Contents lists available at **SCCE**

Journal of Soft Computing in Civil Engineering

Application of GEP, M5-TREE, ANFIS, and MARS for Predicting Scour Depth in Live Bed Conditions around Bridge Piers

Shalini[1*] iD **, Thendiyath Roshni**[2]

1. Research Scholar, Department of Civil Engineering, National Institute of Technology Patna, Patna, 800005, India
2. Associate Professor, Department of Civil Engineering, National Institute of Technology Patna, Patna, 800005, India
Corresponding author: *shalini.ce17@nitp.ac.in*

doi *https://doi.org/10.22115/SCCE.2023.369213.1559*

ARTICLE INFO

ABSTRACT

This paper presents the use of data-driven models, namely Gene expression programming (GEP), M5 model tree (M5-TREE), Multivariate adaptive regression spline (MARS), and Adaptive neuro-fuzzy inference system (ANFIS) to predict bridge pier scour depth. Only 213 data sets of the live bed conditions from laboratory tests and field data measurements were considered for the present analysis. The gamma test has been performed to determine the ideal input combinations for model development. Five main non-dimensional parameters: Sediment Coarseness ratios, Froude number, flow intensity, gradation coefficient of the bed material, and shape factor, were found to be the vital input parameters for scour depth model development. The results of these 4 data-driven models were compared with the results of nine conventional empirical equations using the performance criterion correlation coefficient (R), root mean squared error (RMSE), mean absolute percentage error (MAPE), Nash-Sutcliffe efficiency (E), and index of agreement (I_d) and graphical analysis. Based on values of the performance indices, ANFIS model was selected with R=0.986, RMSE=0.062, MAPE=6.767, E=0.975 and I_d=0.987. The results also show the outperformance of ANFIS model over the other selected data driven models and conventional empirical equations. This model can also be applied to the modelling of bridge pier scour in clear water conditions and can provide insight into the efficacy of modelling approaches in hydraulic properties.

1. Introduction

Bridge scour is a natural occurrence brought on by the erosive activity of water, which removes and excavates materials from the area around the piers of a bridge. The primary cause of bridge failure is scouring at the pier, which can result in huge financial losses and even human causalities [1]. Bridge scour is a dynamic process that fluctuates depending on variables like flow depth, angle and velocity of flow, pier size and shape, bed material gradation, etc [2–4]. Scouring around bridge piers includes complexity of three-dimensional flow and sediment transport processes [5]. Based on the mode of sediment transportation in the scour hole, scouring of bridge pier is studied under two different conditions, namely clear-water conditions and live bed condition [6]. In clear water conditions, no significant sedimentation is done in the scour hole by flowing water where as in the live-bed condition the upstream flow deposits significant amount of sediments in the scour hole [7].Various researchers have conducted numerous studies to comprehend the flow mechanism and anticipate the depth of scour in bridge piers [8–13]. Johnson [14] compared seven bridge pier scour equations which were laboratory-based using a large dataset of field data, for both live-bed and clear water conditions. It was found that most of these equations have uncertainties and over-predict the scour depth when applied to practice. Very comprehensive laboratory and field data sets were used to evaluate 23 pier-scour equations [15]. They concluded that the results varied even for the same case due to variation in parameters involved in these equations. Under the influences of flow, bed materials, and pier, it is challenging to establish and develop mathematical models of the scour process [16]. Moreover, there are not enough acceptable models to anticipate the scour depth to account for all potential variations from the abovementioned techniques [17]. Researchers have developed most of these empirical equations using data from the field and laboratory, and they varied from one another in terms of the variables taken into account when developing the scour model, the parameters involved in the equation, the circumstance in the lab etc [18]. A precise estimate of the scour depth is necessary for designing the bridge foundation securely; underestimating it could result in bridge failure, while overestimating it would result in exorbitant construction expenses [19]. Thus, numerous researchers have been interested in investigating and developing techniques for enhancing conventional physical-based analysis due to their recognition of these challenges and the significance of enhancing prediction abilities. Recently, soft computing methods have offered reasonably impressive solutions for hydrological systems and hydraulic engineering challenges when there is a highly complicated and nonlinear relation between the input-output pairs in the associated data [17,20–22]. To estimate scour depth, data-driven models (ANN, ANFIS, GEP, SVM GA, GP, MARS, FFNN, PSO, and M5-TREE) are now applied widely. Azmathullah et al. [23] employed genetic programming (GP) for predict the scour depth in the bridge pier. They demonstrated that the GP model was better in prediction of scour depth than the regression equations and artificial neural network (ANN). Pal et al. [24] used field data to investigate the potential of M5-Tree in calculating the local scour around bridge piers. The outcome showed that M5-Tree outperformed traditional equations in terms of performance. Akib et al. [25] applied an adaptive neuro-fuzzy inference system (ANFIS) and classical linear regression (LR) in the prediction of scour depth in the bridge. They illustrated that by comparing ANFIS with LR, the former showed comparatively greater accuracy and precision. Sreedhara et al. [26] have tried to investigate the use ANFIS and particle swarm optimization tuned support vector machine (PSO-

SVM) in prediction of scour depth around various pier shapes using experimental data and found that PSO-SVM model is an effective and reliable strategy for estimating the scour depth of a pier. Majedi-Asl et al. [27] examined the support vector machine (SVM) algorithm's capacity to estimate bridge scour depth depending on the pier shape. They also demonstrated that SVM outperformed the nonlinear regression model and expression programming (GEP) in terms of performance. Roshni and Prakash [28] investigated the use of the feedforward neural network (FFNN) and multivariate adaptive regression spline (MARS) models to predict the depth of scour around a bridge pier. The outcomes of the soft computing models were compared with those of empirical models, revealed that soft computing models were superior to other empirical models. Hassan and Jalal [29] applied GEP to predict local scour depth at a bridge pier. The findings suggest that, compared to nonlinear regression (NLR) and conventional regression models, the GEP predicts the local scour depth better. Hassan et al. [30] investigated GEP and ANN based on the PyTroch approach to estimate local scour depth near the bridge pier. They concluded that the equation produced by the ANN-based PyTroch approach performs better than GEP and NLR in estimating the scour depth. Daneshfaraz et al. [31] have experimented on the effect of cables in the local scouring of bridge piers. They claimed that increasing cable diameter might decrease the starting and ending scouring depths. Additionally, they reported that the ANN and ANFIS algorithms have great capabilities for estimating scour depth. Although several data-driven models have been used to estimate scour depth, the mode of sediment transport conditions, the complexity of their hydraulic characteristics, and the sediment properties themselves highlighted the need for developing new models to tackle the related challenges based on their specific characteristics and unique circumstances. Khalid et al. [32] suggested that most bridge failures occur under live-bed conditions during floods but only a few studies have been conducted using data-driven methodologies to determine the scour depth in bridge piers in live-bed conditions. These facts motivated the current research using data-driven models to estimate the scour depth in a live-bed condition. Choosing an ideal combination of input variables enhances the accuracy of data-driven models [33]. As a result, this research has three folds (i) to identify the best set of input variables for predicting the scour depth around bridge piers in live-bed conditions (ii) to assess how well four data-driven models GEP.M5-Tree, MARS, and ANFIS perform in estimating the depth of scour and the final step (iii), to compare the outcomes of all data driven models employed in this research with those of conventional empirical equations. Fig. 1 illustrated the flowchart of the methodology of present research work.

2. Methodology and data collection

This section explains the specifics of scour depth modeling utilizing data-driven models and conventional empirical models.

2.1. Gamma test

Scouring is exceptionally dynamic, nonlinear, and complex. The researcher have to follow tedious and laborious trial-and-error process for finding the optimal input combinations. The Gamma test (GT), which is a non-parametric test is used for the evaluations of best input variables, that are competent enough to build a reliable and smooth model for this problem. Agalbjorn et al. [34] described, the GT approach measures the base mean square error(MSE)

which affects the choices made regarding information input. The results of GT can be sorted by considering of another term, V-ratio, that provides a scale constant value evaluation within 0 and 1. Because it is independent of the yield range, the V-ratio is a good number to consider when comparing yields or yields from different informational collections. A smooth model shows a high consistency of the particular yield when the V-ratio is close to zero. For present study, the GT have been performed using the winGamma software [35].
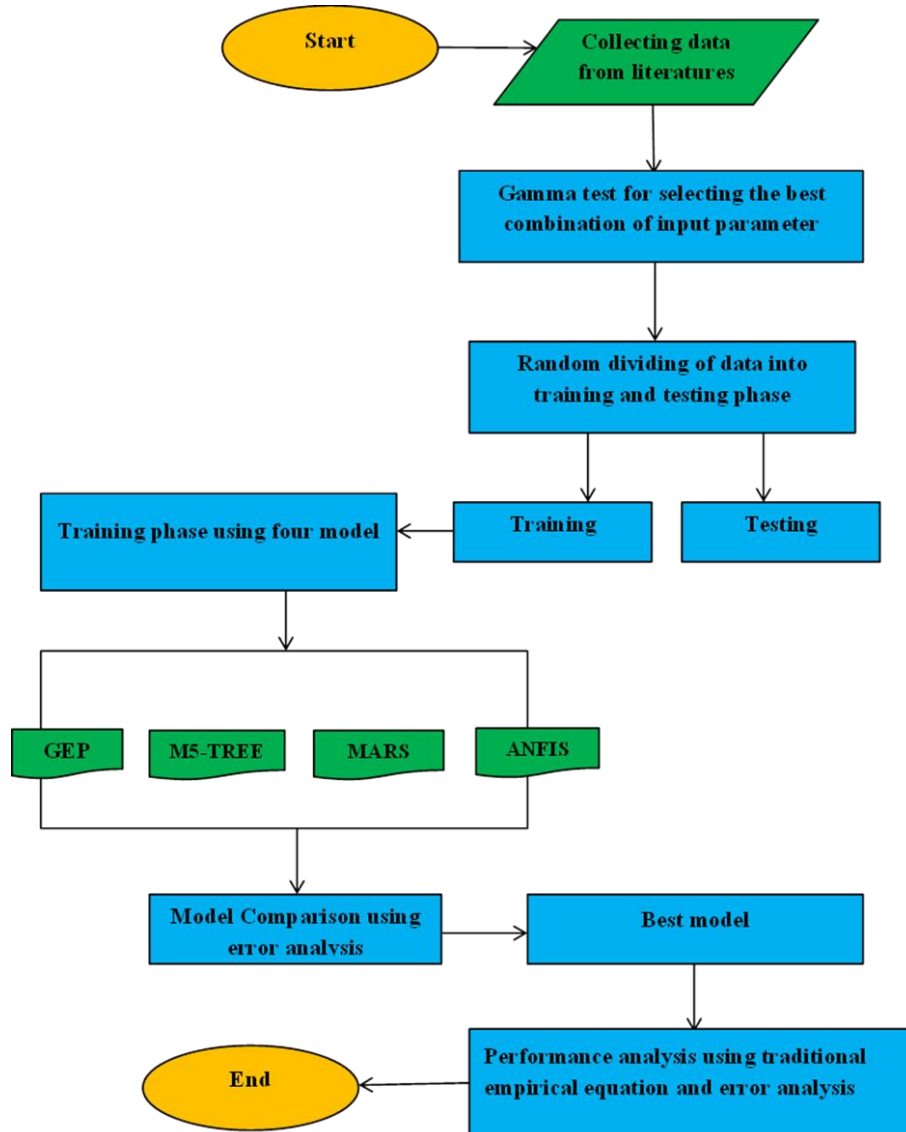
Fig.1. The overall methodology to predict the scour depth near bridge piers in live-bed conditions.

