CrossMark

RESEARCH PAPER

Prediction of Energy Dissipation of Flow Over Stepped Spillways Using Data-Driven Models

Abbas Parsaie[1] · Amir Hamzeh Haghiabi[1] · Mojtaba Saneie[2] · Hasan Torabi[1]

Abstract Stepped spillway is an effective approach to remove the potential occurrences of cavitation in chute of spillways and also to significantly reduce the size of energy dissipators at the toe of dam. In this study, to predict the energy dissipation ratio of flow over stepped spillways, artificial neural network, support vector machine, genetic programming (GP), group method of data handling (GMDH), and multivariate adaptive regression splines (MARS) were developed. MARS, GMDH, and GP are smart function fitting methods that assign more weight to the most effective parameters on the output. These models, in addition to predicting the desired phenomena, present a mathematical expression between independent and dependent variables. Results of applied models indicated that all models have suitable performance; however, MARS model with coefficient of determination close to 0.99 in training and testing stages is more accurate compared to others. This model also has a high ability to present the mathematical expression between involved parameters in energy dissipation. To derive the most influential parameters on efficiency of stepped spillways in terms of energy dissipation of flow, a review on the structure of models derived from GP, GMDH, and MARS was carried out. Results indicated that drop number, ratio of critical depth to the height of steps, and Froude number are the most effective parameters on energy dissipation of flow over stepped spillways.

List of symbols

ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network
AVEG	Average
bi	Bias
DN	Drop number
E	Specific energy
EDR	Energy dissipation ratio
Fr	Froude number
g	Gravity acceleration
GA	Genetic algorithm
GEP	Gene expression programming
GMDH	Group method of data handling
GP	Genetic programming
h	Height of steps
H_w	Dam height
l	Length of steps
LM	Least square
MARS	Multivariate adaptive regression splines
Max	Maximum
Min	Minimum
MLP	Multilayer perceptron neural networks
PSO	Partcle swarm optimization
RBF	Radial basis function

✉ Amir Hamzeh Haghiabi
  haghiabi.a@lu.ac.ir

  Abbas Parsaie
  Abbas_Parsaie@yahoo.com

  Mojtaba Saneie
  Saneie_m@scwmri.ac.ir

  Hasan Torabi
  Torabi.H@lu.ac.ir

[1] Water Engineering Department, Lorestan University, Khorramabad, Iran

[2] Soil Conservation and Watershed Management Research Institute, Tehran, Iran

Shiraz University

Springer

S	Slope of stepped spillway
STDEV	Standard deviation
SVM	Support vector machine
V	Flow velocity
wi	Weight
y	Flow depth
Y_c/h	Critical depth to the height of steps
C	Error penalty factor
w	Normal vector
ε	Loss function

1 Introduction

Stepped spillway is a spillway where steps are installed on the surface. Installing steps starts close to the crest and continues to the toe of dams (Chen 2015). Stepped spillways have two main features: high efficiency for energy dissipation and dramatic decrease in the probability of cavitation occurrence (Chanson 2002; Frizell et al. 2013; Pfister and Hager 2011). In terms of energy dissipation, the feature causes a significant decrease in the size of energy dissipator structure at the toe of dam (Felder and Chanson 2011). This characteristic considerably reduces the cost of construction, since energy dissipators are among the most costly parts of dam construction projects (Sorensen 1985). Due to the high efficiency of stepped spillways in terms of energy dissipation and removing cavitation, several investigations including experimental and numerical methods have been conducted on hydraulic behavior of flow over them (Dehdar-Behbahani and Parsaie 2016; Husain et al. 2014; Morovati et al. 2016; Nikseresht et al. 2013; Parsaie et al. 2015; Zhan et al. 2016). In experimental studies, flow pattern on stepped spillways and the effect of geometry of steps on energy dissipation have been evaluated (Mohammad Rezapour Tabari and Tavakoli 2016). By observing the pattern of flow over stepped spillways, investigators have proposed three classes for flow regime on these structures. They proposed that flow regimes can be divided into three classes as napped flow, transition flow and skimming flow. Napped flow occurs at low discharge values. In this condition, flow leaves the upper step and falls down onto the lower step (Fen et al. 2016; Tabbara et al. 2005). Energy dissipation in this condition is due to the collision of the jet of flow with steps and hydraulic jump that may occur completely or incompletely. Skimming flow occurs in large discharge value, and in this status a pseudo-bottom is created between steps and passing flow. Transition regime is a condition between napped and skimming flow. For more information on flow regime over stepped spillways, refer to Boes et al. 2000; Tatewar and Ingle 1996. Nowadays, by advances in computer facilities and due to the high cost of experiments,

investigators have encouraged the use of numerical method for simulation of hydraulic phenomena. In this regard, computational fluid dynamic (CFD) methods for simulation of flow over steeped spillways have been conducted (Chatila and Jurdi 2004; Parsaie and Haghiabi 2015a, b; Zare and Doering 2012). Using CFD techniques is required to solve Navier–Stokes equations along turbulence models. Fortunately, powerful open source codes such as Open Foam and commercial packages such as Flow3D and Fluent have been proposed (Attarian et al. 2014; Cheng et al. 2006). Recently, by developing soft computing techniques in most areas of engineering, investigators have tried to use them for accurate presentation of results of experiments (Noori et al. 2015; Samadi et al. 2015; Zahiri and Azamathulla 2014). Some of the soft computing techniques such as artificial neural networks (ANNs) (Noori et al. 2010b), adaptive neuro-fuzzy inference system (ANFIS) (Noori et al. 2010a) and support vector machine (SVM) (Azamathulla and Wu 2011; Noori et al. 2009) developed a network for modeling and predicting desired phenomena. Other types of soft computing techniques also have been proposed. In the last few decades, in addition to developing a network, smart functions have also been presented. In this regard, genetic programming (GP) (Azamathulla and Ghani 2011), gene expression programming (GEP) (Azamathulla 2013; Azamathulla and Mohd. Yusoff 2013; Emamgholizadeh et al. 2016; Emamgolizadeh et al. 2015; Guven and Kişi 2011; Sattar and Gharabaghi 2015), group method of data handling (GMDH), and multivariate adaptive regression splines (MARS) technique can be mentioned. In these methods, during development process, more weight is attributed to inputs that have more influence on the output. Using multilayer perceptron neural network (MLP), ANFIS (Salmasi and Özger 2014) and GEP (Roushangar et al. 2014) have been reported for modeling energy dissipation of flow over stepped spillways. In this paper, mathematical expression of the relation between parameters involved in energy dissipation of flow over stepped spillways using the GMDH, MARS, and GP is considered. To developed mentioned techniques, results of a series of experiments that were conducted by authors on the hydraulic laboratory of soil conservation and watershed management research institute (Tehran, Iran) are used. To increase the reliability of modeling the results of similar experiments were collected and used, as well.

2 Materials and Methods

To define the effect of geometrical parameters involved in efficiency of stepped spillways, several investigations have been conducted. Figure 1 shows a sketch of stepped spillways. In this figure, height and length of steps are shown
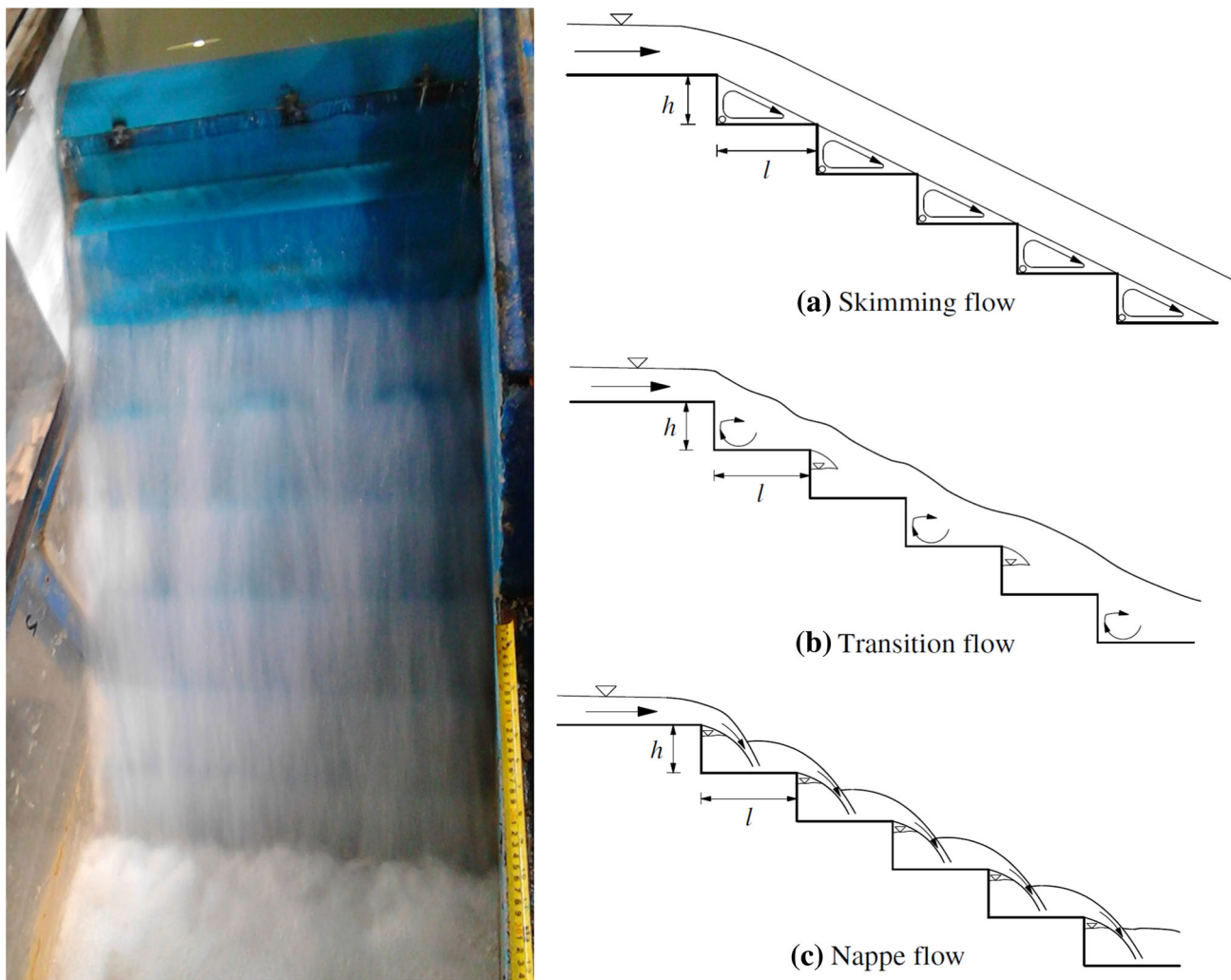
(a) Skimming flow

