

## GEP PREDICTION OF SCOUR AROUND A SIDE WEIR IN CURVED CHANNEL

Fevzi Onen

*Civil Engineering Department, Engineering Faculty, Dicle University, 21280 Diyarbakir, Turkey*

Submitted 21 Mar. 2013; accepted 11 Nov. 2013

**Abstract.** Side-weirs have been widely used in hydraulic and environmental engineering applications. Side-weir is known as a lateral intake structure, which are significant parts of the distribution channel in irrigation, land drainage, and urban sewerage system, by flow diversion device. Local scour involves the removal of material around piers, abutments, side-weir, spurs, and embankments. Clearwater scour depth based on five dimensional parameters: approach flow velocity ( $V_1/V_c$ ), water head ratio  $(h_1-p)/h_1$ , side-weir length ( $L/r$ ), side-weir crest height ( $b/p$ ) and angle of bend  $\theta$ . The aim of this study is to develop a new formulation for prediction of clear-water scour of side-weir intersection along curved channel using Gene Expression Programming (GEP) which is an algorithm based on genetic algorithms (GA) and genetic programming (GP). In addition, the explicit formulations of the developed GEP models are presented. Also equations are obtained using multiple linear regressions (MLR) and multiple nonlinear regressions (MNRL). The performance of GEP is found more influential than multiple linear regression equation for predicting the clear-water scour depth at side-weir intersection along curved channel. Multiple nonlinear regression equation was quite close to GEP, which serve much simpler model with explicit formulation.

**Keywords:** side-weir, clear-water scour, gene expression programming, environmental processes modelling.

### Introduction

A side-weir is an overflow weir set into the side of a channel. Side-weirs are used for water level control, diverting excess water into relief channel during floods, as storm overflows from urban sewage systems and as head regulators of distributaries. The hydraulic behavior and the discharge coefficient of side-weirs for the different type of weirs, main channel, flow and bed conditions have been studied by many researchers such as: Muslu (2002); Cosar, Agaccioğlu (2004); Emiroğlu *et al.* (2010); Agaccioğlu *et al.* (2012). Only a few investigators have been considered for side-weirs flow on an alluvial bed such as: Fares (2000); Agaccioğlu, Onen (2005); Onen, Agaccioğlu (2007a, b).

As can be seen from Figure 1, at sub-critical flow conditions the water surface profile along the weir plane increases from the upstream end towards the downstream end of the side-weir. The flow characteristics are very complicated due to the helicoidally flow pattern when the side-weir is located in a curved channel. In a curved

channel, the inertia of the water keeps the water flowing in a straight path. The water near the surface, which has a velocity somewhat greater than the average velocity, possesses

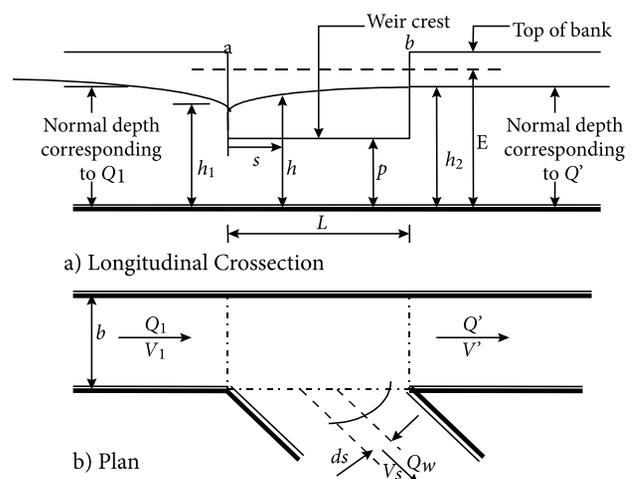


Fig. 1. Definition sketch of subcritical flow over a rectangular side-weir

greater inertia than water near the bottom, which has a velocity less than the average. This greater inertia directs the higher-velocity water toward the outer channel bank, while the slower water moves inward because of the lateral pressure gradient set up by the outward-moving water at the surface. This causes a spiral flow that may persist for some distance downstream. The path of the thread of maximum velocity in an 180° curved channel also deviates from its normal course towards the inner bank at an upstream section of the bend. The path gradually shifts outward from 30° to 60°, and then follows the other bank to about a 150°. After that section, it again returns toward its normal course (Agaccioglu *et al.* 2007).

Scours around the structures such as the bridge piers, the obstacles and lateral intakes are a natural phenomenon caused by the flow of water in rivers and streams. Around these hydraulic structures, scouring introduces the possibilities of a reduction in the support given to foundations of these structures. Scour depths can be seen to depend on the properties of the bed material in the stream (grading, layering, particle shape and size, alluvial or cohesive), the foundation geometries and the properties of flow (Agaccioglu, Onen 2005). According to many investigators, the depth of scour around the structure is related to flow velocity or intensity. Melville and Chiew (1999) mentioned that clear-water scour occurs for mean flow velocities,  $V$  up to the threshold velocity for bed sediment entrainment,  $V_c$ , i.e.  $V \leq V_c$ , while live bed scour occurs for  $V > V_c$ . The maximum depth of scour occurs at  $V = V_c$ .

Under clear water conditions, the local depth of scour in uniform sediment increases almost linearly with flow velocity to a maximum value at the threshold velocity ( $V = V_c$ ). For mean velocity larger than this value ( $V_c$ ), the depth of scour fluctuates periodically with time and about an equilibrium value due to the effect of bed form migration. In addition, Chiew (1984) gave  $H_{d_{max}}/D$  versus  $V/V_c$  with  $d_{50}$  in Figure 2 for bridge piers.

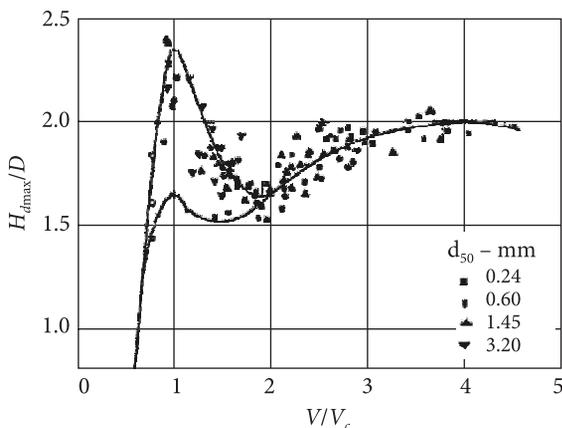


Fig. 2.  $H_{d_{max}}/D$  versus  $V/V_c$  with  $d_{50}$  (Chiew 1984)

The dimensionless parameters of the equilibrium depth of scour at a cylindrical pier of diameter  $D$  in uniform sediment are given by Melville and Chiew (1999) the following:

$$\frac{d_{se}}{D} = f\left(\frac{V}{V_c}, \frac{y}{D}, \frac{d_{50}}{D}, \frac{t}{t_e}\right), \quad (1)$$

where:  $d$  = local scour depth at time  $t$ ;  $d_{se}$  = equilibrium depth of scour;  $t_e$  = time to develop the equilibrium scour depth;  $V/V_c$  = flow intensity;  $y/D$  = flow shallowness and  $d_{50}/D$  = sediment coarseness.

