Journal of Hydroinformatics

Journal of Hydroinformatics Vol 26 No 1, 337 doi: 10.2166/hydro.2023.260

Analysis of discharge characteristics of a symmetrical stepped labyrinth side weir based on global sensitivity

Wuyi Wan[a], Guiying Shen[a,*], Shanshan Li[b], Abbas Parsaie[c], Yuhang Wang[a] and Yu Zhou[a]

[a] Department of Hydraulic Engineering, College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China
[b] State Key Laboratory of Eco-hydraulics in Northwest Arid Region of China, Xi'an University of Technology, Xi'an 710048, China
[c] Faculty of Water Sciences Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran
*Corresponding author. E-mail: sgy2023@zju.edu.cn

ABSTRACT

In this paper, the discharge coefficient prediction model for this structure in a subcritical flow regime is first established by extreme learning machine (ELM) and Bayesian network, and the model's performance is analyzed and verified in detail. In addition, the global sensitivity analysis method is introduced to the optimal prediction model to analyze the sensitivity for the dimensionless parameters affecting the discharge coefficient. The results show that the Bayesian extreme learning machine (BELM) can effectively predict the discharge coefficients of the symmetric stepped labyrinth side weir. The range of 95% confidence interval [−0.055,0.040] is also significantly smaller than that of the ELM ([−0.089,0.076]) and the Kernel extreme learning machine (KELM) ([−0.091,0.081]) at the testing stage. The dimensionless parameter ratio of upstream water depth of stepped labyrinth side weir p/y_1 has the greatest effect on the discharge coefficient C_d, accounting for 55.57 and 54.17% under single action and other parameter interactions, respectively. Dimensionless step number b_s/L has little effect on C_d, which can be ignored. Meanwhile, when the number of steps is less ($N = 4$) and the internal head angle is smaller ($\theta = 45°$), a larger discharge coefficient value can be obtained.

Key words: discharge coefficient, Labyrinth side weir, machine learning, sensitivity analysis

HIGHLIGHTS

- The discharge coefficient model of a symmetrical stepped labyrinth side weir is developed.
- The quantitative model of the discharge coefficient is established.
- The discharge characteristics of a symmetrical stepped labyrinth side weir are analyzed comprehensively.
- The effects of different dimensionless parameters on the discharge coefficient are compared.

1. INTRODUCTION

As a new type of weir, the labyrinth weir can greatly increase the discharge for hydraulic structures with a limited overflow front (Karimi *et al.* 2019). Its excellent discharge capacity and low project cost have made it popular for reservoir reinforcement, landscape engineering, flood control engineering, and spillway reconstruction in recent years. Also, considering factors such as power supply guarantee and management cost in practical engineering, compared with the broad-crested weir controlled by gate control, the labyrinth weir can greatly reduce the operation and management cost. However, in practical engineering, the design parameters of labyrinth weir are generally estimated based on experience, and then verified by model experiment. The whole design process lasts for a long time. If the design parameters are not properly formulated, it will lead to repeated optimization and adjustment of engineering design and modeling experiments, resulting in a waste of time and cost (Yan 2023).

At present, there are mainly rectangular, triangular, circular, and other shapes for the crest shape of labyrinth weirs (Mahmoud *et al.* 2021). Alfatlawi *et al.* (2023) presented a symmetrical stepped labyrinth side weir. The discharge capacity of the weir can be increased by 15–35% using the larger length of the weir, small head angles, and steps for the triangular labyrinth side weir. The discharge coefficient (C_d) is an important hydraulic parameter of the structure. With the change in the number of steps, the water flows through the crest of the weir, there are both positive water flow and lateral water flow, which makes the overflow nappe collide with each other to produce secondary flow and vortex. Its discharge

characteristics are more complex than the traditional weir type. Its discharge characteristics are more complex than the traditional labyrinth side weir. Thus, it is important to systematically study the discharge characteristics of the structure and reveal the interaction mechanism of hydraulic variables for the application of the structure in practical engineering.

Many scholars have studied the C_d for labyrinth side weirs through modeling experiments and numerical simulations (Ben Said & Ouamane 2022; Alfatlawi & Alkafaji 2023; Saffar *et al.* 2023; Zare *et al.* 2023). So far, some researchers have adopted soft computing methods to predict C_d for efficient calculation. Bijanvand *et al.* (2023) assessed the discharge in compound open using the soft computing method, and the most important parameters affecting the discharge were the relative roughness, the ratio in terms of flow dimensions, the relative depth, the convergence angle, and the relative radius. Simsek *et al.* (2023) developed a C_d prediction model for trapezoidal broad-crested weirs, and the Artificial Neural Network model outperformed M5Tree and the Support Vector Machines model. Roushangar *et al.* (2023) established a C_d prediction model using different flow conditions with radial gates. The values of $R = 0.940$, RMSE $= 0.022$, and $R = 0.927$, RMSE $= 0.018$ were obtained for KELM-GWO and Support Vector Machines models for flow under submerged and free conditions, respectively. Azma *et al.* (2023) developed the Gabion weir flow model, and while estimating the ratio of flow discharge, the CatBoost model achieves the best performance; when calculating the C_d, the performance of the XGBoost and CatBoost models is very close to each other. Emami *et al.* (2022) evaluated the Labyrinth Weir's C_d using the Walnut algorithm and support vector regression (Walnut-SVR), and the Walnut-SVR model has higher accuracy than other counterparts. Seyedian *et al.* (2023) state that Gaussian process regression (GPR) provides improved performance and robustness (RMSE $= 0.009$ and $R^2 = 0.986$) for a triangular labyrinth weir. Zaji *et al.* (2020) developed the neural network and genetic algorithm to evaluate the labyrinth side weir's C_d, and the Genetic Algorithm Radial Basis neural network model could still predict the C_d for this structure with the limited dataset. Zounemat-Kermani *et al.* (2019) predicted the C_d of the triangular arc labyrinth weir, and the Multi-Layer Perceptron Neural Network has the highest precision and generalization capability ($R^2 = 0.999$, RMSE $= 0.00385$, and bias coefficient bias < 0.0001). Akbari *et al.* (2019) developed a C_d calculation model for the Piano Key weir from 156 experimental datasets, and they showed that the GPR model significantly outperforms other intelligent models. Karami *et al.* (2018) demonstrated that the ELM model outperforms ANN and genetic programming (GP) in predicting the C_d for a triangular labyrinth weir.