(b) Transition flow

(c) Nappe flow

Fig. 1 Main parameters involved in energy dissipation in skimming flow over stepped spillways

by h and l, respectively. Depth of flow over stepped spillways is shown by y_0; upstream specific energy is defined with E_0. Depth of flow at the toe of the dam is shown by y_1, and y_2 is the conjugate depth of flow at hydraulic jump. To calculate energy dissipation of flow over stepped spillways, Bernoulli equation is applied between the upstream and downstream of stepped spillway. Equations (1, 2) are used for calculation of upstream and downstream specific energies of flow.

$$E_0 = H_w + y_0 + \frac{V_0^2}{2g} = H_w + y_0 + \frac{q^2}{2g(H_w + y_0)} \quad (1)$$

$$E_0 = y_1 + \frac{V_1^2}{2g} = y_1 + \frac{q^2}{2gy_1^2} \quad (2)$$

In these equations, q is the discharge per weir length, V_0 is the velocity of approached flow, and g is the acceleration gravity. Energy dissipation ratio (EDR) is calculated using Eq. (3).

$$\frac{\Delta E}{E_0} = \frac{E_0 - E_1}{E_0} = 1 - \frac{E_1}{E_0} \quad (3)$$

Geometrical and hydraulic parameters involved in energy dissipation are arranged in Eq. (4) to determine those that affect EDR.

$$\frac{\Delta E}{E_0} = f(q, l, h, H_w, g, N) \quad (4)$$

where H_w is the height of dam and N is the number of steps. Salmasi and Özger (2014) used the Buckingham Π theory as the most famous dimensional analysis technique and derived dimensionless parameters effective on EDR as Eq. (5).

$$\frac{\Delta E}{E_0} = f\left(\frac{q^2}{gH_w^3}, \frac{h}{l}, N, \frac{y_c}{h}, Fr_1\right) \quad (5)$$

In Eq. (5), q^2/gH_w^3 is named drop number and shown by DN and h/l is declared as the slope of stepped spillway, and

therefore, it is shown by *S*. Equation (5) can be rewritten as:

$$\frac{\Delta E}{E_0} = f\left(DN, S, N, \frac{y_c}{h}, Fr_1\right) \quad (6)$$

Equation (6) is the foundation for developing soft computing methods. To develop soft computing methods including ANN, SVM, GMDH, MARS, and GP, in addition to obtained results of experiments conducted by authors, related datasets were collected from Salmasi and Özger (2014). The histogram of collected dataset is shown in Fig. 2.

Physical laboratory models of stepped spillways were constructed from the galvanized iron sheets. The main channel was 12 m long whose cross section was rectangular with 0.90 m depth and 0.60 width. The side walls of the channel were made of Plexiglas and its bed was made from well-pointed steel sheet. In order to control the formation of hydraulic jump, a sluice gate was set at downstream. The longitudinal slope of main channel was equal to 0.001. The depths of flow at upstream and downstream of the structure were measured by point gage with ±0.1 mm sensitivity. Discharge of flow was measured with a V notch weir that was installed at downstream for this
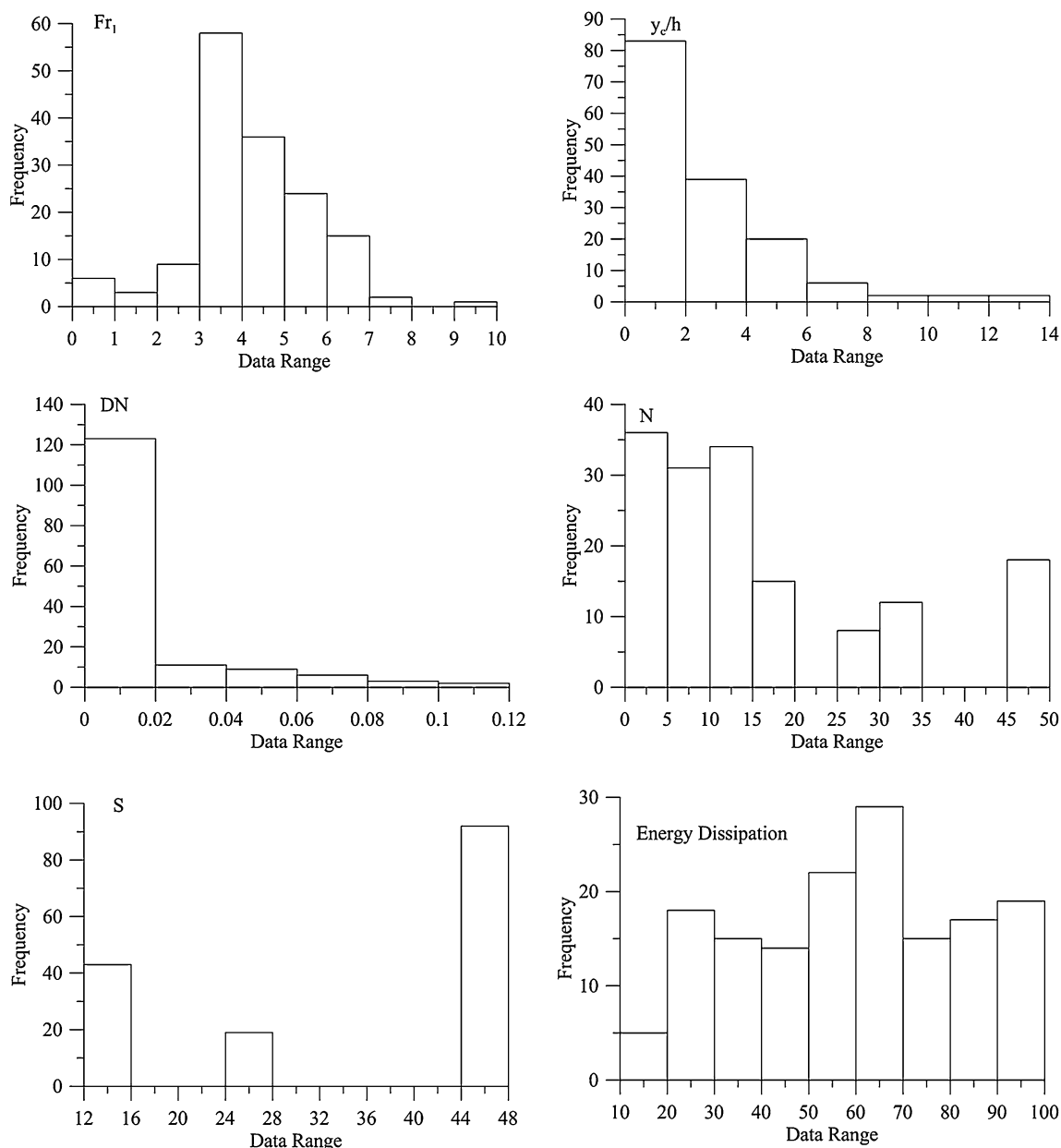

Fig. 2 Histogram of collected datasets related to energy dissipation

Table 1 Summary of models of stepped spillways and flow rates

Models	hs (m)	Ls (m)	N	Q (l/s)
Model 1	0.05	0.05	16	1–40
Model 2	0.05	0.10	16	1–40
Model 3	0.05	0.15	16	1–40
Model 4	0.08	0.5	10	1–40
Model 5	0.10	0.10	8	1–40
Model 6	0.10	0.15	8	1–40
Model 7	0.10	0.20	8	1–40

purpose. Figure 1 shows a laboratory model. The properties of models are given in Table 1.

