Water Supply

Water Supply Vol 23 No 9, 3775 doi: 10.2166/ws.2023.218

Estimating discharge coefficient of triangular free overfall using the GMDH technique

Ahmed Y. Mohammed [a,*] and Ali Sharifi [b]

[a] Department of Dams and Water Resources Engineering, University of Mosul, Mosul, Iraq
[b] Monetary and Banking Research Institute, Tehran, Iran
*Corresponding author. E-mail: ahmedymaltaee@gmail.com

AYM, 0000-0002-2274-9202; AS, 0000-0001-8313-7122

ABSTRACT

Free overfalls are hydraulic structures used in flood control, water supply, irrigation, and flow measurements. The hydraulic systems of free overfall depend on rectangular end shape. The studies that dealt with triangular crest are few and almost non-existent. In this study, a triangular end-shape design uses multiple linear regression (MLR) and group method of data handling (GMDH) methods for four models with six sub-models. Then, 24 scenarios were chosen and compared. The discharge coefficient (C_d) of a free overfall with a triangular terminal was predicted using experimental data. The triangular end edge shape increased crest length, the discharge coefficient, and discharge passing over free overfall. To this goal, 180 triangular free fall tests were performed. Data were collected for two triangular free overfalls with an opposite flow direction with three angles 600, 750, and 900. Results of C_d acquired using the two ways discussed above show that the algorithm GMDH outperforms the other method. Values for the GMDH approach mod46 testing variables: RMSE, MARE, SI, R^2, and NSE are 6.08×10^{-17}, 2.65×10^{-17}, 6.00×10^{-17}, 100.00%, and 1.00, respectively, while these values for MLR are 0.06332, 0.05970, 0.06624, 15.431%, and −3.0419, respectively. The GMDH technique shows the best results concerning MLR and then chooses the best four scenarios from 24 with a C_d percentage error not exceeding ±2%.

Key words: coefficient of discharge, free overfall, GMDH, MLR, open channel hydraulic

HIGHLIGHTS

- A new triangular end-shape design uses multiple linear regression (MLR) and group method of data handling (GMDH) methods for four models with six sub-models.
- Twenty-four scenarios were chosen and compared.
- The discharge coefficient (C_d) of a free overfall with a triangular terminal was predicted.
- The GMDH technique shows the best results concerning MLR and then chooses the best four scenarios from 24 with a C_d percentage error not exceeding ±2%.

LIST OF NOTATIONS

θ	the angle of the brink edge in rad
C_d	coefficient of discharge
C_1 and C_2	constants
$\bar{x}$ and $\bar{y}$	the mean both natural and examined C_d values, respectively
b	canal breadth (L)
C_c	coefficient of contraction
E	specific energy (L)
EDR	end depth ratio
Fr_b	Froude number
g	the acceleration of gravitational forces (L/T^2)
GMDH	group method of data handling
H	head over the standard weir (L)
L	crest breadth (L)
MLR	multiple linear regression

MARE	mean absolute relative error
NSE	Nash–Sutcliffe model efficiency coefficient
P	height of the fall (L)
$Q_{act.}$	actual discharge (L^3/T)
R^2	determinant coefficient
RMSE	root mean square error
SI	scatter index
V	speed of flow (L/T)
v_n	the flow normal velocity (L/T)
x_i and y_i	both natural and examined C_d values, respectively
y_b	brink depth (L)
y_{bcl}	brink depth at the centre line (L)
y_{brl}	brink depth at the left or right centre line (L)
y_n	the normal depth (L)
α	upstream face slope
μ	viscosity (M/LT)
ρ	the density of the water (M/L^3)

1. INTRODUCTION

Free overfalls or drops are a hydraulic structure widely used in an open channel when there are changes in channel bed levels; they can also use as a measurement device in irrigational channels. Many studies dealt with the hydraulics of free fall with different shapes. Beirami *et al.* (2006) studied free overfall theoretically for circular, semicircular, triangular, rectangular, and trapezoidal cross-sections to predict end depth ratio (EDR). The results refer to EDR values as 0.7016 in the rectangular shape, 0.8051 in the triangular shape, and 0.7642 in the exponential channel. Dey *et al.* (2004) investigated free overfall in inverted semicircular canals to predict EDR. The results showed that these values were near 0.695. Rajaratnam & Chamani (1996) and Dey & Kumar (2002) presented energy loss at drop and analysis of EDR in triangular free overfall in two methods. The first depends on the assumption of pseudo-uniform flow, and the second depends on assumed to be similar to the flow over a sharp-crested weir; the results show that the EDR value 0.695. Mohammed (2012) presented the theoretical study and showed equations for EDR and end depth discharge relationship for free overfall with different end shapes. Mohammed (2009) studied the hydraulic characteristics of triangular end lip free overfall with different angles, the EDR and water surface profile (WSP) were studied.

The theoretical data showed a result error of $\pm 5\%$ compared with the experimental. Some of the studies dealt with air entrainment from free overfall, such as Chanson (1993, 2004, 2007). In contrast, other studies refer to the turbulent diffusivity of the air bubble increasing with increasing distance from the point of singularity (Chanson & Toombes 1998; Gualtieri & Chanson 2004; Toombes & Chanson 2005). The studies dealt with erosion and scouring in free overfall and gave a technique for estimating discharge from known end depth equivalent sand roughness (van der Poel & Schwab 1985; Peter 1999). Ghodsian *et al.* (1999) investigate the effect of Froude number, total head to tailwater depth, and bed material size to tailwater depth parameters on scour backside of a free overfall and predicted equations to scour hole calculation. The effectiveness of free overfall roughness was studied by Dey (2000), Firat (2004), Guo *et al.* (2006), and Mohammed *et al.* (2011); these studies show that the EDR was substantially influenced by the spread, intensity, and distribution style of bed materials. The free overfall with various slopes and bed rough spread was investigated in Mohammed (2013), and the EDR and discharge equations with bed rough were predicted. The results showed that EDR values increased when bed rough increased. Other studies dealt with free overfall theoretical analysis (Ferro 1999) and using volume-of-fluid (VOF) (Sousa *et al.* 2009) as well as using artificial neural network (ANN) (Raikar *et al.* 2004; Öztürk 2005; Mohammed 2018); all these studies showed a good agreement between theoretical and experimental data. The discharge coefficient of the combined weir and gate structure was studied experimentally and theoretically by Balouchi & Rakhshandehroo (2018). The theoretical model was compared between multi-layer perceptron (MLP) and support vector regression (SVR) methods and given good agreement of C_d values calculated experimentally and theoretically. The ANNs and the M5P model tree application for calculating scour holes at river convergence were studied by Balouchi *et al.* (2015). The results showed that the MLP model was the most effective in the low scour depth range, while the radial basis function (RBF) model was more effective in the more excellent scour depth range, as measured by various statistical indices (RMSE, MAE, MARE, and R^2). Recently, there have been many studies in different fields dealing with group method of data handling (GMDH) to forecasting the behaviours of a complex system, such as Amanifard *et al.* (2008) and Kaveh *et al.* (2018) which is

