Gonbad Kavous University

Journal of New Approaches in Water Engineering and Environment

Online ISSN: 2981-0280

Homepage: https://nawee.gonbad.ac.ir

# Comparative Evaluation of Soft Computing Models for Predicting the Discharge Coefficient of Arced Labyrinth Weirs

**Mahdi Esmaeilzadeh Fereydani**[1], **Alireza Yazdani**[2,*], **Mehdi Fuladipanah**[3]

[1] Department of Civil Engineering, Na.C., Islamic Azad University, Najafabad, Iran

## Article Info

**Article type:**
Research Article

**Article history:**
Received 23 December 2025
Revised 17 February 2026
Accepted 05 March 2026
Published online 22 December 2026

**Keywords**:
Arced Labyrinth Weir
Discharge Coefficient
Machine Learning Models
Predictability

## ABSTRACT

**Objective**: The accurate prediction of the discharge coefficient (e hydraulic design of labyrinth weirs. While existing literature extensively covers conventional vertical-sidewall geometries, there is a critical knowledge gap regarding the synergistic effect of sloped sidewalls and arced planform geometry on discharge capacity. This study addresses this limitation by investigating and predicting $C_d$ for arced labyrinth weirs with varying sidewall slopes.

**Methods**: Using a comprehensive dataset of 277 experimental runs from a laboratory flume, five dimensionless parameters—namely$\frac{H_0}{P}\frac{A}{t}\frac{L_C}{W}$sional analysis to capture both geometric and hydraulic complexity- were used as inputs for three soft computing approaches: Support Vector Machine (SVM), Gene Expression Programming (GEP), and Artificial Neural Network (ANN).

**Results**: All three data-driven approaches produced satisfactory predictions of Cd. However, the ANN model based on a multilayer perceptron (MLP) demonstrated the best performance. The optimal architecture was MLP 5-9-1 trained using the BFGS algorithm, with the SOS error function, a logistic activation function in the hidden layer, and an identity activation function in the output layer. During training and validation, the performance metrics (RMSE, MAE, $R^2$, $C_{d(DDR(max)))}$ were (0.022, 0.0145, 0.8402, 11.436) and (0.0199, 0.0145, 0.8961, 12.805), respectively. Graphical analyses also showed strong agreement between observed and predicted Cd values across different dataset quartiles.

**Conclusions**: The findings confirm that the MLP framework provides a robust and reliable alternative for hydraulic engineers, filling the current modeling void for non-conventional labyrinth weirs. This approach offers a practical tool for spillway optimization where complex geometric interactions preclude the use of simplified linear regressions.

*Corresponding author, Email: alirezayazdani@iau.ac.ir

Cite this article: Esmaeilzadeh Fereydani M., Yazdani, AR., & Fuladipanah, M. (2026). Comparative evaluation of soft computing models for predicting the discharge coefficient of arced-labyrinth weir. Journal of New Approaches in Water Engineering and Environment, 5(4), 121-147. https://doi.org/10.22034/nawee.2026.585562.1232

© The Author(s).

Publisher: Gonbad Kavous University.

DOI: https://doi.org/10.22034/nawee.2026.585562.1232

## Introduction

Labyrinth weirs are increasingly used in dams and open channels because their folded planform provides a substantially longer crest length than a linear weir for the same width, enabling higher discharge capacity at relatively low heads and within a limited footprint (Idrees and Al-Ameri, 2022; Mattos-Villarroel et al., 2023; Crookston and Tullis, 2013; Emami et al., 2018). This makes them especially attractive for spillway rehabilitation, dam safety upgrades, and flow regulation in narrow valleys where enlarging the structure in plan is constrained (Mattos-Villarroel et al., 2023; Emiroglu and Kaya, 2010). However, the complex three-dimensional flow over labyrinth weirs—characterized by nappe interference, local submergence, and sensitivity to crest and sidewall geometry—introducesificant uncertainty into predicting the discharge coefficient, which is the key parameter controlling their hydraulic performance (Crookston and Tullis, 2013; Mattos-Villarroel et al., 2021; Aydın et al., 2024; Emami et al., 2018). A wide range of plan shapes (trapezoidal, triangular, rectangular, arced, and semicircular) and crest profiles (sharp, quarter-round, half-round, and WES)s have shown that relatively small geometric modifications can change the discharge coefficient by tens of percent, particularly at low head ratios where labyrinth weirs are most efficient (Mattos-Villarroel et al., 2025; Mattos-Villarroel et al., 2021; Alfatlawi et al., 2023; Namazi and Mozaffari, 2023; Crookston and Tullis, 2012). Despite extensive experimental and numerical workfully capture the combined effects of geometry, flow conditions, and site constraints, and further refinement of discharge-coefficient relationships remains essential for reliable and economical labyrinth-weir designrees and Al-Ameri, 2022; Mattos-Villarroel et al., 2023; Aydın et al., 2024).

Physical-model studies in laboratory flumes remain the primary basis for understanding how labyrinth-weir geometry and hydraulic conditions control discharge coefficient ($C_d$). Experiments on triangular, trapezoidal and rectangular labyrinth weirs consistently show that increasing crest length (via smaller vertex angle or longer sidewalls) increases discharge capacity, while $C_d$ typically decreases with increasing head ratio $\frac{H_T}{P}$ or $\frac{h}{P}$ (Seyedian et al., 2023; Ghasemvand et al., 2024; Hussain et al., 2022; Masoudi et al., 2024). For triangular labyrinth weirs, reducing the vertex angle (e.g., to 60°) and decreasing weir height significantly increases $C_d$ relative to linear weirs, confirming the key role of length magnification and P in performance (Hussain et al., 2022). Comparative tests on linear versus arc-plan trapezoidal and triangular labyrinth weirs indicate arc plans can reduce $C_d$ but greatly increase hydraulic loss; rectangular and arc-linear labyrinth variants show better trade-offs between $C_d$ and energy dissipation (Ghasemvand et al., 2024). Studies on dentate trapezoidal labyrinth weirs demonstrate that adding rectangular, triangular or trapezoidal teeth reduces upstream afflux and can increase effective discharge for a given footprint, although $C_d$ still declines with $\frac{H_T}{P}$ and with increasing P. Experiments on inclined circular labyrinth weirs show that lower inclination angles (e.g., 60° from vertical) maximize $C_d$, whereas diameter has only minor influence; free flow tends to become submerged as discharge rises (Zare et al., 2023). Recent work on rectangular and

trapezoidal labyrinth weirs with non-uniform cycle lengthsonger middle or side cycles) reports $C_d$ improvements of about 35–41% compared with equal layouts, emphasizing the importance of non-uniform plan optimization (Masoudi et al., 2024). For multi-cycle W-form and sharp-crested circular arc weirs, systematic variation of height, apex angle, and number of cycles documents large gains in discharge relative to linear weirs and provides high-quality datasets later reused in data-driven modeling (Dutta et al., 2020). New experimental configurations, such as semicircular labyrinth weirs with multiple cycle numbers and heights, and triangular-arced asymmetric labyrinth weirs with variable intermediate-cycle width and arc angle, also show that carefully tuned arcs and widened middle cycles can produce $C_d$ values approaching ~0.95 at optimal Froude numbers, especially for $\frac{H_T}{P}<0.25$ where labyrinths are most efficient (Asgharzadeh-Bonab et al., 2025; Emami et al., 2023; Tabatabai et al., 2025; Karami et al., 2016; Mattos-Villarroel et al., 2025). Across this body of work, dimensional analysis consistently identifies $\frac{H_T}{P}$ or $\frac{h}{P}$, length magnification $\frac{L_C}{P}$ or $\frac{L}{P}$, weir height P, shape factor, and sometimes Froude number as the governing dimensionless parametersr experimental $C_d$ characterization (Norouzi et al., 2019; Majedi-Asl et al., 2024; Seyedian et al., 2023; Ghasemvand et al., 2024; Hussain et al., 2022; Shafiei et al., 2020; Masoudi et al., 2024).

To complement experiments, numerical approaches based on CFD and machine-learning models have been widely deployed to predict $C_d$ for complex labyrinth geometries and related devices. CFD studies using VoF free-surface tracking and k–ε turbulence modelling in ANSYS Fluent successfully replicate two-phase flow over semicircular labyrinth weirs and quantify $C_d$ as a function of $\frac{H_T}{P}$, showing up to ~88% greater discharge capacity than comparable trapezoidal labyrinth weirs, especially for $\frac{H_T}{P}\leq 0.25$ and sidewall angles $\alpha\leq 12°$ (Mattos-Villarroel et al., 2025). However, most recent studies have focused on machine-learning models trained on experimental datasets.sed multilayer perceptron (MLP) and support vector machines (SVM) to estimate $C_d$ of trapezoidal labyrinth weirs with quarter-round crests from 454 lab measurements; with inputs $\frac{L}{W}$, α, $\frac{W}{P}$, and $\frac{H_T}{P}$, MLP achieved R≈0.985 and RMSE≈0.019, outperforming radial-basis networks and SVM variants (Norouzi et al., 2019). Meta-model comparisons for labyrinth weirs show that ANN and SVM generally surpass GEP and commercial tools, with an optimal ANN model reaching $R^2$=0.998 and RMSE=0.006 using $\frac{H_T}{P}$, $\frac{L_C}{P}$ and α, and sensitivity analysis confirming $\frac{H_T}{P}$ as the dominant predictor (Majedi-Asl et al., 2024). Other soft-computing approaches, such as outlier-robust extreme learning machine (ORELM), identify $\frac{H_T}{P}$and $\frac{w}{P}$ as the most influential inputs and yield better accuracy and uncertainty properties than standard ELM, ANN and SVM, while also providing an explicit empirical Cd equation for practitioners (Shafiei et al., 2020). For triangular labyrinth weirs, Gaussian process regression (GPR), quantile regression forest and least-squares SVM have been benchmarked; GPR attains $R^2$≈0.986 and RMSE≈0.009, offers probabilistic prediction intervals, and clearly outperforms classical empirical relations (Seyedian et al., 2023). In

triangular labyrinth side-weirs, extreme learning machine (ELM) surpasses ANN and genetic programming, with $R^2$ up to 0.993 in training and ~0.971 in testing and MAPE below 1% (Karami et al., 2016). Data-mining studies combining SVM and GEP on piano-key, rectangular and trapezoidal labyrinth weirs show both methods are accurate but GEP slightly outperforms SVM and yields interpretable symbolic relations; again, $\frac{H_T}{P}$, $\frac{L_C}{P}$ and Froude number emerge as the key variables depending on weir type (Majedi-Asl et al., 2021). Hybrid and advanced ML architectures have recently been introduced for specialized labyrinth geometries. For pseudo-cosine labyrinth weirs, an LXGB meta-model (LSHADE-optimized XGBoost) trained on 132 experimental data points outperforms SAELM, ANFIS-FFA, GEP and ANN, achieving RMSE as low as 0.009 and NSE≈0.98 when using $\frac{H}{W}$, $\frac{L}{W}$, and $\frac{L}{W}$ as inputs (Emami et al., 2023). For semicircular labyrinth weirs, TabNet-MFO and ELM-JFO, along with LightGBM and decision trees, provide very high accuracy (testing PI≈82 for TabNet-MFO, normalized RMSE E′≈0.012), with SHAP-based interpretability showing h/P, number of cycles N and l/P as the most important features (Asgharzadeh-Bonab et al., 2025). Several studies generalize ML to *labyrinth sluice gates*, where hybrid SVR-PSO and SVR-GA yield $R^2$≈0.97 and clearly beat traditional nonlinear regressions, and gradient boosting (GBM) models reach $R^2$≈0.997 with MAPE ~0.4%, providing robust tools for gate design (Azma et al., 2025; Hashem et al., 2025). Other ANN-based frameworks have targeted specific planforms such as type-D labyrinth weirs, highlighting how using approach-condition descriptors rather than just geometric ratios, and carefully selecting a 70/30 train–test split, can materially improve predictive skill (Belaabed and Arabet, 2024). Additional SVM models for arced and standard labyrinth weirs demonstrate good accuracy across 527 test points with inputs $\frac{H_T}{P}$, $\frac{\alpha}{\theta}$ and $\frac{L_C}{W}$, with sensitivity analysis indicating Froude number and $\frac{H_T}{P}$ as the most influential factors for standard and arced layouts, respectively (Roushangar et al., 2018). Overall, across weirs and gates, modern ML and hybrid models (ANN/MLP, ELM/ORELM, SVM/SVR, GPR, GEP, GBM, LXGB, TabNet-MFO) routinely attain R or $R^2$ above 0.97 and error metrics (RMSE, MAPE, NSE) consistent with design-level precision, while also enabling sensitivity and explainability analyses that align closely with classical hydraulic theory (Norouzi et al., 2019; Asgharzadeh-Bonab et al., 2025; Dutta et al., 2020; Azma et al., 2025; Emami et al., 2023; Seyedian et al., 2023; Karami et al., 2016; Mattos-Villarroel et al., 2025; Belaabed and Arabet, 2024; Majedi-Asl et al., 2021; Hashem et al., 2025; Shafiei et al., 2020; Roushangar et al., 2018).