2.1 Review on ANN

Artificial neural network is a common type of soft computing technique composed of a number of neurons arranged in input, hidden, and output layers. Each neuron is an independent unit in the network where each input is multiplied by a specific weight and then introduced to it. Inputs for the first layer are the original dataset, and inputs for the second and next layers are the output of each neuron in previous layer. As stated earlier, inputs are multiplied by the weights and passed through transfer function that is given on each neuron. The most common types of transfer functions are Gaussian, sigmoidal and tansing. The output of each neuron is summed by a constant value called bias. The most famous types of ANNs are multilayer perceptron neural networks (MLP) that have been widely used in most areas of engineering, especially in hydraulic engineering. MLP usually includes three layers: the input layer used for introducing dataset, hidden layer(s) where main network computation is conducted, and output layer where the results of computation in hidden layer(s) are accumulated and presented. As stated, each input is multiplied by a weight for being introduced to a neuron, and then, results of acting transfer function on them are summed by a bias; the values of weights and biases for all neurons available in

networks are validated using training algorithm. Training means adjusting the variables (weights and biases) to achieve the lowest difference between model output and observed data. Training MLP model can be performed using conventional methods such as Levenberg–Marquardt method. This subject can also be assumed as an optimization problem where modern optimization techniques can be applied to solve it. In this study, an MLP model was developed; the structure of which is shown in Fig. 3. In this model, tangent sigmoid and pure line functions were chosen as governing functions on neurons of hidden and output layers, respectively (Emamgholizadeh et al. 2014a, b), Emamgholizadeh et al. 2016.

2.2 Review on SVM

Support vector machine is a Kernel-based technique that represents a major advance in machine learning algorithm. Support vector machine (SVM) is based on machine learning concept to maximize predictive accuracy; that is,

$$\text{Minimize:} \quad R_{\text{svm}}(\omega, \xi^*) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

$$\begin{aligned}\text{Subject} \\ d_i - \omega\varphi(x_i) + b_i \leq \varepsilon + \xi_i \\ \omega\varphi(x_i) + b_i - d_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \ldots, l,\end{aligned} \quad (7)$$

where w is a normal vector, $(1/2)\|\omega\|^2$ is the regularization term, C is the error penalty factor, b is a bias, ε is the loss function, xi is the input vector, di is the target value, l is the number of elements in the training dataset, $\varphi(xi)$ is a feature space, and ξ_i and ξ_i^* are upper and lower excess deviations. The architecture of SVM is shown in Fig. 4. Famous kernel functions are denoted as follows.

1. Linear kernel: $K(x_i, x_j) = x_i^{\text{T}}x_j$
2. Polynomial kernel: $K(x_i, x_j) = (x_i^{\text{T}}x_j + \gamma)^d, \quad \gamma > 0$
3. RBF kernel: $K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|)^2, \quad \gamma > 0$

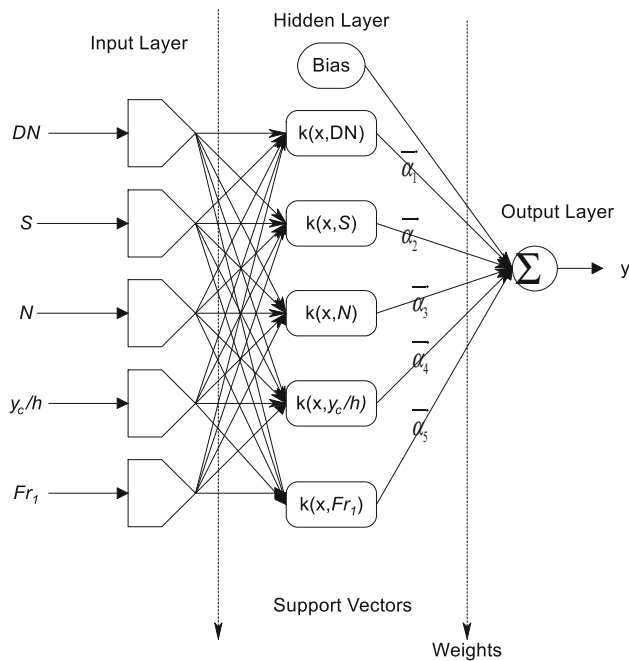**Fig. 3** Structure of multilayer perceptron neural network
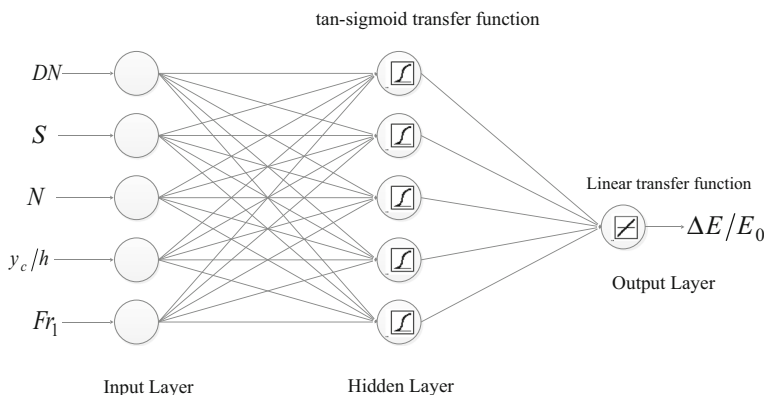
Fig. 4 Network architecture of SVM

4. Sigmoid kernel: $k(x_i, x_j) = \tanh(\gamma x_i^T x_j + r), \quad \gamma > 0$

where variables xi and xj are vectors in the input space, and γ is the regularization parameter. Lagrange multipliers are presented as $\alpha i = \alpha i - \alpha i_*$. The accuracy of prediction is based on selection of three parameters, i.e., γ, ε, and C: the values of which are determined using firefly algorithm (Parsaie et al. 2016; Azamathulla and Wu 2011).

2.3 Review on GP

Genetic programming (GP) technique is a machine learning approach used for modeling input–output complex nonlinear systems that are based on dataset. Developing GP is based on the concept of genetic algorithm (GA). It means that the concepts used in genetic algorithm (GA) are repeated in GP such as genes, multigene, mutation. GP is also used to build a semiempirical formula from the input–output dataset; therefore, it often known as symbolic regression. GP creates the formula that consists of inputs variables and several mathematical operators, namely (+, −, /, and *) and functions, namely (ex, x, sin, cos, tan, lg, sqrt, ln, power). GP performs this process by randomly generating a population of computer programs (represented by tree structures) and then mutating and crossing over the best performing trees to create a new population. This process is continued until the formula is achieved with a suitable accuracy. Unlike classical regression analysis where the designer defined the structure of the empirical formula, GP automatically generates both the structure and the parameters of empirical formula. An individual

multigene is comprised of one or more genes and is called GP tree. To improve the performance of fitness (e.g., to reduce a model's sum of squared errors on a dataset), the genes are obtained incrementally. The final formula may be a weighted linear or nonlinear combination of each gene. The optimal weights for the genes are automatically obtained using ordinary least squares to regress genes against the output data (Azamathulla et al. 2010; Azamathulla et al. 2008).

2.4 Review on MARS

MARS, proposed by Friedman (1991), is a pliable method to amp the relationship between independent and dependent variables in a desired system. MARS method is used to recognize the hidden pattern in dataset in complex systems. Recognition of pattern is defined via proposing a number of coefficients and basic functions. These coefficients and basic functions are justified during regression operation on the used dataset. The main advantage of MARS includes its high ability for mapping input parameters and desired outputs, developing a simple but robust model and its being rational in terms of computational cost. MARS technique is based on simple basis functions defined as follows:

$$|x - t|_+ = \max(0, x - t) = \begin{cases} x - t & x > t \\ 0 & x \leq t \end{cases} \tag{8a}$$

$$|t - x|_+ = \max(0, t - x) = \begin{cases} t - x & x < t \\ 0 & x \geq t \end{cases} \tag{8b}$$

where t denotes the knot. Basic functions are sometimes called mirrored pair functions. These functions are defined for each input variable such as Xj at observed dataset related to them. Sets of basic functions are defined as

$$C = \left\{ (x_j - t)_+, (t - x_j)_+ \right\}; \{x_{1j}, x_{2j}, \ldots, x_{nj}\}, \\ j = 1, \ldots, p \tag{9}$$

The general form of function derived from MARS model is written as an adaptive function as

$$y = \beta_0 + \sum_{i=1}^{M} \beta_i \mathrm{BF}_i(X) \tag{10}$$

where β_0 is constant, $\mathrm{BF}_i(X)$ is known as basic function, and β_i is the coefficient of basic functions. The constant and coefficient of derived function in MARS model are justified using least square error technique. M is the number of basic functions derived from the final stage of model development. Developing MARS model includes two stages. The first one is forward stage. In this stage, the number of basic functions increases to decrease the

difference between the results of model and observed data. In the next step of model development, to avoid over-parameterization and over-fitting, pruning some of the basic functions is considered. In this stage, regarding cross-validation (GCV) criteria given below, basic functions are pruned.

$$GCV = \frac{SSE}{n\left(1 - \left(\frac{C(B)}{n}\right)\right)^2} \tag{11.a}$$

$$C(B) = (B + 1) + dB \tag{11.b}$$

where SSE is the sum of square of residuals, n denotes the number of records, and $C(B)$ defines difficulty criteria, which increases by the number of basic functions. For more information, see (Parsaie et al. 2016; Emamgolizadeh et al. 2015; Haghiabi 2016)