Genetic programming (GP) is a new technique used in the field of water resources engineering. Babovic and Keijzer (2002) applied GP to rainfall-runoff modelling. Harris *et al.* (2003), studied velocity predictions in compound channels with vegetated floodplains using genetic programming. Dorado *et al.* (2003) studied prediction and modelling of the rainfall-runoff transformation of a typical urban basin using artificial neural network (ANN) and GP. Giustolisi (2004) determined Chezy resistance coefficient in corrugated channels by using GP. Khorchani and Blanpain (2005) investigated development of a discharge equation for side-weirs using artificial neural network. Aytok and Kişi (2008) studied a genetic programming approach to suspended sediment modelling. Guven and Gunal (2008) studied a genetic programming approach for prediction of local scour downstream hydraulic structures. Guven and Kisi (2010) investigated linear genetic programming for time series modelling of daily flow rate. Azamathulla *et al.* (2008) studied genetic programming to predict ski-jump bucket spill-way scour. Guven and Aytok (2009) studied new approach for stage-discharge relationship: Gene-expression programming. Eldrandaly (2009) studied integrating gene expression programming and geographic information systems for solving a multi-site land use allocation problem. Azamathulla *et al.* (2010) applied GP to predict bridge pier scour. Bilhan *et al.* (2011) applied neural Networks for estimation of discharge capacity of triangular labyrinth side-weir located on a straight channel. Kisi *et al.* (2012) investigated prediction of lateral out flow over triangular labyrinth side weirs under subcritical conditions using soft computing approaches. Onen (2013) predicted penetration depth in a plunging water jet using soft computing approaches. Bagatur and Onen (2013) studied a predictive model on air entrainment by plunging water jets using GEP and ANN.

The objective of this current study is to develop a new formulation technique for prediction of Clear-water scour of side-weir intersection along curved channel using GEP. For this aim, experimental data are used. The performance of the GEP models was compared with multiple linear regressions and multiple nonlinear regressions.

### 2. Experimental set-up

Experiments to study the development of scour hole around side-weir, a model located on a rectangular flume were conducted in a Plexiglas flume at the Hydraulic Laboratory of Yildiz Technical University, Istanbul, Turkey. The main and discharge flume located between two straight channel are central angle of 180° with 2.95 m and 3.40 m radius to centreline, respectively. The bend is connected with an upstream straight reach of 14 m and a downstream straight reach of 3 m. The main flume is 0.40 m wide and a 0.55 m deep with a well-finished aluminium bottom. The discharge collection flume is 0.55 m deep, 0.50 m wide and situated parallel to the main flume. Details of the flume are shown in Figure 3.

Uniform bed materials were placed on as a 20 cm thick layer in the flume with a bed slope of 0.001. The bed of flume was firstly mixed to prevent sediment grading and layering, and then flattened before each run. During the experiments, the upstream valve was slowly adjusted without causing any disturbance to bed material until the desired flow condition was obtained in the flume. No sediment inflow was allowed from upstream due to clear-water conditions. Three different lengths (25, 40 and 50 cm) and three different weir crest height (7, 12 and 17 cm) of the side weir were located at the 30°, 60°, 90°, 120° and 150° section of bend. The dimensions of side-weirs and the ranges of the parameters tested are given in Table 1. Side-weirs were sharp edged and fully aerated on the overflowed side. The water depth is controlled by a sluice gate located at the end of the main flume to produce the uniform flow. Water is supplied from an overhead tank with a constant head and measured by a calibrated 90° V notched weir ( $Q_1$ ) at the beginning of the main channel.

$Q_1 = \left(\frac{h}{5.11}\right)^{2.447}$  was used for measuring the flow with 90° V notched weir ( $h$ , cm). Another calibrated 90° V

notched weir measured the side-weirs flow rate ( $Q_w$ ), which was located at the downstream end of the collection channel.  $Q_w = 0.0134 \times h^{2.5896}$  was used for side-weirs flow rate. Then, the downstream discharge ( $Q'$ ) in the main channel according to the side-weir location was determined as follows:

$$Q' = Q_1 - Q_w \tag{2}$$

The measurements of water surfaces and bed levels were determined using point gauges capable of reading to the nearest 0.1 mm. The mean particle size of the bed material ( $d_{50}$ ) is 1.15 mm. The critical velocity of initiation of motion, ( $V_c$ ) was determined using Shields diagram with a series of test run and given the following:

$$V_c = 0.1556 \log(4809h_1) \tag{3}$$

where:  $h_1$  = flow depth at the upstream end of side-weir at the flume centre as in m and  $V_c$  is in m/s. Clear water scour experiments were done in the duration between 300–720 minutes (Agaccioglu, Onen 2005).

The changes in the main channel bottom topography around side-weirs are briefly mentioned according to flow visualization and the obtained bottom topography at 30° section of the bend as given in Figure 4, when the side-weir has  $L = 40$  cm and  $p = 7$  cm. The reduction in velocities and shear stresses causes to the development of stagnation along the inner bank and separation at the upstream corner of the side overflow at the downstream end of the side-weir in the main channel. They cause scour at the centre of reverse flow area. The accumulation in front of the upstream of the side-weir and the sand bar behind of the scour hole is also observed. The location of maximum scour hole moves toward the downstream with an increasing of the dimensionless parameters of approach flow intensity  $V_1/V_c$  as well as the water head ratio  $(h_1-p)/h_1$  the weir crest height  $b/p$  and the weir length  $L/r$ . As mentioned below when the approach flow velocity or remaining momentum towards the downstream direction in

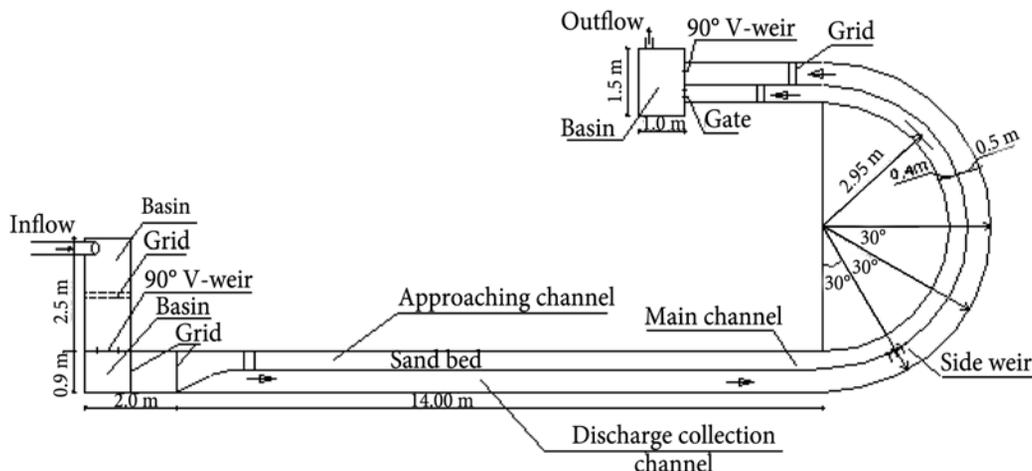


Fig. 3. Plan of test flume

the main channel increases then reverse flow area and the location of scour move toward the downstream end of the side-weir (Agaccioğlu *et al.* 2007).

