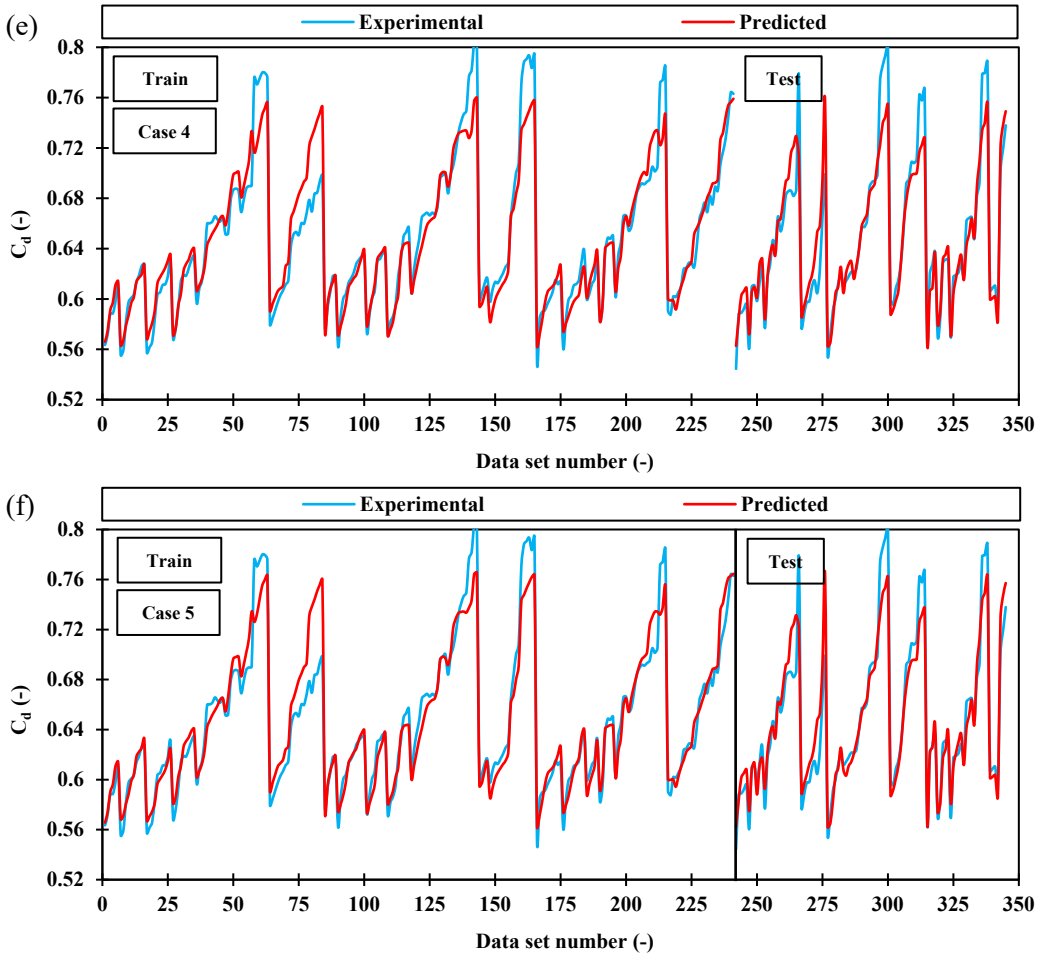


Figure 3. Experimental and predicted values of C_d in the train and test phases for different data

Table (5) shows the results of statistical indexes for different input cases in the ANN method. In this method, like the SVM model, the software randomly selected 70% of the data for training and 30% of the data for the test. The minimum and maximum number of hidden layers were chosen as 3 and 21, respectively, to perform more operations to find the best model. According to Table (5), when all dimensionless parameters are used, the accuracy of the solution increases. Although the results of other cases are close to each other, case no.6 produces better results. The results of R, RMSE, Mean RE, and KGE statistical indexes for case no.6 in the training phase are 0.992, 0.0075, 0.0089, and 0.992, respectively. These values for the test phase are 0.984, 0.0098, 0.0115, and 0.976, respectively. Figure (4) shows the architecture of the ANN method for the superior case consisting of 5 input variables, 21 hidden layers, and one output variable. The scatter plot diagram for the training and test phase of case no.6 is shown in Figure (5).

