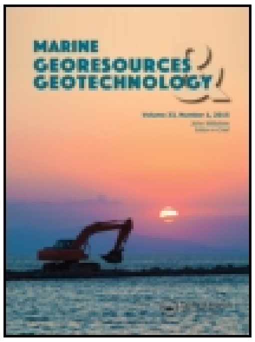

Marine Georesources & Geotechnology

ISSN: 1064-119X (Print) 1521-0618 (Online) Journal homepage: http://www.tandfonline.com/loi/umgt20

New Expressions-Based Models to Estimate Scour Depth at Clear Water Conditions in Rectangular Channels

Mohammad Najafzadeh, Jalal Shiri & Mohammad Rezaie Balf

New Expressions-Based Models to Estimate Scour Depth at Clear Water Conditions in Rectangular Channels

Mohammad Najafzadeh

Assistant Professor, Department of Civil Engineering, Graduate University of Advanced Technology-Kerman, Kerman, Iran

Jalal Shiri

Assistant Professor, Department of Water Engineering, University of Tabriz, Tabriz, Iran

Mohammad Rezaie Balf

Department of Civil Engineering, Graduate University of Advanced Technology-Kerman, Kerman, Iran

Address correspondence to Mohammad Najafzadeh, Assistant Professor, Department of Civil Engineering, Graduate University of Advanced Technology-Kerman, P.O. Box 76315-116, Kerman, Iran. Email: moha.najafzadeh@gmail.com

Abstract

Scouring in the channel contractions occurs due to the flow concentration within them inducing excessive bed shear stress. This is a complex process, so it is difficult to describe it through a general empirical model, the present research work describes contemporary conceptual relationships to estimate the local scour depth under equilibrium and clear water conditions in rectangular channels. Incidentally, Gene-Expression Programming (GEP), Evolutionary Polynomial Regression (EPR) and Model Tree (MT) based formulations were utilized to predict the scour depth at long contractions. The input variables comprising average flow velocity, critical threshold velocity of sediment movement, flow depth, median particle diameter, geometric standard deviation, and un-contracted and contracted channel widths, were used to feed the applied models. The performances of the developed were then compared with those calculated using existing scour prediction equations. The results showed that the developed MT approach in terms of linear relationship could predict the scour depth more precisely than GEP, EPR and the traditional equations. What is more, dimensionless parameter of h_1/b_1 (ratio of upstream flow depth to un-contracted channel width) was determined as the most influential variable in predicting the scour depth in long contractions.

Keywords: Evolutionary polynomial regression, Gene-Expression programming, Long contraction, Model tree, Scouring process

INTRODUCTION

Scouring phenomena occurs when cross-section of a waterway or a channel is declined due to the existence of natural features or man-made structures. In fact, the reduction of cross sectional area leads to significant rise in flow velocity due to the flow conservation. This can cause to provide the conditions of more sediment transport than entering the area and, consequently it is the leading factor of decreasing the bed sediment level in the contracted cross-section (Dey and Raikar 2004). This process is rudimentary introduced as “contraction scour.” Contraction of flow is appeared where bridges, weirs, and barrages, etc are performed. The channel contraction might be designed as long and short contractions. In fact, definition of the type of contractions is contingent upon the length (L) and width (b_1) of contracted zone. According to Komura (1966), a contracted cross section would be considered as long if $L/b_1 > 1$

, whereas Webby (1984) proposed it as $L / b_1 > 2$. Furthermore, Raikar (2004) suggests $L / b_1 \geq 1$ as long contraction criteria. The flow velocity in the contracted cross sections increases due to flow area reduction and bed shear stress increasing. Hence, bed sediment in contracted zone is exposed to the scour process.

Scouring phenomena is classified into two different groups, namely clear-water and live-bed processes (Dey and Raikar 2005). In clear-water condition, sediment particles are removed from the scour hole but not supplied by the approaching flow field, whereas live-bed scour is observed when there is a remarkable sediment transport process caused by the approaching flow fields. Scouring in a contracted zone is occasionally investigated regarding a configuration of rectangular long contraction, as sketched schematically in **Figure 1**. From **Figure 1**, owing to the straightforward geometrical feathers of the problem, a wide range of analytical researches to acquire the maximum scour depth with a permissible degree of accuracy in long contractions have triumphantly been conducted.

Review of literature about scouring process in long contractions shows that a lot of laboratory works have been conducted to characterize effective variables on the scouring phenomena in channel contraction (Straub 1934; Komura 1966; Gill 1981; Webby 1984; Lim 1993; Lim and Cheng 1998; Raikar 2004; Dey and Raikar 2005; Ghazvinei et al. 2016). Although extractions of these experimental studies were documented as empirical equations including effective parameters, these formulations cannot be taken into account as general approaches bringing an adequate prediction of the scour depth with a permissible precision. This is crucially the mercy of highly complex scour process.

On the verge of evolutionary computing and intelligent models, different heuristic approaches e.g. artificial neural networks (ANNs), adaptive neuro-fuzzy inference system

(ANFIS), support vector machines (SVM), gene-expression programming (GEP), group method of data handling (GMDH), model tree (MT), and evolutionary polynomial regression (EPR) have been prosperously employed to simulate the scour around hydraulic structures (e.g. Guven and Gunal 2008a; Ayoubloo, Etemad-Shahidi, and Mahjoobi 2010; Azamathulla et al. 2010; Etemad-Shahidi and Ghaemi 2011; Najafzadeh, Barani, and Azamathulla 2013; Najafzadeh, Barani, and Hessami-Kermani 2014). In case of the scour phenomena in long contractions, Najafzadeh, Etemad-Shahidi, and Lim (2016) employed the SVM and ANFIS methods for predicting the maximum scour depth at long contractions. They found that the proposed models showed more efficient performances compared with the empirical equations.

In this way, to achieve general equations simplifying evaluation of the scour depth in long contraction, the GEP, MT, and EPR models are developed based on experimental datasets. Moreover, efficiency equations extracted by the proposed models are compared with several empirical equations based regressive methods.

EXISTING EQUATIONS FOR PREDICTING THE SCOUR DEPTH IN LONG CONTRACTIONS

Conventional regression-based approaches involve the observations of the scour depths at existing bridge sites, sharp bends, and contractions or, other locations where the natural channel configuration is comparable to the expected configuration at the bridges site. Since 1934, experimental studies in this area have been initially introduced by Straub (1934). He investigated the sediment transportation process for prediction of scour depth in contracted and un-contracted widths of rectangular channels. Eventually, he put forward an empirical equation for evaluation of the scour depth based on the Manning and Du Boy equations:

$$\frac{d_s}{h_1}+1=(\frac{b_1}{b_2})^{6/7}.\left[\frac{\tau_c}{2\tau_1}+\sqrt{\left(\frac{\tau_c}{2\tau_1}\right)^2+\left(1-\frac{\tau_c}{\tau_1}\right)\left(\frac{b_1}{b_2}\right)}\right]^{-3/7} \quad (1)$$

where d_s, h, b, and τ denote the scour depth, flow depth, channel width, and bed shear stress, respectively. Additionally, the subscripts 1 and 2 indicate the un-contracted and contracted cross sections, respectively; and subscript c stands for the critical condition for incipient motion of sediment particles.

