Mathematical Biosciences and Engineering

MBE, 22 (12): 3236–3261.
DOI: 10.3934/mbe.2025119
Received: 24 May 2025
Revised: 06 September 2025
Accepted: 10 October 2025
Published: 11 November 2025

https://www.aimspress.com/journal/MBE

Research article

Artificial Intelligence for Hydraulic Engineering: Predicting discharge coefficients in trapezoidal side weirs

Mehdi Fuladipanah[1], Saleema Panda[2], Namal Rathnayake[3], Upaka Rathnayake[2,*], Hazi Md. Azamathulla[4] and Yukinobu Hoshino[5]

[1] Department of Civil Engineering, Ramh. C., Islamic Azad University, Ramhormoz, Iran.
[2] Department of Civil Engineering and Construction, Faculty of Engineering and Design, Atlantic Technological University, F91 YW50, Sligo, Ireland.
[3] Marine-Earth System Analytics Unit, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama 236-0001, Japan.
[4] Department of Civil and Environmental Engineering, Faculty of Engineering, The University of the West Indies, St. Augustine 331310, Trinidad and Tobago.
[5] School of Systems Engineering, Kochi University of Technology, 185 Miyanokuchi, Tosayamada, Kami, 782-8502, Kochi, Japan.

* **Correspondence:** Email: upaka.rathnayake@atu.ie; Tel: +353- 719155222.

Abstract: Accurately predicting the discharge coefficient (C_d) is fundamental to the hydraulic design and performance of side weirs. In this study, we introduced a novel artificial intelligence (AI) framework to enhance the prediction accuracy of C_d for two-cycle trapezoidal labyrinth side weirs. Using a comprehensive laboratory dataset, three distinct machine learning models (MLMs), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Gene Expression Programming (GEP), were developed and rigorously compared with application of the Γ-test technique for sensitivity analysis, systematically identifying the five most influential geometric and hydraulic parameters (Fr, $\frac{L}{B}$, $\frac{L_e}{L}$, $\frac{Y_1-P}{P}$, α) to serve as model inputs. The model's efficacy was evaluated across training, testing, and validation phases using multiple statistical metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and the Maximum Developed Discrepancy Ratio ($C_{d(DDRmax)}$). The results demonstrated that the three MLMs are effective predictive tools. However,

the ANN model, specifically an MLP5-7-1 architecture utilizing Atan and Identity activation functions optimized with the BFGS 385 algorithm, significantly outperformed the others. It achieved superior results (e.g., validation phase: RMSE = 0.0061, MAE = 0.0003, R^2 = 0.9301, $C_{d(DDRmax)}$ = 5.22), confirming its highest predictive accuracy and robustness. This research conclusively shows that MLMs, particularly ANN, offer a highly precise and efficient method for predicting Cd in complex hydraulic structures.

Keywords: Artificial intelligence; discharge coefficient; lateral weir; non-linear weir; water management

1. Introduction

Lateral weirs, commonly referred to as side weirs, are immobile hydraulic structures strategically positioned alongside channels to divert a portion of flow away from the main open channel [1]. The inclusion of side weirs arises from the imperative to discharge surplus flow from channels while minimizing significant alterations in water levels [2–4]. They find extensive utilization in the diversion of surplus water during flood event regulation of water levels in canal system integration into water and wastewater treatment facilities' integration with sewer networks and incorporation into the delivery and distribution networks of irrigation systems [5]. In a conventional side weir, lateral discharge is triggered once the water height within the channel attains the elevation of the weir crest [4]. Under subcritical flow conditions, the water depth progressively rises along the channel until it surpasses the initial water height at the downstream end of the weir [6].

The performance of side weirs exerts a direct influence on the environmental quality of their vicinity as they fulfill the dual objectives of water diversion and flow rate measurement [7]. Side weirs are characterized by diverse geometrical configurations and are designed in various sizes to suit the distinct cross-sectional profiles of open channels [8]. The discharge through the side weir once initiated can be ascertained by incorporating an appropriate discharge coefficient C_d into the head-discharge relationship [9]. Each type of side weir exhibits unique hydraulic characteristics underscoring the intricacies associated with their design and selection [1,10,11]. The hydraulic design of side weirs typically encompasses two major concerns: First, the estimation of the appropriate length of the weir for a given flow discharge and specific weir height, and second, the determination of the flow discharge corresponding to a given side weir length and height. In the design of the latter type of problems, C_d assumes a crucial role in determining the capacity of lateral overflows. Consequently, extensive research has been conducted in this domain to assess C_d of such weirs. These investigations encompass linear and non-linear weirs that have resulted in the formulation of relationships for C_d within each of these two groups. Table 1 and Table 2 compile a selection of equations proposed for estimating and calculating C_d for linear and non-linear overflows, respectively.

Table 1. A brief list of equations for C_d in linear side weirs.

References	Equation	Comments
Nandesamoorthy and Thomson [12]	$C_d=0.432\sqrt{\frac{2-Fr_1^2}{1+2Fr_1^2}}$	• Canal: Rectangular • Side weir: Rectangular • $Fr_1\leq 0.60$
Subramanya and Awasthy [13]	$C_d=0.864\sqrt{\frac{1-Fr_1^2}{2+Fr_1^2}}$	• Canal: Rectangular • Side weir: Rectangular • $0.02\leq Fr_1\leq 0.85$
Yu-Tech [14]	$C_d=0.623-0.222Fr_1$	• Canal: Rectangular • Side weir: Rectangular
Kumar and Pathak [15]	$C_d=0.642-0.042Fr_1$	• Canal: Rectangular • Side weir: Triangle
Raju et al. [16]	$C_d=0.81-0.6Fr_1$	• Canal: Rectangular • Side weir: Rectangular
Hager [17]	$C_d=0.485\sqrt{\frac{2+Fr_1^2}{2+3Fr_1^2}}$	• Canal: Rectangular • Side weir: Rectangular • $0.0\leq Fr_1\leq 0.87$
Cheong [18]	$C_d=0.45-0.221Fr_1^2$	• Canal: Trapezoidal • Side weir: Rectangular • $0.28\leq Fr_1\leq 0.78$
Singh et al. [19]	$C_d=0.33-0.18Fr_1+0.49[\frac{W}{y_1}]$	• Canal: Rectangular • Side weir: Rectangular • $0.23\leq Fr_1\leq 0.43$

Continued on next page

References	Equation	Comments
Swamee et al. [20]	$C_d=1.06[(\frac{14.14W}{8.15W+y_1})^{10}+(\frac{y_1}{y_1+W})^{15}]^{-0.1}$	• Canal: Rectangular • Side weir: Rectangular
Jalili and Borghei [21]	$C_d=0.71-0.41Fr_1-0.22(\frac{W}{y_1})$	• Canal: Trapezoidal • Side weir: Rectangular • $0.19\leq Fr_1\leq 0.80$
Borghei et al. [22]	$C_d=0.7-0.48Fr_1-0.3(\frac{W}{y_1})+0.06(\frac{L}{B})$	• Canal: Rectangular • Side weir: Rectangular • $0.10\leq Fr_1\leq 0.90$
Ghodsian [23]	$C_d=0.5973-0.1834Fr_1$	• Canal: Rectangular • Side weir: Rectangular
Durga et al. [24]	$C_d=0.9236-0.3274Fr_1-0.0521Fr_1^2$	• Canal: Rectangular • Side weir: Rectangular • $1.5\leq Fr_1\leq 3.0$
Bilhan et al. [25,26]	$C_d=[0.662+(0.307+0.119(\frac{W}{y_1})^{7.135}+0.241(\frac{L}{B})^{0.413}+0.004(\frac{L}{y_1})^{0.827}+0.21Fr_1^{2.338})^{6.164}]^{2.472}$	• Canal: Rectangular • Side weir: Rectangular
Emiroglu et al. [27,28]	$C_d=[0.836+(-0.035+0.39(\frac{W}{y_1})^{12.69}+0.158(\frac{L}{B})^{0.59}+0.049(\frac{L}{y_1})^{0.42}+0.244Fr_1^{2.125})^{3.018}]^{5.36}$	• Canal: Rectangular • $0.08\leq Fr_1\leq 0.92$
Bagheri et al. [10,11]	$C_d=0.468Fr_1^{-0.045}[\frac{y_1}{B}]^{-0.136}[\frac{y_1}{W}]^{0.039}[\frac{L}{B}]^{-0.116}$	• Canal: Rectangular • $0.08\leq Fr_1\leq 0.91$
Keshavarzi and Ball [29]	$C_d=0.7-0.452Fr_1-0.157\left(\frac{W}{y_1}\right)+0.045Z$	• Canal: Trapezoidal • $0.08\leq Fr_1\leq 0.80$

Table 2. A brief list of equations for C_d in non-linear side weirs.

References	Equation	Comments
Ura et al. [30]	$C_d=0.611\left[\cos\theta\sqrt{\frac{3Fr_1^2}{2+Fr_1^2}}+\sin\theta\sqrt{1-\frac{3Fr_1^2}{2+Fr_1^2}}\right]\sin\theta$	• Type: oblique
Emiroglu et al. [31]	$C_d=[18.6-23.535(\frac{L}{B})^{0.012}+6.769(\frac{L}{l})^{0.112}-0.502(\frac{W}{y_1})^{4.024}+0.094\sin\theta-0.393Fr_1^{2.155}]^{-1.431}$	• Type: Labyrinth triangle • One cycle
Emiroglu and Kaya [28]	$C_d=[1.063(\frac{L}{B})^{0.002}-0.069(\frac{L}{L_{ef}})^{0.074}+0.0045(\frac{W}{y_1})^{5.209}-5.196(\sin\alpha)^{38.754}+0.003Fr_1^{2.582}]^{222.104}$	• Type: Labyrinth trapezoidal • One cycle
Borghei et al. [6]	$C_d=\frac{-0.269[\frac{L}{B\sin\frac{\delta}{2}}]^{-1\text{-}188}+[\frac{W}{y_1\text{-}W}]^{0.18}}{1+0.649[\frac{1}{\sin\frac{\delta}{2}}]^{-0.505}+0.056[\frac{Fr_1}{\sin\frac{\delta}{2}}]^{-1.275}}$	• Type: Labyrinth triangle • Two cycles
Borghei and Parvaneh [32]	$C_d=[-0.18(\frac{Fr_1}{\sin\theta'})^{0.71}-0.15(Fr_1)^{0.44})+(\frac{W}{y_1})^{0.7}]\times[-2.37+2.58(\frac{W\sin\theta'}{y_1})^{-0.36}]$	• Modified oblique
Emiroglu et al. [1]	$C_d=[-0.001Fr_1^{-1.78}+0.10(\frac{L}{B})^{0.22}-2.036(\frac{y_1\text{-}W}{W})^{0.03}+2.82(\frac{L_{ef}}{L\sin\alpha})^{0.02}]^{5.77}$	• Type: Labyrinth trapezoidal • One and two cycles
Karimi et al. [33]	$C_d=0.012(\frac{y_1}{W})^{-2.5}+0.881r_\theta^{0.248}+2.97Fr_2^{1.79}$	• Type: Asymmetric triangular • One cycle
Hussein and Jalil [34]	$C_d=0.393[\frac{\frac{l}{L}^{-0.857}}{(\frac{y_1}{W})^{0.635}Fr_1^{0.321}}]$	• Type: Labyrinth triangle • Two cycles