The proposed GEP model formulations are trained with experimental data taken from experimental study (Table 1). The statistical parameters such as; the arithmetic mean ( $x_{mean}$ ), standard deviation ( $S_x$ ) and uncertainty ( $u$ ) reflects the distribution properties of any data set. The uncertainty in measurements ( $x_{mean} = 0.382, S_x = 0.31, u =$

$S_x/n^{0.5}, \% u = ((u/x_{mean}) \times 100)$  is about  $\pm 0.0231$  (6.05%) for Clear-water scour of side-weir intersection along curved channel.

### 3. Dimensionless parameters

The variables that affect the phenomenon of scour depth,  $H_p$ , produced by a side-weir located on a channel having an alluvial bed are the mean upstream velocity of flow,  $V_1$ ;

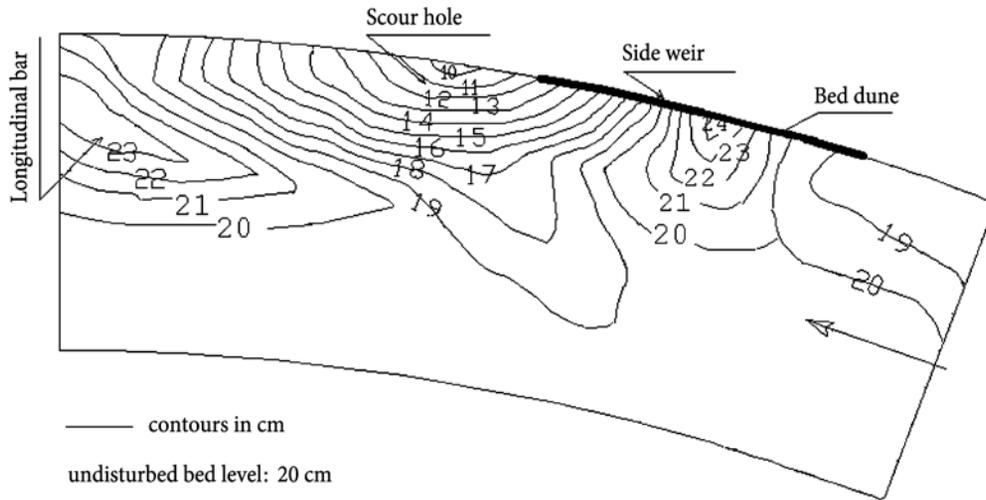


Fig. 4. Bed topography at a side-weir intersection when  $V_1/V_c = 0.98$

Table 1. The dimensions of side-weirs and the ranges of parameters tested for curved channel

$\theta$	$p$ (cm)	$L$ (cm)	$h_1$ (cm)	$(h_1-p)/h_1$	$b/p$	$L/r$	$V_1$ (m/s)	$V_c$ (m/s)	$V_1/V_c$
(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)	(10)	(11)
30°	7	25	15–21	0.50–0.70	5.71	0.085	0.25–0.47	0.44–0.47	0.57–1.0
	12	25	21–27	0.45–0.60	3.33	0.085	0.31–0.49	0.47–0.49	0.66–1.0
	17	25	27–30	0.35–0.45	2.35	0.085	0.34–0.49	0.47–0.49	0.69–1.0
	7	40	15–21	0.50–0.70	5.71	0.136	0.25–0.49	0.44–0.47	0.57–1.0
	12	40	21–27	0.45–0.60	3.33	0.136	0.31–0.49	0.47–0.49	0.58–1.0
	17	40	27–30	0.35–0.45	2.35	0.136	0.34–0.49	0.47–0.49	0.63–1.0
	7	50	14–16	0.50–0.60	5.71	0.170	0.25–0.45	0.44–0.46	0.60–0.98
	12	50	20–25	0.40–0.55	3.33	0.170	0.31–0.48	0.45–0.48	0.66–1.0
	17	50	27–30	0.35–0.45	2.35	0.170	0.34–0.49	0.47–0.49	0.69–1.0
60°	7	40	15–20	0.50–0.65	5.71	0.136	0.25–0.47	0.44–0.47	0.58–1.0
	12	40	21–27	0.45–0.60	3.33	0.136	0.28–0.48	0.46–0.48	0.60–1.0
90°	7	40	15–20	0.50–0.65	5.71	0.136	0.25–0.47	0.44–0.47	0.58–1.0
	12	40	21–27	0.45–0.60	3.33	0.136	0.31–0.48	0.47–0.49	0.66–1.0
120°	7	25	15–21	0.50–0.70	5.71	0.085	0.25–0.47	0.44–0.47	0.57–1.0
	12	25	21–27	0.45–0.60	3.33	0.085	0.31–0.48	0.47–0.49	0.66–1.0
	17	25	27–30	0.40–0.45	2.35	0.085	0.34–0.48	0.48–0.49	0.69–1.0
	7	40	15–21	0.50–0.65	5.71	0.136	0.25–0.47	0.44–0.47	0.57–1.0
	12	40	21–27	0.45–0.55	3.33	0.136	0.27–0.48	0.47–0.49	0.58–1.0
	17	40	27–30	0.37–0.45	2.35	0.136	0.34–0.48	0.48–0.49	0.63–1.0
	7	50	14–16	0.50–0.62	5.71	0.170	0.27–0.45	0.44–0.45	0.60–1.0
	12	50	20–25	0.40–0.55	3.33	0.170	0.31–0.48	0.46–0.48	0.66–1.0
	17	50	27–30	0.35–0.45	2.35	0.170	0.34–0.49	0.47–0.49	0.69–1.0
150°	7	40	15–20	0.50–0.65	5.71	0.136	0.27–0.47	0.44–0.47	0.57–1.0
	12	40	21–27	0.45–0.60	3.33	0.136	0.31–0.48	0.47–0.49	0.66–1.0

the critical mean flow velocity for entrainment of bed sediment,  $V_c$ ; the upstream depth of flow,  $h_1$ ; the width of side-weir,  $L$ ; the crest height of side-weir,  $p$ ; the width of channel,  $b$ ; the radius of curved channel,  $r$ ; the angle of the bend,  $\theta$ . Dimensionless analysis defines the equilibrium depth of scour at the side-weir intersection on an alluvial channel through the following function:

$$\frac{H_d}{p} = f(V_1/V_c, (h_1 - p)/h_1, b/p, L/r, \theta), \quad (4)$$

where: dimensionless parameters on the right-hand side  $V_1/V_c$ ,  $(h_1 - p)/h_1$ ,  $b/p$ ,  $L/r$ , and  $\theta$  represent, respectively, approach flow velocity, water head ratio, side-weir crest height, side-weir length, and the angle of bend (Agaccioğlu *et al.* 2007).