Laursen (1963) proposed following formulation:

$$\frac{d_s}{h_1}+1=(\frac{Q_2}{Q_1})^{6/7}(\frac{b_1}{b_2})^{\alpha}(\frac{n_2}{n_1})^{\alpha_1} \quad (2)$$

in which Q =flow rate; n =Manning coefficient; α =0.59-0.69; and α_1 =0.066-0.367.

Laursen (1963), assumed that the effects of the flow rate over the flood plains is negligible in a compound channel, so, identified that $Q_1=Q_2$ and $n_1=n_2$

$$\frac{d_s}{h_1}+1=(\frac{b_2}{b_1})^{-0.75} \quad (3)$$

Later, Komura (1966) conducted experimental studies to appraise the impacts of sediment particles grading on the scouring process in long contractions, so, proposed an equation based on properties of upstream flow conditions, bed sediment size, and channel geometry as,

$$\frac{d_s}{h_1}+1=1.6Fr_C^{0.2}(\sigma_g)^{-0.5}(\frac{b_2}{b_1})^{-0.67} \quad (4)$$

where, d_{50} = median particle diameter, Fr_C = Froude number due to the approaching flow velocity to initiate sediment particles motion within un-contracted cross section of channel, and σ_g = geometric standard deviation. What is more, the Fr_C parameter is computed as,

$$Fr_C = \frac{U_C}{\sqrt{g.h_1}} \quad (5)$$

where U_C and g are the critical velocity of sediments and gravitational acceleration, respectively.

Gill (1981) presented a formula by assuming that the sediment rate is proportional to the bed and critical shear stress:

$$\frac{d_s}{h_1} + 1 = 1.58(\frac{b_2}{b_1})^{-0.857} \quad (6)$$

Lim (1993) proposed an empirical equation of equilibrium scour depth in long contractions for both the clear-water and live-bed scour conditions, as:

$$\frac{d_s}{h_1} + 1 = 1.545 Fr_0^{0.75} (\frac{d_{50}}{h_1})^{0.25} (\frac{b_2}{b_1})^{-0.75} \quad (7)$$

where Fr_0 is the densimetric flow Froude number defined as,

$$Fr_0 = \frac{U_1}{\sqrt{g.\left(\frac{\rho_S}{\rho_w} - 1\right) d_{50}}} \quad (8)$$

in which U_1, ρ_S, and ρ_w are the average flow velocity, and mass density of sediments and water, respectively.

Lim and Cheng (1998) proposed a simple analytical-based equation for both live-bed and clear water conditions regarding the continuity condition between flow and sediment transport. They have professed that the maximum scour depth is only a function of b_2/b_1. Li (2002) carried out laboratory experiments to detect the effective variables on the maximum scour depth in long contractions with cohesive bed sediments. The results showed that the Froude number is the most influential parameter on the scouring, while the geometric properties of the contracted zone, e.g. length and width presented less effect in this regard. Overall, Dey and Raikar (2005) planned comprehensively laboratory investigations of the scour depth in long contractions. From their experiments, it was found that the scour depth gradually plummets with an increase in the densimetric Froude number for greater values of b_2/b_1.

Moreover, a new empirical equation for the prediction of the maximum scour depth under clear-water condition was proposed as (Dey and Raikar 2005),

$$\frac{d_s}{h_1} = 0.368 Fr_{1ec}^{\;0.55} \left(\frac{d_{50}}{h_1}\right)^{-0.19} \left(\frac{b_2}{b_1}\right)^{-1.26} \qquad (9)$$

in which, Fr_{1ec} = excess approaching flow Froude number:

$$Fr_{1ec} = \frac{U_{1ec}}{\sqrt{g.\left(\frac{\rho_s}{\rho_w} - 1\right)h_1}} \qquad (10)$$

where $U_{1ec}\left(U_1\big|_{U_1=U_c} - U_1\big|_{\substack{d_s=0\\ U_2=U_c}}\right)$ is the excess approaching flow velocity.

DEFINITION OF INPUT-OUTPUT VARIABLES

In experimental or field studies of the scour depth in long contractions, the chief parameters affecting the process are the characteristics of bed sediments, approaching flow conditions, and geometry of contracted and un-contracted cross sections (Straub 1934; Komura 1966; Gill 1981; Webby 1984; Lim 1993; Lim and Cheng 1998; Li 2002; Raikar 2004; Dey and Raikar 2005; Oh 2009). So, the functional relationship between the scour depth and effective parameters can be expressed as follows:

$$d_s = f(d_{50}, U_1, U_c, \nu, g, \sigma_g, h_1, b_1, b_2, \rho_w, \rho_s) \quad (11)$$

where ν is the kinematic viscosity of water. Recent studies show that the artificial intelligence models feeding with non-dimensional parameters present more accurate results than those obtained using dimensional parameters, so, dimensional analysis is carried out using the Buckingham π theorem for reducing the influential parameters (e.g., Azmathullah, Deo, and Deolalikar 2005; Guven, Azamathulla, and Zakaria 2009; Azamathulla et al. 2010; Etemad-Shahidi, Yasa, and Kazeminezhad 2011; Najafzadeh, Barani, and Hessami-Kermani 2014). So, the scour depth, d_s, was normalized using b_1 as follows:

$$d_s/b_1 = f(d_{50}/b_1, h_1/b_1, b_2/b_1, \sigma_g, U_1/U_c, U_1/\sqrt{((\rho_s/\rho_w)-1).g.d_{50}}) \quad (12)$$

Where $U_1/\sqrt{((\rho_s/\rho_w)-1).g.d_{50}}$ is the densimetric Froude number (Fr_0).

The non-dimensional parameters of Eq. (12) were used as input and output parameters for feeding the applied models.

DATASETS DESCRIPTION

204 data patterns collected from the Komura (1966), Gill (1981), Webby (1984), Lim (1993), Dey and Raikar (2005), and Lim (2013) (unpublished data sets) were used for establishing and testing the applied models. The used data set purely belong to the clear-water circumstances $(U_1/U_c \leq 1)$ and rectangular channels with the presence of long contractions. Out of the total data set, roughly 67% and 33% were selected randomly to perform training and testing stages, respectively. The ranges of parameters used for the scour depth modeling are presented in **Table 1**.

MODEL DESCRIPTIONS

Development of the GEP Model

Gene expression programming (GEP) is known as extension of the genetic programming (GP) approach, that evolves computer programs in forms of mathematical expressions, decision trees, and logical expressions (Ferreria 2001; Ferreria 2006; Azamathulla 2012; Azamathulla and Haque 2012). In addition, the GEP model has attracted the ferociously great attention of designers in hydraulic issues. This study represents GEP-based formulation of maximum scour depth predicting in long contractions. The GEP approach is coded in forms of linear chromosomes, which are then expressed into Expression Trees (ETs).

