Flow Discharge Prediction Study Using a CFD-Based Numerical Model and Gene Expression Programming

Sevda Mozaffari [1,*] , Erfan Amini [2] , Hossein Mehdipour [2] and Mehdi Neshat [3]

1 Department of Water Engineering, Urmia University, Urmia 57561, Iran
2 School of Civil Engineering, College of Engineering, University of Tehran, Tehran 13145, Iran; erfan.amini@ut.ac.ir (E.A.); hossein.mehdipour@ut.ac.ir (H.M.)
3 Center for Artificial Intelligence Research and Optimization, Torrens University Australia, Brisbane 4006, Australia; neshat.mehdi@gmail.com
* Correspondence: st_s.mozaffari@urmia.ac.ir

Abstract: The significance of spillways is to allow the flood to be safely discharged from downstream. There is a strong correlation between the poor design of spillways and the failures of dams. In order to address this concern, the present study investigates the flow over the Nazloo-ogee spillway using the CFD 3D numerical model and an artificial intelligence method called Gene Expression Programming (GEP). In a physical model, discharge and flow depths were calculated for 21 different total heads. Among different turbulence models, the RNG turbulence model achieved the maximum compatibility in computational fluid dynamic simulation. In addition, GEP was used to estimate Q, in which 70% of collected data was dedicated to training and 30% to testing. R^2, $RMSE$, and MAE were obtained as performance criteria, and the new mathematical equation for the prediction of discharge was obtained using this model. Finally, the numerical model and GEP outputs were compared with the experimental data. According to the results, the numerical model and GEP exhibited a high level of correspondence in simulating flow over an ogee-crested spillway.

Keywords: ogee spillway; flow prediction; numerical modeling; Flow-3D; CFD method; gene expression programming artificial intelligence model

1. Introduction

Flood control, reliable water supply, navigation, recreation, and hydroelectric power generation are the most important reasons for dam development around the world [1–4]. Spillways are hydraulic structures built on dams to convey excess flood flow beyond the dam's capacity. In order to avoid serious damage, spillways should be strong, reliable, and high-structures. Accordingly, design and construction play a role in dam spillways. Flood risks can be prevented if a dam spillway is properly designed and constructed [5–7]. Given their excellent hydraulic features, ogee-crested spillways are among the most investigated hydraulic structures. The U.S. Bureau of Reclamation, USBR, (1987) performed detailed laboratory experiments to investigate the behavior of water flow over a spillway, which resulted in the development and publication of spillway design manuals [8]. Physical studies have a number of problems, including high construction costs, the considerable time required for development and testing, and the presence of errors in results due to scaling effects [9–11]. The available computing capacity and algorithm improvements helped in finding solutions for a variety of problems, such as flow over the spillway. Numerical models provide a tool the fundamental design of spillways to identify operational concerns at a lower cost and in a shorter time [12–14]. Savage and Johnson investigated both pressure and discharge in the physical and numerical models in an ogee spillway, which indicated a satisfactory performance [15]. Kim and Parkn investigated the pressure and velocity distribution over an ogee spillway using the $RNG\ K - \epsilon$ turbulence model in the CFD model, taking into account the surface roughness effects [16]. Peltier et al. used

experimental modeling of an ogee spillway to estimate velocity and pressure for heads that were much higher than the design head and validated the data with experimental data [17]. Nowadays, numerous researchers have utilized the $K - \epsilon$ turbulence model in their studies, including (Jahad et al. [18]; Aydin et al. [19]; Wan et al. [20]). Moreover, $RNG\ K - \epsilon$ model has been applied in (Ghanbari and Heidarnejad [21]; Bayon et al. [22]; Valero et al. [23]), and remarkable accuracy in stimulating the overflow spillway. Artificial intelligence (AI) is a field of computer science that focuses on developing machines that can engage in intelligent behaviors. In recent years, machine learning techniques (e.g., gene expression programming (GEP), adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN_s), have been used in a wide range of publications. These methods are outstanding forecasting tools that have been used in a variety of civil engineering, hydraulic, and hydrological research in the last decade. Ferreira proposed gene-expression programming as a new adaptive approach for solving issues [24]. Yildiz et al. investigated discharge and flow depth over an ogee spillway using the ANFIS andCFD models in Flow-3D software. According to their results, therewas a reasonable agreement between the physical, numerical, and ANFIS artificial neural networks models [25]. The GEP technique has been used to represent a variety of water resource system components. Ebtehaj et al. utilized this method to predict the discharge coefficient in rectangular side weirs [26].

Using soft computing techniques, Kisi et al. examined the prediction of lateral outflow over triangular labyrinth side weirs under subcritical conditions [27]. Salmasi found a new equation for predicting discharge coefficients in an ogee weir, using gene expression programming and multiple regression techniques. The results demonstrated that the GEP technique was more successful than the regression equations [28]. Khan et al. employed a GEP method, to establish a functional relationship for bridge pier scour. The performanceof GEP was compared to that of other Al-based techniques, such as artificial neural networks (ANNs) and conventional reqression-based techniques [29]. Roushangar et al. used GEP and ANN techniques to estimate energy dissipation over stepped spillways, and the findings showed that GEP and ANN were extremely beneficial and encouraging in these circumstances [30]. Using GEP, Baxgatur and Onen created flood routing prediction models [31]. Bertonse et al. used GEP to simulate the concentration of dissolved oxygen (Do) in lakes [32]. New algorithms and models, particularly those based on soft computing, enable researchers to address the most complex systems in a number of ways [33,34]. In different engineering fields, the forecast methods that are not dependent on physics equations, including remote sensing methods and AI-based approaches likethe GEP method, are becoming more common [35,36]. The GEP has the advantage of being explicit in its formulation. This provides some insight into the nature of the phenomenon under investigation. It is simple to apply in reality [37]. The GEP is an artificial intelligence technique that utilizes key principles of genetic algorithms (GA) and genetic programming (GP) to create a calculation algorithm for forecasting a certain phenomenon. It mimics biological evolution [38]. Genetic programming is based on evolutionary principles developed for mathematical modeling. For solving regression and classification problems, GP provides several computer programs. The optimal values of some predefined parameters are obtained by GA, while finding both the best models and best parameters for a set of variables produced by GP is based on the Darwinian evolution theory [39]. There are also several other applications of such techniques in the literature that have been investigated insightfully in recent years [40,41].

A review of the literature indicates a lack of exclusive comprehensive research on the use of CFD and GEP models to estimate discharge values over an ogee-crested spillway with a pier. In the present study, the statistical performance of the model of an ogee-crested spillway with a pier was evaluated using error criteria, such as the root-mean-square error ($RMSE$), which is a significant criterion, as well as the determination coefficient (R^2) and mean absolute error (MAE) and a new mathematical equation was developed for predicting the spillway's discharge. The results of the GEP were then compared to those

obtained from the physical model. In addition to this analysis, knowledge of discharge and flow depth can be valuable for hydraulic engineering research as well as determining the GEP approach's capability of forecasting a new output variable. The CFD model was validated with the experimental and Flow-3D data. The results showed that all of the model's outputs overlap and are accurate when it comes to tackling fluid problems.