Over the past forty years, extensive investigations have been undertaken to ascertain precise correlations for C_d in various hydraulic structures through laboratory experimentation and the development of empirical relations or diagrams. While laboratory experiments offer valuable insights into flow conditions in prototypes of hydraulic structures, they represent costly and time-intensive approaches in this domain [35]. In the last two decades, subsequent to the advent of computing technologies, research communities have endeavored to explore alternative methods for simulating flow in open channels. Among these alternatives, MLMs have gained prominence incorporating

diverse models [36–41]. Literature extensively documents the efficacy of such computational models in tackling nonlinear and intricate problems [42]. Consequently, numerous hydraulic research groups have adopted these computational models to derive accurate estimations of C_d in hydraulic structures (refer to Table 3).

Table 3. A brief literature of MLMs application to predict C_d.

Reference	Included MLMs	Weir type
Bilhan et al. [25]	FFNN, RBNN	Rectangular side weirs
Bilhan et al. [26]	ANN	Triangular labyrinth side weir in curved channels
Dursun et al. [43]	ANFIS, NLR, MLR	Semi-elliptical side weir
Bagheri et al. [10,11]	ANN	Rectangular sharp crested side weirs
Zaji et al. [44]	MLPNN-PSO, RBNN-PSO	Triangular labyrinth weirs
Parsaie and Haghiabi [45]	MLP	Side weir
Ebtehaj et al. [46]	GEP	Rectangular side weir
Eghbalzadeh et al. [47]	FFBP, RBF, GRNN, MNLR	Circular and Rectangular Side Orifices
Khoshbin et al. [48]	ANFIS-GA, SVD, MLP	Rectangular sharp-crested side weirs
Azamathulla et al. [49]	SVM, ANFIS, ANN	Standard rectangular side weir
Yasi and Gholami [50]	ANN	Labyrinth Fuse gate
Karami et al. [51]	SVR-FA, SVR, PCA, RSM	Triangular labyrinth weir
Shabanlou [52]	SaE-ELM	Sharp-Crested Weirs on the end of Circular
Azimi et al. [53]	ANFIS-PSO- GA, ANFIS, ANFIS-GA, ANFIS-PSO	Side orifices
Ezzeldin et al. [54]	NARX-NN, GMDH, FFBP, RBF, GRNN	Sharp crested circular and square orifices
Parsaie et al. [55]	GMDH, MARS, SFF, GP	Non-linear Weirs with Triangular Plan
Majedi Asl and Fuladipanah [56]	SVM, LR, NLR	Sharp-crested triangular labyrinth weirs
Parsaie and Haghiabi [57]	GEP, GMDH, MARS	Weir-gate
Mohammed and Sharifi [58]	GEP, MLR	Oblique side weir
Hussain et al. [59]	GEP, GMDH	Rectangular side weir
Hu et al. [60]	ANFIS, SVM, M5	Curved labyrinth overflows
Mustafa et al. [61]	SVM	Triangular labyrinth weir

Continued on next page

Reference	Included MLMs	Weir type
Fuladipanah and Majedi-Asl [62]	GEP, SVM	Side Rectangular Weirs
Majedi-Asl et al. [63]	SVM, GEP	Rectangular labyrinth weir and Piano-key weir
Li et al. [64]	SVM with three optimization algorithms	Semi-circular side weir
Tian et al. [65]	MLR, SVM, GEP, ANN, RF, AdaBoost, XGBoost, CatBoost	Type-A piano key weirs
Hamed and Elshaarawy [37]	ANN, GEP	Piano-key weir

Multiple-Linear-Regression (MLR); Random-Forest (RF); Adaptive-Boosting (Adaboost); Extreme-Gradient-Boosting (Xgboost); Categorical-Boosting (Catboost); Feed Forward Neural Network (FFNN); Radial Basis Neural Network(RBNN); Adaptive-Neuro Fuzzy Inference System (ANFIS); Multiple Linear Regression (MLR); Nonlinear Regression (NLR); Multi-Layer Perceptron Neural Network (MLPNN); Radial Basis Neural Network (RBNN); Particle Swarm Optimization (PSO); Multilayer Perceptron (MLP); Feed-Forward Back-Propagation (FFBP); Radial Basis Function (RBF); Generalized Regression Neural Network (GRNN); Multiple Non Linear Regression (MNLR); Singular Value Decomposition (SVD); Support Vector Regression (SVR), Support Vector Regression–Firefly (SVR- Firefly); Response Surface Methodology (RSM); Principal Component Analysis (PCA); Self-Adaptive Optimization Of Extreme Learning Algorithm Using The Differential Evolutionary Algorithm (Sade-ELM); Nonlinear Neural Network Technique (NARX); Group Of Method Data Handling (GMDH), Genetic Programming (GP); Multivariate Adaptive Regression Splines (MARS); Smart Function Fitting (SFF);

The literature underscores the critical importance of optimizing the performance of non-linear weirs and highlights the utilization of machine learning methodologies to predict the discharge coefficient. However, a significant research gap remains in the application of a robust, data-driven sensitivity analysis to objectively determine the optimal input parameters for predicting C_d in two-cycle trapezoidal labyrinth side weirs. Therefore, the novelty of this study is twofold: First, it presents a comprehensive comparative analysis of three distinct MLMs (SVM, ANN, and GEP) for this specific and hydraulically complex structure. Second, and most significantly, it introduces the novel application of a Γ-test sensitivity analysis to this domain, which systematically and objectively identifies the most critical influencing parameters, thereby optimizing the model inputs and enhancing prediction reliability. This methodology directly addresses a crucial gap in hydraulic engineering, offering a novel framework for developing precise and efficient tools for the design and management of water infrastructure, thereby contributing to sustainable water resource management.

2. Materials and methods

2.1. Data series analysis

Our primary objective was to employ three MLMs namely ANN, SVM, and GEP to simulate and predict the C_d in a labyrinth trapezoidal lateral weir with two cycles under varying hydraulic and geometric conditions. We utilized the laboratory data collected by Emiroglu et al. [1]. Figure 1 provides

a schematic view of the lateral weir, detailing its geometry and the laboratory equipment setup. The collected data is summarized in Table A1.

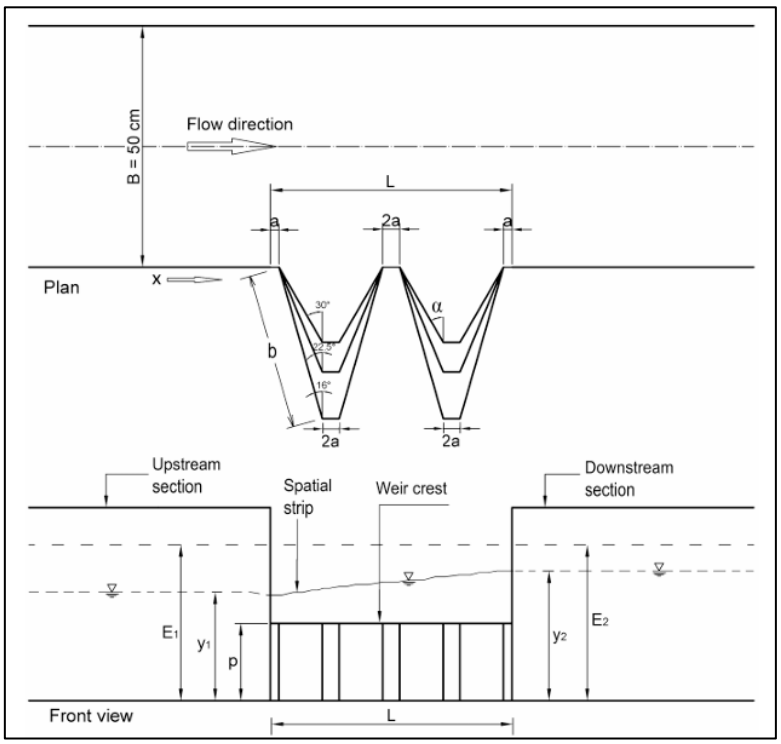


Figure 1. Schematic view of plan and front view of the labyrinth side weir [1].

2.2. The basics of MLMs included

The laboratory data compiled for this study were partitioned into training, testing, and validation subsets for each of the three MLMs employed. All three MLMs were developed in MALAB Simulink model. Based on the results of the sensitivity analysis, the influential input parameters were selected. Subsequently, through a trial-and-error approach, the optimal tuning parameter values for each MLM were determined. For further elaboration, the following section introduces the fundamental concepts of each of the three MLMs utilized in the present research.

2.2.1. An overview of SVM

SVM is a powerful supervised MLM used for classification and regression tasks. Given a training dataset $\{(\mathbf{X}_i\ y_i)\}_{i=1}^{N}$ where $\mathbf{X}_i \in R^d$ are feature vectors and $y_i \in \{-1\ 1\}$ are class labels the goal of SVM is to find an optimal hyperplane defined by Equation 1.

$$\mathbf{W}.\mathbf{X}+b=0, \tag{1}$$

where W is the normal vector to the hyperplane and b is the bias term. The optimization problem for a hard-margin SVM (linearly separable case) is shown in Equation 2.

$$\min_{W\,b}\frac{1}{2}\|\mathbf{W}\|^2, \tag{2}$$

subject to

$$y_i(\mathbf{W}.\mathbf{X}_i+b)\geq 1, \tag{3}$$

For non-linearly separable data, a soft-margin SVM introduces slack variables $\xi_i \geq 0$ to enable some misclassification.

$$\min_{W\,b\,\xi}\frac{1}{2}\|\mathbf{W}\|^2+C\sum_{i=1}^{N}\xi_i, \tag{4}$$

subject to

$$y_i(\mathbf{W}.\mathbf{X}_i+b)\geq 1-\xi_i, \tag{5}$$

$$\xi_i \geq 0, \tag{6}$$

where C is a regularization parameter balancing the trade-off between maximizing the margin and minimizing the classification error. For data that is not linearly separable in the original feature space, SVM can be extended using the kernel trick. A kernel function $\boldsymbol{K}(\mathbf{X}_i\ \mathbf{X}_j)$ implicitly maps the data into a higher-dimensional space where a linear separation is possible. More information on these kernel functions is given in Table A2.