In fact, the ETs are sophisticated computer programming which are usually evolved to solve a practical problem, and are selected accordingly to their fitness at solving that problem (Ferreria 2006).

Development of the GEP approach includes five steps (Ferreria 2006):

In the first instance, the fitness function (f_i), of an individual program (*i*), is selected. This item is evaluated as follows:

$$f_i = \sum_{j=1}^{C_t} (M - |C_{(i,j)} - T_j|) \tag{13}$$

in which M, $C_{(i,j)}$, and T_j are the selection range, value returned by the individual chromosome *i* for fitness case *j*, the largest value for fitness case j.

In the second step, the set of terminals *T* and the set of function *F* were selected to generate the chromosomes. In this study, the terminal set includes six independent parameters in form of

$$T(d_s/b_1) = \{d_{50}/b_1, h_1/b_1, b_2/b_1, \sigma_g, U_1/U_c, Fr_0\} \tag{14}$$

In this way, it is necessary to peer review previous investigations of scour problems at long contractions. Additionally, four basic operators (+,-,*,/) and basic mathematical functions (√, power, sin, cos, exp) were applied to predict the scour depth in long contractions. In fact, selection of basic operators and mathematical function are at the mercy of basic form of empirical equations. The third step is to configure the chromosomal architecture. In the further step, liking function is selected. Eventually, for the fifth stage, the sets of genetic operators which case variation and their rate are selected. Further details about the GEP model development might be found in e.g. Ferreria (2006).

Furthermore, the functional set and the operational parameters applied in the proposed GEP models are presented in **Table 2**. The best formulation of GEP model for evaluation of equilibrium scour depth at long contractions, as a function of input parameters, is acquired as,

$$\frac{d_S}{b_1}=\left(\left(\frac{Fr_0}{e^{\left(Fr_0+\left(\frac{h_1}{b_1}\right)\right)}+\left(-0.61*\left(\sigma_g/\left(d_{50}/b_1\right)\right)\right)}\right)+\left(\frac{\sigma_g}{\left(\left(\left(\frac{h_1}{b_1}\right)*\left(\frac{d_{50}}{b_1}\right)\right)*\left(\sigma_g+8.24\right)\right)-\left(\left(\frac{h_1}{b_1}\right)+\left(\frac{b_2}{b_1}\right)-5.56\right)}\right)+\left(\frac{\left(\frac{b_2}{b_1}\right)}{\left(\frac{-492.74*\sigma_g}{Fr_0-5.34}\right)*\left(\left(\left(\frac{d_{50}}{b_1}\right)+0.173\right)-\left(\frac{h_1}{b_1}\right)\right)}\right)\right) \tag{15}$$

In addition, the expression tree of the above formulation was illustrated in **Figure 2**. From **Figure 2**, it is conceivable that each tree structure introduces a gene and consequently, each gene was composed of mathematical operators, input variables, as well as constant values. Overall, summation of expression trees yields the most recent equation [Eq.(15)] and incidentally includes three genes.

Development of the MT Model

Among the data mining techniques, model trees generalize the concept of the classification and regression tree. They are used to solve the problem by dividing it into several sub-problems (sub-domains) tasks and the result is a combination of these sub-problems (Etemad-Shahidi, Yasa, and Kazeminezhad 2011). Classification trees classify data records by sorting them down the tree from the root node to some leaf nodes (Quinlan 1992; Wang and Witten 1997; Etemad-shahidi and Ghaemi 2011). In this way, MT models can be applied to solve continuous class problems and yield a structural representation of the data sets using the piecewise linear models to approximate nonlinear relationships. As a classic example, the tree-building procedure within four linear regression models and knowledge extraction from the structure for corresponding sub-domains was illustrated in **Figure 3A**. Furthermore, a general tree structure of MT approach was sketched in **Figure 3B**. From **Figure 3**, input space has been separately divided into six segments. In each section, there are several points which a linear regression can be produced. In fact, all models were circumstantially generated. As a classic example, there are two input variables, as X_1 and X_2 seen in **Figure 3**, If X_2 becomes bigger than

2.5 and X_1 becomes smaller than 2, then model3 is made. Through MT approach, the basic tree is firstly generated using the splitting criterion of the standard deviation reduction (SDR) factor:

$$SDR = sd(E) - \sum_i \frac{|E_i|}{|E|} sd(E_i) \quad (16)$$

in which E, sd, and E_i are the set of examples (data records) that reach the node, the set that results from splitting the node according to the chosen attribute (parameter), and standard deviation, respectively. The M5 utilizes the sd parameter as an error measure of the class values that reach a node. Testing all parameters at a node, it calculates the expected reduction in error and then selects the parameter that maximizes SDR. This process stops when the standard deviation reduction becomes less than a certain percent of the standard deviation of the original dataset or when only a few data records remain (Quinlan 1992; Wang and Witten 1997). Then, a linear regression model is developed for each sub-domain. Merely the data in connection with the variables tested in that sub-domain are applied in the regression. Other descriptions of the MT model were presented in literature (e.g., Etemad-shahidi and Ghaemi 2011; Pal et al. 2012).

In this way, the proposed MT approach has six input and one output parameters.

MT technique was developed using 3 rules in form of linear equations. Meantime, schematic diagram of tree-building of the MT approach in form of rules for prediction of the scour depth at long contractions was illustrated in **Figure 4**.

These linear equations and their corresponding rules were given **Table 3**. As seen in **Table 3**, Eqs.(17)-(18) include two splitting parameters of b_2/b_1 and U/U_c. b_2/b_1 variable is the splitting parameter for Eq.(17) and its value was fixed 0.55. In addition, for Eq.(17), value of splitting is 0.724.

Development of the EPR Model

EPR is generally introduced as a non-linear stepwise regressive approach, producing mathematical formulas with persuasive level of accuracy in order to find either mathematical patterns or dependency between input and output of complex systems (Giustolisi and Savic 2006).

In term of mathematical function, Giustolisi and Savic (2006) proposed the general expressions obtained by EPR which are exhaustively included those of a number of additive terms multiplied by as many coefficients mentioned as,

$$\hat{\mathbf{Y}} = a_0 + \sum_{j=1}^{m} a_j \cdot (\mathbf{X}_1)^{\mathbf{ES}(j,1)} \cdot \ldots \cdot (\mathbf{X}_\mathrm{k})^{\mathbf{ES}(j,\mathrm{k})} \cdot f\left((\mathbf{X}_1)^{\mathbf{ES}(j,\mathrm{k}+1)} \cdot \ldots \cdot (\mathbf{X}_\mathrm{k})^{\mathbf{ES}(j,2\mathrm{k})}\right) \qquad (20)$$

where m, $\mathbf{X}_\mathrm{i}$, and $\hat{\mathbf{Y}}$ are the maximum number of additive terms, input variables, and final output of EPR, respectively. Additionally, both of function f and exponents of variables are opted by the user.