Table 5. Results of statistical indicators of different input cases in ANN method

cases no.	Train				Test			
	R (-)	RMSE (-)	KGE (-)	Mean RE (-)	R (-)	RMSE (-)	KGE (-)	Mean RE (-)
1	0.973	0.0140	0.963	0.0161	0.95	0.0172	0.95	0.0172
2	0.974	0.0138	0.961	0.0146	0.962	0.0151	0.961	0.0144
3	0.978	0.0125	0.969	0.0130	0.953	0.0168	0.953	0.0161
4	0.976	0.0133	0.965	0.0149	0.953	0.0168	0.952	0.0163
5	0.977	0.0129	0.964	0.0138	0.953	0.0167	0.953	0.0156
6	0.992	0.0075	0.992	0.0089	0.984	0.0098	0.976	0.0115

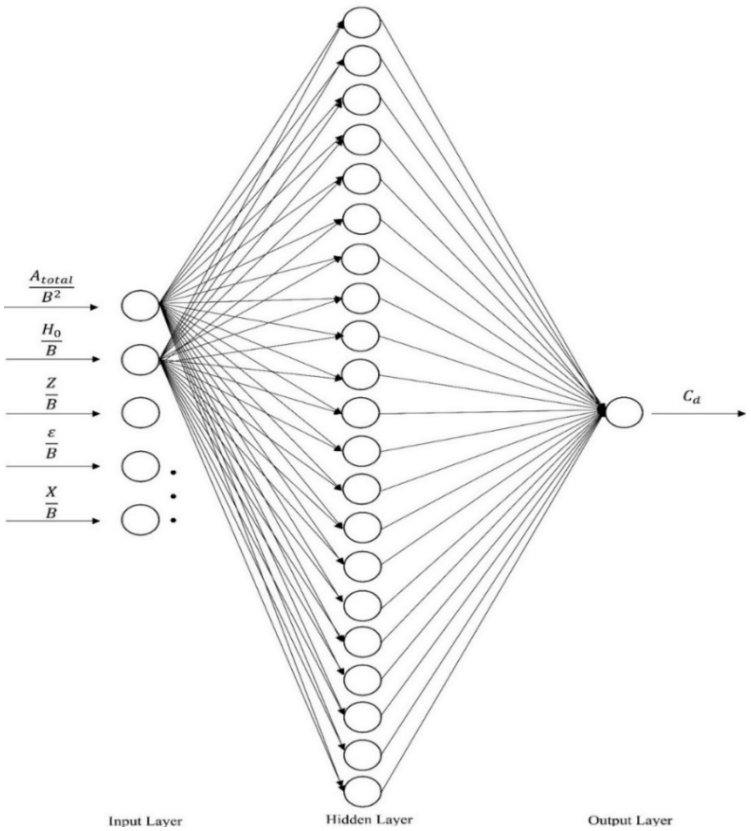


Figure 4. Architecture of ANN method for the superior case

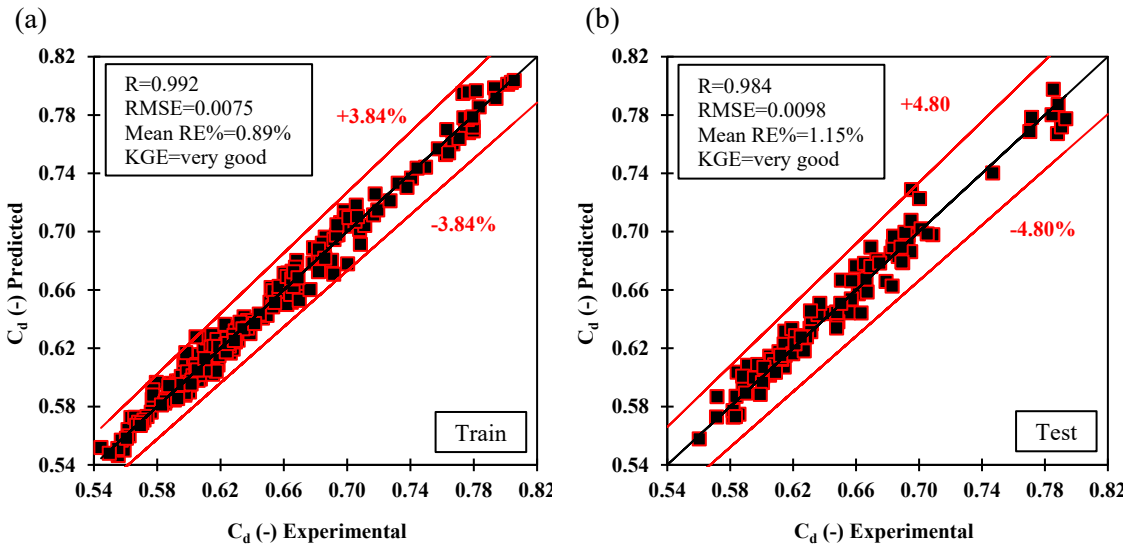


Figure 5. Experimental C_d against predicted a) train phase b) test phase

In Figure (6), the RMSE and Mean RE values for the test phase at different values of K coefficient are given. The right amount of K is the one that can determine the boundary that predicts the data with the least error. Examining the coefficient of close neighbor showed that the best result is obtained when the coefficient 2 is used. In K equal 2, the RMSE and Mean RE have the lowest values. In Figure (7), the C_d values for different distance measure criteria are presented. In the Manhattan distance criteria, the results are close to the experimental results, and the results of the statistical indexes R, RMSE, Mean RE%, and KGE are 0.965, 0.016, 1.70%, and 0.96, respectively. The results are close to each other in Euclidean and Euclidean Squared distance measures. Both mentioned distance measures are within the relative error range of ±8.39%. However, the statistical indexes mentioned on each chart have differences from each other. The Chebychev distance measure criteria has the worst results of statistical indexes.

