Journal of AI and Data Mining
Vol 5, No 1, 2017, 55-66

DOI: 10.22044/jadm.2016.757

Estimating scour below inverted siphon structures using stochastic and soft computing approaches

M. Fatahi and B. Lashkar-Ara*

Civil Engineering Department, Jundi-Shapur University of Technology, Dezful, Iran.

Received 27 January 2016; Accepted 04 October 2016
Corresponding author: Lashkarara@jsu.ac.ir (B. Lashkar-Ara).

Abstract
In this work, we use the non-linear regression, artificial neural network (ANN), and genetic programming (GP) approaches in order to predict an important tangible issue, i.e. the scour dimension downstream of inverted siphon structures. Dimensional analysis and non-linear regression-based equations are proposed for the estimation of the maximum scour depth, location of the scour hole, and location and height of the dune downstream of the structures. In addition, The GP-based formulation results are compared with the experimental results and other accurate equations. The analysis results show that the equations derived from the forward stepwise non-linear regression method have the correlation coefficients R^2= 0.962, 0.971, and 0.991, respectively. This correlates the relative parameter of the maximum scour depth (*s*/*z*) in comparison with the GP and ANN models. Furthermore, the slope of the fitted line extracted from computations and observations for dimensionless parameters generally presents a new achievement for sediment engineering and scientific community, indicating the superiority of the ANN model.

Keywords: *Scour, Inverted Siphon, Neural Network, Genetic Programming.*

1. Introduction

Scour is a worldwide natural phenomenon caused by the flowing stream on the sediment beds. The local scour downstream of a hydraulic structure poses an immense problem in designing the foundation and stability of the hydraulic structure of Sarkar and Dey [1]. If the scour depth becomes significant, the stability of the foundation of the structure may be endangered, with a consequent risk of damage and failure. Therefore, the prediction and control of scour is necessary. During the formative phase of the scour profile, the local sediment transport is rather active, while by approaching the equilibrium condition, the phenomenon tends to a ‘‘purely hydraulic” Ghetti and Zanovello [2] mechanism, in which the hole profile is the result of a mass balance between the removed and deposited particles inside the pool. The difference in height between the upstream and downstream bed levels of the river-intersecting structures will form a vertical waterfall in the tail-water that plays an important role in grade-control structures. An example of these structures is the Balaroud inverted siphon structure in Dez irrigation and drainage network in the south of Andimeshk county, Khozestan province, Iran (Figure 1-a). The Balaroud inverted siphon is one of the largest national water structures that is situated in the irrigation and drainage network of Dez. Having a length of 990 m and a capacity of 156 m^3/s, this structure transfers water of the Dez river through the beneath of the Balarod river. The Balaroud inverted siphon has an intensive general and local scour, as a consequence of the anomalous usage of the materials in the Balaroud River. Figure 1-b shows the conditions of the structure. A sketch of the scour downstream of Balaroud inverted siphon is shown in figure 2, having a weir width of *b* and a fall height of *z*.

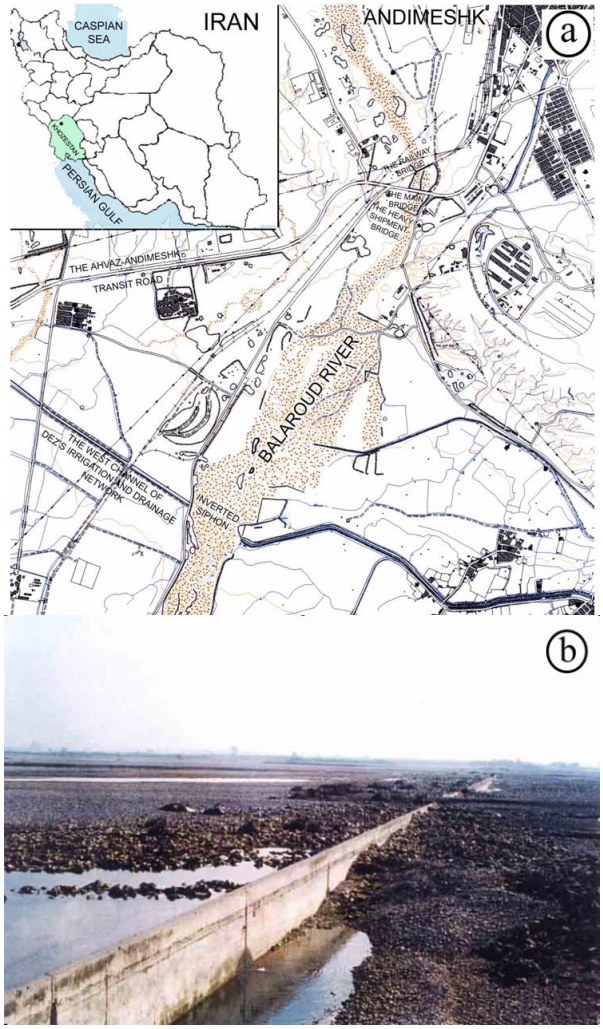

Figure 1. a) Location of studied area. b) Formation of grade-control structure in Balaroud river bed by protrusion of Balaroud inverted siphon structures located in Dez west irrigation system in SW of Iran.

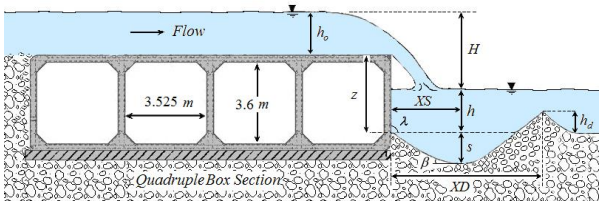

Figure 2. A sketch of scour of an alluvial bed downstream of an inverted siphon structure.

It is worth mentioning that the erosive action of the flowing water causes a significant downstream local scour, and may cause problematic issues for these structures. Thus the structural design of the inverted siphon structures must include sufficient protective provisions against local scours.

An appropriate structural design of these structures requires full comprehension, and, somehow, prediction of the nature of the downstream scour, namely the location and extent of the scour.

Many experimental studies on scour downstream of hydraulic structures are available in the literature. Among these, we can refer to the studies carried out by Rouse [3], Doddiah et al. [4], Mason and Arumugam [5], D'Agostino [6], Robinson et al. [7], Bormann and Julien [8], and Bennett et al. [9].

Rouse [3] and Doddiah et al. [4] have shown that the scour depth s increases with time T, and its changes accord to the following relationship:

$$s/h = k_1 + k_2 \log\left(QT/bz^2\right) \quad (1)$$

where, k_1 and k_2 are constants; and h is the tail-water depth above the non-scoured bed level (Figure 2).

Bennett et Al. [9] have observed that a jet is separated from the over-fall, and diffuses downstream of the structure.

This jet is split into two wall-jets, forming two counter-rotating eddies (rollers), downstream and upstream of the diffuse jet, and eroding and forming the scour hole. The upstream roller is captive between the over-fall and the impinging jet. The circulation within the downstream roller causes a region of significant upwelling, and sediment deposition occurs as the flow directs toward the water surface. The D'Agostino [6] studies have neglected the influence of the bed grain-size, and he suggested equation (2) due to estimating XD; it is a distance between the downstream of the structure and the accumulated depositions crest.

$$XD/z = 3.55 \times \sqrt[3]{q^2/g}\,/z + 0.34 \quad (2)$$

where, $q = Q/b$ is the discharged per unit weir width.