In this study, the dimensionless parameters in equation were used as input and output parameters in GEP modelling. Table 1 reports the ranges of input and output parameters.

#### 4. Multiple linear regressions

Multiple linear regressions (MLR) is a method used to model the linear relationship between a dependent variable and one or more independent variables. The dependent variable is sometimes also called the predict, and the independent variables the predictors. MLR is based on least squares: the model is fit such that the sum-of-squares of differences of observed and predicted values is minimized. The general Multiple linear regression (MLR) models with response  $y$  and terms  $x_1, \dots, x_i$  have the form:

$$y = b_0 + b_1x_1 + \dots + b_ix_i, \quad (5)$$

where:  $b_i$  = model coefficients;  $y$  = predicted variable;  $x_i$  = independent parameters. The model coefficients ( $b_i$ ) were calculated by least square regression. Values of root mean square error (RMSE) were used to find the fitting degrees of linear models with experimental data. The best model selection process stopped when the lowest RMSE value was reached. Multiple linear regression analysis yielded the following equations:

$$\begin{aligned} H_d/p &= 2.2 + 1.67V_1/V_c + 0.000056L/r + \\ &1.83(h_1 - p)/h_1 + 0.085b/p; \\ R^2 &= 0.84; \end{aligned} \quad (6)$$

$$\begin{aligned} H_d/p &= -1.68 + 1.86V_1/V_c - 0.0058L/r - \\ &0.164(h_1 - p)/h_1 + 0.125b/p; \\ R^2 &= 0.81; \end{aligned} \quad (7)$$

$$\begin{aligned} H_d/p &= -1.59 + 1.85V_1/V_c - 0.0054L/r - 0.139 \\ &(h_1 - p)/h_1 + 0.127b/p - 0.118\sin\theta; \\ R^2 &= 0.82; \end{aligned} \quad (8)$$

Eqn (6) contains four independent variables and used data of 30°. Eqn (7) contains four independent variables

and used data from 30° to 150°. Eqn (8) contains five independent variables and used data from 30° to 150°.

#### 5. Multiple non-linear regression

Multiple non-linear functions of interest are as follows:

$$Y = a X_1^b X_2^c X_3^d. \quad (9)$$

Nonlinear relationships between the dependent variables and the independent variables based on multivariate power function were considered. The following equations were derived:

$$\begin{aligned} H_d/p &= 0.87(V_1/V_c)^{2.7}(L/r)^{0.356}(h_1 - p)/h_1^{3.07}(b/p)^{0.58}; \\ R^2 &= 0.94; \end{aligned} \quad (10)$$

$$\begin{aligned} H_d/p &= 0.18(V_1/V_c)^{3.94}(L/r)^{-0.12^*}(h_1 - p)/h_1^{0.16}(b/p)^{1.27}; \\ R^2 &= 0.87; \end{aligned} \quad (11)$$

$$\begin{aligned} H_d/p &= 0.158(V_1/V_c)^{3.9}(L/r)^{-0.107^*}(h_1 - p)/h_1^{0.154} \times \\ &(b/p)^{1.28}\sin\theta^{-0.263}; \\ R^2 &= 0.88. \end{aligned} \quad (12)$$

Eqn (10) obtained for 30°, Eqn (11) and Eqn (12) obtained from 30° to 150°.

#### 6. Gene expression programming (GEP)

Ferreira proposed a new algorithm based on genetic algorithms (GA) and genetic programming and incorporates both the simple, linear chromosomes of fixed length similar to the ones used in genetic algorithms and the ramified structures of different sizes and shapes similar to the parse trees of genetic programming. So, the phenotype of GEP is made of the same kind of ramified structures used in genetic programming, but the ramified structures created by GEP are the expression of a totally autonomous genome. The main aim of GEP find a mathematical function principal using a set of data presented. For the mathematical equation the GEP process performed the symbolic regression by means of the most of the genetic operators of GA. The brief flowchart of GEP is given in Figure 5. Initially, chromosomes of each individual of the population are randomly generated. Then chromosomes are mean and each individual is appreciated based on fitness function and selected to reproduce with modification. The individuals of new generation are subjected to some developmental processes such as expression of the genomes, confrontation of the selection environment and reproduction with modification. The process is repeated for a certain number of generations or until a good solution is found (Ferreira 2002).

The two main players of GEP are the chromosomes and expression trees (ETs). The expression of the genetic information's encoded in the chromosome. As in nature, the process of information decoding is named translation and this translation implies a code and a set of rules. The

genetic code of GEP is very simple; a one-to-one relation between the symbols of the chromosome and the nodes they represent in the trees. The rules determine the spatial organization of nodes in the expression trees and the type of interaction between sub-ETs. Thus, there are two languages in GEP; the language of the genes and the language of expression trees and, thanks to ETs structures which are determined from simple rules and their interactions. It is possible to at once infer the expression tree given the sequence of a gene and vice versa. This is termed Karva language (Ferreira 2002). The GEP genes have two main parts called the head and tail. The head includes some mathematical operators, variables and constants (+, -, \*, /, sin, cos, 1, a, b, c) and they are used to encode a mathematical expression. The tail includes variables and constants (1, a, b, c) named as terminal symbols. In GEP, some operators such as selection, transposition, mutation and cross-over (recombination) are used to modify chromosomes for the next generation.

## 7. GEP model development

To generate the mathematical function for the prediction of the clear-water scour of a side-weir intersection along curved channel was the main aim of development of GEP models. For that reason, a development of GEP model was realized. The GEP model has five input parameters (approach flow velocity, water head ratio, side-weir crest

height, side-weir length and angle of bend). The parameters of GEP models were presented in Table 2.

Table 2. Parameters of the GEP model

Chromosomes	30
Fitness function error type	R <sup>2</sup>
Number of the genes	3
Head size	8
Linking function	*
Function set	+, -, *, /, √, x <sup>2</sup>
Mutation rate	0.044
One-point recombination rate	0.3
Two-point recombination rate	0.3
Inversion rate	0.1
Transposition rate	0.1

There are five major preliminary steps to solve a problem using GEP. These are the determination of: (1) the set of terminals; (2) the set of functions; (3) the fitness measure; (4) the values of numerical parameters and qualitative for controlling the run; and (5) the criterion for designating a results and terminating a run (Ferreira 2006).