2.2. GEP

A search technique called GEP, or genetic programming extension [36], develops computer programs. Ferreira [37], Teodorescu and Sherwood [38] were the first to encode linear chromosomes before being represented or converted into expression trees (ETs). In contrast to GP, which combines genotype and Phenotype functions, GEP is a highly effective gene. There are five steps involved in the formulation of the GEP. Making a sample population group was the

first step. Any population size at this point can be used, however a study by Ferreira [37] suggested that a population range of 30 to 100 produced the best results. Second step involves the determination of fitness function for an individual chromosome. In the third step, we select the terminal and set of functions for constructing chromosomes. The fourth step is choosing the chromosome architecture through head size and gene count. In the fifth step, we adjust genetic operators like mutation, inversion, transport of insertion sequences (IS), root insertion sequences (RIS), transport of genes, double or single crossover along with gene crossover to achieve the required accuracy. Fig.2 depicts the methodology used to model the scour using the GEP approach [39].
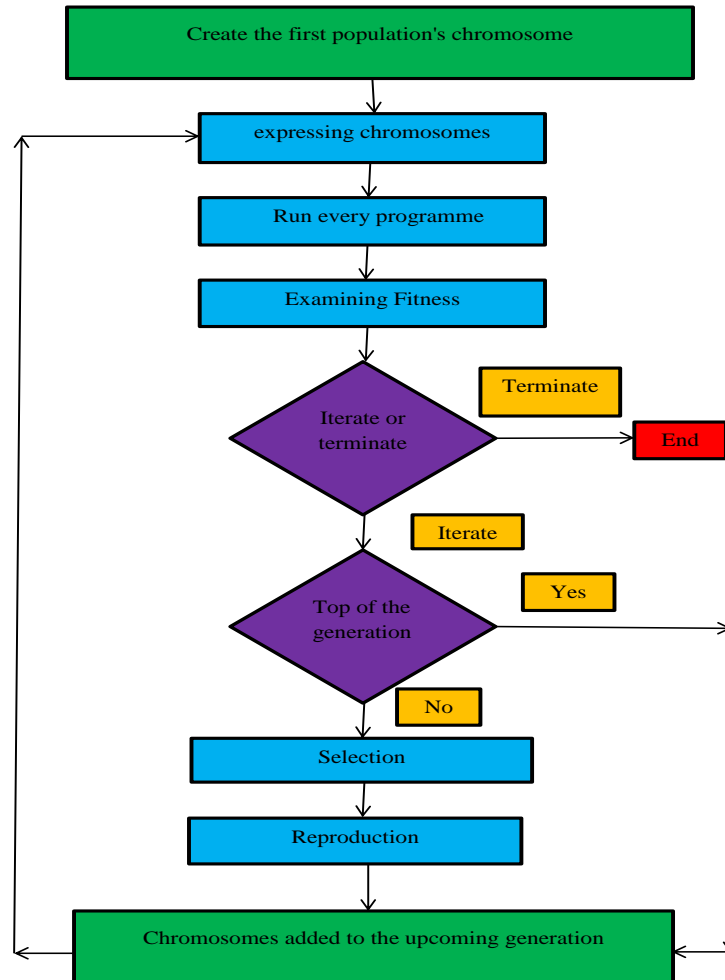

Fig.2. Flowchart of GEP modelling procedure [39].

2.3. M5-TREE

The M5-Tree model, that combines a linear regression and a conventional decision tree which segregates a dataset into sub-datasets, was first introduced by Quinlan [40]. The nomenclature of regression trees with fixed values at their leaves is enhanced by this model [41]. In the next step, pruning, and then splitting has been done to build the M5-Tree model. The standard deviation (SD) values are utilized as the measure of error at the nodes when computing the expected reduction in error using the splitting criterion. When only a few instances are left or differ

slightly, we stop further splitting in M5-Tree. This process results in splitting the M5-Tree into a sizeable tree-like structure by replacing a sub-tree with a leaf in the next step. This pruning may sometimes lead to abrupt discontinuities, which must be removed by smoothing the pruned tree in the final step. Fig.3 depicts the methodology used to model the scour using the M5-Tree approach [42].
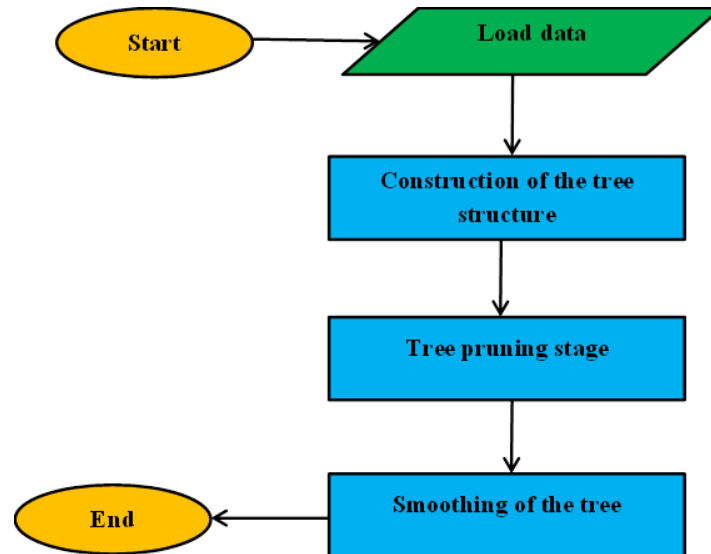

Fig.3. Flowchart of M5-Tree modelling procedure [42].

2.4. MARS

The multivariate adaptive regression splines (MARS) starting accesses the nonlinear relationship between a collection of input and dependent variables in high-dimensional datasets using a sequence of piecewise segments known as splines Friedman [43]. The input data has been divided further into subgroups with equal intervals for each spline. Knots describe the beginning and ending points of these segments [44]. These piecewise curves, known as basis functions (BFs), can identify nonlinearities, making this model more adaptable. MARS produced BFs by doing a step by step search for all conceivable univariate Knot positions and across all interactions between the factors. MARS development has different phases. The model has been generated in the forward phase in which BFs and the regression coefficients are constant. In the backward phase, MARS eliminates the terms with the lowest efficacy and the over-fitting components to improve the generalizability of the created model Zhang et al. [45]. Fig.4 depicts the methodology used to model the scour using the MARS approach.

2.5. ANFIS

Fuzzy inference system (FIS) and ANN techniques are incorporated in a hybrid scheme known as adaptive neuro-fuzzy inference system (ANFIS), according to Jang [46]. Fundamentally, ANFIS uses ANN for improving the benefits of FIS membership functions (MFs) which is done by adapting a learning process involving two method: back-propagation gradient descent and least-squares. Utilizing different algorithms, such as Takagi-sugeno, Mamdai, and Tsukamoto fuzzy, ANFIS can be effectively implemented [47]. The four building blocks of fuzzy inference systems (FIS), namely the fuzzifier, fuzzy inference engine, Knowledge base, and defuzzifier, incorporate the competence of an expert into the system design (Fig. 5) [48].
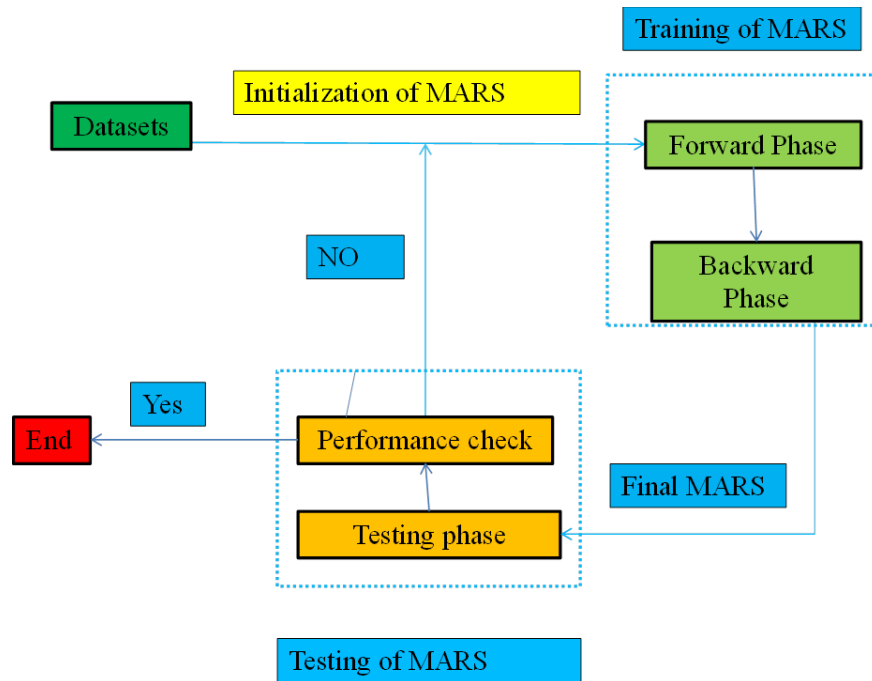
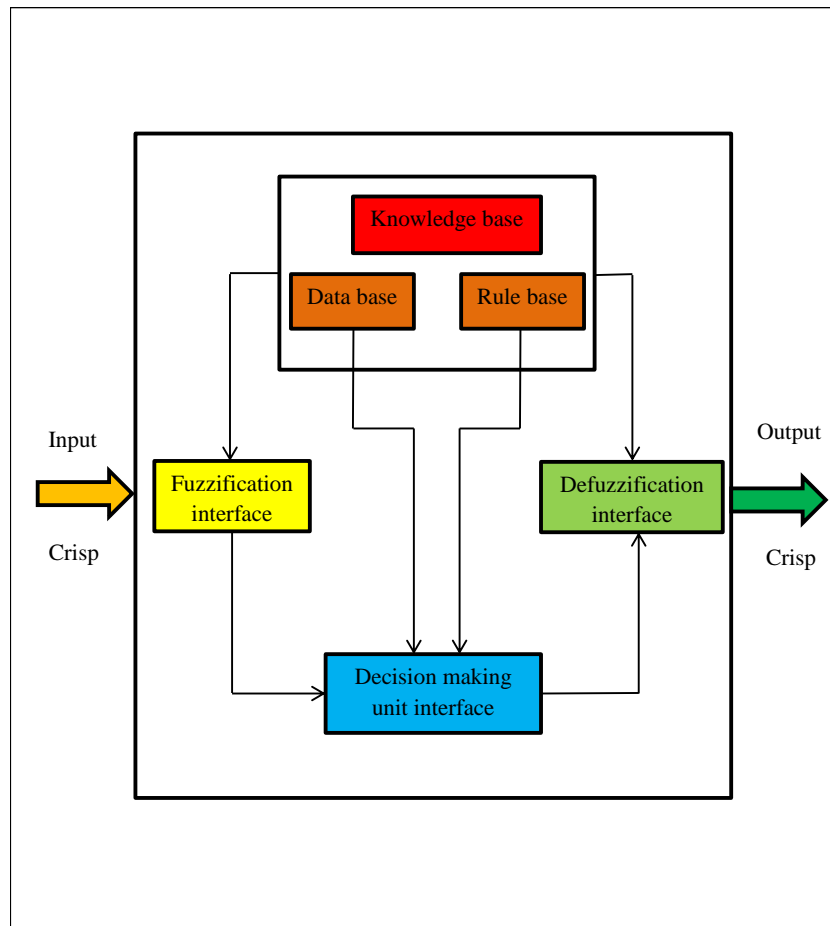**Fig.4.** Flowchart of MARS modelling procedure.