A large-scale model research work carried out by Bormann and Julien [8] has enabled the calibration of an equilibrium equation based on particle stability and its validation in a variety of conditions such as wall and vertical jets, free over-fall jets, submerged jets, and flow-over large-scale grade-control structures. According to the results obtained by Bormann and Julien [8], the relationship for estimating s is in the following form:

$$s = \left[\frac{0.611}{\left[\sin(0.436+\beta')\right]^{0.8}} q^{0.6} \frac{U_0}{g^{0.8} d_{90}^{0.4}} \sin\beta'\right] - z \quad (3)$$

where, g is the acceleration due to gravity; z the difference in height between the crest of the grade-control structure and the bottom of the downstream undisturbed bed level; U_o is the mean flow velocity at the weir crest (equal to the jet entering velocity); and β' is the maximum side-

angle of scour hole. The angle β' (Figure 2) is approximately equal to the jet angle, and has been experimentally inferred by Bormann and Julien [8]:

$$\beta' = 0.316\sin\lambda + 0.15\ln\left(\frac{z+y_0}{y_0}\right) + 0.13\ln\left(\frac{h}{y_0}\right) - 0.05\ln\left(\frac{U_0}{\sqrt{gy_0}}\right) \quad (4)$$

in which, λ is the downstream face angle of the grade-control structure (rad); and y_o is the water depth at the crest.

Mason and Arumugam [5] have tested some formulas for the scours under free-falling jets using model and prototype data. The authors obtained the best agreement between the selected equations and measurements for the model data using a representative diameter d_s equal to the mean particle size d_m. They have proposed a comprehensive model and prototype equation, which can be rewritten according to the suggestions made by Yen [10] in the following form:

$$\frac{s}{\sqrt[3]{q^2/g}} = \left(6.42 - 3.10H^{0.1}\right)g^{-H/600} \times \left(\frac{gH^3}{q^2}\right)^{(20+H)/600}\left(\frac{H}{d_s}\right)^{0.1}\left(\frac{h}{H}\right)^{3/20} \quad (5)$$

where, H is the difference in height between upstream and tail-water level.

Azmathulla et al. [11] have predicted the relative scour depth downstream of ski-jump bucket spillway by genetic programming (GP) and artificial neural network (ANN). The GP-based estimations were found to be equally and more accurate than the ANN-based ones, especially when the underlying cause-effect relationship became more uncertain to model.

Lee et al. [12] have predicted the scour depth around bridge piers by the back-propagation neural network (BPNN) and non-linear relationships. They have shown that the scour depth around bridge piers can be efficiently predicted using BPNN.

The main objectives of this study were to investigate the scour process, estimating the maximum depth and location of the scour hole, and evaluating the maximum height and location of the sedimentary mound at the downstream of the grade-control structure. In this study, the experimental data obtained by the previous researchers was used, and the equations were reviewed and re-written using the D'Agostino and Ferro [13] studies in order to improve the accuracy of the existing relationships. In the next step, the hydroinformatic science and the soft computing technique were used to achieve more accuracy for the relationships of the hole's characteristic and the sedimentary mound in alluvial ducts containing non-cohesive sediments.

Zhang et al. [14] have provided a collection of high-quality research articles that address the broad challenges in bioinformatics and biomedicine of SCs and reflect the emerging trends in the state-of- the-art SC algorithms.

Asghari Esfandani and Nematzadeh [15] carried out research using the Genetic algorithm and Neural Network to make a hybrid method to predict air pollution in Tehran. The result show that the proposed method has a good agreement with field observations.

2. Research method

After evaluating the studies carried out by Veronese [16], Mossa [17], D'Agostino [18], Falciai and Giacomin [19], Lenzi et al. [20], and D'Agostino and Ferro [13], the study carried out by D'Agostino [18] was chosen for our study because it had a favorable situation for the experimental data analysis and processing. The data range used in the form of the revised effective relationships of the scour hole and sedimentary mound under grade-control structures with sharp-crested weir are shown in table 1.

Table 1. Changes in range of parameters used in this study

Parameter	Symbol	Unit	Range
Channel width	B	m	0.5
Weir width	B	m	0.15-0.3
Fall height	Z	m	0.41- 0.71
Total head above the weir crest	h_0	m	0.043-0.2006
Tail water depth	H	m	0.083-0.435
Water discharge	Q	L/s	8.35-83.35
Diameter of which 50-percent is finer	D_{50}	mm	4.1, 11.5
Diameter of which 90-percent is finer	D_{90}	mm	7, 17.6
Maximum scour depth	s	m	0.045-0.285
Location of the maximum scour depth to weir	XS	m	0.215-0. 705
Maximum height of the mound above the undisturbed bed level	h_d	m	0.0250.255
Location of the maximum height of stockpiling sediments	XD	m	0.24-1.705

According to the theory of dimensional analysis and also the characteristics shown in figure 2, all the affecting parameters in this research work are as what follow:

- Kinematic characteristics: Q is the discharge, and g is the acceleration due to gravity.
- Dynamic characteristics: ρ_s-ρ is the submerged weight of sediment particles; ρ is the mass density of fluid; and μ is the water viscosity.
- Geometric characteristics: z is the fall height; b is the weir width; B is the channel width; h is the tail water depth; H is the difference in height from the water level upstream of the weir to the tail-water level; D_{50} is the diameter for which 50% of particles are finer; and D_{90} is the diameter for which 90% of particles are finer.

Therefore, the following functional relationship can be expressed:

$$f(\varphi, z, b, B, h, H, Q, \rho_s - \rho, \rho, g, \mu, D_{50}, D_{90}) = 0 \quad (6)$$

in which φ is representative scour hole parameters containing as: s is the maximum scour depth; XS is the horizontal distance between the weir crest and the section of maximum scour depth; h_d is the maximum height of the mound above the undisturbed bed level and XD is the location of the maximum height of stockpiling sediments.

The dimensionless parameters were obtained according to (7) using the Buckingham π theorem, and the independent variables z and ρ, as the repeated variables, were selected.

$$f(\pi_1 = \frac{\varphi}{z}, \pi_2 = \frac{b}{z}, \pi_3 = \frac{B}{z}, \pi_4 = \frac{h}{2}, \pi_5 = \frac{H}{z}, \quad (7)$$
$$\pi_6 = \frac{D_{50}}{z}, \pi_7 = \frac{D_{90}}{z}, \pi_8 = \frac{\rho - \rho_S}{\rho}, \pi_9 = \frac{gz^5}{Qz^2}) = 0$$

The dimensionless parameters b/B, h/H, and D_{50}/D_{90} are the results of the combination of the dimensionless parameters π_2, π_3, π_4, π_5, π_6, and π_7. The dimensionless parameter $Q/\left(bz\sqrt{gD_{50}(\rho_s - \rho)/\rho}\right)$ is obtained through combination with the parameters π_2, π_6 , π_8 , and π_9.

$$f\left(\frac{\varphi}{z}, \frac{b}{z}, \frac{b}{B}, \frac{h}{H}, \frac{D_{90}}{D_{50}}, \frac{Q}{bz\sqrt{gD_{50}(\rho_s - \rho)/\rho}}, \frac{z\mu}{\rho Q}\right) = 0 \quad (8)$$

Parameter $z\mu/\rho Q$ is the Reynolds number that is neglected from its effect in the equations because of flow turbulence. As a result, (8) is summarized in the form of (9).