This paper consists of four sections. Section 2 contains a brief discussion of the methods and components utilized to simulate the ogee-crested spillway as well as the provided gene expression programming and a Flow-3D numerical model. Section 3 discusses the findings of the investigation regarding numerical and experimental data comparisons, and mathematical equations. Finally, Section 4 summarizes the paper's main findings.

2. Materials and Methods Methods

The head, the inclination of the upstream face of the overflow section, and the height of the overflow section above the entrance channel's floor determine the shape of the nappe profile over an ogee crest, which influences the velocity of approach to the crest. The general discharge equation for an ogee-crested spillway is as follows:

$$Q = C_o L H_o^{\frac{3}{2}} \tag{1}$$

where Q is the design discharge ($\frac{m^3}{s}$), C_o is the variable coefficient of discharge for free-flow condition ($\frac{m^{0.5}}{s}$), L is the effective length of the crest (m), H_o is the design head on the crest(m), including the velocity of approach, h_a (m). Figure 1 demonstrates the ogee profile.
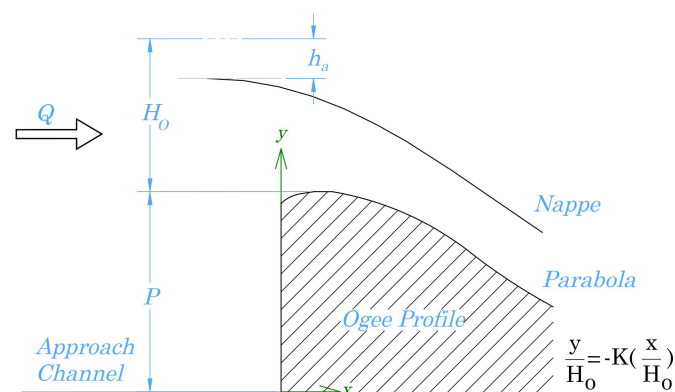

Figure 1. The ogee−crested spillway.

In this study, a physical model of a standard ogee spillway was developed at the Water Research Institute of the Ministry of Energy, Iran. On the west side of lake Urmia , on the Nazloo river, the Nazloo dam spillway is under construction with a width of 42 m, a crest height of 6 m, and a crest level of 1492 m above sea level. The hydraulic performance of the Nazloo spillway was studied using a physical model with a scale of 1:40 made of Plexiglas Figure 2. The effective highest measured at a distance of 4 m from the top of the crest of the water surface. The pier over the crest had a height of 3.51 m, a width of 2 m, and a length of 10 m.

Figure 2. Physical model of Nazloo Spillway.

Crest piers cause the flow to be constructed, reduce the effective length of the crest, and decrease the discharge compared to an uncontrolled crest. The shape and location of the pier nose are shown in Figure 3, as well as the thickness of the pier, the head relative to the design head, and the approach velocity, which all influence the pier construction coefficient, K_p [42].

$$L_e = L - 2(N * K_p + K_a)H_o \qquad (2)$$

where L_e is the effective length of the crest for calculatingthe discharge, L is the net clear length of the spillway crest, N is the number of piers, K_p is the pier construction coefficient, K_a is the abutment contraction coefficient and H_o is the total design head on the crest including velocity head. For round-nosed piers, the K_p and K_a are set to 0.01 and 0.05, respectively. Dimensional analysis is a fundamental tool used in experimental research to determine dimensionless quantities. For this purpose, characteristics of the effect of the ogee spillway discharge should be identified at first. Then, Buckingham's theory should be utilized to estimate the dimensionless parameters, which can be used to analyze their impact on the spillway's discharge and identify the logical correlation between them [21]. The geometric parameters and flow characteristics might be used to calculate the discharge capacity of an ogee spillway. The quantities influencing ogee crest discharge can be summarised as follows:

$$f(H_o, L_e, H_d, P, g, \mu, \rho, \sigma) = 0 \qquad (3)$$

where f is a functional symbol; H_e and H_o are the total upstream head and the design head, respectively; P is the ogee crest upstream spillway height; μ and ρ are dynamic viscosity and density, respectively; and σ is the surface tension. The possible effects of surface tension on discharge were small in all experiments $H > 30$ mm. The Weber number was excluded from the analysis. Furthermore, since the flow was turbulent, the viscosity impact was minor compared to gravity. As a result, the Reynolds Number effect was also removed from the analysis. In this study, the ultimate relationship between the dimensionless parameters impacting the ogee spillway discharge would be stated as:

$$Q = f(\frac{H_e}{H_o}, \frac{L_e}{H_o}, \frac{H_e}{p}) \qquad (4)$$

The experiments were performed with 21 total heads over the spillway, ranging from $\frac{H_e}{H_o} = 0.123$ to $\frac{H_e}{H_o} = 1.22$.
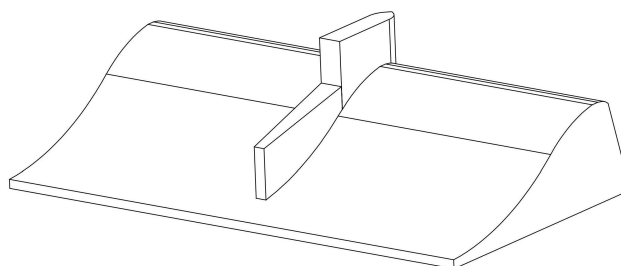
Figure 3. Isometric View of an ogee-crested.

2.1. Flow-3D Numerical Modeling

FLOW-3D software is a useful tool for analyzing complex fluid issues such as free-surface transient three-dimensional flow problems with complex geometries. This software uses the finite volume method along with regular rectangular grids. In order to simulate turbulence in hydraulic problems, two-equation models are commonly utilized. To derive time-averaged Reynolds equations, the Renormalization Group *RNG* turbulence model was used in this study. Furthermore, the numerical problems were solved using the FLOW-3D software, and the transient governing equations were numerically solved using the finite volume model. The geometry of the finite volume approach was defined in this software utilizing the Fractional Area-Volume Obstacle Representation (FAVOR) algorithm, as shown in Figure 4. The in-field obstacles in the computational cells were treated as a fractional value between 0 and 1. Therefore, if the entire cell was filled with obstacles, the fractional value of the area-value was equal to 1. The volume-of-fluid (VOF) algorithm was used to determine the flow's free surface [43].