However, according to the current literature research, no scholars have proposed an intelligent calculation model for the C_d of the symmetrical stepped labyrinth side weir. Also, the C_d for the structure is affected by the slope of the steps, the height of the steps, the width of the steps, and the number of steps. In addition, many scholars have explored the impact of hydraulic parameters on the C_d through the sensitivity method. Samadi *et al.* (2022) researched the influence of crest height, cycle number, and plane weir configuration on the C_d for a labyrinth weir through the model experiments, and results show that under the same geometric parameters, the rectangular has a higher discharge for each head as compared to trapezoidal and triangular. Dogan & Kaya (2023) experimentally explored the influence of effective crest length on the C_d for the labyrinth side weir, and by reducing the upstream effective crest length by 1/3, 2/3, and 3/3, the discharge reduced about 10, 23, and 48%, respectively. It can be seen that the influence on C_d of hydraulic parameters has also been the focus of attention for many researchers. For the stepped labyrinth side weir, hydraulic parameters are obviously more than other labyrinth weirs. Furthermore, the stability of the model can be improved by reducing parameter uncertainties and analyzing the influence on Cd from parameter interactions. However, for the impact of hydraulic parameters on C_d, most scholars use local sensitivity studies to obtain the influence of individual parameters on C_d.

This paper is designed to address the discharge characteristics for the stepped labyrinth side weir in a subcritical flow regime. Firstly, the C_d calculation model is developed based on the experimental dataset and artificial intelligence algorithm. The dimensional analysis of the hydraulic parameters affecting the C_d is carried out to define the model input parameters and output parameters. The optimal intelligent model is synthesized by different evaluation indexes, and the model uncertainty is further analyzed. On this basis, the global sensitivity method is introduced to quantify the sensitivity coefficients between the dimensionless parameters and the discharge coefficients, and the insensitive parameters are removed. Finally, the variation of important parameters about the discharge coefficients is analyzed.

2. MATERIAL AND METHODS

2.1. Experimental model

The experimental data were from Alfatlawi *et al.* (2023). The experimental model is composed of a rectangular flume of 10 m in length, 0.3 m in width, and 0.45 m in height, and the structure consists of a metal bed and glass side walls. The main and

evacuation channels have a calibrated rectangular gauging weir positioned at each channel's end, and a point gauge with ± 0.1 mm accuracy is fixed at the upstream end of the weir, to provide for precise discharge measurement, 0.40 m away from the gauging weir. The stepped labyrinth side weir is located 5.50 m above the main channel. This experiment was performed on different stepped labyrinth side weirs with different lengths, heights, number of steps, and internal head angles for a total of 275 experiments. The experiment was in a subcritical flow regime, and the data are characterized as shown in Figure 1.

2.2. Dimensionless parameters

For a symmetric stepped labyrinth side weir, the discharge is calculated as:

$$Qw = \frac{2}{3} C_d \cdot L_f \cdot \sqrt{2g}(y_1 - p)^{\frac{3}{2}} \tag{1}$$

where L_f is the weir crest length, m; C_d is the discharge coefficient; y_1 is the upstream water depth of the channel center, m; g is the acceleration of gravity, m/s^2.

As shown in Figure 2, the parameters that affect the value of C_d are as follows:

$$C_d = f_1(L, B, p, y_1, bs, as, v, g, \rho, \sigma, \mu) \tag{2}$$

According to Buckingham π theorem, the dimensional analysis shows that:

$$C_d = f_2\left(\frac{L}{B}, \frac{L}{y_1}, \frac{p}{y_1}, \frac{b_s}{a_s}, \frac{b_s}{L}, F_r, R_e, W_e\right) \tag{3}$$

where L is the side weir width, m; B is the channel width, m; p is the side weir height, m; b_s is the step width, m; a_s is the step height, m; b_s/a_s is the dimensionless step slope (θ); b_s/L is the dimensionless steps number (N); F_r is the upstream Froude number; W_e is the Weber number; R_e is the Reynolds number. The experiment was carried out under subcritical flow and steady flow, and the water head above the weir was higher than 30 mm. Thus, the influence of the R_e and W_e on the C_d can be neglected (Mohammed & Golijanek-Jedrzejczyk 2020). Refer to Subramanya & Awasthy (1972) and Borghei *et al.*

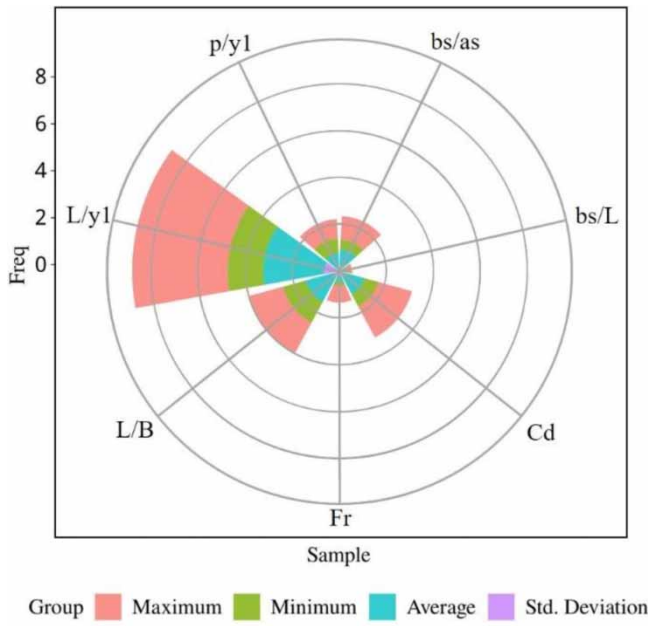

Figure 1 | Characteristics of experimental data.

