Gonbad Kavous University

Journal of New Approaches in Water Engineering and Environment

Online ISSN: 2981-0280

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

Machine Learning models for High-Accuracy Prediction of Energy Dissipation Through Gabion Sills Downstream of Spillways

Shahram Shakeri yousefi[1], Mohsen Najarchi [2*], Mehdi Fuladipanah[3], Mahmood Rabani Bidgoli[4]

[1] Department of Civil Engineering, Ar. C., Islamic Azad University, Arak, Iran.
[2] Department of Civil Engineering, Ar. C., Islamic Azad University, Arak, Iran.
[3]Department of Civil Engineering Ramh. C., Islamic AzadUniversity,Ramhormoz, Iran.
[4] Department of Civil Engineering, Jasb. C., Islamic Azad University, Jasb, Iran.

Article Info

Article type:
Research Article

Article history:
Received 22 January 2025
Revised 14 February 2025
Accepted 17 February 2025
Published online 14 June 2025

Keywords:
Artificial Intelligence
Energy Loss
Sensitivity Analysis
Performance Assessment

ABSTRACT

Objective: the objective of this paper is to develop and compare three MLMs, SVR, GEP and ANN- for the high-accuracy prediction of energy dissipation downstream of gabion sills in spillways. Through dimensional analysis and sensitivity evaluation using the Γ-test, the most influential hydraulic and geometric parameters are identified. The performance of each model is rigorously assessed using statistical metrics to determine their predictive reliability and accuracy, with the aim of identifying the most effective computational approach for optimizing energy dissipation in gabion-structured spillway systems.

Material and Methods: experimental data from a lab flume was used. Dimensional analysis identified key parameters affecting energy dissipation. Sensitivity analysis via the Gamma Test selected the most influential inputs. These were used to train and compare three machine learning models: SVR, GEP, and ANN.

Results and Discussion: the GEP model demonstrated superior performance, achieving the highest R^2 (0.936) and lowest errors (RMSE=0.003) in predicting energy dissipation. It outperformed both ANN and SVR. Sensitivity analysis identified four key hydraulic parameters as the most influential inputs.

Conclusions: the study conclusively found the Gene Expression Programming (GEP) model to be the most accurate and reliable for predicting energy dissipation over gabion sills, significantly outperforming ANN and SVR. The four key hydraulic parameters identified were crucial for model success, demonstrating the effectiveness of this machine learning approach.

*Corresponding author, Email: mohsen.najarchi@iau.ac.ir

Cite this article: Last Name, Initial., Last Name, Initial., & Last Name, Initial. (2025). Title of paper in lower case letters (except for initial letter of first word, initial of first word after a colon, and proper nouns). *Journal of New Approaches in Water Engineering and Environment*, 4(2), 91-106. https://doi.org/ 10.22034/nawee.2025.507107.1140

Publisher: Gonbad Kavous University.
DOI: https://doi.org/ 10.22034/nawee.2025.507107.1140

1. Introduction

Spillways are hydraulic structures designed to convey excess water from the upstream to the downstream of rivers during flood conditions. The unique geometry of ogee spillways enhances their discharge capacity by aligning the flow nappe with the spillway surface. Consequently, implementing energy dissipation structures downstream of spillways is crucial to mitigate destructive energy. Modern energy dissipation techniques include using gabion structures, artificial roughness, blocks, and perforated plates along the flow path as alternative solutions to conventional stilling basins (Fuladipanah and Jafarinia, 2011; Iranpour et al., 2024).

Spillways are critical hydraulic structures designed to safely manage excess water during flood events by channeling flow from upstream to downstream. A key challenge in spillway design lies in mitigating the destructive energy of high-velocity flows, which necessitates effective energy dissipation mechanisms. Gabion structures—porous, rock-filled cages—have emerged as a sustainable and cost-effective solution, offering advantages such as enhanced energy dissipation, reduced erosion, and environmental adaptability compared to traditional stilling basins. Recent studies have increasingly employed machine learning (ML) models to optimize gabion performance, demonstrating success in predicting energy loss. However, existing research lacks a systematic comparison of advanced ML frameworks for gabion-sill applications, particularly in evaluating their interpretability, robustness, and sensitivity to hydraulic parameters. Rezi et al. (2023) found that in gabion step weirs, the downstream slope had minimal impact on energy loss, increased step numbers reduced energy dissipation, and the highest energy loss occurred with 10 mm and 40 mm particles. Mahjoubi and Kashefipour (2023) observed that gabion stepped spillways enhanced energy dissipation by up to 13% compared to non-gabion spillways and 72.1% over simple chutes, with maximum energy loss at four steps. Naseri and Kashefipour (2022) reported that gabion step spillways with 45% porosity achieved 86% maximum energy loss, 8% higher than impermeable models. Mobin et al. (2023) showed that the MARS model outperformed EPR for energy loss estimation in gabion weirs, achieving RMSE=0.054, MAPE=0.017, and CC=0.99. Daneshfaraz et al. (2022) found that gabions in vertical drops increased energy dissipation by 57% and reduced the Froude number range from 3.8–5.7 to 0.52–2.5, with porosity playing a significant role. Salehi et al. (2022) observed that larger gabion particles increased downstream scour depth, while in-passage flow reduced scour, and predictive models had ~17% error in flow rate estimation. Daneshfaraz et al. (2021) reported that gabion slope breakers enhanced energy dissipation by 561% and downstream depth by 50.1%, reducing bed erosion. Kheyraei et al. (2017) noted that steeper slopes in gabion spillways increased hydraulic head and that larger particles raised discharge coefficients, with a model correlation of 0.97. Pourhosein-Ghadi et al. (2023) similarly found minimal impact of downstream slope on energy loss but confirmed higher energy dissipation with 10 mm and 40 mm particles. Kheyraei and Fathi Moghaddam (2016) showed that steeper slopes decreased discharge coefficients, while larger materials increased them, developing a predictive equation with a 0.97 correlation. Srinivas and Tiwari (2022) concluded that DNN models outperformed others (ANFIS, BPNN, MVLR, MVNLR) for predicting oxygen aeration efficiency (OAE20) in gabion spillways (CC=0.9757), with porosity as the most influential parameter. Hussain et al. (2022) indicated that GMDH models outperformed GEP and traditional regressions for predicting inverse relative energy dissipation (IRED) over gabion weirs (R=0.979, RMSE=0.314), emphasizing the pre-jump Froude number as a critical variable. Alsubih et al. (2023) reviewed the sustainability, flexibility, and environmental benefits of gabion water barriers. Through numerical simulations, Zuhaira et al. (2021) showed that EPR models outperformed nonlinear regressions (NMR) for predicting the inception point in non-uniform gabion stepped spillways, achieving R^2=0.93 and MAE=1.66.

Based on the findings of prior research in artificial intelligence, the current study addresses the prediction of energy dissipation in a Gabion-structured stepped spillway with a water base using three machine learning models: SVR, GEP, and ANN. Sensitivity analysis was employed to identify parameters influencing energy dissipation, and these parameters were subsequently used to evaluate the outputs of the machine learning models through performance evaluation metrics.