The first major step in preparing to employ the GEP paradigm is to identify the set of terminals to be used in the individual computer programs. The major steps of terminal sets contain the independent variables of the problem, the state variables of the system and the functions with no arguments. The terminal set used in present study is  $\left\{V_1/V_c, L/r, (h_1 - p)/h_1, b/p, \theta\right\}$ .

The second major step is to define the set of functions and arithmetic operations. We used the four basic arithmetic operators (+, -, \*, /), and some basic mathematical functions (√, x<sup>2</sup>). The third major step is fitness measure, which specifies the way of evaluating how good a given problem solves a particular problem. In the present study, root mean square error (RMSE) of the training set is taken as fitness function. The terminals and the functions are the components of the programs that form the junctions in the tree. The four major steps is the selection of certain parameters to control the runs. The control parameters contain the size of the population, the rate of crossover, etc. the last step is the determination of the criteria to terminate the run. If the RMSE between the model results got reasonably close to zero, then the model was considered acceptable (Kisi *et al.* 2012).

We used a length of the head, h = 8 and three genes per chromosome. In this case, we linked the sub-ETs by multiplication.

Figure 6 shows the ET of the formulation which actually is:

$$\text{pow}((d[0] \times G1C0), 2) \times (((d[4] - d[2]) / (d[0] + G2C0)) \times d[0]) \times (((d[4] - G3C1) - (d[3] - d[2])) / ((d[2] \times d[2]) \times d[4])) + d[0]), \quad (13)$$

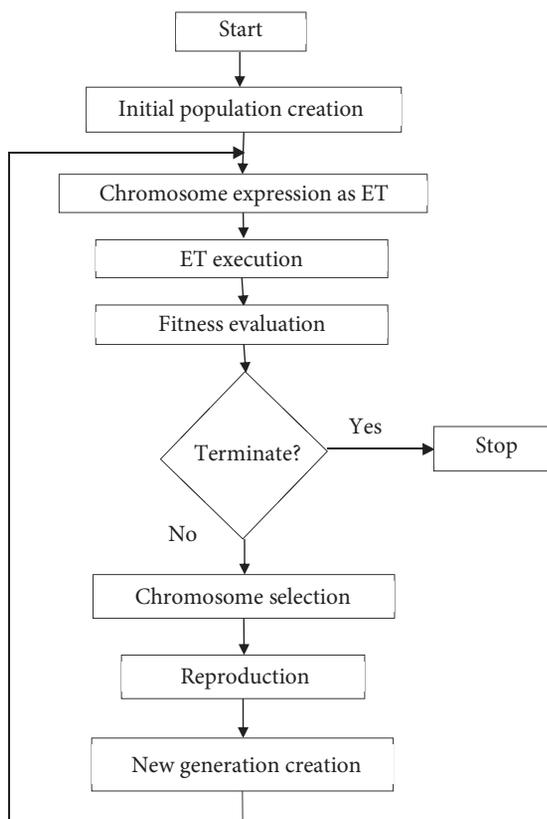


Fig. 5. GEP algorithm

where: the constants in the formulation are G1C0 = -0.989; G2C0 = -5.78 and G3C1 = 2.86 and the actual parameters are d[0] = V<sub>1</sub>/V<sub>c</sub>, d[1] = L/r, d[2] = b/p, d[3] = (h<sub>1</sub> - p)/h<sub>1</sub> and d[4] = sin. After putting the corresponding values, the final equation becomes for 30°-150° of bend:

$$H_d / p = \left[ -0.989 \frac{V_1}{V_c} \right]^2 \left[ \frac{\sin\theta - (b/p)}{(V_1/V_c) - 5.78} \left( \frac{V_1}{V_c} \right) \right] \times \left[ \frac{\sin\theta - (h_1 - p)/h_1 + (b/p) - 2.86}{(b/p)^2 \times \sin\theta} + \frac{V_1}{V_c} \right]. \quad (14)$$

It should be noted that the proposed GEP formulation is valid for the ranges of the training set given in Table 2. The prediction of the proposed GEP formulation versus actual experimental values for training and testing sets are given in Figures 7 and 8. Eqn (14) contains five independent variables and used data from 30° to 150°. Figure 7 shows how good GEP learned the nonlinear relation between parameters and Figure 8 proves the high generalization capacity of the proposed GEP formulation. Eqn (15) contains four independent variables and used data from 30° to 150°.

$$H_d / p = \left[ \frac{V_1}{V_c} - \frac{(V_1/V_c) + 2.3 + (b/p)^{0.5}}{3.3(b/p)} \right] \times \left[ \frac{-2.09}{6.44 + (V_1/V_c) + (b/p) - (h_1 - p)/h_1} + \frac{V_1}{V_c} \right]^2 \times [b/p]^{0.5}; \quad (15)$$

$$H_d / p = \left[ \frac{V_1}{V_c} - \left( (V_1/V_c + 7.75 - 7.75/L/r \times (V_1/V_c)(h_1 - p)/h_1) \right)^2 \right] \times \left[ (L/r - V_1/V_c) \frac{b}{p} + (L/r)^2 - (V_1/V_c \times \frac{b}{p} + 74.13) \right] \times \left[ \frac{(h_1 - p)/h_1}{L/r} \times \frac{(h_1 - p)/h_1}{64 - 6.44 \times b/p} \right]. \quad (16)$$

Eqn (15) contains four independent variables and used data from 30° to 150°. Eqn (16) contains four independent variables and used data for 30° of bend.