Parameter $Q/\left(bz\sqrt{gD_{50}(\rho_S - \rho)/\rho}\right)$ is emanated from the densimetric Froude number, and it is shown by $Fr_{D_{50}}$. Thus we can write:

$$\frac{\varphi}{z} = f\left(\frac{b}{B}, \frac{D_{90}}{D_{50}}, \frac{h}{H}, \frac{b}{z}, Fr_{D_{50}}\right) = 0 \quad (9)$$

In addition, the effect of the independent parameters b/z, h/H, Fr_{D50}, D_{90}/D_{50}, and b/B on the dependent parameters φ/z is introduced in the form of (10):

$$\frac{\varphi}{z} = a \times \left(\frac{b}{B}\right)^b \left(\frac{D_{90}}{D_{50}}\right)^c \left(\frac{h}{H}\right)^d \left(\frac{b}{z}\right)^e \left(Fr_{D_{50}}\right)^f \quad (10)$$

In this equation, the coefficients a, b, c, d, e, and f are constant numbers, and their effects are determined using the statistical analysis of the experimental observations made by the SPSS software using the non-linear regression technique by the forward stepwise regression method.

3. Results and discussion

3.1. Non-linear regression method

The SPSS software was used for determining the effective equations in this research work. The observed values for the independent dimensionless relative parameters b/z, h/H, Fr_{D50}, D_{90}/D_{50}, and b/B were evaluated versus the dependent parameters maximum scour relative depth s/z, maximum relative distance of maximum scour depth XS/z, relative height of sedimentary mound h_d/z, and maximum relative distance accumulation of sediments to weir toe XD/z in order to determine the mapping space between the independent and dependent parameters mentioned in (10). The mapping space between the independent and dependent parameters can be shown as (11)-(14):

$$\frac{s}{z} = 0.5292 \times \left(\frac{b}{B}\right)^{0.3104} \times \left(\frac{D_{90}}{D_{50}}\right)^{-0.0651} \times \left(\frac{h}{H}\right)^{0.0849} \times \left(\frac{b}{z}\right)^{0.5052} \times \left(Fr_{D_{50}}\right)^{0.5302} \quad (11)$$

$$\frac{XS}{z} = 1.8113 \times \left(\frac{b}{B}\right)^{0.0333} \times \left(\frac{D_{90}}{D_{50}}\right)^{-0.0839} \times \left(\frac{h}{H}\right)^{0.1161} \times \left(\frac{b}{z}\right)^{0.3583} \times \left(Fr_{D_{50}}\right)^{0.3601} \quad (12)$$

$$\frac{h_d}{z} = 1.369 \times \left(\frac{b}{B}\right)^{1.1387} \times \left(\frac{D_{90}}{D_{50}}\right)^{-1.5679} \times \left(\frac{h}{H}\right)^{-0.0573} \times \left(\frac{b}{z}\right)^{0.3413} \times \left(Fr_{D_{50}}\right)^{0.7772} \quad (13)$$

$$\frac{XD}{z} = 4.5856 \times \left(\frac{b}{B}\right)^{0.4987} \times \left(\frac{D_{90}}{D_{50}}\right)^{-0.6068} \times \left(\frac{h}{H}\right)^{0.05} \times \left(\frac{b}{z}\right)^{0.3136} \times \left(Fr_{D_{50}}\right)^{0.5035} \quad (14)$$

The fitting method of (11)-(14) extracted from the experimental results are shown in figures 3-6. The error analysis functions were used in order to evaluate the results obtained by the proposed equations.

A summary of the results is shown in table 2.

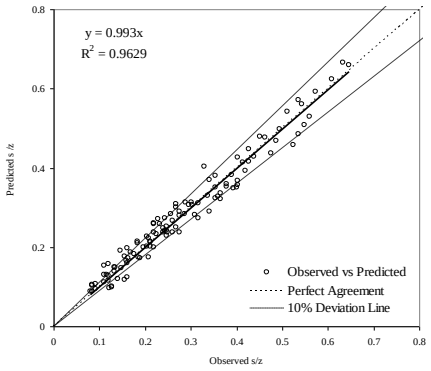

Figure 3. Comparison between observed and predicted equation (11) to estimate *s/z*.

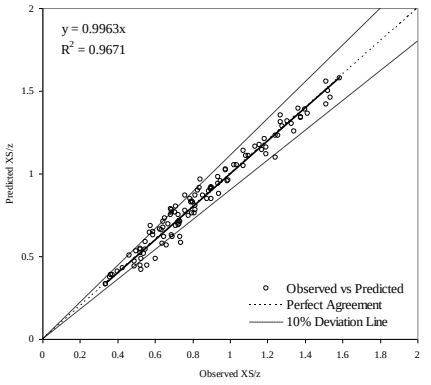

Figure 4. Comparison between observed and predicted equation (12) to estimate *XS/z*.

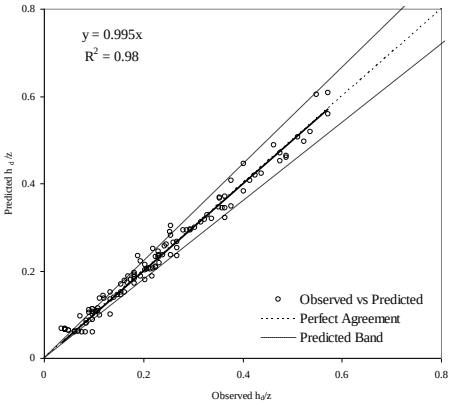

Figure 5. Comparison between observed and predicted equation (13) to estimate h_d/z.

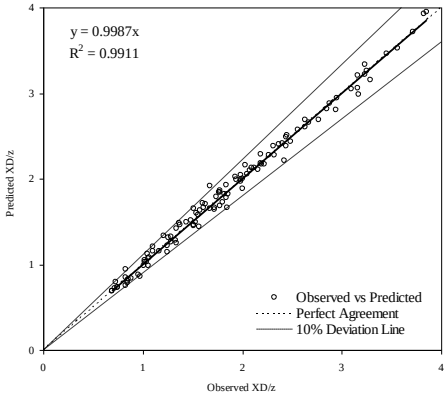

Figure 6. Comparison between observed and predicted equation (14) to estimate *XD/z*.

Table 2. Error functions from results suggested by equations (11)-(14) against experimental observations.