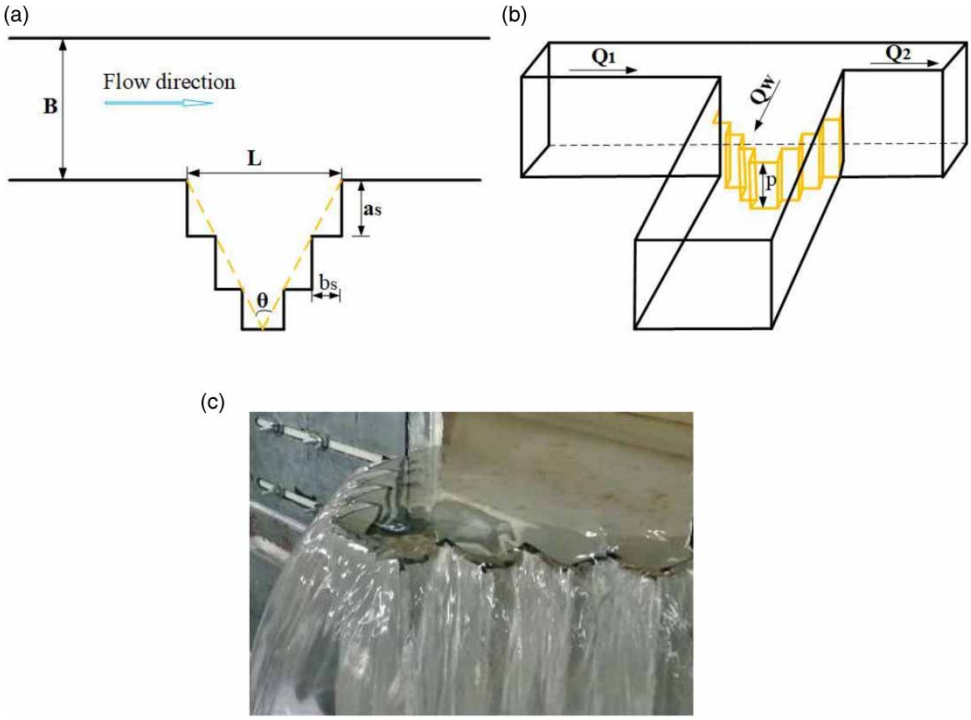

Figure 2 | Symmetrical stepped labyrinth side weir: (a) layout plan, (b) three-dimensional structure diagram, and (c) flow regime diagram (Alfatlawi *et al.* 2023).

(1999) for the dimensional analysis of discharge coefficients. Equation (3) can be expressed as the following equation.

$$C_d = f_2\left(\frac{L}{B}, \frac{L}{y_1}, \frac{p}{y_1}, \frac{b_s}{a_s}, \frac{b_s}{L}, F_r\right) \tag{4}$$

2.3. Models

Considering that the dataset used is the experimental dataset of the physical model. The dimensional analysis shows that the model has six input parameters, while the extreme learning machine (ELM) has only one hidden layer, so its learning speed is faster. Additionally, in the process of using the ELM, the output weight of a Moore–Penrose generalized inverse matrix is a least squares problem, it is easy to overfit. Therefore, the kernel function is introduced to avoid randomly generating weights and thresholds. Secondly, ELM accuracy is very sensitive to the number of hidden neurons, which leads to a very time-consuming and labor-intensive process by manually tuning the parameters in the training process. So far, the Bayesian network optimization ELM has shown good prospects in many engineering fields (Olyaie *et al.* 2019; Quilty *et al.* 2023), and the ability of this method to solve hydraulic engineering problems has not been explored. Therefore, in this study, a Bayesian network is chosen to find the optimal hyperparameters for ELM automatically.

2.3.1. Extreme learning machine

ELM as an improved neural network model using the least squares method for network development. ELM has a very fast learning speed because it is a simple untuned algorithm, which avoids the slow iterative learning process and easily falls into the local minimum (Luo & Zhang 2014). Compared with traditional intelligent algorithms, ELM has better scalability, generalization performance, lower computational complexity, and less manual intervention (Huang *et al.* 2011).

2.3.2. Kernel extreme learning machine (KELM)

The kernel function mainly relies on Mercer's theorem, and kernel-based learning methods are widely used for their powerful generalization performance. The KELM introduces the concept of kernel matrix and adopts the kernel function matrix

instead of the random matrix HH^T for the ELM model, which transforms the linearly indivisible data in the low-dimensional space into linearly divisible data, which not only preserves the speed of the ELM operation but also abandons the principle of randomly generating weights and biases of the ELM (Jalil-Masir *et al.* 2022). After determining the kernel function, KELM only needs to adjust the penalty factor, and the training results are stable and unchanged. The same accuracy can be obtained when the same samples are trained (Parida *et al.* 2021).

2.3.3. Bayesian extreme learning machine

The Bayesian extreme learning machine (BELM) uses Bayesian networks for optimizing the output layer weights for ELM models, which helps to overcome the model falling into overfitting. The joint probability distribution is represented by encoding the independent relationships of random variables using a graph. Generally, Bayesian modeling is divided into two steps (Soria-Olivas *et al.* 2011; Olyaie *et al.* 2019)

(1) The model parameter inference from the posterior distribution. In order to scale it, multiply the likelihood function by the prior distribution:

$$P(w|D) \propto P(w) \cdot P(D|w) \tag{5}$$

where P represents the probability distribution, w represents the parameter set, and D represents the dataset.

(2) Calculate a new distribution for the input x_{new} and the output y_{new}. It is defined as the integration over the posterior distribution for parameter w.

$$p(y_{new}|x_{new}, D) = \int p(y_{new}|x_{new}, w) \cdot p(w|D) \cdot dw \tag{6}$$

The linear modeling follows $t = h\beta + \varepsilon$

where $h = [h\ (x, o_1, r_2), ..., (x, o_{\text{M}}, r_{\text{M}})$ is the hidden layer's output vector for the input x; β is an output weight vector; ε obeys a Gaussian distribution with zero mean and σ^2 variance; and then, the output conditional distribution is as follows:

$$p(t|h, \beta, \sigma^2) = N(h\beta; \sigma^2) \tag{7}$$

For most applications, the parameters are distributed as shown in the following equation:

$$p(h|\alpha) = N(0; \alpha^{-1} \cdot I) \tag{8}$$

where I represent the unit matrix; α represents the hyperparameter. Assuming that the distribution of the prior and the likelihood function follow a Gaussian distribution and that the posterior distribution is also Gaussian, and the mean m and variance S are defined as follows:

$$\begin{aligned} m &= \sigma^{-2} \cdot S \cdot x^T \cdot y \\ S &= (\alpha I + \sigma^{-2} \cdot x^T \cdot x)^{-1} \end{aligned} \tag{9}$$

where $y = [y_1, y_2, ..., y_n]$, $x = [x_1, x_2, ..., x_N]$ are the model input vector and output vector matrices, respectively.