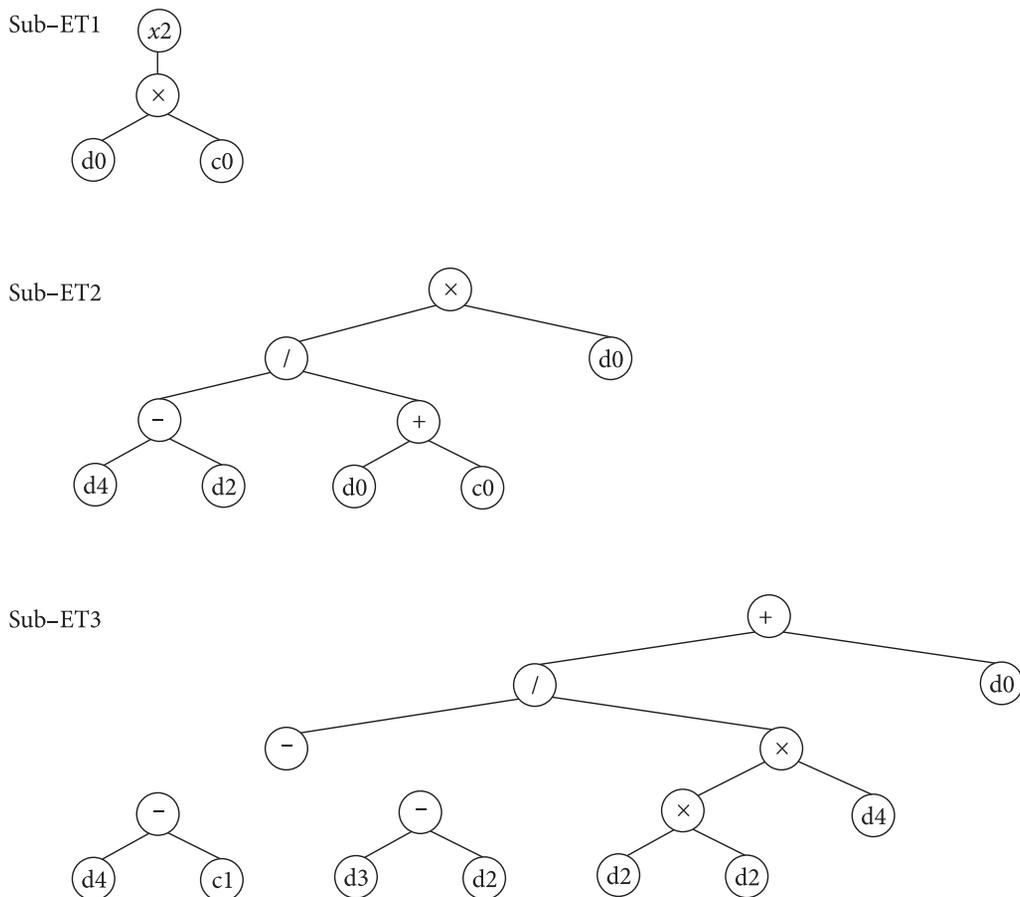


Fig. 6. Expression tree (ET) for the proposed GEP formulation

**8. Comparison of GEP with regression model**

The statistical performances of both GEP and regression models are summarized in Table 3. As far as Table 3 is concerned, satisfactory agreement between the model predictions and experimental data is observed for most of models. In statistics, the overall error performances of the relationship between two groups can be interpreted from coefficient of correlation ( $R$ ) values. If a proposed model gives  $R > 0.8$ , there is strong correlation between measured and prediction values for overall the data available in the database. In addition to this, the statistical performance of any model is evaluated in terms of some error criteria such as root-mean-square (RMSE), which is significant

criteria as well as  $R$  value, since sometimes a model with high  $R^2$  value may exhibit high RMSE (Taskiran 2010).

Both multiple linear regressions and multiple nonlinear regressions are applied between  $H_d/p$  with  $V_1/V_c$ ,  $L/r$ ,  $(h_1-p)/h_1$ ,  $b/p$  for  $30^\circ$ . It is applied between  $H_d/p$  with  $V_1/V_c$ ,  $L/r$ ,  $(h_1-p)/h_1$  and  $b/p$  for  $30^\circ-150^\circ$ . Then it is applied between  $H_d/p$  with  $V_1/V_c$ ,  $L/r$ ,  $(h_1-p)/h_1$ ,  $b/p$  and  $\theta$  for  $30^\circ-150^\circ$ . We obtained six regression formulations (Eqns (6), (7), (8), (10), (11) and (12)). In the GEP model, first five independent parameters ( $V_1/V_c$ ,  $L/r$ ,  $(h_1-p)/h_1$ ,  $b/p$ ,  $\theta$ ) used for  $30^\circ-150^\circ$  of bend, four independent parameters ( $V_1/V_c$ ,  $L/r$ ,  $(h_1-p)/h_1$ ,  $b/p$ ) used for  $30^\circ-150^\circ$  of bend and then four independent parameters ( $V_1/V_c$ ,  $L/r$ ,  $(h_1-p)/h_1$ ,

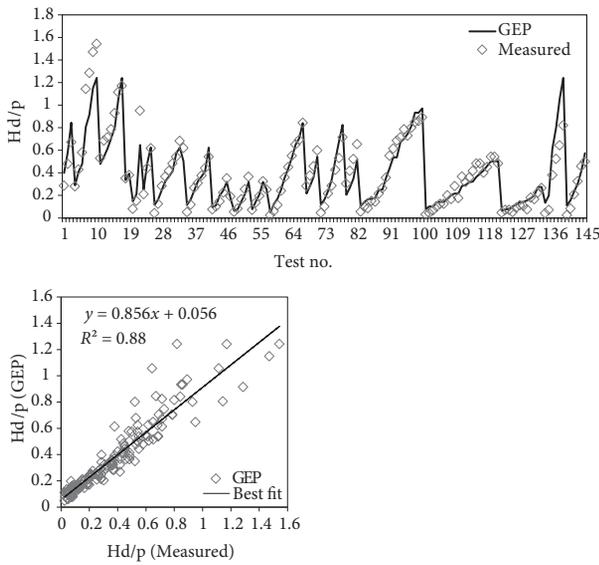


Fig. 7. GEP estimations of  $H_d/p$  versus measured ones for train set (model 1)

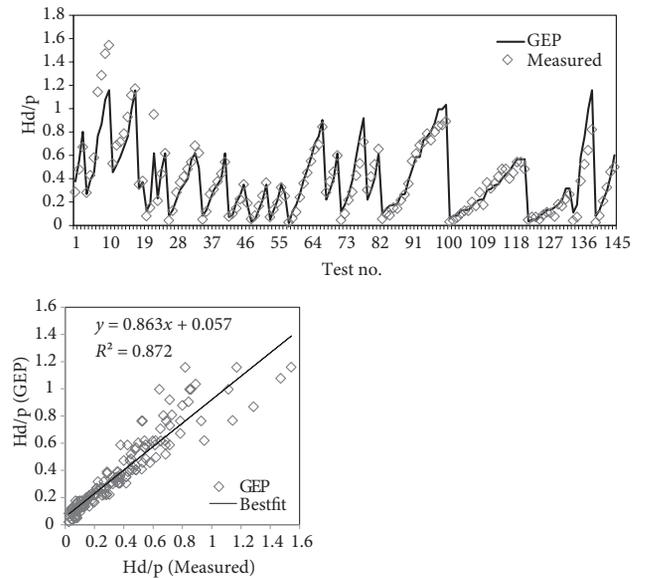


Fig. 9. GEP estimations of  $H_d/p$  versus measured ones for train set (model 2)