**Table 1- Summary of literature review**

| Focus | Typical methods and contributions | Citations |
|---|---|---|
| Experimental hydraulics | Flume studies on planar shape, height, inclination, dentation, non-uniform cycles; Cd vs HT/P, h/P | (Dutta et al., 2020; Tabatabai et al., 2025; Seyedian et al., 2023; Zare et al., 2023; Ghasemvand et al., 2024; Hussain et al., 2022; Masoudi et al., 2024) |
| CFD/mathematical | VoF, k–ε modelling of semicircular labyrinths; regression of $C_d$(HT/P) | (Mattos-Villarroel et al., 2025) |
| Classical ML | ANN, SVM, ELM, GEP, ORELM, GPR for Cd prediction using dimensionless inputs | (Norouzi et al., 2019; Majedi-Asl et al., 2024; Seyedian et al., 2023; Karami et al., 2016; Majedi-Asl et al., 2021; Shafiei et al., 2020; Roushangar et al., 2018) |
| Hybrid / advanced ML | LXGB, TabNet-MFO, ELM-JFO, GBM, SVR-PSO/GA, ANN with optimized data splitting and inputs | (Asgharzadeh-Bonab et al., 2025; Azma et al., 2025; Emami et al., 2023; Belaabed and Arabet, 2024; Hashem et al., 2025) |

Although numerous studies have employed machine-learning techniques to estimate the discharge coefficient of labyrinth weirs, most have concentrated on conventional geometries and vertical-sidewall configurations. A comparative review of previous investigations further indicates that only limited attention has been given to arced labyrinth weirs, and the available studies have mainly focused on general hydraulic performance, geometric effects, or model accuracy evaluations under specific conditions. As a result, the available knowledge remains limited regarding arced labyrinth weirs with sloped sidewalls, despite the fact that sidewall inclination may significantly influence flow patterns, effective crest behavior, and discharge characteristics. Moreover, the combined effects of arced planform geometry and sloped sidewalls on discharge coefficient prediction have not been systematically investigated. Previous machine-learning-based studies have also largely emphasized predictive accuracy comparisons rather than assessing the applicability of such models to underexplored hydraulic configurations with practical design relevance. Therefore, the present study addresses an important gap by investigating the discharge coefficient of arced labyrinth weirs with sloped sidewalls through experimental data and data-driven modeling. The contribution of this work lies not in proposing a new machine-learning algorithm, but in extending predictive modeling to a specific hydraulic configuration that has not been sufficiently examined in earlier studies. In this way, the study not only evaluates the performance of selected machine-learning models for this geometry, but also improves understanding of the hydraulic behavior of this weir type for prediction and design purposes.

## Materials and Methods

### *Data gathering*

Experimental data were collected in the Hydraulic Laboratory of the University of Maragheh using a 13 m long, 1.20 m wide, and 0.80 m deep rectangular open-channel flume operated with a smooth horizontal bed (zero slope). Discharge was supplied by a variable-speed centrifugal

pump (7-52 L/s) and measured with an ultrasonic flow meter (≈0.01 L/s resolution). Approach-flow uniformity was improved using upstream flow conditioners (perforated screens and a surface-stabilizing plate), and a downstream tailgate was used to regulate tailwater level to achieve both free and submerged overflow conditions. For each run, measurements were recorded after approximately 10 minutes of stabilization; water levels upstream and downstream of the weir were measured with a point gauge (0.1 mm accuracy). The test program included multiple arced labyrinth weir models with different geometric configurations (cycle number, arc angle, and wall slope); representative photographs are provided, and the complete set of geometric and hydraulic test conditions is summarized in Table 2. All geometric and hydraulic parameters listed in Table 2 are illustrated in Fig. 1.

**Table 2- Summary of experimental dataset**

| No. | Weir height | U/S nose length (cm) | D/S nose length (cm) | Cycle number | Scale ratio | Wall slope | Cycle arc angle | Wall length (cm) |
|---|---|---|---|---|---|---|---|---|
| | P(cm) | | | N | $\frac{L_C}{W}$ | $\varphi$ | θ | |
| 1 | 10 | 5 | 5 | 3 | 2.81 | 0 | 20 | 34 |
| 2 | 10 | 5 | 5 | 3 | 2.81 | 0 | 30 | 34 |
| 3 | 10 | 5 | 5 | 3 | 2.81 | 0 | 45 | 34 |
| 4 | 9.85 | 5 | 8 | 3 | 2.94 | 10 | 20 | 34 |
| 5 | 9.85 | 8 | 5 | 3 | 2.94 | 10 | 20 | 34 |
| 6 | 9.4 | 5 | 11.5 | 3 | 3.06 | 20 | 20 | 34 |
| 7 | 9.4 | 11.5 | 5 | 3 | 3.06 | 20 | 20 | 34 |
| 8 | 9.8 | 5 | 8 | 3 | 2.94 | 10 | 30 | 34 |
| 9 | 9.8 | 8 | 5 | 3 | 2.94 | 10 | 30 | 34 |
| 10 | 9.4 | 5 | 11.5 | 3 | 3.06 | 20 | 30 | 34 |
| 11 | 10 | 5 | 5 | 3 | 2.21 | 0 | 20 | 26 |
| 12 | 10 | 5 | 5 | 3 | 2.21 | 0 | 30 | 26 |
| 13 | 10 | 5 | 5 | 3 | 2.21 | 0 | 45 | 26 |
| 14 | 10 | 5 | 5 | 3 | 2.21 | 0 | 50 | 26 |
| 15 | 10 | 5 | 5 | 4 | 3.9 | 0 | 20 | 34 |
| 16 | 10 | 5 | 5 | 5 | 4.22 | 0 | 20 | 34 |

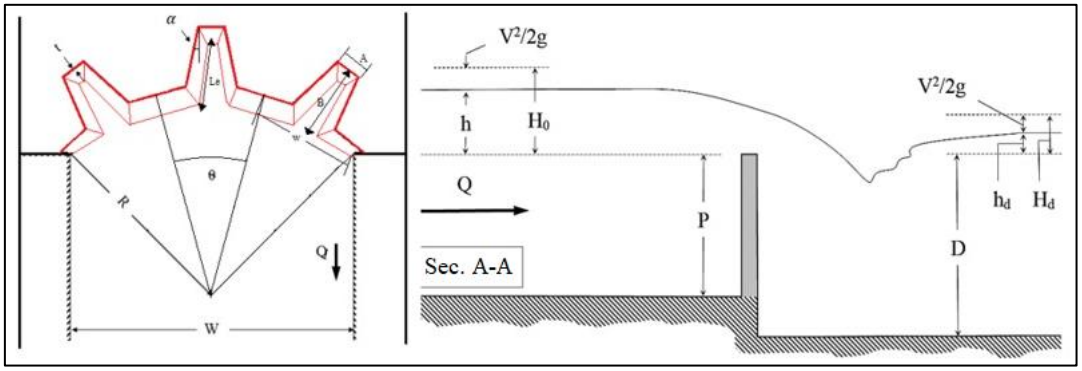


**Figure 1- Hydraulic and geometric parameters in a schematic view**

***Dimensional analysis***

The discharge capacity of the weir can be expressed by the general equation:

$$Q=\frac{2}{3}C_d L_c\sqrt{2g}H_0^{1.5} \quad (1)$$

where $L_c$ is the effective crest length. According to the longitudinal section and plan view of the arced labyrinth weir shown in Fig. 1, the discharge coefficient is influenced by a set of geometric and hydraulic parameters, which may be written as:

$$Q=f(H_0,P,D,L_c,t,\theta,S_0,W,H_d,g,\sigma,\mu,\rho,S_e,N,\alpha,A,\varphi) \quad (2)$$

Here Q denotes the flow discharge, $H_0$ is the upstream total head over the crest, P is the weir height, D represents the downstream depth, Lc is the effective crest length, t is the crest or wall thickness, θ is the arc angle, $S_0$ is the channel bed slope, W is the flume width, $H_d$ is the downstream head, g is the gravitational acceleration, σ is the surface tension, μ is the dynamic viscosity, ρ is the fluid density, Se is the crest shape factor, N is the number of cycles, α is the geometric angle of the crest section, A is the cross-sectional area, and φ is the wall slope angle.

Using dimensional analysis based on the Buckingham π-theorem, the governing relationship in Eq. (2) can be expressed in dimensionless form as:

$$f\left(\frac{Q}{L_c\sqrt{g}H_0^{1.5}},\frac{H_0}{P},\frac{W}{P},\frac{t}{P},\frac{D}{P},\frac{H_d}{H_0},\frac{A}{t},\frac{L_C}{W},\theta,S_e,S_0,N,\alpha,\varphi,\frac{\rho VL}{\mu},\frac{\rho V^2L}{\sigma}\right)=0 \quad (3)$$

Here, $\frac{\rho VL}{\mu}$ and $\frac{\rho V^2L}{\sigma}$ represent Reynolds and Weber numbers, respectively. For the experimental conditions considered in this study, the flow regime was fully turbulent, resulting in sufficiently high Re numbers. Moreover, the upstream water head over the weir was large enough to produce high We numbers; therefore, the effects of viscosity and surface tension were negligible. Consequently, the fluid properties were omitted from the dimensional relationship. Since the bed slope and crest geometry were constant in all experiments, $S_0$ and $S_e$ were omitted. In addition, because all tests were conducted under free-flow conditions, $\frac{H_d}{H_0}$ was excluded. Parameters $\frac{W}{P}$, α, $\frac{t}{P}$, N, and $\frac{D}{P}$ were also removed due to their constant values in all models. Hence, the simplified form becomes:

$$f\left(\frac{Q}{L_c\sqrt{g}H_0^{1.5}},\frac{H_0}{P},\frac{L_C}{W},\frac{A}{t},\theta,\varphi\right)=0 \quad (4)$$

Thus, the discharge coefficient can be expressed as:

$$C_d=f\left(\frac{H_0}{P},\frac{A}{t},\frac{L_C}{W},\theta,\varphi\right) \quad (5)$$

***Review of machine-learning models***

***Overview of GEP***

GEP, introduced by Ferreira (2001, 2002), is an evolutionary computation method that differs from GA and GP by clearly separating genotype and phenotype. In GEP, individuals are encoded as fixed-length linear chromosomes and then expressed as nonlinear Expression Trees (ETs). This genotype–phenotype mapping, implemented through the Karva language, ensures that genetic modifications always produce valid executable structures. Compared with GA and GP, this gives GEP greater flexibility and robustness in evolutionary modeling (Ahmadzadeh-Kaleybar et al., 2024).