For the considered field, the FLOW-3D numerical model provided a three-dimensional structural grid made up of cuboid cells as shown in Figure 5. As a result, a three-dimensional model was created in AutoCAD software to build up the geometry of the models with a stereolithography (STL) file, based on the laboratory models' specifications. The data was then loaded into FLOW-3D, where the gird was generated using VOF and FAVOR, and the boundaries and computational network were determined. The considered area was constructed using VOF and FAVOR methods after loading the geometric data into the software.
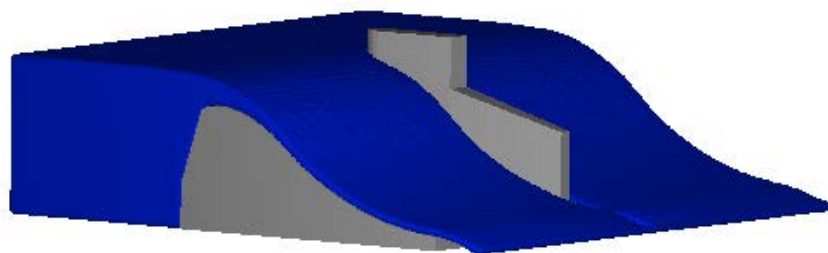

Figure 4. Depiction of the fractional area volume obstacle (FAVOR) approach that is applied to the solid contours.

The boundary conditions' most critical aspect is to create a flow situation similar to the physical status. Each type of boundary condition in the Flow-3D software can be used for the unique condition of the models. The X-minimum boundary condition, specified pressure, was applied in this study. An outflow was chosen for the X-maximum boundary condition since there was nothing to calculate at the flume's end. Both (Y-min) and (Y-max) were computed as symmetry, with the bottom (Z-min) as a wall boundary condition and the top (Z-max) as a symmetry boundary condition. The mesh sensitivity evaluation comes second. In order to identify and ensure the representation of all the phenomena involved, a computational mesh sensitivity analysis must be done. The mesh sizes were examined from the largest to the smallest in a trial-and-error approach. The model's precision was improved by fine-tuning the mesh size. A mesh size of 0.007 was

used for these simulations. The determination of the correct mesh plays an indispensable role in any numerical model simulation. It is important to minimize the number of cells due to having sufficient resolution as well as this mesh and cell size affect simulation time. The mesh and cell size were tested in a sensitivity analysis, and the 0.007 cell size performed best in terms of results and time spent solving equations.A further decrease in this value in cell size does not affect the accuracy of results for the Q tests. Different cell sizes were used in this study (e.g., 0.06, 0.02, 0.01 and 0.006). These were then reduced to 0.007 to provide more accurate results for the flow tests, as shown in Figure 6. After simulating several different models, the numerical model was set to 40 s, which is a sufficient time to obtain a stable result. The SI system was used to determine the length unit, and degrees Celsius was chosen as the temperature unit. The fluid database was utilized to choose water with a temperature of 20 °C and a viscosity of 10 pa s in the model. A review of the previous research indicated that a Renormalized group (RNG) turbulence model was appropriate for the numerical model. The numerical model functioned as a no-slip condition over the ogee's surface, No specific material characteristics were defined for the ogee spillway, and the roughness was 0.002.
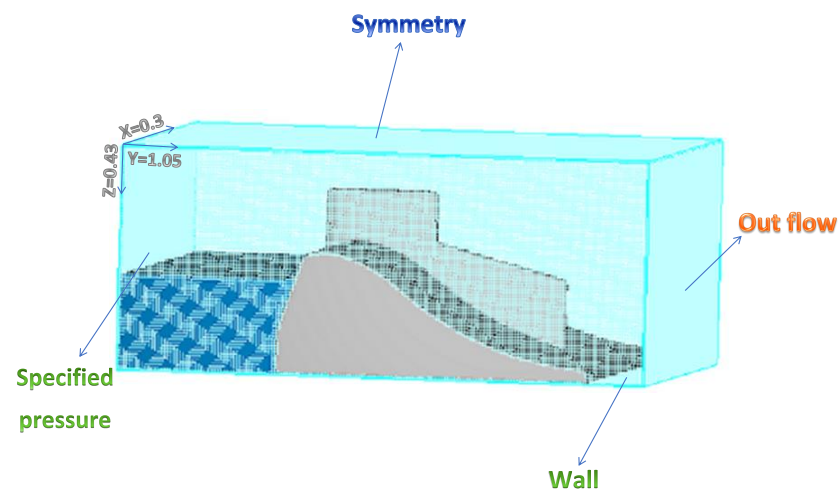

Figure 5. Boundary conditions and a three-dimensional structural grid model made up of cuboid cells.
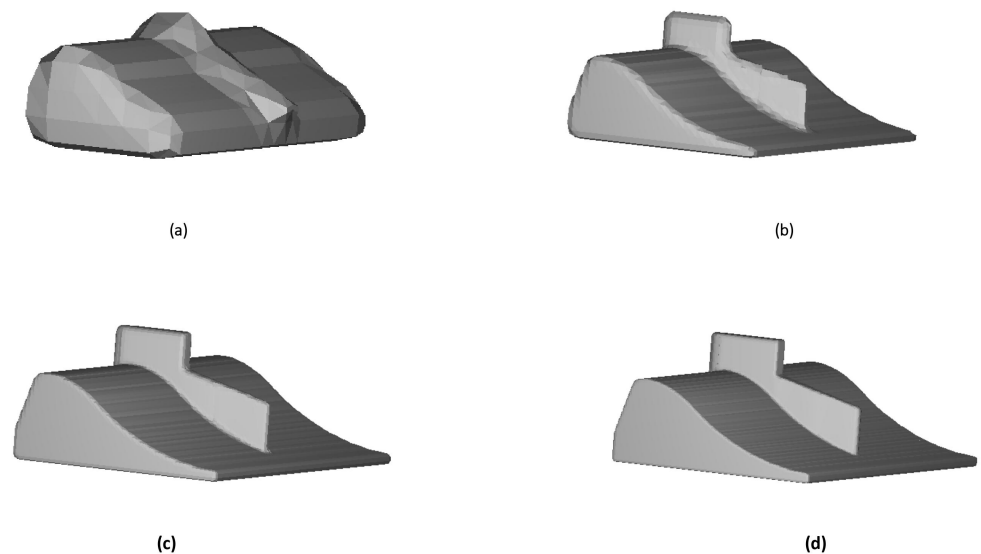

(a)

(b)

(c)

(d)

Figure 6. Different resolutions used for sensitivity analysis, (**a**)—cell size is 0.06, (**b**)—cell size is 0.02, (**c**)—cell size is 0.01, and (**d**)—cell size is 0.006.

Estimation of Uncertainty in a CFD Application

Computational Fluid Dynamics (CFD) analysis provided a more detailed evaluation of flow characteristics than an experiment. The complexity of the phenomenon makes computer modeling difficult, and its reliability must be evaluated [44]. In this study, error estimates such as an approximate relative error (e_a^{21}), an extrapolated relative error (e_{ext}^{21}) and the fine-grid convergence index (GCI_{fine}^{21}) were used to evaluate the CFD model.

$$e_a^{21} = \left|\frac{Q_1 - Q_2}{Q_1}\right| \tag{5}$$

$$e_{ext}^{21} = \left|\frac{Q_{ext}^{21} - Q_1}{Q_{ext}^{21}}\right| \tag{6}$$

$$GCI_{fine}^{21} = \frac{1.25 e_a^{21}}{r_p^{21} - 1} \tag{7}$$

The calculation procedure for three selected grids showed in Table 1. Where Q represents the discharge, *r* indicates the grid refinement factor, which is greater than 1.3 [45].