2.4. Global sensitivity analysis method

The Sobol method is used as a global sensitivity analysis method, which can quantitatively give the parameter sensitivity of the model (Sobol 1990). Sobol's method is centered on decomposing the overall variance in the objective function into the variance for individual parameters and the variance between combinations of parameters, and the sensitivity coefficient is obtained. More importantly, this method can simultaneously reflect the direct and interactive effects of parameters. Compared with other methods, the Sobol method has higher stability. The specific calculation principle can be found in our previous research (Shanshan *et al.* 2023).

2.5. Uncertainty analysis

Model uncertainty analysis can effectively evaluate the prediction error of the model and validate the data relationship for the model. Through the mean error ($\bar{e}$) and standard error (S_e) of the model, a 95% confidence interval was created around the

error value (Pradeep *et al.* 2022). Positive and negative mean error values represent overestimation and underestimation of the model, respectively.

$$\bar{e} = \frac{\sum_{i=1}^{N} e_i}{N} \tag{10}$$

$$S_e = \sqrt{\frac{\sum_{i=1}^{N} (e_i - \bar{e})^2}{N-1}} \tag{11}$$

$$e_i = P_i - O_i \tag{12}$$

where e_i represents the prediction error of individual data; N is the total number of samples in the dataset.

2.6. Model evaluation indices

Seven statistical methods were used in this study to evaluate model quality. The indices are root mean square error (RMSE), variance account factor (VAF), Nash–Sutcliffe efficiency (NSE), Willmott's agreement index (WIA), mean absolute percentage error (MAPE), coefficient of determination (R^2), bias coefficient (Bias), respectively. The specific calculation process for each indicator is described in the following equations (Pradeep & Samui 2022; Shen *et al.* 2022).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (O_i - P_i)^2}{N}} \tag{13}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{O_i - P_i}{O_i} \right| \times 100 \tag{14}$$

$$R^2 = \frac{\left(\sum_{i=1}^{N} (O_i - O_a)(P_i - P_a) \right)^2}{\sum_{i=1}^{N} (O_i - O_a)^2 \sum_{i=1}^{N} (P_i - P_a)^2} \tag{15}$$

$$WIA = 1 - \frac{\sum_{i=1}^{N} (P_i - O_i)^2}{\sum_{i=1}^{N} (|P_i - O_a| + |O_i - O_a|)^2} \tag{16}$$

$$Bias = \frac{1}{n} \sum_{i=1}^{n} (P_i - O_i)^2 \tag{17}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N} (O_i - P_i)^2}{\sum_{i=1}^{N} (O_i - O_a)^2} \tag{18}$$

$$VAF = \left(1 - \frac{var(O_i - P_i)}{var(O_i)} \right) \times 100 \tag{19}$$

where O_i represents the experimental value of C_d, P_i represents the predicted value of C_d, O_a represents the mean of the experimental value of C_d, and P_a represents the mean of the predicted value of C_d.

3. RESULTS AND DISCUSSION

According to the dimensional analysis, the dimensionless parameters affecting the symmetrical stepped labyrinth side weir are F_r, L/B, L/y_1, p/y_1, b_s/a_s, and b_s/L, respectively. The dimensionless parameters above are used as parameters for the inputs of ELM, KELM, and BELM, respectively, and C_d is used as the output parameter for the model. Randomly selected 70% (192) and 30% (83) of the model experimental datasets were used as the training and testing sets for all models, respectively. The performance of the ELM is mainly affected by neuron quantity, and when neurons are 20 in number, the model performs best. For the KELM model, the best performance is achieved when the kernel parameter $C = 1$, the regular coefficient $\lambda = 10$, and the weights and thresholds do not change randomly. For the BELM model, as shown in Figure 3, firstly, some hyperparameters are randomly tried to obtain their outputs, and then the next observation point is obtained in the search space, and finally, the minimum feasible point of the model is obtained. At this time, the hyperparameters $k = 1.23$, $\lambda = 830.07$.

3.1. Model performance comparison

Tables 1 and 2 show the evaluation indicators for all the models in the training and testing phases. The smaller the values of MAPE, RMSE, and bias, the higher the values of R^2, WIA, NSE, and VAF, indicating that the model has less error and better computing efficiency and accuracy. The ELM model has the worst performance in all evaluation indicators. The KELM has slightly better performance indicators than the ELM. Among them, in the training phase, all evaluation indicators are better than the ELM. In the testing phase, the values of WIA and VAF are very close to those of the ELM model, which indicates that the prediction abilities of the ELM and KELM models are close to each other. For the BELM model, compared with KELM, the MAPE is reduced by about 8.95 and 50.77% in the training and testing phases, respectively, which indicates that the hyperparameters obtained through Bayesian network optimization are suitable for the model, and improved modeling accuracy. The value of VAF is significantly better than that of KELM and ELM, indicating that the prediction ability and generalization ability of BELM are higher than those of KELM and ELM.

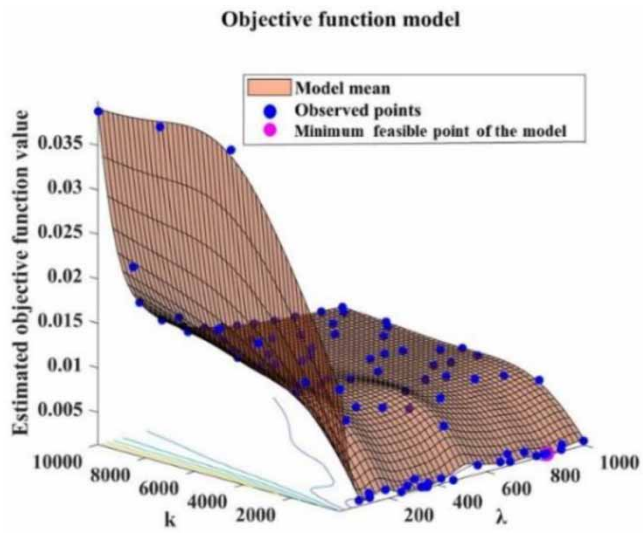

Figure 3 | Bayesian network performance diagram.