The GEP evolutionary process begins with the random generation of an initial population of fixed-length chromosomes. Each chromosome features a head-tail gene architecture that guarantees syntactic validity and is deterministically decoded into Expression Trees (ETs). These ETs function as symbolic programs and are evaluated using a predefined fitness function. Based on these evaluations, the best-performing individuals are selected to reproduce, introducing variation through a diverse set of genetic operators- including replication, mutation, transposition (IS, RIS, and gene), and recombination. Due to the inherent head-tail structural organization, these operators preserve syntactic correctness at every generation. This cycle of expression, evaluation, and reproduction iterates until a termination criterion- such as a maximum number of generations or a satisfactory fitness level- is achieved.

***Overview of SVM***

The SVMs are supervised learning techniques based on statistical learning theory that aim to identify an optimal decision boundary (Vapnik, 1995). In regression tasks, the algorithm seeks a hyperplane that maximizes the margin between data points, effectively minimizing the structural risk to enhance generalization. The points closest to this hyperplane, known as support vectors, are critical in defining the model's predictive boundaries (Vapnik, 1995).

The SVR, an extension of the SVM framework for continuous-valued prediction, utilizes an ϵ-insensitive loss function to handle regression tasks (Vapnik, 1995; Smola, 1996). The objective is to find a regression function that deviates from the observed targets by no more than a predefined threshold εg errors smaller than this value as insignificant. This approach seeks to balance the accuracy of the fit with the smoothness of the function, effectively maximizing the margin within an ϵ-tube around the regression line as follows:

$$f(x) = \omega . x + b \tag{6}$$

where $\omega \in \mathbb{R}^N$ is the weight vector and $b \in \mathbb{R}$ is the bias term. The dot product $\omega . x$ is computed in an *N*-dimensional feature space. To ensure model smoothness and penalize deviations larger than ϵ, the SVR problem is formulated to minimize the weight vector norm $\|\omega\|^2$ alongside slack variables ($\xi_i$, $\xi_i'$) using a regularization parameter C (Smola, 1996). This parameter controls the trade-off between model complexity and training error tolerance. By transforming the primal formulation into its dual representation using Lagrange multipliers, the final regression function is expressed as r, 1984):

$$f(x) = \sum_{i=1}^{k} (\lambda_i - \lambda_i')(x_i . x) + b \tag{7}$$

where $\lambda_i$ and $\lambda_i'$ are non-negative Lagrange multipliers. To handle nonlinear relationships between input and output variables, SVR employs kernel functions that implicitly map the input data into a higher-dimensional feature space. In this case, the inner product $x_i.x_j$ is replaced by a Kernel function $K(x_i, x_j)$, defined as

$$K(x_i, x_j) = \Phi(x_i) . \Phi(x_j) \tag{8}$$

Where $\Phi(.)$ denotes the nonlinear mapping (Vapnik, 1995). The final nonlinear SVR model can therefore be written as

$$f(x) = \sum_{i=1}^{k} (\lambda_i - \lambda_i') K(x_i, x) + b \tag{9}$$

This formulation enables SVR to effectively capture complex and nonlinear dependencies in the data while retaining strong generalization capabilities. The definition of kernel functions is presented in Table fi et al., 2025).

**Table 3- Kernel functions used in SVRs**

| Kernel function | Definition |
|---|---|
| Linear | $K(x_i, x_j) = (x_i, x_j)$ |
| Polynomial | $K(x_i, x_j)=((x_i, x_j)+1)^d$ |
| Radial based function | $K(x_i, x_j) = exp(\frac{\|x_i, x_j\|^2}{2\gamma^2})$ |
| Spline | $K(x_i, x_j) = 1 + (x_i, x_j) + \frac{1}{2}\min(x_i, x_j) - \frac{1}{6}\min(x_i, x_j)^3$ |

Support Vector Machines were evaluated using correlation coefficients and RMSE within a cross-validation framework. The dataset was partitioned into multiple folds, where each fold was used once for validation and the remaining folds for training. This procedure provides an unbiased estimate of predictive performance by repeatedly training the SVR model on different subsets of the data.

***Overview on ANN***

The ANNs were employed as nonlinear data-driven models to estimate the discharge coefficient. The adopted architecture consisted of an input layer, one or more hidden layers, and an output layer (Tamizi et al., 2026). Neurons were connected through adjustable weights ($\omega_{ij}$), which were iteratively updated during training to minimize prediction error. For each neuron, the weighted sum of inputs was transformed using an activation function (σ), and the output of neuron j in layer l is expressed as:

$$z_j^l = \sum_i \omega_{ij}^l a_i^{l-1} + b_j^l \quad (10)$$

$$\alpha_j^l = \sigma(z_j^l) \quad (11)$$

In this formulation, $a_i^{l-1}$ denotes the activation values passed from layer $l$-1, $\omega_{ij}^l$ represents the trainable connection weights, and $b_j^l$ is the bias associated with neuron j in layer $l$. Nonlinearity in the network is introduced through an activation function σ. The model parameters are learned using the backpropagation algorithm, in which the loss function L(y, ŷ) between the observed target y and the predicted value ŷ is minimized through iterative gradient-based updates of the connection weights. The general weight update rule can be written as:

$$\omega_{ij}^l \leftarrow \omega_{ij}^l - \eta \frac{\partial L}{\partial \omega_{ij}^l} \quad (12)$$

Here, η is the learning rate (Aichouri et al., 2015).

***Assessment metrics***

To assess model accuracy and generalization ability, four statistical indices were used: root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (Ry ratio (DDR) as follows (Kisi et al., 2024):

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i\hat{y}_i)^2} \quad (13)$$

$$MAE=\frac{\sum_{i=1}^{n}|y_i\hat{y}_i|}{n} \quad (14)$$

$$R^2=1\frac{\sum_{i=1}^{n}(y_i\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i\bar{y})^2} \quad (15)$$

$$DDR=\frac{\text{Predicted value}}{\text{Observed value}}-1 \quad (16)$$

Here, n is the number of observations, $y_i$ is the actual value, and $\hat{y}_i$ is the predicted value for the $i^{th}$ observation. These complementary metrics provide a concise evaluation of prediction accuracy, error magnitude, explained variance, and relative bias. DDR is particularly useful for identifying systematic overestimation or underestimation patterns. In advanced applications, normalized forms such as $X_{DDR}$ and $Z_{DDR}$ can further improve bias diagnosis and visual interpretation of model behavior (Noori et al., 2010).

To provide a comprehensive overview of the data-driven predictive modeling process, the systematic workflow implemented in this study is illustrated in Fig. 2. The methodology is structured into three main phases: data preparation, model development, and performance evaluation. Initially, experimental data comprising 277 data points of geometric and hydraulic variables are gathered. Five primary independent parameters $(\frac{H_0}{P},\frac{A}{t},\frac{L_C}{W},\theta,\varphi)$d as inputs, while the $C_d$ is designated as the target output. Prior to model training, the dataset is pre-processed to check for consistency, remove outliers, and normalize the variables to enhance training convergence. The processed data is then randomly partitioned into training and testing subsets. In the model development phase, three distinct data-driven paradigms—namely, ANN, SVR, and GEP—are configured and trained. Finally, the predictive accuracy of the models is rigorously assessed using multiple statistical indicators (Rness and lowest error is selected to predict the discharge coefficient of the labyrinth weirs under various configurations.

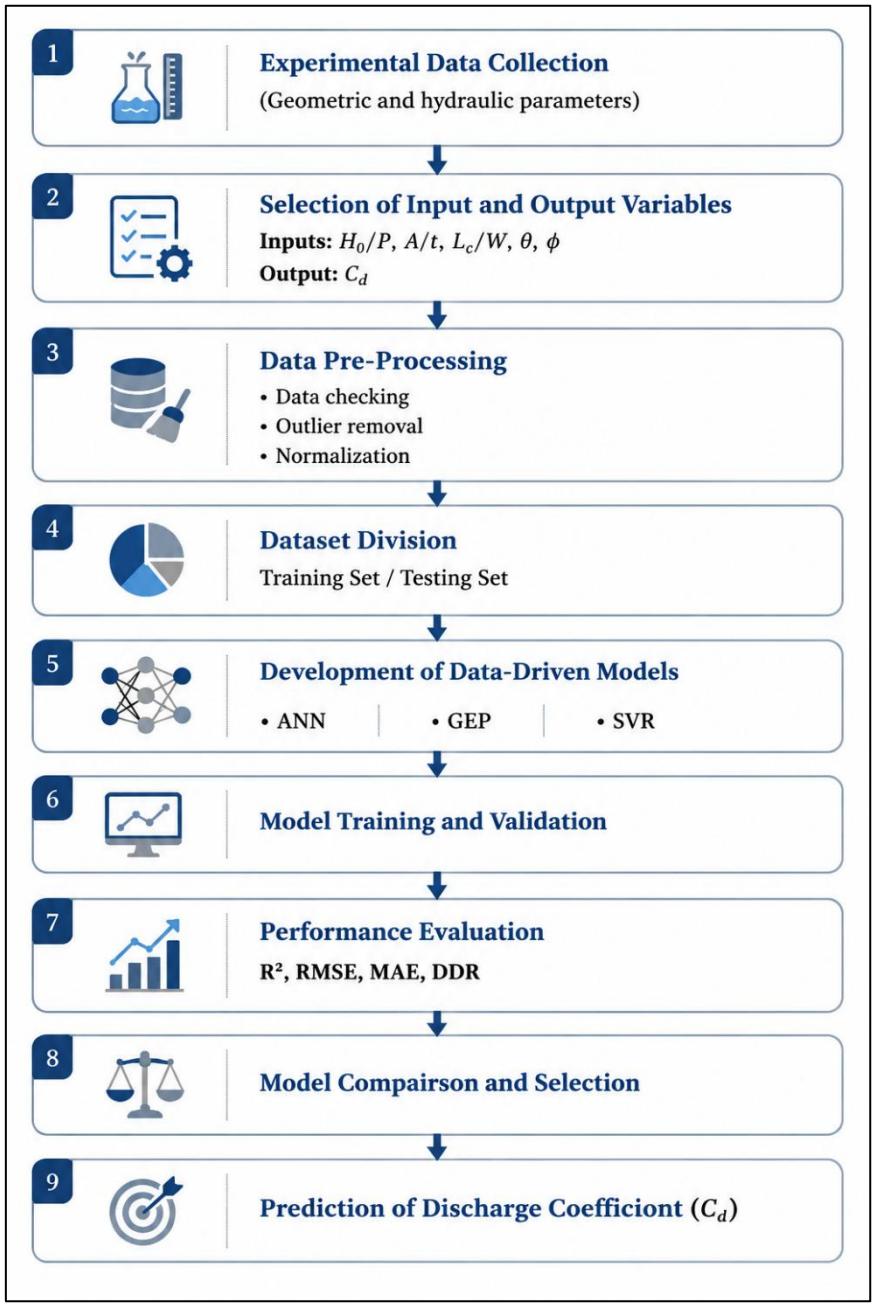


**Figure 2- Workflow of the present research**

## Results

### *Data Preprocessing and Normalization*

Prior to conducting the model training and evaluation, a data preprocessing step was performed to eliminate scale effects and ensure stable convergence across all three data-driven models (SVM, GEP, and ANN). Because the input variables possess different physical units and ranges, and some independent variables contain zero values, all raw experimental datasets were normalized. The normalization process maps all variables into a uniform range of [0, 1] using the Min-Max scaling technique, expressed as follows:

$$X_{normal} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (17)$$

Here, $X_{normal}$ is the normalized value of the variable, X is the original experimental value, $X_{min}$ and $X_{max}$ represent the minimum and maximum values of the corresponding parameter in the dataset, respectively. This normalization method successfully handles independent variables with minimum values of zero by mapping them directly to the lower boundary of the normalized scale (0) without any mathematical singularity, thereby preserving the physical integrity of the input feature space. The normalized datasets were subsequently used as the inputs for the training and testing phases of the ANN, GEP, and SVM models.