2. Material and Methods

2.1. Experimental setup

This study was conducted in a specialized experimental channel at Maragheh University. The channel measured 1300 cm in length, 120 cm in width, and 80 cm in depth, featuring a metal floor and transparent 10-mm-thick glass walls to enable precise observation of fluid dynamics. Water was supplied to the system via a pump that transferred it from a storage reservoir into the channel. At the inlet section, two steel mesh screens and a floating foam layer (200 cm × 100 cm × 20 cm) were installed to minimize hydraulic turbulence. The experimental weir was constructed from stainless steel sheets with a width of 120 cm, a height of 50 cm, and a terminal curvature radius of 28 cm. It was designed to accommodate a maximum flow rate of 0.045 cubic meters per second (162 cubic meters per hour). The gabion system is comprised of steel mesh cages measuring 120 cm in width and 10 cm in length (along the flow direction), enclosed by 1-mm-thick metal mesh grids with 10-mm apertures. All experiments were performed under a horizontal channel configuration (zero longitudinal slope). The gabion structure was positioned 1.38 meters downstream from the main weir, a distance calculated to initiate a hydraulic jump phenomenon at the closest possible point to the weir's base. Figure 1 clearly illustrates the overall experimental setup, flow patterns around the structures, and the formation of turbulent flow zones upstream of the gabion. Supplementary tables provide detailed geometric and hydraulic parameters of the models. Schematic diagrams depict the spatial layout of the channel, the positioning of the semi-circular weir, and the gabion structure, along with visualizations of flow interactions with these components. Table 1 presents the values of measured parameters in the current study.

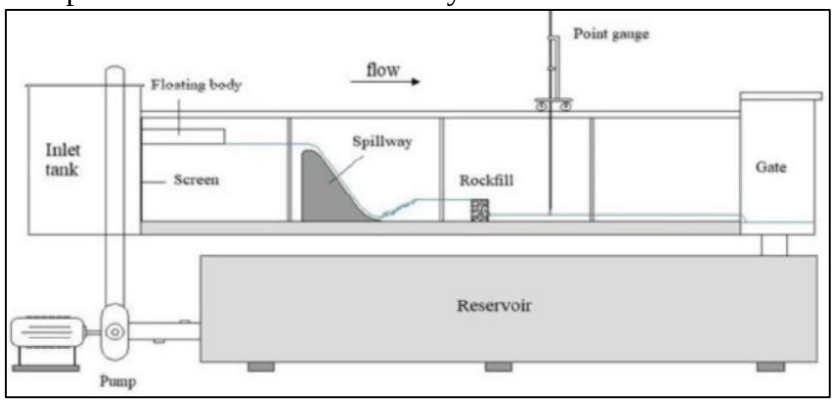

Figure 1- Schematic view of flume and equipment

Table 1- The values of hydraulic and physical parameters included in the present study

Parameter	Flow discharge	Gabion height	Porosity	Mean size of gabion gradation	Initial depth of hydraulic jump	Sequent depth of hydraulic jump	Flow depth after the gabion
Symbol	Q (lit/s)	h (mm)	ε (mm)	d (mm)	y1 (mm)	y2 (mm)	y3 (mm)
Range	20 to 40	50, 100	10	15, 22 and 30	11 to 475	124 to 485	60 to 98

2.2. Dimensional analysis

Dimensional analysis enables the identification of scale-invariant dimensionless parameters by combining variables, reducing system complexity, and establishing universal relationships for experimental scaling, theoretical modeling, and validating physical similitude across diverse systems. Through the investigation of the mechanical and fluid properties of the flow passing over the weir and the permeable gabion structure, the key parameters influencing energy dissipation levels in the downstream region were identified and presented as Eq. (1):

$$F(\Delta E, E_0, \rho, g, \mu, d, Q, b, L, h, z, y_1, y_2, y_3, \varepsilon, d, L_s)=0 \quad (1)$$

Wherein h represents the upstream water depth, y_1 the downstream water depth at the weir, y_2 the water depth upstream of the gabion structure, y_3 the water depth downstream of the gabion, μ the dynamic viscosity, ρ the water density, E_o the upstream energy, z the weir height, Q the flow discharge, b the flume width, g the gravitational acceleration, $\Delta E=E_2-E_3$ the energy dissipation, ε the porosity, L the gabion length, and d the average aggregate diameter (referring to Figure 2).

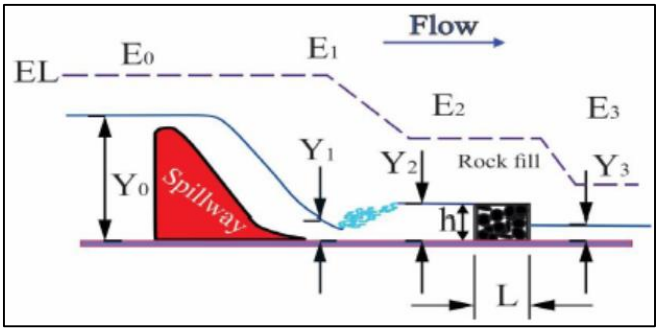

Figure 2-Parameters measured in experiments

By applying Buckingham's Pi-theorem and selecting μ, y_1, and Q as repeating parameters, the following dimensionless equation was derived:

$$F\left(\frac{\rho qL}{\mu}, \frac{q}{\sqrt{gy_1^3}}, \frac{y_2}{y_1}, \frac{y_3}{y_1}, \frac{h}{y_1}, \frac{\varepsilon}{y_1}, \frac{d}{y_1}, \frac{z}{y_1}, \frac{E_o}{y_1}, \frac{\Delta E}{y_1}\right)=0 \quad (2)$$

The parameters $\frac{w}{y_1}$ and $\frac{z}{y_1}$ had constant values in all experiments and were therefore excluded from the calculations. Additionally, the parameter $\frac{\rho qL}{\mu}$, which represents the Reynolds number, did not influence the computations due to the turbulent nature of the flow. For this reason, its effect was neglected. By combining the two parameters $\frac{E_o}{y_1}$ and $\frac{\Delta E}{y_1}$, the parameter $\frac{\Delta E}{E_o}$ was derived. The final form of the equation under investigation can be expressed as follows:

$$\frac{\Delta E}{E_o} = F\left(\frac{y_2}{y_1}, \frac{y_3}{y_1}, \frac{h}{y_1}, \frac{\varepsilon}{y_1}, \frac{d}{y_1}, Fr_1\right) \quad (3)$$

2.3. Sensitivity analysis

Sensitivity analysis is a crucial step in modeling complex systems, allowing for the evaluation of the impact of input variables on model output. The Gamma Test (GT) is a nonparametric statistical method used to estimate the minimum mean square error (MSE) in nonlinear models. This method estimates the noise level in data and identifies significant input variables without requiring model construction. WIN-GAMMA software is an effective tool for

performing the Gamma Test, processing data, and determining the most influential input variables. One of the key metrics in this software is the Variance Ratio (V-ratio)**,** which plays a critical role in assessing data predictability.