The genetic algorithm is employed to select the exponents $\mathbf{ES}(j,i)$ from among the values in set **EX**. This means that an integer coding of possible alternative exponents $\mathbf{ES}(j,i)$ is adopted to obtain non-linear expression. Moreover, It is worthwhile to consider which, if the set of exponents contains zero and $\mathbf{ES}(j,i) = 0$, the relevant input eliminates from the final mathematical function. Ergo, simple structure of Eq.(19) is able to benefit exceedingly from persuasive level of generalization to perceive either patterns or physical meaning of observed datasets.

An improvement of EPR model, as an ingredient approach, is implemented into linear equations with respect to coefficient (a_j) in a way that they are predicted making use of classical numerical regression such as least squares techniques. The search space for EPR-MOGA is fundamentally defined by user in terms of the basic structure of mathematical expressions, the maximum number of additive terms *(m)*, adjusting a set of exponents (**EX**) and number of candidate explanatory variables (i.e., *k*) (Laucelli and Giustolisi 2011).

Apparently, Laucelli and Giustolisi (2011) have designed the OPTImized Multi-Objective Genetic Algorithm as a model search which is on the basis of Pareto dominance criterion so as to carry out multi-objective optimization approach (Pareto 1896; Van Veldhuizen and Lamont 2000).

EPR-MOGA explores the space of *m*-term mathematical expressions with the different degree of complexity by taking into account three objectives. The first objective is that accuracy of proposed model is maximized and the minimization of the number for model coefficients is the second objective. Ultimately, the number of actually applied input variables for mathematical model should possibly be minimized (Laucelli and Giustolisi 2011).

In term of application of EPR, a functional relationship in Eq. (20) for predicting the scour depth in long contraction is generated. Therefore, $d_{50}/b_1, h_1/b_1, b_2/b_1, U_1/U_c$, Fr_0 and σ_g have deliberately been assigned as input parameters. Furthermore, the range of exponents **(EX**) used to develop an optimal expression is between -2 and 2 [-2; -1.5; -1; -0.5; 0; 0.5; 1; 1.5; 2]. Three polynomial terms (m = 3) is adjusted without considering a bias *(a_0=0)* and in addition to the integrating only positive coefficients (a_j > 0).

Among good many of models extracted by EPR-MOGA-XL, the following equation was choosen in order to meet an acceptable level of model accuracy:

$$d_s / b_1 = 0.14308 \frac{(h_1/b_1)^{0.5} . U / U_c}{(b_2/b_1)^{1.5}} + 0.00047141 \frac{(Fr_0)^{0.5} . \sigma_g}{(h_1/b_1)^{1.5}} \ln\left((b_2/b_1)^{1.5} . \sigma_g . (U/U_c)^2\right)$$
$$+ 0.26838 (b_2/b_1) . (\sigma_g)^{1.5} . (h_1/b_1)^2 \ln\left(\frac{(d_{50}/b_1)^{0.5} (U/U_c)^{1.5}}{(b_2/b_1)^{0.5}} \right) \quad (21)$$

PERFORMANCE EVALUATION CRITERIA

Correlation coefficient (R), root mean square error (RMSE), coefficient of determination (CoD), and discrepancy ratio (DR) were used to evaluate the performances of models:

$$R = \frac{\sum_{i=1}^{M} \left(d_s/b_{1i(Observed)} - \overline{d_s/b_1}_{(Observed)}\right)\left(d_s/b_{1i(\mathrm{Predicted})} - \overline{d_s/b_1}_{(\mathrm{Predicted})}\right)}{\sqrt{\sum_{i=1}^{M} \left(d_s/b_{1i(Observed)} - \overline{d_s/b_1}_{(Observed)}\right)^2 . \sum_{i=1}^{M} \left(d_s/b_{1i(\mathrm{Predicted})} - \overline{d_s/b_1}_{(\mathrm{Predicted})}\right)^2}} \quad (22)$$

$$RMSE = \left[\frac{\sum_{i=1}^{M} \left(d_s/b_{1i(\mathrm{Predicted})} - d_s/b_{1i(Observed)}\right)^2}{M} \right]^{1/2} \quad (23)$$

$$CoD = 1 - \frac{\sum_{i=1}^{M} \left(d_s/b_{1i(predicted)} - d_s/b_{1i(observed)}\right)^2}{\sum_{i=1}^{M} \left(d_s/b_{1i(observed)} - \overline{d_s/b_1}_{(observed)}\right)^2} \quad (24)$$

$$DR = \frac{1}{M} \sum_{i=1}^{M} \frac{d_s/b_{1i(\mathrm{Predicted})}}{d_s/b_{1i(Observed)}} \quad (25)$$

where M is the number of data and $\overline{d_s/b_1}$ is the average value of observations.

RESULTS AND DISCUSSIONS

In this section, outperformances of training and testing stages would be discussed. Conceivably, an argument in terms of evaluation of empirical equations are made.

MODELS IMPLEMENTATION

Error statistics computed by applied models are presented for the both training and testing stages in **Table 4**. Attending to the training results, in terms of R and CoD assessment, it is seen that MT model produces more accurate results (CoD=0.631 and R=0.797), compared with EPR (CoD=0.620 and R=0.788) and GEP (CoD=0.29 and R=0.77) approaches. Moreover, the equation given by MT model predicted the scour depth with RMSE of 0.036 and DR of 1.283 than those obtained using EPR (RMSE=0.0367 and DR=1.301) and GEP (RMSE=0.037 and DR=1.28) models. **Figure 5** illustrates the scatter plots of the observed scour depths versus predicted ones for the EPR, MT, and GEP models during the training stage. In terms of qualitative comparisons, all intelligent models indicated a pretty illustrative over-prediction for observed ds/b_1 ranged between 0 and 0.05. By the way, predicted values of ds/b_1 are out of +20% range. What is more, Eq.(15), produced by GEP, has the highest level of under-prediction being seen out of -20% line when compared with other approaches.

MODELS TESTING

In the testing stage, it can be noted that the EPR model predicted the scour depth with higher accuracy (CoD=0.026 and R= 0.903) than those obtained by MT (CoD=0.756 and R= 0.874) and GEP (CoD=0.8 and R= 0.89) techniques. Furthermore, RMSE and DR values fixed by the EPR model are 0.0263 and 1.13, for MT approach are 0.0296 and 1.25, and for GEP are 0.027 and 1.18, respectively. The scatter plots of the observed vs. predicted scour depths are presented in **Figure 6** for the test period. Through the testing stage, **Figure 6** provided the readers with the information about the obvious over-prediction for MT and EPR when observed ds/b_1 is between 0 and 0.05. In fact, several points of predicted ds/b_1 are out of upper limit of error band. Put it another way, it is crystal clear that non-dimensional scour depth prognosticated by MT, observed out of lower error band, had more comparatively under-estimation than EPR and GEP models.