***The outcomes of SVM***

Based on dimensional analysis, the input space was defined according to Eq. 6. Since the hydraulic response is inherently nonlinear, the RBF kernel was adopted to map the inputs into a higher-dimensional feature space. Model performance was governed by the hyperparameters C, ε, and γ (Table 4).

**Table 4- Samples of tuning parameters of SVM model**

| | C | Γ | ε |
|---|---|---|---|
| Group 1 | 10 | 0.01 | 0.25 |
| Group 2 | 0.5 | 0.001 | 0.001 |
| Group 3 | 1 | 0.05 | 0.05 |
| Group 4 | 25 | 0.2 | 0.1 |
| Group 5 | 50 | 1 | 0.5 |

Parameter C balanced training error and regression smoothness, ε regulated sensitivity to minor fluctuations, and γ determined the locality of the RBF kernel. The dataset was split into 193 training and 84 testing samples (n = 277sed on cross-validation, the optimal hyperparameters were determined as C=10, ε=0.25, and γ=0.01, indicating a smooth regression structure with balanced generalization.

**Table 5- The values of assessment metrics of SVM of different values of tuning parameters**

| | Training phase | | | | Testing phase | | | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ |
| Group 1 | 0.0379 | 0.0271 | 0.6024 | 6.7222 | 0.0242 | 0.0207 | 0.7281 | 10.0475 |
| Group 2 | 0.4427 | 0.3671 | 0.0438 | 0.6309 | 0.3182 | 0.2674 | 0.0225 | 0.7834 |
| Group 3 | 0.2970 | 0.2511 | 0.0314 | 0.8697 | 0.1321 | 0.1110 | 0.1653 | 1.8667 |
| Group 4 | 5.3720 | 4.5348 | 0.0002 | 0.0821 | 0.1327 | 0.1123 | 0.1533 | 2.0364 |
| Group 5 | 0.5976 | 0.4885 | 0.0168 | 0.4775 | 0.2773 | 0.2419 | 0.0312 | 0.8827 |

The predictive capability of the optimized SVM model is illustrated in Figs. 3, 4, and 5. The predicted $C_d$ values closely track the observed trends in both phases (Fig. 3), with minor deviations at peak values. This accuracy is further supported by the standardized DDR index ($Z_{DDR}$) distributions (Fig. 4), which are symmetrically centered at $Z_{DDR} \approx 0$ across the groups, indicating unbiased predictions. Additionally, the 2D Kernel Density Estimation (KDE) plots

(Fig. 5) display high-density cores (red regions) tightly aligned with the 45° diagonal, confirming minimal dispersion and strong generalization without overfitting.

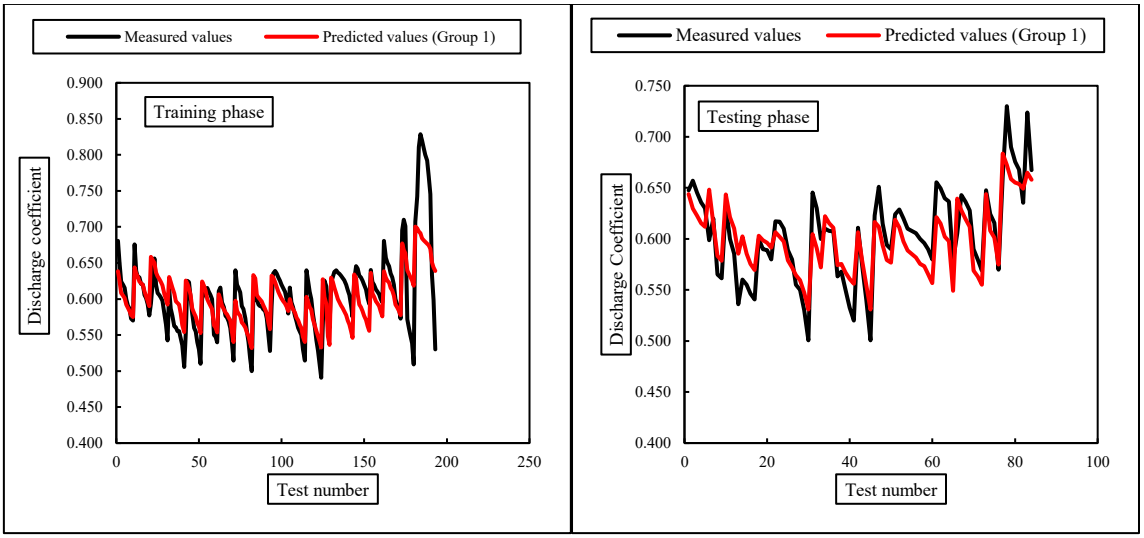


**Figure 3-Alignment between observed and predicted data in the SVM model**

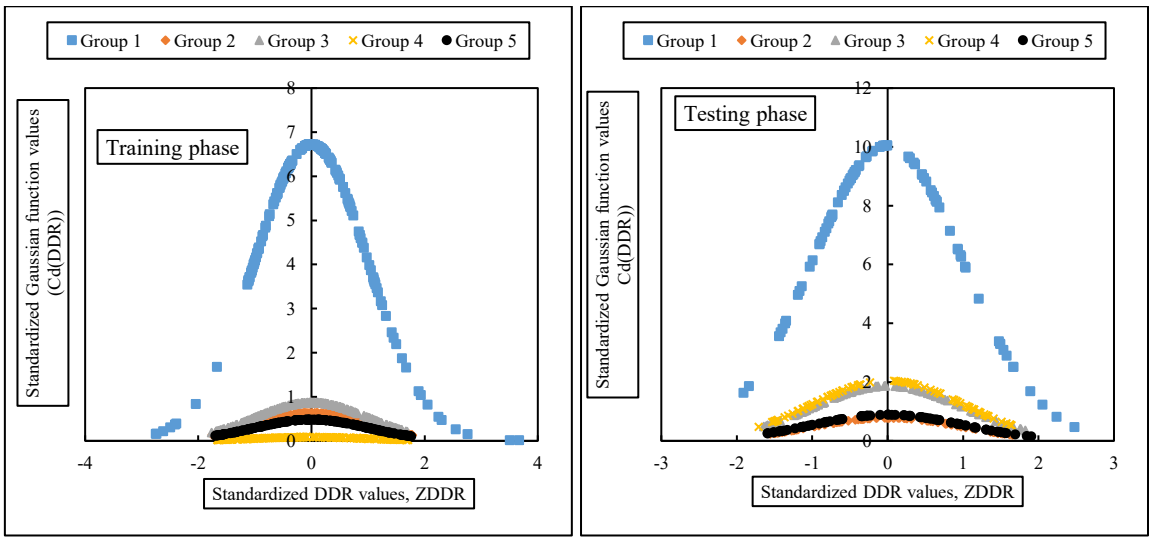


**Figure 4-Distribution of DDR index for the SVM model**

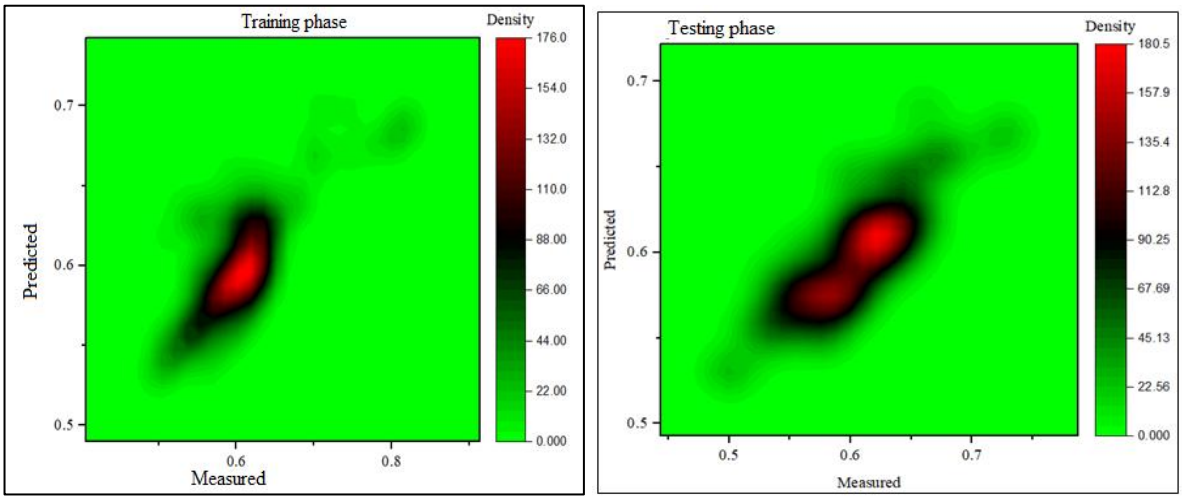


**Figure 5- 2D Kernel Density Estimation of observed and predicted $C_d$ in SVM Model**

***The outcomes of GEP***

The GEP was employed to derive an explicit equation for the discharge coefficient. Various configurations of chromosomes, head size, and gene numbers were evaluated (Table 6).

**Table 6- Samples of tuning parameters of GEP model**

| | Chromosome number | Head size | Gene number | Combination |
|---|---|---|---|---|
| Group 1 | 15 | 8 | 3 | + |
| Group 2 | 23 | 16 | 5 | + |
| Group 3 | 30 | 21 | 5 | × |
| Group 4 | 42 | 30 | 6 | × |
| Group 5 | 50 | 32 | 6 | + |

While highly complex structures improved training accuracy, they led to overfitting and overly complex equations. The optimal GEP structure (Group 1) comprised 15 chromosomes, a head size of 8, and 3 genes linked by addition (+), balancing accuracy and physical interpretability (Table 7).

**Table 7- The values of assessment metrics of GEP of different values of tuning parameters**

| | Training phase | | | | Testing phase | | | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ |
| Group 1 | 0.0299 | 0.0220 | 0.6864 | 7.9320 | 0.0267 | 0.0206 | 0.7642 | 8.8892 |
| Group 2 | 0.4907 | 0.4054 | 0.0180 | 0.5884 | 0.2333 | 0.2001 | 0.0650 | 1.0525 |
| Group 3 | 0.1815 | 0.1547 | 0.1127 | 1.4552 | 0.1287 | 0.1085 | 0.2940 | 1.9046 |
| Group 4 | 0.1617 | 0.1391 | 0.1111 | 1.5440 | 0.1141 | 0.0959 | 0.1109 | 2.3789 |
| Group 5 | 0.1833 | 0.1523 | 0.0820 | 1.4193 | 0.1265 | 0.1006 | 0.1126 | 2.1187 |

Figs. 6, 7, and 8 present the performance of the GEP model. The high overlap between observed and predicted values (Fig. 6) demonstrates the model's capacity to capture the physical trends. The $Z_{DDR}$ curves (Fig. 7) are centered at $Z_{DDR}$=0, with the Group 1 structure exhibiting a sharp, high peak that reflects a high concentration of low-error predictions. The

alignment of the KDE density cores with the 1:1 line (Fig. 8) confirms low prediction scatter and robust generalization in both training and testing phases.