Table 1 | Performance indicators for all models in the training phase

Model	MAPE	RMSE	R^2	WIA	NSE	VAF	Bias
ELM	3.164	0.036	0.957	0.989	0.956	95.733	0.0013
KELM	2.815	0.035	0.963	0.990	0.961	96.140	0.0012
BELM	2.563	0.024	0.990	0.998	0.990	99.027	0.0003

Table 2 | Performance indicators for all models in the testing phase

Model	MAPE	RMSE	R^2	WIA	NSE	VAF	Bias
ELM	3.701	0.043	0.942	0.985	0.941	96.258	0.0019
KELM	3.455	0.042	0.943	0.984	0.941	94.272	0.0018
BELM	1.701	0.023	0.981	0.995	0.979	98.035	0.0006

Figure 4 shows the scatter fitting plots for ELM, KELM, and BELM in the training and testing phases. As can be seen from Figure 4, the experimental and predicted values for BELM are closer to the 1:1 line. The R^2 values of BELM are greater than those of ELM and KELM in the training and testing phases, indicating that the predicted values of the discharge coefficient obtained by BELM have less error with the experimental values, and the fitting ability of the function is significantly better than that of KELM and ELM in the running process.

To further visualize the comprehensive performance for different models. Taylor diagram can statistically quantify the similarity between different models. In the Taylor diagram, the three different evaluation indexes of standard deviation, RMSE, and correlation coefficient are represented by one point in the two-dimensional space, and a model that is closest to the target point is the best. It can be seen from Figure 5 that the black line represents the standard deviation, the green line indicates the RMSE and the blue line indicates the correlation coefficient. In the training and testing phase, the points representing the BELM are closer to the target points. Therefore, from the perspective of visualization, BELM is the best model.

3.2. Model uncertainty analysis

Tables 3 and 4 show uncertainty analysis results for all models. The mean error values for all the models are negative, indicating that all the models are underestimated in the calculation process of C_d of the symmetric stepped labyrinth side weir, but in the training and testing phases, the average error values are very smaller, and it can be considered that the models are more stable. In addition, the BELM has the smallest average error and the narrowest confidence bandwidth at the same confidence level. Meanwhile, it can be seen that the BELM has the smallest range of uncertainty widths in different phases, and the range of the 95% confidence intervals is also significantly smaller than that of the ELM and the KELM.

3.3. Sensitivity analysis for the parameters

In summary, it can be seen that the BELM has the best performance, and it can be used as the C_d prediction model for a symmetric stepped labyrinth side weir. However, six dimensionless parameters affect the discharge coefficient. Sensitivity analysis for the input parameters of the BELM model is performed by the Sobol method. For comparing the influence on

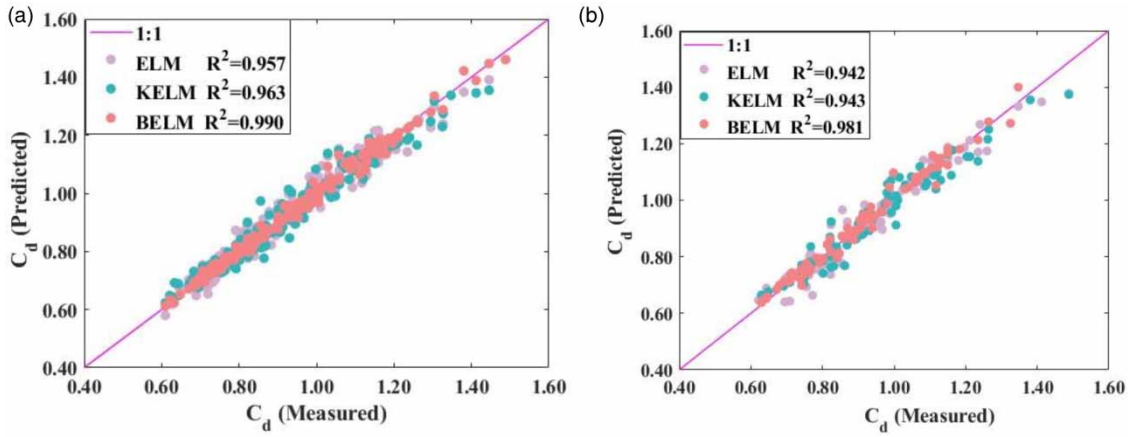

Figure 4 | Scatter plot of all models (a) Training phase (b) Testing phase.

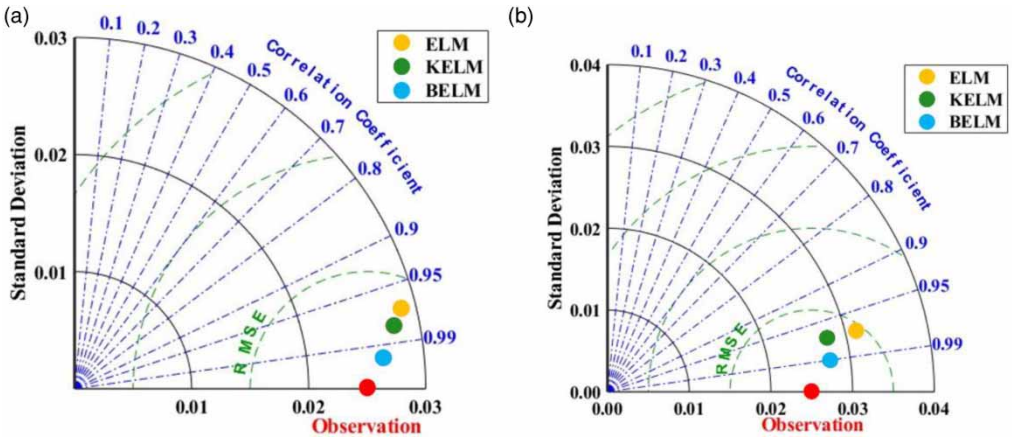

Figure 5 | Taylor diagram for ELM, KELM, and BELM: (a) training phase and (b) testing phase.

