ENGINEERING APPLICATIONS OF COMPUTATIONAL FLUID MECHANICS
2021, VOL. 15, NO. 1, 1761–1774
https://doi.org/10.1080/19942060.2021.1976280

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Forecasting the discharge capacity of inflatable rubber dams using hybrid machine learning models

Wenyu Zheng[a,b], Shahab S. Band[c], Hojat Karami[d], Sohrab Karimi[d], Saeed Samadianfard[e], Sadra Shadkani[e], Kwok-Wing Chau[f] and Amir H. Mosavi[g,h]

[a] Department of Civil Engineering and Architecture, Nanyang Normal University, Nanyang, Henan, People's Republic of China; [b] Nanyang Lingyu Machinery Co., Ltd., Nanyang, Henan, People's Republic of China; [c]Future Technology Research Center, National Yunlin University of Science and Technology, Douliou, Taiwan; [d]Department of Civil Engineering, Semnan University, Semnan, Iran; [e]Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran; [f]Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, People's Republic of China; [g]Faculty of Civil Engineering, Technische Universität Dresden,, Dresden, Germany; [h] John von Neumann Faculty of Informatics, Obuda University, Budapest, Hungary

ABSTRACT

Inflatable dams are flexible hydraulic structures that are constructed on rivers and are inflated by fluids such as air or water. This research investigates the effects of influential dimensionless factors on estimating one of the critical hydraulic characteristics of inflatable dams, namely the discharge capacity. Various parameters such as the proportion of total upstream head to dam height (H_1/D_h), the ratio of overflowing head to dam height (h/D_h), the ratio of discharge per unit width to its maximum value (q/q_{max}), the ratio of the internal pressure of the tube to its maximum value (p/p_{max}) and the ratio of the longitudinal coordinate placement of each element to x_{max} are used. A hybrid model based on the Particle Swarm Optimization (PSO) and the Genetic Algorithm (GA), PSO-GA, is proposed to improve the accuracy of the estimation by combining the advantages of both algorithms. Moreover, the performance of the model is compared with available hybrid models, including the Artificial Neural Networks (ANNs) optimized by Stochastic Gradient Descent (SGD) model (ANN-SGD) and the ANN-PSO and ANN-GA models. Finally, the performance of the algorithms is evaluated using statistical indicators such as the coefficient of determination (R^2), root mean square error (RMSE), mean absolute percentage error (MAPE) and the scatter index (SI). The results show that the internal pressure plays a vital role with respect to forecasting the discharge coefficient, and omitting it degrades the accuracy by 2.12%. In comparison with other models, the proposed PSO-GA hybrid model provides the most accurate results ($R^2 = 0.999$, MAPE $= 0.04$). Finally, comparing the results of the proposed PSO-GA with the benchmarked ANN-GA, ANN-PSO and ANN-SGD methods proves the superiority of the hybrid PSO-GA method.

ARTICLE HISTORY
Received 23 March 2021
Accepted 30 August 2021

KEYWORDS
Inflatable dams; particle swarm optimization; genetic algorithm; machine learning; artificial intelligence

1. Introduction

The measurement of discharge and fluid depth in rivers and open channels is one of the topics that often interest engineers and researchers. More often, engineers use different types of hydraulic structures to control flood and water levels (Akoz et al., 2014). One of them is an inflatable dam, which is made of rubberized tubes that are typically filled with fluids, air or water, and hooked to two ends of concrete foundations (Doty et al., 1986; Waldow & Bystrom, 2002). As an engineering solution, inflatable dams are favoured by many practitioners in the industry owing to multiple advantages that they offer, such as flexibility of height adjustment of the dam, environment-friendliness, economic structure, readiness for use for flood control, readiness for adjusting water levels for agricultural purposes, usefulness in hydropower projects, and a good alternative to steel structures in low-temperature areas and cold countries such as Japan and Canada (Chanson & Tam, 1998; ul Islam & Kumar, 2009, 2016; Waldow & Bystrom, 2002). The literature has numerous studies that focused on inflated dams. For example, Kim (2003) conducted a 2D analysis of water-filled inflatable dams to investigate critical external water levels on the tube due to internal pressure head. An explicit finite difference program, FLAC, was used and the results of numerical analysis corresponded well with experimental observations. El-Jumaily and Salih (2005) investigated the performance of air and water

CONTACT Amir H. Mosavi amir.mosavi@mailbox.tu-dresden.de; Shahab S. Band shamshirbands@yuntech.edu.tw

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

inflated dams under different upstream and downstream water depths and compared their empirical findings with theoretical results. They observed acceptable agreement between the two outcomes. Cheraghi-Shirazi et al. (2014) presented a numerical model for inflatable dams under different internal pressures and upstream/downstream water heads. They found that the effect of some parameters such as the coefficient of elasticity, interior pressure, channel width and the dam's thickness were in equilibrium with the height of the inflatable dam. Recently, many novel artificial intelligence methods have evolved for parameter estimation in engineering problems. For example, Karimi et al. (2015b) determined the discharge capacity of a triangular labyrinth side weir utilizing the Multi-Layer Neural Network (ANN-MLP) method. Abdin and Abdeen (2007) introduced an ANN method with a high degree of proficiency to simulate the efficacy of submerged lentic weeds on the hydraulic efficiency of branched open channel systems.

Different soft computing methods, on the basis of empirically collected datasets, have been used to simulate and optimize the required parameters in many engineering problems with high accuracy. For instance, Parsaie (2016) estimated the discharge coefficient of side weirs utilizing empirical functions, MLP and radial basis function (RBF) neural networks. Their results indicated that the RBF model had the best accuracy amongst empirical functions and artificial intelligence methods. The adoption of computational intelligence algorithms has been extended to different sectors of engineering sciences including water and hydraulic structures in order to estimate the mean velocity of flow, and to predict the discharge coefficient, sediment transport and scour. Karimi et al. (2015a) estimated the mean flow velocity in different widths of intake channel through the use of Gene Expression Programming (GEP), and their results showed highly efficient prediction by the model. Ebtehaj et al. (2015) simulated the discharge coefficient in a side weir using the Group Method of Data Handling (GMDH) and demonstrated the high efficiency of this method. Ebtehaj et al. (2016) estimated sediment transport in open channels using a combined Feed-Forward Neural Networks (FFNNs) and an Extreme Learning Machine (ELM) algorithm (FFNN-ELM) and proved its high capability for sediment estimation. Goel (2008) modelled scour downstream of spillways by the support vector machine (SVM) method and proved the appropriate potential of this method for predicting scour downstream of spillways. Several research works have shown that combining more than one optimization techniques can be effective and beneficial. For example, Zaji and Bonakdari (2002) tried to estimate the discharge coefficient of rectified side weirs using various neural network and particle swarm optimization approaches. They found that all methods used in that study had a low estimation error, but equations obtained from PSO could be used in practical projects. Zaji et al. (2015) predicted the discharge coefficient of a side weir using radial basis neural networks and particle swarm optimization. Their results indicated that the MNLPSO model showed lower error compared to MLPSO and RBNN methods. Jung and Karney (1991, 2006) and Chang et al. (2013) used PSO and GAs for the optimization of transient hydraulic protection devices and water resources utilization, respectively. They introduced PSO and GA techniques as powerful optimization techniques that were capable of solving and modelling many complex problems.

In the present research, both approaches are combined in order to produce a hybrid modelling approach that can be used to improve the accuracy of predicting the discharge coefficient of inflatable dams. To this end, dimensional analysis is performed in aiding to determine factors that can influence the estimation of the discharge capacity with six different models, considering each of the dimensionless parameters under different conditions with the aim of underpinning the most influential input variables through training and testing phases. Four statistical indicators (R^2, RMSE, MAPE and SI) are utilized to appraise the accuracy of the results of the models. The most important drawbacks of this type of work based on previous articles presented in this field were often the high errors of results using experimental formulas, and moderate amounts of error using artificial intelligence methods. In this study, the performance of an ANN model is strengthened for the prediction of the discharge capacity of inflatable rubber dams using hybrid algorithms including PSO, GA and SGD. Additionally, to the best knowledge of the authors, an SGD algorithm is used for the first time for forecasting the discharge capacity of inflatable rubber dams.