The optimization problem in the dual form with kernel becomes as shown in Equation 7.

$$\mathrm{Max}_{\alpha}\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i\,j=1}^{N}\alpha_i\alpha_j y_i y_j \boldsymbol{K}(\mathbf{X}_i\ \mathbf{X}_j), \tag{7}$$

subject to

$$0\leq\alpha_i\leq C, \tag{8}$$

$$\sum_{i=1}^{N}\alpha_i y_i=0, \tag{9}$$

where α_i are the Lagrange multipliers.

When implementing SVM, the initial phase involves precisely characterizing the nature of the problem, whether it pertains to classification, regression, or anomaly identification. Subsequently, relevant data must be gathered and preprocessed to address inconsistencies, scale features appropriately, and ensure overall quality for modeling. Depending on the data characteristics, a suitable kernel function is chosen to map the input into a higher-dimensional space if linear separation is not feasible. The prepared dataset is then partitioned into training and testing subsets, often following a conventional split for evaluation. Critical hyperparameters must also be configured, including the

regularization constant C, which balances margin width against misclassification penalties, and kernel-specific settings like the bandwidth γ in radial basis functions or the degree in polynomial kernels. Using the training subset, the SVM optimization process proceeds by identifying the optimal separating hyperplane that maximizes the margin between classes for classification or minimizes deviation within an epsilon-tube for regression. Model validation is performed via cross-validation to gauge robustness and direct further refinement. Hyperparameter optimization is carried out using systematic approaches such as grid or random search. Ultimately, the model's performance is assessed on the held-out test set using relevant metrics; for instance, classification accuracy, F1-score, and confusion matrices for categorical outcomes, or root mean squared error and mean absolute error for continuous targets. This final assessment indicates the model's ability to generalize and its practical utility.

2.2.2. An overview of ANN

ANNs are a fundamental component of modern machine learning inspired by the structure and function of biological neural networks in the human brain. They are composed of layers of interconnected nodes or neurons which process input data to produce an output. Each neuron in an ANN performs a weighted sum of its inputs adds a bias term and passes the result through an activation function. Mathematically, for a neuron ***j*** in layer ***l*,** this can be showcased in Equation 10.

$$z_j^l = \sum_{i=1}^{n} \omega_{ij}^l a_i^{l-1} + b_j^l, \tag{10}$$

where z_j^l is the weighted input to neuron j in layer *l*. ω_{ij}^l is the weight connecting neuron i in layer *l*-1 to neuron j in layer *l*, a_i^{l-1} is the activation of neuron i in layer *l*-1, and b_j^l is the bias term for neuron j in layer *l*. The activation a_i^{l-1} of neuron j in layer *l* is then computed using an activation function σ, which is given in Equation 11.

$$a_i^{l-1}=\sigma(z_j^l)\,, \tag{11}$$

Sigmoid Tanh and ReLU are common activation functions. The architecture of an ANN defines the arrangement of neurons in layers and the connectivity between them. Common architectures include Feed forward Neural Networks, Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs).

The development of the ANN model follows a structured, multi-phase procedure aimed at optimizing architectural and learning components. First, the number of hidden layers is evaluated to balance model capacity and complexity. Next, the quantity of neurons within each hidden layer is determined to sufficiently capture patterns without overfitting. Subsequently, activation functions, such as sigmoid, hyperbolic tangent, or linear types, are selected to introduce non-linearity and shape model behavior. Finally, a training algorithm is chosen to effectively minimize error during the weight optimization process. A common strategy for configuring the MLP begins with a single hidden layer containing neurons equal to the number of input features. Multiple activation functions are empirically tested to identify the most suitable one. After establishing a performance baseline, the architecture is

iteratively refined by incrementally increasing the number of hidden layers and neurons within them. This gradual expansion enhances the model's ability to represent complex relationships in the data, improving predictive accuracy through systematic experimentation and validation.

2.2.3. An overview of GEP

The GEP developed by Ferreira in 1999 [66] is an evolutionary algorithm that extends genetic algorithms (GA) and genetic programming (GP). It uses a fixed-length linear genome to encode variable-length tree structures (expression trees, ETs), representing solutions to a problem. This separation of genotype (chromosome) and phenotype (ET) enables GEP to efficiently explore the solution space and evolve complex structures. A gene consists of a head that contains functions and terminals, and a tail that contains only terminals. The length of the head (h) determines the length of the tail (t), as given in Equation 12:

$$t = h(n-1) + 1, \tag{12}$$

where n is the maximum arity of the functions used. GEP uses several genetic operators to evolve the population like selection, recombination, mutation, inversion, and transposition.

To initiate a modeling process using GEP, one must first establish a precise problem definition, specifying whether the task involves regression, classification, or symbolic representation. Subsequently, an appropriate dataset is assembled and refined through preprocessing steps that address missing values and adjust feature scales as needed to ensure data integrity and model readiness. The dataset is then divided into training and testing portions, with a common practice being to reserve 70–80% of the data for training and the rest for performance evaluation. Next, the GEP architecture is configured by selecting structural parameters such as the quantity of genes, the length of the head section, a set of functions (e.g., arithmetic operators), and terminal elements (variables and constants). The algorithm initializes a population of candidate solutions, which undergoes iterative improvement via genetic operations, including mutation, recombination, and transposition, across multiple generations. Fitness is assessed using problem-specific criteria, such as the mean squared error for regression or accuracy for classification. Through successive iterations, the population evolves toward more fit individuals. Upon completion of the evolutionary process, the best-performing symbolic expression is selected and evaluated on the testing subset using relevant metrics. The final model may then be interpreted or simplified for practical deployment or further analysis.

2.2.4. Evaluation metrics

Several statistical metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and the Developed Discrepancy Ratio (DDR), are employed to evaluate the performance of predictive models. Each metric provides distinct insights into the accuracy and reliability of the model predictions. More information about these indices can be found in Fuladipanah et al. [67].

2.2.5. Sensitivity analysis

According to Koncar [68], the Gamma test is a non-parametric statistical technique employed to

estimate outputs by identifying optimal input-output dataset configurations that yield minimal mean square error values. This method is recommended for determining the most effective combination of various input variables to accurately characterize the output. In this approach, the dataset is represented as $\{(x_i\ y_i)\ 1\leq i\leq M)\}$, where the input vectors $x_i\in R^m$ are m-dimensional vectors, and the corresponding outputs $y_i\in R$ are scalars. The input vectors x influences the output y. The relationship between the input and output variables is defined by Equation 13:

$$y=Gx+\Gamma, \tag{13}$$

where G and Γ denote the gradient and the intercept of the regression line where x=0, respectively, and y is the output. Lower values of G and Γ suggest that the respective input variables are more appropriate. In addition to these two criteria, an indicator denoted as V-Ratio=$\frac{\Gamma}{\sigma^2(y)}$ where Γ represents the gamma function and $\sigma^2(y)$ is the output variance is utilized to identify the optimal input parameters. The values of the V-Ratio range from 0 to 1. A V-Ratio value closer to zero for each input parameter indicates the effectiveness of that particular input. The combinations of input variables have been delineated following the format introduced by Mask [69]. The relationship between the independent variables and the dependent variable can be expressed in Equation 14:

$$C_d=F(Fr, \frac{L}{B}, \frac{L_e}{L}, \frac{Y_1\text{-}P}{P}, \alpha), \tag{14}$$

To conduct a sensitivity analysis using the gamma test, the presence or absence of each independent variable is denoted by the binary digits 1 and 0, respectively. For instance, when all independent variables are incorporated into the sensitivity analysis, the Mask column in Table 4, which summarizes the sensitivity analysis, displays the combination "11111". Alternatively, the combination "00111" in the seventh row of Table 4 indicates that the sensitivity analysis in this scenario was performed without considering the variables Fr and $\frac{L}{B}$. As evidenced by the results in Table 4, the lowest value of the V-Ratio parameter was observed when all independent variables were included (rows 1 and 2 of Table 4). Additionally, the values of the parameters Γ, G, and the standard error were also obtained for these two rows. The variations pertaining to the V-Ratio-Gradient-standard error variable are depicted in Figure 2a. Figure 2b illustrates the concurrent changes in V-Ratio and Standard Error for each of the tests listed in Table 4. Considering the influence of effective length on the flow rate, all independent variables were incorporated into the calculations related to the machine learning models. Although the Gamma test evaluation indices are identical in the first and second rows, all independent variables were included in the calculations related to the machine learning models due to the impact of effective length, L_e, on the flow rate. Figure 2c and Figure 2d present the distribution of Γ-σ and frequency-Γ-σ for the opted combination of input variables.

The Γ-test sensitivity analysis, Table 4, objectively identified several input combinations that minimized the mean squared error, with the masks "11111" (Fr, $\frac{L}{B}$, $\frac{L_e}{L}$, $\frac{Y_1\text{-}P}{P}$ and α) yielding identical and optimal values for the Gamma (Γ = 0.0330) and V-ratio (0.2072) metrics.