Table 3 | Uncertainty analysis results for all models in the training phase

Model	Mean error	Standard error	Bandwidth	95% confidence interval	
ELM	−0.0020	0.036	0.070	−0.070	0.070
KELM	−0.0012	0.035	0.068	−0.069	0.067
BELM	−0.0002	0.017	0.034	−0.034	0.034

Table 4 | Uncertainty analysis results for all models in the testing phase

Model	Mean error	Standard error	Bandwidth	95% confidence interval	
ELM	−0.0051	0.044	0.086	−0.091	0.081
KELM	−0.0063	0.042	0.082	−0.089	0.076
BELM	−0.0045	0.024	0.048	−0.055	0.040

sensitivity analysis results from different sampling methods, firstly, the sampling method is verified. In this study, Halton sequence (Ermakov 2017), and Sobol sequence (Wang *et al.* 2022) are used. Halton sequence is a quasi-random low difference sequence, which uses different bases on each dimension, and the bases on each dimension must be coprime, resulting in Halton sequence. Sobol sequence is a numerical sequence method for generating efficient uniformly distributed values, which uses a set of predefined dummy points, which are referred to as the generating elements. Each value in the sequence is computed by bitwise operations in combination with the original polynomial. Figure 6 shows the sampling results of the two sampling methods for sample size $N = 1{,}000$ with two-dimensional sampling and value space [0,1]. It can be seen that the Sobol sequence obtains a more uniform sampling distribution.

Figure 7 shows the sensitivity coefficient of dimensionless parameters calculated by the Sobol method for the BELM model. It is observed that the coefficient of first-order sensitivity S_1 of F_r, L/B, L/y_1, p/y_1, b_s/a_s, and b_s/L are 0.35503, 0.0276, 0.00242, 0.54203, 0.05110, and 0.00100, respectively. It means that $p/y_1 > F_r > b_s/a_s > L/B > L/y_1 > b_s/L$ when only considering the influence of individual parameters on the model output C_d. The global sensitivity coefficients of F_r, L/B, L/y_1, p/y_1, b_s/a_s, and b_s/L are 0.36082, 0.03467, 0.01622, 0.55653, 0.05822, and 0.00100, respectively. The results show that when considering the influence of individual parameters on the model output parameter C_d after interacting with other parameters, $p/y_1 > F_r > b_s/a_s > L/B > L/y_1 > b_s/L$. Tang *et al.* (2007) used 0.01 as the sensitivity threshold of the Sobol method and considered that when the sensitivity of the parameter is greater than 0.01, it is a sensitive parameter, and when it is less than 0.01, it is an insensitive parameter. Therefore, p/y_1 had the greatest impact on C_d, accounting for 55.57 and 54.17% under single and interaction, respectively. The effect of b_s/L on C_d is very small and can be ignored.

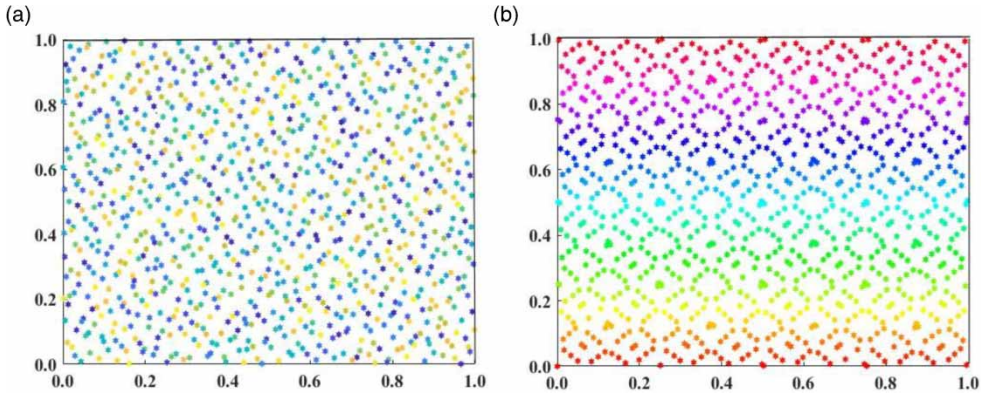

Figure 6 | Data point distribution from different sampling methods: (a) Halton sequence and (b) Sobol sequence.

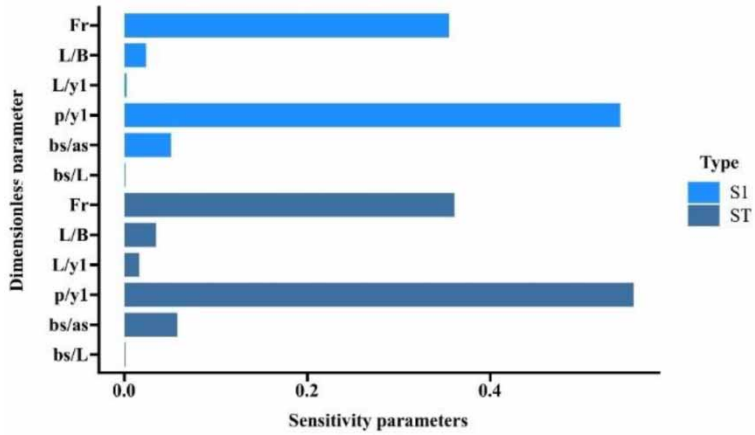

Figure 7 | Sensitivity coefficients for input parameters.

Figure 8 depicts the relationship between C_d and p/y_1 for different geometric parameters. It can be seen that C_d increases with the increase of p/y_1. In particular, when the number of steps is low, the discharge coefficient is larger. Meanwhile, the smaller θ is, the larger the corresponding C_d value is. As the θ increases, the trend of C_d increase also decreases for the same p/y_1, when $N = 10$, the corresponding C_d value is greater than or close to $N = 8$. This may be because as the N increases, the vortex above the step and the overflow nappe collide with each other, increasing the secondary flow along the lateral flow, and making the discharge coefficient larger.

4. CONCLUSION

The flow characterization of this structure is complicated due to the interaction of multiple factors affecting the symmetric stepped labyrinth side weir. In this study, the discharge coefficient estimation model was developed by ELM and Bayesian Network to comprehensively evaluate the stability and generalization ability from multiple perspectives. Also, uncertainty and sensitivity analyses of the model are analyzed, and the relationship between the variations of the geometric parameters. The following conclusions are obtained:

(1) The model performance of BELM in evaluating the C_d for the symmetric stepped labyrinth side weir is better than that of ELM and KELM, which can be used as the discharge coefficient prediction model for this structure. In the testing phase, Bias = 0.0006, $R^2 = 0.981$, with the smallest average error and 95% confidence interval range [−0.055, 0.040], indicating that the model has robust performance, data consistency, and high confidence.