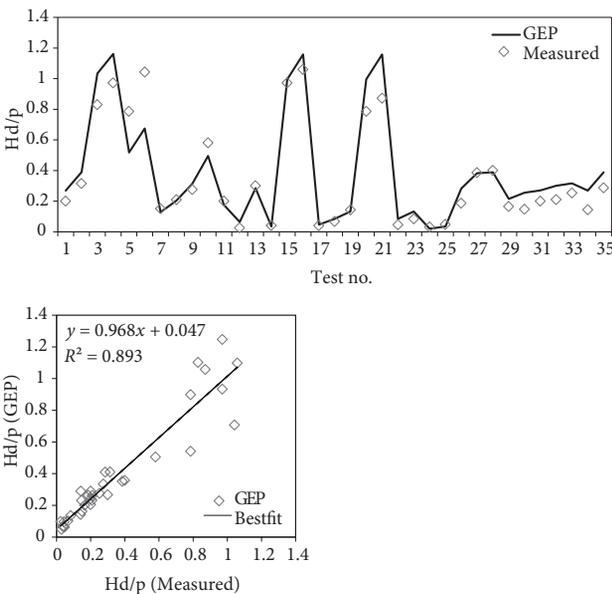


Fig. 8. GEP estimations of  $H_d/p$  versus measured ones for test set (model 1)

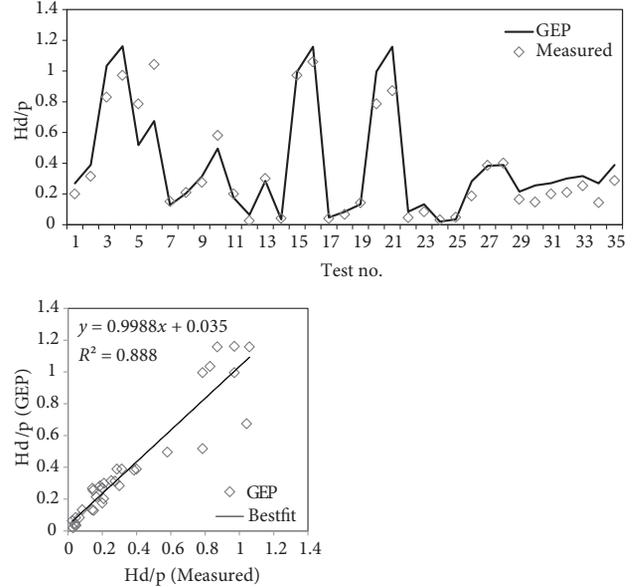


Fig. 10. GEP estimations of  $H_d/p$  versus measured ones for test set (model 2)

$b/p$ ) used for 30° of bend. Also, explicit formulations of developed GEP models are obtained (Eqns (14), (15) and (16)). The performance of GEP is found more influential when compared to multiple linear regressions (MLR) for predicting the clear-water scour of side-weir intersection along curved channel in Table 3.

Table 3. Statistical performance of GEP and regression models (MLR/MNLR)

Method	RMSE	R <sup>2</sup>
MLR Eqn (6)	0.146	0.840
MLR Eqn (7)	0.221	0.810
MLR Eqn (8)	0.200	0.816
MNLR Eqn (10)	0.093	0.938
MNLR Eqn (11)	0.114	0.866
MNLR Eqn (12)	0.108	0.881
GEP Eqn (14)	0.105	0.882 (train)
	0.117	0.893 (test)
GEP Eqn (15)	0.109	0.872 (train)
	0.121	0.888 (test)
GEP Eqn (16)	0.101	0.927 (train)
	0.096	0.928 (test)

**9. Training and testing results of GEP modelling**

The performance of GEP in training and testing sets is validated in terms of the common statistical measures R<sup>2</sup> (coefficient of determination), root-mean-square (RMSE):

$$R^2 = \left( \frac{\sum xy}{\sum x^2 \sum y^2} \right)^2 ; \tag{17}$$

$$RMSE = \left[ \frac{\sum (X - Y)^2}{n} \right]^{\frac{1}{2}}, \tag{18}$$

where:  $x$  = observed values;  $y$  = predicted value and  $n$  = number of samples (Azamathulla *et al.* 2010).

Table 4. Sensitivity analysis for independent parameters for the testing set

Model	RMSE	R <sup>2</sup>
1 $H_d/p = f(V/V_c, L/r, h_1-p/h_1, b/p, \theta)$	0.117	0.893
2 $H_d/p = f(V/V_c, L/r, h_1-p/h_1, b/p)$	0.121	0.888
3 $H_d/p = f(V/V_c, L/r, b/p)$	0.125	0.878
4 $H_d/p = f(V/V_c, L/r, h_1-p/h_1)$	0.120	0.867
5 $H_d/p = f(V/V_c, h_1-p/h_1, b/p)$	0.126	0.882
6 $H_d/p = f(L/r, h_1-p/h_1, b/p)$	0.214	0.612

First, to assess the significance of each input parameter, attempt was performed on  $H_d/p$ . The GEP models are compared with one of the independent parameters removed in each case in Table 4, and deleting any independent

parameter from the input set yielded larger RMSE and lower R<sup>2</sup> values. These five independent parameters have non-negligible influence on  $H_d/p$  and thus in this study the functional relationship given in Eqn (14) is used for GEP modelling. Also  $V/V_c$  is very effective on the scour ( $R^2 = 0.612$  and  $RMSE = 0.214$ ) in model 6. The testing performance of proposed GEP model showed a high generalization capacity with  $R^2 = 0.893$  and  $RMSE = 0.117$ .

**Conclusions**

1. A new and efficient approach for the formulation of clear-water scour of a side-weir intersection along curved channel using GEP for the first time in the literature is presented in this study.

2. The proposed GEP formulation is empirical and based on experimental results. For maximum scour depth, regression formulations are proposed. The testing performance of proposed GEP model showed a high generalization capacity with  $R^2 = 0.893$  and  $RMSE = 0.117$  for maximum scour depth.

3. GEP formulations are compared with experimental results, multiple linear regression equation (MLR), multiple nonlinear regression equation (MNLR). MNLR were quite close to GEP.

4. The results of this study recommend that GEP modelling is a more sophisticated and improved than the corresponding conventional approaches for predicting clear-water scour of a side-weir intersection along curved channel.

**References**

Agaccioglu, H.; Onen, F. 2005. Clear water scour at a side-weir intersection along the bend, *Irrigation and Drainage* 54(5): 553–569. <http://dx.doi.org/10.1002/ird.209>

Agaccioglu, H.; Onen, F.; Toprak, Z. F. 2007. Scour around a side-weir at a 30° section of a 180° alluvial curved channel, *Irrigation and Drainage* 56(4): 423–438. <http://dx.doi.org/10.1002/ird.304>

Agaccioglu, H.; Emiroglu, M. E.; Kaya, N. 2012. Discharge coefficient of side weirs in curved channels, *Proceedings of the Institution of Civil Engineers Water Management* 165(6): 339–352.