Table 1. Estimation of discretization error for three cell sizes (0.01, 0.018, 0.02).

	Values	Values	Values
N_1, N_2, N_3	1780000, 305212, 222500	1780000, 305212, 222500,	1780000, 305212, 222500
r_{21}	1.72	1.72	1.72
r_{32}	1.56	1.56	1.56
Q_1	5.43	10.29	11.805
Q_2	5.51	10.2	11.69
Q_3	5.37	10.08	11.58
P	2.48	1.79	1.2
Q_{ext}^{21}	5.38	10.34	11.93
Q_a^{21}	1.47%	0.87%	0.97%
Q_{ext}^{21}	0.5%	0.48%	1.04%
Q_{fine}^{21}	0.66%	0.67%	1.31%

2.2. Gene Expression Programming Approaches

The gene expression programming method , which is based on genetic programming and genetic algorithms, was created by Ferreira. GEP is a heuristic search and optimization technique that uses biological evolution to create computer software that can forecast a certain event [24].

One of the GEP's advantages is that it is multigenic, allowing for the creation of increasingly complex programs with many subprograms. From genetic algorithms, it inherited fixed-length linear chromosomes, as well as expressive parse trees of various sizes and shapes via genetic programming. The *GA* is a probabilistic search approach that is based on evolution in nature. The following provides the *GA*'s general procedure [46]: Step 1: a population is created by randomly selecting *n* chromosomes (potential problem solutions). Step 2: determine the compatibility of each chromosome (x) using the fitness function $[f(x)]$. Step 3: The following stages are iterated to create new ones. (i) (Selection): Based on their compatibility, two parent chromosomes are chosen from the population. (ii) (Crossover): Step 1's parent chromosomes intercross with a particular probability and create two children. (iii) (Mutation): The chromosomes of the created children mutate at a random rate with a given probability. (iv) The children generated as a result of the genetic operators (selection, crossover, and mutation) are added to the new population. Step 4: in the new generation, the parents of the created population are replaced. Step 5: the algorithm terminates and the existing population shows the desired response if favorable circumstances, such as the desired accuracy or the number of iterations stated in the

problem,are reached. Step 6: if the algorithm does not end in Step 5, Steps 2–6 are repeated until the desired results are reached. The chromosomes and Expression Trees (*ETs*) are two primary components of GEP. One or more genes representing a mathematical expression may be found on the chromosomes. A gene's mathematical code can be expressed in two languages: the language of genes and the language of Expression Trees (ET). The GEP genes are divided into two sections: the head and the tail. The head contains mathematical operators, variables, and constants $(+, -, *, /, \ln, exp, 1, a, b, c)$ used to represent a mathematical expression. Variables and constants $(1, a, b, c)$ are included as terminal symbols in the tail. Additional symbols are used if the terminal symbols in the head are insufficient to explain a mathematical equation. Starting at the top line of the tree and reading left to right, top to bottom, the translation of the expression tree is performed. This method's gene sequences are similar to biological gene sequences [31].

Three input parameters were used in the GEP model (flow head ratio, crest height, crest length). Table 2 shows the parameters of the GEP models. The fitness function used in this study was the root mean square error (*RMSE*)of the training set. The powerful soft computing software package GeneXproTools (Version 5.0.3926) was applied to generate GEP-based discharge prediction models. The program was run for a number of generations before being stopped since the fitness function value did not change. In the GEP model, five options were considered using the various operators listed in Table 3 , as well as the output results.

Table 2. Parameters of the GEP Model.

Setting Parameters	Value
Functions set	$+, -, /, x^2$, exp, ln, cube root, Atan, Tanh
Chromosomes	30
Head size	8
Number of genes	3
Linking function	Addition
Fitness function error type	RMSE
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

Table 3. Different GEP operators with errors of prediction for the Training and Testing datasets.

Type	Operator	Train Phasing			Test Phasing		
		R^2	RMSE	MAE	R^2	RMSE	MAE
Option 1	$+, -, *, /$	0.9	1.57	1.07	0.86	1.73	1.13
Option 2	$+, -, *, /, x^2$	0.91	1.53	1.05	0.86	1.73	1.12
Option 3	$+, -, *, /, x^2$, exp	0.9	1.65	1.03	0.86	1.59	1.01
Option 4	$+, -, *, /, x^2$, exp, ln, cube root	0.9	1.59	1.06	0.87	1.61	1.06
Option 5	$+, -, *, /, x^2$, exp, ln, cube root, Atan,Tanh, min, max	0.972	0.85	0.64	0.912	1.42	1.12

Performance Criteria

A number of evaluation indicators can be used to assess the created models' forecasting performance using statistical metrics of the goodness of fit [47]. In this study, Performance measures such as root mean square error (*RMSE*), determination coefficient (R^2) and mean absolute error (*MAE*) were used to evaluate the GEP model.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Q_p - Q_o)^2}{N}}$$

$$R^2 = \frac{\sum_{i=1}^{N}[(Q_o - \bar{Q}_o)(Q_p - \bar{Q}_p)]^2}{\sum_{i=1}^{N}(Q_o - \bar{Q}_o)^2(Q_p - \bar{Q}_p)^2} \quad (8)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Q_o - Q_p|$$

where N is the number of observations, Q_o indicates the observed data, Q_p indicates the predicted data, $\bar{Q}_o$ refers to observed data, and $\bar{Q}_p$ refers to the mean of the predicted data. The mean of the errors is represented by the RMSE, which varies from 0 to ∞, with lower values indicating better model performance. The simulation accuracy of the model is described by R^2, ranging from 0 to 1. The MAE is a comparison of two continuous variables. The average vertical distance between each point and the y = x line, commonly known as the one-to-one line, is called MAE.

3. Discussion

The numerical model had to be calibrated with experimental data in the first stage. The experimental investigation conducted at the Ministry of Energy's Water Research Institute, provided discharge values and flow depth that were compared with those of the numerical model evaluated by Flow-3D and GEP models.

To reach steady-state conditions, it is crucial to extract the exact values from the data of a physical or numerical model. In this research, after running numerous different models with the existing numerical model, the acceptable time to extract the results was determined to be 40 s. There is no separation between the flow and the crest bottom, demonstrating the solid boundary condition's true simulation.