Parameter	RMSE	MPE	SEE	EF	m	R^2
s/z	0.0264	1.6378	0.026	0.9643	0.993	0.962
XS/z	0.0553	0.065	0.055	0.9674	0.9963	0.967
h_d/z	0.019	2.839	0.018	0.9804	0.995	0.98
XD/z	0.0743	0.028	0.074	0.9911	0.9987	0.991

The angular coefficient of the fitted line extracted from the results of (11)-(14) indicates that the non-linear regression estimates of the dimensionless parameter values *s/z*, *XS/z*, h_d/z, and *XD/z* are, respectively, 0.7%, 0.37%, 0.5%, and 0.13% lower than the observed values. The estimating bands of the above-mentioned four parameters used to determine the scour hole dimension are shown in figures 3-6. The skewness results obtained from the statistical prediction dimensionless parameters *s/z*, *XS/z*, h_d/z, and *XD/z* had desirable distributions.

3.2. Genetic programming

Genetic programming (GP) is used as one of the evolutionary algorithm techniques in order to flourish the presented relationship accuracy in the second part of this work. GP is an automatic programming technique used for evolving computer programs to solve problems. GP is frequently applied to the model structure identification problems in engineering applications. In such applications, GP is used to infer the underlying structure of either a natural or an experimental process in order to model the process numerically. GP is a member of the evolutionary algorithm (EA) family. EAs are based upon the Darwin's natural selection theory of evolution, where a population is progressively improved by selectively discarding the not-so-fit populations and breeding from the better populations. EAs work by defining a goal in the form of a quality criterion, and then using this goal to evaluate the solution candidates in a stepwise refinement of a dataset structures and return an optimal solution after a number of generations. GP can optimize both the structure of the model and its parameters. Since GP evolves an equation relating the output and input variables, it has the advantage of providing the inherent functional relationship explicitly over techniques like ANN [21].

After extracting the model by the GP method, the results obtained were analyzed using the error function, and compared with the experimental

observations. All programmings were done in MATLAB (version 8.2.0.29).

The best input should be a considered pattern in order to determine the best response. Therefore, in the first step, various parameters that are effective in modeling such as the population members, number of generations, size of tree structures, and method of generation of initial population should be determined carefully with regards to the investigated data user.

The size of each tree structure has a significant role in the accuracy of the final model. determining the larger numbers than optimal value leads in reduction of the accuracy of the model the models are not presented mainly because the models made by GP in order to estimate the scour hole dimensions in the downstream grade-control structures were very long-scale.

The root mean square error (RMSE) is used to represent the fitness function. The $RMSE_i$ of an individual program *i* is evaluated using the following equation:

$$RMSE_i = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(P_{ij} - T_j)^2} \quad (15)$$

where, $P_{(ij)}$ is the value predicted by the individual program *i* for the fitness case *j* (out of *n* fitness cases), and T_j is the target value for the fitness case *j*.

For a perfect fit, $P_{(ij)} = T_j$ and $RMSE_i = 0$. Thus the RMSE index ranges from 0 to infinity, with 0 corresponding to the ideal. As it stands, $RMSE_i$ cannot be used directly as fitness since for the fitness proportionate selection, the fitness value must increase with efficiency. Thus to evaluate the fitness f_i of an individual program *i*, the following equation is used:

$$f_i = 1000\frac{1}{1+RMSE_i} \quad (16)$$

which obviously ranges from 0 to 1000, with 1000 corresponding to the ideal. In order to evolve a model with GP, the function set and the characteristic of the employed GP must be introduced. In this case, after setting composed functions of the operators ithave been used to achieve the best model evolved by GP.The final characteristics of the employed GP for all the scour parameters *(ds/z)*, *(XS/z)*, *(h_d/z)*, and *(XD/z)* are shown in table 3.

GP against the experimental results for the parameters *s/z*, *XS/z*, *h_d/z*, and *XD/z* are shown in figures 7-10. The error analysis functions were used in order to evaluate the GP results, tabulating them in table 4.

Table 3. Characteristics of employed GP for *(ds/z)*, *(XS/z)*, *(h_d/z)*, and *(XD/z)*.

Parameter	Definition	Value *(ds/z), (XS/z)*	Value *(h_d/z), (XD/z)*
P_1	Function set	+, −, *, √, ^2, cos, exp	+, −, *, √, ^2
P_2	Terminal set	b/B, D_{90}/D_{50}, h/H, b/Z, Fr d	b/B, D_{90}/D_{50}, h/H, b/Z, Fr d
P_3	Number of inputs	5	5
P_4	Fitness function	RMSE	RMSE
P_5	Error type	error function	error function
P_6	Crossover rate	0.85%	0.85%
P_7	Mutation rate	0.1%	0.1%
P_8	Gene reproduction rate	0.05%	0.05%
P_9	Population size	250	350
P_{10}	Number of generation	120	150
P_{11}	Tournament type	regular	regular
P_{12}	Tournament size	6	6
P_{13}	Max tree depth	4	4
P_{14}	Max node per tree	Inf	Inf
P_{15}	Constants range	[-10, +10]	[-10, +10]

The angular coefficient of the fitted line extracted from the results of the model made indicated that GP estimated the values for the dimensionless parameters *s/z*, *XS/z*, *h_d/z*, and *XD/z* to be, respectively, 0.78%, 0.9%, 1.2%, and 0.65% lower than the observed values. Figures 7-10 show the estimating bands for the above-mentioned four parameters to determine the scour hole dimensions by GP. The skewness results obtained from the predicted dimensionless parameters *s/z*, *XS/z*, *h_d/z*, and *XD/z* using the GP data mining system was satisfactory.

Table 4.. Error function of GP model against experimental observations.

Parameter	RMSE	MPE	SEE	EF	m	R^2
s/z	0.024	–2.06	0.024	0.969	0.992	0.97
XS/z	0.051	–0.79	0.051	0.972	0.991	0.97
h_d/z	0.020	–4.35	0.020	0.976	0.987	0.97
XD/z	0.105	–0.36	0.105	0.982	0.993	0.98

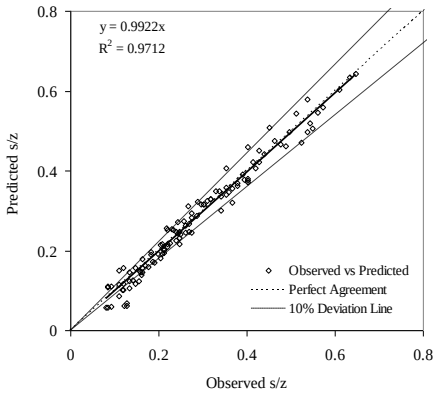

Figure 7. Comparison between observed and predicted GP to estimate *s/z*.

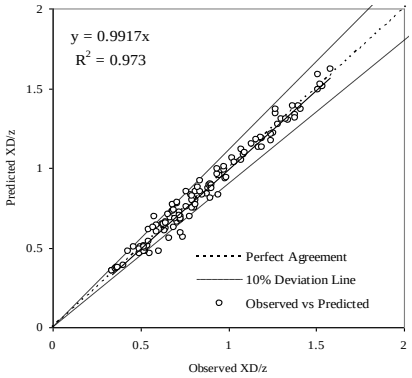

Figure 8. Comparison between observed and predicted GP to estimate *XS/z*.