2. Materials and methods

In this research, a hybrid particle swarm optimization—genetic algorithm (PSO-GA) model is designed to estimate the discharge coefficient (C_d) of an air-inflated dam. Some brief explanations of this model are given as follows.

2.1. The Genetic Algorithm (GA)

The GA is an intelligent search method based on genetic and chromosome structures (Ebtehaj & Bonakdari, 2014). It is a computational optimization algorithm that considers a set of points in solution space at each computational iteration and searches different areas of the

solution space efficiently. In this search mechanism, the fitness function is not computed at all points of the solution space but the target function value for each point is inferred to get an average statistical target function in all subspaces that are interdependent with that point (Ebtehaj & Bonakdari, 2014). These subspaces are averaged statistically. The search space leads to areas where the statistical mean is high, and the existence of an absolute optimal point is more possible. In this method, in contrast to a one-way procedure, the solution space is searched comprehensively, and therefore the possibility of converging to a local optimum point is low. Additionally, this algorithm has no limitation for optimizing functions such as derivative and continuity functions (Ebtehaj & Bonakdari, 2014). This algorithm only requires the amount of target function at different points to be determined and does not use additional information such as the derivative of the target function (Ebtehaj & Bonakdari, 2014). Genetic algorithm can be considered in a number of problems including linear, nonlinear, continuous and separated. Moreover, it is easily compatible with different problems. In general, it simulates the biologic evolution process using selection, crossover and mutation operators.

The selection operator is one of the remarkable processes for selecting parents to produce new populations that affect the convergence of genetic algorithms (Ebtehaj & Bonakdari, 2014). Goldberg and Deb first evaluated the speed of convergence of different selection states (Goldberg & Deb, 2008). In the roulette wheel or probable sampling method, which is applied in this research, chromosomes with higher fitness are given more chances to be considered for being parents. The probability for each chromosome i is indicated in Equation (1) (Ebtehaj & Bonakdari, 2014):

$$p_i = f_i / \sum_{j=1}^{N} f_i \tag{1}$$

where f_i and N denote the fitness of chromosome i and the population size, respectively. Crossover and mutation are two other operators in genetic algorithms. Crossover simultaneously operates on two chromosomes and combines features of the chromosomes to produce a new generation (Ebtehaj & Bonakdari, 2014). The movement of people among the sub-population in which the best individual of a sub-population is replaced by the worst people in the population is called migration. Random selection of a cut point and producing a new generation via combining a fraction of one of the parents to the left side of the cut point with a fraction of the other parent to the right side is a simple method to obtain crossover (Ebtehaj & Bonakdari, 2014). In mutation, some parts of the chromosome are changed randomly to obtain a better response and escape from the optimization area. In fact, in this mode, characteristics are generated that are not found in the father.

2.2. Particle Swarm Optimization (PSO)

The PSO algorithm is inspired by the concurrent flight of birds, the shoalling of fish, and their social lives. This algorithm uses a series of simple equations. In this algorithm, each particle deputizes a probable answer to move randomly in the problem space. The experience of neighbouring particles and the knowledge of each particle affect the changing location of each particle in the search space (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). A particle's position influences the search process of the particle itself. Modelling such socio-biological instincts creates a search procedure in which particles desire and tend towards suitable areas. Each particle in the group learns from the others and, according to the knowledge obtained, tends to go to the best of their neighbours (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015).

The operation of the algorithm is based on the theory that each particle will regulate its position in the search space according to the best place located there and the best location within the entire neighbourhood (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). It is assumed that there is a D-dimensional space and that the ith particle of the population is designated a velocity vector and a position vector (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). Any change in the location of each particle is the result of changing its velocity and position in the previous conditions (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). The best value ever reached (*pbest*) and the position of x_i are presented by each particle. Each particle comes to determine the best answer by identifying the best results ever obtained from *pbest* (*gbest*) (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). Each particle then tries to change its position to gain the best answer using the information below:

$$V_i^{k+1} = wv_i^k + c_1 r_1 \cdot (pbest_i - x_i^k) + c_2 r_2 \cdot (gbest - x_i^k) \tag{2}$$

where V_i and x_i denote the current position and current speed, respectively (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). Equation (2) shows the speed changes of each particle. In Equation (2), v_i^k is the speed of each representative in repetition k_i; w, c_1, c_2, r_1 and r_2 denote weighting parameter, weight factors, and random numbers between zero and one, respectively; x_i^k denotes the location of each particle i at iteration k; $pbest_i$ denotes

pbest in particle *i* and; *gbest* denotes *gbest* in the group. Equation (3) demonstrates the position of each particle (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015):

$$x_i^{k+1} = x_i^k + v_i^{k+1} \tag{3}$$

In Equation (2), the weighting parameter w guarantees the convergence of the algorithm. The weighting parameter controls the effect of previous speeds on current rates. The balance between functionality and comprehensive exploration of the local group is represented by an appropriate weighting parameter value (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). A suitable choice of weighting factor will decrease the number of repetitions to find the optimal solution. Inertia constant factor greater than one increases the effect of the previous rate in developing the algorithm search space, but renders the algorithm unstable. The amount of W was first constant. Shi and Eberhart's researches (Taddy, 2019; Tsuruoka et al., 2009) created great changes to this algorithm. They changed the weighting parameter linearly from the maximum amount to its minimum amount as shown in Equation (4):

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{iter_{\max}} \cdot iter \tag{4}$$

where $w_{\max}$ and $w_{\min}$ denote initial and final weights, *iter* and $iter_{\max}$ illustrate a number of repetitions and the largest number of repetitions. In Equation (2), C_1 and C_2 parameters are critical for the convergence of the algorithm (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015). An appropriate answer may converge sooner and avoid the possibility of being stuck in local minimum points. The parameters r_1 and r_2 affect the uniformity of solutions and this parameter is randomly selected between zero and one. The value of the particles allows them to move in random paces in the range from gbest to pbest (Ebtehaj & Bonakdari, 2016; Kennedy, 2015b; Zaji et al., 2015).

2.3. Stochastic Gradient Descent (SGD)

SGD is a method for improving an objective function with suitable softness properties (Taddy, 2019). SGD method, unlike the massive gradient descent method, changes one step for each data such as $x^{(i)}$ and its corresponding value $y^{(i)}$:

$$\theta = \theta - \eta \cdot \nabla_\theta J(\theta; x^{(i)}; y^{(i)}) \tag{5}$$

The massive descending gradient method performs additional and unnecessary computations for large data sets because the gradient computes the same data repeatedly before making any changes. The random method eliminates these additional computations by making a change with each gradient. For this reason, this method is much faster and can have instant learning. The random method makes changes with high variance, which causes the objective function to fluctuate sharply.

While the mass gradient method always converges to the minimum of the same arc that encloses the parameters, the fluctuations of the random method on the one hand enable it to find better relative minimum and on the other hand cause the final convergence to be accurate. Minimization is harder because this method has a large deviation from the original value in each iteration. However, it has been proven that when the learning rate is slowly increased, this method shows the same convergence of the mass method with high reliability, converges to absolute minimum at convex levels and to relative minimum at non-convex levels. Besides, in each iteration, the order of the data is changed randomly (Tsuruoka et al., 2009).

SGD uses a randomly-selected subsection of the training examples to predict the gradient of the target function given by Equation (5) (Tsuruoka et al., 2009).

$$\alpha_w = \sum_{i=1}^{N} K(j \cdot w) - C \sum_i |\omega_i| \tag{6}$$

where N is the batch size, C is the meta-parameter and ω_t is the weight of the feature. The updated weights of the features at training instance j are presented as follows:

$$\omega^{k+1} = \omega^k + \eta_K \frac{\partial}{\partial \omega}(K(j \cdot \omega) - \frac{C}{N} \sum_i |\omega_i|) \tag{7}$$

In Equations (6) and (7), K is the iteration numerator and η_K is the learning rate, which decreases as the iterated son advances.