Baffles are two-dimensional surfaces that come in a variety of shapes, including planes, cylinders, cones, and spheres. They come in both porous and non-porous varieties, and they may be used to quantify mass flows, heat streams, and applied forces. The number of elements in the computational mesh depends on the transit flow, which is directly acquired from the section where the baffle is positioned [48]. The flow rate was measured using baffle, and the X-minimum boundary condition was used with a specified pressure. The design head was set to 0.19 m, and the design discharge coefficient and discharge were calculated at 2.15 and 0.0175 m^3/s, respectively, based on the scale of 1:40. For all data, the physical and numerical models are in good agreement. To make a more objective comparison between models, the results are non-dimensionalized. $\frac{H_e}{H_o}$ and $\frac{Q}{Q_d}$ values are calculated and shown on graphs in Figure 7 for all models. The experiments conducted in water institute studies were simulated in three dimensions by Flow-3D, as shown in Figure 8.
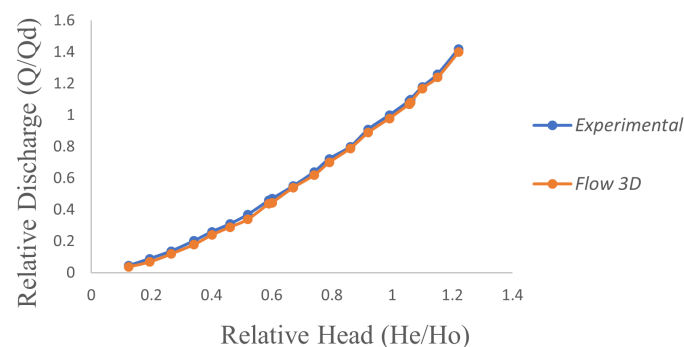

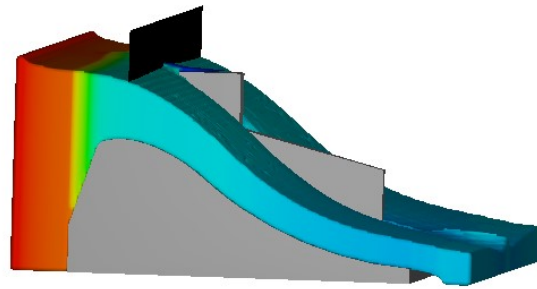**Figure 7.** Numerical and physical model comparison.

Figure 8. Numerical modeling for simulating flow depth.

In order to verify the numerical simulation results the grid size and boundary condition are critical, and it should be remembered that a wrong boundary condition results in a completely different result. Darwin's theory of evolution is the foundation for gene expression programming [49]. This technique can select the input variables that have the greatest impact on the model automatically. The GEP model formulations were used to extract the most effective indicators. Ferreira [49] provided more information on the technical formulation of the GEP approach.

The GEP is the newest evolutionary algorithm approach, and it is becoming more prominent due to its high accuracy [50]. The main advantage of gene expression programming is that it may express the link between variables explicitly. It should be noted that each run of the GEP method produces a different formula. Accordingly, each model was run multiple times with various GEP parameter settings, and the best models (with the lowest amount of error) were chosen. 70% and 30% of the data were used to train and test discharge estimating models, respectively. The GEP approach demonstratedthat the predicted discharge and experimental discharge were in good agreement. The head over the spillway crest was $H_o = \text{h} + V_a^2/2$ g, in which V_a was the approach velocity. The crest height and crest length were utilized as inputs, while discharge was used as an output.

In the GEP model, five options were explored utilizing different operators listed in Table 3, in addition to the ultimate results. The four basic arithmetic operators (+, −,*, /), were used as well as several basic mathematical functions (x^2, exp, ln, cube root, Atan, Tanh, min, max). The GEP's mathematical equation for option 5 is as follows (9). The parameter d in this equation represents the input parameters, while the parameter c donates constant values determined by GEP. For $R^2 = 0.972$, RMSE = 0.85, and MAE = 0.64 , option 5 was the best alternative. Before 5 was chosen as the best option, a variety of performances were carried out, some of which are listed in Table 4. As well as this, different equations of these runs are shown based on changing the test and train stages. Moreover, the scatter plot in Figure 9 indicates that the training and testing data points are close to the $y = x$ line and, providing that, GEP can properly predict relative discharge values. The GEP model's output is typically displayed as mathematical equations, and decision trees Figure 10. The following is the optimal model obtained through GEP modeling for discharge prediction:

$$Y = 0.0$$

$$Y = (d(2) - (1.0/(((((d(1) + d(1))/G1C1) - (\frac{d(0) + d(0)}{2})))))$$

$$Y = Y + atan(((((1.0/d(0))) + (1.0/d(0))))/2.0) + ((1.0/(d(0))) - d(0)))) \quad (9)$$

$$Y = Y + ((atan((atan(G3C6) - d(2))) + tanh(\frac{d(0)}{d(2)})))/2.0)$$

$$result = Y$$

where d_0, d_1 and d_2 correspond to the model input variables and G1C1 and G3C6 are constant values. The numerical constants are shown in Table 5 and the equation derived

from the expression trees is shown in Figure 10. The following GEP prediction approach is highlighted in the Materials and Methods section.
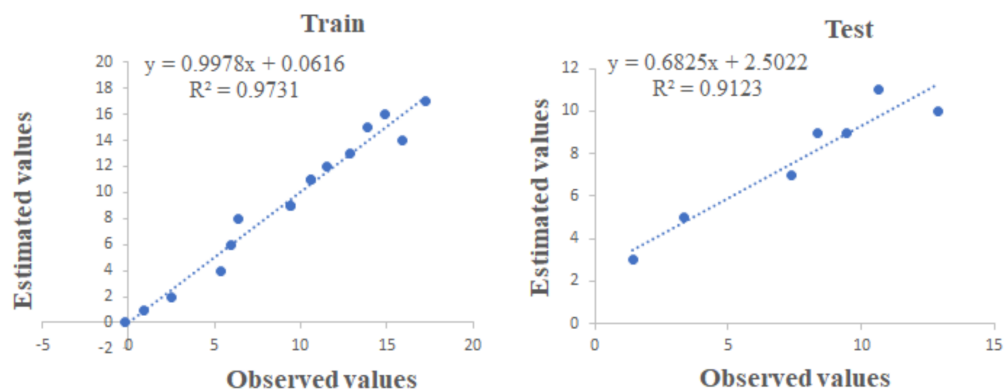

Figure 9. Comparison between estimated and observed values of $\frac{Q}{Q_d}$ for Train set and Test set.

Table 4. Different run cases of option 5 (equations are provided in Appendix A, Table A1).