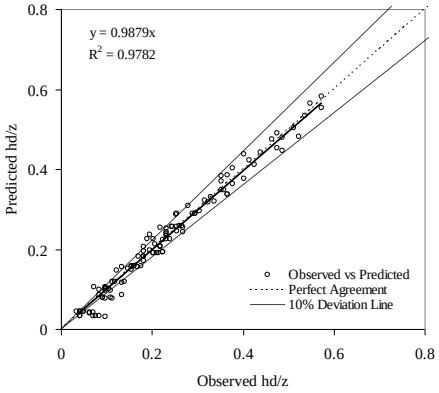

Figure 9. Comparison between observed and predicted GP to estimate h_d/z.

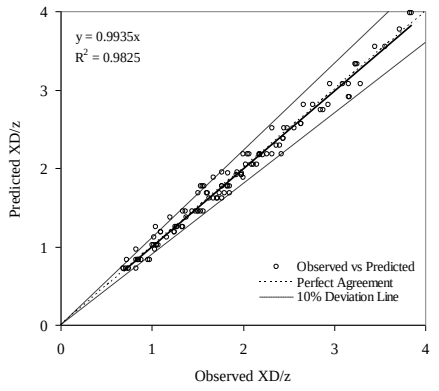

Figure. 10. Comparison between observed and predicted GP to estimate *XD/z*.

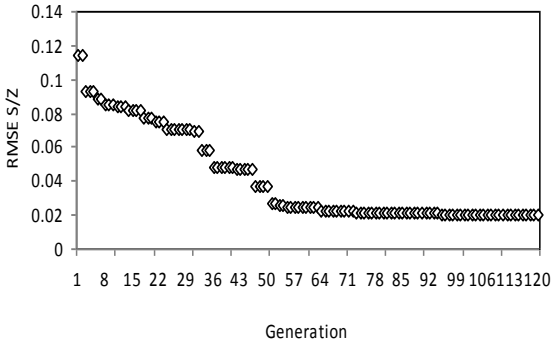

Figure 11. Root mean square error versus generation of *s/z*.

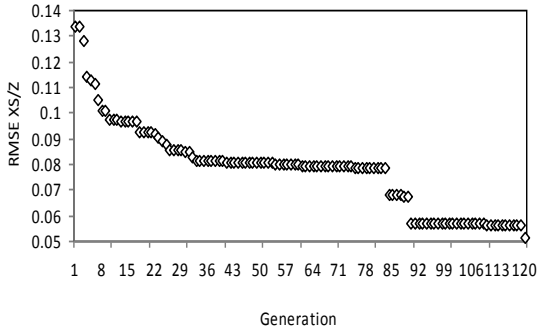

Figure 12. Root mean square error versus generation of *XS/z*.

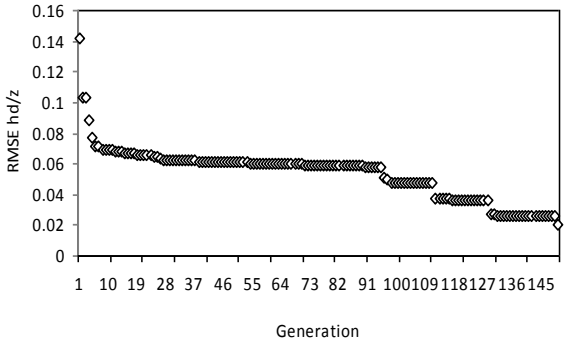

Figure 13. Root mean square error versus generation of h_d/z.

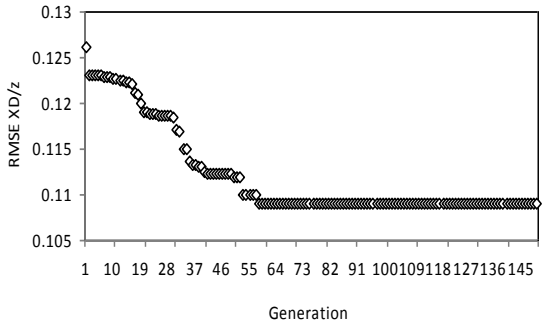

Figure 14. Root mean square error versus generation of *XD/z*.

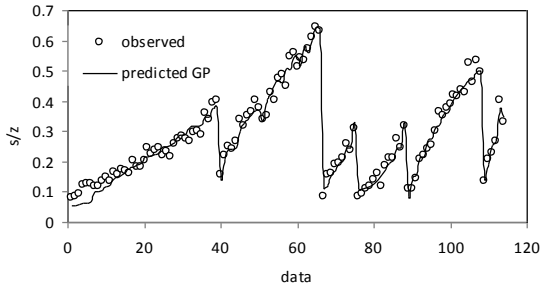

Figure 15. GP estimations of *s/z* versus measured ones

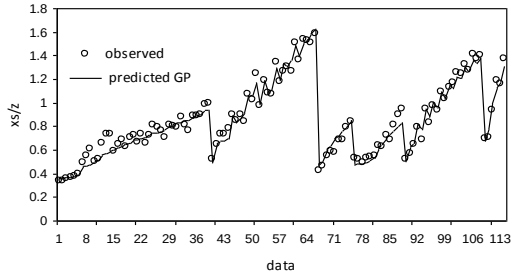

Figure 16. GP estimations of *XS/z* versus measured ones

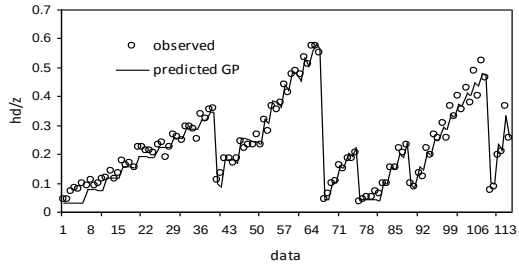

Figure 17. GP estimations of h_d/z versus measured ones

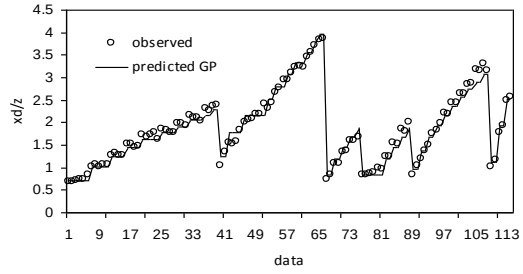

Figure 18. GP estimations of *XD/z* versus measured ones

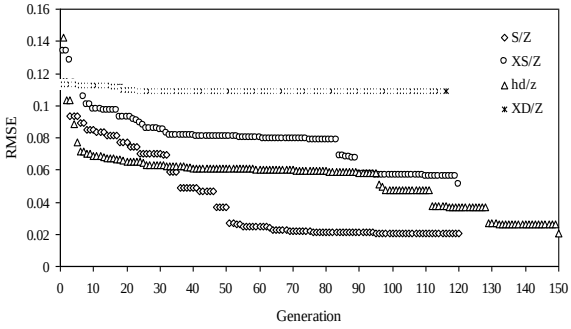

Figure 19. Root mean square error versus generation in GP.