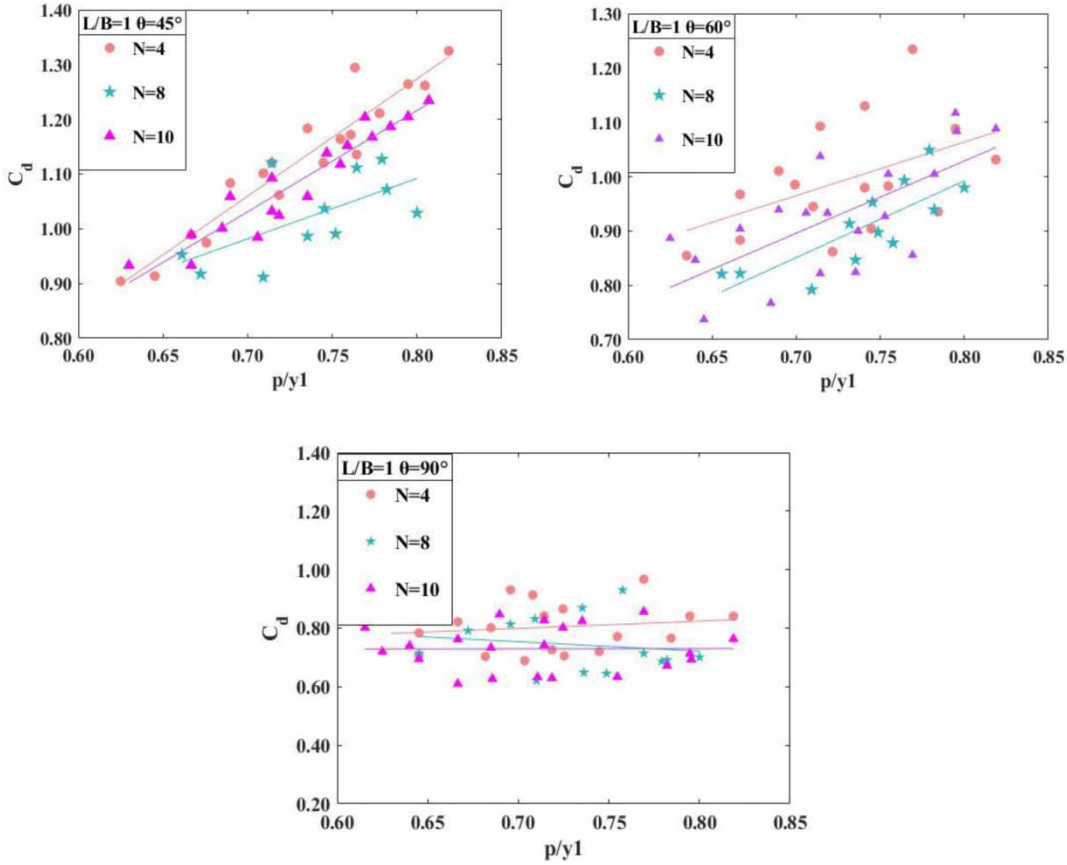

Figure 8 | The relationship between C_d and p/y_1.

(2) The global sensitivity indicates that the first-order sensitivity coefficient S_1 for F_r, L/B, L/y_1, p/y_1, b_s/a_s, and b_s/L are 0.35503, 0.0276, 0.00242, 0.54203, 0.05110, and 0.00100, respectively. Global sensitivity coefficient S_T is 0.36082, 0.03467, 0.01622, 0.55653, 0.05822, and 0.00100, respectively. Therefore, the influence of each dimensionless parameter on the C_d is $p/y_1 > F_r > b_s/a_s > L/B > L/y_1 > b_s/L$.

(3) C_d increases with increasing p/y_1. In particular, the discharge coefficient is larger when the number of steps $N = 4$. Also, the smaller internal head angle $\theta = 45°$ corresponds to a larger value of C_d, and the tendency of increasing C_d decreases as the internal angle increases.

In summary, the dataset for this study is small, and only flow characteristics under subcritical flow are analyzed. Secondly, this paper is carried out under the condition of clear water and does not involve the change of flow characteristics of the structure when the flow contains sediment. In future research, it can be further studied.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant Nos. 52079122, 52379080).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Akbari, M., Salmasi, F., Arvanaghi, H., Karbasi, M. & Farsadizadeh, D. 2019 Application of Gaussian process regression model to predict discharge coefficient of gated piano key weir. *Water Resources Management* **33**, 3929–3947.

Alfatlawi, T. J. & Alkafaji, R. D. M. 2023 Effect of anti-vortex structures installed on stepped labyrinth side weirs on discharge capacity. *Flow Measurement and Instrumentation* **89**, 102307.

Alfatlawi, T. J., Hashem, T. & Hasan, Z. H. 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.

Azma, A., Tavakol Sadrabadi, M., Liu, Y., Azma, M., Zhang, D., Cao, Z. & Li, Z. 2023 Boosting ensembles for estimation of discharge coefficient and through flow discharge in broad-crested gabion weirs. *Applied Water Science* **13** (2), 45.

Ben Said, M. & Ouamane, A. 2022 Performance of rectangular labyrinth weir – an experimental and numerical study. *Water Supply* **22** (4), 3628–3644.

Bijanvand, S., Mohammadi, M., Parsaie, A. & Mandala, V. 2023 Modeling of discharge in compound open channels with convergent and divergent floodplains using soft computing methods. *Journal of Hydroinformatics* **25** (5), 1713–1727.

Borghei, S. M., Jalili, M. R. & Ghodsian, M. 1999 Discharge coefficient for sharp-crested side weir in subcritical flow. *Journal of Hydraulic Engineering* **125** (10), 1051–1056.

Dogan, Y. & Kaya, N. 2023 The effects of changing the effective crest length of labyrinth side weir on discharge capacity. *Arabian Journal for Science and Engineering* **48** (4), 5289–5304.

Emami, H., Emami, S. & Parsa, J. 2022 A Walnut optimization algorithm applied to discharge coefficient prediction on labyrinth weirs. *Soft Computing* **26** (22), 12197–12215.

Ermakov, S. M. 2017 On randomization of Halton quasi-random sequences. *Vestnik St. Petersburg University, Mathematics* **50**, 337–341.

Huang, G. B., Zhou, H., Ding, X. & Zhang, R. 2011 Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **42** (2), 513–529.