The Gamma Test is based on the principle that if a smooth function f(x) maps input variables X to output Y, then the squared differences between output values for closely spaced inputs should be small. This method estimates the noise level in data and assesses model predictability. For each sample x_i, a set of its kk nearest neighbors is identified, and their mean Euclidean distance is computed as follows:

$$\Delta_m = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{m}\sum_{i=1}^{m}\left\|x_i - x_{i(j)}\right\|^2 \quad (4)$$

Where $x_{i(j)}$ represents the ij-th nearest neighbor of x_i, m is the number of nearest neighbors considered, and N is the total number of samples. Similarly, the mean squared difference of output values for neighboring data points is computed as:

$$\Gamma_m = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{m}\sum_{i=1}^{m}\left\|y_i - y_{i(j)}\right\|^2 \quad (5)$$

Wherein y_i represents the output corresponding to x_i. Using linear regression on the points (Δ_m, Γ_m), the gamma value (γ), which indicates data noise, is estimated as follows:

$$\Gamma_m = A\Delta_m + \gamma \quad (6)$$

Wherein Slope A represents the rate of change in mean squared output differences based on input distances**,** Intercept γ denotes the gamma value, which reflects the inherent noise level in the data.

The Variance Ratio (V-ratio) is a key metric in the Gamma Test, computed as follows:

$$\text{V-Ratio} = \frac{\gamma}{\sigma^2(y)} \quad (7)$$

Where γ is the gamma value, $\sigma^2(y)$ is the variance of the output data. The small values of the V-ratio indicate that the model can accurately predict the output. Values close to 1 suggest high noise levels in the data and weak modeling capability. The Gamma Value (γ) plays a crucial role in assessing model predictability, where a value less than 0.01 indicates high model predictability, while a value greater than 0.1 suggests significant noise, making modeling less reliable. The Correlation Coefficient (R^2) between Δ_m and Γ_m provides further insights into data quality; when it is close to 1, it indicates smooth data with high predictability, whereas low values suggest significant errors in the data. The Variance Ratio (V-ratio) is another key metric, where a V-ratio below 0.1 suggests that input variables have high predictability, a V-ratio between 0.1 and 0.5 indicates that modeling is feasible but may be challenging, and a V-ratio above 0.5 implies low model accuracy due to significant data noise. Additionally, the Sensitivity Ranking of Input Variables is determined using WIN-GAMMA, which identifies the most influential variables by comparing gamma values for different input combinations, with the variable causing the highest increase in gamma value being considered the most important. When conducting sensitivity evaluation through the gamma method, each predictor variable is assigned a binary indicator (1 indicating its presence and 0 reflecting its absence). For instance, if all predictors are evaluated, the Mask field in Table 1-which summarizes select sensitivity outcomes—would display the seven-character binary code "1111111".

Overview of SVR

The SVRs are powerful supervised learning models used for classification and regression tasks. They are based on the principle of finding an optimal hyperplane that maximizes the margin between different classes in a dataset. SVR is particularly effective in high-dimensional spaces and cases where the number of dimensions exceeds the number of samples. For

classification problems, SVR aims to separate data points belonging to different classes using a decision boundary that minimizes classification errors. For regression problems, SVR attempts to fit a model within a certain error margin while minimizing complexity (Vapnik, 1995). Given a dataset of N training samples, where each sample consists of an input vector $x_i \in \mathbb{R}^d$ and a corresponding class label $y_i \in \{-1,1\}$, the objective of SVR is to find the optimal separating hyperplane:

$$w^T x+b=0 \quad (8)$$

Wherein w is the weight vector, b is the bias term, x represents input features. To ensure maximum separation between classes, SVR maximizes the margin $\frac{2}{\|w\|}$ while satisfying the constraints:

$$y_i(w^T x_i+b) \geq 1, \qquad \forall i=1,\ldots, N. \quad (9)$$

This constrained optimization problem is solved using Lagrange multipliers:

$$L(w,b,\alpha)=\frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i[y_i(w^T x_i+b)-1] \quad (10)$$

Where α_i are the Lagrange multipliers. The Karush-Kuhn-Tucker (KKT) conditions are applied to derive the final solution. For non-linearly separable data, SVR uses a soft margin approach by introducing slack variables ξ_i:

$$y_i(w^T x_i+b) \geq 1-\xi_i \qquad \xi_i \geq 0, \qquad \sum_{i=1}^{N} \xi_i \text{ minimized} \quad (11)$$

This results in the regularized objective function:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \xi_i \quad (12)$$

Wherein C is a penalty parameter that controls the trade-off between margin maximization and misclassification. For nonlinear problems, SVR employs a kernel function $K(x_i,x_j)$ to transform data into a higher-dimensional space where a linear separator can be applied. Common kernel functions include (Fuladipanah et al., 2024):

- Linear Kernel: $K(x_i, x_j) = x_i . x_j$.
- Polynomial Kernel: $K(x_i, x_j) = (x_i . x_j + c)^d$
- Radial Basis Function RBF Kernel: $K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2)$,
- Sigmoid Kernel: $K(x_i, x_j) = \tanh(\kappa x_i . x_j + c)$

Figure 3 displays the steps of the SVR model simulation process.

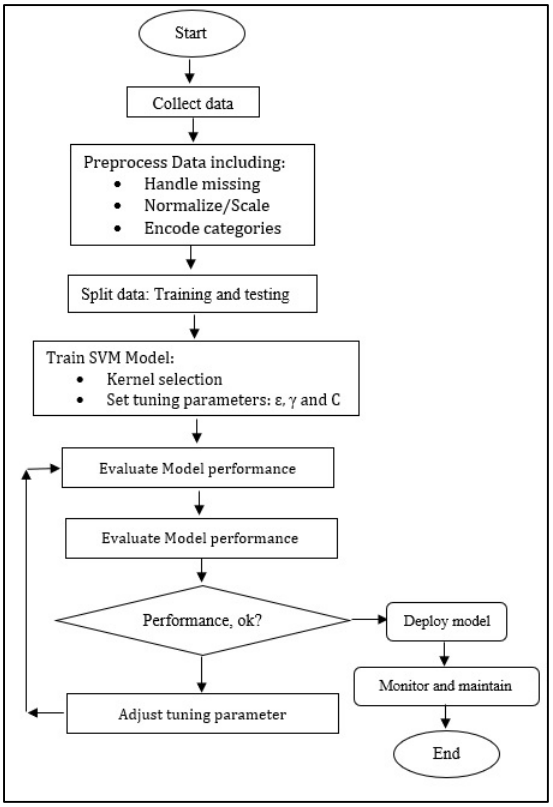

Figure 3- The SVR flowchart

2.4. Overview of GEP

The GEP is an evolutionary algorithm that combines the strengths of Genetic Algorithms (GA) and Genetic Programming (GP). It is widely used for symbolic regression, classification, time-series prediction, and optimization problems. GEP encodes mathematical expressions as linear chromosomes, which are then mapped into nonlinear expression trees for evaluation. GEP operates by evolving a population of candidate solutions through selection, mutation, crossover, and transposition. A GEP chromosome consists of genes, which encode mathematical expressions using function sets *(e.g., +, −, , /, sin, log)* and terminal sets (input variables and constants)**.** Each individual solution in GEP is represented as a fixed-length linear genome, which is later translated into a parse tree (Expression Tree, ET). The fitness of each individual is determined based on a predefined objective function, such as minimizing the error between predicted and actual values. The general mathematical representation of a function evolved by GEP is (Ferreira, 2001):