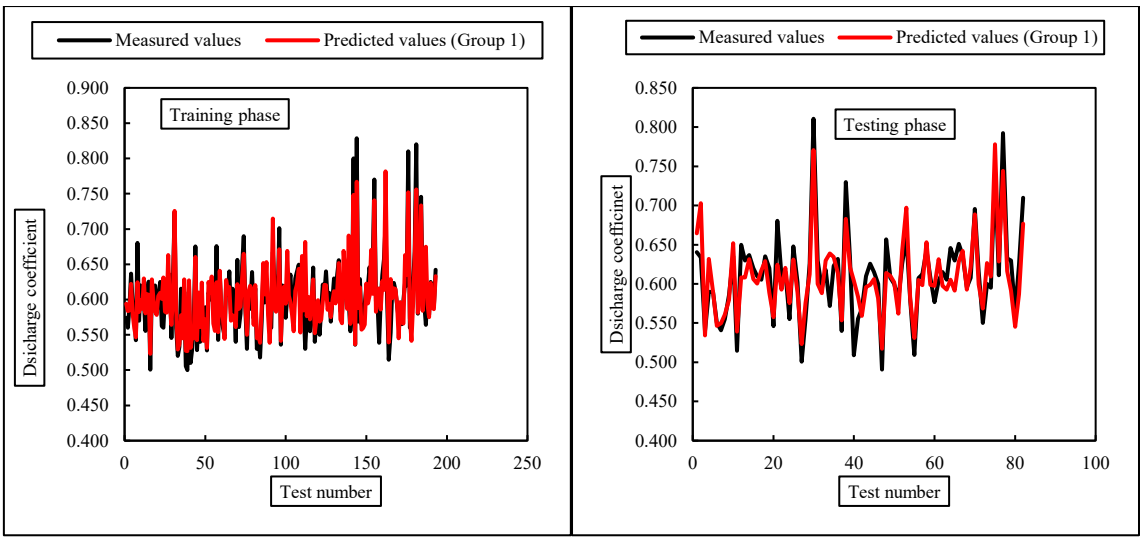


**Figure 6- Alignment between Observed and Predicted Data in the GEP Model**

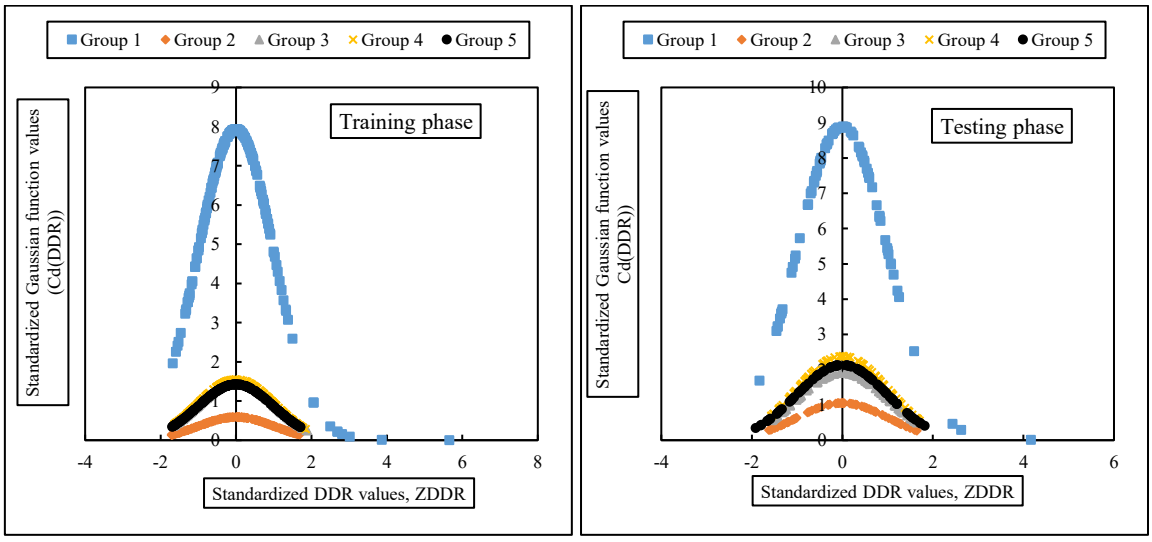


**Figure 7- Distribution of DDR index for the GEP model**

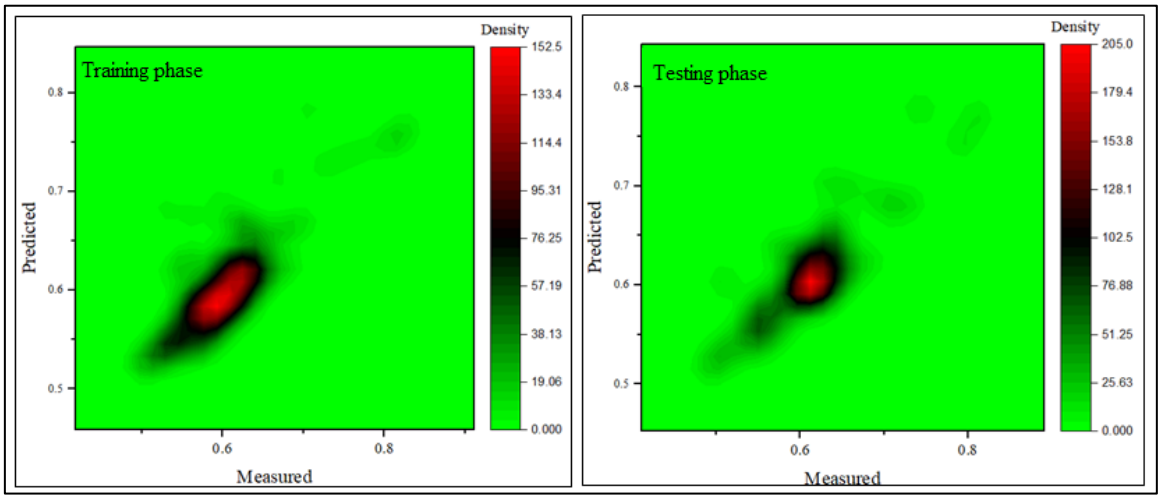


**Figure 8- 2D Kernel Density Estimation of observed and predicted $C_d$ in GEP Model**

The GEP tree structure (Fig. 9) represents the final formulation using three independent genes (Sub-ET1 to Sub-ET3), where $\frac{H_0}{P}$=d0, $\frac{A}{t}$=d1, φ=d2, θ=d3 and $\frac{L_C}{W}$=d4. The constants (c0,c1) for Genes 1 to 3 are (6.4967,5.7701), (−3.2491,9.1468), and (6.6322,8.5803), respectively.

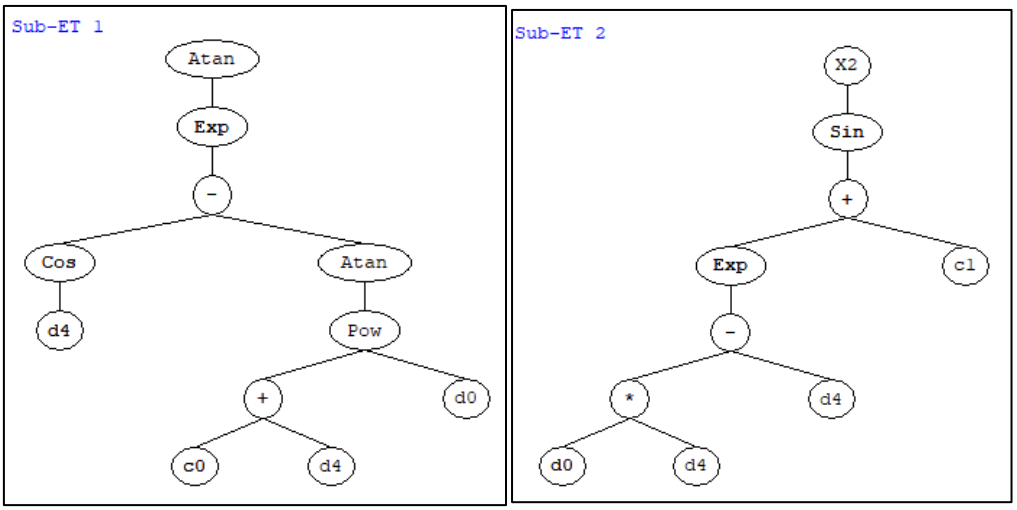


**Figure 9- Expression tree for the GEP output**

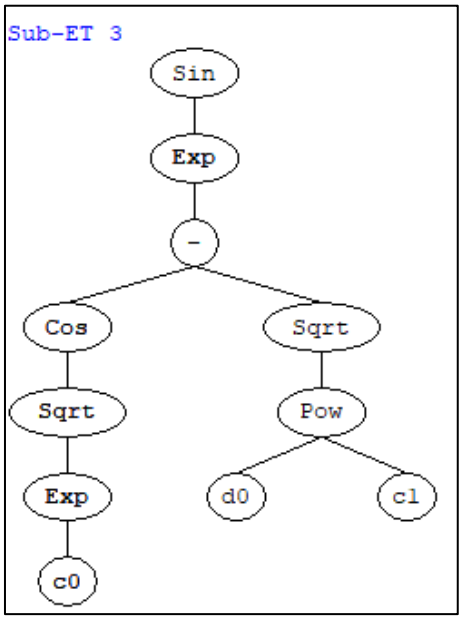


**Figure 10 (continued)**

***ANN results***

An ANN based on a Multilayer Perceptron (MLP) was developed, testing various hidden nodes, training algorithms, and activation functions. To ensure robust model training and avoid bias, the total dataset (277 data points) was randomly partitioned using a pseudo-random split into 70% for training (193 data points), 15% for testing (42 data points), and 15% for validation (42 data points). To overcome the network's ts and avoid local minima, each configuration was executed multiple times with different random seeds. Selected configurations are summarized in Table 8. All selected MLP networks were trained using the quasi-Newton BFGS algorithm, achieving fast convergence. Table 9 presents the detailed evaluation metrics of these configurations during the training, testing, and validation phases. Among the configurations, MLP 5-9-1 achieved the highest Ralidation), while MLP 5-3-1 variants yielded minimum errors, indicating that both compact and larger networks can successfully capture the nonlinear hydraulic patterns.

**Table 8- Selected ANN Models with Satisfactory Performance**

| No. | Network name | $R^2$ | | | Error | | | Training algorithm | Error function | Activation function | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Training | Testing | Verifying | Training | Testing | Verifying | | | Hidden layer | Output layer |
| **1** | MLP 5-9-1 | 0.9166 | 0.9041 | 0.9466 | 0.00024 | 0.00019 | 0.00020 | BFGS 50 | SOS | Logistic | Identity |
| **2** | MLP 5-3-1 | 0.9189 | 0.8914 | 0.9323 | 0.00023 | 0.00020 | 0.00024 | BFGS 79 | | Exponential | Identity |
| **3** | MLP 5-3-1 | 0.9276 | 0.8963 | 0.9372 | 0.00021 | 0.000190 | 0.00021 | BFGS 47 | | Tanh | Tanh |
| **4** | MLP 5-10-1 | 0.9308 | 0.9111 | 0.9390 | 0.00020 | 0.000160 | 0.00021 | BFGS 23 | | Exponential | Tanh |
| **5** | MLP 5-3-1 | 0.9293 | 0.9133 | 0.9426 | 0.00021 | 0.000160 | 0.00020 | BFGS 296 | | Logistic | Logistic |

**Table 9-The values of assessment metrics of MLP of different values of tuning parameters**

| | | Training phase | | | | Testing phase | | | | Verifying phase | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **RMSE** | MAE | $R^2$ | $C_{d(DDR(max))}$ | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ |
| **MLP** | **5-9-1** | 0.0220 | 0.0145 | 0.8402 | 11.4632 | 0.0135 | 0.0074 | 0.8174 | 12.2179 | 0.0199 | 0.0145 | 0.8961 | 12.8054 |
| **MLP** | **5-3-1** | 0.0217 | 0.0149 | 0.8444 | 11.5453 | 0.1663 | 0.1160 | 0.9501 | 9.7419 | 0.0218 | 0.0160 | 0.8692 | 11.0836 |
| **MLP** | **5-3-1** | 0.0205 | 0.0145 | 0.8605 | 12.1515 | 0.1668 | 0.1164 | 0.9493 | 9.1240 | 0.0207 | 0.0163 | 0.8782 | 11.4885 |
| **MLP** | **5-10-1** | 0.0201 | 0.0140 | 0.8664 | 12.5235 | 0.1669 | 0.1165 | 0.9545 | 9.2780 | 0.0207 | 0.0163 | 0.8818 | 11.8614 |
| **MLP** | **5-3-1** | 0.0203 | 0.0145 | 0.8637 | 12.2857 | 0.1670 | 0.1166 | 0.9598 | 9.6387 | 0.0202 | 0.0162 | 0.8885 | 11.7358 |

The predictive quality of the ANN is visualized in Figs. 10, 11, and 12. The MLP 5-9-1 predictions show tight agreement with the experimental data across all three phases (Fig. 10). The $Z_{DDR}$ distributions (Fig. 11) are centered around zero, confirming unbiased predictions, while the KDE maps (Fig. 12) show a narrow, elongated high-density core along the ideal 1:1 line, verifying the absence of overfitting.