2.5 Review on GMDH

GMDH is a soft computing approach categorized in self-organizing methods developed by Ivakhnenko (1971). In this model, complex networks are gradually developed with regard to the performance of a combination of pairs of inputs and desired output. Each pair of inputs is introduced to a neuron in GMDH network. In the first hidden layer, all combinations of input pairs are evaluated. The number of neurons in the first hidden layer is calculated as shown below.

$$\frac{N_{inp}\left(N_{inp} - 1\right)}{2} \tag{12}$$

As stated, in the first hidden layer, each pair of inputs is introduced to a neuron in the governing equation on which is a quadratic polynomial function. In other words, each pair of inputs is passed through a quadratic polynomial function as

$$\bar{y} = G\left(x_i, x_j\right) = w_0 + w_1 x_i + w_2 x_j + w_3 x_i^2 + w_4 x_j^2 + w_5 x_i x_j \tag{13}$$

where $\bar{y}$ is the output of each neuron; x_i and x_j are the inputs; $w_{1,\dots,5}$ are the weights (coefficients); and w_0 is a bias (constant). $G(x_i, x_j)$ means that the governing function on neurons is only a proportional pair of inputs. Values of weights and biases are justified in training stage. Training means minimizing the difference between output of each neuron with observed data by adapting coefficients and constant of governing equation. To this end, conventional algorithms such as least square (LM) method can be applied. Training can be assumed as an optimization problem, and advanced modern optimization algorithms such as genetic algorithm (GA), particle swarm optimization (PSO) can be used to this end. The idea of using the

quadratic polynomial function as transfer function governing neurons was taken from the Volterra functional series, which states a complete system can be estimated via infinite series of polynomial of inputs. This series is also known as Kolmogorov–Gabor polynomial; the general form of which is given below.

$$y = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{k} w_{ij} x_i x_j x_k + \cdots \tag{14}$$

In development of GMDH, some concepts of GA algorithm, namely seeding, rearing, crossbreeding, selection, and rejection, have been used. In other words, only for developing the first hidden layer do all inputs participate. For developing the second hidden layer in GMDH network, inputs are selected based on their performance. This means that neurons with more accurate answer are selected. Figure 5 shows a sketch of GMDH model. As shown here, for developing the second and next hidden layer(s), neuron(s) with suitable performance in the previous layer are selected (Karbasi and Azamathulla 2016; Najafzadeh 2016; Najafzadeh and Azamathulla 2015; Najafzadeh and Barani 2011; Najafzadeh and Bonakdari 2016; Najafzadeh et al. 2016; Najafzadeh and Sattar 2015; Najafzadeh and Tafarojnoruz 2016; Najafzadeh and Zahiri 2015).

3 Results and Discussion

Development of soft computing techniques is based on dataset. This means that the basic stage of modeling is data preparation. In this stage, dataset should be divided into two groups as training and testing. Assigning dataset to each group is performed with regard to random approach. It is notable that it is better to choose a range of groups near each other. The range of training and testing dataset is given in Table 2. Training encompasses 80% of dataset, and the rest (20%) is considered as testing dataset. Training and testing datasets are used for developing models given in materials and methods section. The next step of preparation of some models such as ANNs is designing the structure of model. Designing the structure includes choosing the number of hidden layer(s), the number of neuron(s) in each hidden layer, types of transfer functions on neuron, and learning algorithm. Learning means justifying the weights and biases in order to minimize the difference between the model outputs and observed desired data.

Developing ANN model is based on designer's experiment. However, recommendations of a researcher who conducted similar studies are also very useful. In this study, preparing the multilayer perceptron neural network
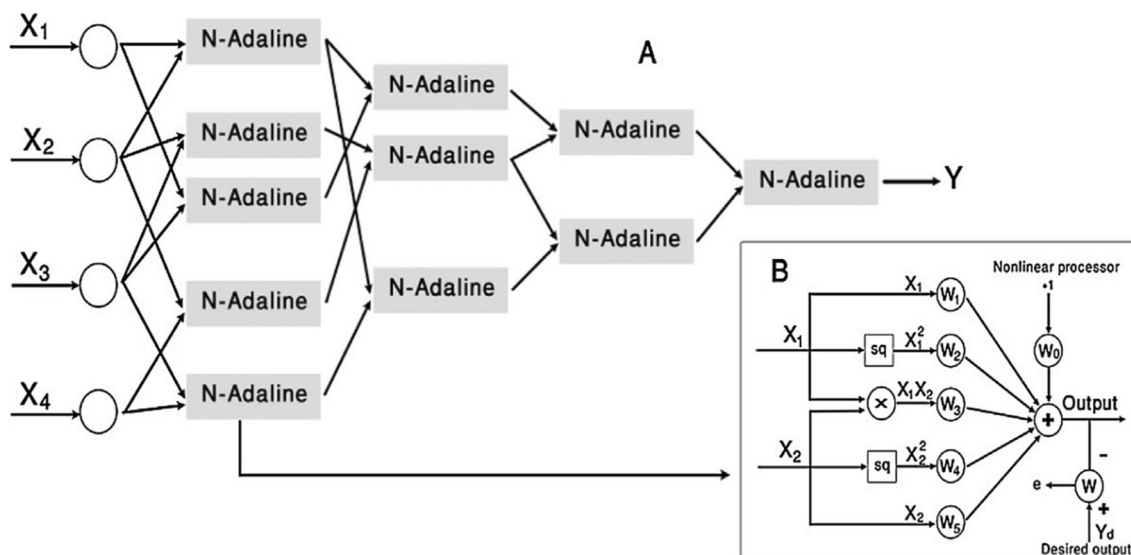
Fig. 5 Schematic of GMDH model development

Table 2 Range of dataset assigned to stages of preparing soft computing models

Stage	Range	Fr_1	Y_c/h	DN	N	S	EDR
Training	Min	0.234	0.094	0.000	3.000	15.000	15.136
	Max	9.339	13.781	0.104	50.000	45.000	96.580
	AVEG	4.128	2.556	0.012	18.309	33.943	59.292
	STDV	1.401	2.566	0.022	14.202	13.465	23.133
Testing	Min	0.307	0.233	0.000	3.000	15.000	13.145
	Max	6.883	6.327	0.109	50.000	45.000	96.441
	AVEG	4.438	2.344	0.014	17.839	34.032	56.412
	STDEV	1.500	1.799	0.028	15.323	14.226	23.534

as a common type of ANN was developed based on recommendations of Azamathulla et al. (2016). They stated that to develop MLP model, after data preparation, designing the structure should be considered step by step. Based on these recommendations, initially one hidden layer is considered and then a number of neurons are chosen. There are different types of transfer functions. In this study, different types of transfer functions created in MATLAB software were tested. In the next step and after finding the proper transfer function, the number of neurons in the same hidden layer may increase in order to improve the performance of developed model. Another way to increase the performance of MLP is to increase the number of hidden layers. A summary of the development process is presented in Table 3. As given in this table, the tansig as transfer function has the best and most suitable performance compared to other tested functions. Models shown in row number four are chosen for predicting energy dissipation of flow over stepped spillways. As shown in this table, increasing the number of hidden

Table 3 Summary of performance of MLP model during development stage

Row	N-H-L	F-HL&TF	R^{2*}	RMSE*	R^{2**}	RMSE**
1	1	5-Purelin	0.63	8.47	0.57	10.85
2	1	5-radbas	0.84	6.58	0.81	7.36
3	1	5-logsig	0.93	4.83	0.92	5.25
4	1	5-tansig	0.987	2.65	0.95	3.98
5	1	9-tansig	0.99	2.52	0.96	3.68

N-H-L number of hidden layer(s), *F-HL&TF* First hidden layer and transfer function

* error indices of MLP during training stage, ** error indices of MLP during testing stage

layers does not have a significant effect on increasing the model accuracy. It is notable that the proposed MLP model was trained using LM method. Results of MLP model in training and testing stages are shown in Figs. 6 and 7. In these figures, the outcome of MLP model is shown versus the observed data.
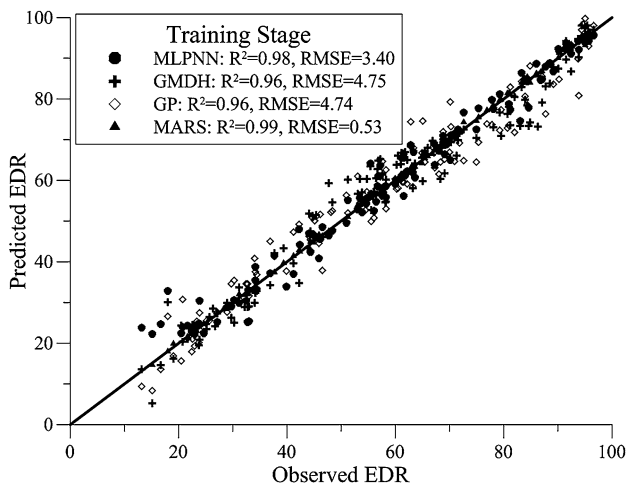
Fig. 6 Results of proposed applied models in training stage
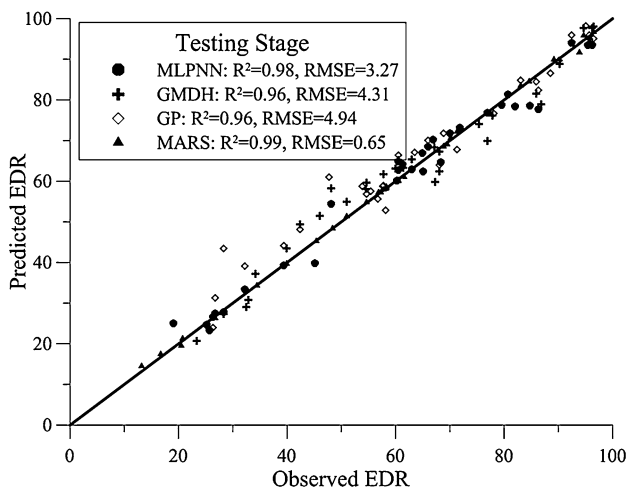