$$Y=F(X_1, X_2,..., X_n) \quad (13)$$

where Y is the target variable, and $X_1, X_2, ..., X_n$ are input variables. The evolutionary process optimizes the function ff to fit the given data best. A typical GEP simulation step of the GEP is illustrated in Figure 4 (Azamathulla et al., 2013; Ahmadzadeh-Kaleybar et al., 2024; Fuladipanah et al., 2024).

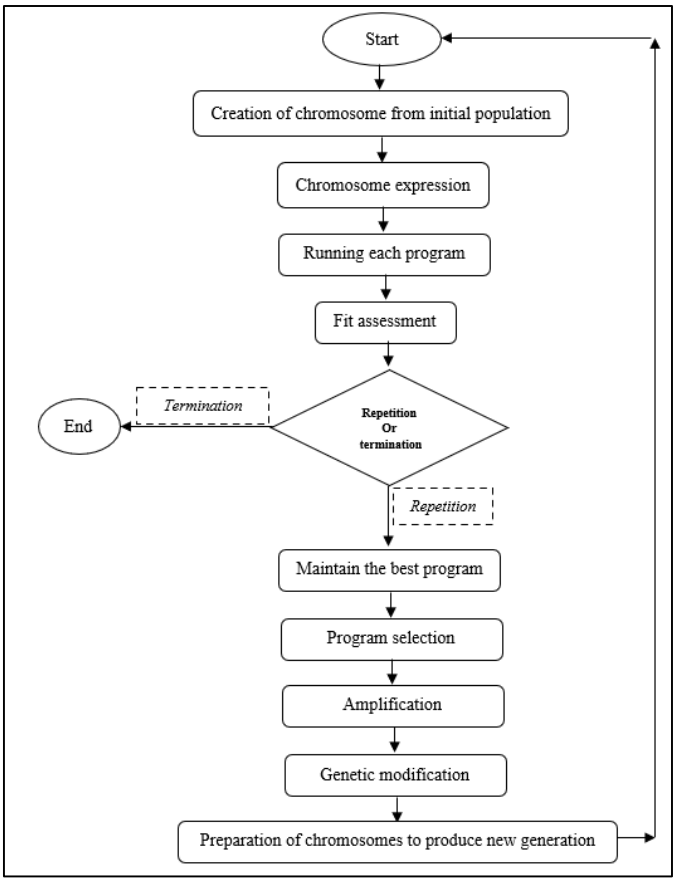

Figure 4-Flowchart of the GEP model

2.5. Overview of ANN

The ANNs are a class of machine learning models inspired by the structure and function of the human brain. They consist of interconnected processing units called neurons, which are organized in layers and trained to recognize patterns, approximate complex functions, and make predictions based on input data. A typical ANN consists of three main layers (Azmathullah et al., 2005):

- Input Layer: Receives raw data and passes it to the network.
- Hidden Layer(s): Process the data using weighted connections and activation functions to capture complex relationships.
- Output Layer: Produces the final predictions based on the transformed data.

The network learns by adjusting the weights and biases of connections between neurons to minimize prediction errors. This learning process is typically achieved through an optimization algorithm such as backpropagation with gradient descent. (Azmathullah et al., 2005; Fuladipanah et al., 2024).

Performance assessment

Evaluating model performance requires a systematic comparison of empirical and simulated results through quantitative indicators. These metrics detect biases, validate reliability, and characterize uncertainty to optimize predictive frameworks and inform decision-making. This study employs five statistical measures—R^2 (coefficient of determination), RMSE (root mean square error), MAE (mean absolute error), SI (scatter index), and DDR (developed discrepancy ratio)—to rigorously test model outputs. Their mathematical formulations are (Fuladipanah et al., 2023; Bagheri Khaneghahi et al., 2025):

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad (14)$$

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

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \quad (16)$$

$$SI=\frac{RMSE}{\frac{1}{n}\sum_{i=1}^{n}y_i} \quad (17)$$

$$DDR=\frac{\text{Predicted value}}{\text{Observed value}}-1 \quad (18)$$

Here, y_i denotes observed values, $\hat{y}_i$ represents predictions, $\hat{y}$ is the mean of observations, and n is the sample size. The R^2 indicates variance explained by the model (higher values preferred). The RMSE reflects average error magnitude, amplified by extreme deviations. The MAE provides outlier-resistant absolute error assessment. The SI standardizes errors for scale-agnostic comparisons. The DDR evaluates precision by benchmarking prediction-to-observation ratios, revealing error consistency across magnitudes. For enhanced interpretability, DDR values undergo Gaussian normalization. First, raw DDR data x is transformed into adjusted values x_{DDR} via distribution alignment. Second, x_{DDR} is plotted against standardized equivalents Z_{DDR} to visualize error patterns. Improved model performance is indicated by tighter clustering around the mean in these plots and higher normalized x_{DDR} values, demonstrating reduced variability and robust agreement between predictions and reality.

3.Results and Discussion

3.1. Opting for sensitive parameters

Table 2, which presents a summary of the sensitivity analysis calculations, indicates that among the seven extracted parameters in Eq. (3), the four parameters $\frac{y_2}{y_1}$, $\frac{h}{y_1}$, $\frac{d}{y_1}$ and Fr_1 have the most significant impact on the energy loss caused by the presence of the gabion. Therefore, it is essential to use these four parameters in machine learning-based simulations. This selection is justified as the lowest values of the parameters γ, A, SE, and V-Ratio correspond to the combination "1010011".