2.4. Particle Swarm Optimization–Genetic Algorithm (PSO-GA)

The PSO and GA algorithms are two kinds of population-based evolutionary algorithms. These two algorithms have similar properties, but each of them has advantages and disadvantages (Bilhan et al., 2010; Chang et al., 2013; Zaji & Bonakdari, 2002). Requiring a lot of functions for examining the model and high computation costs are some of the disadvantages of GA. Besides, different GA operators, such as selection, mutation, and crossover, convergence, decrease the speed of the algorithm (Bilhan et al., 2010; Chang et al., 2013; Zaji & Bonakdari, 2002). PSO algorithm has memory, yet causes loss of information of individual not chosen. In PSO, common action and reaction increase the search for an optimal solution. However, in GA, investigating optimal solutions in a

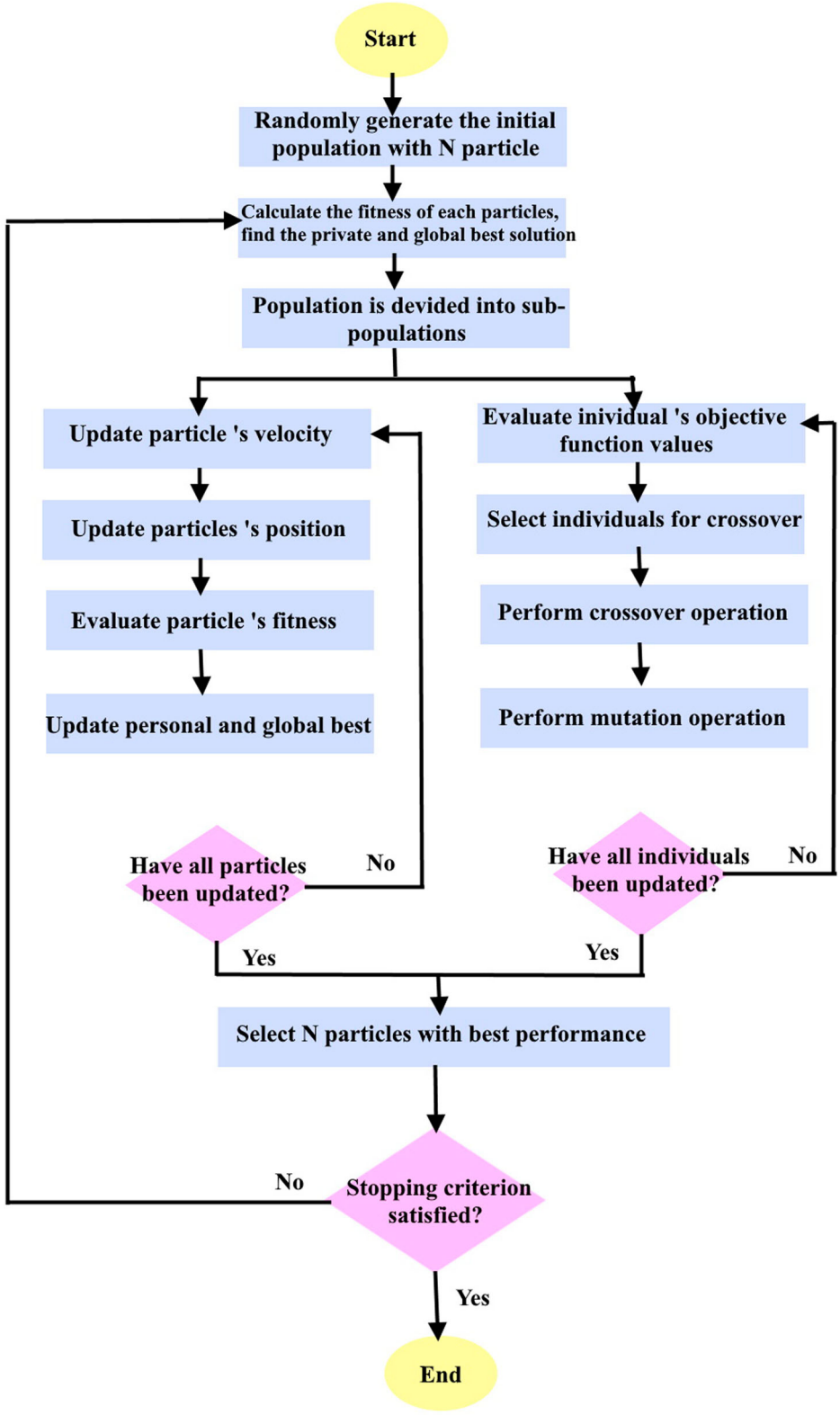


Figure 1. Flowchart of the PSO-GA hybrid algorithm.

globally optimal area have encountered problems (Bilhan et al., 2010; Chang et al., 2013; Zaji & Bonakdari, 2002).

In this research, a new hybrid algorithm with advantages of both PSO and GA algorithms, termed PSO-GA, is created to boost the accuracy of the search (Zaji & Bonakdari, 2002). PSO-GA method is used for computing global solutions. For this purpose, this method considers the entire search space that combines

standard and fast regulations updated the position in the PSO with the concepts of selection, mutation, and crossover in GA (Figure 1) (Chang et al., 2013; Smith, 1993; Zaji & Bonakdari, 2002). Some of the individuals having similar properties with those in the previous generation, but in a different position in the search space, are replaced using alternative operators by applying GA operators according to PSO, produced individuals, and remaining individuals. This updating mechanism constitutes the evolution in PSO standard algorithm. The updating principle of PSO algorithm is shown in Equations (8) and (9) (Chang et al., 2013; Smith, 1993; Zaji & Bonakdari, 2002):

$$v_j^{t+1} = wv_j^t + c_1 rand_1(pbest_j(\mathrm{GA}) - k_j^t) + c_2 rand_2(gbest(\mathrm{GA}) - k_j^t) \quad (8)$$

$$k_j^{t+1} = k_j^t + v_j^{t+1}, \qquad k^{\min_j^{t+1}\max} \quad (9)$$

In Equations (8) and (9), $pbest_j$ (GA) indicates the best answer obtained by particle j after action by the GA operators, and *gbest* (GA) indicates the best solution in the group after action by the GA operators (Zaji & Bonakdari, 2002). A setting method is utilized in the optimal design of performance upon PSO-GA algorithm (Figure 1) (Chang et al., 2013; Smith, 1993; Zaji & Bonakdari, 2002). For allocating priority to PSO or GA, two driving parameters are added to the algorithm. K_1 and K_2 are two effective terms that are used in PSO and GA, respectively ($K_1 + K_2 = 1$). K_1 determines the number of individuals that can be transmitted to the next generation and K_2 expresses the number of individuals that can replace the current generation. If K_1 or K_2 is equal to zero, GA or PSO (respectively) does not have any effect on individuals. To explain the effectiveness of PSO-GA, the following points are mentioned: in PSO-GA, the quality of the solution (competence) in the upgrade method is considered. Factors that are close to a good solution attract other factors that are exploring various sections of the search space. When all factors are close to an appropriate result, they move very slowly. Therefore, *gbest* supports them to find the best overall solution. Using memory, the PSO-GA algorithm retains the best *gbest* solution ever discovered so that it is available at all times. Each factor can see the best *gbest* solution and lean towards it. The above points make the PSO-GA algorithm powerful enough to solve a wide range of optimization problems.

2.5. Artificial Neural Networks (ANNs)

The advantages of the ANN method include good performance in analysing complex flows and nonlinear flow patterns. ANN's flexible structure enables it to model complex and/or nonlinear patterns between input and output data. Neural network training means gaining network weights. The classification of ANN types is based on methods used to obtain weights as well as transfer functions used. A multilayer perceptron neural network, as one of the most widely used types of ANN, consists of an input layer, one or more hidden layers and an output layer, see Figure 2 (Kisi et al., 2012; Smith, 1993; Steinfeld et al., 2015). Each layer consists of a number of neurons. In the neural network studied in here, sigmoid and linear activation functions are used for hidden and output layer neurons, respectively. Various functions can be considered as sigmoid functions and, in this study, the hyperbolic tangent is utilized as an activation function in hidden layers. The Lunberg–Marquardt method is used to train the ANN. In this method, the back propagation algorithm is utilized to find the weights and biases of the neural network. The implemented ANN model has two hidden layers. Moreover, the number of neurons in each hidden layer is determined utilizing a trial-and-error procedure (Kisi et al., 2012; Smith, 1993; Steinfeld et al., 2015).