Fig. 7 Results of proposed applied models in testing stage

Development of SVM is similar to MLP model. This means that to prepare SVM, the first step is data preparation. In this study, the same dataset used for developing MLP was used for preparation of SVM. The next step of preparing SVM is designing the structure. One subroutine of designing is choosing transfer function. In this study, the structure of SVM developed for prediction of energy dissipation is given in Fig. 4. Four kernel functions are given in the review on SVM section were tested. Training SVM can be used as optimization problem. To perform this task, quadratic optimization method is used. A summary of testing kernel functions is given in Table 4. As illustrated in this table, RBF has the best performance among tested transfer functions. The value of parameters of kernel function including Gamma value = 33,391.97 and $C = 47.61$ was obtained in preparation stages. Results of SVM in training and testing stages are given in Figs. 6 and 7.

Table 4 Summary of performance of kernel functions

Kernel function	Train		Test	
	R^2	RMSE	R^2	RMSE
Linear	0.78	10.70	0.85	10.16
Polynomial(3)	0.98	3.13	0.85	9.14
RBF	0.99	1.22	0.98	2.61
Sigmoid	0.96	4.18	0.93	6.17

As stated in the review on GP section, GP is a smart function fitting method. The main point in smart function fitting is related to assigning more weight to inputs that are more effective on output. To develop GP model for mathematical expression of involved parameters on energy dissipation ratio, input parameters regarding Eq. (6) (i.e., inputs: DN, S, N, y_c/h, Fr and output: $\Delta E/E_0$) were used. It is notable that in this study, training dataset used for development of ANN and SVM was used for developing GP and testing dataset is used for evaluating the derived model. To develop GP, mathematical operations including [summation (+), mines (−), multiply (×), and division (÷)] and a number of mathematical functions such as (times, minus, plus, square, tanh, exp) were applied. To derive a suitable model, several generations were performed. To derive a suitable model based on GP, the same approach considered for developing ANN was considered for using GP. The number of genes increased one by one and mathematical functions were added one by one. Results of values of justified parameters in GP are given in Table 5.

General form of derived model from GP is $\Delta E/E_0 = w + \sum_{i=1}^{n} \alpha_i \text{gene}_i$. In this equation, w is the bias and α_i is the weight of each gene. Results of derived model are expressed as Eq. (15). The structure derived from GP is given in Fig. 8. As shown in this figure, there are five genes available in derived model. Results of derived model from GP in training and testing stages are given in Figs. 6 and 7. As stated in GP model description, GP is a smart fitting function method. Reviewing the structure of derived model genes shows that the ratio of critical depth to step height, drop number, and the number of steps and drop number appear in most genes. Comparing the performance of MLP and SVM shows that GP has accuracy close to MLP; however, the performance of SVM is slightly better than GP.

$$\frac{\Delta E}{E_0} = 87.75 + 3.095\frac{y_c}{h} + 3.095N$$
$$- 0.02761 \, \text{square} \, (N + 9.532)\ldots$$
$$- 49.14 \tanh\left(\frac{y_c}{h}\right) + 3.095\text{DN} \times S - 8.666Fr_1\text{DN}$$
$$\times \text{square} \, (Fr_1)\ldots$$
$$- 1.963Fr_1\frac{y_c}{h}\tanh(Fr_1) \tag{15}$$

Table 5 Justified parameters of GP model

Parameter	Description of parameter	Setting of parameter
P1	Function set	Times, minus, plus, square, tanh, exp
P2	Population size	200
P3	Mutation frequency %	0.94
P4	Crossover frequency %	50
P5	Number of replications	10
P6	Block mutation rate %	30
P7	Instruction mutation rate %	30
P8	Instruction data mutation rate %	40
P9	Homologous crossover %	95
P10	Program size	Initial 64, maximum 256
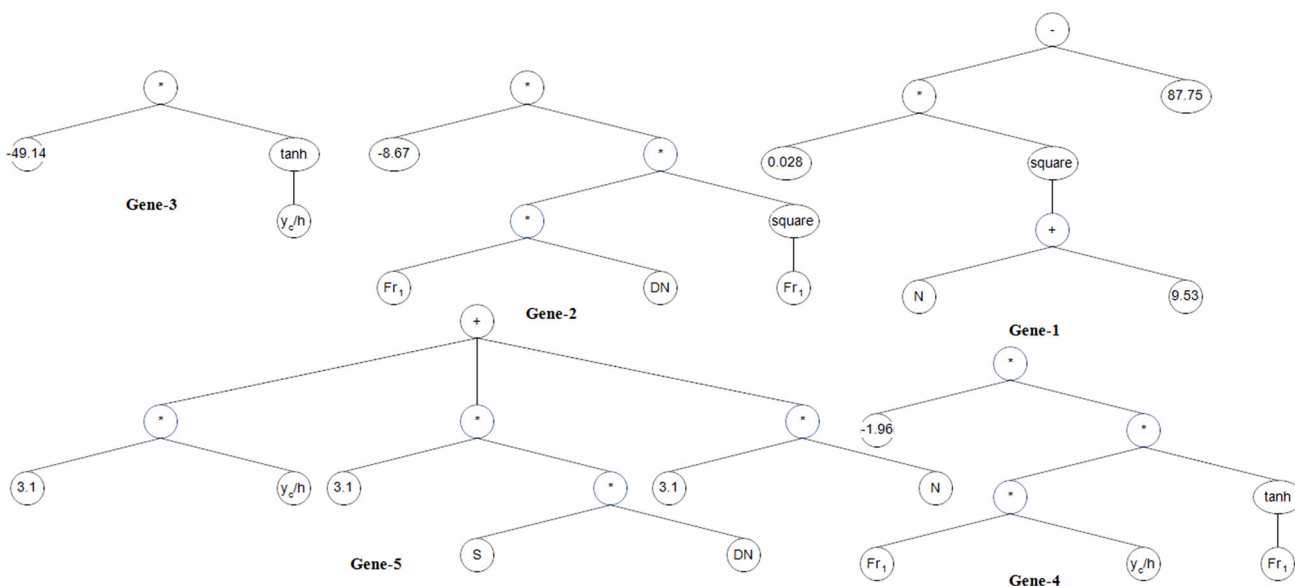
Fig. 8 Structure of derived model from GP

Developing GMDH as smart fitting function similar to other types of soft computing techniques such as MLP, SVM, and GP is based on dataset. The same dataset used for preparation of MLP, SVM, and GP was used to develop GMDH. As stated in the review on GMDH section, the number of neurons in the first layer is equal to 10. However, some of them do not deserve to attend the next layer. Each neuron is trained using training dataset and then assessed using testing dataset. Being selected to attend in the next layer is according to performance of neurons in the previous layer. Results of selected neuron in each layer are shown in Fig. 9. Values of justified coefficients and biases of equation of selected neurons are presented in Table 6. As shown in Fig. 9, proposed model includes three hidden layers. Reviewing the structure of developed model indicated that the most effective parameters for modeling energy dissipation of

flow over stepped spillways are Fr_1, DN, and y_c/h. Results of developed GMDH model in training and testing stages are presented in Figs. 6 and 7. Comparing the performance of developed GMDH model with MLP, SVM, and GP shows that the accuracy of GMDH is slightly better than MLP and GP; SVM is more accurate compared to GMDH.