Fig. 5. A Flow diagram of a fuzzy inference system (FIS) [48].

A neural fuzzy network which have two inputs, one output and two laws is shown in Fig. 6 [25]. Fig. 6 illustrates the five layers that make up the architecture of ANFIS: a fuzzified layer (layer 1), an implication layer (layer2), a normalizing layer (layer3),a defuzzifying layer (layer4), and combined layers (layer5). Nodes in a layer can be of two kinds: while layers 2, 3 and 5 make the nodes which are fixed, layers 1 and 4 nodes that are adaptable. To determine the ideal ANFIS architecture, a trial-and-error approach using various membership functions with various input parameters and shapes, numbers, and types should be used.

Fig. 6. Architecture of ANFIS [25].

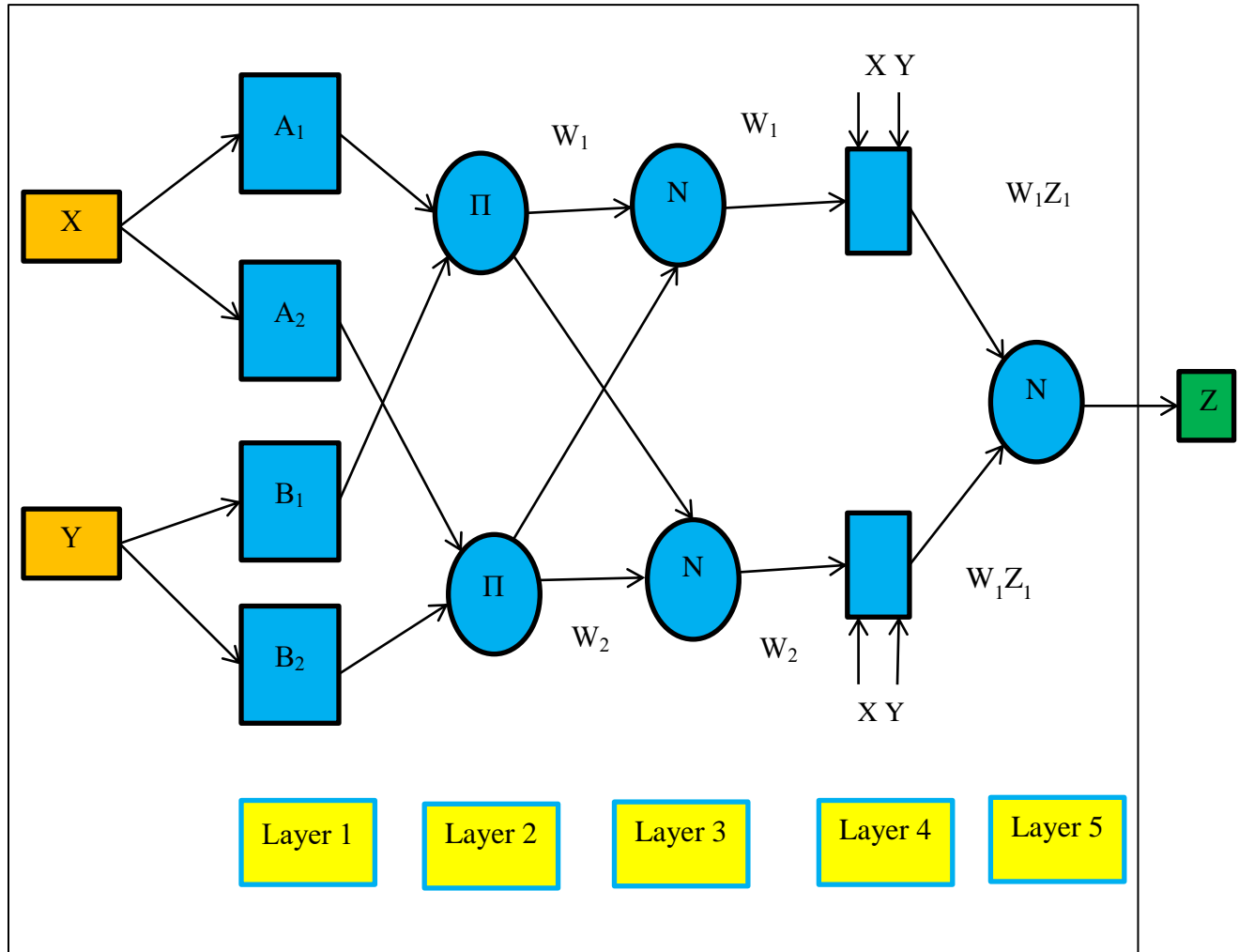## 2.6. Conventional empirical equations

Based on previous experimental studies, different empirical equations have been developed for estimating scour depth around bridge piers in live-bed conditions. Among these, nine pier scouring equations have been selected for this study in order to evaluate the performance of data-driven models. Table 1 displays the chosen conventional empirical equations.

Table 1
The existing conventional empirical equations for determination of scour depth in live bed condition.

Authors	Conventional empirical equations	Eqn nos.
Laursen and Toch [8]	$\dfrac{D_{sl}}{Y} = 1.35\left(\dfrac{D_P}{Y}\right)^{0.7}$	(1)
Larras [49]	$D_s = 1.05\,D_p^{0.75}$	(2)
Breusers [50]	$D_s = 1.4\,D_p$	(3)
Shen et al. [9]	$\dfrac{D_{sl}}{Y} = 3.4(Fr)^{0.67}\left(\dfrac{D_p}{Y}\right)^{0.67}$	(4)
Hancu [51]	$\dfrac{D_{sl}}{D_p} = 2.24\left(2\dfrac{U}{U_c} - 1\right)\left(\dfrac{U_c^2}{gD_p}\right)^{\frac{1}{3}}$	(5)
Breusers [10]	$\dfrac{D_{sl}}{D_P} = f\left(\dfrac{U}{U_c}\right)\left[2\tanh\left(\dfrac{Y}{D_p}\right)\right]$ $f\left(\dfrac{U}{U_c}\right) = 0,\ for\ \dfrac{U}{U_c} < 0.5$ $f\left(\dfrac{U}{U_c}\right) = \left(2\dfrac{U}{U_c} - 1\right),\ for\ 0.5 \le \dfrac{U}{U_c} < 1.0$ $f\left(\dfrac{U}{U_c}\right) = 1,\ for\ \dfrac{U}{U_c} \ge 1.0$	(6)
Melville and Sutherland [11]	$\dfrac{D_{sl}}{Y} = \dfrac{2.4\,D_P}{Y}$	(7)
Melville [52]	$\dfrac{D_{sl}}{D_P} = K_1 K_d K_{yD}$ $K_1 = \dfrac{U}{U_C}\ if\ \dfrac{U}{U_C} < 1\ and\ K_1 = 1\ otherwise$ $K_d = 0.57\log\left(2.24\,D_p/d_{50}\right) if\ D/D_{50} \le 25\ and\ K_d = 1\ otherwise,\ and$ $K_{yD} = 2.4\ if\ D_p/Y < 0.7,\ K_{yD} = 2\sqrt{Y/D_p}\ if\ 0.7 \le D_p/Y \le 5,\ and$ $K_{yD} = 4.5\,Y/D_P\ if\ D_P/Y > 5$	(8)
Richardson and Davis [5]	$\dfrac{D_{sl}}{Y} = 2.1\left(\dfrac{D_P}{Y}\right)^{0.65} Fr^{0.43}$	(9)

2.7. Dataset

For this research, live bed pier scour data were compiled from studies of Chabert and Engeldinger [53], Shen et al. [9], Norman [54], Jain and Fischer [55], Chee [56], Chiew [57], Butch [58], Wilson [59], US Geological Survey [60], Sheppard and Miller [61], and Holnbeck [62]. Scour of a bridge, according to Kothyari [63], is dependent on the hydraulic parameters approach mean velocity (U), the critical velocity of the sediment (Uc), pier diameter (Dp), median diameter of the sediment (d_{50}), approach flow depth (y), and Froude number (Fr). Currently 213 datasets have been used for model development. The range of the input datasets is displayed in Table 2. About 75% (160 sets) of the total 213 input-output pairs were randomly chosen and utilized for the purpose of training, where as the remaining 25% (53 sets) were used for testing [64].