studied shear strength predicted for fibre-reinforced plastic (FRP) reinforced concrete, Srinivasan (2008) and Najafzadeh *et al.* (2013), which they studied the GMDH in energy conservation and engineering geology. Ebtehaj *et al.* (2015) studied predicting the discharge coefficient in a side weir using GMDH by four input parameters then one output parameter predicted (C_d) experimentally and theoretically, their results lead to an acceptable degree of accuracy (R^2 0.779, MAPE 5.263, RMSE 0.038, E 0.757, and SI 0.069). The flowrate coefficient for the piano key lateral weir was investigated using the GMDH and developed group method of data handling (DGMDH) techniques by Mehri *et al.* (2019), the results of R^2, MAE, and RMSE values for the GMDH technique are 0.875, 0.0501, and 0.0792, while these values of the DGMDH technique are 0.938, 0.0372, and 0.053, this results given successfully of the DGMDH technique comparison with the GMDH technique. Teng *et al.* (2017) submitted a study for compared four methods, ANN, GMDH, multiple linear regression (MLR), and support vector machine (SVM), for forecasting China's transport energy demand. The results are compared according to R^2 and RMSE, and then the GMDH technique gives better performance compared with other techniques. The present study aims to a predicted coefficient of discharge experimentally and compares these results using MLR and GMDH, by using four models with six common algorithms that may be created, then there are 24 outputs. A group of GMDH models are used in this work, with the free overfall modified to triangular end edge shaped, which led to increase the edge length then increased the discharge coefficient and then increased discharge passing over free overfall.

2. METHODOLOGY

2.1. Experimental methodology

The experiment dose carries out in a rectangular facility channel 10 m long, 0.3 m broad, and 0.45 m deep. A typical thin weir (0.15*0.3*0.01) m was placed at the channel's terminus to measure the flow rate (Figure 1).

The free overfall made from wood in six models fixed at 6 m from the upstream channel entrance has dimensions (1*0.3*0.15) m, with triangular end lip inclined outside and inside at (60°, 75°, and 90°) (Figure 2).

Five values of upstream water depth were measured using a point gauge, and the discharges ranged from 7.3 to 16.5 l/s to the actual runoff. It was calculated using Equation (1):

$$Q_{\text{act.}} = 0.58 \times H^{1.5} \tag{1}$$

where $Q_{\text{act.}}$ is the real discharge (Sutopo *et al.* 2022), and H is the water depth over the weir (Widyastuti *et al.* 2022).

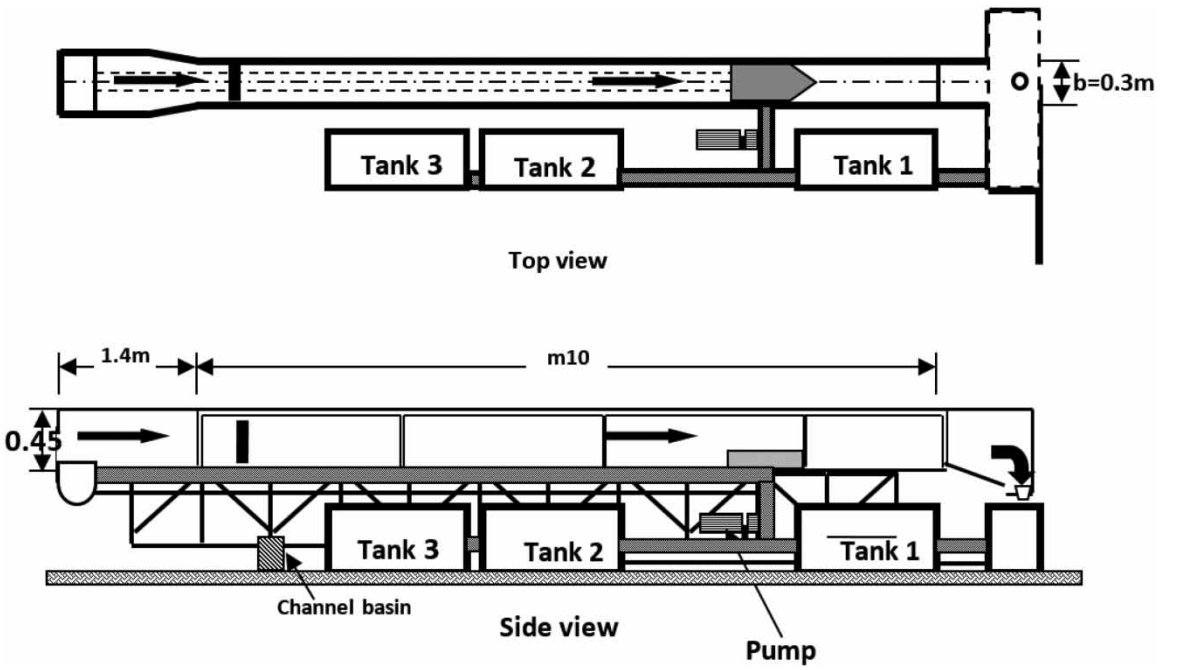

Figure 1 | The laboratory channel.