Preparation of MARS model similar to MLP, SVM, GP, and GMDH is based on dataset. The same dataset used for developing applied models was used to prepare MARS model. Developing MARS model, as stated in the section of review on MARS, includes two stages of growing and pruning. In growing stage, 30 basic functions were considered and in the next stage (pruning stage) 17 basic functions were pruned. At the end, optimal MARS model with 13 basic functions was derived. The inclusive form of obtained MARS model is given in Eq. (16). Extended form of MARS model is given in Table 7. The pruning criteria
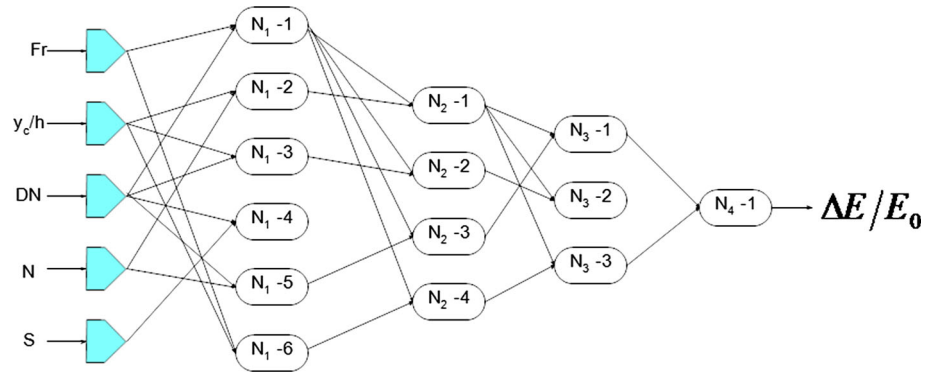
Fig. 9 Structure of developed GMDH model

Table 6 Results of adjusting parameters of transfer function of GMDH model

Layer	Neuron	b_0	b_1	b_2	b_3	b_4	b_5	RMSE
Layer-1	N_1-1	102.999	−7.096	630.329	0.012	9487.0912	−637.271	8.0246
	N_1-2	65.173	−19.268	2.384	1.170	−0.023	−0.053	10.429
	N_1-3	83.209	−8.004	−1429.212	0.429	9570.685	35.425	12.39
	N_1-4	44.125	−1657.141	2.185	13,193.853	−0.035	−1.896	13.143
	N_1-5	76.830	−1782.860	−0.649	13,414.217	0.012	4.163	13.187
	N_1-6	116.779	−11.968	−24.768	0.904	0.983	2.342	14.198
Layer-2	N_2-1	10.401	0.330	0.101	−0.003	−0.001	0.010	5.33
	N_2-2	24.276	0.436	−0.766	−0.001	0.007	0.008	7.099
	N_2-3	44.209	0.349	−1.902	−0.009	0.008	0.029	7.237
	N_2-4	−0.854	0.073	0.777	0.006	−0.008	0.005	7.994
Layer-3	N_3-1	−0.445	0.512	0.486	−0.005	−0.007	0.012	4.987
	N_3-2	−0.786	0.888	0.121	−0.005	−0.004	0.009	5.203
	N_3-3	−2.699	2.446	−1.323	−0.002	0.015	−0.014	5.697
Layer-4	N_4-1	−0.462	1.809	−0.793	−0.018	0.004	0.014	4.979

Table 7 Basic functions and related coefficients of MARS model

Basic function	Equation	Coefficient (β_m)
$h_1(x)$	BF1 = max(0, 0.001 −DN)	7615.681
$h_2(x)$	BF2 = max(0, Fr$_1$ −2.791)	−6.162
$h_3(x)$	BF3 = BF2 * max(0, DN −0.016)	435.211
$h_4(x)$	BF4 = max(0, DN −0.047)	237.326
$h_5(x)$	BF5 = BF2 * max(0, 0.0002 −DN)	13,201.877
$h_6(x)$	BF6 = max(0, 0.006 −DN)	1672.668
$h_7(x)$	BF7 = max(0, DN −0.0005)	82,585.675
$h_8(x)$	BF8 = BF2 * max(0, DN −0.003)	1705.908
$h_9(x)$	BF9 = max(0, DN −0.024)	−99,006.358
$h_{10}(x)$	BF10 = max(0, 0.024 −DN)	99,310.281
$h_{11}(x)$	BF11 = max(0, DN −0.0005)	10,659.370
$h_{12}(x)$	BF12 = BF7 * max(0, Fr$_1$ −5.225)	−4324.478
$h_{13}(x)$	BF13 = BF7 * max(0, 5.225 − Fr$_1$)	2194.232

introduced with GVC parameter in development of MARS model were derived equal to 0.0021. As mentioned in the review on MARS model, each basic function has a coefficient and a constant which is adjusted in MARS model development process and derived using least square method.

$$EPR = -2321.246 + \sum_{M=1}^{13} \beta_m h_m(x) \qquad (16)$$
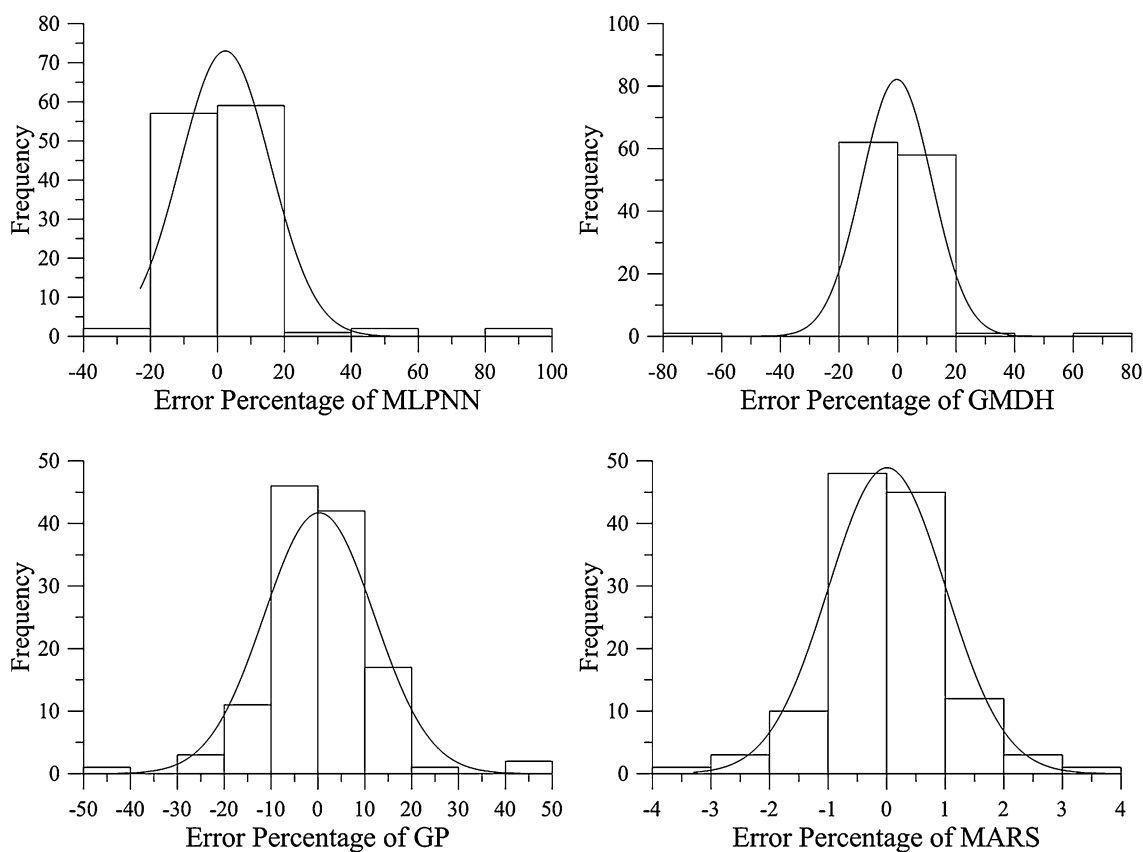
Fig. 10 Histogram of errors of applied models in training stage

Input parameters were considered regarding Eq. (5). This implies that DN, S, N, y_c/h, and Fr_1 were considered as inputs and $\Delta E/E_0$ as output. Results of MARS model during model development (preparation and testing) are shown in Figs. 6 and 7. As seen in these figures, MARS model has a high ability for modeling energy dissipation over stepped spillways. Comparing the results of MARS with MLP, SVM, GP, and GMDH shows that MARS is more accurate. Reviewing Table 7 shows that Fr_1 and DN are the most important parameters for modeling and predicting energy waste of flow over stepped spillways. This obtained point also upheld the results of GP.

Evaluation of performance of applied models regarding standard error indices provided an average value of errors of models. To provide more information about the distribution of errors through the dataset two famous approaches including histogram of errors and calculating the developed discrepancy ratio (DDR) index have been proposed. In this study, the histogram of errors of each model in development and evaluation stages of preparation are proposed, as shown in Figs 10 and 11. Histogram of errors of MLPNN shows that most errors are accumulated in the range of -20

to 20 percent in training and -15 to 25 percent for testing stage. Results of histogram of errors of SVM show that most errors are accumulated in the range of -20 to 20 percent for training and -10 to 10 percent for testing. Histogram of errors of GMDH shows that errors are accumulated in the range of -20 to 20 percent in training and -12 to 10 percent in testing stage. Our assessment of histogram of errors of applied models shows that the minimum range of error is related to MARS model. To provide an introductory view about the reliability of applied models, it is better to calculate the DDR index. The DDR index is defined as the ratio of outcome of models to the observed data. Ranges of DDR index for applied models in training and testing are given in Figs 12 and 13. As shown, the lowest DDR index is for the MARS model. This indicates that the results of MARS model are more reliable compared with others.