In order to compare the results of the applied models with traditional ones, an argument was released. The empirical equations presented by Eq.(3) (Laursen 1963), Eq.(4) (Komura 1966), Eq.(6) (Gill, 1981) and Eq.(7) (Lim 1993) were used in this regard, to evaluate the scour depth under the equilibrium and clear water conditions in the studied rectangular channels. From **Table 4**, it is vividly evident that the Eq. (3) proposed by Laursen (1963), produced the lower error of scour depth evaluation with RMSE of 0.0543 and DR of 1.074 compared with the other conventional models. On the other hand, Eq.(3) including three input parameters, as Fr_C, σ_g, and b_2/b_1, prognosticated the scour depth with more permissible accuracy (RMSE=0.0833 and DR=1.678) than Eq.(6) (RMSE=0.2 and DR=3.307), Eq.(7) (RMSE=0.134 and DR=2.32). Overall, comparison of the proposed models with traditional models reveals that the intelligent approaches presenting more accurate results than the traditional relationships is occasionally encountered with crucial drawbacks corresponded to the datasets range. It is conceivable that different scales of experimental investigations affect the efficiency of the models. In contrast, one of the most reasonable assumptions to apply intelligent approaches in hydraulic engineering fields is that the influences of various scales on performing the predictive data-driven approaches and regression-based techniques are neglected. By the way, it would presumably plummet an acceptable level of precision. The scatter plots between the observed and predicted scour depths of the empirical equations are displayed in **Figure 7**. From **Figure 7**, it can be vividly professed that merely a good many of the scour depth predicted by Eq.(3), is out of upper error band. Conversely, ds/b_1 predicted by other intelligent models have provided an extremely aggressive increase in over-prediction, as points seen in out of upper error band line.

RELATIVE SIGNIFICANCE OF INPUT VARIABLES

To assign the comparative influence of each input variable on the scour depth, the EPR model was selected to perform a sensitivity analysis. The analysis was conducted such that, one parameter of Eq. (12) was eliminated each time to evaluate the impact of that input on output. Results of the analysis demonstrated that h_1/b_1 is the most effective parameter on the maximum scour depth with R of 0.385 and RMSE of 0.0557. Furthermore, with regarding the other statistical parameters, EPR has produced a considerably large error (CoD = -5.248 and DR = 1.667) by neglecting $h_1/b1$as a input parameter. Conversely, b_2/b_1 has the lowest level of impacts on the d_s/b_1 due to the fact that coefficient correlation (R) and root mean square error (RMSE) yielded 0.524 and 0.0513, respectively. What is more, CoD (-1.289) and DR (1.458), values are an indicative of this trend. The other influential parameters on the d_s/b_1 include d_{50}/b_1, σ_g , U_1/U_c and Fr_0 which are ranked from higher impacts to lower ones, respectively. The error statistics yielded using the sensitivity analysis are given in **Table 5**.

CONCLUSION

In the present research, the EPR, MT, and GEP approaches were developed to evaluate scour depth at equilibrium and clear water conditions in rectangular channels. Performances of the proposed techniques for training and testing stages were carried out using experimental datasets collecting from literature. Beside, empirical equations, as proposed by Laursen, Komura, Gill and Lim were utilized to compare results with the proposed models. Regarding the EPR, MT, and GEP, to obtain the optimum functions on the basis of the best formulations, a dimensional analysis was used to extract parameters affecting the scour process at long contractions under clear water conditions.

From the statistical error parameters, it can be concluded that MT method with 3 rules in form of 3 linear equations for training stage captured the scour depth at long contractions with more efficient performance compared with EPR and GEP approaches. In addition, performance of the testing stages demonstrated that EPR method predicted the scour depth with more precise estimation in terms of RMSE (0.026) and CoD (0.903) than MT and GEP techniques. Apparently, performances of empirical equations illustrated that Eq. (3), proposed by Laursen (1963), provided relatively lower error of scour depth predictions in terms of RMSE and CoD compared with the other traditional models. Accordingly, it was generally professed that evaluation of the scour depth at contracted cross-section making use of empirical equations, on the basis of CoD and R parameters, indicated remarkably higher level of error than the proposed techniques. Conspicuously, the quantitative results of the sensitivity analysis indicated that h_1/b_1 is the most important parameter in the modeling of d_s/b_1 by the EPR model.

To come up with a conclusion, the trend is towards making a wise decision whether taking head of pros and cons of applying the proposed approaches had been successfully persuasive or not. It is about time equations extracted by intelligent models were recruited in the right circumstances in a way that misunderstanding of outperformances are not inevitably emerged even when field datasets of scour depth in rectangular channels with long contracted zones have been applied to validate empirical equations. By the way, applying the proposed intelligent techniques in form of the best formulations has a prominent role to play in the achieving astonishing and outstanding successes for practical uses. Another plus side of this research is that explicit mathematical expressions extracted in the current research work can be taken into account as reasonably valuable approaches for specialists who intend to design rectangular channels being exposed to the scouring phenomena. From a logical viewpoint, to gain a permissible precision of the scour depth, ranges of variables, conditions of upstream flow and bed sediment embedded in contracted zones were strongly advised to be considered. Even though, empirical equations are extracted in the controlled conditions and consequently suffer relatively from covering limited ranges of experimental datasets but conventional models can be considered as a straightforward pathway to predict the scour depth at long contractions as well as proposed intelligent approaches.

References

Ayoubloo, M. K., A. Etemad-Shahidi, and J. Mahjoobi. 2010. Evaluation of regular wave scour around a circular pile using data mining approaches. *Applied Ocean Research* 32 (1), 34–39. doi:10.1016/j.apor.2010.05.003

Azamathulla, H. M. 2012. Gene expression programming for prediction of scour depth downstream of sills. *Journal of Hydrology* 460–461, 156–59. doi:10.1016/j.jhydrol.2012.06.034

Azamathulla, H. M., A. Ab Ghani, N. A. Zakaria, and A. Guven. 2010. Genetic programming to predict bridge pier scour. *Journal of Hydraulic Engineering, ASCE* 136 (3), 165–69.

Azmathullah, H. M., M. C. Deo, and P. B. Deolalikar. 2005. Neural networks for estimation of scour downstream of a ski-jump bucket. *Journal of Hydraulic Engineering, ASCE* 131 (10), 898–908. doi:10.1061/(asce)0733-9429(2005)131:10(898)

Azamathulla, H. M., and A. A. M. Haque. 2012. Prediction of scour depth at culvert outlets using gene-expression programming. *International Journal of Innovative Computing, Information and Control* 8 (7B), 5045–54.