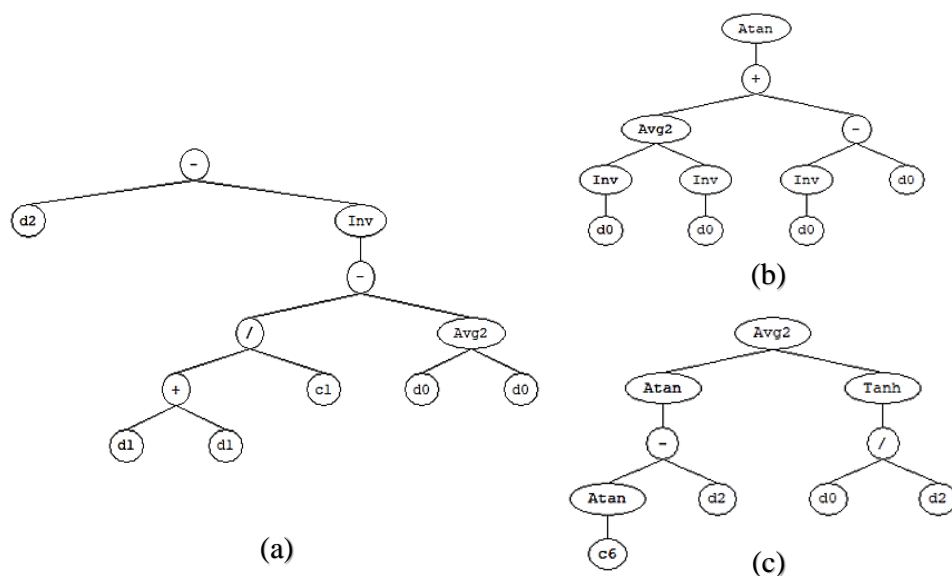| Case Number	Train Phasing			Test Phasing		
	R^2	RMSE	MAE	R^2	RMSE	MAE
Case 1	0.9	1.611	0.95	0.86	1.49	0.97
Case 2	0.92	1.43	0.85	0.89	1.71	1.28
Case 3	0.96	0.94	0.74	0.86	1.19	0.87
Case 4	0.94	1.21	0.95	0.71	2.11	1.3
Case 5	0.91	1.48	0.83	0.87	1.89	1.01
Case 6	0.92	1.47	0.94	0.85	1.49	0.93
Case 7	0.96	0.97	0.7	0.87	1.47	1.07
Case 8	0.93	1.32	0.86	0.82	1.73	1.016
Case 9	0.91	1.48	0.95	0.89	1.51	1.02
Case 10	0.91	1.54	1.02	0.86	1.69	1.06

Figure 10. Expression tree (ET) for presented model (Equation (8)) and the overall discharge is Q = a+b+c. The first line of an equation is represented by (**a**,**b**) stands for the equation's second line, and the third line of the equation is shown by (**c**).

Table 5. The values of the input and constant parameters used in ET.

Parameters	Value	Coefficients	Value
d_0	$\frac{L_e}{H_o}$	G1C1	2.02
d_1	$\frac{H_e}{H_o}$	G3C6	0.06
d_2	$\frac{H_e}{p}$	-	-

4. Conclusions

An ogee spillway is an essential hydraulic structure that can be developed on a variety of sites where concurrent flooding is a concern. Numerous studies based on numerical and analytical relationships have been undertaken to evaluate the discharge values in ogee spillways. The analysis of the behavior and hydraulic parameters of flow over a spillway dam is a difficult and time-consuming task. Flow-3D was utilized for modeling 21 various water levels. In computational fluid dynamics modeling, *RNG* turbulence model has the highest level of compatibility. Artificial intelligence approaches were also applied to estimate the flow over the spillway. The input (independent) variables in the proposed GEP technique included $\frac{H_e}{H_o}$, $\frac{L_e}{H_e}$ and $\frac{H_e}{p}$, whereas the output (dependent) variable was the discharge of the ogee spillway. R^2, RMSE and MAE were employed as evaluation indicators. Using these tools aids a designer in establishing the best condition for a hydraulic construction in which the flow pattern is critical. The modeling findings suggest that the CFD methods and GEP models are useful tools for simulating the flow pattern in the ogee spillways. They can be utilized to estimate Q with no need for complicated and time-consuming laboratory techniques. The proposed method is simple to use and accurate enough for practical usage.

Author Contributions: Conceptualization, S.M. and E.A.; data preparation, S.M., E.A. and M.N.; formal analysis, S.M. and E.A.; investigation, E.A. and M.N.; methodology, S.M.; resources, S.F.; supervision, M.N.; validation, S.M., E.A.; visualization, S.M., H.M. and E.A.; writing—original draft, S.M., E.A. and M.N.; writing—review and editing, M.N. and H.M.; correspondence, S.M. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GEP	Gene expression Programming
RMSE	Root Mean Square Error
R^2	Coefficient of Determination
MAE	Mean Absolute Error

Appendix A

Mathematical equations from GEP model for estimating Q showed in Table A1, performance results of GEP model for option 5 described in Table 4. where d_1, d_2 and d_3 correspond to the model input variables and numbers related to coefficients are also included in the equations.

Table A1. Each case's formulation for option 5, described in Table 4.

Case Number	Equation	
Case 1	$Y = -3.87$ $Y = Y + (d(3)/atan((((d(1)^2)*d(2)) + (-3.4))$ $Y = Y + gep3Rt(((((d(3) + d(3))^2) + ((-8.61*-8.61) + d(3)))/2.0))$ $result = Y$	(A1)
Case 2	$Y = min(((min((3.95 + 9.42), min(d(3), 3.95)) + (4.6 - d(2)))/2.0), (-2.4))$ $Y = Y + max(atan(exp(atan(d(1)))), (((d(3) + d(2)) + (d(2)/d(1)))/2.0))$ $Y = Y + tanh((((min(-4.18, d(2)) + (d(1)*d(1)))/2.0) - ((d(3) - 3.3) + 2.52)))$ $result = Y$	(A2)
Case 3	$Y = (1.0/((((((d(1) - d(2))*(d(1)^2) + reallog(0.51))/2.0)^2)))$ $Y = Y + (((((7.2 + d(3))/2.0) + (d(2) - d(2)))/2.0) - ((7.57 + 2.09) - ((d(2) + 7.57)/2.0)))$ $Y = Y + reallog((((reallog(d(2)) + (4.157 + d(2))) + (((3.74 + d(3))/2.0))/2.0))$ $result = Y$	(A3)
Case 4	$Y = (1.0/(atan(gep3Rt(((reallog(d(2) + (-6.03)) + ((d(2) + (-6.03))/2.0))))))$ $Y = Y + ((1.0 - max(reallog(d(2)), (-5.78)))*((d(2) - d(3)) + (1.0/(2.28))))$ $Y = Y + d(3)$ $result = Y$	(A4)
Case 5	$Y = (d(3) - ((atan(2.39) - (d(1) - d(1))) - (1.0/(d(1)))))$ $Y = Y + tanh(((max(2.86, d(1)) + (-7.93 + d(1))) + ((d(1) + d(1)) - d(2))))$ $Y = Y + tanh((((6.7 - (-1.05)) + (-5.52*(-7.54)))/2.0) - ((d(1)^2) - d(3))))$ $result = Y$	(A5)
Case 6	$Y = ((1.0/(d(2))) - 1.58)$ $Y = Y + d(3)$ $Y = Y + (1.0/(((((d(1) - d(3))*(8.1 + (-1.55)) + (1.0 - ((-3.74 + 1.30)/2.0)))$ $result = Y$	(A6)
Case 7	$Y = gep3Rt(((min(d(3), d(1))/tanh(d(1))) + (min(2.68, 2.34) - d(1))))$ $Y = Y + tanh(tanh((((-3.98 - d(3)) - (d(3) - (-7.59)) + (d(1)*d(1)))))$ $Y = Y + (((((d(2)/d(1)) + ((-7.26 + d(1))/2.0))/2.0) + ((d(3)/0.97)*atan(d(3)))/2.0)$ $result = Y$	(A7)
Case 8	$Y = ((gep3Rt((1.0 - tanh(d(2)))) + ((d(2) + reallog(d(3)))/2.0))/2.0)$ $Y = Y + ((min(gep3Rt((d(1) - 6.6)), ((d(3) + (-8.98)/2.0)) + d(3)))/2.0)$ $Y = Y + ((gep3Rt(d(3))/(d(1))) + atan(tanh(3.3)))$ $result = Y$	(A8)
Case 9	$Y = ((((d(3) - 7.21) - (-4.07 - d(3))) + ((1.0/(d(1))) + (1.0/(d(2))))/2.0)$ $Y = Y + atan((((tanh(-1.7)*d(2)) + d(1))*d(1)))$ $Y = Y + atan(((9tanh(-1.706) - (d(1) - d(3))) + atan((d(3) - d(1)))))$ $result = Y$	(A9)
Case 10	$Y = gep3Rt(-6.92)$ $Y = Y + (1.0 - (((((d(2)*(-0.327) + reallog(d(2))/2.0)*(d(3) - d(2)))^2))$ $Y = Y + ((d(2) + d(2))/2.0)$ $result = Y$	(A10)