4 Conclusion

In this paper, energy dissipation of flow over stepped spillways was modeled and predicted using powerful soft computing techniques including multilayer perceptron
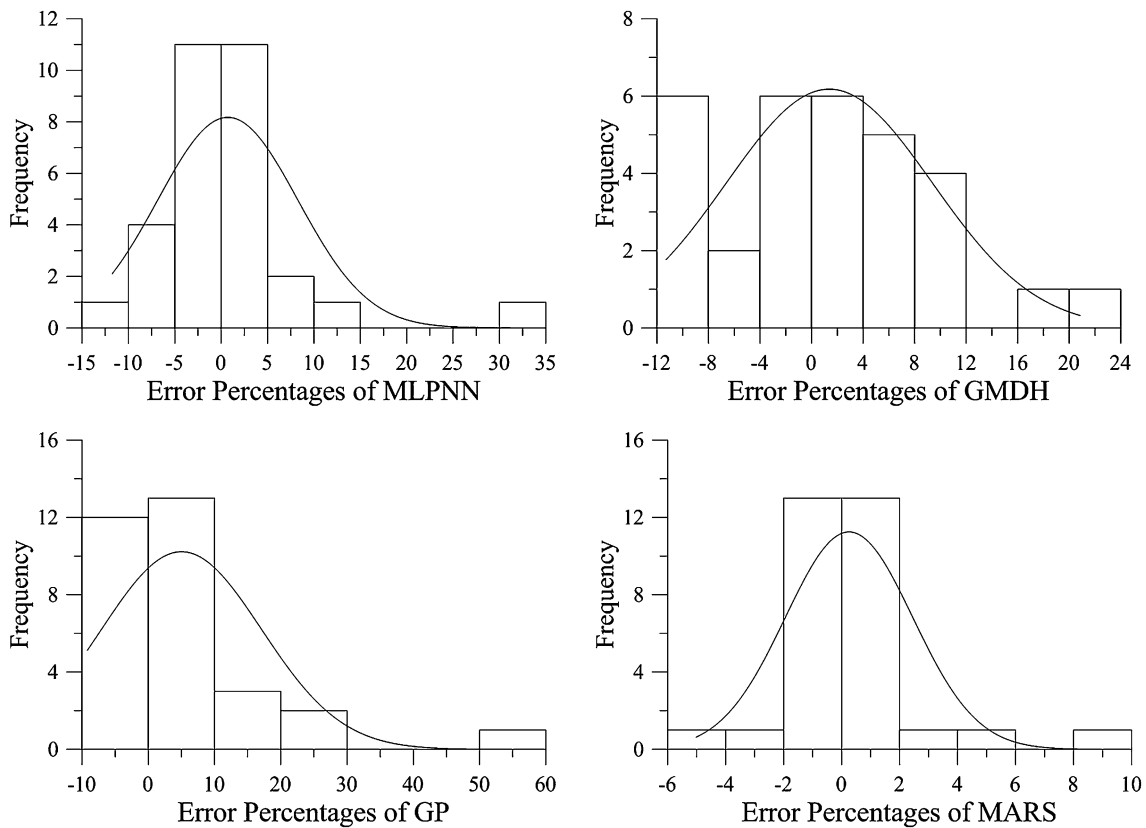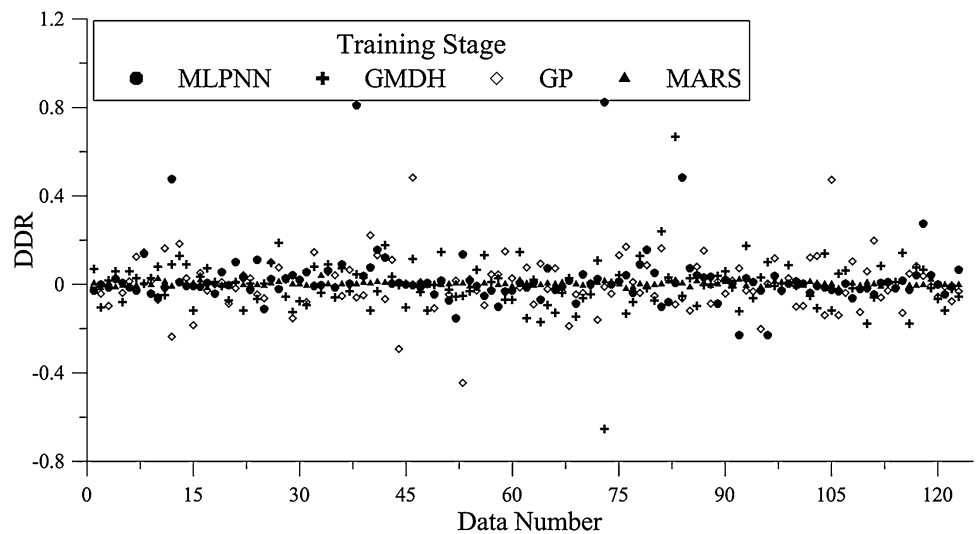
Fig. 11 Histogram of errors of applied models in testing stage

Fig. 12 Results of DDR index of applied models in training stages

neural network, support vector machine (SVM), genetic programming (GP), group method of data handling (GMDH), and multivariate adaptive regression splines (MARS). MARS, GP, and GMDH were categorized as smart function fitting techniques which, in addition to developing a network, also propose a smart function. Properties of MARS, GP, and GMDH in terms of smart

fitting function cause more weight to be assigned to inputs that are more affective on output. Results of all developed models indicated that those with suitable performance predict the ratio of energy dissipation. However, MARS model is more accurate compared to others. Reviewing the structure of obtained models from GP, GMDH, and MARS techniques revealed that drop

Fig. 13 Results of DDR index of applied models in testing stages
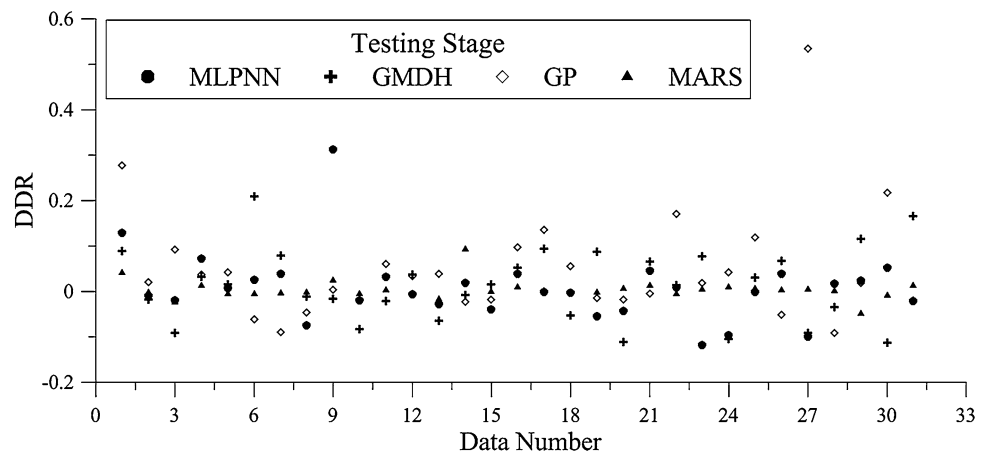
number, number of steps, and ratio of the critical depth to the height of steps are the most effective parameters on energy dissipation ratio.

References

Attarian A, Hosseini K, Abdi H, Hosseini M (2014) The effect of the step height on energy dissipation in stepped spillways using numerical simulation. Arab J Sci Eng 39:2587–2594. doi:10.1007/s13369-013-0900-y

Azamathulla HM (2013) A review on application of soft computing methods in water resources engineering. In: Gandomi AH, Talatahari S, Alavi AH (eds) Metaheuristics in water, geotechnical and transport engineering. Elsevier, Oxford, pp 27–41

Azamathulla H, Ghani A (2011) Genetic programming for predicting longitudinal dispersion coefficients in streams. Water Resour Manag 25:1537–1544. doi:10.1007/s11269-010-9759-9

Azamathulla HM, Mohd. Yusoff MA (2013) Soft computing for prediction of river pipeline scour depth. Neural Comput Appl 23:2465–2469. doi:10.1007/s00521-012-1205-x

Azamathulla HM, Wu F-C (2011) Support vector machine approach for longitudinal dispersion coefficients in natural streams. Appl Soft Comput 11:2902–2905. doi:10.1016/j.asoc.2010.11.026

Azamathulla H, Ghani A, Zakaria NA, Lai SH, Chang CK, Leow CS, Abuhasan Z (2008) Genetic programming to predict ski-jump bucket spill-way scour. J Hydrodyn Ser B 20:477–484. doi:10.1016/S1001-6058(08)60083-9

Azamathulla HM, Ghani AA, Zakaria NA, Guven A (2010) Genetic programming to predict bridge pier scour. J Hydraul Eng 136:165–169. doi:10.1061/(ASCE)HY.1943-7900.0000133

Azamathulla HM, Haghiabi AH, Parsaie A (2016) Prediction of side weir discharge coefficient by support vector machine technique. Water Sci Technol Water Supply. doi:10.2166/ws.2016.014

Boes RM et al (2000) Characteristics of skimming flow over stepped spillways. J Hydraul Eng 126:860–873. doi:10.1061/(asce)0733-9429(2000)126:11(860)

Chanson H (2002) Hydraulics of stepped chutes and spillways. Taylor & Francis, London

Chatila JG, Jurdi BR (2004) Stepped spillway as an energy dissipater. Can Water Resour 29:147–158. doi:10.4296/cwrj147

Chen SH (2015) Hydraulic structures. Springer, Berlin

Cheng X, Chen Y, Luo L (2006) Numerical simulation of air–water two-phase flow over stepped spillways. Sci China Ser E: Technol Sci 49:674–684. doi:10.1007/s10288-006-2029-2