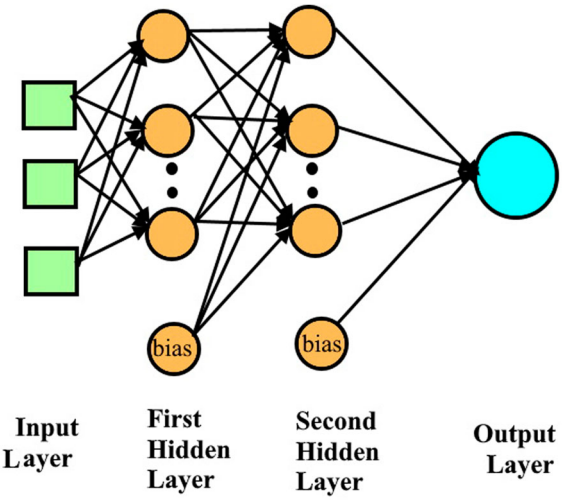


Figure 2. Architecture of a multi-layer perceptron network.

2.6. Genetic Programming (GP)

GP, as one of the evolutionary algorithms, has a tree structure. The recursive form of assessment and estimation of trees is easy. Any node of a GP structure has an operator function mode, and each terminal node is contained with an operand that causes improvement as well as an assessment of mathematical statements (Koza & Koza, 1992; Khan et al., 1998; Steinfeld et al., 2015). The process of GP is such that functions required to create the final model and set are selected first. In the next

step, the available data set is called to predict the desired parameters and compare them with the actual amounts. Chromosomes are then generated randomly to represent the initial population. In the next step, for the population produced using the available chromosomes, the program is executed, and the suitability of the objective function is checked. If a program stopping criterion is fulfilled, the program is stopped; otherwise, new chromosomes that are genetically modified are used in the new population, and the target function is reappraised. This operation will continue until the conditions for stopping the program are met (Ferreira, 2001a; Koza & Koza, 1992; Khan et al., 1998; Steinfeld et al., 2015).

The fitness of a unique program i for the fitness of model j by Ferreira (2001) is presented as follows:

if $E(ij) \leq p$, then $f(ij) = 1$; else $f(ij) = 0$

In the above relation, p is the precision and $E(ij)$ is the error of program i for the fitness of model j, which is expressed as follows for absolute error:

$E(ij) = |p(ij) - T_j|$

The fitness value of a unique program (f_i) is also expressed as follows: $f_i = \sum \left(R - |p_{(ij)} - T_j\right)$

In the above, the relation R of the selection domain $p(ij)$ is the value predicted by the unique program i for the fitness of model j and T_j is the target value for the fitness of model j. It should be noted that the parameters of GEP are determined utilizing a trial-and-error procedure to obtain the most accurate predictions.

3. Experiments

3.1. Experimental setup

Empirical data provided by Alhamati et al. (2005) are utilized for estimating the discharge coefficient of an inflatable dam. The laboratory studies are performed on a 20 m long, 0.9 m wide and 0.6 m deep channel. The inflatable dam model is placed at 8.5 m from the upstream end of the channel. A rectangular weir is fixed at the beginning of the channel to measure the discharge and an air compressor is utilized to inflate the air in the dam. To prevent leakage between the channel walls, an extra length (20 cm from both sides of the model) is added to the model length. Figure 3 shows a schematic sketch of the inflatable dam used in this experiment, and Table 1 lists the statistical values of the input parameters used to estimate the discharge coefficient.

Table 1. Ranges of parameters used to estimate the discharge coefficient.

	p/p_{max}	q/q_{max}	H_1/D_h	h/D_h	x/x_{max}	C_d
Minimum	0.3	0.2565	1.1194	0.1194	0.11191	0.298
Maximum	1	1	1.28	0.276	1	0.364
Average	0.6111	0.4298	1.1618	0.1618	0.44	0.3231
St. dev.	0.1918	0.2502	0.0523	0.0523	0.4301	0.02134
Variance	0.0368	0.0626	0.0027	0.0027	0.1851	0.0004

The following dependent parameters are utilized for estimating the discharge coefficient (C_d):

- (H_1/D_h): the ratio of total upstream head to dam height,
- (h/D_h): the ratio of overflow head to dam height,
- (q/q_{max}): the ratio of discharge per width unit to its maximum value,
- (p/p_{max}): the ratio of internal pressure of the tube to its maximum value,
- (x/x_{max}): the ratio of longitudinal coordinate placement of each element to x_{max}.

To specify the efficacy of each input parameter (q/q_{max}, p/p_{max}, H_1/D_h, h/D_h, x/x_{max}) on the output parameter (C_d), the coefficient of determination (R^2) is used (Sharafi et al., 2016). The closer this coefficient is to one, the more effective is the input parameter. The computed R^2 for q/q_{max}, p/p_{max}, H_1/D_h, h/D_h and x/x_{max} data are 0.88, 0.59, 0.860, 0.86 and 0.15, respectively, indicating that q/q_{max} and x/x_{max} range from the highest to the

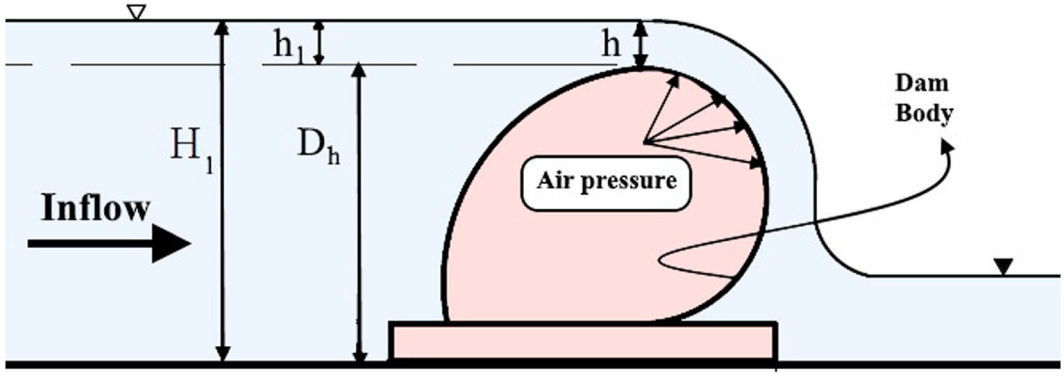


Figure 3. A section view of the air-inflated rubber dam.

lowest correlation. From the results of a regression correlation test, it can be expected that x/x_{max} is the least effective input data for forecasting C_d, and q/q_{max} is the most effective input data.

Six models are utilized to predict the discharge coefficient (C_d). As shown in Equations (10)–(15), Model 1 uses all the parameters, while each of the remaining models opts out one of the input parameters to observe its effect on the prediction outcome:

$$\text{Model 1:} \quad C_d = f\left(\frac{p}{p_{max}}.\frac{q}{q_{max}}.\frac{h}{D_h}.\frac{H_1}{D_h}.\frac{x}{x_{max}}\right) \tag{10}$$

$$\text{Model 2:} \quad C_d = f\left(\frac{p}{p_{max}}.\frac{q}{q_{max}}.\frac{h}{D_h}.\frac{H_1}{D_h}\right) \tag{11}$$

$$\text{Model 3:} \quad C_d = f\left(\frac{p}{p_{max}}.\frac{q}{q_{max}}.\frac{h}{D_h}.\frac{x}{x_{max}}\right) \tag{12}$$

$$\text{Model 4:} \quad C_d = f\left(\frac{p}{p_{max}}.\frac{q}{q_{max}}.\frac{H_1}{D_h}.\frac{x}{x_{max}}\right) \tag{13}$$

$$\text{Model 5:} \quad C_d = f\left(\frac{p}{p_{max}}.\frac{h}{D_h}.\frac{H_1}{D_h}.\frac{x}{x_{max}}\right) \tag{14}$$

$$\text{Model 6:} \quad C_d = f\left(\frac{q}{q_{max}}.\frac{h}{D_h}.\frac{H_1}{D_h}.\frac{x}{x_{max}}\right) \tag{15}$$

3.2. Evaluation criteria

To assess the accuracy of the measured C_d by the hybrid PSO-GA model, different statistical indices, such as the coefficient of determination (R^2), the Root Mean Square Error (RMSE), the Mean Absolute Percentage Error (MAPE), SI and δ, are utilized as shown in Equations (16)–(20):

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

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

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i-y_i|}{x_i} \tag{18}$$

$$\text{SI} = \frac{\text{RMSE}}{\bar{x}} \tag{19}$$

$$\delta\% = \left(\sum_{i=1}^{N}|(y_i-x_i)|/\sum_{i=1}^{N}y_i\right)\times 100 \tag{20}$$

Table 2. Adopted optimum parameters of PSO-GA.