The characteristics of fluid flow, bed sediments along with pier characteristics that effects the scour depth of bridge piers. Scour depth is indicated by the functional relationship Eq. (10):

$$D_{sl} = f\left(U, U_c, Y, g, D_{50}, D_p, \rho, k_s\right)$$

(10)

All pier had a zero angle of alignment with the flow. Circular, square, rectangular, and cylindrical pier shapes were all employed in this investigation.

Table 2
Details of selected input dataset range for the model development.

Author	D_P/D_{50}	Fr	U/U_C	σ_g	K_s	D_{sl}/Y	No. of datasets
Chabert and Engeldinger [53]	96.129-192.258	0.189-0.378	1.028-1.308	1.3	1	0.254-1.329	12
Shen et al. [9]	331.304	0.289	1.262	2.2	0.9	0.941	1
Norman [54]	128.336-365.76	0.522-0.863	1.009-2.202	3-5.5	0.9-1	0.194-0.5	2
Jain and Fischer [55]	20.360-405.993	0.499-1.498	1.025-4.690	1.3	1	0.351-1.811	30
Chee [56]	36.358-425.45	0.301-1.212	1.113-4.297	1.2-1.3	1	0.548-1.439	35
Chiew [57]	10.001-187.96	0.238-0.881	1.008-3.101	1.2-5.5	1	0.087-0.569	92
Butch [58]	50.8-67.733	0.349-0.517	1.006-1.337	1.5-2.7	0.9-1	0.049-0.209	12
Wilson [59]	324.687-499.165	0.380-0.461	1.034-1.291	6.2-6.9	1-1.1	0.093-0.735	11
U.S.Geological Survey [60]	56.444	0.452	1.419	2.3	1	0.186	1
Sheppard and Miller [61]	181.428	0.299-1.260	1.386-5.363	1.3	1	0.448-0.989	11
Holnbeck [62]	5.720-80.920	0.439-1.184	1.002-1.238	1.6-2.4	0.9-1.1	0.123-0.754	6

3. Model performance evaluation

The effectiveness of developed models were evaluated in the current study using a variety of statistical performance measures. As listed in Table 3, the following performance indices were

used: correlation coefficient (R), root mean squared error (RMSE), mean absolute percentage error (MAPE), Nash-Sutcliffe efficiency (E), and index of agreement (I_d).

Table 3
List of Statistical performance measures.

Statistical performance measures	Expressions	Eqn nos.
Correlation coefficient (R)	$R = \dfrac{\sum_{i=1}^{n}(P-P_a)\times(O-O_a)}{\sqrt{\sum_{i}^{n}(P-P_a)^2 \sum_{i}^{n}(O-O_a)^2}}$	(11)
Root mean squared error (RMSE)	$RMSE = \sqrt{\dfrac{1}{n}\sum_{i}^{n}(P-O)^2}$	(12)
Mean absolute percentage error (MAPE)	$MAPE = \dfrac{1}{n}\sum_{i}^{n}\dfrac{|P-O|}{O}\times 100$	(13)
Nash-Sutcliffe efficiency (E)	$E = 1-\dfrac{\sum_{i}^{n}(O-P)^2}{\sum_{i}^{n}(O-O_a)^2}$	(14)
Index of agreement (I_d)	$I_d = 1-\dfrac{\sum_{i}^{n}(O-P)^2}{\sum_{i}^{n}\left(|P-O_a|+|O-O_a|\right)^2}$	(15)

Where, O=observed value; P=predicted value; O_a= average of the observed value, and P_a=mean of the predicted value, i is the number of observations. The quality of the relationship between the predicted and observed data is expressed by the correlation correlation coefficient (R) .The discrepancy between values predicted by a model and the actual values observed is measured by the RMSE. MAPE is a measure of model prediction accuracy expressed as relative error. It should be noticed that the ideal model has an RMSE of 0 and R of 1, as well as MAPE values of lesser than 10% [65]. The ratio of mean square error to potential error is represented by the index of agreement(I_d).I_d have ranges lies between 0 stands for no correlation and 1 for perfect fit for predicted value and observed value [66] .To evaluate a model's predictive ability, Nash and Sutcliffe(1970) [67] suggested the Nash-Sutcliffe efficiency (E). E has a range between -∞ and 1. The more closely the model efficiency approaches 1, the more precise the model.

3. Results

3.1. Gamma Test for selection of inputs parameters

Gamma tests have been performed for all input combinations and the detailed test results are shown in Table 4. The input combination which yield the lowest absolute gamma value is considered as the ideal combination. There are 2m-1 combinations possible for m scalar inputs, but this can result in many irrational input combinations. Because the values of G and V-ratio are low (0.047 and 0.306, respectively), which are extremely close to zero in comparison to other blends, Table 4 demonstrates that the combination of five parameters with mask (11111) can

provide an appropriate model in contrast to other possible blends. For the present study, a combination of five non-dimensional parameters were used to develop models.

Table 4
Determining the best combination for scour depth modelling.

Sl.No.	Input Combination	Gamma	Standard Error	V-ratio	Mask
1	$D_P/D_{50}, Fr, U/U_C, \sigma_g, K_s$	0.047	0.024	0.306	11111
2	$D_P/D_{50}, Fr, U/U_C, \sigma_g$	0.047	0.024	0.307	11110
3	$D_P/D_{50}, Fr, U/U_C$	0.056	0.022	0.365	11100
4	$D_P/D_{50}, Fr$	0.057	0.027	0.370	11000
5	$Fr, U/U_C, \sigma_g, K_s$	0.075	0.006	0.487	01111
6	D_P/D_{50}	0.144	0.029	0.934	10000
7	$D_P/D_{50}, U/U_C, \sigma_g, K_s$	0.048	0.023	0.314	10111
8	$D_P/D_{50}, \sigma_g, K_s$	0.143	0.018	0.929	10011
9	$D_P/D_{50}, K_s$	0.146	0.026	0.949	10001

The gamma test findings has been used to create the model for this study, which uses five non-dimensional parameters as input and the ratio of scour depth to flow depth as output, listed in Eq. (16). The final non -dimensional function is as follows, and it depicts how the variables effects the scour depth at a bridge pier in live-bed condition:

$$\frac{D_{sl}}{Y} = f\left(\frac{D_p}{D_{50}}, Fr, \frac{U}{U_c}, \sigma_g, K_s\right) \tag{16}$$

D_P/D_{50} represents the sediment coarseness, Fr is the Froude number, U/U_c is flow intensity, σ_g is the sediment gradation coefficient of the bed material, and K_s is the shape factor.

Fig.7 depicts non-dimensional relationship between the selected input and desired output. It summaries the correlation matrix plots, which use Person's correlation coefficient to explain the linear association of the aforementioned variables. The Person correlation value ranges from [1 to -1], where a positive value of one depicts direct-proportionality of input and output variables to one another and vice-versa.

$$r_p = \frac{\sum_{i=1}^{n}(X - X_a)(Y - Y_a)}{\sqrt{\sum_{i=1}^{n}(X - X_a)^2 \sum_{i=1}^{n}(Y - Y_a)^2}} \tag{17}$$

Where X is the value of predictors, Y is the value of target, X_a is the average value of predictors, and Y_a is the average value target. According to the Fig.7, Fr (r_p=0.61) have the highest linear impact on the scour depth ratio (D_{sl}/Y) followed by U/Uc (r_p=0.45), D_P/D_{50} (r_p=0.29), and K_s (r_p=0.029), respectively.
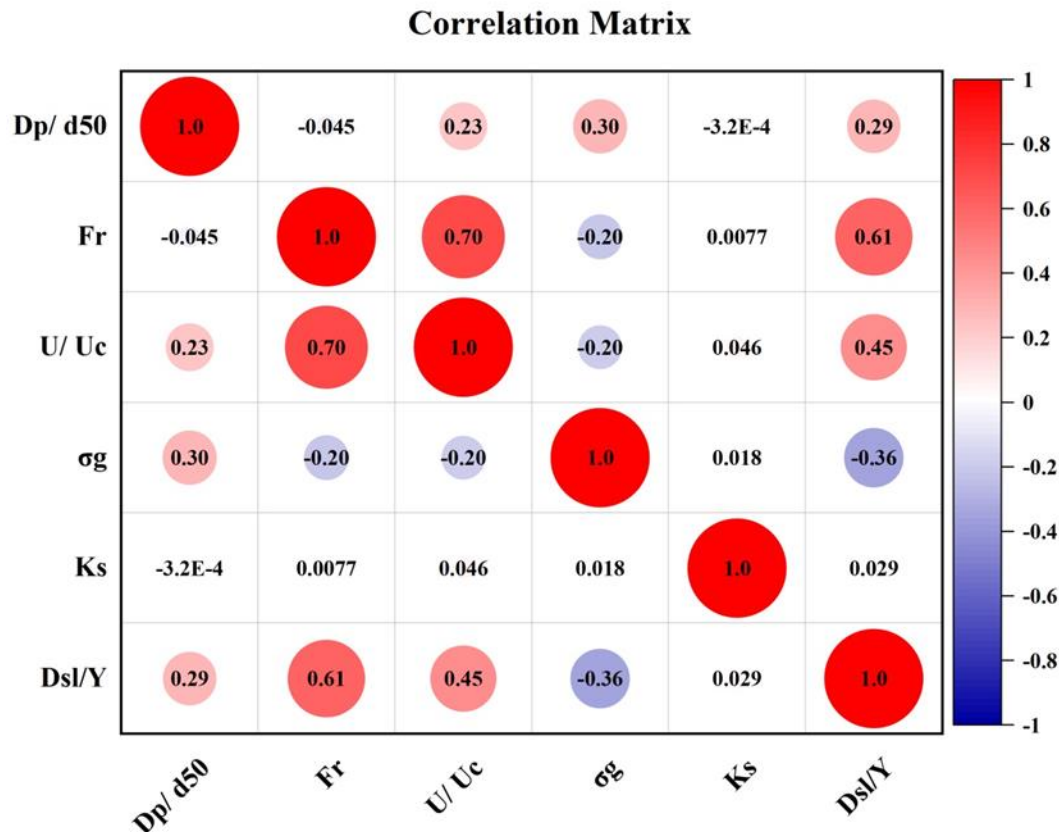
Correlation Matrix

Fig. 7. The plot of the correlation matrix between the input and output variable.