3.3. Artificial neural network

The artificial neural network (ANN) is one of the most common network models, which generally presents a system of inter-connected neurons that can compute values from inputs. A neuron consists of multiple inputs and a single output. There is an input layer that acts as a distribution structure for the data being presented to the networks. This layer is not used for any type of processing. After this layer, one or more processing layers follow, called the hidden layers. The final processing layer is called the output layer in a network. This process is repeated until the error rate is minimized or reaches an acceptable level or until a specified number of iterations have been accomplished. In the ANN models, the sigmoid function is used. Here, in this work, we used the multi-layer perceptron (MLP) neural network model. An MLP is a feed-forward ANN model that maps sets of input data onto a set of appropriate outputs. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. The MLP-ANN models were used to estimate the dimensionless parameter values $\left(b/B, D_{90}/D_{50}, h/H, b/z, Fr_{D_{50}}\right)$. For this purpose, 80% of the experimental data was used for network training, and the remaining 20% was used for testing the results obtained. This procedure was repeated for 1,000 times to achieve the best performance. The parameters *s/z*, *XS/z*, h_d/z, and *XD/z* were introduced as the input parameters to the model. The details of the MLP-ANN architecture is shown in table5.

Figures 20-23 show the performance of the MLP-ANN model to estimate the scour hole dimensions after training. Then 20% of the data that was not used in the training stage would be used to evaluate the performance of the model. The error function results of the neural network used for estimating the scour hole dimensions are summarized in table 6.

Table 5. Details of MLP-ANN architecture

parameter	Input layer	Hidden layer	Hidden layer	Output layer
s/z	5	6	4	1
XS/z	5	4	2	1
h_d/z	5	4	3	1
XD/z	5	3	1	1

Table 6. Statistical error functions of ANN for estimating scour hole dimensions for best network architecture

Parameter	Training				Testing			
	MPE	**RMSE**	**EF**	**R^2**	**MPE**	**RMSE**	**EF**	**R^2**
s/z	0.2514	0.0082	0.9811	0.9853	0.0345	0.0133	0.9908	0.9911
XS/z	1.6163	0.0326	0.9492	0.9737	–1.9160	0.0494	0.9793	0.9837
h_d/z	2.8802	0.0081	0.9756	0.9899	–2.1214	0.0186	0.9864	0.9885
XD/z	0.4463	0.0258	0.9937	0.9933	1.8970	0.0737	0.9935	0.9933

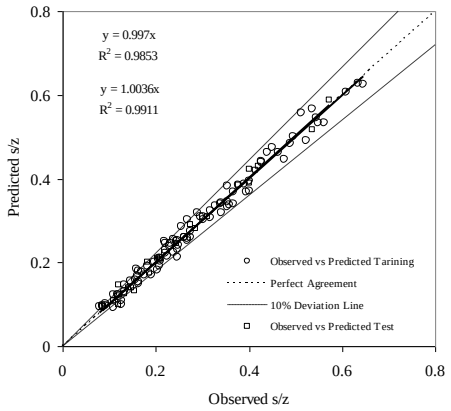

Figure 20. Comparison between observed and predicted *s/z* of ANN in training and testing stages.

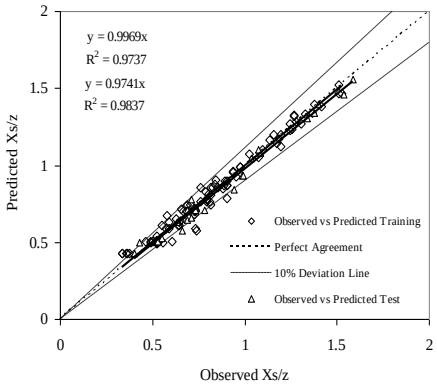

Figure 21. Comparison between observed and predicted *XS/z* of ANN in training and testing stages.

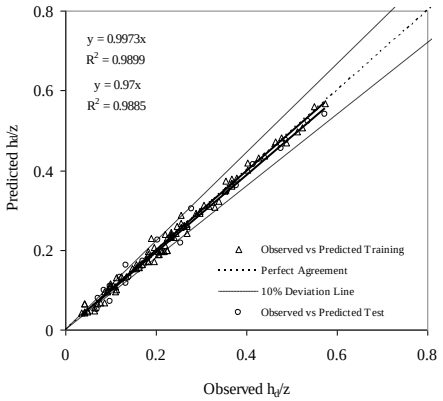

Figure 22. Comparison between observed and predicted *h_d/z* of ANN in training and testing stages

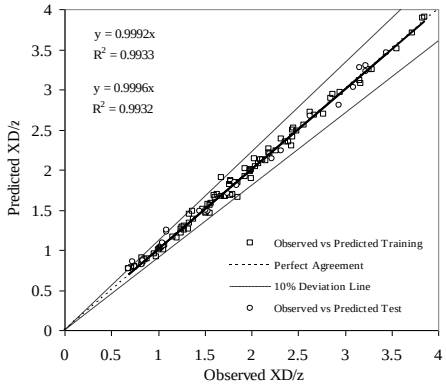

Figure 23. Comparison between observed and predicted *XD/z* of ANN in training and testing stages.

Comparing tables 2, 4, and 6 showed that the neural network model performed better in term of R^2, compared to the non-linear regression and GP methods. Also the estimated bands of the above-mentioned four parameters used to determine the scour hole dimension are shown in figures 20-23. The results of the statistical analysis conducted at various stages of training and testing is shown in figures 20-23. The skewness results obtained using the statistical prediction dimensionless parameters, i.e. *s/z*, *XS/z*, h_d/z, and *XD/z*, had desirable distributions.The angular coefficient of the fitted line extracted from the results of the model made indicated that ANN estimated the values for the dimensionless parameters, i.e. *s/z*, *XS/z*, h_d/z, and *XD/z*, to be 0.3%, 0.4%, 0.3%, and 0.08%, respectively, lower than the observed values in the training phase, the dimensionless parameter *s/z*, 0.2% more, and the dimensionless parameters *XS/z*, h_d/z, *XD/z*, 2.6%, 3%, and 0.04%, respectively, lower than the values observed in the testing phase. The skewness results obtained from the statistical prediction of the dimensionless parameters *s/z*, *XS/z*, h_d/z, and *XD/z* had desirable distributions.

3.4. Sensitivity analysis methods

Sensitivity analysis approaches can be classified in different ways. In this paper, they have been

classified as mathematical, statistical or graphical. Other classifications focus on the capability, rather than the methodology, of a specific technique [22]. The classification approaches aim to understand the applicability of a specific method to a particular model and analysis objective.