Table 2- Sensitivity analysis via Gamma-test

No.	Gamma	Gradian	Standard Error	V-ratio	MASK
1	5.965E-08	6.045E-04	2.399E-07	5.462E-04	1010011
2	3.543E-07	4.306E-04	3.884E-07	3.244E-03	0010001
3	4.219E-07	4.584E-04	2.263E-07	3.863E-03	1010000
4	2.743E-06	3.186E-05	5.185E-06	2.511E-02	1011110
5	1.647E-06	5.952E-05	9.091E-06	1.508E-02	0001101
6	3.238E-06	7.581E-05	3.881E-06	2.964E-02	1011011
7	1.495E-06	1.922E-04	1.062E-06	1.369E-02	0011000
8	2.546E-06	8.413E-05	7.297E-06	2.331E-02	0001110
9	3.487E-06	3.375E-05	5.506E-06	3.193E-02	1001110
10	2.629E-06	8.266E-05	1.291E-05	2.407E-02	0011100
11	4.051E-06	7.557E-05	4.382E-06	3.709E-02	1011010
12	7.404E-06	2.263E-05	3.892E-06	6.779E-02	1011111
13	6.188E-06	2.594E-05	4.013E-06	5.666E-02	1001111
14	5.193E-06	2.274E-04	4.960E-06	4.755E-02	1000011

15	1.647E-05	6.224E-06	4.983E-06	1.508E-01	1111111

Figure 5 presents the Gamma-test plot, which evaluates the relationship between input distances and squared output differences. A lower intercept (gamma value) indicates minimal noise, suggesting that the dataset is well-suited for predictive modeling. A steep gradient in the plot implies a strong dependency between input and output variables. Figure 6 ranks the input parameters based on their influence on energy dissipation, determined through the Gamma-test. Parameters with lower gamma values exhibit higher predictive power, whereas those with higher values contribute less to the model's accuracy. This sensitivity ranking aids in selecting the most relevant input features for machine learning-based simulations. Figure 7 illustrates the Variance Ratio (V-ratio), which quantifies the predictability of the dataset. A V-ratio close to zero indicates that the selected input variables effectively capture the variance in the output, making them suitable for modeling. Conversely, a higher V-ratio suggests increased noise and reduced predictive capability. By analyzing these figures, it becomes evident that a subset of the extracted parameters significantly impacts energy dissipation, reinforcing the necessity of including them in machine learning simulations to enhance model accuracy and reliability.

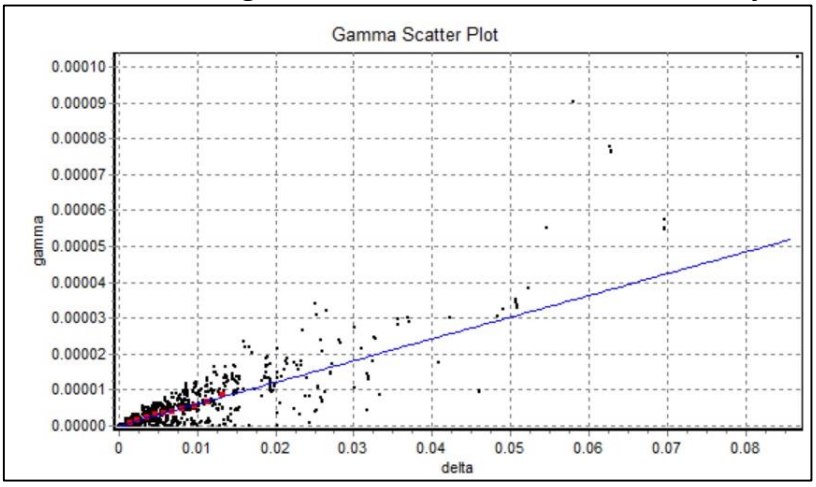

Figure 5- Gamma-test plot illustrating the relationship between γ and Δ

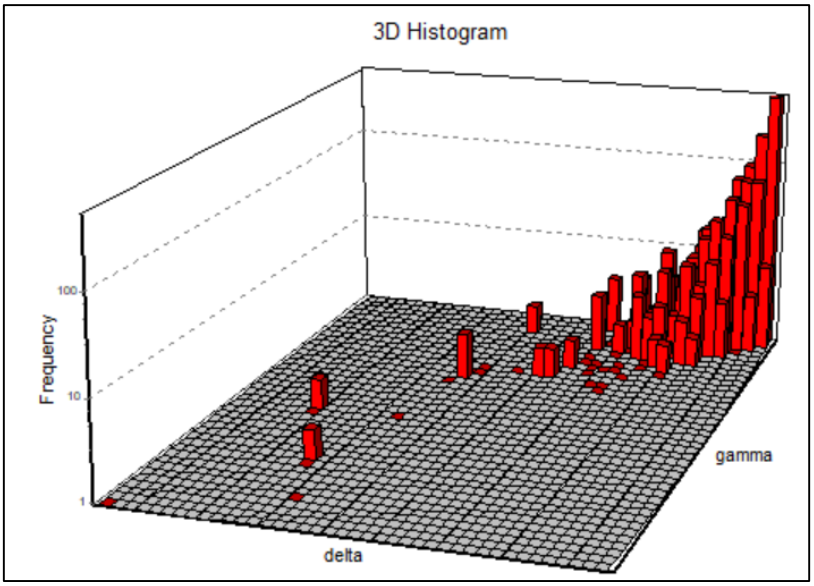

Figure 6- Sensitivity ranking of input parameters based on the Γ-test

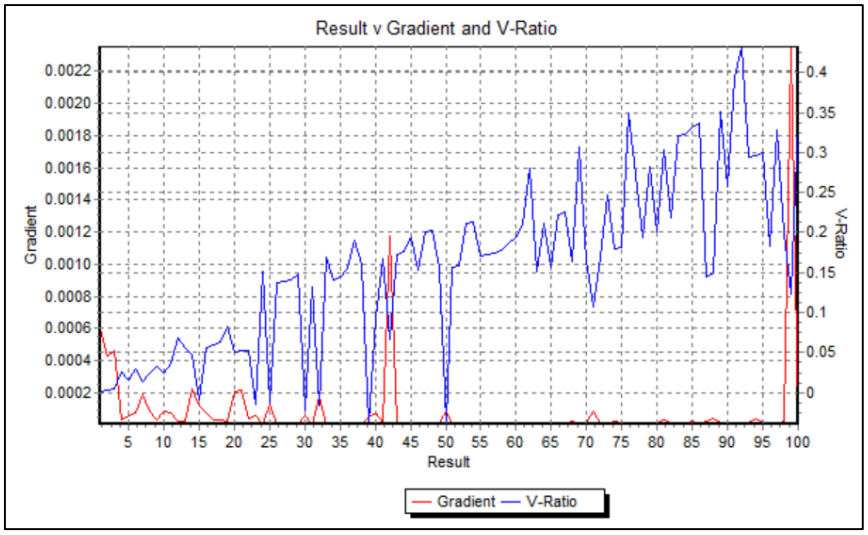

Figure 7- Variation of gradient and V-ration against the dataset

3.2. The MLMs outcomes analysis

In the next phase of the study, 155 measured laboratory data were used to simulate the percentage of energy loss caused by the porous gabion structure downstream of the flip bucket spillway using three machine learning models: SVR, GEP, and ANN. The training and testing processes for the models were set at 70% (including 110 data points) and 30% (including 45 data points), respectively. The results obtained from these three machine learning models were evaluated using performance assessment indices in both training and testing stages. The obtained results are presented in Table 3.