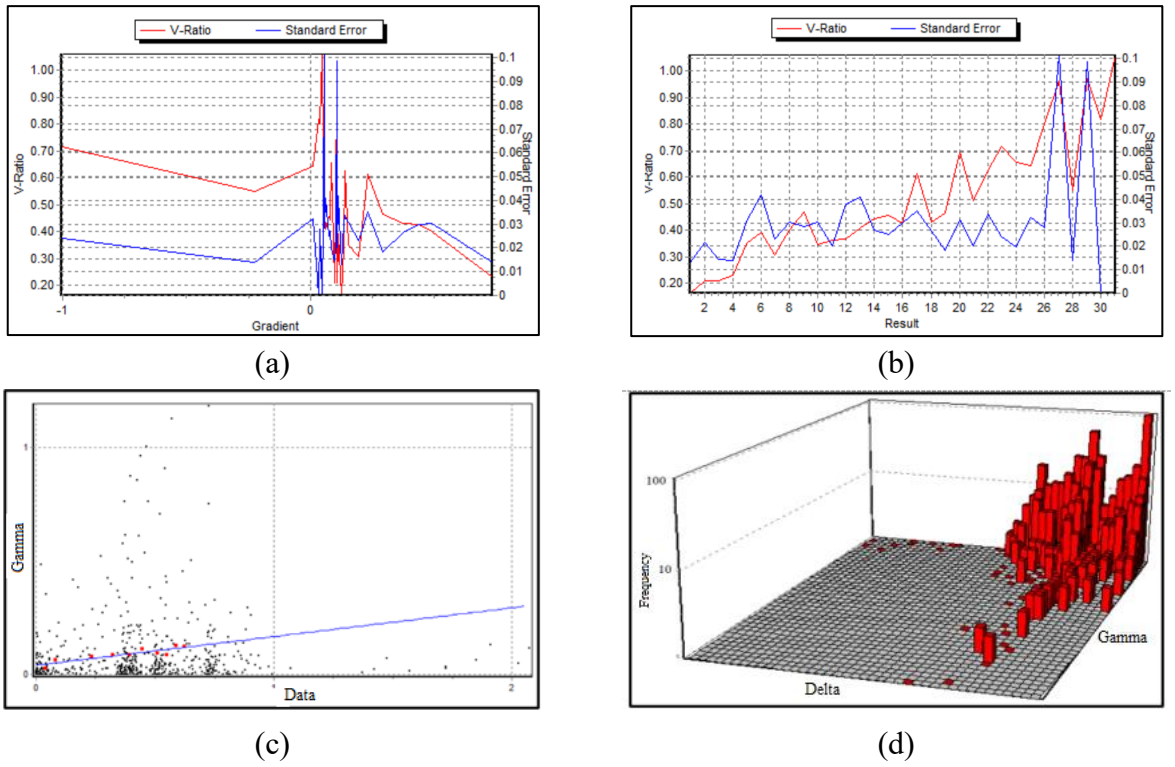


Figure 2. Sensitivity of the data set: (a) Variation of G vs. V-Ratio and standard error during sensitivity analysis; (b) variations of V-Ratio and standard error for different masks; (c) scatter plot of Γ vs. σ for opted mask (11011); and (d) as a 3D histogram of σ vs. Γ for selected mask (11011).

Table 4. Summary of sensitivity analysis using the Γ-test.

Test No.	Gamma	Gradient	Standard error	V-Ratio	Mask	Test No.	Gamma	Gradient	Standard error	V-Ratio	Mask
1	0.0330	0.2421	0.0161	0.2072	11111	17	0.0722	0.1399	0.0131	0.4533	01011
2	0.0392	0.654	0.0172	0.2354	11011	18	0.0766	0.0687	0.0205	0.4806	01110
3	0.0438	0.1414	0.0184	0.2747	11110	19	0.1141	-2.8340	0.0152	0.7161	10000
4	0.0404	1.0267	0.0110	0.2537	10010	20	0.0967	0.7698	0.0229	0.6069	10100
5	0.0579	0.2669	0.0179	0.3631	10111	21	0.0743	0.3952	0.0120	0.4665	01010
6	0.0730	0.3204	0.0263	0.4578	11101	22	0.0974	0.7036	0.0221	0.6115	10001
7	0.0640	0.6447	0.0199	0.4015	00111	23	0.0810	0.3930	0.0128	0.5082	11000
8	0.0722	0.1399	0.0131	0.4533	01111	24	0.1486	0.0003	0.0198	0.9324	01101
9	0.0579	0.2669	0.0179	0.3631	10011	25	0.1043	0.0730	0.0126	0.6542	01100
10	0.0593	0.2606	0.0178	0.3718	10110	26	0.0878	-0.3171	0.0088	0.5510	00010
11	0.0571	0.1865	0.0122	0.3585	11010	27	0.1440	0.0002	0.0222	0.9034	01001
12	0.0730	0.3204	0.0263	0.4578	11001	28	0.2078	0.0000	0.0454	1.3041	00101
13	0.0967	0.7698	0.0229	0.6069	10101	29	0.1550	0.0918	0.0631	0.9726	00100
14	0.0659	0.6989	0.0202	0.4132	00110	30	0.1306	0.0274	0.0000	0.8192	01000
15	0.0640	0.6442	0.0199	0.4017	00011	31	0.1691	0.0003	0.0000	1.0611	00001
16	0.0756	0.2211	0.0282	0.4743	11100						

3. Results and discussion

Table 5 presents a summary of the performance evaluation indicators of the MLMs utilized in this research. These indicators were extracted and calculated in three phases: Training, testing, and validation. The allocation of data for these phases to get the optimum outcome was 70% for training, 15% for testing, and 15% for validation.

Based on the data presented in Table 5, the ANN-MLP 5-7-1 model demonstrated superior performance across all three stages. The most effective neural network, characterized by the MLP 5-7-1 architecture and utilizing Tanh and Identity activation functions for the hidden and output layers, respectively, was achieved through the BFGS 385 optimization algorithm. The performance evaluation metrics (RMSE, MAE, R^2, $C_{d(DDRmax)}$) during the testing phase are (0.0087, 0.0005, 0.9133, 4.00), respectively, during the training phase are (0.0110, 0.0006, 0.8716, 4.14), respectively, and during the validation phase are (0.0061, 0.0003, 0.9301, 5.22), respectively. These metrics indicate that the neural network model achieves the lowest RMSE and MAE values, alongside the highest R^2 and $C_{d(DDRmax)}$ values.

Table 5. Statistical metrics of MLMs performance.

Model name	Training phase			
	RMSE	MAE	R^2	$C_{d(DDR\ max)}$
SVM	0.0129	0.0007	0.8139	3.0515
GEP	0.0124	0.0007	0.8222	2.6724
MLP (5-7-1)	0.0087	0.0005	0.9133	4.0000
Model name	Testing phase			
SVM	0.0145	0.0009	0.7816	2.5416
GEP	0.0139	0.0009	0.7985	2.8032
MLP (5-7-1)	0.0110	0.0006	0.8716	4.1421
Model name	Validation phase			
SVM	0.0121	0.0008	0.7840	2.8202
GEP	0.0113	0.0007	0.7527	2.5713
MLP (5-7-1)	0.0061	0.0003	0.9301	5.2216

There is no significant disparity in the performance between the SVM and GEP models. However, overall, the SVM model exhibits superior output compared to the GEP model. The optimal SVM model, utilizing the RBF kernel function, was determined with adjustment parameters of C=16, ε=0.1, and γ=0.2, and comprised 31 support vectors. The indices (RMSE, MAE, R^2, $C_{d(DDRmax)}$) in the training and testing phases are (0.0129, 0.0007, 0.8139, 305) and (0.0145, 0.0009, 0.7816, 2.54), respectively. During the validation phase, the indices (RMSE, MAE, R^2, $C_{d(DDRmax)}$) are (0.0121, 0.0008, 0.7840, 2.82), respectively.