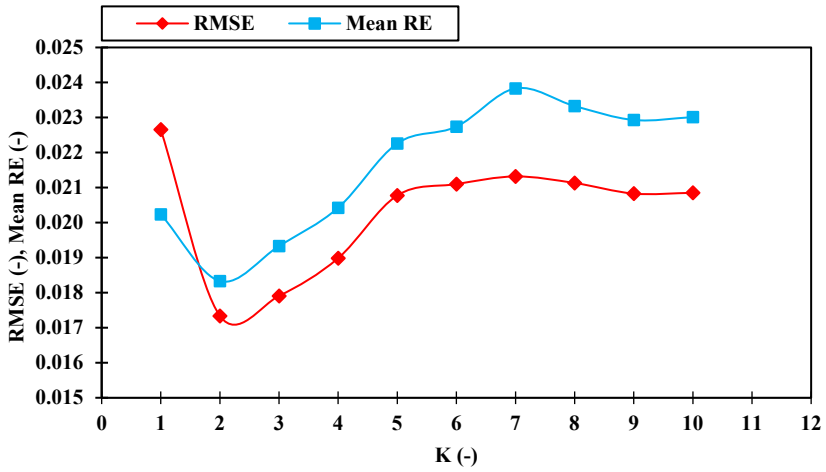


Figure 6. The values of RMSE and Mean RE for different K coefficient