Table 3- Values of assessment metrics of MLMs

	Training phase				
	RMSE	MAE	R^2	SI	DDR_{max}
SVR	0.00661	0.00545	0.80383	0.02523	20.579
GEP	0.00308	0.00250	0.93648	0.01188	37.964
ANN (MLP 6-10-1)	0.00504	0.00414	0.86784	0.01935	24.783
	Testing phase				
	RMSE	MAE	R^2	SI	DDR_{max}
SVR	0.0007	0.00014	0.8818	0.00281	28.090
GEP	0.0006	0.0001	0.9311	0.00232	39.961
ANN (MLP 6-10-1)	0.0006	0.00012	0.8976	0.00233	32.875

In the training phase, GEP demonstrated the highest accuracy, achieving the lowest RMSE (0.00308) and MAE (0.00250), which indicates minimal deviation from actual energy loss values. It also recorded the highest R^2 (0.93648), signifying a strong correlation between predicted and observed values. The SI was the lowest for GEP (0.01188), reinforcing its robustness, while the highest DDR_{max} (37.964) suggests its ability to effectively capture extreme variations in energy loss. The values of tuning parameters of the GEP model are presented in Table 4. Figure 8 presents tree expression of the GEP model. The values of constants for the first to the third gene are (G1C0 = -7.71875, G1C1 = 9.77649), (G2C0 = -4.846497, G2C1 = 3.405975) and (G3C0 = 8.144043, G3C1 = 9.671326), respectively. The parameters d0 to d4 stand for $\frac{y_2}{y_1}$, $\frac{h}{y_1}$, $\frac{d}{y_1}$ and Fr_1, respectively.

Table 4- Values of tuning parameters for the GEP model

Parameters	Value
Head size	7
Chromosomes numbers	45
Number of genes	3
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 the transposition rate	0.1
RIS transposition rate	0.1
Fitness function error type	RMSE
Linking function	+

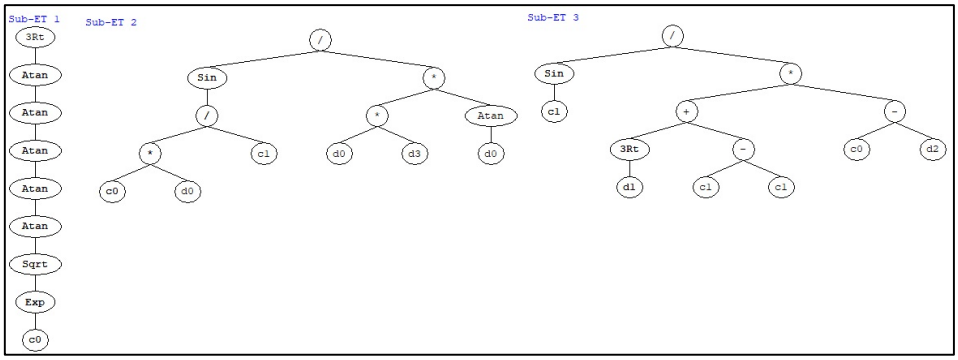

Figure 8- Tree expression of GEP model

The optimum outcomes of ANN model were obtained for ANN (MLP 6-10-1) model. The validation performance, optimizing algorithm, error function and hidden layer activation function were BFGS 344, SOS, Tanh and Logistic, respectively. The ANN (MLP 6-10-1) followed GEP in performance, with an RMSE of 0.00504 and an MAE of 0.00414, indicating a slightly higher level of error. The R^2 value for ANN was 0.86784, which is lower than GEP but still indicates a good fit to the data. The SI value for ANN (0.01935) was higher than GEP, suggesting that its predictions had slightly more dispersion. The DDR_{max} value for ANN was 24.783, showing that while it captured some extreme variations, it was less effective in this regard than GEP.

The SVR exhibited the lowest predictive accuracy among the three models, with the highest RMSE (0.00661) and MAE (0.00545), indicating greater deviation from actual values. The R^2 value of 0.80383 suggests that SVR explained less variance in the dataset compared to the other models. The SI value for SVR (0.02523) was the highest, indicating more dispersion in predictions, while the DDR_{max} value (20.579) was the lowest, suggesting that it was the least effective in capturing extreme variations in energy loss. The values of tuning parameters of SVR model were C=57, γ=2.3, and ε=0.21.

In the testing phase, GEP again outperformed the other models, achieving the lowest RMSE (0.0006) and MAE (0.0001), confirming its reliability in predicting unseen data. It also had the highest R^2 value (0.9311), demonstrating its strong generalization capability. The lowest SI value (0.00232) for GEP indicated high stability, while the highest DDR_{max} value (39.961) confirmed its ability to handle extreme cases effectively. ANN maintained good predictive accuracy in the testing phase, with an RMSE of 0.0006 and an MAE of 0.00012. Its R^2 value of 0.8976, though slightly lower than GEP, still showed a strong correlation with measured values. The SI value for ANN (0.00233) was close to that of GEP, indicating similar predictive stability, while the DDR_{max} value (32.875) suggested that it was capable of capturing significant

variations but not as effectively as GEP. SVR showed improvement in the testing phase compared to training, with an RMSE of 0.0007 and an MAE of 0.00014. However, its R^2 value (0.8818) remained the lowest among the three models, indicating that it was still less effective in explaining the variance in energy loss. The SI value for SVR (0.00281) was slightly higher than those of GEP and ANN, indicating that its predictions were more dispersed. The DDR_{max} value for SVR (28.090) was higher than in the training phase, suggesting an improvement in capturing extreme variations, but it was still lower than those of GEP and ANN.

Figure 9 presents scatter plots comparing the measured and predicted values of energy dissipation for the three MLMs. The scatter plot provides a visual representation of each model's predictive accuracy by plotting observed values against their corresponding predictions. A strong predictive model should exhibit data points closely aligned with the 45-degree reference line (y=x)**,** which indicates a perfect match between measured and predicted values. The degree of dispersion around this line determines how well the model captures the variability in energy dissipation. The GEP model exhibits the highest alignment with the 45-degree line, demonstrating a strong correlation between measured and predicted values. The clustering of points around this reference line indicates minimal error, reinforcing GEP's high R^2 values observed in Table 3. This suggests that GEP provides the most reliable predictions with minimal deviations. The ANN model also shows a good correlation between measured and predicted values, with most data points falling close to the reference line. However, there is slightly more dispersion compared to GEP, particularly for extreme values. This aligns with the slightly lower R^2 value of ANN relative to GEP, indicating that while ANN performs well, it is marginally less accurate. The SVR model exhibits the highest dispersion of points around the 45-degree line, indicating a weaker correlation between predicted and measured values. The increased spread of points suggests a higher error rate, aligning with SVR's lower R^2 value and higher RMSE and MAE values in Table 3. This confirms that SVR struggles to predict energy dissipation accurately compared to the other two models. Overall, Figure 9 visually confirms the numerical results in Table 3, highlighting that GEP is the most effective model for predicting energy dissipation, followed by ANN, while SVR shows the least accuracy.