Parameter	Value
Maximum number of iterations	1200
Population size	400
Maximum number of sub-iterations for GA	50
Maximum number of sub-iterations for PSO	50
Weight factor in PSO	2

where y_i and x_i denote the forecasted and real discharge coefficients, respectively. $\bar{y}$ and $\bar{x}$ denote the average forecasted and real discharge coefficients, respectively.

Moreover, Taylor diagrams are utilized to compare and appraise the accuracy of the studied models. Moreover, to understand the difference between the measured and estimated values, different points with different colours at the polar pole are used in Taylor diagrams.

4. Results and discussion

The accuracies of six different models for estimating the discharge coefficient of the inflatable dam are reported at this section. Furthermore, given that there is no specific way to split training and test data in data-driven methods, different researchers have used different ratios, e.g. Diop et al. (2020), Esmaeilzadeh et al. (2017), Kargar et al. (2019) and Samadianfard et al. (2014). For developing the studied ANN-PSO, ANN-GA and PSO-GA models, the data are split into 75% training section and 25% testing section. In Models 1–6, the dimensionless parameters—the ratio of total upstream head to dam height (H_1/D_h), the ratio of overflow head to dam height (h/D_h), the ratio of discharge per width unit to its maximum value (q/q_{max}), the ratio of the internal pressure of the tube to its maximum value (p/p_{max}) and the ratio of the longitudinal coordinate placement of each element to x_{max}—are used as independent parameters to forecast the discharge coefficient. To have a controlled experiment, the impacts of dropping out one of the parameters on the forecasted discharge coefficient are investigated using the PSO-GA algorithm and compared to empirically observed ones as shown in Figure 4. Additionally, according to previous studies (Gharabaghi et al., 2018; Gholami et al., 2018), the adopted optimum parameters of PSO-GA are presented in Table 2.

The best result is spawned by Model 1 ($R^2 = 0.993$), followed by Model 2 ($R^2 = 0.973$). In Model 2, the discharge coefficient of the inflatable dam is estimated with a relative approximate error of 0.74% and an acceptable R^2 of 0.973. This model suggests that removing the dimensionless parameter x/x_{max} still leads to good results and hence it has no significant impact.

In Model 4, omitting the ratio of overflow head to dam height (h/D_h) has a considerable efficacy on the

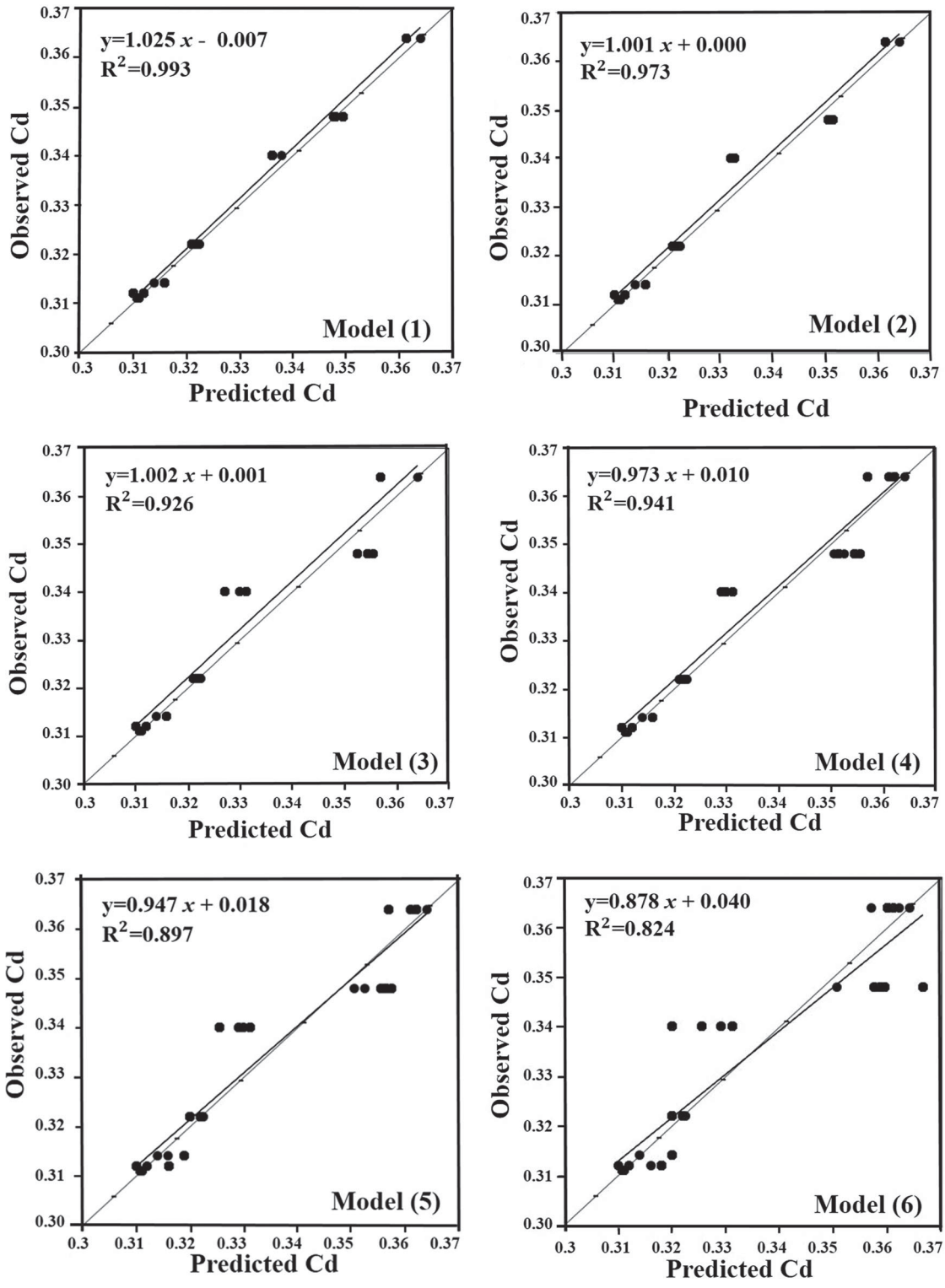


Figure 4. Comparing forecasted C_d with observed outcome in the training phase.

estimation of the discharge coefficient, as the accuracy decreases and the determination coefficient is nearly $R^2 = 0.941$.

Model 3 evaluates the effect of the ratio of total upstream head to dam height (H_1/D_h) on the discharge coefficient, where a determination coefficient of $R^2 = 0.926$ is obtained. The outcomes reveal that the calculation accuracy of Model 3 is lower than that of Models 1, 2 and 4. Model 3 also produces lower values compared to empirical values at some points.

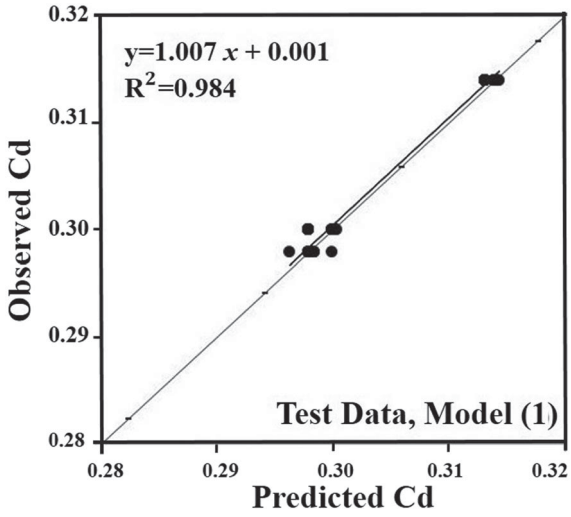


Figure 5. Comparing forecasted C_d with observed one under Model 1 in the testing phase.