3.2. Modelling with GEP

For the GEP formulation, the laboratory and field data set from the current study have been considered. To begin the formulation, entire data set has been split into the training set (75%) and the testing set (25%). The parameters and procedures contained in GEP were subsequently established in five steps to enable the development of the mathematical equation which was required for estimation of scour depth. In first step several tests have been performed to determine the ideal population size, and ultimately the population size to be used in the study was taken 30 as this population size produce the most optimum results. The fitness function of an individual chromosome, as determined by RMSE, was measured in the subsequent step. In the third step, the terminal and set of functions for constructing chromosomes for the current study have been selected and listed in Table 5. In this study, gene number is selected as three and head size as eight in fourth step. In the fifth step, genetic operators were selected for making allowances for variation in both type and rate of expression through the final equation of these sub-expression trees (sub-ETs), which is linked by addition (+). There were three genes per chromosome. The software programme Gene Xpro Tools 5.0 was used after specifying all the necessary parameters. An explicit and concise mathematical equation for estimating the depth of scour around bridge piers is provided by this software. Table 5 shows the most optimum genetic operator values among all the tried genetic operators. The resultant scour depth formula is shown as an expression tree (ET) in Fig. 8 while the associated equation is written as Eq. (18). The

values of constant shown in Eq. (18) are G1C7=0.968, G1C4=-3.686, G2=-8.303, G3C9=3.529 and G3C4=10.775. Fig.8 show the ETs for GEP model in which d0= D_P/D_{50}, d1= Fr, d2= U/U_c, d3= σ_g and d4= K_s.

Table 5
Parameters of the optimized GEP model.

Serial number	Description of parameter	Parameter setting
1	Chromosomes	30
2	Genes	3
3	Head size	8
4	Number of generation	150979
5	Mutation rate	0.044
6	Inversion rate	0.1
7	Function set	+,-,*,/,power,Exp,Ln
8	One point recombination rate	0.3
9	Two- point recombination rate	0.3
10	Gene recombination rate	0.1
11	Gene transposition rate	0.1
12	Linking function	Addition
13	Program Size	38
14	Fitness function	RMSE

The explicit equation derived from Sub-ETs to create the GEP model may be written as:

$$\frac{D_{sl}}{Y} = \left\{ \left[\frac{-3.69\dfrac{U}{U_C}\dfrac{U}{U_C} - F_r + 3.69}{\dfrac{D_P}{D_{50}}} \times \frac{1}{0.97 + F_r} \right] + \frac{-0.12}{(K_s - 8.30)F_r} + \frac{\left(2.38 \times K_S^{3.53}\right)^{\left(\sigma_g - F_r\right)}}{\dfrac{U}{U_C}} \right\} \tag{18}$$

3.3. Modeling with M5-tree

M5-TREE modeling has been carried out using the Waikato Environment for Knowledge Analysis, or WEKA [68], a well-known software of machine learning tools developed at the University of Waikato. The M5- TREE is a tree based regression technique that requires only one parameter, the minimum number of training instances allowed at a terminal node, to be chosen for a specific dataset. Six training instances were found to work best with this input data after several tests. Using the M5 modeling approach, Table 6 shows the dataset's correlation coefficient and RMSE values. The availability of four simple linear relations (Eq. (19), (20), (21), and (22)), makes it easy to estimate pier scour using laboratory and field data, which is one of important advantages of the M5-TREE. The M5-TREE (8), produced the best outcome among all the tried M5-TREE models, as shown in Fig. 9. The use of linear models for various input parameters have been shown in Fig 9.
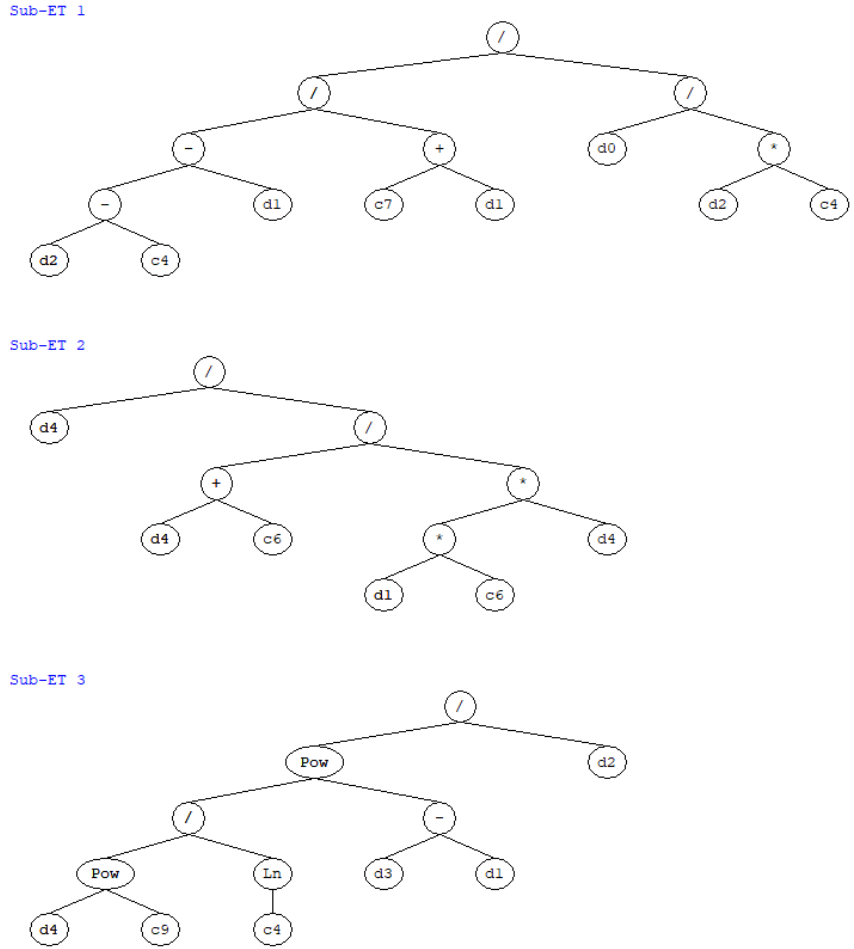
Fig. 8. Expression-Tree of GEP.

Table 6
Parameters variation with M5-TREE Model.

Model	Instances Number	Percentage split	Number of rules	Correlation coefficient	RMSE
M5-TREE(1)	1	10	1	0.613	0.329
M5-TREE(2)	1	20	1	0.851	0.217
M5-TREE(3)	1	30	1	0.846	0.216
M5-TREE(4)	1	40	1	0.777	0.245
M5-TREE(5)	1	50	1	0.768	0.257
M5-TREE(6)	1	25	1	0.856	0.219
M5-TREE(7)	5	25	1	0.856	0.219
M5-TREE(8)	6	25	4	0.856	0.219
M5-TREE(9)	15	25	4	0.779	0.206
M5-TREE(10)	20	25	1	0.779	0.206
M5-TREE(11)	25	25	1	0.755	0.293
M5-TREE(12)	5	20	1	0.851	0.217
M5-TREE(13)	6	20	4	0.851	0.217
M5-TREE(14)	10	20	4	0.852	0.216
M5-TREE(15)	15	20	4	0.812	0.246

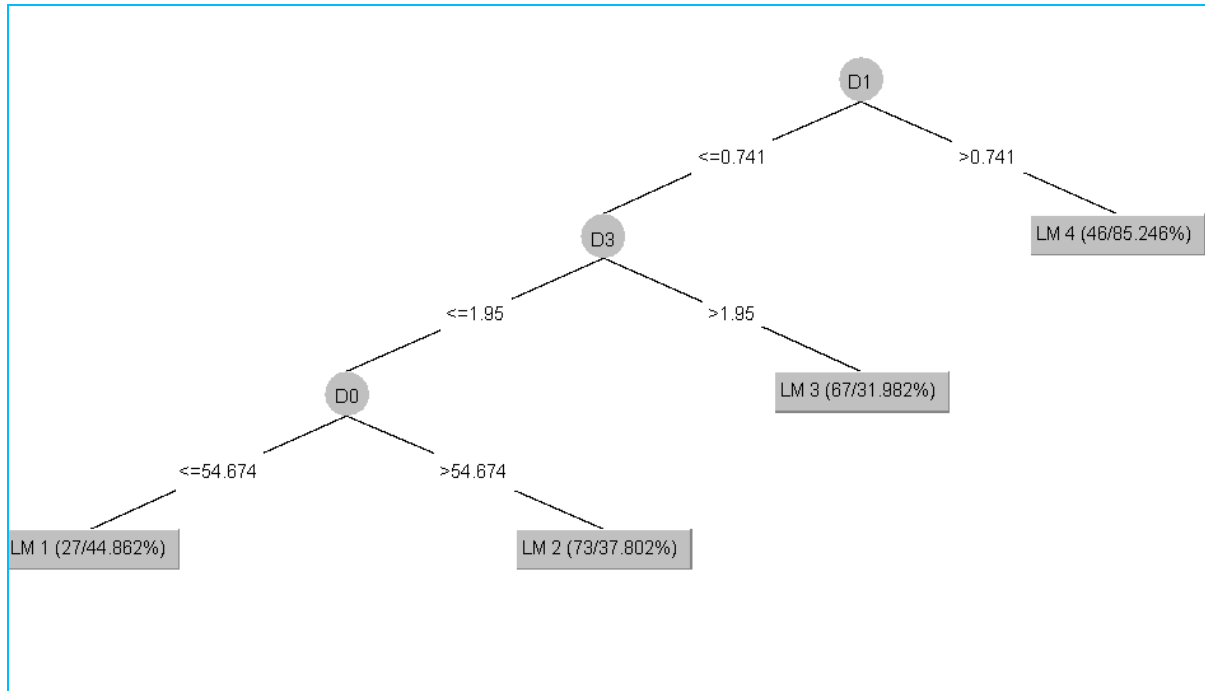**Fig. 9.** Tree representation of the best model (M5-TREE (8) in Table 6).