Dey, S., and R. V. Raikar. 2005. Scour in long contractions. *Journal of Hydraulic Engineering, ASCE* 131 (12), 1036–49. doi:10.1061/(asce)0733-9429(2005)131:12(1036)

Etemad-Shahidi, A., and N. Ghaemi. 2011. Model tree approach for prediction of pile groups scour due to waves. *Ocean Engineering.* 38, 1522–27. doi:10.1016/j.oceaneng.2011.07.012

Etemad-Shahidi, A., R. Yasa, and M. H. Kazeminezhad. 2011. Prediction of wave-induced scour depth under submarine pipelines using machine learning approach. *Applied Ocean Research* 33, 54–59. doi:10.1016/j.apor.2010.11.002

Ferreria, C. 2001. Gene-expression programming. A new adaptive algorithm for solving problems. *Complex Systems* 13 (2), 87–129.

Ferreria, C. 2006. *Gene-expression programming Mathematical Modeling by an Artificial Intelligence*. Berling, Heidelberg, New York: Spriner, 21.

Ghazvinei, P. T., A. Junaidah, A. M. Thamer, A. Amini, M. M. Aghajani, S. Saheri, and S. Ansarimoghaddam. 2016. Contraction scour analysis at protruding bridge abutments. *Proceedings of the Institution of Civil Engineers - Bridge Engineering*, 169 (1), 39–53. doi:10.1680/bren.14.00011

Gill, M. A. 1981. Bed erosion in rectangular long contraction. *Journal of the Hydraulics Division ASCE* 107 (3), 273–84.

Giustolisi, O., and D. A. Savic. 2006. A symbolic data-driven technique based on evolutionary polynomial regression. *Journal of Hydroinformatics*, 8, 207–22.

Giustolisi, O., and D. A. Savic. 2009. Advances in data-driven analyses and modelling using EPR-MOGA. *Journal of Hydroinformatics*, 11, 225–36. doi:10.2166/hydro.2009.017

Guven, A., H. M. Azamathulla, and N. A. Zakaria. 2009. Linear genetic programming for prediction of circular pile scour. *Ocean Engineering* 36 (12–13), 985–991.

Guven, A., and M. Gunal. 2008a. Genetic programming approach for prediction of local scour downstream hydraulic structures. *Journal of Irrigation and Drainage Engineering*, 134 (2), 241–49. doi:10.1061/(asce)0733-9437(2008)134:2(241)

Guven, A., and M. Gunal. 2008b. Prediction of scour downstream of grade-control structures using neural networks. *Journal of Hydraulic Engineering, ASCE* 134 (11), 1656–60. doi:10.1061/(asce)0733-9429(2008)134:11(1656)

Haykin, S. 1999. *Neural networks: A comprehensive foundation*. 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall Inc.

Komura, S. 1966. Equilibrium depth of scour in long constrictions. *Journal of the Hydraulics Division ASCE* 92 (5), 17–38.

Laucelli, D., L. Berardi, A. Doglioni, and O. Giustolisi. 2012. EPR-MOGA-XL: An excel based paradigm to enhance transfer of research achievements on data-driven modeling, Proceedings of 10th International Conference on Hydroinformatics HIC 2012, 14-18 July, Hamburg, Germany, R. Hinkelmann, M.H. Nasermoaddeli, S.Y. Li-ong, D. Savic, P. Fröhle (Eds).

Laucelli, D., and O. Giustolisi. 2011. Scour depth modelling by a multi-objective evolutionary paradigm. *Environmental Modeling & Software* 26 (4), 498–509. doi:10.1016/j.envsoft.2010.10.013

Laursen, E. M. 1963. An analysis of relief bridge scour. *Journal of the Hydraulics Division ASCE* 89 (3), 93–118.

Li, Y. 2002. Bridge pier scour and contraction scour in cohesive soils on the basis of flume tests. Ph. D. Dissertation, Texas A&M University, College Station, TX, USA.

Lim, S. Y. 1993. Clear water scour in long contractions. *Proceeding of the Institution Civil Engineerings-Waters, Maritime and Energy*, London, UK 101 (6), 93–98.

Lim, S. Y., and Cheng, N. S. 1998. Scouring in long contractions. *Journal of Irrigation and Drainage Engineering ASCE* 124 (5), 258–61. doi:10.1061/(asce)0733-9437(1998)124:5(258)

Najafzadeh, M., G. A. Barani, and H. M. Azamathulla. 2013. GMDH to prediction of scour depth around vertical piers in cohesive soils. *Applied Ocean Research* 40, 35–41. doi:10.1016/j.apor.2012.12.004

Najafzadeh, M., G. A. Barani, and M. R. Hessami-Kermani. 2014. Estimation of pipeline scour due to waves by the group method of data handling. *Journal of Pipeline Systems Engineering and Practice*, *ASCE* 5 (3), 06014002. doi:10.1061/(asce)ps.1949-1204.0000171

Najafzadeh, M., A. Etemad-Shahidi, and S.-Y. Lim. 2016. Prediction of scour depth in long contractions using ANFIS and SVM. *Ocean Engineering, Elsevier* 111, 128–35.

Pareto, V. 1896. e. Rouge and Cic, Lausanne, Switzerland.

Quinlan, J. R. 1992. Learning with continuous classes. In *Proceedings of AI'92. World scientific*, ed. Adams, Sterling, 343–48.

Raikar, R. V. 2004. Local and general scour of gravel beds. PhD thesis, Dept. of Civil Engineering, Indian Institute of Technology, Kharagpur, India.

Straub, L. G. 1934. Effect of channel contraction works upon regimen of movable bed streams. *Transactions, American Geophysical Union* 2, 454–463. doi:10.1029/tr015i002p00454

Wang, Y., and I. H. Witten. 1997. Induction of model trees for predicting continuous lasses. In *Proceedings of the poster papers of the European conference on machine learning*. Prague: University of Economics, Faculty of Informatics and Statistics.

Webby, M. G. 1984. General scour at contraction. RRU Bulletin 73, National Roads Board, Bridge Design and Research Seminar, New Zealand, 109-118.

Table 1. Ranges of grouped input and output parameters for scour depth modeling

Parameters	Training	Testing
d_{50}/b_1	0.000875–0.0237	0.000875–0.2375
h_1/b_1	0.0509–0.442	0.0360–0.32
b_2/b_1	0.25–4.062	0.33–4.062
Fr_0	0.878–3.504	1.08–3.28
U_1/U_c	0.321–1	0.431–1
σ_g	1.065–3.61	1.065–3.61
ρ_s/ρ_w	2.65	2.65
d_s/b_1	0.00131–0.29	0.0118–0.256

Table 2. Parameters of the optimized GEP model

Parameters	Description of parameters	Setting of parameters
P_1	Function set	+,–,×,/,exp, power
P_2	Mutation rate	0.138
P_3	Inversion rate	0.546
P_4	One point and two point recombination rate respectively (%)	0.277
P_5	Gene recombination rate	0.277
P_6	Gene transportation rate	0.277
P_7	Maximum tree depth	6
P_8	Number of Gene	3
P_9	Number of Chromosomes	30