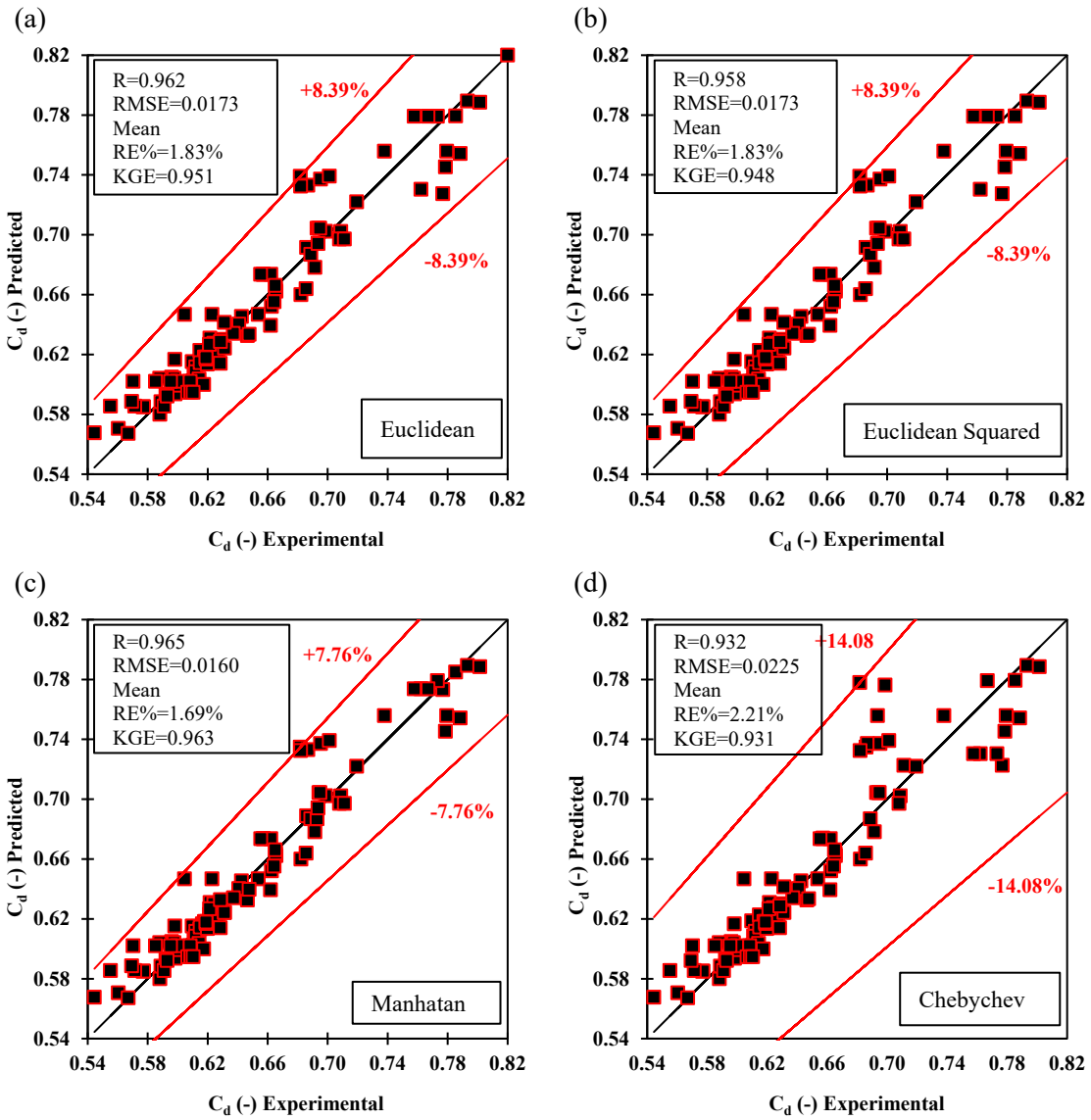
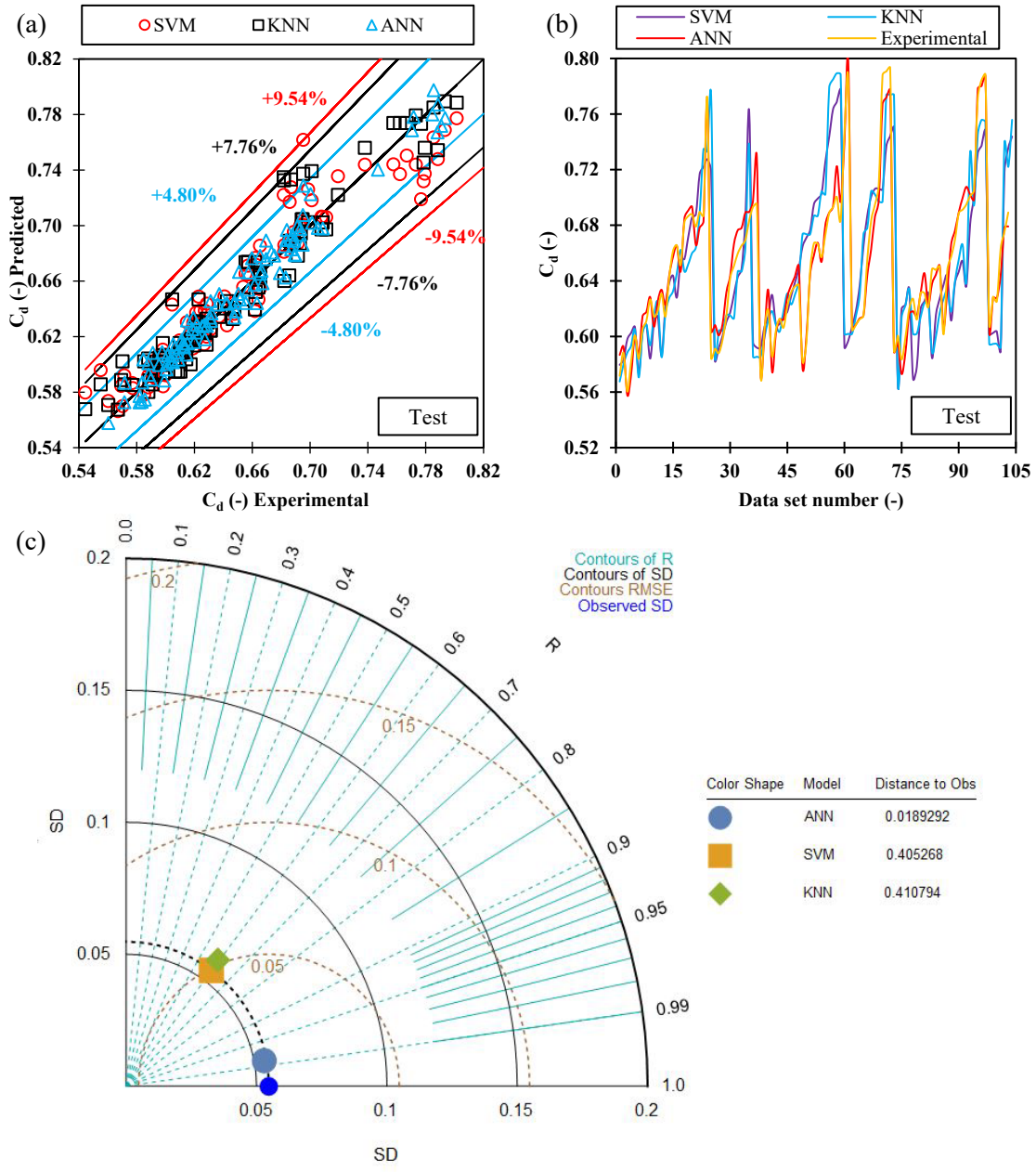