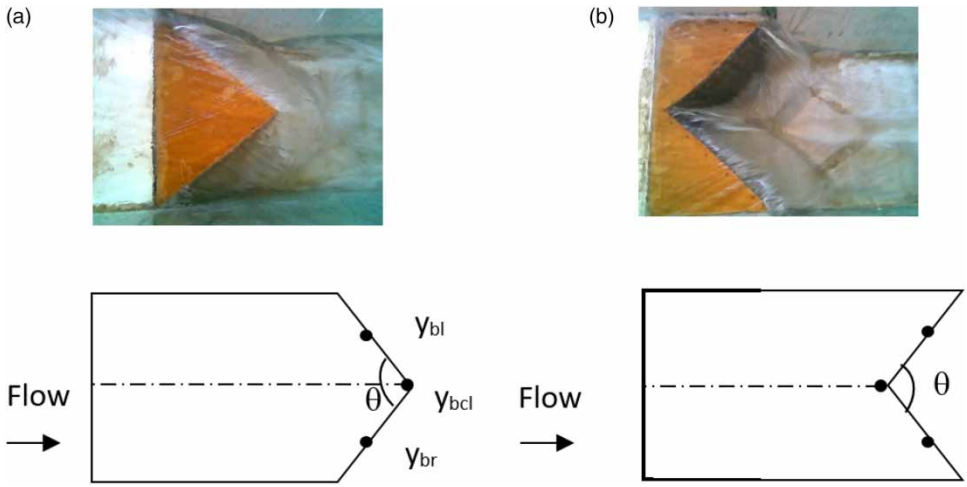

Figure 2 | The experimental modes of free overfall used in the present study. (a) The triangular edge of the notch concerning the flow path. (b) The triangular edge of the notch is angled in the reverse flow path.

The other parameters measured in this study are water depth over the brink of the free overfall (y_b) at channel centre (y_{bcl}), water depth at the mid-distance from the centre of the model referring to the angle of the template lip (y_{brl}), and normal water level over the free overfall (y_n).

First the free overfall was fixed with 60° end lip edge with flow direction (Figure 2(a)) and five flow discharges then measured typical depth, brink depth at centre, and brink depth left and right of the centre line, then another model in the opposite flow direction was fixed (Figure 2(b)) and measured the previous parameters, then repeat procedures with 75° and 90°.

2.2. Theoretical methodology

According to prior research by Ferro (1999), flow through an overfall in a rectangular channel is similar to flow over a thin weir for free rectangular overfall.

The velocity of flow is defined by applying the energy equation (Chow 1959; Ferro 1999):

$$H = E = y_n + \frac{V^2}{2g} \tag{2}$$

where E and H are specific energy and energy head; V is the velocity of flow; and g is the acceleration due to gravity.

The conception of the runoff over a thin weir assumes a zero-pressure variation and parallel streamlines at the edge, ignoring the nappe contraction.

The discharge dQ can be determined using the following continuity equation using a basic cross-section dA at a vertical height z (Chow 1959; Ferro 1999):

$$dQ = \sqrt{2g(H - z)}dA \tag{3}$$

The arrangement of zero pressure and uniform standardizes at the apex of the cliff is explained by these hypotheses; it was possible to integrate Equation (2) from the channel bed ($z = 0$) to the normal depth ($z = y_n$). Then, the runoff Q is evaluated as (Chow 1959; Ferro 1999):

$$Q = \int_0^{y_n} C_c\sqrt{2g}(H - z)^{1/2} b \cdot dz \tag{4}$$

By integrating Equation (3), the given equation may be written as follows (Chow 1959; Ferro 1999):

$$Q = \frac{2}{3} b \sqrt{2g} C_c [H^{3/2} - (H - y_n)^{3/2}] \tag{5}$$

where b is the breadth of the canal and C_c is the contraction coefficient.

2.3. Dimensional analysis

The factors that influence the flow of water at a free overfall are: y_b is the water depth over the edge (L); Q is flowrate (L^3/T); y_n is the consistent height of flow (L); P is the height of the fall (L); v_n is the normal velocity of flow (L/T); g is the acceleration due to gravity (L/T^2); μ is the viscosity in a dynamic state (M/LT); ρ is the water's density (M/L^3); and angle of the brink edge in rad.

Then the equation is written as

$$f(Q, y_n, y_b, v_n, P, L, g, \rho, \theta, \mu) = 0 \tag{6}$$

These non-dimensional parameters can be reduced to seven non-dimensional definitions shown in Equation (7) by utilizing Buckingham's argument with g, ρ, and y_b as the repeatable factors.

$$\frac{Q}{y_b^{3/2}\sqrt{g y_b}} = f\left(\frac{V}{\sqrt{g y_b}}, \frac{y_n}{y_b}, \frac{P}{y_b}, \frac{L}{y_b}, \frac{\mu}{\rho y_b v_n}, \theta\right) \tag{7}$$

where $V/\sqrt{g y_b}$ is Froude number (Fr) (Qasim *et al.* 2022); and $\mu/\rho y_b v_n$ is Reynolds number (Re).

Since the experiments dose conducted under a fully turbulent regime, the Reynolds number dropped, and Equation (7) became:

$$C_d = f\left(\mathrm{Fr}_b, \frac{P}{y_b}, \frac{L}{y_b}, \frac{y_n}{y_b}, \theta\right) \tag{8}$$

Besides, the Buckingham theorem allows us to use arithmetic operations to combine dimensionless constants; it is also possible to acquire particular new non-dimensional quantities, then Equation (8) can be described as

$$C_d = f\left(\mathrm{Fr}_b, \frac{y_n - P}{L}, \theta\right) \tag{9}$$

2.4. Modelling of discharge coefficient using MLR

Regression analysis is engaged in many applications. These applications' linear and non-linear analysis depends on the variables involved for instances of these equations obtained when more than one variable is used in the MLR model. Several experiments with several equation models were studied using the SPSS user guide to get a general equation for the coefficient of discharge in free overfall with a triangular edge.

Equation (9) produces an MLR with a correlation coefficient 0.392 using several SPSS models:

$$C_d = C_1 + \left(\frac{\mathrm{Fr}_b + \theta}{(y_n - P)/L}\right)^{C_2} \tag{10}$$

where C_1 and C_2 are constants, and θ is measured in radians.

$$C_d = 0.002 + \left(\frac{\mathrm{Fr}_b + \theta}{(y_n - P)/L}\right)^{-0.021} \tag{11}$$

The percentage error for C_d values is calculated using Equation (11), and experimental values do not exceed 10%.

3. GMDH TECHNIQUE FOR FREE OVERFALL

This section details the results of the GMDH networks and the suggested GMDH model for the discharge coefficient of triangular free overfall. In GMDH, a model may be visualized as a succession of neurons linked to one another by quadratic and tri-quadratic polynomials, with each subsequent step resulting in the formation of a new neuron. Based on the input parameters, this combination might anticipate the output values. Therefore, utilizing the observed effective parameters of the triangular form free overfall, the primary goal is to find the output parameter C_d. Consequently, using 'n' parameters, a single output parameter C_d may be determined using Equation (12) (Najafzadeh & Barani 2011):

$$y = a_o + \sum_{i=1}^{s} a_i x_1^{j1} x_2^{j2} + \cdots + x_k^{jk} \tag{12}$$

where y is the output parameter (C_d) and a, x are the input parameters.