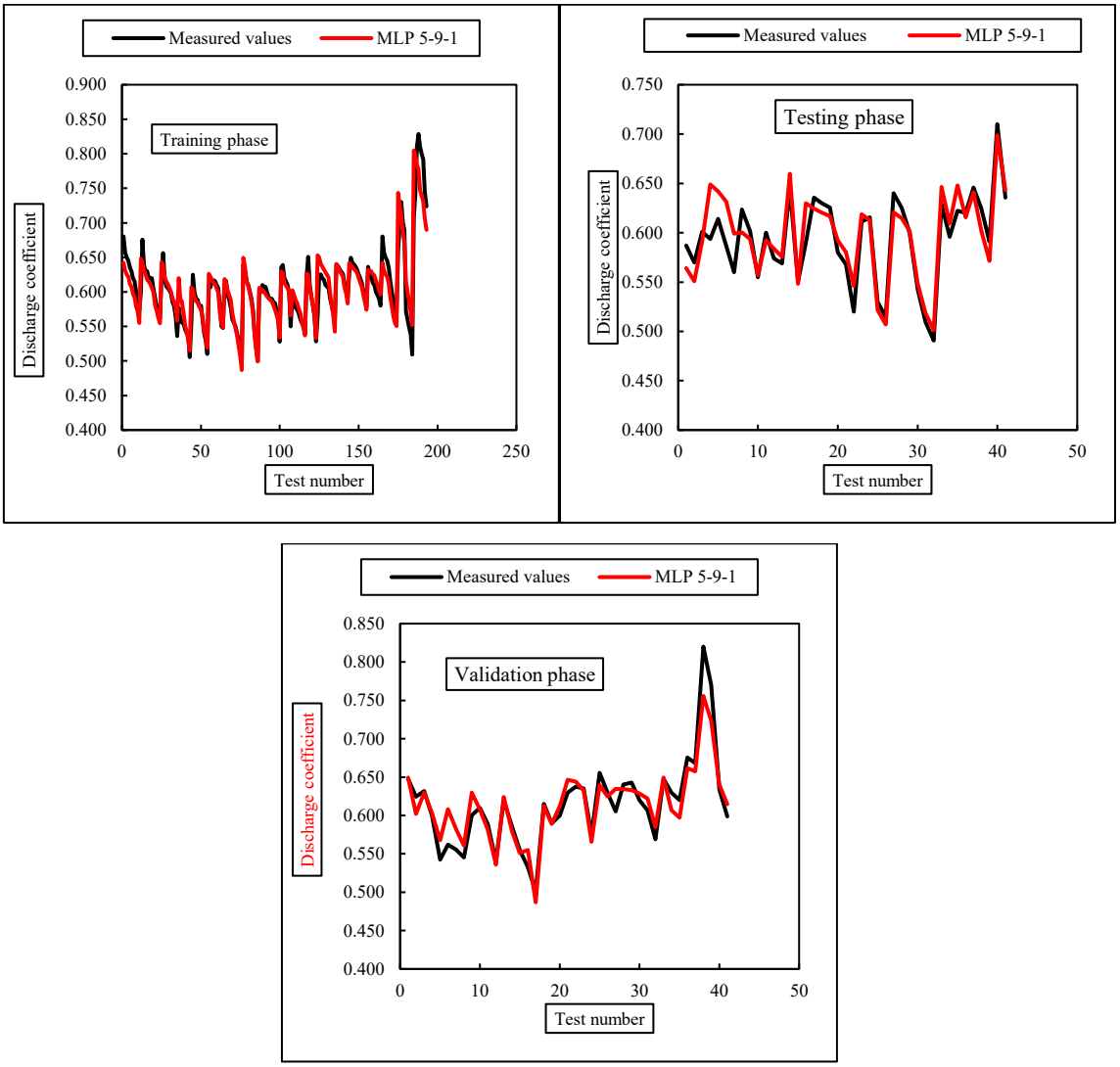


**Figure 11- Alignment between Observed and Predicted Data in the ANN Model**

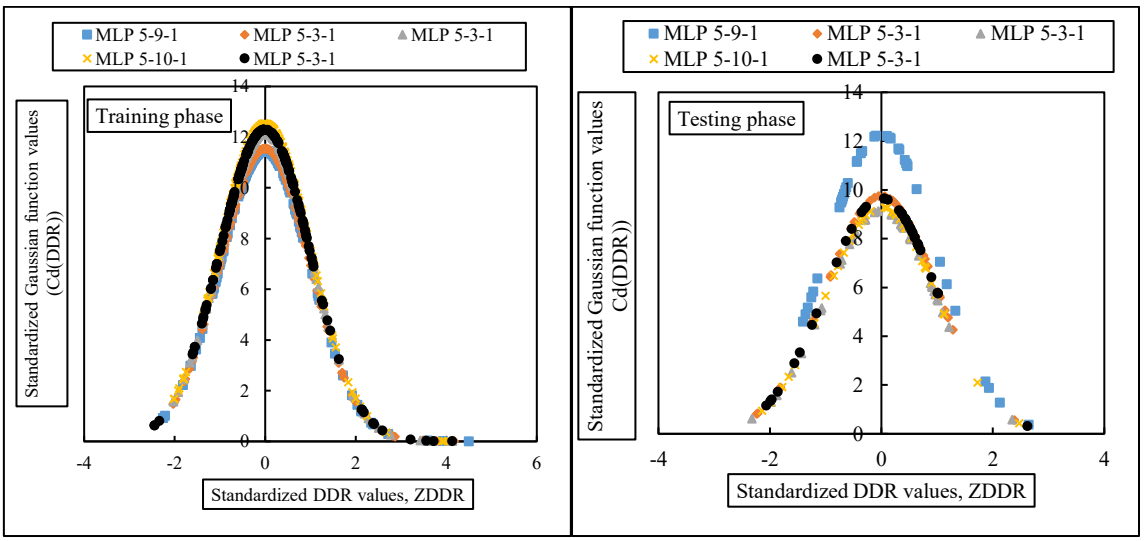


**Figure 12- Distribution of DDR index for the MLP model**

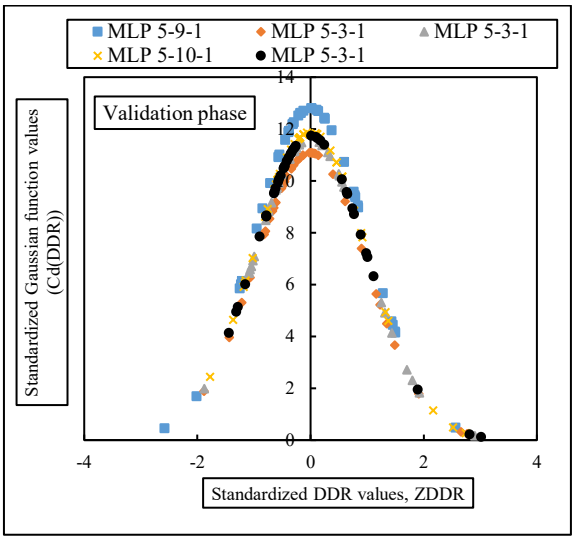


**Figure 13 (continued)**

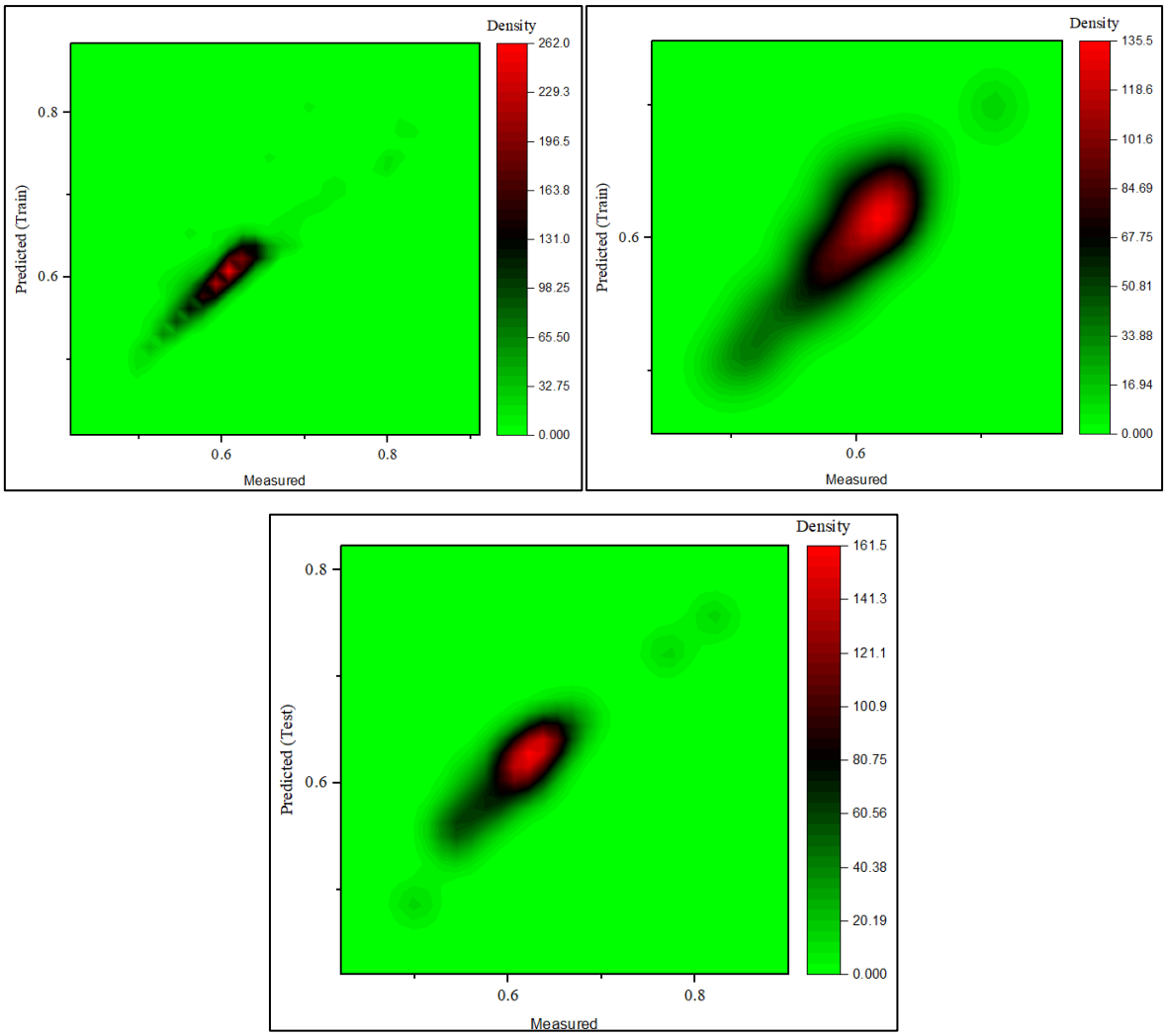


**Figure 14-2D Kernel Density Estimation of observed and predicted $C_d$ in MLP Model**

Table 10 compares the performance of the optimized SVM, GEP, and MLP models.