ConditionsLM 1 :

$D1 <= 0.741, D3 <= 1.95, D0 <= 54.674$

$$LM1: \frac{D_{sl}}{Y} = 0.001 * \frac{D_p}{D_{50}} + 0.4697 * F_r - 0.0925 * \frac{U}{U_c} - 0.1784 * \sigma_g + 0.5049$$

(19)

ConditionsLM 2 :

$D1 <= 0.741, D3 <= 1.95, D0 > 54.674$

$$LM2: \frac{D_{sl}}{Y} = 0.0018 * \frac{D_p}{D_{50}} + 2.0185 * F_r - 0.4339 * \frac{U}{U_c} - 0.8123 * \sigma_g + 1.269$$

(20)

ConditionsLM 3 :

$D1 <= 0.741, D3 > 1.95$

$$LM3: \frac{D_{sl}}{Y} = 0.0011 * \frac{D_p}{D_{50}} - 2.888 * F_r + 0.1554 * \frac{U}{U_c} - 0.0487 * \sigma_g + 0.244$$

(21)

ConditionsLM 4 :

$D1 > 0.741$

$$LM4: \frac{D_{sl}}{Y} = 0.0011 * \frac{D_p}{D_{50}} + 0.6772 * F_r - 0.0297 * \frac{U}{U_c} - 0.0268 * \sigma_g + 0.3758$$

(22)

Where, D0= Sediment Coarseness ratios, D1=Froude number and D3= gradation coefficient of the bed material.

3.4. Modeling with MARS

The adaptive regression splines (ARES) lab toolbox (version 1.13.0) in MATLAB were used to construct the MARS model in the present work. The type of function between the inputs and outputs does not have to be established a priori by MARS [69]. The number of basis functions in the forward step (BF (F)) and the degree of interaction (DOI) influences the MARS model performance [70]. In this study, the ranges for (BF (F)) and (DOI) were 1 to 50 and 0 to 4, respectively. Finally, we select the (BF (F)) and (DOI) values that produced the most accurate estimate of equilibrium scour depth. Compared to models with different DOI values, the MARS model with a DOI of 4 performed well for the present dataset. The MARS model is developed in the second phase, as was previously explained. Fifty basis functions were used for MARS development in the forward phase, and in the backward phase, it has been pruned to 2 basis functions. The best MARS equation is developed with 37 basis functions. Generalized cross-validation (GCV), which is based on the process of forward selection and backward deletion process, is used to develop the final MARS models [44]. Table-7 provides the variations of the different parameter for each MARS models.MARS-10 provided the most optimum values of different among all the MARS models used in the current dataset.

Table 7
Parameter variation with MARS model.

Model	Max-BF	Max-IN	BF(F)	BF(B)	BF(M)	DOI	NVM	R	RMSE
MARS-1	1	5	1	0	1	0	0	1.34E-31	0.388
MARS-2	5	5	5	0	5	1	2	0.596	0.246
MARS-3	10	5	9	1	8	2	4	0.763	0.188
MARS-4	15	5	15	3	12	3	4	0.890	0.128
MARS-5	20	5	18	2	15	3	5	0.909	0.117
MARS-6	25	5	24	1	18	3	5	0.927	0.104
MARS-7	30	5	28	2	22	3	5	0.940	0.094
MARS-8	35	5	34	1	25	3	5	0.948	0.088
MARS-9	40	5	38	2	33	4	5	0.955	0.082
MARS-10	50	5	48	2	37	4	5	0.957	0.080

Where, Max-BF=Maximum Basis function, Max-IN=Maximum number of interaction, BF (F)= Basis function in forward step, BF (B)= Basis function in backward,BF (M)=Basis function in final model, DOI=Degree of interaction, NVM=Number of variables used in model, R= Correlation coefficient and (RMSE)=Root mean square error.

3.5. Modeling with ANFIS

In this study, the ANFIS model has been developed in relation with chosen input parameters for estimating scour depth. The ANFIS model was developed in MATLAB using a variety of membership functions (MFs), including the triangular membership function (trimf), trapezoidal membership function (trapmf), generalised bell-shaped membership function (Gbellmf), gaussian membership function (gaussmf), and gaussian combination membership function (gauss2mf). There are two options for running a fuzzy model: subtractive fuzzy clustering (which requires

less computational effort) and grid partitioning (requiring more computational effort).To find the best ANFIS model we have run the MATLAB with different type MFs for each input paramter and tried different combinations on a number of members functions from 2 to 3.In contrast to other MFs, the grid partition (GP) technique of grid generation with the triangular-shaped MF (trimf) has a best performance during the development of the ANFIS model. Table 8 provides the specifics of the configuration of best ANFIS model for prediction of bridge pier scour depth.

Table 8
Parameters of the best ANFIS model.

Serial number	Architecture of ANFIS	Parameter setting
1	Number of membership function	33333
2	Algorithm selected	Hybrid
3	Number of epoch runs given	100
4	Generated fuzzy inference system	Grid partition
5	Membership function(MF) type	linear
6	Type of membership function(MF)used	trimf
7	Number of nodes	524
8	Number of linear parameters	1458
9	Number of nonlinear parameters	45
10	Total number of parameters	1503
11	Number of fuzzy rules	243

3.6. Performance evaluation of data-driven models with conventional empirical equations

Here,we have used nine conventional empirical equations to estimate scour depth in a bridge pier under live-bed conditions.In addition to this, results of the four data-driven models of the four data-driven models used for the same purpose of bridge pier scour estimation have been compared with the results of these selected nine conventional equations for evaluation of the effectiveness of suggested model (Table 9). The comparison shows that the ANFIS model outperformed all four data-driven models and other empirical equations. The highest and lowest correlation coefficients were achieved by the Larras [49] and Hancu [51] equations as 0.913 and 0.717, respectively. Also, the highest and lowest RMSE values for Shen et al. [9] and Breusers [50] equation were 0.919 and 0.244, respectively. For the equation Shen et al. [9], the maximum value of MAPE =168.54 is achieved. The MAPE parameter in Breusers [10] equation has a minimum value of 44.525. The maximum and lowest values of E, 1.503 and 0.371, were found by the Shen et al. [9] equation and Larras [49], respectively. The highest and minimum values of the Index of Agreement (Id) are 0.813 and 0.521 for the Larras [49] equation.

4. Discussion

4.1. Graphical Analysis of Models

The exactness of the proposed data-driven models in predicting the depth of scour around piers of a bridge in live-bed conditions have been analysed using graphical analysis. Error distribution of four data-driven models in the shape of violin plots, have been demonstrated in Fig.10. From Fig.10, the maximum and minimum relative deviation of the GEP model are 65.55 and 53.06, for the M5-Tree model 57.38 and -111.19, for MARS 26.29 and -42.15, ANFIS 24.57 and -48.70,

respectively. These results shows, the distribution of errors in the ANFIS model is better than the other three data-driven models. Furthermore, it is evident that the MARS model resembles the ANFIS model and is similarly close to it (Fig. 10). Therefore, the MARS model can be regarded as the best model in predicting scour depth after the ANFIS model. Which also performs significantly better than the GEP and M5-Tree model.
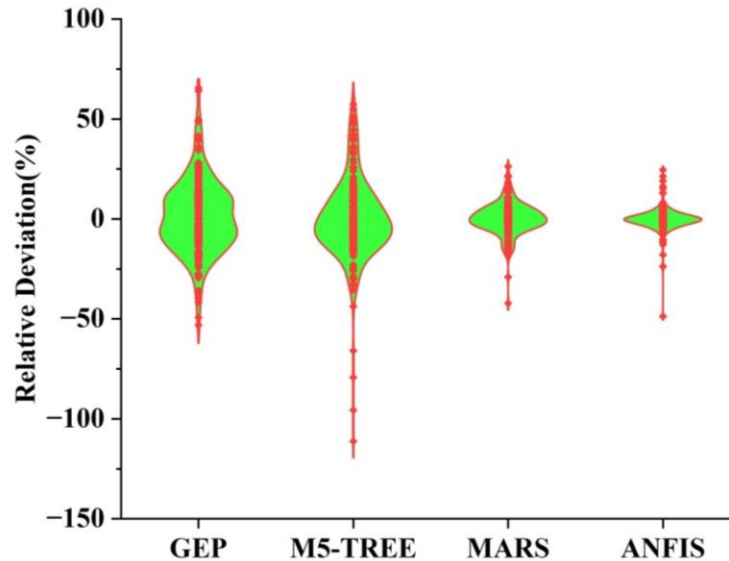

Fig. 10. Violin plot of error distribution for data-driven model.

Fig.11 illustrates the scatter plots between observed versus predicted scour depth ratio in live-bed conditions for all data and models in the present research. From Fig.11, the value of the coefficient of determination (R^2) for the ANFIS model equals 0.9741, the highest among all four data-driven models and nine conventional equations. The MARS model also performed well quantitatively value of R^2 is equal to 0.9573, which was lower after the ANFIS model, and the MARS model under-predicted more scour depth ratio measurement compared to the ANFIS model. GEP and M5-Tree models have a value of R^2 equal to 0.7855 and 0.6997, respectively, which is lesser than the value of R^2 for ANFIS and MARS. It was observed that the conventional equation, [5,8,9,11,49,52] over-predict the scour depth ratio, which would result in overdesigned bridge foundations. However, [10,50,51] equations contain a mix of under-prediction and over-prediction values. From fig.11, maximum and minimum over prediction occurs in Shen et al. [9] and Breusers [10], having R^2 equal to 0.7596 and 0.6523, respectively. Fig.11 showed that the Larras [49] equation performed quantitatively well as the coefficient of determination value is 0.8339, the highest among the nine conventional equations used in the present study. However, it was observed that predicted values of Larras [49] equation shows slightly variance owing to the fact it is based solely on pier characterstics and is not sensitive to hydraulic or sediment factors. None among of the conventional empirical equations estimated the scour depth consistently as done by the data-driven models for the live-bed condition, as illustrated in Fig.11. It can be seen that ANFIS model results are closest to the best-fit line and thus, indicates better accuracy. So, the ANFIS model effectively adjusts to the complex non-linear relationship between the