Eliminating the discharge parameter in Model 5 results in a lower prediction accuracy for the discharge coefficient and degrades the fitness of R^2, which drops to 0.897. Consequently, the discharge parameter plays an important role in estimating the discharge coefficient.

Model 6 examines the effect of the absence of internal pressure of the tube on the prediction accuracy. The model produces estimations with a higher degree of variation, above and below, around the actual values of the discharge coefficient. The results show that removing this input parameter decreases the relative error by 1.82% and decreases accuracy compared to other models.

As depicted in Figure 5, the predicted values of the discharge coefficients are compared to the actual ones under the testing phase of Model 1. The determination coefficient shows a high fitness value close to 1.0 ($R^2 = 0.984$) as the predicted values are close to the real ones as summarized in Table 3 with a relative error of 0.57%.

Table 4 provides an overall summary of the performance of all models against one another using various statistical tools. Models 1 and 2 show top results with $R^2 = 0.993$ and $R^2 = 0.73$, respectively. Using RMSE to evaluate errors in the model's prediction, Model 1 poses the lowest RMSE value as it is closest to zero. Similarly, the RMSE of Model 2 is as low as 0.002, which confirms that the absence of the dimensionless parameter x/x_{max} has no impact on the estimation of the discharge coefficient. MAPE, as an error-index, reflects the gap between the estimated and observed data as a percentage of the observed values. The maximum error (1.82%) is posed in Model 6, where dropping out the internal pressure of the tube parameter increases the error notably. On the other hand, Model 1 exhibits the best MAPE value, as low as 0.46%, which reflects a higher accuracy of the model. The SI index is another error measurement indicator that is computed by dividing the RMSE index by the mean observation values. Hence, the closer it is to zero, the more accurate the prediction is. With reference to the SI value, Model 6 produces the worst SI value (0.019) while Model 1 has the best one (0.004).

Table 3 shows error indicators for different models in the testing phase. The least value of R^2, which is around 0.811, is found for Model 6, denoting lower prediction accuracy. Model 1 gives the closest estimates to real values and thus its R^2 is high and near 1.0 (0.984). Model 1 obtains an RMSE index equal to 0.001, which indicates a low prediction error. The highest MAPE error index is attributed to Model 6, which is approximately 2.12%. The results indicate that Models 1 and 2 have top performance for estimating the discharge coefficient, and Model 1 has the best prediction accuracy. Table 5 summarizes the results obtained under Model 1.

In Table 6, the performance of the hybrid algorithms. including ANN-GA, ANN-PSO and ANN-SGD, are assessed. According to Table 6, the SGD algorithm,

Table 3. Performance evaluation of various models in the testing phase.

	PSO-GA						GP	ANN
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 1	Model 1
R^2	0.984	0.961	0.918	0.937	0.876	0.811	0.857	0.853
RMSE	0.001	0.002	0.004	0.003	0.005	0.008	0.002	0.004
MAPE (%)	0.57	0.87	1.34	1.03	1.58	2.12	0.69	2.00
SI	0.004	0.007	0.012	0.010	0.015	0.019	0.009	0.022

Table 4. Performance evaluation of various models in the training phase.

	PSO-GA						GP	ANN
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 1	Model 1
R^2	0.993	0.973	0.926	0.941	0.897	0.824	0.818	0.878
RMSE	0.001	0.002	0.004	0.003	0.004	0.006	0.007	0.005
MAPE (%)	0.46	0.74	1.22	0.99	1.36	1.82	2.05	1.91
SI	0.004	0.007	0.012	0.010	0.013	0.019	0.031	0.023

Table 5. Estimated discharge coefficient using the hybrid PSO-GA under Model 1.

p (KN/m^2)	q (m^2/s/m)	$H1/D1$	$h/D1$	x(m)	C_d (EXP)	C_d (PSO-GA)
4	0.005,833,333	1.121,156	0.121,156	−2.15E-03	0.314	0.314,001
4	0.005,833,333	1.121,156	0.121,156	−2.58E-02	0.314	0.313,96
4	0.005,833,333	1.121,156	0.121,156	−4.30E-02	0.314	0.313,929
4	0.005,833,333	1.121,156	0.121,156	8.51E-01	0.314	0.314,045
4	0.005,833,333	1.121,156	0.121,156	7.87E-01	0.314	0.314,138
4	0.005,833,333	1.121,156	0.121,156	7.19E-01	0.314	0.314,221
4	0.005,833,333	1.121,156	0.121,156	6.38E-01	0.314	0.314,296
2	0.005,833,333	1.130,964	0.130,964	−2.95E-17	0.298	0.297,973
2	0.005,833,333	1.130,964	0.130,964	8.60E-03	0.298	0.297,987
2	0.005,833,333	1.130,964	0.130,964	5.15E-01	0.298	0.298,337
2	0.005,833,333	1.130,964	0.130,964	8.13E-01	0.298	0.298,079
2	0.005,833,333	1.130,964	0.130,964	7.15E-01	0.298	0.298,202
2	0.005,833,333	1.130,964	0.130,964	6.38E-01	0.298	0.298,272
1.5	0.005,833,333	1.133,799	0.133,799	−2.15E-03	0.3	0.299,895
1.5	0.005,833,333	1.133,799	0.133,799	−2.15E-03	0.3	0.299,895
1.5	0.005,833,333	1.133,799	0.133,799	−2.15E-03	0.3	0.299,895
1.5	0.005,833,333	1.133,799	0.133,799	8.17E-01	0.3	0.300,079
1.5	0.005,833,333	1.133,799	0.133,799	7.32E-01	0.3	0.300,147
1.5	0.005,833,333	1.133,799	0.133,799	6.38E-01	0.3	0.300,197

Table 6. Performance of the hybrid algorithms.

	PSO-GA	ANN-GA	ANN-PSO	ANN-SGD
R^2	0.999	0.942	0.933	0.957
RMSE	0.0001	0.004	0.003	0.001
MAPE (%)	0.04	1.29	0.78	0.43
SI	0.0004	0.015	0.009	0.004

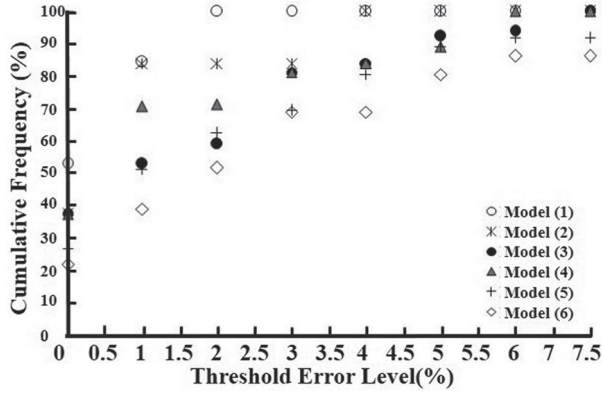


Figure 6. Error distribution of the hybrid PSO-GA algorithm across all models.

which is used for the first time for forecasting the discharge capacity of inflatable rubber dams, has the most significant impact on improving the ANN results. Finally, by comparing the error metrics presented in Tables 3 and 6, it can be observed that the hybrid PSO-GA model has a higher capability for accurately forecasting the discharge capacity of inflatable rubber dams.

Moreover, Figure 6 gives the overall distribution of errors among all models presented in this study. It shows that Model 1 has the best performance in forecasting the discharge coefficient. According to Figure 5, 90% of the predicted discharge coefficients in Model 1 have a relative error of less than 3%, while 80% of the data have a relative error lower than 3%. It is worth mentioning that Model 2 also shows a sufficiently good performance, and Model 6 has a relative error of less than 4%.