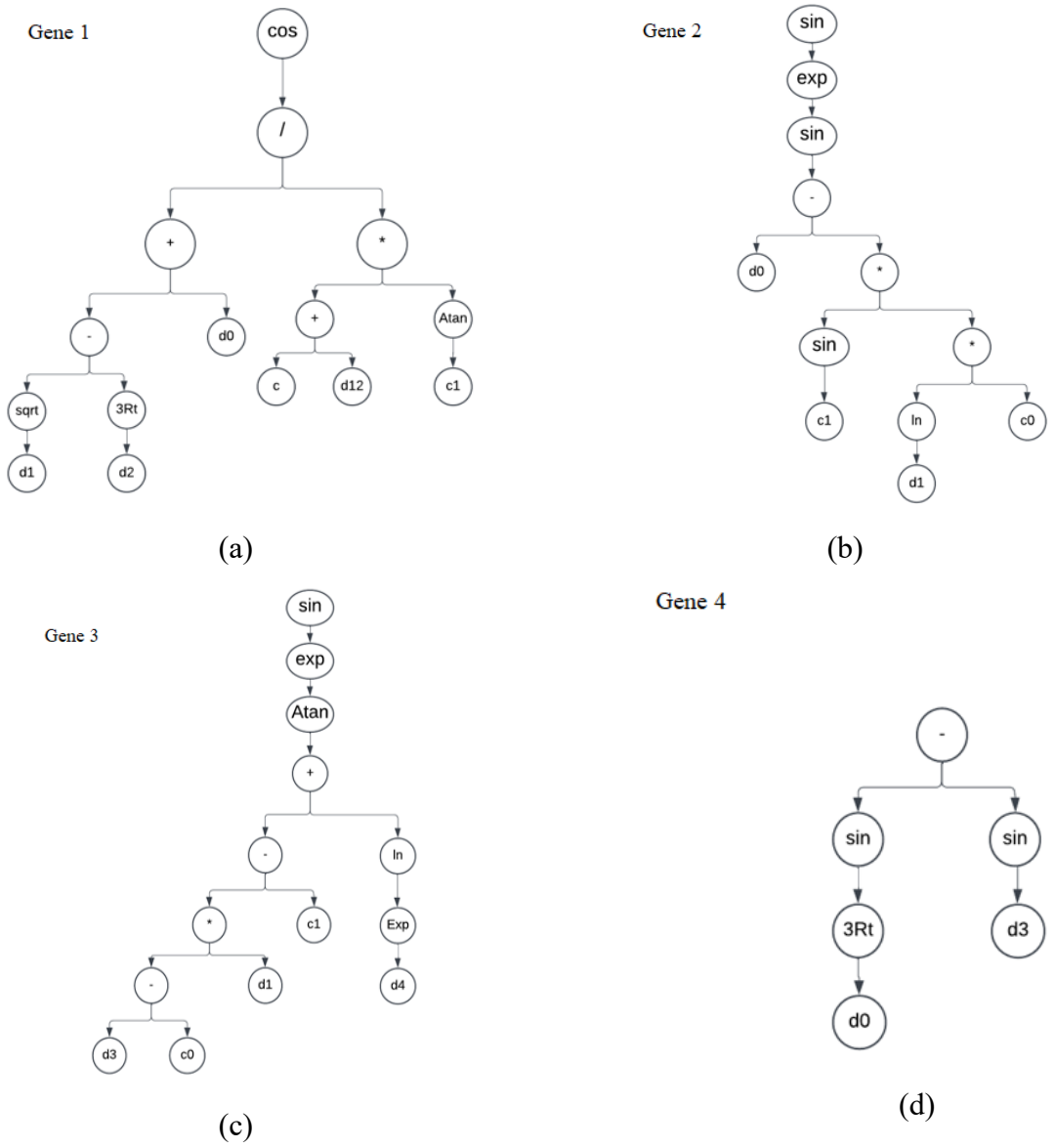


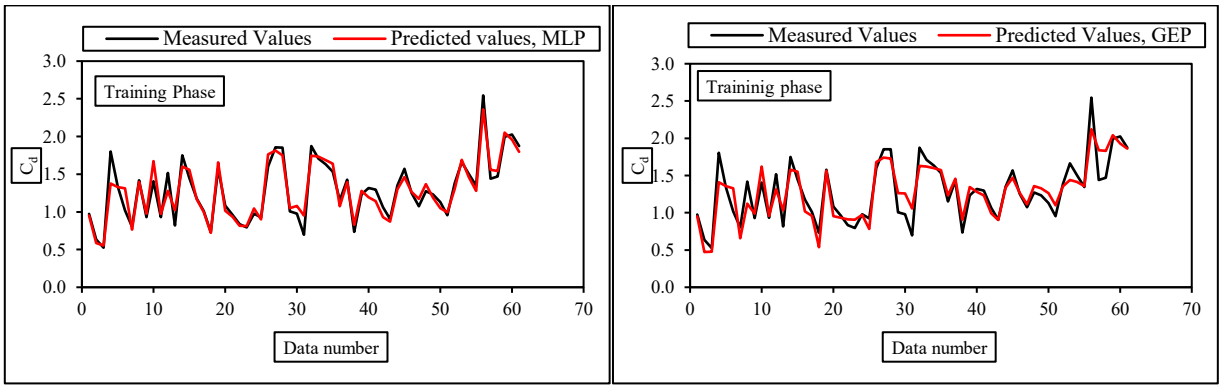
Figure 3. The ET of the GEP outcome.

Table 6 presents the tuning parameter values of the optimal GEP model. Notably, the GEP model with four genes produces the most optimal outcomes across the training, testing, and validation phases. Additionally, the ET of the GEP model's output is illustrated in Figure 3. The constant values in the first gene are c0=5.003967 and c1=-3.52298; in the second gene, c0=-7.677979 and c1=0.898163; in the third gene, c0=5.212982 and c1=5.506561; and in the fourth gene, c0=7.216675 and c1= 9.622864. The calculations indicate that the values of the indices (RMSE, MAE, R^2, $C_{d(DDRmax)}$) during the training, testing, and validation phases are (0.0124, 0.0007, 0.8222, 2.67), (0.0139, 0.0009, 0.7985, 2.8), and (0.0113, 0.0007, 0.7527, 2.57), respectively.

Table 6. Tuning parameters of the GEP model.

Parameters	Values
Head size	10
Chromosomes numbers	45
Number of genes	4
Mutation rate	0.044
Inversion rate	0.1
One-point recombination rate	0.3
Two-point recombination rate	0.3
Gene recombination rate	0.1
Gene transposition rate	0.1
IS transposition rate	0.1
RIS transposition rate	0.1
Fitness function error type	RMSE
Linking function	+

Figure 4 is a critical visual evaluation of the predictive performance and generalization capabilities of the three MLMs- SVM, GEP, and ANN-MLP5-7-1 architecture- across the training, testing, and validation phases. This figure is a qualitative assessment based on the degree of data adaptability, most likely presented as a scatter plot of measured versus predicted values for each model, divided by phase and centered around a 1:1 perfect prediction line. The textual description indicates that the ANN model demonstrates a significantly higher output adaptation rate across all three phases. This translates visually to a tight clustering of its data points around the 1:1 line, demonstrating not only exceptional accuracy but also a superior ability to generalize beyond the training data to new, unseen data. In contrast, the SVM and GEP models, while effective, would exhibit a more dispersed cloud of points, indicating higher prediction errors and slightly weaker generalization, with the SVM's performance likely appearing marginally better than GEP's. The profound significance of Figure 4 lies in its role as a visual validation of the quantitative metrics in Table 5; it provides an intuitive and powerful demonstration of the ANN's robustness and consistency, making it the most reliable and accurate predictive tool for the discharge coefficient in this complex hydraulic application and solidifying our central conclusion.



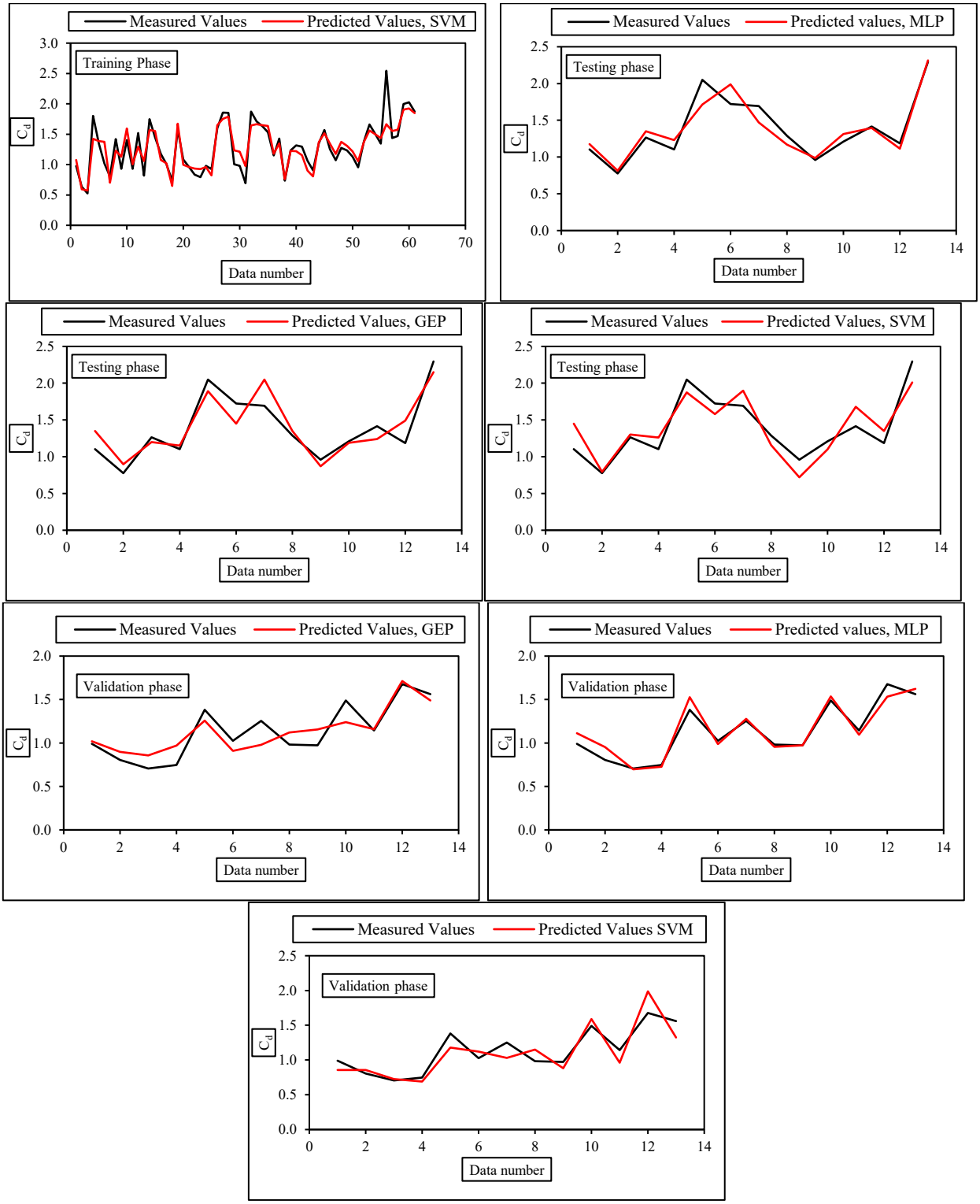


Figure 4. Performance Evaluation of MLMs Across Training, Testing, and Validation Phases.

Figure 5 is a critical visual performance metric for comparing the three MLMs. As mentioned, the DDR index, calculated as the ratio of predicted to measured C_d values, provides a direct measure of prediction accuracy, where an ideal value of 1.0 indicates a perfect prediction. The figure plots the frequency distribution of these DDR values for the SVM, GEP, and MLP models across the training, testing, and validation phases. Superior model performance is visually characterized by a curve that is

narrow, indicating consistent predictions with low error variance, and tall, signifying a high concentration of accurate predictions near the ideal DDR of 1.0. Based on our description, the ANN model (MLP5-7-1) demonstrably outperforms the others, meaning its curve would be the narrowest and have the highest peak, tightly clustered around 1.0, which visually confirms its superior statistical results from Table 5. The SVM and GEP models follow with lesser accuracy, evidenced by their curves being wider and flatter, indicating a greater spread of errors and less prediction consistency. The profound significance of this figure lies in its demonstration of the ANN's robustness and generalization capability, as its optimal DDR distribution remains consistent across all three phases, proving it reliably captures the complex hydraulic relationships without overfitting and is the most trustworthy model for practical engineering applications.