3.4.1. Graphical methods for sensitivity analysis

Graphical methods give a representation of the sensitivity in the form of graphs, charts or surfaces. Generally, graphical methods are used to give a visual indication of how an output is affected by the variation in inputs [23]. Graphical methods can be used as a screening method before further analysis of a model or to represent complex dependencies between inputs and outputs [24, 25]. Graphical methods can be used to complement the results of mathematical and statistical methods for a better representation [26]. The sensitivity analysis was performed considering negligible *h/H* parameter impact, and thus we removed it from the equation, and (11)-(14) can be formed as follows:

$$\frac{s}{z}=0.541\times\left(\frac{b}{B}\right)^{0.294}\times\left(\frac{D_{90}}{D_{50}}\right)^{-0.008}\times\left(\frac{b}{z}\right)^{0.608}\times\left(Fr_{D_{50}}\right)^{0.561} \quad (17)$$

$$\frac{XS}{z}=1.833\times\left(\frac{b}{B}\right)^{-0.002}\times\left(\frac{D_{90}}{D_{50}}\right)^{-0.734}\times\left(\frac{b}{z}\right)^{0.500}\times\left(Fr_{D_{50}}\right)^{0.395} \quad (18)$$

$$\frac{h_d}{z}=1.344\times\left(\frac{b}{B}\right)^{1.141}\times\left(\frac{D_{90}}{D_{50}}\right)^{-1.601}\times\left(\frac{b}{z}\right)^{0.276}\times\left(Fr_{D_{50}}\right)^{0.753} \quad (19)$$

$$\frac{XD}{z}=4.629\times\left(\frac{b}{B}\right)^{0.489}\times\left(\frac{D_{90}}{D_{50}}\right)^{-0.567}\times\left(\frac{b}{z}\right)^{0.373}\times\left(Fr_{D_{50}}\right)^{0.522} \quad (20)$$

The fitting method for equations (17)-(20), extracted from the experimental results, are shown in figures 3-6. The error analysis functions were used in order to evaluate the results of the proposed equations. A summary of the results obtained are shown in table 7. The angular coefficient of the fitted line extracting the results of (17)-(20) indicates that the non-linear regression estimates the dimensionless parameter value *s/z*, *XS/z*, h_d/z, and *XD/z* to be, respectively, 0.85%, 0.74%, 0.9%, and 0.21% lower than the observed values. The estimating bands of the above-mentioned four parameters used to determine the scour hole dimension are shown in figures 15-18. The fitting method for equations (17)-(20), extracted from the experimental results, are shown in figures 24-27. The error analysis functions were used in order to evaluate the results of the proposed equations. A summary of the results obtained are shown in table 7.

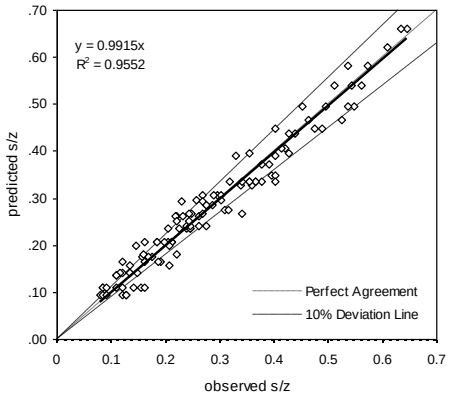

Figure 24. Comparison between observed and predicted equation (17) to estimate *s/z*.

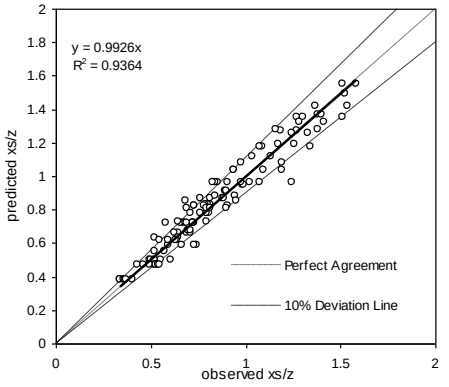

Figure 25. Comparison between observed and predicted equation (18) to estimate *XS/z*.

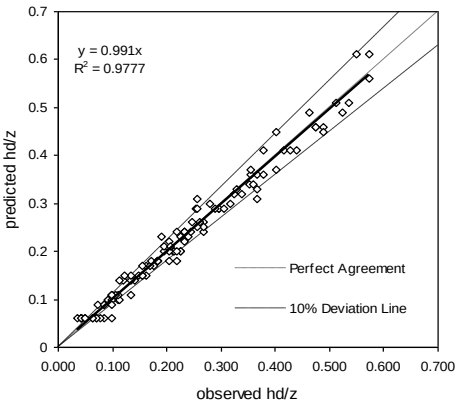

Figure 26. Comparison between observed and predicted equation (19) to estimate h_d/z.

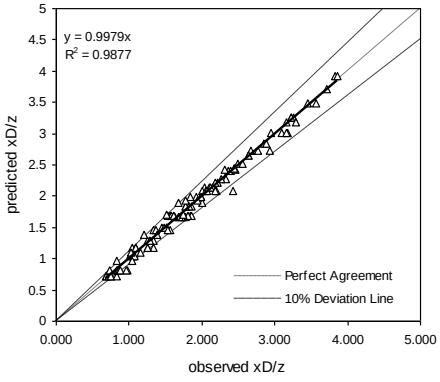

Figure 27. Comparison between observed and predicted equation (20) to estimate *XD/z*.

Table 5. Error functions from results suggested by equations (17)-(20) against experimental observations.

Parameter	RMSE	MPE	SEE	EF	m	R^2
s/z	0.028	1.843	0.028	0.957	0.991	0.955
XS/z	0.075	0.480	0.074	0.939	0.992	0.936
h_d/z	0.020	1.965	0.020	0.978	0.991	0.977
XD/z	0.087	0.066	0.087	0.987	0.997	0.987

4. Conclusion

By comparing the results tabulated in tables 2, 4, and6, it can be seen that the angular coefficient of the fitted line extracted from the results of the predicted parameters *s/z*, *XS/z*, h_d/z, and *XD/z* resulting from ANN is 45 degrees closer to the slope of the line of the non-linear regression and GP comparing to the predicted values. This indicates that the ANN model was more successful in estimating these parameters. The root mean square error had fewer values in predicting the parameter *s/z*, *XS/z*, h_d/z, and *XD/z* by ANN than non-linear regression and GP, and this indicates the advantage of this approach in estimation of these parameters. GP may serve as a robust approach, and it may open a new area for an accurate and effective explicit formulation of many water engineering problems. Generally, with regard to this point that since using the presented non-linear regression for estimating scour parameters does not require a computer, it can, therefore, be claimed that using the non-linear regression compared to GP and ANN in estimating the scour hole dimensions in the downstream grade-control structure is better and more tangible.

Acknowledgment

The authors would like to thanks Jundi-Shapur University of Technology, Dezul, Iran for the financial support of this work.

References

[1] Sarkar, A. & Dey, S. (2004). Review on local scour due to jets. International Journal of Sediment Research, vol. 19, no. 3, pp. 210-239.

[2] Ghetti, A. & Zanovello, A. (1954). Esame delle escavazioni d'alveo a valle di traverse mediante esperienze su modelli in piccola scala. Proc., 1st Convegno di Costruzioni Idrauliche, pp. 1-9

[3] Rouse, H. (1940). Criteria for similarity in the transportation of sediment, University of Iowa Studies in Engineering of Iowa, vol. 20, pp. 33-49.

[4] Doddiah, D., Albertson, M. L. & Thomas, R. (1953). Scour from jets. Proceedings: Minnesota International Hydraulic Convention, pp. 161-169.

[5] Mason, P. J. & Arumugam, K. (1985). Free jet scour below dams and flip buckets. Journal of Hydraulic Engineering, vol. 111, no. 2, pp. 220-235.