**Table 10- Comparison of model outputs**

| | Training phase | | | | Testing phase | | | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ | RMSE | MAE | $R^2$ | $C_{d(DDR(max))}$ |
| SVM | 0.0379 | 0.0271 | 0.6024 | 6.7222 | 0.0242 | 0.0207 | 0.7281 | 10.0475 |
| GEP | 0.0299 | 0.0220 | 0.6864 | 7.9320 | 0.0267 | 0.0206 | 0.7642 | 8.8892 |
| MLP | 0.0220 | 0.0145 | 0.8402 | 11.4632 | 0.0199 | 0.0145 | 0.8961 | 12.8054 |

The MLP model outperformed the other approaches, achieving the lowest errors (RMSE=0.0220, MAE=0.0145 in training; RMSE=0.0199, MAE=0.0145 in testing) and the highest goodness of fit ($R^2$=0.8402 training, $R^2$=0.8961 testing). Furthermore, the DDR index ($C_{d(DDRmax)}$) for MLP was the highest (11.4632 in training, 12.8054 in testing), compared to SVM (6.7222, 10.0475) and GEP (7.9320, 8.8892).

This superiority is visually confirmed by the scatter plots (Fig. 13) and the boxplots (Fig. 14). The MLP predictions cluster tightly along the 1:1 trend line, and the boxplots demonstrate a near-identical distribution of predicted and measured $C_d$ values in terms of median, interquartile range, and overall spread, with minor deviations occurring only at extremely high values.

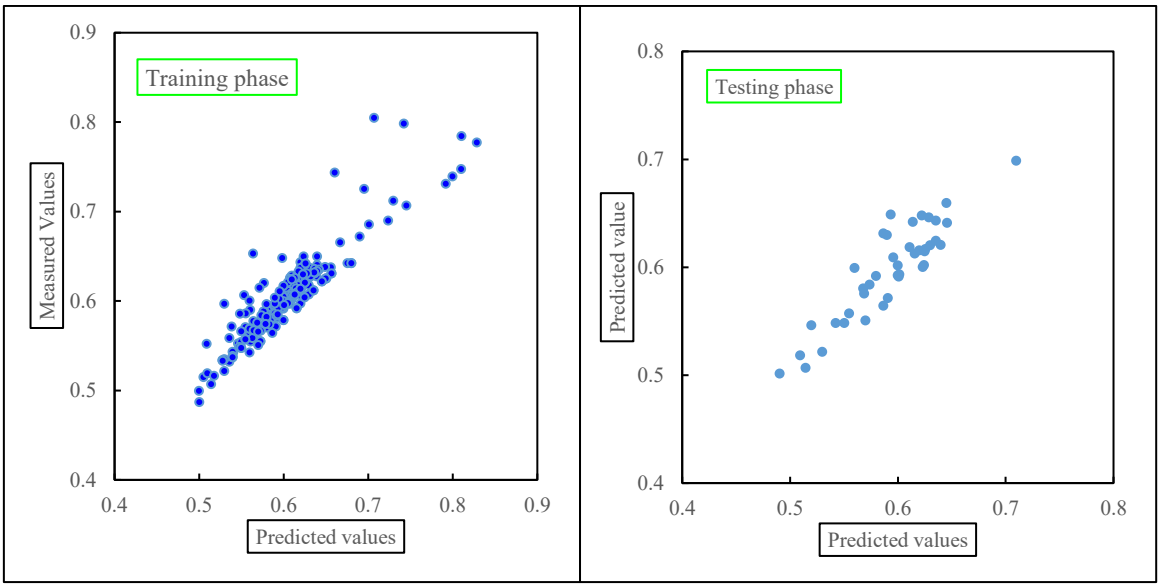


**Figure 15-Scatter plot of observed and predicted data for the MLP model**

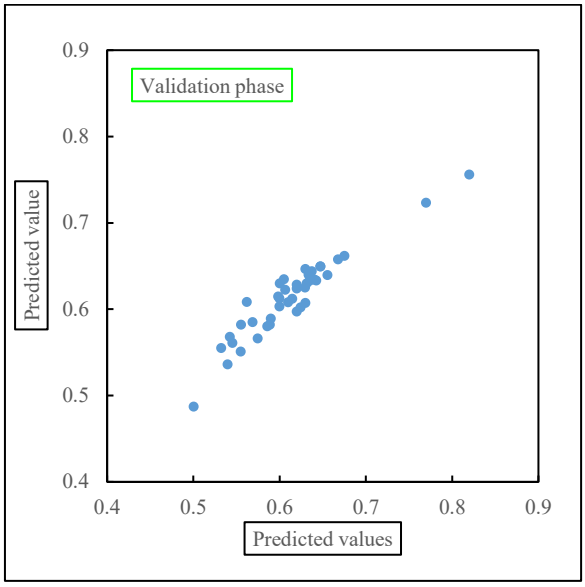


**Figure 16 (continued)**

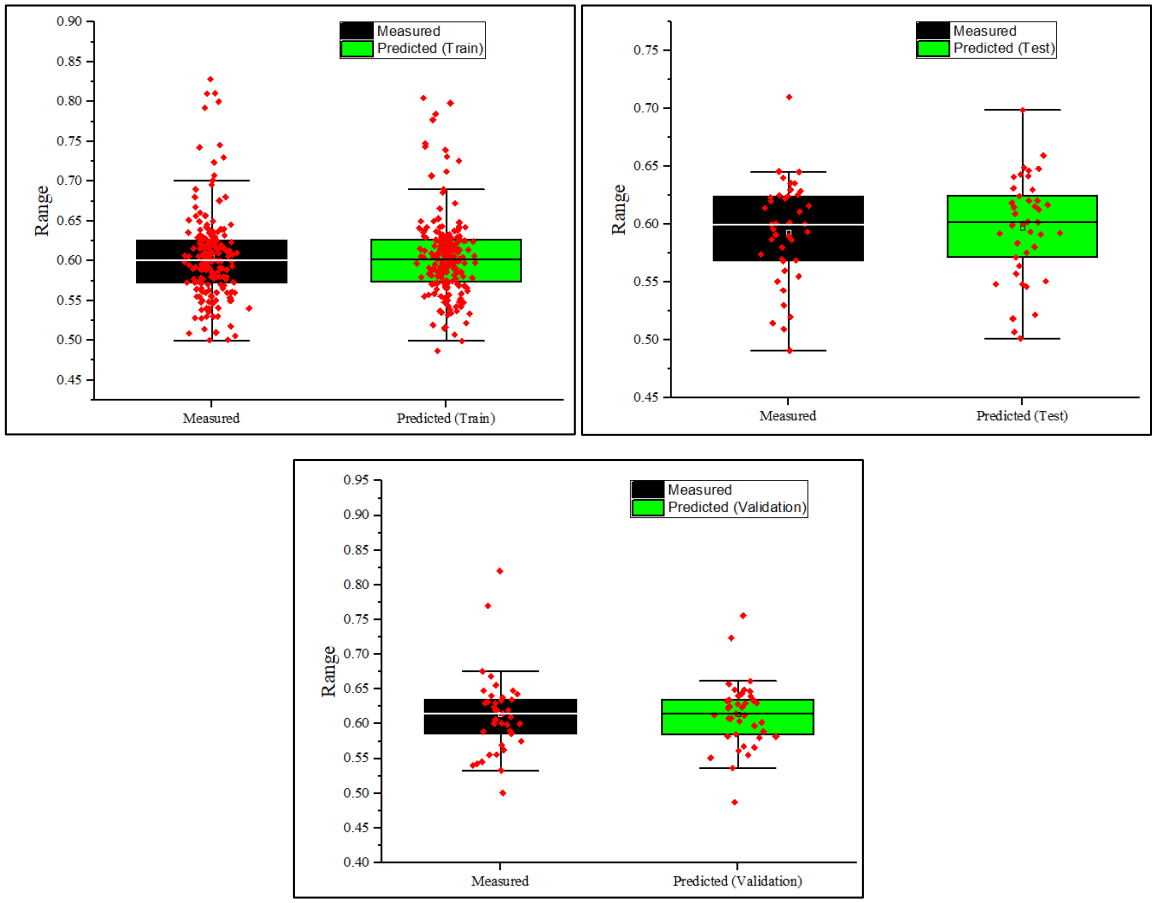


**Figure 17- Overlaid curves of observed and predicted Cd values for the ANN model**

**Conclusions**

This study compared the predictive performance of three data-driven approaches, namely SVM/SVR, GEP, and ANN/MLP, for estimating $C_d$ of an arced labyrinth weir. The models were developed using 277 experimental observations collected from the Hydraulic Laboratory of the University of Maragheh, with 193 samples used for training and 84 samples for testing. The problem was formulated using five input variables related to the hydraulic and geometric characteristics of the weir, and the final results demonstrated that the ANN/MLP framework provided the most accurate and reliable estimation of $C_d$ among the investigated methods. In the SVR model, the optimal RBF kernel parameters were identified as C=10, ε=0.25, and γ=0.01, indicating a balance between model flexibility and generalization capability. For the ANN approach, several multilayer perceptron architectures were evaluated using the quasi-Newton BFGS training algorithm, and the best results were obtained from the MLP 5-3-1 and MLP 5-10-1 configurations. In particular, MLP 5-3-1 achieved a testing $R^2$ of 0.9133 and the lowest reported error of 0.00016, while MLP 5-10-1 produced the highest training $R^2$ of 0.9308 and a testing $R^2$ of 0.9111. The study also reported that all selected ANN structures achieved coefficients of determination greater than 0.89 across the training, testing, and validation stages, with validation errors decreasing to as low as 0.0002, which clearly indicates strong learning ability and stable generalization.

The final comparative analysis, summarized in Table 10, confirmed the dominance of the ANN/MLP model over both GEP and SVM when all four evaluation indices were considered simultaneously. In the training phase, MLP achieved the lowest RMSE of 0.0220 and the lowest MAE of 0.0145, while also recording the highest $R^2$ of 0.8402 and the highest $C_{d(DDR(max))}$value of 11.4632. In the testing phase, the superiority of MLP remained evident, with RMSE=0.0199, MAE=0.0145, $R^2$=0.8961, and $C_{d(DDR(max))}$=12.8054. By comparison, GEP yielded intermediate performance with training and testing RMSE values of 0.0299 and 0.0267, respectively, while SVM showed comparatively weaker results, particularly in the training stage where its RMSE and MAE were 0.0379 and 0.0271. Although SVM provided a useful nonlinear regression framework and GEP offered the advantage of an explicit symbolic expression that improves interpretability, neither method could match the predictive accuracy and robustness of ANN. The reported ranking of MLP > GEP > SVM reflects this clear difference in performance and confirms that the ANN-based model is the most effective tool for capturing the complex nonlinear relationship between the governing variables and the discharge coefficient of arced labyrinth weirs.

Overall, the findings of this research demonstrate that machine learning methods can successfully model the hydraulics of arced labyrinth weirs within the investigated experimental ranges**,** and that the ANN/MLP approach provided the best performance among the examined models for the present dataset**.** Its lower error values, higher coefficients of determination, and stronger generalization capability for the studied conditions make it a promising approach for estimating Cd within the tested ranges of hydraulic and geometric parameters. Therefore, the proposed ANN model can serve as a useful complementary tool to traditional empirical

relationships for the configurations investigated here**,** offering improved accuracy for hydraulic analysis and design optimization of arced labyrinth weirs with sloped sidewalls**.** Caution should be exercised when applying the developed models outside the investigated ranges, and further validation with independent datasets is recommended.