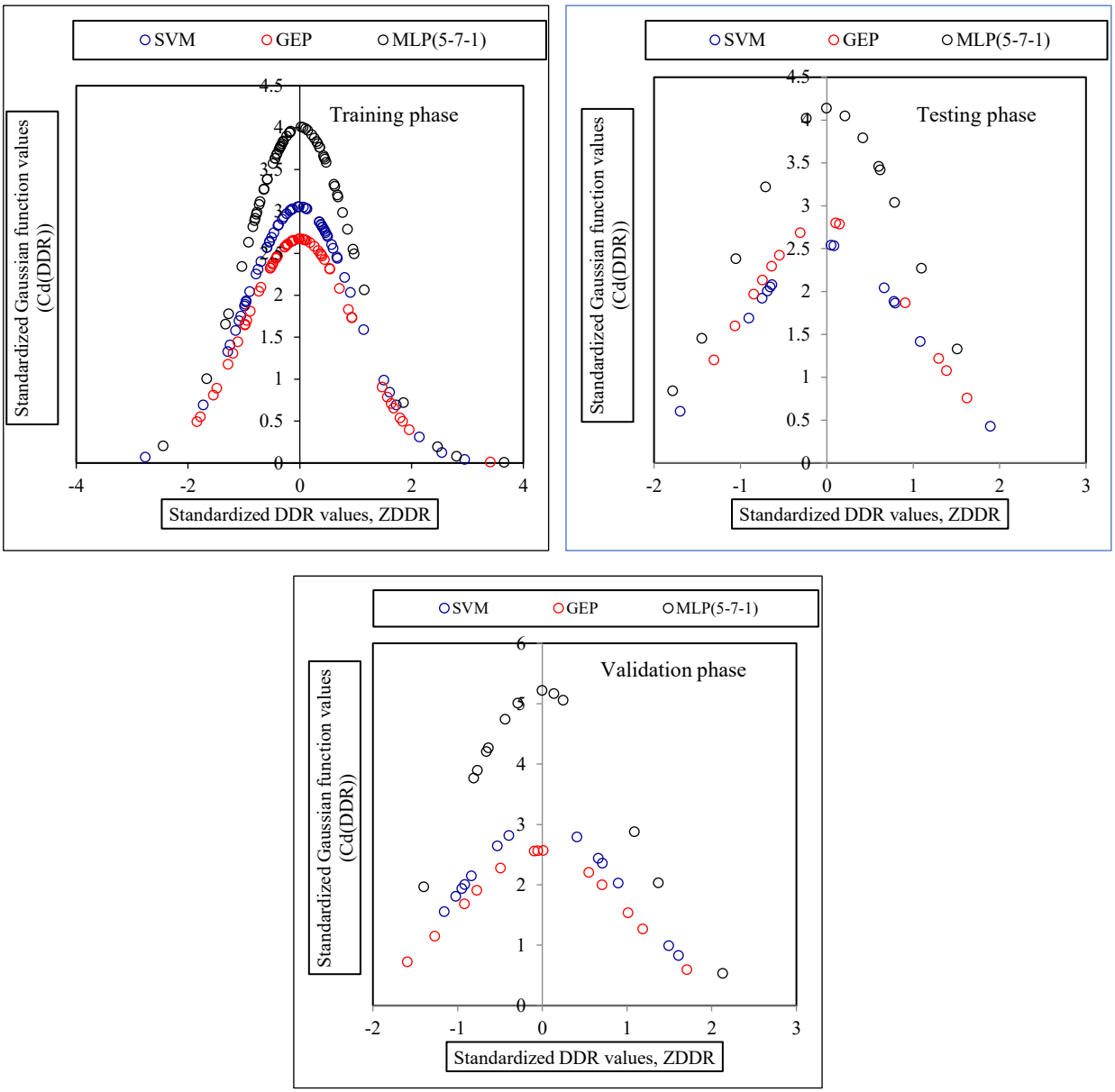


Figure 5. Distribution of DDR index for MLMs involved during the training, testing, and validation phases.

Figure 6 provides a direct and compelling graphical comparison of the predictive accuracy of the three MLMs by plotting their computed C_d values against the experimentally measured values, serving as a visual confirmation of the models' performance that complements the statistical metrics. This scatter plot, built around a central 1:1 line representing perfect prediction, enables an immediate and intuitive assessment of each model's efficacy, where the proximity of the data points to this ideal line directly correlates with the model's accuracy. The text explicitly states that the superiority of the MLP model over the other two is evident from these graphs, indicating that its data points form the tightest and most concentrated cluster around the 1:1 line, a visual manifestation of its lowest RMSE and highest R^2 values. In contrast, the data points for the SVM and GEP models demonstrate greater dispersion and scatter farther from the line, visually quantifying their higher error rates and lower correlation coefficients, as reported in Table 5. The profound significance of Figure 6 is that it transcends numerical metrics to offer a clear, instantaneously understandable illustration of the ANN model's superior capability to capture the complex underlying hydraulic relationships, thereby providing undeniable visual evidence that solidifies its selection as the most precise and reliable predictive tool for the discharge coefficient in two-cycle trapezoidal labyrinth side weirs.

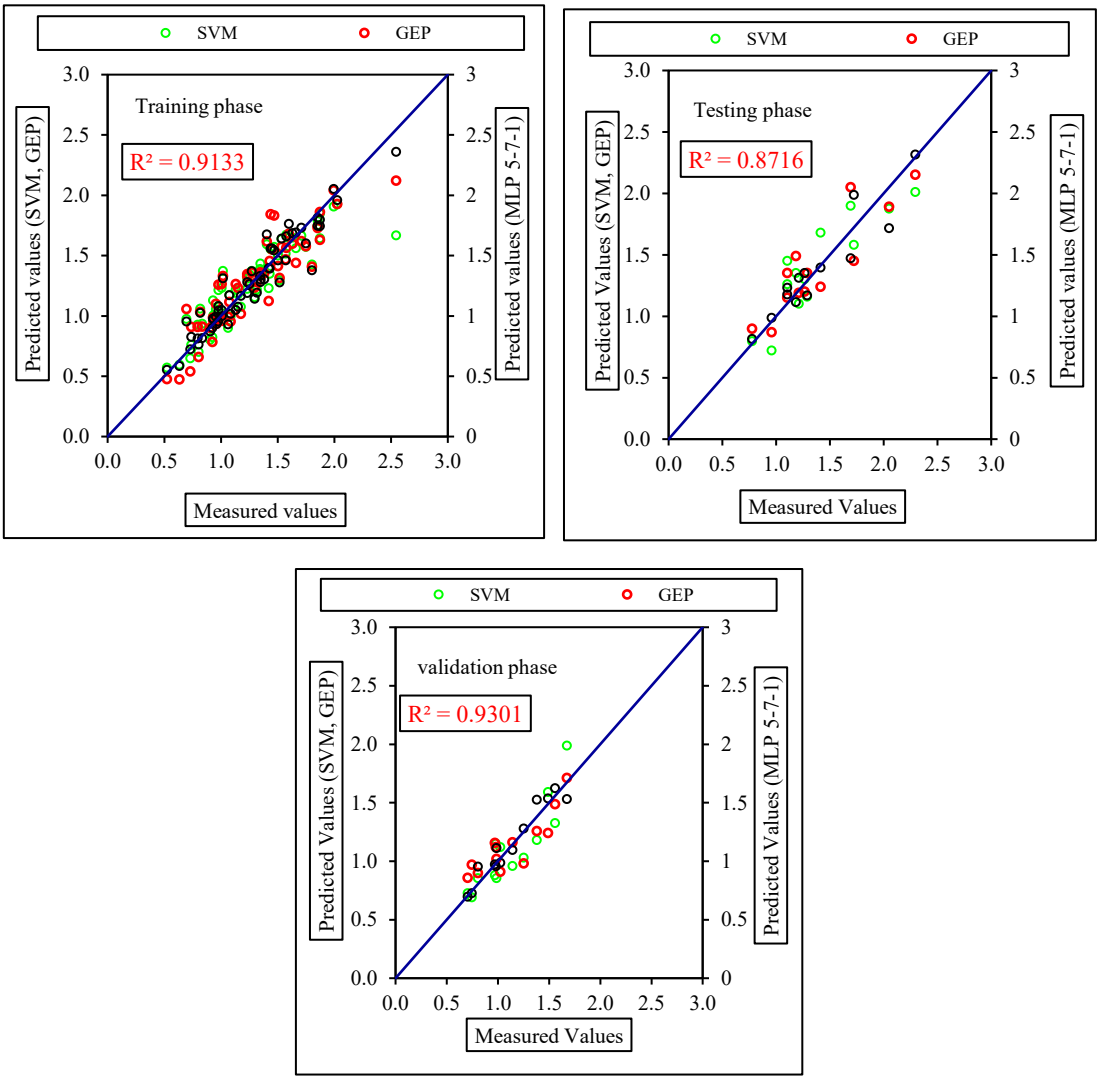


Figure 6. Scatter plot of measured vs. predicted values of the outcome of the MLP.

In this study, we showcase the potential of artificial intelligence in predicting the C_d for trapezoidal labyrinth side weirs. These weirs directly impact the sustainability of water resource management. It is highly important to predict the discharge coefficient accurately to optimize the hydraulics of water systems. Irrigation systems, flood control, and wastewater systems are a few of the most important water sectors that are benefited. As it is well understood, climate change has imposed frequent flash floods in many parts of the world and a quick response may not be available in some of the areas for such disasters. Economical solutions would be highly essential in developing countries for such scenarios. Therefore, trapezoidal labyrinth side weirs can be useful in such cases with minimum human intervention. The efficiency of these weirs, as discussed, is solely dependent on the discharge coefficient. Therefore, the research gap that has been addressed in this study showcases a greater potential for implementation in the real-world.

Among the three models tested, ANN-MLP 5-7-1 demonstrates the highest predictive accuracy across training, testing, and validation phases. The ANN model's superior performance underscores the potential of neural networks in handling complex, nonlinear hydraulic behaviours. While the SVM and GEP models provided reasonable accuracy, their limitations suggest that further refinement, such as hybridizing machine learning approaches, may yield better results. Future research could explore the application of deep learning techniques, hybrid models, and real-time data integration to improve prediction accuracy further. Additionally, policymakers and engineers should consider adopting machine learning-based predictive tools for infrastructure planning, aligning with global efforts toward sustainable and resilient water management practices.

4. Conclusions

We successfully developed and compared three machine learning models (SVM, GEP, and ANN) for predicting the C_d in two-cycle trapezoidal labyrinth side weirs using experimental data collected from Emiroglou et al. [1]. The primary contribution of this work is twofold. First, we introduce the novel application of a Γ-test sensitivity analysis to this domain, providing a robust, data-driven methodology to objectively identify the most influential input parameters Fr, $\frac{L}{B}$, $\frac{L_e}{L}$, $\frac{Y_1\text{-}P}{P}$ and α, thereby optimizing model architecture and enhancing predictive reliability. Second, it demonstrates that artificial intelligence, particularly an optimized ANN (MLP5-7-1), can achieve superior accuracy in modeling the complex, non-linear hydraulics of these structures, outperforming SVM and GEP models. The exceptional performance of the ANN model ($R^2 = 0.9301$) aligns with the results of Azamathulla et al. [49], who also found ANN and SVM to be highly effective for standard rectangular side weirs, and Bagheri et al. [10,11], who successfully applied ANN to sharp-crested side weirs. Furthermore, the strong predictive capability of the developed GEP model, which provides an explicit formulation, corroborates the work of Ebtehaj et al. [46] and Hussain et al. [59], who demonstrated the utility of GEP for rectangular side weirs.