As case studies of the generalized architecture, six standard algorithms may be created:

(1) Only the middle components of the multilayered iterative algorithm (MIA) contain pairs.
(2) The middle and beginning factors occur in pairs in the relaxation iterative algorithm (RIA).
(3) CIA stands for combined iterative algorithm, which allows for pairing middle and starting factors.
(4) Iterative-combinational multilayered (MICA).
(5) Iterative-combinational relaxation (RICA).
(6) Combined iterative-combinational (CICA) (practically equivalent to GIA).

The 24 outputs group of the GMDH models (models 11–16, models 21–26, models 31–36, and models 41–46) compared to the regression models suggested by Bos (1985), Fritz & Hager (1998), and Goodarzi *et al.* (2012). The recommended GMDH network (best from all 24 groups) under consideration yielded a coefficient of determination of 92.168, 97.117, 100, and 100% for four GMDH models (mod11, mod21, mod36, and mod46) (Figure 3).

These four selected models can be expressed as

mod11 < - gmdh.combi(X = x, y = y, G = 2, criteria = "ICOMP")
mod21 < - gmdh.combi.twice(X = x, y = y, G = 2, criteria = "ICOMP")
mod36 < - gmdh.mia(X = x, y = y, prune = 150, criteria = "PRESS")
mod46 < - gmdh.mia(X = x, y = y, prune = 200, criteria = "PRESS")
where,
G = 2: polynomial degree.

3.1. GMDH external standard

Values:

Criteria = "PRESS"; Predicted Residual Error Sum of Squares. It calculates without processing a system per reference node and considering all data into a standard sample.

ICOMP: Index of Informational Complexity. Like PRESS, it calculates without the need to recalculate the system.

Prune:

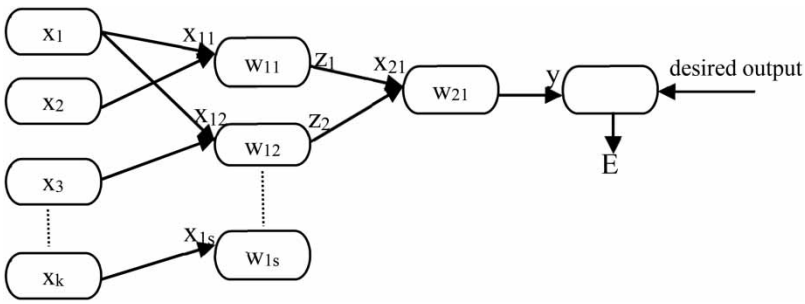

Figure 3 | Algorithm GMDH (Najafzadeh & Barani 2011).

The number of prior predictors is the suggested minimum value for this integer.

The amount of RAM available will determine the highest values. It advises working at the highest value, although it can be time-consuming and computationally.

From layer i, the number of neurons pruned to layer $i+1$ is the number chosen. The resultant layer $i+1$ contains prune (prune $-$ 1)/2 neurons; for example, if prune = 190 use, the causing neurons are 11.175. Table 1 shows the resulting layers used in all GMDH models. Given many fitted models to select the best algorithm and model with the highest fitting accuracy, Table 1 is the most concise description of the inner layers and internal functions used in the fits.

The flowchart methodology of the present study is shown in Figure 4.

4. RESULTS AND DISCUSSION

Table 2 describes the parameters for all models using the GMDH technique.

The formulas for computation free overfall before the coefficient of runoff studies are presented in Table 3.

Figure 5 shows line plots for the goodness of fit in best-selected models vs. the unbiasedness line and the MLR model (Residuals $\times 10^{06}$ sorted by observation). The excellent fit between models according to observation values and estrangement of the MLR values can be seen.

The residuals in Figure 6 show that mod36 and mod46 are almost on the unbiasedness line. To further explore and increase accuracy, the unbiasedness line examines two models, mod36 and mod46, separately by removing the other models.

According to Figure 6, the residuals of both models are minimal, and the choice between these two models depends on other conditions such as difficulty.

Figure 7 shows the relation between (C_d) observed from experimental data and that predicted using MLR and GMDH. The best four models used are mod (11, 21, 36, and 46). The figure observed that the error between observed values and GMDH models does not exceed $\pm 2\%$. At the same time, there are cleared unexpected data points for the MLR method, which prove that this method gives poor efficiency compared with the GMDH technique. Thereby, recommend the GMDH models usage generally to anticipate C_d of free overfall.

4.1. Indexes for results

Below is a calculation and definition of the subjective output of the accessible equations in terms of (R^2) determination coefficient, root mean square error (RMSE), scatter index (SI), and mean absolute relative error (MARE).

The determination coefficient is the proportion of the variance that is foreseeable from the independent parameters in the dependent variable, which usually ranges from 0 to 1.

$$R^2 = \left[\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) / \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2} \right]^2 \tag{16}$$

where x_i and y_i are the corresponding typical and investigated C_d values; $\bar{x}$ and $\bar{y}$ are the average of the measured amounts of C_d and the typical values, respectively.

RMSE is also used to calculate the discrepancy between a model's predicted values and those observed from the modelled thing. RMSE is one of the widely used statistics for error index is:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2} \tag{17}$$

The SI determines by dividing RMSE by the mean of the observations. Concerning the mean remark, it presents the percentage of RMSE difference or gives the rate of expected error with the argument is

$$\text{SI} = \frac{\text{RMSE}}{\bar{x}} \tag{18}$$

Table 1 | The layers and neurons of all models selected

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	Layer	Neurons																			
mod11																					
mod12																					
mod13																					
mod14																					
mod15																					
mod16																					
mod21	8																				
mod22	6																				
mod23	4																				
mod24	10																				
mod25	4																				
mod26	6																				
mod31	6	3	7	7	7	7	7														
mod32	14	3	12	41	102	102	102	102	102	102	102	102	102	102	102						
mod33	20	3	12	42	228	4,092	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,023	11,023	11,023
mod34	6	3	7	7	7	7	7														
mod35	12	3	12	42	102	102	102	102	102	102	102	102	102								
mod36	196	3	12	42	250	5,562	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	11,172	….			
mod41	4	3	3	3	3																
mod42	4	3	3	3	3																
mod43	16	3	12	42	228	4,092	19,897	…													
mod44	10	3	12	42	42	42	42	42	42	42	42										
mod45	74	3	12	42	250	1,222	…														
mod46	190	3	12	42	250	5,562	19,897	…													