Using a Taylor diagram, the correlation and standard deviation values between estimated and observed values are investigated. In Figure 7, a Taylor diagram is presented for all methods implemented. The RMSE parameter in the diagram is defined as the distance from the reference point (green dot) to any other point. The most accurate model is considered as the minimum space between the green and corresponding dots (Taylor, 2001). According to Figure 7, the red point (PSO-GA) is the closest point to the green point and has a minimum distance from the reference point; therefore it yields accurate results with minimum error.

5. Conclusion

Inflatable dams are popular solutions for many engineering and environmental problems, such as irrigation, power generation, flood control and environmental improvement. In this research, the discharge coefficient of an inflatable dam (C_d) is estimated by a new hybrid PSO-GA algorithm. Five dimensionless input parameters are utilized in this research: the ratio of total upstream head to dam height (H_1/D_h), the ratio of overflowing head to dam height (h/D_h), the ratio of discharge per unit width to its maximum value (q/q_{max}), the ratio of the internal pressure of the tube to its maximum value (p/p_{max}) and the ratio of the longitudinal coordinate placement of each element to x_{max}. The proposed hybrid algorithm is a combination of PSO and GA so as to render the new algorithm more efficient (faster convergence) and effective (higher accuracy). The developed hybrid PSO-GA is characterized by its quick convergence, yet it slows down sharply near the optimal point during the

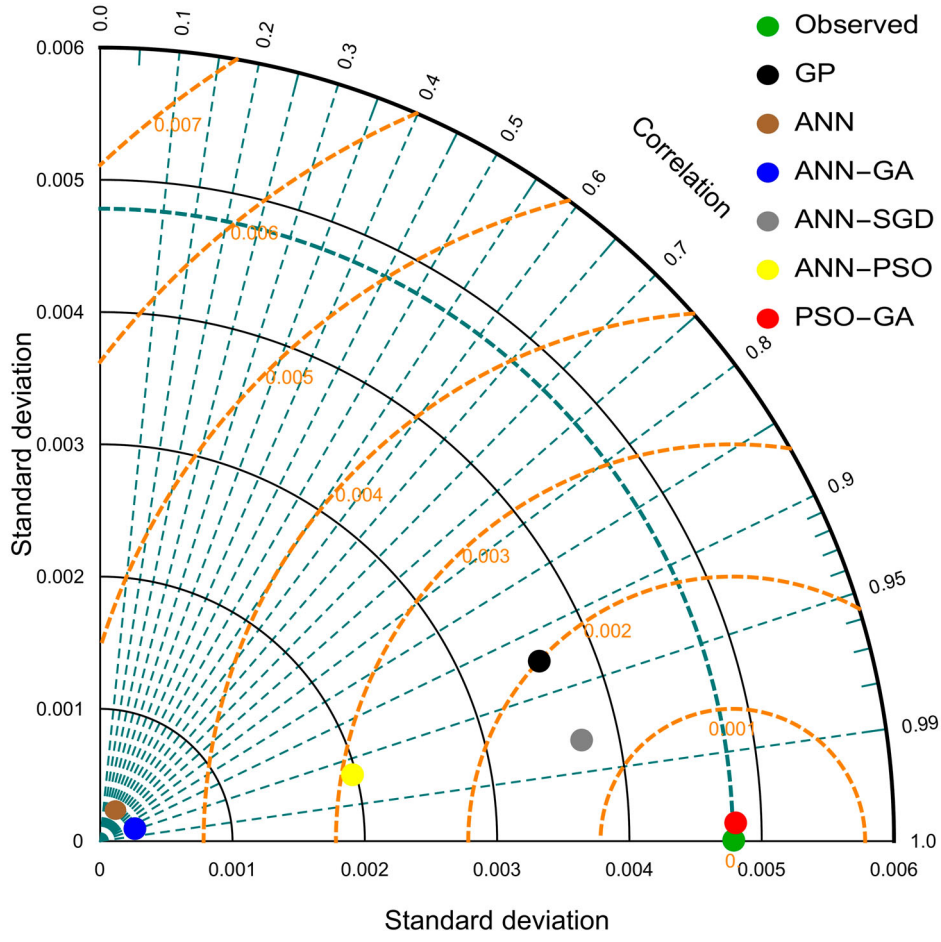


Figure 7. Taylor diagram of observed and estimated values.

search process. The genetic algorithm is very sensitive to the initial conditions. The random nature of genetic operators makes the algorithm sensitive to the initial population. This may lead to no convergence if the initial population is not well selected. Furthermore, experiments with the hybrid PSO-GA algorithm show that the developed algorithm responds faster, is not dependent on the initial population and finds more accurate results. This paper experiments with six different models and conducts a sensitivity analysis by observing the impact of excluding one input parameter at a time on the final prediction. This comprehensive investigation demonstrates that the model that utilizes all dimensionless parameters (Model 1) for estimating the discharge coefficient produces better results in comparison with other models, and also has the highest R^2 values, reaching 0.984. Model 1 provides the least MAPE value, which is almost 0.57% in comparison with the rest. The results indicate that internal pressure is introduced as the most effective parameter for determining the discharge coefficient, and neglecting this parameter decreases the estimation accuracy and relative error of estimation up to 2.12% in this model. Finally, the performance of some existing hybrid algorithms, including ANN-GA, ANN-PSO and ANN-SGD, are compared against PSO-GA. According to the results, the SGD algorithm has the most significant impact on improving the ANN results. As a final conclusion, comparing the error metrics indicates that the proposed PSO-GA model has a high capability for accurately forecasting the discharge capacity of inflatable rubber dams.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Henan Natural Science Foundation Project of China (No.182300410291), Nanyang science and technology project of China (No. KJGG219, No. KJGG004). The open access funding is by the publication fund of the TU Dresden.

ORCID

Hojat Karami http://orcid.org/0000-0002-2017-7204
Saeed Samadianfard http://orcid.org/0000-0002-6876-7182
Kwok-Wing Chau http://orcid.org/0000-0001-6457-161X
Amir H. Mosavi http://orcid.org/0000-0003-4842-0613

References

Abdin, A. E., & Abdeen, M. A. M. (2007). Ann model for predicting the impact of submerged aquatic weeds existence on the hydraulic performance of branched open channel system accompanied by water structures. *Journal of Mechanical Science and Technology*, *21*(7), 1139–1149. https://doi.org/10.1007/BF03027664

Akoz, M. S., Gumus, V., & Kirkgoz, M. S. (2014). Numerical simulation of flow over a semicylinder weir. *Journal of Irrigation and Drainage Engineering*, *140*(6), 4014016. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000717

Alhamati, A. A. N., Mohammed, T. A., Ghazali, A. H., Norzaie, J., & Al-Jumaily, K. K. (2005). Determination of coefficient of discharge for air-inflated dam using physical model. *Suranaree Journal of Science and Technology*, *12*(1), 19–27.

Bilhan, O., Emiroglu, M. E., & Kisi, O. (2010). Application of two different neural network techniques to lateral outflow over rectangular side weirs located on a straight channel. *Advances in Engineering Software*, *41*(6), 831–837. https://doi.org/10.1016/j.advengsoft.2010.03.001

Chang, J., Bai, T., Huang, Q., & Yang, D. (2013). Optimization of water resources utilization by PSO-GA. *Water Resources Management*, *27*(10), 3525–3540. https://doi.org/10.1007/s11269-013-0362-8

Chanson, H., & Tam, P. W. M. (1998). Discussion and closure: Use of rubber dams for flood mitigation in Hong Kong. *Journal of Irrigation and Drainage Engineering*, *124*(3), 181–184. https://doi.org/10.1061/(ASCE)0733-9437(1998)124:3(181)

Cheraghi-Shirazi, N., Kabiri-Samani, A. R., & Boroomand, B. (2014). Numerical analysis of rubber dams using fluid–structure interactions. *Flow Measurement and Instrumentation*, *40*, 91–98. https://doi.org/10.1016/j.flowmeasinst.2014.08.006

Diop, L., Samadianfard, S., Bodian, A., Yaseen, Z. M., Ghorbani, M. A., & Salimi, H. (2020). Annual rainfall forecasting using hybrid artificial intelligence model: Integration of multilayer perceptron with whale optimization algorithm. *Water Resources Management*, *34*(2), 733–746. https://doi.org/10.1007/s11269-019-02473-8