The findings confirm that MLMs are powerful tools for hydraulic prediction. The ANN model's exceptional performance (evidenced by the lowest RMSE and MAE and highest R^2 across all phases) underscores its capability to capture the intricate relationships governing flow over trapezoidal labyrinth weirs. This offers a highly precise and efficient computational alternative to traditional empirical formulas and costly physical modeling.

Despite the promising results, this study has certain limitations. The models are developed and

validated on a comprehensive yet laboratory-scale dataset. Their performance in real-world, large-scale field conditions with additional complexities like sediment transport, debris, or non-uniform inflow distributions remains to be verified. Furthermore, while the ANN model is a powerful predictive black box, the GEP model provides a distinct advantage by offering an explicit mathematical formula, highlighting a trade-off between predictive accuracy and model interpretability. These limitations pave the way for several future research directions. In the future, researchers could: (1) Explore hybrid ML models that combine the predictive power of ANNs with the interpretability of GEP; (2) investigate more advanced deep learning architectures and other potent algorithms like Gradient Boosting Machines; (3) validate and adapt these models using field data from operational irrigation or drainage systems, (4) expand the input parameters to include factors like viscosity, sediment load, and weir surface roughness; and (5) develop user-friendly software or graphical interfaces based on the optimal ANN model to facilitate its adoption by practicing engineers. In broader implications, this research provides a validated AI framework that can contribute significantly to the sustainable design and management of water infrastructure. By enabling more accurate prediction of side weir performance, it aids in optimizing irrigation networks, enhancing flood mitigation structures, and improving urban drainage systems. This work aligns with the global need for intelligent, data-driven solutions in water resource engineering, supporting more resilient and adaptive water management practices in the face of growing environmental challenges.

The presented models, particularly the optimal ANN architecture with its detailed parameters, provide a ready foundation for developing user-friendly software tools, such as a graphical user interface (GUI), to facilitate adoption by practicing engineers in the design and analysis of trapezoidal labyrinth side weirs.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. E. M. Emiroglu, M. Cihan Aydin, N. Kaya, Discharge characteristics of a trapezoidal labyrinth side weir with one and two cycles in subcritical flow, *J. Irrig. Drain. Eng.,* **140** (2014), 04014007. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000709
2. X. Han, H. Zhang, N. Yi, G. Gao, Experimental study on flow capacity of a typical side weir, In: *Proceedings of the 5th International Conference on Advances in Civil and Ecological Engineering Research*, Springer, (2024), 245–255. https://doi.org/10.1007/978-981-99-5716-3_21
3. H. F. Isleem, M. K. Elshaarawy, A. K. Hamed, Analysis of flow dynamics and energy dissipation in piano key and labyrinth weirs using computational fluid dynamics, In: *Computational Fluid Dynamics-Analysis, Simulations, and Applications*, IntechOpen, (2024). https://doi.org/10.5772/intechopen.1006332
4. S. M. Seyedian, O. Kisi, Uncertainty analysis of discharge coefficient predicted for rectangular

side weir using machine learning methods, *J. Hydrol. Hydromechan.*, **72** (2024), 113–130. https://doi.org/10.2478/johh-2023-0043

5. Z. M. Hadi, H. Q. Majeed, Experimental and numerical study of the discharge capacity of a labyrinth side weir in a straight channel, *AIP Conference Proceedings*, **3091** (2024), 020038. https://doi.org/10.1063/5.0204400
6. S. M. Borghei, M. A. Nekooie, H. Sadeghian, M. R. J. Ghazizadeh, Triangular labyrinth side weirs with one and two cycles, *Proc. Inst. Civ. Eng. Water Manag.*, **166** (2011), 27–42. https://doi.org/10.1680/wama.11.00032
7. M. R. Namaee, An investigation of flow over side weir by numerical model, *J. De L'hydraulique*, **36** (2014), 1–11. Available from: https://www.persee.fr/doc/jhydr_0000-0001_2014_act_36_1_2313
8. S. Balahang, M. Ghodsian, Evaluating performance of various methods in predicting triangular sharp-crested side weir discharge, *Appl. Water Sci.*, **13** (2023), 171. https://doi.org/10.1007/s13201-023-01971-w
9. A. Lindermuth, T. St Pierre Ostrander, S. Achleitner, B. Gems, M. Aufleger, Discharge calculation of side weirs with several weir fields considering the undisturbed normal flow depth in the channel, *Water*, **13** (2021), 1717. https://doi.org/10.3390/w13131717
10. S. Bagheri, A. H. Kabiri-Samani, M. Heidarpour, Discharge coefficient of rectangular sharp-crested side weirs Part I: Traditional weir equation, *Flow Meas. Instrum.*, **35** (2014), 109–115. https://doi.org/10.1016/j.flowmeasinst.2013.11.005
11. S. Bagheri, A. R. Kabiri-Samani, M. Heidarpour, Discharge coefficient of rectangular sharp-crested side weirs part II: Domınguez's method, *Flow Meas. Instrum.*, **35** (2014), 116–121. https://doi.org/10.1016/j.flowmeasinst.2013.10.006
12. T. Nandesamoorthy, A. Thomson, Discussion of spatially varied flow over side weir, *ASCE J. Hydraul. Division*, **98** (1972), 2234–2235. https://doi.org/10.1061/JYCEAJ.0003529
13. K. Subramanya, S. C. Awasthy, Spatially varied flow over side weirs, *ASCE J. Hydraul. Division*, **98** (1972), 1–10. https://doi.org/10.1061/JYCEAJ.0003188
14. L. Yu-Tech, Discussion of spatially varied flow over side weir, *ASCE J. Hydraul. Division*, **98** (1972), 2046–2048. https://doi.org/10.1061/JYCEAJ.0003489
15. C. P. Kumar, S. K. Pathak, Triangular side weirs, *ASCE J. Irrig. Drain. Eng.*, **113** (1987), 98–105. https://doi.org/10.1061/(ASCE)0733-9437(1987)113:1(98)
16. K. G. R. Raju, B. Prasad, S. K. Grupta, Side weir in rectangular channel, *ASCE J. Hydraul. Division*, **105** (1979), 547–554. https://doi.org/10.1061/JYCEAJ.0005207
17. W. H. Hager, Lateral outflow over side weirs, *ASCE J. Hydraul. Eng.*, **113** (1987), 491–504. https://doi.org/10.1061/(ASCE)0733-9429(1987)113:4(491)
18. H. F. Cheong, Discharge coefficient of lateral diversion from trapezoidal channel, *ASCE J. Irrig. Drain. Eng.*, **117** (1991), 321–333. https://doi.org/10.1061/(ASCE)0733-9437(1991)117:4(461)
19. R. Singh, D. Manivannan, T. Satyanarayana, Discharge coefficient of rectangular side weirs, *ASCE J. Irrig. Drain. Eng.*, **20** (1994), 814–819. https://doi.org/10.1061/(ASCE)0733-9437(1994)120:4(814)
20. P. K. Swamee, S. K. P. Santosh, S. A. Masoud, Side weir analysis using elementary discharge coefficient, *ASCE J. Irrig. Drain. Eng.*, **20** (1994), 742–755. https://doi.org/10.1061/(ASCE)0733-9437(1994)120:4(742)
21. M. R. Jalili, S. M. Borghei, Discussion: Discharge coefficient of rectangular side weirs, *ASCE J.*

Irrigat. Drain. Eng., **122** (1996), 132. https://doi.org/10.1061/(ASCE)0733-9437(1996)122:2(132)
22. M. Borghei, M. R. Jalili, M. Ghodsian, Discharge coefficient for sharp-crested side weir in subcritical flow, *ASCE J. Hydraul. Eng.*, **125** (11999), 1051–1056. https://doi.org/10.1061/(ASCE)0733-9429(1999)125:10(1051)
23. M. Ghodsian, Flow over triangular side weir, *Sci. Iran. Sharif Univ. Technol.*, **11** (2004), 114–120.
24. K. H. V. Durga Rao, C. R. S. Pillai, Study of flow over side weirs under supercritical conditions, *Water Resour. Manag.*, **22** (2008), 131–143. https://doi.org/10.1007/s11269-007-9153-4
25. O. Bilhan, M. E. Emiroglu, O. Kisi, Application of two different neural network techniques to lateral outflow over rectangular side weirs located on a straight channel, *Adv. Eng. Software*, **41** (2010), 831–837. https://doi.org/10.1016/j.advengsoft.2010.03.001
26. O. Bilhan, M. E. Emiroglu, O. Kisi, Use of artificial neural networks for prediction of discharge coefficient of triangular labyrinth side weir in curved channels, *Adv. Eng. Software*, **42** (2011), 208–214. https://doi.org/10.1016/j.advengsoft.2011.02.006
27. M. E. Emiroglu, H. Agaccioglu, N. Kaya, Discharging capacity of rectangular side weirs in straight open channels, *Flow Meas. Instrum.*, **22** (2011), 319–330. https://doi.org/10.1016/j.flowmeasinst.2011.04.003
28. M. E. Emiroglu, N. Kaya, Discharge coefficient for trapezoidal labyrinth side weir in subcritical flow, *Water Resour. Manag.*, **25** (2011), 1037–1058. https://doi.org/10.1007/s11269-010-9740-7
29. A. Keshavarzi, J. Ball, Discharge coefficient of sharp-crested side weir in trapezoidal channel with different side-wall slopes under subcritical flow conditions, *Irrig. Drain.*, **63** (2014), 512–522. https://doi.org/10.1002/ird.1856
30. M. Ura, Y. Kita, J. Akiyama, H. Moriyama, J. A. Kumar, Discharge coefficient of oblique side weirs, *J. Hydrosci. Hydraul. Eng.*, **19** (2001), 85–96.
31. M. E. Emiroglu, N. Kaya, H. Agaccioglu, Discharge capacity of labyrinth side weir located on a straight channel, *J. Irrig. Drain. Eng.*, **136** (2010), 37–46. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000112
32. S. M. Borghei, A. Parvaneh, Discharge characteristics of a modified oblique side weir in subcritical flow, *Flow Meas. Instrum.*, **22** (2011), 370–376. https://doi.org/10.1016/j.flowmeasinst.2011.04.009
33. M. Karimi, M. J. Ghazizadeh, M. Saneie, J. Attari, Flow characteristics over asymmetric triangular labyrinth side weirs, *Flow Meas. Instrum.*, **68**, (2019), 101574. https://doi.org/10.1016/j.flowmeasinst.2019.101574
34. B. S. Hussein, S. A. Jalil, Hydrodynamic behavior simulation of flow performance over labyrinth side weir, *Polish J. Environ. Studies*, **33** (2024), 1159–1171. https://doi.org/10.15244/pjoes/173442
35. Y. Qian, P. Guo, Y. Wang, Y. Zhao, H. Lin, Y. Liu, Advances in laboratory-scale hydraulic fracturing experiments, *Adv. Civil Eng.*, (2020), 1–18. https://doi.org/10.1155/2020/1386581
36. A. M. Armanuos, M. K. Elshaarawy, Estimating saltwater wedge length in sloping coastal aquifers using explainable machine learning models, *Earth Sci. Inform.*, **18** (2025), 405. https://doi.org/10.1007/s12145-025-01900-2
37. A. K. Hamed, M. K. Elshaarawy, Soft computing approaches for forecasting discharge over symmetrical piano key weirs, *AI Civil Eng.*, **4** (2025), 6. https://doi.org/10.1007/s43503-024-00048-0