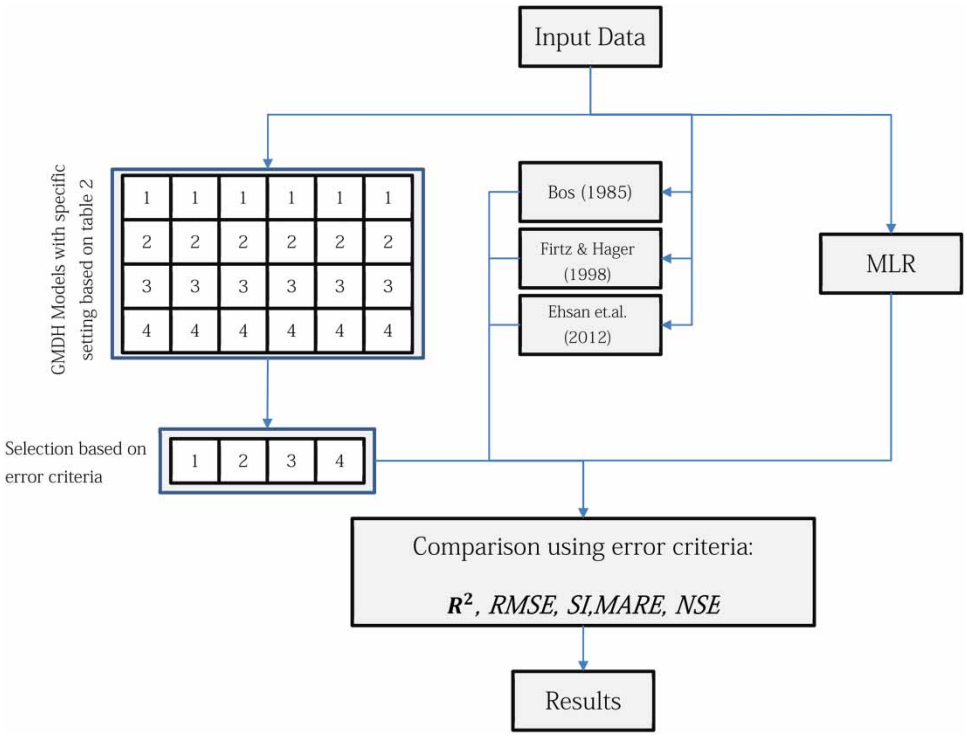

Figure 4 | The flowchart methodology of the present study.

In statistics, the MARE is a loss formula for artificial intelligence regression issues that use typically to quantify the validity of an estimation system's forecast. Usually, accuracy indicates as a ratio specified by the formula:

$$\text{MARE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - y_i|}{x_i} \tag{19}$$

The Nash–Sutcliffe efficiency relationship is employed to assess the predictive ability of hydrological analyses. Compared to the calculated data variance, it is mathematically validated that specifies the amount of the response values ('noise') and shows how well the plot of observed versus expected data matches the 1:1 axis.

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)}{\sum_{i=1}^{n}(x_i - \bar{x})} \tag{20}$$

$E = 1$ corresponds to an ideal combination of the expected coefficient of discharge with the data gathered; $E = 0$ denotes that the model is just as potent as the current data's average; and $-\infty < E < 0$ indicates that the identified median value is preferable than the model, meaning that the findings are undesirable.

Table 4 compares performance parameters and percentage error of GMDH, MLR model, and available equation of free overfall coefficient of discharge in literature.

GMDH and MLR models predicted outcomes that are adequately related to the accessible equations of C_d for free overfall. The qualitative performance of the present MLR has low RMSE (0.06332), MARE (0.05970), SI (0.06624), NSE (−3.0419), and R^2 (15.431%), respectively. The GMDH model has lower RMSE (6.08×10^{-17}), MARE (2.65×10^{-17}), S.I. (6.00×10^{-17}), higher NSE and R^2 (1.000) and (100.00%), respectively, this demonstrates that it performs better than other available predictors.

Table 2 | Describes the parameters of model 11 to model 46 for the GMDH model setting, count of fitted equations different according to models

Model name	Polynomial degree	Criteria	Prune
Build a regression model performing GMDH combinatorial			
Model 11	Original Ivakhnenko quadratic polynomial	ICOMP	
Model 12	Linear regression with interaction terms	ICOMP	
Model 13	Linear regression without quadratic and interaction terms	ICOMP	
Model 14	Original Ivakhnenko quadratic polynomial	PRESS	
Model 15	Linear regression with interaction terms	PRESS	
Model 16	Linear regression without quadratic and interaction terms	PRESS	
Create a linear regression using the GMDH twice-multilayered combinatorial (TMC) method			
Model 21	Original Ivakhnenko quadratic polynomial	ICOMP	8
Model 22	Linear regression with interaction terms	ICOMP	6
Model 23	Linear regression without quadratic and interaction terms	ICOMP	4
Model 24	Original Ivakhnenko quadratic polynomial	PRESS	10
Model 25	Linear regression with interaction terms	PRESS	4
Model 26	Linear regression without quadratic and interaction terms	PRESS	6
Create a proposed model using ActiveNeurons and the GMDH GIA (generalized iterative algorithm) (combinatorial algorithm)			
Model 31		ICOMP	6
Model 32		ICOMP	14
Model 33		ICOMP	20
Model 34		PRESS	6
Model 35		PRESS	12
Model 36		PRESS	196
Create an estimation technique for GMDH MIA (Multilayered Iterative Algorithm)			
Model 41		ICOMP	4
Model 42		ICOMP	4
Model 43		ICOMP	16
Model 44		PRESS	10
Model 45		PRESS	74
Model 46		PRESS	190

Table 3 | Formulas for computation of free overfall discharge coefficient in the prior lecturer

Authors	Equations	
Bos (1985)	$0.93 + 0.1*(12 - 15)/30C_d = 0.93 + 0.1\left(\left(\frac{h-P}{L}\right)\right)$	13
Fritz & Hager (1998)	$C_d = 1 - \left(\frac{2\sin(45)}{g\left(1 + \left(\frac{h-P}{L}\right)^4\right)}\right)^{0.5}$	14
Goodarzi *et al.* (2012)	$C_d = 1 + \frac{4.63\cos^{3/2}(\alpha)}{g\left(2.33 + \left(\frac{h-P}{L}\right)^4\right)}$	15

α is the upstream face slope; P is the weir height; l is the crest length, h is a depth of flow over the crest.