Table 3. General Feature of the proposed MT approach

Rules of MT approach	Linear Models(LM)	Eq.no
Rule:1 IF b_2/b_1 > 0.55 THEN LM1	$LM1: \quad d_s/b_1 = -0.0908 - 0.0031(b_2/b_1) + 0.0785Fr_0 - 0.0157\sigma_g - 0.0366(U/U_c) + 4.733(d_{50}/b_1)$	(17)
Rule:2 IF U/Uc >0.724 THEN LM2	$LM2: \quad d_s/b_1 = 0.1374 + 0.6584(h_1/b_1) + 0.0633(U/U_c) - 0.3218(b_2/b_1)$	(18)
Rule: 3	$LM3: \quad d_s/b_1 = 0.0937$	(19)

Table 4. Statistical parameters of the training and testing data set for difference models

Training Stage	R	CoD	RMSE	DR
GEP	0.77	0.29	0.037	1.28
MT	0.797	0.631	0.0362	1.283
EPR	0.788	0.621	0.0367	1.301
Testing Stage				
	R	CoD	RMSE	DR
GEP	0.89	0.8	0.027	1.18
MT	0.874	0.756	0.0296	1.25
EPR	0.903	0.0263	0.0263	1.13
Eq.(3) (Laursen, 1963)	0.685	0.646	0.0543	1.074
Eq.(4) (Komura, 1966)	0.75	0.633	0.0833	1.678
Eq.(6) (Gill, 1981)	0.693	0.138	0.200	3.307
Eq.(7) (Lim, 1993)	0.720	0.443	0.134	2.32

Table 5. Sensitivity analysis for independent parameters

Model	Input parameters	R	CoD	RMSE	DR
	$d_s/b_1 = f(h_1/b_1, b_2/b_1, \sigma_g, Fr_0, U_1/U_c)$	0.442	–4.626	0.0548	1.662
h_1/b	$d_s/b_1 = f(d_{50}/b_1, b_2/b_1, \sigma_g, Fr_0, U_1/U_c)$	0.385	–5.248	0.0557	1.667
b_2/b_1	$d_s/b_1 = f(d_{50}/b_1, h_1/b_1, \sigma_g, Fr_0, U_1/U_c)$	0.524	–1.289	0.0513	1.458
	$d_s/b_1 = f(d_{50}/b_1, h_1/b_1, b_2/b_1, Fr_0, U_1/U_c)$	0.461	–2.576	0.0534	1.509
	$d_s/b_1 = f(d_{50}/b_1, h_1/b_1, b_2/b_1, \sigma_g, U_1/U_c)$	0.485	–2.682	0.0528	1.523
	$d_s/b_1 = f(d_{50}/b_1, h_1/b_1, b_2/b_1, \sigma_g, Fr_0)$	0.479	–2.215	0.0529	1.530

Figure 1. Schematic of a long rectangular channel contraction at equilibrium scour condition: (a) top view; and (b) side view.

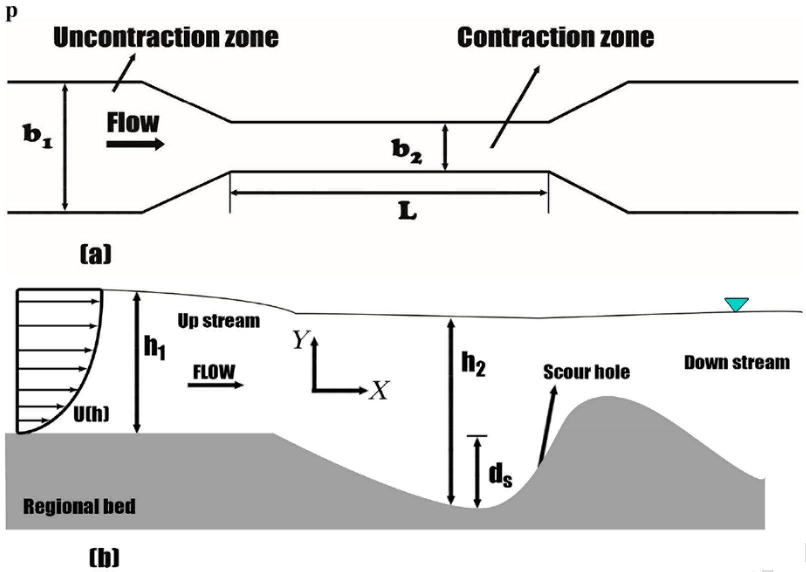

Figure 2. Optimal expression tree (ET) structures for the GEP model for prediction of the scour depth at long contractions with the actual input variables are the d0=d_{50}/b_1, d1= b_2/b_1, d2= Fr_0, d3= h_1/b_1, d4=σ_g, and d5=U/Uc; (A) C9=5.77, C4=-6.38; (B) C1=-8.24, C5=-5.56; (C) C8= -492.74, C9= -5.34, C4= 0.17.

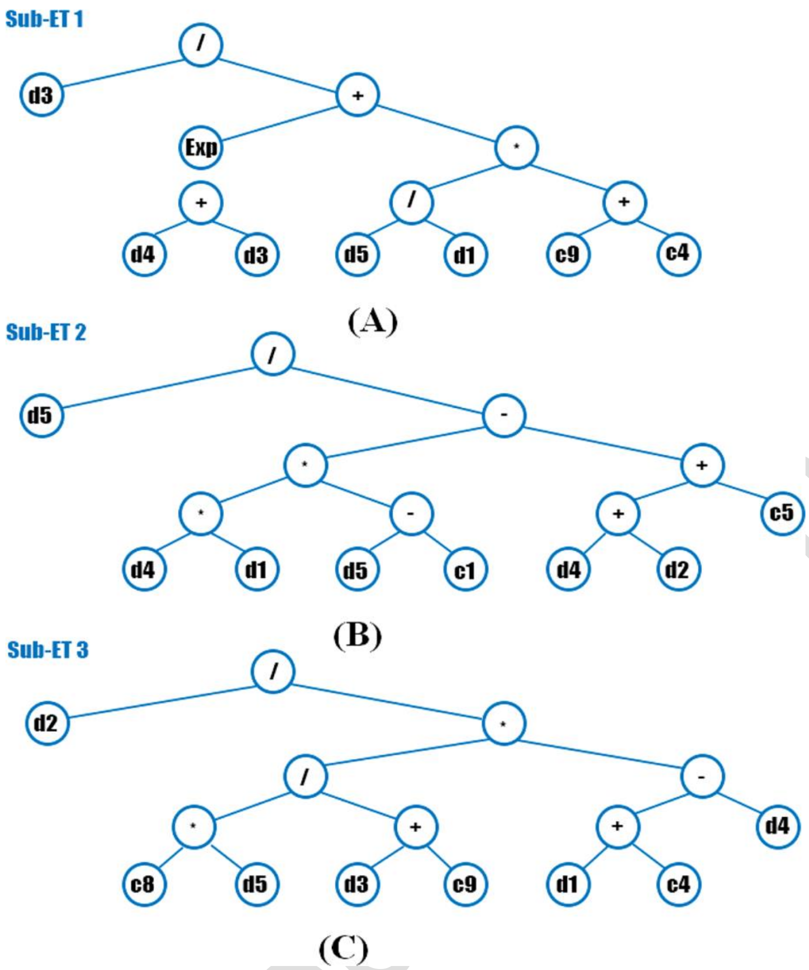

Figure 3. Splitting the input space and prediction by the model tree for a new data record: (A) splitting of the input space ($X_1 \times X_2$) by the M5 model tree algorithm; (B) predicting a new data record by the model tree.