parameters and hence it can be adopted for prediction of scour, which has to be considered in design of hydraulic structures.
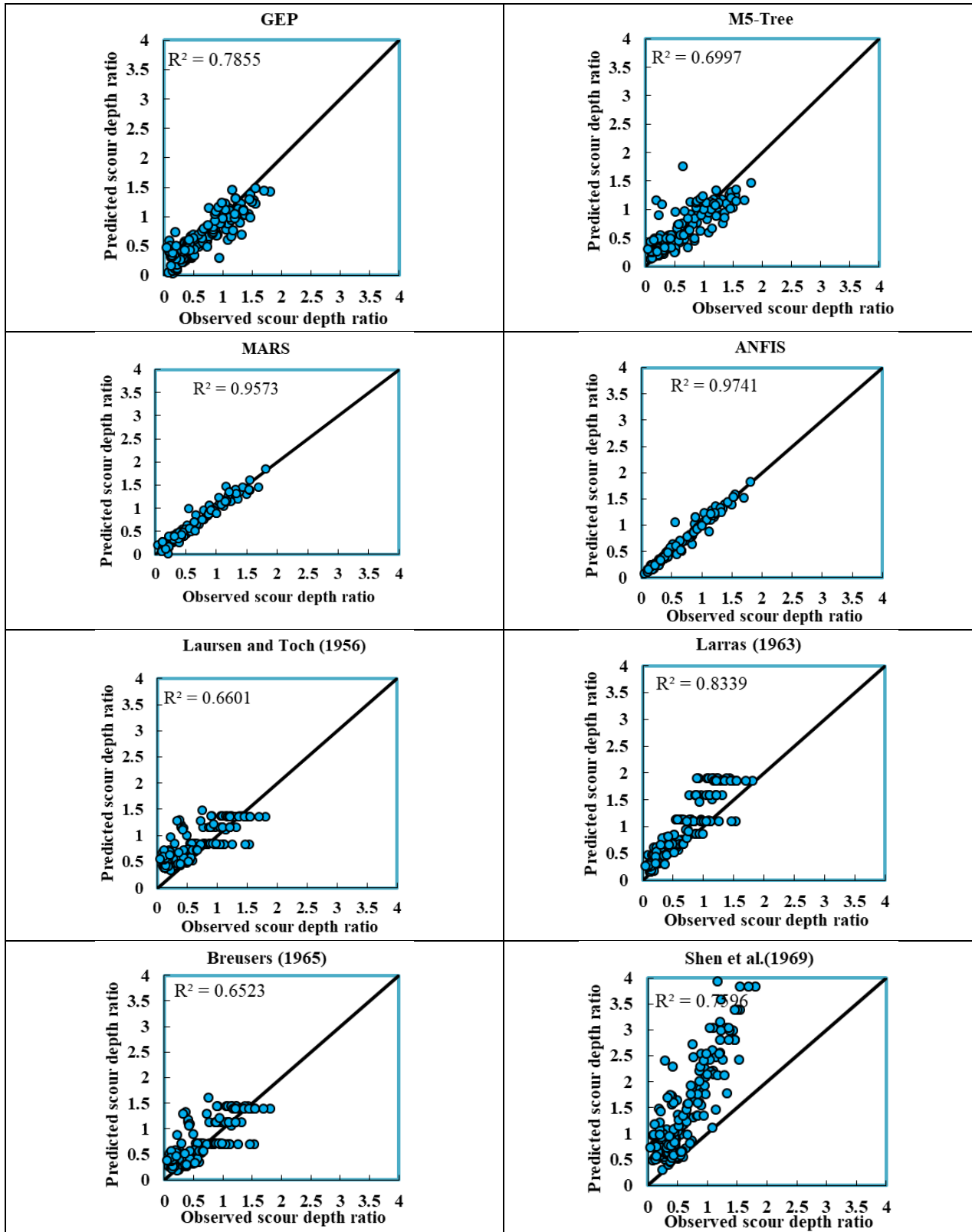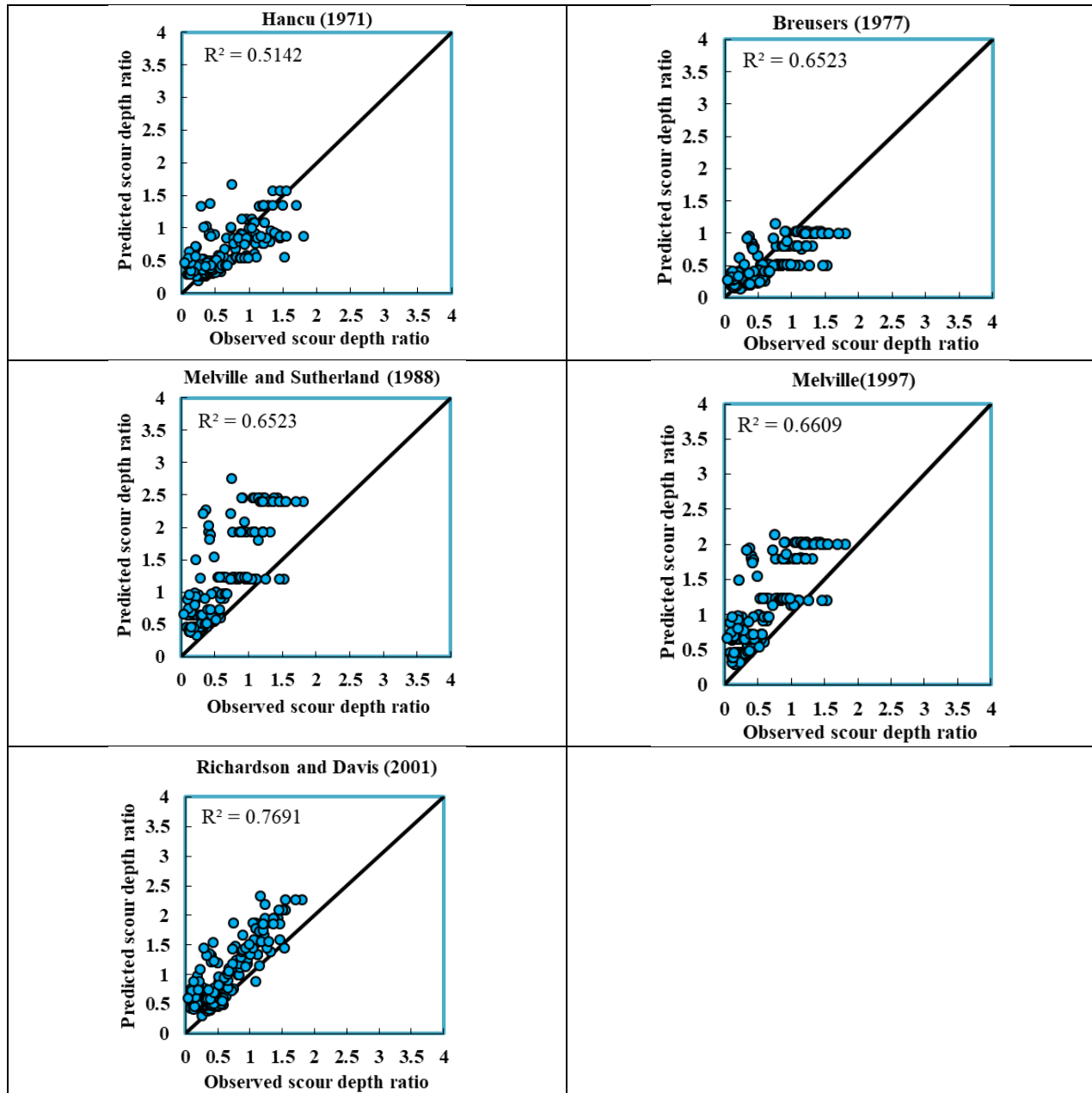
Fig. 11. Scatter plots of observed versus predicted scour depth ratio in live-bed condition for all data and models in present research.

4.2. Comparison of the Data-driven models with existing conventional empirical equations

In this section the data-driven models of GEP, M5-Tree, MARS, and ANFIS have been used to predict bridge scour depth using non-dimensional dataset configurations in live-bed conditions. Table 9 indicates the statistical results obtained from all data-driven models and previous models. The results of these calculation clearly demonstrates that ANFIS is the best model, with R =0.986, RMSE = 0.062, MAPE = 6.767, E = 0.975, and I_d = 0.9831 for the overall dataset, followed by MARS and GEP.

The statistical indices R, RMSE, MAPE, E, and I_d, have been calculated for all the developed models and the empirical expressions and the results are listed in Table 9. Additionally, the performance index (PI_m), a single multi-index criterion, is used for precise validation [42].

$$PI_m = 1/3\left(\frac{R_{min}}{R_m} + \frac{RMSE_m}{RMSE_{max}} + \frac{MAPE_a}{MAPE_{max}}\right)$$

(23)

Each predictive scour depth model prediction for the live bed is indicated by the subscript "m." The statistical performance criteria is described in Table 9 which highlights that all data-driven GEP (PI_m=0.417), M5-Tree (PI_m=0.438), MARS (PI_m=0.303), and ANFIS (PI_m=0.278) models are more precise than selected existing models in evaluating the scour depth in live-bed conditions. However, the best performing among all models especially with respect to relationships for the existing database is Breusers [50], which is ranked fifth and has the lowest(RMSE=0.244 and PI_m=0.474) and highest(E=0.624) values.

Shen et al. [9] also have the worst statistical performance indices, with RMSE=0.919, MAPE=168.54, E=-4.292, and PI_m=0.940, ranking thirteen as the least accurate in evaluating scour depth in live bed conditions when taking data sets into account.

Table 9
Performance of data driven model and conventional empirical equation.

MODEL	R	RMSE	MAPE	E	I_d	PI_m	Rank
GEP	0.886	0.183	40.885	0.79	0.863	0.417	3
M5-TREE	0.836	0.215	37.903	0.711	0.818	0.438	4
MARS	0.978	0.08	15.225	0.958	0.978	0.303	2
ANFIS	0.986	0.062	6.767	0.975	0.987	0.278	1
Laursen and Toch [8]	0.812	0.266	66.533	0.554	0.793	0.522	8
Larras [49]	0.913	0.316	55.364	0.371	0.831	0.486	7
Breusers [50]	0.807	0.244	45.463	0.624	0.815	0.474	5
Shen et al. [9]	0.871	0.919	168.54	4.292	0.561	0.940	13
Hancu [51]	0.717	0.277	54.288	0.515	0.686	0.541	9
Breusers [10]	0.807	0.279	44.525	0.509	0.521	0.485	6
Melville and Sutherland [11]	0.807	0.632	112.97	1.503	0.641	0.748	12
Melville [52]	0.812	0.517	104.18	0.678	0.685	0.687	11
Richardson and Davis [5]	0.877	0.421	98.181	0.107	0.761	0.619	10

It is noteworthy that each prior relationship is present for specific conditions, including variable characteristics of the fluid, flow, bed sediment, and pier type. These comparisons based on a particular data set do not establish the inability of the conventional empirical equations.