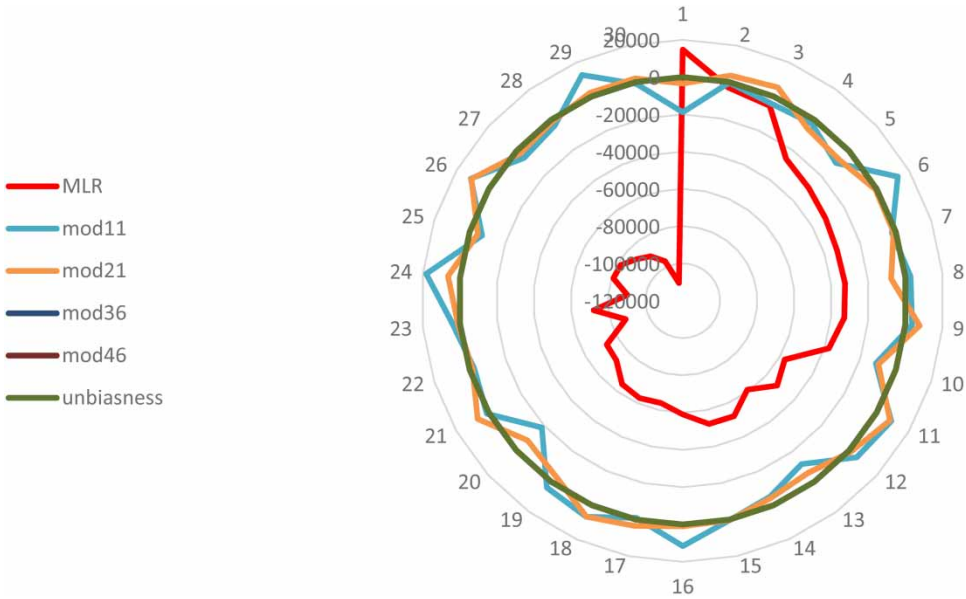

Figure 5 | The goodness of fit in best-selected models vs. observation and MLR models (Residuals $\times$ 10^6 sorted by observation).

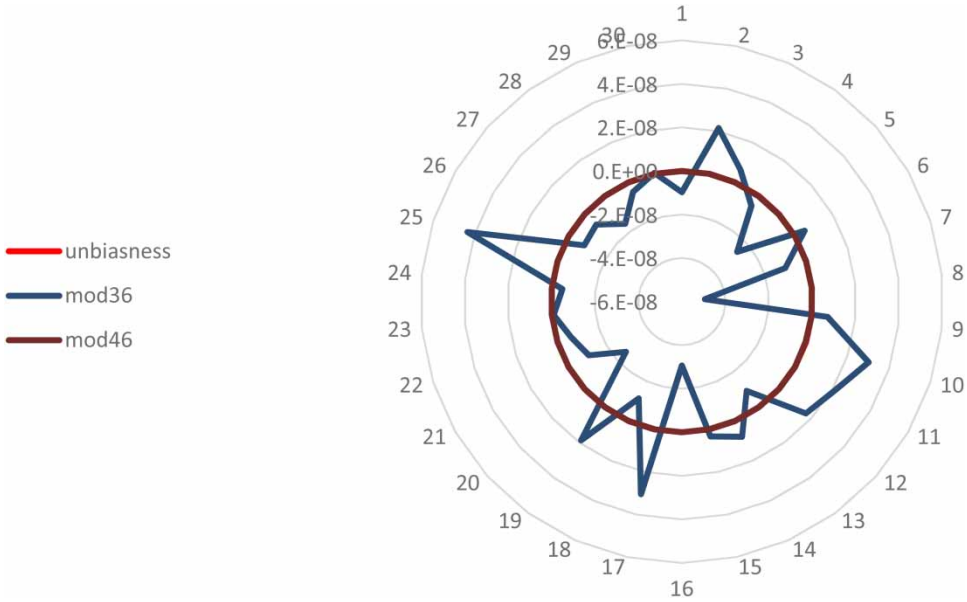

Figure 6 | The goodness of fit between mod36 vs. mod46 (Residuals $\times$ 10^6 sorted by observation).

Radar plot for the goodness of fit in selected models estimated from the present work vs. other models calculated from equations shown in Table 4 (Residuals $\times$ 10^6 sorted by observation values shown in Figure 8). The results refer to the existing agreement between observed values, and that data are calculated from theoretical methods (GMDH), and the best models 36 and 46, while model 46 took 28 h running. In contrast, model 36 took 18 h, so model 36 is recommended because it took less time and gave better values than other models.

5. CONCLUSION

This study conducted the experimental analysis of C_d for triangular end-shape free overfall using 180 collected data based on data collected. Numerical models (MLR and GMDH) apply to predicting outcomes. Data were collected for two models (with

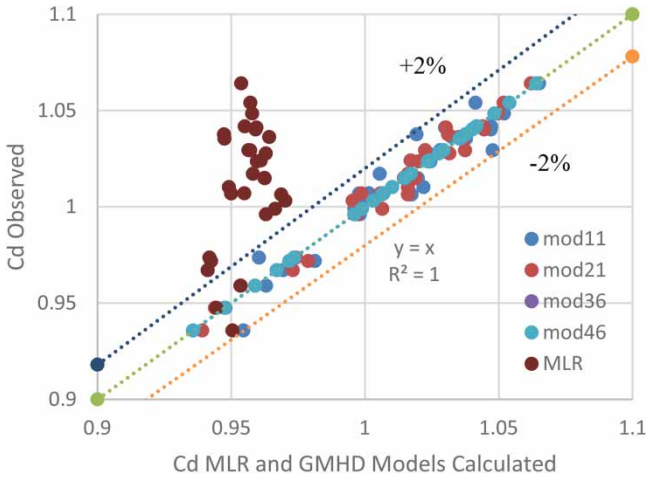

Figure 7 | An examination of C_d computed using GMDH and C_d observed.