Doty, C. W., Moore, R. M., & Foutz, T. L. (1986). Performance of an inflatable dam during extreme events. *Applied Engineering in Agriculture*, *2*(2), 108–113. https://doi.org/10.13031/2013.26722

Ebtehaj, I., & Bonakdari, H. (2014). Comparison of genetic algorithm and imperialist competitive algorithms in predicting bed load transport in clean pipe. *Water Science and Technology*, *70*(10), 1695–1701. https://doi.org/10.2166/wst.2014.434

Ebtehaj, I., & Bonakdari, H. (2016). Assessment of evolutionary algorithms in predicting non-deposition sediment transport. *Urban Water Journal*, *13*(5), 499–510. https://doi.org/10.1080/1573062X.2014.994003

Ebtehaj, I., Bonakdari, H., & Shamshirband, S. (2016). Extreme learning machine assessment for estimating sediment transport in open channels. *Engineering with Computers*, *32*(4), 691–704. https://doi.org/10.1007/s00366-016-0446-1

Ebtehaj, I., Bonakdari, H., Zaji, A. H., Azimi, H., & Khoshbin, F. (2015). GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs. *Engineering Science and Technology, an International Journal*, *18*(4), 746–757. https://doi.org/10.1016/j.jestch.2015.04.012

El-Jumaily, K. K., & Salih, A. A. (2005). Analysis of inflatable dams under hydrostatic conditions. *Journal of Engineering and Sustainable Development*, *9*(3).

Esmaeilzadeh, B., Sattari, M. T., & Samadianfard, S. (2017). Performance evaluation of ANNs and an M5 model tree in Sattarkhan Reservoir inflow prediction. *ISH Journal of Hydraulic Engineering*, *23*(3), 283–292. https://doi.org/10.1080/09715010.2017.1308277

Ferreira, C. (2001a, September 10–24). Gene expression programming in problem solving. *6th online World Conf. on Soft Computing in Industrial Applications (invited tutorial)*. Springer, Berlin (Germany).

Ferreira, C. (2001b). Gene expression programming: A new adaptive algorithm for solving problems. *ArXiv Preprint Cs/0102027*.

Gharabaghi, B., Bonakdari, H., & Ebtehaj, I. (2018, July 10–12). Hybrid evolutionary algorithm based on PSOGA for ANFIS designing in prediction of no-deposition bed load sediment transport in sewer pipe. *Science and Information Conference* (pp. 106–118). Springer.

Gholami, A., Bonakdari, H., Ebtehaj, I., Mohammadian, M., Gharabaghi, B., & Khodashenas, S. R. (2018). Uncertainty analysis of intelligent model of hybrid genetic algorithm and particle swarm optimization with ANFIS to predict threshold bank profile shape based on digital laser approach sensing. *Measurement*, *121*, 294–303. https://doi.org/10.1016/j.measurement.2018.02.070

Goel, A. (2008, October 22–24). Estimation of scour downstream of spillways using SVM modeling. *Proceedings of the World Congress on Engineering and Computer Science WCECS* (pp. 22–24), San Francisco, CA, USA.

Goldberg, D. E., & Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. In Gregory J. E. Rawlins (Ed.), *Foundations of genetic algorithms* (Vol. 1, pp. 69–93). Elsevier.

Jung, B. S., & Karney, B. W. (2006). Hydraulic optimization of transient protection devices using GA and PSO approaches. *Journal of Water Resources Planning and Management*, *132*(1), 44–52. https://doi.org/10.1061/(ASCE)0733-9496(2006)132:1(44)

Kargar, K., Safari, M. J. S., Mohammadi, M., & Samadianfard, S. (2019). Sediment transport modeling in open channels using neuro-fuzzy and gene expression programming techniques. *Water Science and Technology*, *79*(12), 2318–2327. https://doi.org/10.2166/wst.2019.229

Karimi, S., Bonakdari, H., Ebtehaj, I., & Zaji, A. H. (2015a). *Prediction of Mean Velocity in Open Channel Intake and Rivers Using Gene Expression Programming*.

Karimi, S., Bonakdari, H., & Gholami, A. (2015b). Determination discharge capacity of triangular labyrinth side weir using multi-layer neural network (ANN-MLP). *Current World Environment*, *Special Volume 10*, 111–119. http://doi.org/10.12944/CWE.10.Special-Issue1.16

Kennedy, J. (1998, March 25–27). The behavior of particles. *International Conference on Evolutionary Programming* (pp. 579–589). Springer.

Khan, M., Azamathulla, H. M., Tufail, M., & Ab Ghani, A. (2012). Bridge pier scour prediction by gene expression programming. *Proceedings of the Institution of Civil Engineers*, *165*(9), 481–493.

Kim, M. (2003). *Two-dimensional analysis of four types of water-filled geomembrane tubes as temporary flood-fighting devices*. Virginia Polytechnic Institute and State University.

Kisi, O., Emiroglu, M. E., Bilhan, O., & Guven, A. (2012). Prediction of lateral outflow over triangular labyrinth side weirs under subcritical conditions using soft computing approaches. *Expert Systems with Applications*, *39*(3), 3454–3460. https://doi.org/10.1016/j.eswa.2011.09.035

Koza, J. R., & Koza, J. R. (1992). *Genetic programming: On the programming of computers by means of natural selection* (Vol. 1). MIT press.

Parsaie, A. (2016). Predictive modeling the side weir discharge coefficient using neural network. *Modeling Earth Systems and Environment*, *2*(2), 63. https://doi.org/10.1007/s40808-016-0123-9

Samadianfard, S., Nazemi, A. H., & Sadraddini, A. A. (2014). M5 model tree and gene expression programming based modeling of sandy soil water movement under surface drip irrigation. *Agriculture Science Developments*, *3*(5), 178–190.

Sharafi, H., Ebtehaj, I., Bonakdari, H., & Zaji, A. H. (2016). Design of a support vector machine with different kernel functions to predict scour depth around bridge piers. *Natural Hazards*, *84*(3), 2145–2162. https://doi.org/10.1007/s11069-016-2540-5

Smith, M. (1993). *Neural networks for statistical modeling*. Thomson Learning.

Steinfeld, B., Scott, J., Vilander, G., Marx, L., Quirk, M., Lindberg, J., & Koerner, K. (2015). The role of lean process improvement in implementation of evidence-based practices in behavioral health care. *The Journal of Behavioral Health Services & Research*, *42*(4), 504–518. https://doi.org/10.1007/s11414-013-9386-3

Taddy, M. (2019). *Business data science: Combining machine learning and economics to optimize, automate, and accelerate business decisions*. McGraw Hill Professional.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719. https://doi.org/10.1029/2000JD900719

Tsuruoka, Y., Tsujii, J., & Ananiadou, S. (2009, August 2–7). Stochastic gradient descent training for l1-regularized log-linear models with cumulative penalty. *Proceedings of the Joint Conference of the 47th Annual meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP* (pp. 477–485).

ul Islam, S., & Kumar, A. (2015). Evaluation of rubber dams for SHP in India. *Indian Journal of Science and Technology*, *8*(28), 1–8. https://doi.org/10.17485/ijst/2015/v8i28/84090

ul Islam, S., & Kumar, A. (2016). Inflatable dams for shp projects. *Renewable and Sustainable Energy Reviews*, *57*, 945–952. https://doi.org/10.1016/j.rser.2015.12.115

Waldow, G., & Bystrom, B. (2002, May 20–22). Inflatable crest gates at Mississippi River lock and Dam 1. *Cold Regions engineering: Cold Regions Impacts on Transportation and Infrastructure* (pp. 828–837).

Zaji, A. H., & Bonakdari, H. (2014). Performance evaluation of two different neural network and particle swarm optimization methods for prediction of discharge capacity of modified triangular side weirs. *Flow Measurement and Instrumentation*, *40*, 149–156. https://doi.org/10.1016/j.flowmeasinst.2014.10.002

Zaji, A. H., Bonakdari, H., & Karimi, S. (2015). Radial basis neural network and particle swarm optimization-based equations for predicting the discharge capacity of triangular labyrinth weirs. *Flow Measurement and Instrumentation*, *45*, 341–347. https://doi.org/10.1016/j.flowmeasinst.2015.08.002