[6] D'Agostino, V. (1996). La progettazione delle controbriglie, Proc., 25th Convegno di Idraulica e Costruzioni Idrauliche,Torino, pp. 107-118 .(in Italian).

[7] Robinson, K. M., Hanson, G. J.& Cook, K. R. (1998). Velocity field measurements at an overfall. Pap. Am. Soc. Agric. Eng., 982063.

[8] Bormann, N. E. & Julien, P. Y. (1991). Scour downstream of grade-control structures. Journal of Hydraulic Engineering, vol. 117, no. 5, pp. 579-594.

[9] Bennett, S. J., Alonso, C. V., Prasad, S. N. & Ro¨mkens, M. J. M. (2000). Experiments on headcut growth migration in concentrated flows typical of upland areas, Water Resources Research, vol. 36, no. 7, pp. 1911–1922.

[10] Yen, C. L. (1987). Discussion on 'Free jet scour below dams and flip buckets, by Peter J. Mason and Kanapathypilly Arumugam. Journal of Hydraulic Engineering, vol. 113, no. 9, pp. 1200-1202.

[11] Azmathulla, H. MD., Ghani, A. AB., Zakaria, N. A., Lai, S. H., Chang, C. K., Leow, C. S. & Abuhasan, Z. (2008). Genetic programming to predict ski-jump bucket spillway scour. Journal of Hydrodynamics, Ser. B, vol. 20, no. 4, pp. 477-484.

[12] Lee, T. L., Jeng, D. S., Zhang, G. H. & Hong, J. H. (2007). Neural network modeling for estimation of scour depth around bridge piers. Journal of Hydrodynamics, Ser. B, vol. 19, no. 3, pp. 378-386.

[13] D'Agostino, V. & Ferro, V. (2004). Scour on alluvial bed downstream of grade-control structures. Journal of Hydraulic Engineering, vol. 130, no. 1, pp. 24-37.

[14] Zhang, Y., Balochian, S. & Bhatnagar, V. (2014). Emerging Trends in Soft Computing Models in Bioinformatics and Biomedicine . The Scientific World Journal, Article ID 683029, 3 pages.

[15] Asghari Esfandani, M. & Nematzadeh, H. (2016). Prediction of Air Pollution in Tehran: Genetic algorithm and back propagation neural network. Journal of AL and Data Mining. Vol. 4, no. 1, pp. 49-54.

[16] Veronese, A. (1937). Erosioni di fondo a valle di uno scarico. Annal. Lavori Pubbl, vol. 75, no. 9, pp. 717-726. (in Italian)

[17] Mossa, M. (1998). Experimental study on the scour downstream of grade-control structures, Proc. 26th Convegno di Idraulica e Costruzioni Idrauliche, Catania, September, vol. 3, pp. 581-594.

[18] D'Agostino, V. (1994). Indagine sullo scavo a valle di opere trasversali mediante modello fisico a fondo mobile. L'Energia elettrica, vol. 71, no. 2, pp. 37-51 (in Italian).

[19] Falciai, M. & Giacomin, A. (1978). Indagine sui gorghi che si formano a valle delle traverse torrentizie. Italia Forestale Montana, vol. 23, no. 3, pp. 111-123.

[20] Lenzi, M. A., Marion, A., Comiti, F. & Gaudio, R. (2000). Riduzione dello scavo a valle di soglie di fondo per effetto dell'interferenza tra le opera, Proc. 27th Convegno di Idraulica e Costruzioni Idrauliche, Genova, September, vol. 3, pp. 271–278 (in Italian).

[21] Jay awardena AW., Muttil N. & Fernando TMKG. (2005). Rainfall-Runoff Modelling using Genetic Programming. International Congress on Modelling and Simulation Society of Australia and New Zealand, pp. 1841-1847

[22] Saltelli, A., K. Chan. & E. M. Scott (Eds). (2000). Sensitivity Analysis. John Wiley and Sons, Ltd.: West Sussex, England.

[23] Geldermann, J. & Rentz, O. (2001). Integrated Technique Assessment with Imprecise Information as a Support for the Identification of Best Available Techniques. OR Spektrum, vol. 23, no. 1, pp. 137-157.

[24] Stiber, N. A., Pantazidou, M. & Small M. J. (1999). Expert System Methodology for Evaluating Reductive Dechlorination at TCE Sites. Environmental Science and Technology, vol. 33, no. 17, pp. 3012-3020.

[25] Critchfield, G. C. & Willard K. E.(1986). Probabilistic Analysis of Decision Trees Using Monte Carlo Simulation. Medical Decision Making, vol. 6, no. 1, pp. 85-92.

[26] McCamley, F. & R. K. Rudel. (1995). Graphical Sensitivity Analysis for Generalized Stochastic Dominance. Journal of Agricultural and Resource Economics, vol. 20, no. 2, pp. 403-403.

نشریه هوش مصنوعی و داده کاوی

تخمین آبشستگی پایین دست سازه سیفون معکوس با استفاده از روش‌های آماری و محاسبات نرم

معصومه فتاحی و بابک لشکرآرا*

گروه مهندسی عمران آب، دانشگاه صنعتی جندی شاپور دزفول، دزفول، ایران.

ارسال ۲۰۱۶/۰۱/۲۷؛ پذیرش ۲۰۱۶/۱۰/۰۴

چکیده:

در این تحقیق، ما از روش‌های رگرسیون غیرخطی، شبکه عصبی وبرنامه‌سازی ژنتیک به منظور پیش‌بینی مساله مهمی نظیر ابعاد آبشستگی پایین دست سازه سیفون معکوس استفاده کرده‌ایم. برای این منظور با استفاده از تحلیل ابعادی و رگرسیون غیرخطی روابطی بدون بعد جهت تخمین حداکثر عمق آبشستگی، مکان حفره آبشستگی و همچنین حداکثر ارتفاع و موقعیت مکانی تپه رسوبی در پایین‌دست این نوع سازه‌ها ارائه شده است. علاوه بر این با استفاده ازسیستم داده‌کاوی برنامه‌سازی ژنتیک و مدل(GP) شبکه عصبی مصنوعی(ANN) نسبت به تدقیق روابط اقدام گردید. تحلیل نتایج نشان داد که رگرسیون غیر خطی به روش گام به گام پیشرو در مقایسه با مدل ارائه شده توسط برنامه‌سازی ژنتیک ومدل شبکه عصبی در تخمین پارامتر نسبی حداکثر عمق آبشستگی s/z بترتیب از ضریب همبستگی ۰/۹۶۲ ، ۰/۹۷۱ و ۰/۹۹۱برخوردار است. از طرفی شیب خط برازش شده از بین نتایج مشاهداتی و محاسباتی در هر سه مدل جهت تخمین پارامترهای بدون بعد s/z، XS/z، h_d/z، XD/z حاکی از برتری پیش‌بینی انجام شده توسط شبکه عصبی مصنوعی می‌باشد.

کلمات کلیدی: آبشستگی، سیفون معکوس، شبکه عصبی، برنامه‌سازی ژنتیک.