Table 4 | Performance parameters of the coefficient of discharge for MLR and GMDH compared with previous literature

		Present (GMDH) results						
	Present (MLR) results	mod11	mod21	mod36	mod46	Bos (1985)	Fritz & Hager (1998)	Goodarzi *et al.* (2012)
RMSE	0.063	0.008	0.005	1.93×10^{-14}	6.08×10^{-17}	0.063	0.089	0.072
MARE	0.059	0.006	0.004	1.48×10^{-14}	2.65×10^{-17}	0.059	0.092	0.060
SI	0.066	0.008	0.005	1.90×10^{-14}	6.00×10^{-17}	0.066	0.096	0.066
R^2	15.431%	92.16%	97.11%	100.00%	100.00%	24.661%	33.834%	5.522%
NSE	−3.041	0.921	0.971	1.000	1.000	−3.036	−7.060	−4.237

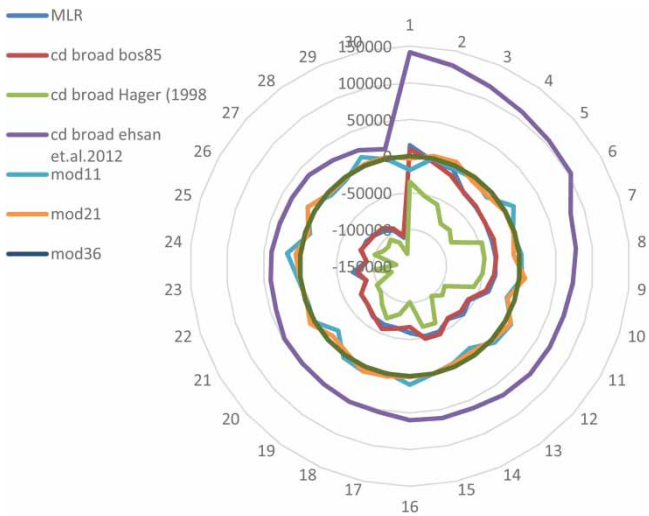

Figure 8 | Radar plot for the goodness of fit in selected models vs. other models (Residuals $\times 10^6$ sorted by observation values).

and opposite) flow directions with three different angles (60°, 75°, and 90°). Results show that RMSE, MARE, SI, R^2, and NSE values for testing data in GMDH technique mod36 are 1.93E-14, 1.48E-14, 1.90E-14, 100.00%, and 1.00, respectively, while these values for MLR are 0.06332, 0.05970, 0.06624, 15.431%, and −3.0419, respectively. From these results, the performance of the GMDH technique gives excellent indications for using this technique in the future because of its accuracy in calculating C_d. Even though the MLR model has an incredible ability to fit and predict, in contrast to the GMDH method, MLR has a much lower power in explaining the response variable. The results show that all four selected GMDH models have much higher levels of prediction. Model 46 took 28 h to run and gave the best results, while model 36 took 18 h, so model 36 was recommended because it took less time and gave better values than the other 24 scenarios. Also, the results show that the worst GMDH estimate has far better results than the best MLR.

ACKNOWLEDGEMENTS

A.Y.M. would like to thank the staff of the hydraulic laboratory in the dams and water resources department for their support in laboratory work.

AUTHOR CONTRIBUTIONS

All authors have read and agreed to the published version of the manuscript.

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

Amanifard, N., Nariman-Zadeh, N., Farahani, M. H. & Khalkhali, A. 2008 Modelling multiple short-length-scale stall cells in an axial compressor using evolved GMDH neural networks. *Energy Convers. Manage.* **49** (10), 2588–2594.

Balouchi, B. & Rakhshandehroo, G. 2018 Using physical and soft computing models to evaluate discharge coefficient for combined weir–gate structures under free flow conditions. *Iran. J. Sci. Technol. - Trans. Civ. Eng.* **42** (4), 427–438. doi:10.1007/s40996-018-0117-0.

Balouchi, B., Nikoo, M. R. & Adamowski, J. 2015 Development of expert systems for predicting scour depth under live-bed conditions at river confluences: application of different types of ANNs and the M5P model tree. *Appl. Soft Comput. J.* **34**, 51–59. doi:10.1016/j.asoc.2015.04.040.

Beirami, M. K., Nabavi, S. V. & Chamani, M. R. 2006 Free overfall in channels with different cross sections and sub-critical flow. *Iran. J. Sci. Technol. Trans. B Eng.* **30** (1), 97–105. Available from: www.SID.ir.

Bos, M. G. 1985 Required head loss over the structure. In: Long-Throated Flumes and Broad-Crested Weirs. Springer, p. 142, ISBN: 978-90-247-3113-8. https://doi.org/10.1007/978-94-009-6225-5.

Chanson, H. 1993 Stepped spillway flows and air entrainment. *Can. J. Civ. Eng.* **20** (3), 422–435. doi:10.1139/l93-057.

Chanson, H. 2004 Understanding air-water mass transfer in rectangular drop shafts. *J. Environ. Eng. Sci.* **3** (5), 319–330. doi:10.1139/S04-014.

Chanson, H. 2007 Air entrainment processes in a full-scale rectangular drop shaft at large flows. *J. Hydraul. Res.* **45** (1), 43–53. doi:10.1080/00221686.2007.9521742.

Chanson, H. & Toombes, L. 1998 Supercritical flow at an abrupt drop: flow patterns and aeration. *Can. J. Civ. Eng.* **25** (5), 956–966. doi:10.1139/l98-013.

Chow, V. T. 1959 Open Channel Hydraulics. McGraw-Hill, New York, p. 680.

Dey, S. 2000 End depth in steeply sloping rough rectangular channels. *Sadhana – Acad. Proc. Eng. Sci.* **25** (1), 1–10. doi:10.1007/BF02703802.

Dey, S. & Kumar, B. R. 2002 Hydraulics of free overfall in Δ-shaped channels. *Sadhana – Acad. Proc. Eng. Sci.* **27** (Part 3), 353–363. doi:10.1007/BF02703656.

Dey, S., Kumar, D. N. & Singh, D. R. 2004 End-depth in inverted semicircular channels: experimental and theoretical studies. *Nord. Hydrol.* **35** (1), 73–79. doi:10.2166/nh.2004.0006.