Figure 7. Predicted against experimental C_d in various distance measure criteria a) Euclidean b) Euclidean Squared c) Manhattan d) Chebychev

In order to choose the best model among SVM, KNN, and ANN models, the best results of each group are shown in Figure (8). According to Figure (8-a), it can be seen that for the SVM-RBF model, the values are within the percentage relative error range of ±9.54%. The value of RMSE and Mean RE% for this model is 0.0182 and 1.92%, respectively. For the KNN-Manhattan model, the data are within the percentage relative error range of ±7.76%. This model has provided favorable results compared to the SVM-RBF model. So that the amount of RMSE and Mean RE% for the mentioned model is 0.160 and 1.70%, respectively. The results of the ANN-MLP method compared to the previous two models have statistically better results and are close to the experimental results. For the ANN-MLP method, the data are within the percentage relative error range of ±4.80%. The values of the above statistical indices for this model are 0.0098 and 1.15%,

respectively. The value of the correlation coefficient for the above models in the test phase is 0.955, 0.965, and 0.984 respectively. The comparison of the C_d obtained from different models and the experimental results indicates a better overlap of the data in the ANN method with the experimental results (Figure 8-b). In Figure (8-c), the distance to the observation data in ANN method is 0.0189, which indicates the high accuracy of ANN method. While in SVM and KNN models, these values are 0.4053 and 0.4108, respectively (Taylor diagram). In addition in Figure (8-d), ANN method is in good agreement with the experimental chart (Density diagram).



(d)

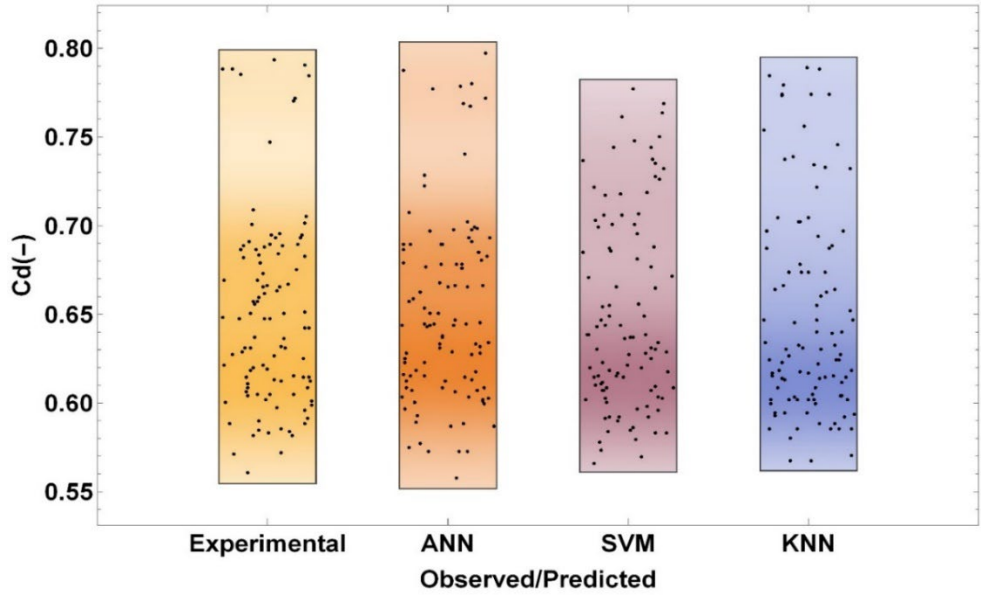


Figure 8. a) Experimental C_d values against predicted b) Comparison of the C_d values for different data in the test phase c) Taylor diagram d) Density diagram

It should be noted that no study has been done in the field of gate-sill using soft computing. The polynomial non-linear regression Equation (18) was presented by Daneshfaraz et al. [18] using Solver in Excel software. Table (6) shows the accuracy of the present research models with Equation (18). The results of the statistical indexes show that Equation (18) has low accuracy compared to the models of this research. One of the reasons for the inappropriate results of Equation (18) can be mentioned in a certain range of data.

$$C_d = 1{,}992\left(\frac{A_{total}}{B^2}\right)^{-0{,}0477} - 6{,}843\left(\frac{H_0}{B}\right)^{-0{,}0062} + 5{,}614\left(\frac{Z}{B}\right)^{0{,}0263} + 0{,}01\left(\frac{X}{B}\right) \quad (18)$$