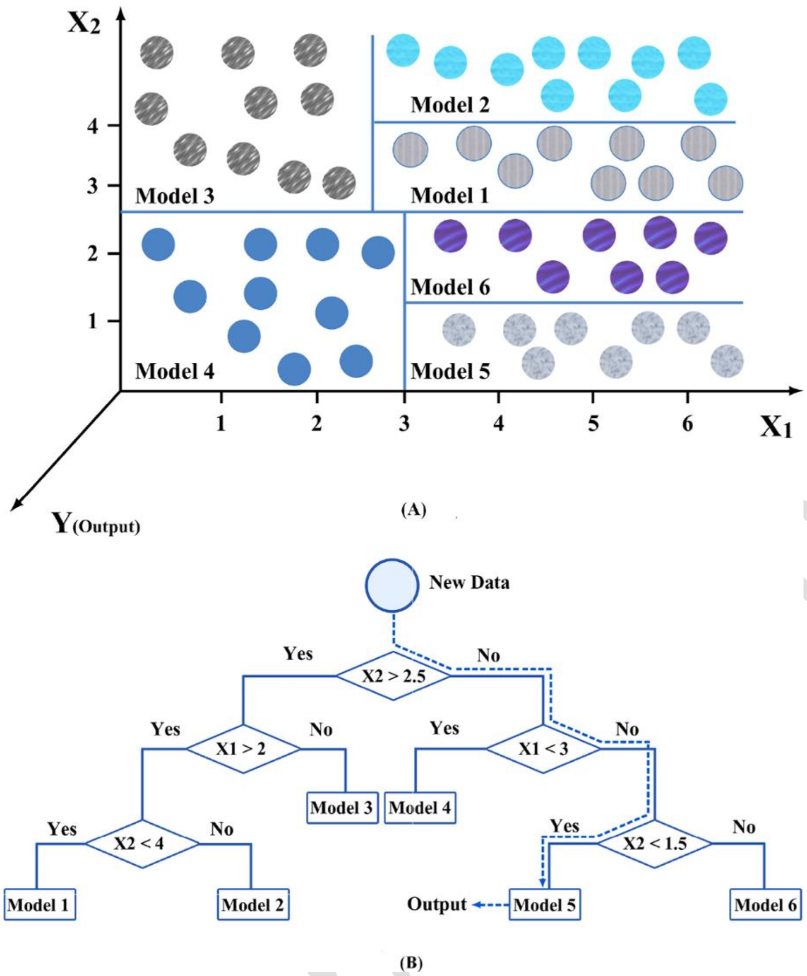

Figure 4. Proposed MT structure for prediction of the local scour depth at long contractions.

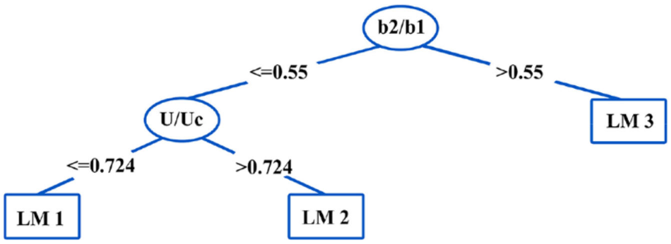

Figure 5. Scatter plot of observed and predicted scour depth at equilibrium and clear water conditions for training of the proposed models.

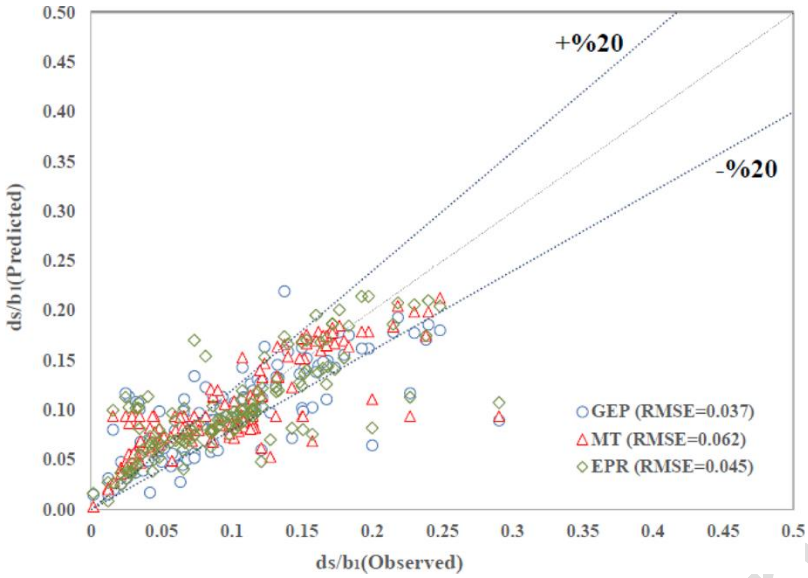

Figure 6. Scatter plot of observed and predicted scour depth at equilibrium and clear water conditions for testing of the proposed models.

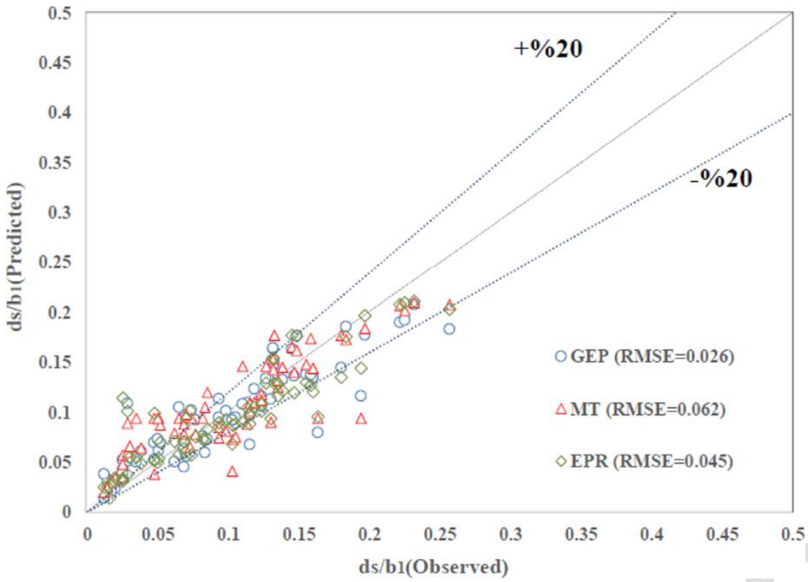

Figure 7. Scatter plot of observed and predicted scour depth at equilibrium and clear water conditions for the traditional equations.