Ebtehaj, I., Bonakdari, H., Zaji, A. H., Azimi, H. & Khoshbin, F. 2015 GMDH-type neural network approach for modelling the discharge coefficient of rectangular sharp-crested side weirs. *Eng. Sci. Technol. Int. J.* **18** (4), 746–757. doi:10.1016/j.jestch.2015.04.012.

Ferro, V. 1999 Theoretical end-depth-discharge relationship for free overfall. *J. Irrig. Drain. Eng.* **125** (1), 40–44. doi:10.1061/(asce)0733-9437(1999)125:1(40).

Firat, C. E. 2004 *Abstract Effect of Roughness on Flow Measurements in Sloping Rectangular Channels with Free Overfall*.

Fritz, H. M. & Hager, W. H. 1998 Hydraulics of embankment weirs. *J. Hydraul. Eng.* **124** (9), 963–971. doi:10.1061/(asce)0733-9429(1998)124:9(963).

Ghodsian, M., Ardeshir, A. F. & Ali, A. A. 1999 *Scour Downstream of Free Over Fall Spillway*.

Goodarzi, E., Farhoudi, J. & Shokri, N. 2012 Flow characteristics of rectangular broad-crested weirs with sloped upstream face. *J. Hydrol. Hydromech.* **60** (2), 87–100. doi:10.2478/v10098-012-0008-1.

Gualtieri, C. & Chanson, H. 2004 *Analysis of air Bubbles Probability Distribution Functions in a Large-Size Dropshaft*.

Guo, Y., Zhang, L. & Zhang, J. 2006 *Numerical Simulation of Free Overfall in a Rough Channel*.

Kaveh, A., Bakhshpoori, T. & Hamze-Ziabari, S. M. 2018 GMDH-based prediction of shear strength of FRP-RC beams with and without stirrups. *Comput. Concr.* **22** (2), 197–207.

Mehri, Y., Soltani, J. & Khashehchi, M. 2019 Predicting the coefficient of discharge for piano key side weirs using GMDH and DGMDH techniques. *Flow Meas. Instrum.* **65**, 1–6. doi:10.1016/j.flowmeasinst.2018.11.002.

Mohammed, A. Y. 2009 Hydraulic characteristics of free overfall with triangular end lip. In *33rd IAHR Congress: Water Engineering for a Sustainable Environment*, January 2009.

Mohammed, A. Y. 2012 Theoretical end depth ratio and end depth discharge relationship for free overfall with different end lip shape. *Jordan J. Civ. Eng.* **6** (4), 410–417.

Mohammed, A. Y. 2013 Effect of bed roughness distribution and channel slope on rectangular free overfall. *AL-Qadisiya J. Eng. Sci.* **6** (2), 115–123.

Mohammed, A. Y. 2018 Artificial neural network (ANN) model for end depth computations. *J. Civ. Environ. Eng.* **8** (3). doi:10.4172/2165-784x.1000316.

Mohammed, M. Y., Al-Taee, A. Y. & Al-Talib, A. N. 2011 Gravel roughness and channel slope effects on rectangular free overfall. *Damascus Univ. J.* **27** (1), 47–54.

Najafzadeh, M. & Barani, G.-A. 2011 Comparison of group method of data handling based genetic programming and backpropagation systems to predict scour depth around bridge piers. *Sci. Iran.* **18** (6), 1207–1213. https://doi.org/10.1016/j.scient.2011.11.017.

Najafzadeh, M., Barani, G.-A. & Kermani, M. R. H. 2013 Abutment scour in clear-water and live-bed conditions by GMDH network. *Water Sci. Technol.* **67** (5), 1121–1128.

Öztürk, H. U. 2005 *Discharge Predictions Using ANN in Sloping Rectangular Channels with Free Overfall. A Thesis Submitted to the Graduate School of Natural and Applied Sciences of Middle East Technical University in Partial Fulfilment of the Requirements for the Degree of*.

Peter, S. 1999 *Ottenstein Dam – Scour at Stilling Basin Caused by Free Overfall Jet Confirmation of Calculation Model by In-Situ Measurement*.

Qasim, R. M., Mohammed, A. A. & Abdulhussein, I. A. 2022 An investigating of the impact of bed flume discordance on the weir-gate hydraulic structure. *HighTech Innov. J.* **3** (3), 341–355. doi:10.28991/HIJ-2022-03-03-09.

Raikar, R. V., Nagesh Kumar, D. & Dey, S. 2004 End depth computation in inverted semicircular channels using ANNs. *Flow Meas. Instrum.* **15** (5–6), 285–293. doi:10.1016/j.flowmeasinst.2004.06.003.

Rajaratnam, N. & Chamani, M. R. 1996 Energy loss at drops – reply. *J. Hydraul. Res.* **34** (2), 278.

Sousa, V., Bombardelli, F. A. & Chanson, H. 2009 Numerical simulation of rectangular drop shafts using a volume-of-fluid (VoF). In: *33rd IAHR Congr.* pp. 7009–7016.

Srinivasan, D. 2008 Energy demand prediction using GMDH networks. *Neurocomputing* **72** (1–3), 625–629.

Sutopo, Y., Utomo, K. S. & Tinov, N. 2022 The effects of spillway width on outflow discharge and flow elevation for the probable maximum flood (PMF). *Civ. Eng. J.* **8** (4), 723–733. doi:10.28991/CEJ-2022-08-04-08.

Teng, G., Xiao, J., He, Y., Zheng, T. & He, C. 2017 Use of group method of data handling for transport energy demand modelling. *Energy Sci. Eng.* **5** (5), 302–317. doi:10.1002/ese3.176.

Toombes, L. & Chanson, H. 2005 Air entrainment and velocity redistribution in a bottom outlet jet flow. In: *31st IAHR Congress 2005: Water Engineering for the Future, Choices and Challenges*. pp. 2716–2726.

van der Poel, P. & Schwab, G. O. 1985 Plunge pool erosion in cohesive channels below a free overfall. *Transactions of the ASAE* **31** (4), 1148–1153. (doi: 10.13031/2013.30836)

Widyastuti, I., Thaha, M. A., Lopa, R. T. & Hatta, M. P. 2022 Dam-break energy of porous structure for scour countermeasure at bridge abutment. *Civ. Eng. J.* **8** (12), 3939–3951. doi:10.28991/CEJ-2022-08-12-019.

First received 28 April 2023; accepted in revised form 11 August 2023. Available online 22 August 2023