Table 6. The results of the accuracy of the present models with the non-linear regression Equation (18) of Daneshfaraz et al. [18] study

model	statistical index (-)			
	R	Mean RE	RMSE	KGE
SVM	0.133	0.0865	0.0744	0.133
ANN	0.009	0.0957	0.0823	-0.009
KNN	0.021	0.0945	0.0833	0.018

4. Conclusion

In the present study, modern data mining methods of Support Vector Machine (SVM), K-Nearest Neighbor (KNN) algorithm, and Artificial Neural Network (ANN) were used in predicting the C_d of the slide gate in the sill state. For this purpose, 345 experimental data were used in the

form of 6 different input cases based on dimensional analysis. For all the mentioned models, 70% of the data were randomly used for the training phase and the rest for the test phase. The results of statistical indexes of R, RMSE, Mean RE%, and KGE showed that the case with all input parameters was recognized as the superior case in all SVM, KNN, and ANN models. The results of the examination of different kernels showed that the radial basis function (RBF) kernel has favorable results compared to the other polynomial, linear, and sigmoid kernels compared to the experimental results. In the KNN model, for different neighbor coefficients (K), at K equal to 2, the results of RMSE and Mean RE are the lowest. In addition, for this model, the examination of different distance measure criteria showed that the Manhattan criteria has better results compared to the Euclidean, Euclidean Squared, and Chebychev criteria and is known as the superior criterion in the KNN model. The results of the above statistical indexes for this model are 0.965, 0.0160, 0.0169, and 0.963, respectively. Compared to the previous two models, the ANN method has better results.

References

1. Negm, A.M., Alhamid, A.A., El-Saiad, A.A, (1998). Submerged Flow Below Sluice Gate with s Sill. Proceedings of International Conference on Hydro-Science and Engineering Hydro-Science and Engineering ICHE98, Advances in Hydro-Science and Engineering, Vol.III, Published on CD-Rom and A Booklet of Abstracts, 31 Aug.-3 Sep. 1998, Cottbus/Berlin, G.
2. Henry H.R, (1950). Discussion on Diffusion of Submerged Jets, "by Albertson, M. L. et al., Trans. Am. Society Civil Engrs., 115: 687.
3. Rajaratnam N, Subramanya K, (1967). Flow Equation for the Sluice Gate. Journal of the Irrigation and Drainage Division, 93(3):167-186.
4. Rajaratnam N, (1977). Free Flow Immediately Below Sluice Gates. Journal of the Hydraulics Division, 103(4): 345-351.
5. Swamee P.K, (1992). Sluice Gate Discharge Equations. Journal of Irrigation and Drainage Engineering, 118(1): 56-60.
6. Hager W.H, (1999). Underflow of Standard Sluice Gate. Experiments in Fluids, 27(4): 339-350. Doi: 10.1007/s003480050358
7. Shivapur A.V, Shesha Prakash M.N, (2005). Inclined Sluice Gate for Flow Measurement. ISH Journal of Hydraulic Engineering, 11(1): 46-56.
8. Khalili Shayan H, Farhoudi J, (2013). Effective Parameters for Calculating Discharge Coefficient of Sluice Gates. Flow Measurement and Instrumentation, 33: 96-105.
9. Salmasi F, Abraham J, (2020) Expert System for Determining Discharge Coefficients for Inclined Slide Gates Using Genetic Programming. Journal of Irrigation and Drainage Engineering, 146(12). https://doi.org/10.1061/(ASCE)IR.1943-4774.0001520
10. Salmasi F, Nouri M, Sihag P, Abraham J, (2021). Application of SVM, ANN, GRNN, RF, GP and RT Models For Predicting Discharge Coefficients of Oblique Sluice Gates Using Experimental Data. Water Supply, 21(1): 232-248.
11. Daneshfaraz R, Norouzi R, Abbaszadeh H, (2022). Experimental Investigation of Hydraulic Parameters of Flow in Sluice Gates with Different Openings. Environment and Water Engineering, 8(4): 923-939. doi: 10.22034/jewe.2022.321259.1700
12. Alhamid A.A, (1999). Coefficient of Discharge for Free Flow Sluice Gates. Journal of King Saud University - Engineering Sciences, 11(1): 33-47.