Aytek, A.; Kisi, O. 2008. Genetic programming approach to suspended sediment modelling, *Journal of Hydraulic* 351: 288–298.

Azamathulla, H. M. D.; Ghani, A. A. B.; Zakaria, N. A.; Lai, S. H.; Chang, C. K.; Leow, C. S.; Abuhasan, Z. 2008. Genetic programming to predict ski-jump bucket spill-way scour, *Journal of Hydrodynamics* 20(4): 477–484. [http://dx.doi.org/10.1016/S1001-6058\(08\)60083-9](http://dx.doi.org/10.1016/S1001-6058(08)60083-9)

Azamathulla, H. M. D.; Ghani, A. A. B.; Zakaria, N. A.; Guven, A. 2010. Genetic programming to predict bridge pier scour, *Journal of Hydraulic Engineering* 136(2): 165–169. [http://dx.doi.org/10.1061/\(ASCE\)HY.1943-7900.0000133](http://dx.doi.org/10.1061/(ASCE)HY.1943-7900.0000133)

Babovic, V.; Keijzer, M. 2002. Rainfall-runoff modeling based on genetic programming, *Nordic Hydrology* 33(5): 331–346.

- Bagatur, T.; Onen, F. 2013. A predictive model on air entrainment by plunging water jets using GEP and ANN, *KSCE Journal of Civil Engineering* 18(1): 304–314. <http://dx.doi.org/10.1007/s12205-013-0210-7>
- Bilhan, O.; Emiroglu, M. E.; Kisi, O. 2011. Use of artificial neural networks for prediction of discharge coefficient of triangular labyrinth side-weir in curved channels, *Advance in Engineering Software* 42(4): 208–214. <http://dx.doi.org/10.1016/j.advengsoft.2011.02.006>
- Chiew, Y. M. 1984. *Local scour at bridge piers*: Ph.D. Thesis, School of Engineering, University of Auckland, New Zealand, No. 355.
- Cosar, A.; Agaccioğlu, H. 2004. Discharge coefficient of a triangular side-weir located on a curved channel, *Journal of Irrigation and Drainage Engineering* ASCE 130(5): 410–423. [http://dx.doi.org/10.1061/\(ASCE\)0733-9437\(2004\)130:5\(410\)](http://dx.doi.org/10.1061/(ASCE)0733-9437(2004)130:5(410))
- Dorado, J.; Rabunal, J. R.; Pazos, A.; Rivero, D.; Santos, A.; Puertas, J. 2003. Prediction and modelling of the rainfall-runoff transformation of a typical urban basin using ANN and GP, *Application Artificial Intelligence* 17(4): 329–343. <http://dx.doi.org/10.1080/713827142>
- Eldrandaly, K. 2009. Integrating gene expression programming and geographic information systems for solving a multi site land use allocation problem, *American Journal of Applied Sciences* 6(5): 1021–1027. <http://dx.doi.org/10.3844/ajassp.2009.1021.1027>
- Emiroglu, M. E.; Kaya, N.; Agaccioğlu, H. 2010. Discharge capacity of labyrinth side weir located on a straight channel, *Journal of Irrigation and Drainage Engineering* ASCE 136(1): 37–46. [http://dx.doi.org/10.1061/\(ASCE\)IR.1943-4774.0000112](http://dx.doi.org/10.1061/(ASCE)IR.1943-4774.0000112)
- Fares, Y. R. 2000. Changes of bed topography in meandering rivers at a neck cutoff intersection, *Journal of Environmental Hydrology* 8(13): 1–18.
- Ferreira, C. 2002. *Gene expression programming: mathematical modelling by an artificial intelligence*. Portugal: Angra do Heroísmo. 478 p.
- Ferreira, C. 2006. *Gene-expression programming: mathematical modelling by an artificial intelligence*. Berlin: Springer. 478 p.
- Giustolisi, O. 2004. Using genetic programming to determine Chezyresistance coefficient in corrugated channels, *Journal of Hydroinformatics* 6(3): 157–173.
- Güven, A.; Günel, M. 2008. Genetic programming approach for prediction of local scour downstream hydraulic structures, *Journal of Irrigation and Drainage Engineering* 134(2): 241–249. [http://dx.doi.org/10.1061/\(ASCE\)0733-9437\(2008\)134:2\(241\)](http://dx.doi.org/10.1061/(ASCE)0733-9437(2008)134:2(241))
- Güven, A.; Aytekin, A. 2009. New approach for stage-discharge relationship: gene-expression programming, *Journal of Hydrologic Engineering* 14(8): 812–820. [http://dx.doi.org/10.1061/\(ASCE\)HE.1943-5584.0000044](http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000044)
- Güven, A.; Kisi, O. 2010. Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming, *Water Resources Management* 25(2): 691–704. <http://dx.doi.org/10.1007/s11269-010-9721-x>
- Harris, E. L.; Babovic, V.; Falconer, R. A. 2003. Velocity predictions in compound channels with vegetated floodplains using genetic programming, *International Journal of River Management* 1(2): 117–123. <http://dx.doi.org/10.1080/15715124.2003.9635198>
- Khorchani, M.; Blanpain, O. 2005. Development of a discharge equation for side-weirs using artificial neural network, *Journal of Hydroinformatics* 7: 31–39
- Kisi, O.; Emiroglu, M. E.; Bilhan, O.; Güven, A. 2012. Prediction of lateral out flow over triangular labyrinth side weirs under subcritical conditions using soft computing approaches, *Expert systems with Applications* 39(3): 3454–3460. <http://dx.doi.org/10.1016/j.eswa.2011.09.035>
- Melville, B. W.; Chiew, Y. M. 1999. Time scale for local scour at bridge piers, *Journal of Hydraulic Engineering* ASCE 125(1): 59–65.
- Onen, F.; Agaccioğlu, H. 2007a. Local scours around side-weirs located on alluvial bends, *Teknik Dergi* 18(1): 4035–4053.
- Onen, F.; Agaccioğlu, H. 2007b. Scour at a side-weir intersection located on an alluvial river, *Nordic Hydrology* 38(2): 165–176. <http://dx.doi.org/10.2166/nh.2007.005>
- Onen, F.; 2013. Prediction of penetration depth in a plunging water jet using soft computing approaches, *Neural Computing and Application*, September 2013. <http://dx.doi.org/10.1007/s00521-013-1475-y>
- Taskiran, T. 2010. Prediction of California bearing ratio (CBR) of fine grained soils by AI methods, *Advances in Engineering Software* 41(9): 1115–1123.

**Fevzi ONEN**, Assistant Professor, Dr, Department of Civil Engineering, Faculty of Engineering, Dicle University, Diyarbakir, Turkey. Publications: author of several scientific publications. Research interests: water resources, scouring in open channel, hydrology and hydroelectric power plant.