Dehdar-behbahani S, Parsaie A (2016) Numerical modeling of flow pattern in dam spillway's guide wall. Case study: Balaroud dam, Iran. Alex Eng J 55:467–473. doi:10.1016/j.aej.2016.01.006

Emamgholizadeh S, Kashi H, Marofpoor I, Zalaghi E (2014a) Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models. Int J Environ Sci Technol 11:645–656. doi:10.1007/s13762-013-0378-x

Emamgholizadeh S, Moslemi K, Karami G (2014b) Prediction the groundwater level of Bastam Plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS). Water Resour Manag 28:5433–5446. doi:10.1007/s11269-014-0810-0

Emamgholizadeh S, Bahman K, Bateni SM, Ghorbani H, Marofpoor I, Nielson JR (2016) Estimation of soil dispersivity using soft computing approaches. Neural Comput Appl. doi:10.1007/s00521-016-2320-x

Emamgolizadeh S, Bateni SM, Shahsavani D, Ashrafi T, Ghorbani H (2015) Estimation of soil cation exchange capacity using genetic expression programming (GEP) and multivariate adaptive regression splines (MARS). J Hydrol 529(3):1590–1600. doi:10.1016/j.jhydrol.2015.08.025

Felder S, Chanson H (2011) Energy dissipation down a stepped spillway with nonuniform step heights. J Hydraul Eng 137:1543–1548. doi:10.1061/(asce)hy.1943-7900.0000455

Fen N, Kozlov DB, Rumyantsev IS (2016) Hydraulic studies of stepped spillways of various design. Power Technol Eng 49:337–344. doi:10.1007/s10749-016-0625-7

Friedman JH (1991) Multivariate adaptive regression splines. The annals of statistics 1-67

Frizell KW, Renna FM, Matos J (2013) Cavitation potential of flow on stepped spillways. J Hydraul Eng 139:630–636. doi:10.1061/(asce)hy.1943-7900.0000715

Guven A, Kişi Ö (2011) Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming. Water Resour Manag 25:691–704. doi:10.1007/s11269-010-9721-x

Haghiabi AH (2016) Prediction of longitudinal dispersion coefficient using multivariate adaptive regression splines. J Earth Syst Sci. doi:10.1007/s12040-016-0708-8

Husain SM, Muhammed JR, Karunarathna HU, Reeve DE (2014) Investigation of pressure variations over stepped spillways using smooth particle hydrodynamics. Adv Water Resour 66:52–69. doi:10.1016/j.advwatres.2013.11.013

Ivakhnenko AG (1971) Polynomial theory of complex systems. IEEE Trans Syst Man Cybern SMC 1:364–378. doi:10.1109/TSMC.1971.4308320

Karbasi M, Azamathulla HM (2016) Prediction of scour caused by 2D horizontal jets using soft computing techniques. Ain Shams Eng J. doi:10.1016/j.asej.2016.04.001

Mohammad Rezapour Tabari M, Tavakoli S (2016) Effects of stepped spillway geometry on flow pattern and energy dissipation. Arab J Sci Eng 41:1215–1224. doi:10.1007/s13369-015-1874-8

Morovati K, Eghbalzadeh A, Javan M (2016) Numerical investigation of the configuration of the pools on the flow pattern passing over pooled stepped spillway in skimming flow regime. Acta Mech 227:353–366. doi:10.1007/s00707-015-1444-x

Najafzadeh M (2016) Neurofuzzy-based GMDH-PSO to predict maximum scour depth at equilibrium at culvert outlets. J Pipeline Syst Eng Pract 7:06015001. doi:10.1061/(ASCE)PS.1949-1204.0000204

Najafzadeh M, Azamathulla HM (2015) Neuro-fuzzy GMDH to predict the scour pile groups due to waves. J Comput Civ Eng 29:04014068. doi:10.1061/(ASCE)CP.1943-5487.0000376

Najafzadeh M, Barani GA (2011) Comparison of group method of data handling based genetic programming and back propagation systems to predict scour depth around bridge piers. Sci Iran 18:1207–1213. doi:10.1016/j.scient.2011.11.017

Najafzadeh M, Bonakdari H (2016) Application of a neuro-fuzzy GMDH model for predicting the velocity at limit of deposition in storm sewers. J Pipeline Syst Eng Pract. doi:10.1061/(asce)ps.1949-1204.0000249

Najafzadeh M, Sattar AA (2015) Neuro-fuzzy GMDH approach to predict longitudinal dispersion in water networks. Water Resour Manag 29:2205–2219. doi:10.1007/s11269-015-0936-8

Najafzadeh M, Tafarojnoruz A (2016) Evaluation of neuro-fuzzy GMDH-based particle swarm optimization to predict longitudinal dispersion coefficient in rivers. Environ Earth Sci 75:1–12. doi:10.1007/s12665-015-4877-6

Najafzadeh M, Zahiri A (2015) Neuro-fuzzy GMDH-based evolutionary algorithms to predict flow discharge in straight compound channels. J Hydrol Eng 20:04015035. doi:10.1061/(asce)he.1943-5584.0001185

Najafzadeh M, Etemad-Shahidi A, Lim SY (2016) Scour prediction in long contractions using ANFIS and SVM. Ocean Eng 111:128–135. doi:10.1016/j.oceaneng.2015.10.053

Nikseresht AH, Talebbeydokhti N, Rezaei MJ (2013) Numerical simulation of two-phase flow on step-pool spillways. Sci Iran 20:222–230. doi:10.1016/j.scient.2012.11.013

Noori R, Karbassi A, Farokhnia A, Dehghani M (2009) Predicting the longitudinal dispersion coefficient using support vector machine and adaptive neuro-fuzzy inference system techniques. Environ Eng Sci 26:1503–1510. doi:10.1089/ees.2008.0360

Noori R, Karbassi A, Salman Sabahi M (2010a) Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. J Environ Manag 91:767–771. doi:10.1016/j.jenvman.2009.10.007

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

Noori R, Yeh H-D, Abbasi M, Kachoosangi FT, Moazami S (2015) Uncertainty analysis of support vector machine for online prediction of five-day biochemical oxygen demand. J Hydrol 527:833–843. doi:10.1016/j.jhydrol.2015.05.046

Parsaie A, Haghiabi A (2015a) The effect of predicting discharge coefficient by neural network on increasing the numerical modeling accuracy of flow over side weir. Water Resour Manag 29:973–985. doi:10.1007/s11269-014-0827-4

Parsaie A, Haghiabi AH (2015b) Computational modeling of pollution transmission in rivers. Appl Water Sci. doi:10.1007/s13201-015-0319-6

Parsaie A, Haghiabi AH, Moradinejad A (2015) CFD modeling of flow pattern in spillway's approach channel. Sustain Water Resour Manag 1(3):245–251. doi:10.1007/s40899-015-0020-9

Parsaie A, Haghiabi AH, Saneie M, Torabi H (2016) Applications of soft computing techniques for prediction of energy dissipation on stepped spillways. Neural Comput Appl 1–17. doi:10.1007/s00521-016-2667-z

Pfister M, Hager WH (2011) Self-entrainment of air on stepped spillways. Int J Multiph Flow 37:99–107. doi:10.1016/j.ijmultiphaseflow.2010.10.007

Roushangar K, Akhgar S, Salmasi F, Shiri J (2014) Modeling energy dissipation over stepped spillways using machine learning approaches. J Hydrol 508:254–265. doi:10.1016/j.jhydrol.2013.10.053

Salmasi F, Özger M (2014) Neuro-fuzzy approach for estimating energy dissipation in skimming flow over stepped spillways. Arab J Sci Eng 39:6099–6108. doi:10.1007/s13369-014-1240-2

Samadi M, Jabbari E, Azamathulla HM, Mojallal M (2015) Estimation of scour depth below free overfall spillways using multivariate adaptive regression splines and artificial neural networks. Eng Appl Comput Fluid Mech 9(1):291–300. doi:10.1080/19942060.2015.1011826

Sattar AMA, Gharabaghi B (2015) Gene expression models for prediction of longitudinal dispersion coefficient in streams. J Hydrol 524:587–596. doi:10.1016/j.jhydrol.2015.03.016

Sorensen RM (1985) Stepped spillway hydraulic model investigation. J Hydraul Eng 111:1461–1472. doi:10.1061/(asce)0733-9429(1985)111:12(1461)

Tabbara M, Chatila J, Awwad R (2005) Computational simulation of flow over stepped spillways. Comput Struct 83:2215–2224. doi:10.1016/j.compstruc.2005.04.005

Tatewar SP, Ingle RN (1996) Energy dissipation in skimming flow over stepped spillways ish. J Hydraul Eng 2:45–51. doi:10.1080/09715010.1996.10514591

Zahiri A, Azamathulla HM (2014) Comparison between linear genetic programming and M5 tree models to predict flow discharge in compound channels. Neural Comput Appl 24:413–420. doi:10.1007/s00521-012-1247-0

Zare HK, Doering JC (2012) Energy dissipation and flow characteristics of baffles and sills on stepped spillways. J Hydraul Res 50:192–199. doi:10.1080/00221686.2012.659840

Zhan J, Zhang J, Gong Y (2016) Numerical investigation of air-entrainment in skimming flow over stepped spillways. Theor Appl Mech Lett 6:139–142. doi:10.1016/j.taml.2016.03.003