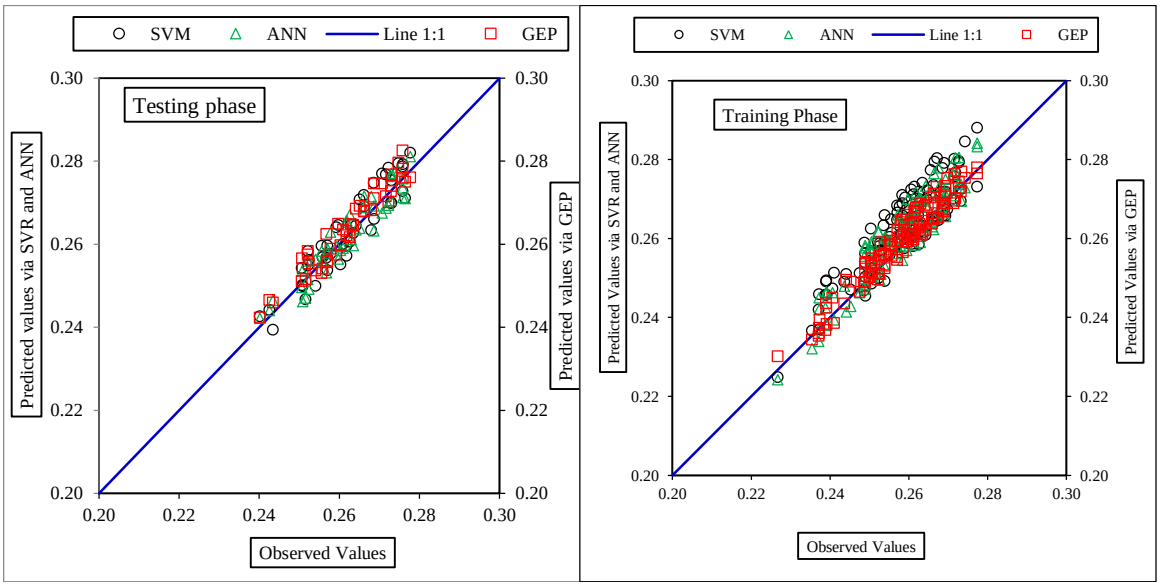

Figure 9- Scatter plot of measured vs. predicted values of energy dissipation of MLMs

Figure 10 illustrates the distribution of the DDR index for the three machine learning models in both the training and testing phases. The DDR index is an important performance metric that

evaluates the consistency and reliability of a model's predictions by comparing the ratio of predicted to observed values. A lower spread and higher concentration around the mean DDR value indicate greater stability and accuracy in predictions. The GEP model exhibits a narrow and well-centered distribution in both training and testing phases. The low variance in the DDR index indicates that GEP maintains high predictive consistency across different datasets. Additionally, the DDR_{max} value for GEP is the highest among the three models, confirming that it effectively captures extreme variations in energy dissipation. The ANN model shows a slightly broader distribution compared to GEP, particularly in the testing phase. While the DDR values remain relatively stable, the increased spread suggests a slightly lower consistency in predictions. This observation aligns with the model's slightly lower R^2 and higher RMSE compared to GEP. The SVR model displays the widest spread in DDR values, indicating greater variability and inconsistency in its predictions. The higher variance suggests that SVR is less reliable in capturing energy dissipation patterns, especially in extreme cases. This aligns with the model's higher error metrics and lower R^2 values observed in Table 3. Figure 10 confirms that GEP provides the most stable and consistent predictions across both training and testing phases, followed closely by ANN, while SVR demonstrates the highest level of uncertainty and error in its predictions. The DDR index distribution further supports the conclusion that GEP is the most suitable model for predicting energy dissipation in the presence of a gabion structure.

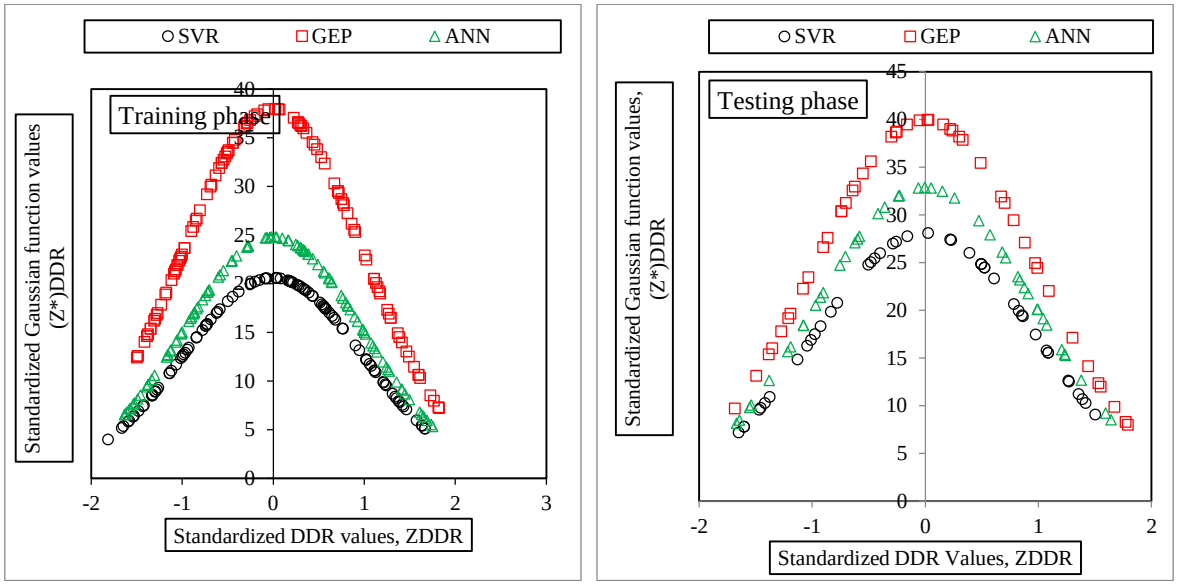

Figure 10- Distribution of DDR index for training and testing phases of MLMs

4.Conclusions

This study evaluated the energy dissipation efficiency of gabion structures downstream of spillways using three MLMs: SVR, GEP, and ANN. The predictive performance of these models was assessed using standard statistical metrics, including RMSE, MAE, R^2, SI, and DDR. The findings demonstrate that the GEP model exhibited superior predictive capability in both training and testing phases, outperforming the ANN and SVR models in terms of accuracy and consistency. The training phase results indicated that GEP provided the most accurate predictions with minimal error, followed by ANN, while SVR demonstrated the highest variability and least accuracy. Similarly, in the testing phase, GEP maintained its superior performance, showing strong generalizability and predictive stability, whereas ANN followed closely with slightly higher error margins. SVR, despite improvements, continued to show the lowest performance among the three models. Sensitivity analysis via the Gamma-test was

conducted to determine the most influential parameters affecting energy dissipation. The analysis identified four key parameters: $\frac{y_2}{y_1}$, $\frac{h}{y_1}$, $\frac{d}{y_1}$ and Fr_1. These parameters exhibited the highest sensitivity scores, indicating their significant impact on the predictive accuracy of the models.