13. Salmasi F, Norouzi Sarkarabad R, (2018). Investigation of Different Geometric Shapes of Sills on Discharge Coefficient of Vertical Sluice Gate. Amirkabir Journal of Civil Engineering, 52(1): 2-2. Doi: 10.22060/ceej.2018.14232.5596
14. Karami S, Heidari M.M, Adib Rad M.H, (2020). Investigation of Free Flow Under the Sluice Gate with the Sill Using Flow-3D Model. Iran J Sci Technol Trans Civ Eng., 44: 317–324.
15. Salmasi F, Abraham J, (2020). Prediction of Discharge Coefficients for Sluice Gates Equipped with Different Geometric Sills under the Gate Using Multiple Non-Linear Regression (MNLR). Journal of Hydrology, 597. https://doi.org/10.1016/j.jhydrol.2020.125728
16. Ghorbani M.A, Salmasi F, Saggi M.K, Bhatia A.S, Kahya E, Norouzi R, (2020). Deep Learning under H_2O Framework: A Novel Approach for Quantitative Analysis of Discharge Coefficient in Sluice Gates. Journal of Hydroinformatics, 22(6): 1603-1619.
17. Daneshfaraz R, Norouzi R, Abbaszadeh H, Kuriqi A, Di Francesco S, (2022). Influence of Sill on the Hydraulic Regime in Sluice Gates: An Experimental and Numerical Analysis. Fluids, 7(7): 244. https://doi.org/10.3390/fluids7070244
18. Daneshfaraz R, Norouzi R, Abbaszadeh H, Azamathulla H.M, (2022). Theoretical and experimental analysis of applicability of sill with different widths on the gate discharge coefficients. Water Supply, 22(10): 7767-7781. doi: https://doi.org/10.2166/ws.2022.354
19. Murzyn F, Chanson H, (2008). Experimental assessment of scale effects affecting two-phase flow properties in hydraulic jumps. Experiments in Fluids, 45: 513-521.
20. Raju, R. (1984). Scale Effects in Analysis of Discharge Characteristics of Weir and Sluice Gates; Kobus: Esslingen am Neckar, Germany.
21. Lauria A, Calomino F, Alfonsi G, D'Ippolito A, (2020). Discharge Coefficients for Sluice Gates Set in Weirs at Different Upstream Wall Inclinations. Water, 12(1): 245.
22. Abbaszadeh H, Daneshfaraz R, Norouzi R, (2023). Experimental Investigation of Hydraulic Jump Parameters in Sill Application Mode with Various Synthesis. Journal of Hydraulic Structures, 9(1): 18-42. doi: 10.22055/jhs.2023.43208.1245
23. Vapnik V.N, The Nature of Statistical Learning Theory. Springer-Verlag, New York, 1995.
24. Norouzi R, Sihag P, Daneshfaraz R, Abraham J, Hasannia V, (2021). Predicting Relative Energy Dissipation for Vertical Drops Equipped with a Horizontal Screen Using Soft Computing Techniques. Water Supply, 21(8): 4493-4513. doi: https://doi.org/10.2166/ws.2021.193
25. Su M.Y, (2011). Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbor classifiers. Expert Systems with Applications, 38(4): 3492-3498.
26. Al-Bulushi N.I, King P.R, Blunt M.J, Kraaijveld M, (2012). Artificial neural networks workflow and its application in the petroleum industry. Neural Computing and Applications, 21: 409-421.
27. Gupta H.V, Kling H, Yilmaz K.K, Martinez G.F, (2009). Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. J. Hydrol., 377(1-2): 80-91. https://doi.org/10.1016/j.jhydrol.2009.08.003.