38. A. K. Hamed, M. K. Elshaarawy, M. M. Alsaadawi, Stacked-based machine learning to predict the uniaxial compressive strength of concrete materials, *Comput. Struct.*, **308** (2025), 107644. https://doi.org/10.1016/j.compstruc.2025.107644
39. R. Fatahi-Alkouhi, E. Afaridegan, N. Amanian, Discharge coefficient estimation of modified semi-cylindrical weirs using machine learning approaches, *Stochast. Environ. Res. Risk Assess.*, **38** (2024), 3177–3198. https://doi.org/10.1007/s00477-024-02739-7
40. S. Emami, H. Emami, J. Parsa, LXGB: A machine learning algorithm for estimating the discharge coefficient of pseudo-cosine labyrinth weir, *Sci. Rep.*, **13** (2023). https://doi.org/10.1038/s41598-023-39272-6
41. W. Chen, D. Sharifrazi, G. Liang, S. S. Band, K. W. Chau, A. Mosavi, Accurate discharge coefficient prediction of streamlined weirs by coupling linear regression and deep convolutional gated recurrent unit, *Eng. Appl. Comput. Fluid Mechan.*, **16** (2022), 965–976. https://doi.org/10.1080/19942060.2022.2053786
42. N. H. Elmasry, M. K. Elshaarawy, Hybrid metaheuristic optimized Catboost models for construction cost estimation of concrete solid slabs, *Sci. Rep.*, **15** (2025), 21612. https://doi.org/10.1038/s41598-025-06380-4
43. D. F. N. Dursun, M. Firat Kaya, Estimating discharge coefficient of semi-elliptical side weir using ANFIS, *J. Hydrol.*, **426** (2012), 55–62. https://doi.org/10.1016/j.jhydrol.2012.01.010
44. A. H. Zaji, H. Bonakdari, Performance evaluation of two different neural network and particle swarm optimization methods for prediction of discharge capacity of modified triangular side weirs, *Flow Meas. Instrum.*, **40** (2014), 149–156. https://doi.org/10.1016/j.flowmeasinst.2014.10.002
45. A. Parsaie, A. Haghiabi, The effect of predicting discharge coefficient by neural network on increasing the numerical modeling accuracy of flow over side weir, *Water Resour. Manag.*, **29** (2015), 973–985. https://doi.org/10.1007/s11269-014-0827-4
46. I. Ebtehaj, H. Bonakdari, A. H. Zaji, H. Azimi, A. Sharifi, Gene expression programming to predict the discharge coefficient in rectangular side weirs, *Appl. Soft Comput.*, **35** (2015), 618–628. https://doi.org/10.1016/j.asoc.2015.07.003
47. A. Eghbalzadeh, M. Javan, M. Hayati, A. Amini, Discharge prediction of circular and rectangular side orifices using artificial neural networks, *KSCE J. Civ. Eng.*, **20** (2016), 990–996. https://doi.org/10.1007/s12205-015-0440-y
48. F. Khoshbin, H. Bonakdari, S. H. Ashraf Talesh, Adaptive neuro-fuzzy inference system multi-objective optimization using the genetic algorithm/singular value decomposition method for modelling the discharge coefficient in rectangular sharp-crested side weirs, *Eng. Optim.*, **48** (2016), 933–948. https://doi.org/10.1080/0305215X.2015.1071807
49. H. M. Azamathulla, A. H. Haghiabi, A. Parsaie, Prediction of side weir discharge coefficient by support vector machine technique, *Water Supply*, **16** (2016), 1002–1016. https://doi.org/10.2166/ws.2016.014
50. Y. Yasi, Z. Gholami, Performance evaluation of discharge coefficient in physical models of labyrinth fuse gate spillways with intellectual and statistical models, *Iran J. Irrig. Drain.*, **11** (2017), 798–809.
51. H. Karami, S. Karimi, M. Rahmanimanesh, S. Farzin, Predicting discharge coefficient of triangular labyrinth weir using support vector regression support vector regression-firefly response surface methodology and principal component analysis, *Flow Meas. Instrum.*, **55** (2017), 75–81. https://doi.org/10.1016/j.flowmeasinst.2016.11.010

52. S. Shabanlou, Improvement of extreme learning machine using self-adaptive evolutionary algorithm for estimating discharge capacity of sharp-crested weirs located on the end of circular channels, *Flow Meas. Instrum.*, **59** (2018), 63–71. https://doi.org/10.1016/j.flowmeasinst.2017.11.003
53. A. H. Azimi, A. Rajabi, S. Shabanlu, Optimized ANFIS-genetic algorithm-particle swarm optimization model for estimation of side orifices discharge coefficient, *J. Numer. Methods Civ. Eng.*, **2** (2018), 27–38. https://doi.org/10.29252/nmce.2.4.27
54. R. Ezzeldin, A. Hatata, Application of NARX neural network model for discharge prediction through lateral orifices, *Alex. Eng. J.*, **57** (2018), 2991–2998. https://doi.org/10.1016/j.aej.2018.04.001
55. A. Parsaie, A. Haghiabi, Z. Shamsi, Intelligent mathematical modeling of discharge coefficient of non-linear weirs with triangular plan, *AUT J. Civ. Eng.*, **3** (2019), 149–156.
56. M. Majedi Asl, M. Fuladipanah, Application of the evolutionary methods in determining the discharge coefficient of triangular labyrinth weirs, *JWSS-Isfahan Univ. Technol.*, **22** (2019), 279–290. https://doi.org/10.29252/jstnar.22.4.279
57. A. Parsaie, A. H. Haghiabi, Mathematical expression for discharge coefficient of Weir-Gate using soft computing techniques, *J. Appl. Water Eng. Res.*, (2020), 1–9. https://doi.org/10.1080/23249676.2020.1787250
58. A. Y. Mohammed, A. Sharifi, Gene Expression Programming (GEP) to predict coefficient of discharge for oblique side weir, *Appl. Water Sci.*, **10** (2020), 1–9. https://doi.org/10.1007/s13201-020-01211-5
59. A. Hussain, A. Shariq, M. Danish, M. A. Ansari, Discharge coefficient estimation for rectangular side weir using GEP and GMDH methods, *Adv. Comput. Design*, **6** (2021), 135–151.
60. Z. Hu, H. Karami, A. Rezaei, Y. DadrasAjirlou, Md. J. Piran, S. S. Band, et al., Using soft computing and machine learning algorithms to predict the discharge coefficient of curved labyrinth overflows, *Eng. Appl. Comput. Fluid Mech.*, **15** (2021), 1002–1015. https://doi.org/10.1080/19942060.2021.1934546
61. M. D. Mustafa, T. Mansoor, M. Muzzammil, Support vector machine (SVM) approach to develop the discharge prediction model for triangular labyrinth weir, *Water Supply*, **22** (2022), 8942–8956. https://doi.org/10.2166/ws.2022.393
62. M. Fuladipanah, M. Majedi-Asl, Soft computing application to amplify discharge coefficient prediction in side rectangular weirs, *Irrig. Water Eng.*, **12** (2022), 213–233.
63. M. Majedi-Asl, M. Fuladipanah, V. Arun, R. P. Tripathi, Using data mining methods to improve discharge coefficient prediction in Piano Key and Labyrinth weirs, *Water Supply*, **22** (2022), 1964–1982. https://doi.org/10.2166/ws.2021.304
64. S. Li, G. Shen, A. Parsaie, G. Li, D. Cao, Discharge modeling and characteristic analysis of semi-circular side weir based on the soft computing method, *J. Hydroinform.*, **26** (2024), 175–188. https://doi.org/10.2166/hydro.2023.268
65. W. Tian, H. F. Isleem, A. K. Hamed, M. K. Elshaarawy, Enhancing discharge prediction over Type-A piano key weirs: An innovative machine learning approach, *Flow Meas. Instrum.*, **100** (2024), 102732. https://doi.org/10.1016/j.flowmeasinst.2024.102732
66. P. Heramb, P. K. Singh, K. V. R. Rao, A. Subeesh, Modelling reference evapotranspiration using gene expression programming and artificial neural network at Pantnagar, India, *Inf. Process. Agric.*, (2022). https://doi.org/10.1016/j.inpa.2022.05.007

67. M. Fuladipanah, A. Shahhosseini, N. Rathnayake, H. M. Azamathulla, U. Rathnayake, D. P. P. Meddage, et al., In-depth simulation of rainfall–runoff relationships using machine learning methods, *Water Pr. Technol.*, **19** (2024), 2442–2459. https://doi.org/10.2166/wpt.2024.147
68. N. Koncar, Optimisation methodologies for direct inverse neurocontrol (Publication No. SW72BZ), Doctoral dissertation, London University, England, (1997).
69. A. Malik, Y. Tikhamarine, N. Al-Ansari, S. Shahid, H. S. Sekhon, R. K. Pal, et al., Daily pan-evaporation estimation in different agroclimatic zones using novel hybrid support vector regression optimized by Salp swarm algorithm in conjunction with gamma test, *Eng. Appl. Comput. Fluid Mech.*, **15** (2021), 1075–1094. https://doi.org/10.1080/19942060.2021.1942990

AIMS Press