References

1. Demeke, G.K.; Asfaw, D.H.; Shiferaw, Y.S. 3D hydrodynamic modelling enhances the design of Tendaho Dam Spillway, Ethiopia. *Water* **2019**, *11*, 82. [CrossRef]
2. Heravi, G.; Salehi, M.M.; Rostami, M. Identifying cost-optimal options for a typical residential nearly zero energy building's design in developing countries. *Clean Technol. Environ. Policy* **2020**, *22*, 2107–2128. [CrossRef]
3. Korzeniowski, A.; Ghorbani, N. Put Options with Linear Investment for Hull-White Interest Rates. *J. Math. Financ.* **2021**, *11*, 152. [CrossRef]
4. Amini, E.; Golbaz, D.; Asadi, R.; Nasiri, M.; Ceylan, O.; Majidi Nezhad, M.; Neshat, M. A Comparative Study of Metaheuristic Algorithms for Wave Energy Converter Power Take-Off Optimisation: A Case Study for Eastern Australia. *J. Mar. Sci. Eng.* **2021**, *9*, 490. [CrossRef]
5. Boiten, W. Flow measurement structures. *Flow Meas. Instrum.* **2002**, *13*, 203–207. [CrossRef]
6. Tafazzoli, S.; Shafaghat, R.; Alamian, R. Numerical investigation on the multi-body hydrodynamic interactions under Caspian Sea environmental conditions. *Ocean. Eng.* **2021**, *232*, 109048. [CrossRef]

7.　Mazaheri, P.; Asgarian, B.; Gholami, H. Assessment of strengthening, modification, and repair techniques for aging fixed offshore steel platforms. *Appl. Ocean. Res.* **2021**, *110*, 102612. [CrossRef]

8.　US Department of the Interior, Bureau of Reclamation. *Design of Small Dams*; US Department of the Interior: Washington, DC, USA, 1987.

9.　Ho, D.; Riddette, K. Application of computational fluid dynamics to evaluate hydraulic performance of spillways in Australia. *Aust. J. Civ. Eng.* **2010**, *6*, 81–104. [CrossRef]

10.　Tafazzoli, S.; Shafaghat, R.; Alamian, R. Optimization study of a catenary mooring system for a spar floating wind turbine based on its hydrodynamic responses. *J. Eng. Marit. Environ.* **2021**, *235*, 657–674. [CrossRef]

11.　Choupin, O.; Henriksen, M.; Etemad-Shahidi, A.; Tomlinson, R. Breaking-down and parameterising wave energy converter costs using the CapEx and similitude methods. *Energies* **2021**, *14*, 902. [CrossRef]

12.　Salazar, F.; San-Mauro, J.; Celigueta, M.Á.; Oñate, E. Shockwaves in spillways with the particle finite element method. *Comput. Part. Mech.* **2020**, *7*, 87–99. [CrossRef]

13.　Ghorbani, N.; Korzeniowski, A. Adaptive Risk Hedging for Call Options under Cox-Ingersoll- Ross Interest Rates. *J. Math. Financ.* **2020**, *10*, 697–704. [CrossRef]

14.　Choupin, O.; Andutta, F.P.; Etemad-Shahidi, A.; Tomlinson, R. A decision-making process for wave energy converter and location pairing. *Renew. Sustain. Energy Rev.* **2021**, *147*, 111225. [CrossRef]

15.　Savage, B.M.; Johnson, M.C. Flow over ogee spillway: Physical and numerical model case study. *J. Hydraul. Eng.* **2001**, *127*, 640–649. [CrossRef]

16.　Kim, D.G.; Park, J.H. Analysis of flow structure over ogee-spillway in consideration of scale and roughness effects by using CFD model. *KSCE J. Civ. Eng.* **2005**, *9*, 161–169. [CrossRef]

17.　Peltier, Y.; Dewals, B.; Archambeau, P.; Pirotton, M.; Erpicum, S. Pressure and velocity on an ogee spillway crest operating at high head ratio: Experimental measurements and validation. *J. Hydro-Environ. Res.* **2018**, *19*, 128–136. [CrossRef]

18.　Jahad, U.; Al-Ameri, R.; Chua, L.; Das, S. Investigating the Effects of Geometry on the Flow Characteristics and Energy Dissipation of Stepped Spillway Using Two-Dimensional Flow Modelling. In Proceedings of the International Association for Hydro-Environment Engineering and Research. Asia Pacific Division Congress, Yogyakarta, Indonesia, 2–5 September 2018; pp. 289–296.

19.　Aydin, M.C.; Isik, E.; Ulu, A.E. Numerical modeling of spillway aerators in high-head dams. *Appl. Water Sci.* **2020**, *10*, 1–9. [CrossRef]

20.　Wan, W.; Liu, B.; Raza, A. Numerical prediction and risk analysis of hydraulic cavitation damage in a High-Speed-Flow spillway. *Shock Vib.* **2018**, *2018*, 1817307. [CrossRef]

21.　Ghanbari, R.; Heidarnejad, M. Experimental and numerical analysis of flow hydraulics in triangular and rectangular piano key weirs. *Water Sci.* **2020**, *34*, 32–38. [CrossRef]

22.　Bayon, A.; Valero, D.; García-Bartual, R.; López-Jiménez, P.A. Performance assessment of OpenFOAM and FLOW-3D in the numerical modeling of a low Reynolds number hydraulic jump. *Environ. Model. Softw.* **2016**, *80*, 322–335. [CrossRef]

23.　Valero, D.; Bung, D.B.; Crookston, B.M. Energy dissipation of a Type III basin under design and adverse conditions for stepped and smooth spillways. *J. Hydraul. Eng.* **2018**, *144*, 04018036. [CrossRef]

24.　Ferreira, C. Gene expression programming: A new adaptive algorithm for solving problems. *arXiv* **2001**, arXiv:cs/0102027.