Scatter plots further validated these findings, illustrating that the GEP model's predicted values aligned closely with measured energy dissipation values, indicating minimal error and high correlation. The ANN model also demonstrated strong alignment but exhibited slightly greater dispersion in its predictions, whereas the SVR model displayed the highest variability, confirming its lower predictive reliability. The DDR index distribution further highlighted the consistency of the GEP model, which exhibited the most concentrated distribution, followed by the ANN and SVR models, reinforcing the superiority of the GEP approach.

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.

Ethical Considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

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.

Alsubih, M., Ahmed, M., Alqadhi, S., Mallick, J. 2023. Gabion water barrier structures as a sustainable approach to water and land conservation. *Environmental Science and Pollution Research*, 30(60):126057-126071. DOI: 10.1007/s11356-023-31040-3.

Azamathulla, H.M., Cuan, Y.C., Ghani, A.A., Chang, C.K., 2013. Suspended sediment load prediction of river systems: GEP approach. *Arabian Journal of Geosciences*, 6, 3469-3480.DOI: 10.1007/s12517-012-0608-4.

Azmathullah, H.M., Deo, M.C. and Deolalikar, P.B. 2005. Neural networks for estimation of scour downstream of a ski-jump bucket. *Journal of Hydraulic Engineering*, 131(10): 898-908. https://doi.org/10.1061/(ASCE)0733-9429(2005)131:10(898).

Bagheri Khaneghahi,M. , Hezarjaribi,A. , Kamali,M. I. and Zamani,F. 2025. Evaluating the accuracy of AquaCrop 7.1 crop model and investigating the effect of changing the planting date on the performance of rainfed wheat in different climates of Iran. *Journal of New Approaches in Water Engineering and Environment*, *4*(1), 129-144. Doi: 10.22034/nawee.2025.493925.1122.

Daneshfaraz, R., Majedi Asl, M., Bagherzadeh, M. 2021. Experimental Investigation of the Energy Dissipation and the Downstream Relative Depth of Pool in the Sloped Gabion Drop and the Sloped Simple Drop. *Amirkabir Journal of Civil Engineering*, 53(9):3665-36778. DOI: 10.22060/ceej.2020.18059.6751.

Daneshfaraz, R., Majedi-Asl, M., Mortazavi, S., Bagherzadeh, M. 2022. Laboratory evaluation of energy dissipation in the combined structure of the vertical drop with gabion. *Civil Infrastructure Researches*, 8(1):145-157. DOI:10.22091/cer.2022.7720.1344 .

Ferreira, C. 2001. Algorithm for solving gene expression programming: a new adaptive problem. *Complex Systems*, 13(2), pp. 87-129. https://doi.org/10.48550/arXiv.cs/0102027.

Fuladipanah, M., Jafarinia, R. 2011. The Derivation of Energy Dissipation Equation for Adverse-Slopped Stepped Spillway. *World Applied Science Journal*, 15(5): 637-642.

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), p. 231. https://doi.org/10.1007/s13201-023-02022-0.

Fuladipanah, M., Shahhosseini, A., Rathnayake, N., Azamathulla, H.M., Rathnayake, U., Meddage, D.P.P., Tota-Maharaj, K. 2024. In-depth simulation of rainfall–runoff relationships using machine learning methods. *Water Practice & Technology*, 19(6), pp. 2442-2459. https://doi.org/10.2166/wpt.2024.147.

Hussain, A., Ansari, M.A., Ahmed, M.N., Ahmad, F., Jahangeer, J. 2022. Model development for energy dissipation over gabion stepped weirs using GEP and GMDH techniques. *Canadian Journal of Civil Engineering*, 49(6): 969-979. DOI: 10.1139/cjce-2021-0197.

Iranpour,A. , Egdernezhad,A., Heidarnejad,M. 2024. Numerical Investigation of Energy Dissipation in Baffled Piano Key Weirs. *Journal of New Approaches in Water Engineering and Environment*, *3*(1): 55-70. Doi: 10.22034/nawee.2024.451817.1069.

Kheyraei, M., Fathi Moghaddam, M. 2016. Hydraulic Characteristics of the Crump Gabion Weirs for Free Flow Conditions. *Water Engineering*, 9(29):75-86. DOI: 20.1001.1.20086377.1395.9.29.6.9

Mahjoubi, A., Kashefipour, S.M., 2023. Investigation the Energy Dissipation on Gabion Stepped Weirs and Downstream Hydraulic Jump Characteristics. *Irrigation Sciences and Engineering*, 46(1): 65-76. DOI: 10.22055/jise.2019.14393.

Mobayen, R., Najafzadeh, M., Farrahi-Moghaddam, K. 2023. Evaluation of Regression-Based Soft Computing Techniques for Estimating Energy Loss in Gabion Spillways. *Journal of Environment and Water Engineering*, 9(2):241-255. DOI: 10.22034/ewe.2022.329153.1724.

Naseri, R., KashefiPour, S.M. 2022. The Effect of Gabion Stepped Spillway Porosity on Energy Dissipation and Characteristics of Downstream Hydraulic Jump of Weir. *Irrigation Sciences and Engineering*, 45(1): 1-17. DOI: 10.22055/jise.2019.18454.1337.

Pourhosein Ghadi, M., Dehghani, A.A., Meftah Halghi, M. 2022. Effect of different hydraulic conditions on the hyporheic flow characteristics around gabion weir structures. *Iranian Journal of Eco Hydrology*, 9(1): 15-33. DOI: 10.22059/ije.2022.329091.1539.

Razi, S., Salmasi, F., Hoseinzade Dalir, A. 2019. Laboratory Study of the Effects of Step Number, Slope and Particle Size on Energy Dissipation in Gabion Stepped Spillways. *Amirkabir Journal of Civil Engineering*, 51(4): 749-756. DOI: 10.22060/ceej.2018.13984.5527.

Salehi, S., Esmaili, S.A., Sadeghi-Namaghi, N., Esmaili, K. 2022. Overflow and underflow within the broad-crested gabion weir in free flow with sand bed conditions. *Sharif Journal Civil Engineering*, 37(4): 3-13. DOI: 10.24200/j30.2021.57425.2907.

Srinivas, R., Tiwari, N.K. 2022. Oxygen aeration efficiency of gabion spillway by soft computing models. *Water Quality Research Journal*, 57(3):215-232. https://doi.org/10.2166/wqrj.2022.009.

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

Zuhaira, A.A., Al-Hamd, R.K.S., Alzabeebee, S., Cunningham, L.S. 2021. Numerical investigation of skimming flow characteristics over non-uniform gabion-stepped spillways. *Innovative Infrastructure Solutions*, 6,1-19. DOI: 10.1007/s41062-021-00579-w.