Jalil-Masir, H., Fattahi, R., Ghanbari-Adivi, E., Asadi Aghbolaghi, M., Ehteram, M., Ahmed, A. N. & El-Shafie, A. 2022 An inclusive multiple model for predicting total sediment transport rate in the presence of coastal vegetation cover based on optimized kernel extreme learning models. *Environmental Science and Pollution Research* **29** (44), 67180–67213.

Karami, H., Karimi, S., Bonakdari, H. & Shamshirband, S. 2018 Predicting discharge coefficient of triangular labyrinth weir using extreme learning machine, artificial neural network and genetic programming. *Neural Computing and Applications* **29**, 983–989.

Karimi, M., Ghazizadeh, M. J., Saneie, M. & Attari, J. 2019 Flow characteristics over asymmetric triangular labyrinth side weirs. *Flow Measurement and Instrumentation* **68**, 101574.

Luo, M. & Zhang, K. 2014 A hybrid approach combining extreme learning machine and sparse representation for image classification. *Engineering Applications of Artificial Intelligence* **27**, 228–235.

Mahmoud, A., Yuan, X., Kheimi, M. & Yuan, Y. 2021 Interpolation accuracy of hybrid soft computing techniques in estimating discharge capacity of triangular labyrinth weir. *IEEE Access* **9**, 6769–6785.

Mohammed, A. Y. & Golijanek-Jedrzejczyk, A. 2020 Estimating the uncertainty of discharge coefficient predicted for oblique side weir using Monte Carlo method. *Flow Measurement and Instrumentation* **73**, 101727.

Olyaie, E., Banejad, H. & Heydari, M. 2019 Estimating discharge coefficient of PK-weir under subcritical conditions based on high-accuracy machine learning approaches. *Iranian Journal of Science and Technology, Transactions of Civil Engineering* **43**, 89–101.

Parida, N., Mishra, D., Das, K. & Rout, N. K. 2021 Development and performance evaluation of hybrid KELM models for forecasting of agro-commodity price. *Evolutionary Intelligence* **14**, 529–544.

Pradeep, T. & Samui, P. 2022 Prediction of rock strain using hybrid approach of ANN and optimization algorithms. *Geotechnical and Geological Engineering* **40** (9), 4617–4643.

Pradeep, T., GuhaRay, A., Bardhan, A., Samui, P., Kumar, S. & Armaghani, D. J. 2022 Reliability and prediction of embedment depth of sheet pile walls using hybrid ANN with optimization techniques. *Arabian Journal for Science and Engineering* **47** (10), 12853–12871.

Quilty, J., Jahangir, M. S., You, J., Hughes, H., Hah, D. & Tzoganakis, I. 2023 Bayesian extreme learning machines for hydrological prediction uncertainty. *Journal of Hydrology* **626**, 130138.

Roushangar, K., Alirezazadeh Sadaghiani, A. & Shahnazi, S. 2023 Novel application of robust GWO-KELM model in predicting discharge coefficient of radial gates: A field data-based analysis. *Journal of Hydroinformatics* **25** (2), 275–299.

Saffar, S., Safaei, A., Aghaee Daneshvar, F. & Solimani Babarsad, M. 2023 Flow-3D numerical modeling of converged side Weir. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 1–10.

Samadi, A., Salmasi, F., Arvanaghi, H. & Mousaviraad, M. 2022 Effects of geometrical parameters on labyrinth weir hydraulics. *Journal of Irrigation and Drainage Engineering* **148** (10), 06022006.

Seyedian, S. M., Haghiabi, A. H. & Parsaie, A. 2023 Reliable prediction of the discharge coefficient of triangular labyrinth weir based on soft computing techniques. *Flow Measurement and Instrumentation* **92**, 102403.

Shanshan, L., Guiying, S., Parsaie, A., Guodong, L. & Dingye, C. 2023 Discharge modeling and characteristic analysis of semi-circular side weir based on the soft computing method. *Journal of Hydroinformatics*, jh2023268.

Shen, G., Li, S., Parsaie, A., Li, G., Cao, D. & Pandey, P. 2022 Prediction and parameter quantitative analysis of side orifice discharge coefficient based on machine learning. *Water Supply* **22** (12), 8880–8892.

Simsek, O., Gumus, V. & Ozluk, A. 2023 Prediction of discharge coefficient of the trapezoidal broad-crested weir flow using soft computing techniques. *Neural Computing and Applications* **35**, 17485–17499.

Sobol, I. M. 1990 On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovanie* **2** (1), 112–118.

Soria-Olivas, E., Gomez-Sanchis, J., Martin, J. D., Vila-Frances, J., Martinez, M., Magdalena, J. R. & Serrano, A. J. 2011 BELM: Bayesian extreme learning machine. *IEEE Transactions on Neural Networks* **22** (3), 505–509.

Subramanya, K. & Awasthy, S. C. 1972 Spatially varied flow over side-weirs. *Journal of the Hydraulics Division* **98** (1), 1–10.

Tang, Y., Reed, P., Van Werkhoven, K. & Wagener, T. 2007 Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis. *Water Resources Research* **43** (6), W06415.

Wang, Q., Li, H., Lu, L., Yang, L. & Wang, S. 2022 Global sensitivity analysis of earth-moon transfer orbit parameters based on Sobol method. *International Journal of Aerospace Engineering* **2022** (1), 22.

Yan, P. 2023 Hydraulic characteristics of a maze weir spillway in Kenya. *Journal of the Yangtze River Academy of Sciences* **40** (8), 92–96.

Zaji, A. H., Bonakdari, H., Khameneh, H. Z. & Khodashenas, S. R. 2020 Application of optimized artificial and radial basis neural networks by using modified genetic algorithm on discharge coefficient prediction of modified labyrinth side weir with two and four cycles. *Measurement* **152**, 107291.

Zare, H., Vaghefi, M., Mahmoudi, A. & Behroozi, A. M. 2023 Experimental exploration of flow hydraulics and discharge coefficient for an inclined circular labyrinth weir. *Water Resources Management* **37** (11), 4521–4536.

Zounemat-Kermani, M., Kermani, S. G., Kiyaninejad, M. & Kisi, O. 2019 Evaluating the application of data-driven intelligent methods to estimate discharge over triangular arced labyrinth weir. *Flow Measurement and Instrumentation* **68**, 101573.

First received 5 November 2023; accepted in revised form 12 December 2023. Available online 21 December 2023