The present study has several limitations. The experimental data were collected from a single laboratory source, which may restrict the generalizability of the proposed models beyond the investigated conditions. In addition, the analysis was limited to arced labyrinth weirs with sloped sidewalls and did not cover other labyrinth geometries or wider ranges of hydraulic and geometric variables. Finally, the model comparison was confined to ANN, SVM, and GEP. Therefore, future studies should consider larger and more diverse datasets, cross-laboratory validation, and additional geometric configurations to improve the robustness and applicability of the findings.

While the developed MLP, SVM, and GEP models demonstrated high accuracy in predicting the discharge coefficient, it is worth noting that other advanced soft computing algorithms, such as ensemble tree-based models (e.g., XGBoost, Random Forest, CatBoost) and Gaussian Process Regression (GPR), could offer alternative benefits in handling non-linear hydraulic relationships. Although implementing these models was beyond the scope of the present study, exploring their predictive capacities represents a promising direction for future research.

### *Author Contributions*

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

### *Data Availability Statement*

Data available on request from the authors.

### *Funding*

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### *Conflict of Interest*

The authors declare no conflict of interest.

## References

Ahmadzadeh-Kaleybar,F., Shahmohammadi Kalalagh,S., Fard Moradinia,S ., 2024. Simulating of Changes in Water Distribution Uniformity Coefficient in Classic Stationary Sprinkler Irrigation Using Data-Mining Models. *Journal of New Approaches in Water Engineering and Environment*, *3*(2), 118-135. doi: 10.22034/nawee.2024.471274.1100

Aichouri,I, Bougherira,N, Hani,A., Lallahem,S., 2015. River flow model using artificial neural networks peer-review under responsibility of the Euro-Mediterranean Institute for Sustainable Development (EUMISD). *Energy Procedia*, 74, 1007-1014.

Alfatlawi,TJ, Hashem, T., Hasan, ZH . , 2023. Discharge coefficient of symmetrical stepped and triangular labyrinth side weirs in a subcritical flow regime. *Journal of Irrigation and Drainage Engineering*, 149(4), 04023004. https://doi.org/10.1061/jidedh.ireng-9627

Asgharzadeh-Bonab,A, Bijanvand, S, Parsaie, A., Afaridegan,E ., 2025. Machine learning-based estimation of discharge coefficient for semicircular labyrinth weirs. *Scientific Reports*, 15(1), 33002. https://doi.org/10.1038/s41598-025-18230-4

Aydın,M., Ulu, A., Işık,E ., 2024. Determination of effective flow behaviors on discharge performance of trapezoidal labyrinth weirs using numerical and physical models. *Modeling Earth Systems and Environment*, 10, 3763 - 3776. https://doi.org/10.1007/s40808-024-01996-3

Azma, A., Borthwick, AG., Ahmadian, R., Liu, Y., Zhang, D., 2025. Modeling the discharge coefficient of labyrinth sluice gates using hybrid support vector regression and metaheuristic algorithms. *Physics of Fluids*, 37(4). https://doi.org/10.1063/5.0260738

Belaabed, F., Arabet, L ., 2024. Prediction of the Discharge Coefficient of a Labyrinth Weir Type D by an Artificial Neural Network Method. *Archives of Hydro-Engineering and Environmental Mechanics*, 71, 59 - 72. https://doi.org/10.2478/heem-2024-0004

Crookston, B., Tullis, B ., 2012. Arced Labyrinth Weirs. *Journal of Hydraulic Engineering*, 138, 555-562. https://doi.org/10.1061/(asce)hy.1943-7900.0000553

Crookston, B., Tullis, B., 2013. Hydraulic Design and Analysis of Labyrinth Weirs. I: Discharge Relationships. *Journal of Irrigation and Drainage Engineering-ASCE*, 139, 363-370. https://doi.org/10.1061/(asce)ir.1943-4774.0000558

Dutta, D., Mandal, A., Afzal, MS., 2020. Discharge performance of plan view of multi-cycle W-form and circular arc labyrinth weir using machine learning. *Flow Measurement and Instrumentation*, 73, 101740. https://doi.org/10.1016/j.flowmeasinst.2020.101740

Emami, S., Arvanaghi, H., Parsa, J., 2018. Numerical Investigation of Geometric Parameters Effect of the Labyrinth Weir on the Discharge Coefficient. *Journal of Rehabilitation in Civil Engineering*, 6, 1-9. https://doi.org/10.22075/jrce.2017.11428.1188

Emami, S., Emami, H., Parsa, J., 2023. LXGB: a machine learning algorithm for estimating the discharge coefficient of pseudo-cosine labyrinth weir. *Scientific Reports*, 13. https://doi.org/10.1038/s41598-023-39272-6

Emiroglu, M., Kaya, N., 2010. Discharge Coefficient for Trapezoidal Labyrinth Side Weir in Subcritical Flow. *Water Resources Management*, 25, 1037-1058. https://doi.org/10.1007/s11269-010-9740-7

Fuladipanah, M., Hazi, M.A., Kisi, O., 2023. An in-depth comparative analysis of data-driven and classic regression models for scour depth prediction around cylindrical bridge piers. *Applied Water Science*, 13(12), 231. https://doi.org/10.1007/s13201-023-02022-0

Ghasemvand, R., Heidarnejad, M., Masjedi, A., Bordbar, A., 2024. Experimental study of hydraulic losses in linear and nonlinear weirs: A comparative analysis. *Water Science*, 38, 378 - 390. https://doi.org/10.1080/23570008.2024.2373954

Hashem, T., Mohammed, A.Y., Sharifi, A., 2025. Machine learning-based modeling of discharge coefficients in labyrinth sluice gates. *Flow Measurement and Instrumentation*, 102, 102823. https://doi.org/10.1016/j.flowmeasinst.2025.102823

Hussain, R.A., Hassan, S.A., Jamel, A.A.J ., 2022. Experimental study on flow over triangular labyrinth weirs. *Int J Des Nat Ecodynamics*, 17(2), 249-255. https://doi.org/10.18280/ijdne.170211

Idrees, A.K., Al-Ameri,R.,2022. A review of hydraulic performance and design methods of labyrinth weirs. *Water Supply*, 22(11), 8120-8138. https://doi.org/10.2166/ws.2022.378

Karami, H., Karimi, S., Bonakdari, H., Shamshirband, S., 2016. Predicting discharge coefficient of triangular labyrinth weir using extreme learning machine, artificial neural network and genetic programming. *Neural Computing and Applications*, 29, 983 - 989. https://doi.org/10.1007/s00521-016-2588-x

Kisi, O., Azamathulla, H.M., Cevat, F., Kulls, C., Kuhdaragh, M., Fuladipanah,M., 2024. Enhancing river flow predictions: comparative analysis of machine learning approaches in modeling stage-discharge relationship, *Results Eng.*, 22(2024), 102017. https://doi.org/10.1016/j.rineng.2024.102017

Leunberger, D ., 1984. Linear and nonlinear programming, Amazon.

Majedi-Asl, M., Fuladipanah, M., Arun, V., Tripathi, R.P., 2022. Using data mining methods to improve discharge coefficient prediction in Piano Key and Labyrinth weirs. *Water supply*, 22(2), 1964-1982. https://doi.org/10.2166/ws.2021.304

Majedi-Asl, M., Ghaderi, A., Kouhdaragh, M., Alavian, T.O., 2024. A performance comparison of the meta model methods for discharge coefficient prediction of labyrinth weirs. *Flow Measurement and Instrumentation*, 96, 102563.. https://doi.org/10.1016/j.flowmeasinst.2024.102563

Masoudi, M., Yari, A., Sadeghian, J., Norouzi, H., 2024. Experimental investigation of the discharge coefficient of the rectangular and trapezoidal labyrinth weirs considering variable congress lengths. *Modeling Earth Systems and Environment*, 10, 2819-2832. https://doi.org/10.1007/s40808-023-01925-w

Mattos-Villarroel, E., Flores-Velázquez, J., Ojeda-Bustamante, W., Díaz-Delgado, C., Salinas-Tapia, H ., 2021. Influence of crest geometric on discharge coefficient efficiency of labyrinth weirs. *Flow Measurement and Instrumentation*, 81, 102031. https://doi.org/10.1016/j.flowmeasinst.2021.102031

Mattos-Villarroel, E.D., Ojeda-Bustamante, W., Díaz-Delgado, C., Salinas-Tapia, H., Bautista-Capetillo, C.F., Flores-Velázquez, J., Aguilar-Rodríguez, C.E ., 2025. Influence of the Semicircular Cycle in a Labyrinth Weir on the Discharge Coefficient. *Water*, 17(21), 3151. https://doi.org/10.3390/w17213151

Mattos-Villarroel, E.D., Ojeda-Bustamante, W., Díaz-Delgado, C., Salinas-Tapia, H., Flores-Velázquez, J., Bautista Capetillo, C ., 2023. Methodological proposal for the hydraulic design of labyrinth weirs. *Water*, 15(4), 722. https://doi.org/10.3390/w15040722

Namazi, FSA., Mozaffari, J ., 2023. Investigation of labyrinth weirs discharge coefficient with the same length. *Flow Measurement and Instrumentation*, 94, 102468. https://doi.org/10.1016/j.flowmeasinst.2023.102468

Noori, R., Khakpour, A., Omidvar, B., Farokhnia, A ., 2010. Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert systems with applications*, 37(8), 5856-5862. 10.1016/j.eswa.2010.02.020

Norouzi, R., Daneshfaraz, R., Ghaderi, A., 2019. Investigation of discharge coefficient of trapezoidal labyrinth weirs using artificial neural networks and support vector machines. *Applied Water Science*, 9(7), 148. https://doi.org/10.1007/s13201-019-1026-5

Roushangar, K., Alami, M., Shiri, J., Majedi-Asl, M ., 2018. Determining discharge coefficient of labyrinth and arced labyrinth weirs using support vector machine. *Hydrology Research*, 49, 924-938. https://doi.org/10.2166/nh.2017.214

Seyedian, S.M., Haghiabi, A., Parsaie, A ., 2023. Reliable prediction of the discharge coefficient of triangular labyrinth weir based on soft computing techniques. *Flow Measurement and Instrumentation*, 92, 102403. https://doi.org/10.1016/j.flowmeasinst.2023.102403

Shafiei, S, Najarchi, M., Shabanlou, S ., 2020. Prognostication of discharge coefficient of labyrinth weir using outlier robust extreme learning machine. *Flow Measurement and Instrumentation*, 71, 101681. https://doi.org/10.1016/j.flowmeasinst.2019.101681

Shakeri yousefi, S., Najarchi, M., Fuladipanah, M., Rabani Bidgoli, M ., 2025. Machine Learning models for High-Accuracy Prediction of Energy Dissipation Through Gabion Sills Downstream of Spillways. *Journal of New Approaches in Water Engineering and Environment*, *4*(2), 91-106. doi: 10.22034/nawee.2025.507107.1140

Smola,AJ ., 1996. Regression estimation with support vector learning machines, Master's Thesis, Technische Universität München, Germany.

Tabatabai, S.M., Heidranejad, M., Masjedi, A., Bordbar, A ., 2026. Tecnología y ciencias del agua. *Tecnología y ciencias del agua*, 17(1), 314-348. https://doi.org/10.24850/j-tyca-17-1-5

Tamizi, H. R. , Vaghefi, M., Farzin, S., Mahdi, V. A ., 2026. Modeling the wastewater treatment plant using artificial neural networks-based modeling (Case study: Bushehr city, Iran). *Journal of New Approaches in Water Engineering and Environment*, *5*(2), 64-83. doi: 10.22034/nawee.2025.555176.1174

Vapnik,VN .,1995. The Nature of Statistical Learning Theory, Springer-Verlag, New York.

Zare, H., Vaghefi, M., Mahmoudi, A., Behroozi, A., 2023. Experimental Exploration of Flow Hydraulics and Discharge Coefficient for an Inclined Circular Labyrinth Weir. *Water Resources Management*, 37, 4521-4536. https://doi.org/10.1007/s11269-023-03577-y

s