25.　Yildiz, A.; Yarar, A.; Kumcu, S.Y.; Marti, A.I. Numerical and ANFIS modeling of flow over an ogee-crested spillway. *Appl. Water Sci.* **2020**, *10*, 1–10. [CrossRef]

26.　Ebtehaj, I.; Bonakdari, H.; Zaji, A.H.; Azimi, H.; Sharifi, A. Gene expression programming to predict the discharge coefficient in rectangular side weirs. *Appl. Soft Comput.* **2015**, *35*, 618–628. [CrossRef]

27.　Kisi, O.; Emiroglu, M.E.; Bilhan, O.; Guven, A. Prediction of lateral outflow over triangular labyrinth side weirs under subcritical conditions using soft computing approaches. *Expert Syst. Appl.* **2012**, *39*, 3454–3460. [CrossRef]

28.　Salmasi, F. Effect of downstream apron elevation and downstream submergence in discharge coefficient of ogee weir. *ISH J. Hydraul. Eng.* **2018**, *27*, 375–384. [CrossRef]

29.　Khan, M.; Azamathulla, H.M.; Tufail, M. Gene-expression programming to predict pier scour depth using laboratory data. *J. Hydroinform.* **2012**, *14*, 628–645. [CrossRef]

30.　Roushangar, K.; Akhgar, S.; Salmasi, F.; Shiri, J. Modeling energy dissipation over stepped spillways using machine learning approaches. *J. Hydrol.* **2014**, *508*, 254–265. [CrossRef]

31.　Bagatur, T.; Onen, F. Computation of design coefficients in ogee-crested spillway structure using GEP and regression models. *KSCE J. Civ. Eng.* **2016**, *20*, 951–959. [CrossRef]

32.　Bertone, E.; Stewart, R.A.; Zhang, H.; Veal, C. Data-driven recursive input–output multivariate statistical forecasting model: Case of DO concentration prediction in Advancetown Lake, Australia. *J. Hydroinform.* **2015**, *17*, 817–833. [CrossRef]

33.　Esmaeilpour, M.; Ghanbarian, B.; Liang, F.; Liu, H.H. Scale-dependent permeability and formation factor in porous media: Applications of percolation theory. *Fuel* **2021**, *301*, 121090. [CrossRef]

34.　Azita, M.; Mohammad, I. A Blockchain Information Management Framework for Construction Safety. In Proceedings of the International Conference on Computing in Civil Engineering, Orlando, FL, USA, 12–14 September 2021.

35.　Rahimi, Z.; Othman, F.; Shariff, A.R.M. Improving geometric construction of high resolution SAR images using Kriging-based surrogate modelling in mountainous terrain of Malaysia. *Int. J. Remote Sens.* **2021**, *42*, 8624–8639. [CrossRef]

36. A GNSS-based weather forecasting approach using Nonlinear Auto Regressive Approach with Exogenous Input (NARX). *J. Atmos. Sol. Terr. Phys.* **2018**, *178*, 74–84. [CrossRef]

37. Mohammadrezapour, O.; Piri, J.; Kisi, O. Comparison of SVM, ANFIS and GEP in modeling monthly potential evapotranspiration in an arid region (Case study: Sistan and Baluchestan Province, Iran). *Water Supply* **2019**, *19*, 392–403. [CrossRef]

38. Roushangar, K.; Alami, M.T.; Saghebian, S.M. Modeling open channel flow resistance with dune bedform via heuristic and nonlinear approaches. *J. Hydroinform.* **2018**, *20*, 356–375. [CrossRef]

39. Dindarloo, S.R.; Siami-Irdemoosa, E. Estimating the unconfined compressive strength of carbonate rocks using gene expression programming. *arXiv* **2016**, arXiv:1602.03854.

40. Amini, E.; Golbaz, D.; Amini, F.; Majidi Nezhad, M.; Neshat, M.; Astiaso Garcia, D. A parametric study of wave energy converter layouts in real wave models. *Energies* **2020**, *13*, 6095. [CrossRef]

41. Neshat, M.; Mirjalili, S.; Sergiienko, N.Y.; Esmaeilzadeh, S.; Amini, E.; Heydari, A.; Garcia, D.A. Layout optimisation of offshore wave energy converters using a novel multi-swarm cooperative algorithm with backtracking strategy: A case study from coasts of Australia. *Energy* **2022**, *239*, 122463. [CrossRef]

42. Venkata Raju Badanapuri, Design Principles that are involved in the Design of Flow over an Ogee Crest Spillway. *Int. J. Sci. Res.* **2019**, *8*, 245–254.

43. Sarkardeh, H.; Marosi, M.; Roshan, R. Stepped spillway optimization through numerical and physical modeling. *Int. J. Energy Environ.* **2015**, *6*, 597.

44. Matozinhos, C.F.; Campagnole dos Santos, A.A. Two-phase CFD simulation of research reactor siphon breakers: A verification, validation and applicability study. *Nucl. Eng. Des.* **2018**, *326*, 7–16. [CrossRef]

45. Celik, I.; Ghia, U.; Roache, P.; Freitas, C.; Coloman, H.; Raad, P. Procedure of Estimation and Reporting of Uncertainty Due to Discretization in CFD Applications. *J. Fluids Eng.* **2008**, *130*, 078001. [CrossRef]

46. Azimi, H.; Shabanlou, S.; Ebtehaj, I.; Bonakdari, H.; Kardar, S. Combination of computational fluid dynamics, adaptive neuro-fuzzy inference system, and genetic algorithm for predicting discharge coefficient of rectangular side orifices. *J. Irrig. Drain. Eng.* **2017**, *143*, 04017015. [CrossRef]

47. Legates, D.R.; McCabe, G.J., Jr. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]

48. Morais, V.H.; Gireli, T.Z.; Vatavuk, P. Numerical and experimental models applied to an ogee crest spillway and roller bucket stilling basin. *RBRH* **2020**, *25*, 1–12. [CrossRef]

49. Ferreira, C. *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2006; Volume 21.

50. Lu, X.; Wang, X.; Zhang, L.; Zhang, T.; Yang, C.; Song, X.; Yang, Q. Improving forecasting accuracy of river flow using gene expression programming based on wavelet decomposition and de-noising. *Hydrol. Res.* **2018**, *49*, 711–723. [CrossRef]