5. Conclusions

To aid in design of a cost effective and secure bridge that could even withstand during flood situation by proper estimation of scour depth around the piers of the bridge. This research uses four data-driven modeling approaches: GEP, M5-TREE, MARS, and ANFIS to evaluate scour depth in live-bed conditions. A total of 213 different datasets from various field studies and laboratory experiments published in the literature are collected to generate a new model. The model created utilizing above mention dataset provides a better model for variable input parameters resulting from the ever changing dynamic situation in live-bed conditions. Gamma test results reveal a combination of five dimensionless input parameters that includes sediment Coarseness Ratio (D_P/D_{50}), Froude Number (Fr), Flow Intensity (U/U_c), Gradation Coefficient of the Bed material (σ_g), and shape factor (K_s) are factors on which scour depth variation depends during live-bed conditions. When the results of these four data-driven model were compared, it was found that ANFIS performed the best, followed by MARS, M5-TREE and GEP models respectively. Then these data-driven models were compared with nine conventional empirical equations. The findings showed that the ANFIS outperformed other models, with Breusers [50] and Shen et al. [9] having the lowest and highest errors in scour depth prediction. Out of all the models chosen for the current database, the ANFIS model with $PI_m=0.417$ ranked first. As shown by the current study's findings, ANFIS has a high capacity for applicability and practicability in predicting scour depth in live-bed conditions around a bridge pier. With similar conditions and a wide variety of input parameters, this can be used effectively for pertinent tasks.

Funding

Conflicts of interest

The authors declare no conflict of interest.

References

[1]   Shepherd R, Frost JD. Failures in civil engineering: structural, foundation and geoenvironmental case studies, ASCE; 1995.

[2]   Melville BW. Local scour at bridge sites 1975.

[3]   J. RA, Robert E. Clear-Water Scour at Cylindrical Piers. J Hydraul Eng 1983;109:338–50. https://doi.org/10.1061/(ASCE)0733-9429(1983)109:3(338).

[4]   Ansari SA, Kothyari UC, Ranga Raju KG. Influence of cohesion on scour around bridge piers. J Hydraul Res 2002;40:717–29. https://doi.org/10.1080/00221680209499918.

[5]   Richardson E V, Davis SR. Evaluating scour at bridges, United States. Fed Highw Adm Off Bridg Technol 2001.

[6]   C. KU, J. GRC, G. RRK. Temporal Variation of Scour Around Circular Bridge Piers. J Hydraul Eng 1992;118:1091–106. https://doi.org/10.1061/(ASCE)0733-9429(1992)118:8(1091).

[7]   Toth E, Brandimarte L. Prediction of local scour depth at bridge piers under clear-water and live-bed conditions: comparison of literature formulae and artificial neural networks. J Hydroinformatics 2011;13:812–24. https://doi.org/10.2166/hydro.2011.065.

[8]   Laursen EM, Toch A. Scour around bridge piers and abutments (Vol. 4). Ames, IA Iowa Highw Res Board 1956.

[9]   W. SH, R. SV, Susumu K. Local Scour Around Bridge Piers. J Hydraul Div 1969;95:1919–40. https://doi.org/10.1061/JYCEAJ.0002197.

[10]   Breusers HNC, Nicollet G, Shen HW. Local Scour Around Cylindrical Piers. J Hydraul Res 1977;15:211–52. https://doi.org/10.1080/00221687709499645.

[11]   W. MB, J. SA. Design Method for Local Scour at Bridge Piers. J Hydraul Eng 1988;114:1210–26. https://doi.org/10.1061/(ASCE)0733-9429(1988)114:10(1210).

[12]   Hoffmans G, Verheij HJ. Scour Manual 1997. Rotterdam/Brookf Balkema 1997.

[13]   Melville BW, Coleman SE. Bridge scour. Water Resources Publication; 2000.

[14]   A. JP. Comparison of Pier-Scour Equations Using Field Data. J Hydraul Eng 1995;121:626–9. https://doi.org/10.1061/(ASCE)0733-9429(1995)121:8(626).

[15]   M. SD, B. M, H. D. Evaluation of Existing Equations for Local Scour at Bridge Piers. J Hydraul Eng 2014;140:14–23. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000800.

[16]   Firat M, Gungor M. Generalized Regression Neural Networks and Feed Forward Neural Networks for prediction of scour depth around bridge piers. Adv Eng Softw 2009;40:731–7. https://doi.org/10.1016/j.advengsoft.2008.12.001.

[17]   Lee TL, Jeng DS, Zhang GH, Hong JH. Neural Network Modeling for Estimation of Scour Depth Around Bridge Piers. J Hydrodyn 2007;19:378–86. https://doi.org/10.1016/S1001-6058(07)60073-0.

[18]   Lu D, S. CC. Bridge Scour: Prediction, Modeling, Monitoring, and Countermeasures—Review. Pract Period Struct Des Constr 2010;15:125–34. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000041.

[19]   Azamathulla HM, Ghani AA. Genetic Programming to Predict River Pipeline Scour. J Pipeline Syst Eng Pract 2010;1:127–32. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000060.

[20]   Sudheer KP, Jain A. Explaining the internal behaviour of artificial neural network river flow models. Hydrol Process 2004;18:833–44. https://doi.org/10.1002/hyp.5517.

[21]   Azmathullah HM, Deo MC, Deolalikar PB. Neural Networks for Estimation of Scour Downstream of a Ski-Jump Bucket. J Hydraul Eng 2005;131:898–908. https://doi.org/10.1061/(ASCE)0733-9429(2005)131:10(898).

[22]   Rahbar A, Mirarabi A, Nakhaei M, Talkhabi M, Jamali M. A Comparative Analysis of Data-Driven Models (SVR, ANFIS, and ANNs) for Daily Karst Spring Discharge Prediction. Water Resour Manag 2022;36:589–609. https://doi.org/10.1007/s11269-021-03041-9.

[23]   Azamathulla HM, Ghani AA, Zakaria NA, Guven A. Genetic Programming to Predict Bridge Pier Scour. J Hydraul Eng 2010;136:165–9. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000133.

[24]   Pal M, Singh NK, Tiwari NK. M5 model tree for pier scour prediction using field dataset. KSCE J Civ Eng 2012;16:1079–84. https://doi.org/10.1007/s12205-012-1472-1.

[25]   Akib S, Mohammadhassani M, Jahangirzadeh A. Application of ANFIS and LR in prediction of scour depth in bridges. Comput Fluids 2014;91:77–86. https://doi.org/10.1016/j.compfluid.2013.12.004.

[26]   Sreedhara BM, Rao M, Mandal S. Application of an evolutionary technique (PSO–SVM) and ANFIS in clear-water scour depth prediction around bridge piers. Neural Comput Appl 2019;31:7335–49. https://doi.org/10.1007/s00521-018-3570-6.

[27]   Majedi-Asl M, Daneshfaraz R, Fuladipanah M, Abraham J, Bagherzadeh M. Simulation of bridge pier scour depth base on geometric characteristics and field data using support vector machine

algorithm. J Appl Res Water Wastewater 2020;7:137–43. https://doi.org/10.22126/arww.2021.5747.1189.

[28] Thendiyath R, Prakash V. Role of Regression Models in Bridge Pier Scour Prediction. Int J Appl Metaheuristic Comput 2020;11:156–70. https://doi.org/10.4018/IJAMC.2020040108.

[29] Hassan WH, Jalal HK. Prediction of the depth of local scouring at a bridge pier using a gene expression programming method. SN Appl Sci 2021;3:159. https://doi.org/10.1007/s42452-020-04124-9.

[30] Hassan WH, Hussein HH, Alshammari MH, Jalal HK, Rasheed SE. Evaluation of gene expression programming and artificial neural networks in PyTorch for the prediction of local scour depth around a bridge pier. Results Eng 2022;13:100353. https://doi.org/10.1016/j.rineng.2022.100353.

[31] Seifollahi M, Abbasi S, Abraham J, Norouzi R, Daneshfaraz R, Lotfollahi-Yaghin MA, et al. Optimization of Gravity Concrete Dams Using the Grasshopper Algorithm (Case Study: Koyna Dam). Geotech Geol Eng 2022:1–16. https://doi.org/10.1007/s10706-022-02227-1.

[32] Khalid M, Muzzammil M, Alam J. A reliability-based assessment of live bed scour at bridge piers. ISH J Hydraul Eng 2021;27:105–12. https://doi.org/10.1080/09715010.2019.1584543.

[33] Ghorbani MA, Zadeh HA, Isazadeh M, Terzi O. A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. Environ Earth Sci 2016;75:476. https://doi.org/10.1007/s12665-015-5096-x.

[34] Agalbjorn S, Koncar N, Jones AJ. A note on the gamma test. Neural Comput Appl 1997;5.3:131–3.

[35] Durrant PJ. "winGamma: A non-linear data analysis and modelling tool with applications to flood prediction." UK: Department of Computer Science, Cardiff University, Wales, UK; 2001.

[36] Koza JR. Genetic programming II: automatic discovery of reusable programs. MIT press; 1994.

[37] Ferreira C. Gene expression programming: a new adaptive algorithm for solving problems. ArXiv Prepr Cs/0102027 2001.

[38] Teodorescu L, Sherwood D. High Energy Physics event selection with Gene Expression Programming. Comput Phys Commun 2008;178:409–19. https://doi.org/10.1016/j.cpc.2007.10.003.

[39] Aytac G, Ali A. New Approach for Stage–Discharge Relationship: Gene-Expression Programming. J Hydrol Eng 2009;14:812–20. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000044.

[40] Adams A, Sterling L. AI '92. AI '92, WORLD SCIENTIFIC; 1992, p. 1–410. https://doi.org/doi:10.1142/9789814536271.

[41] Witten, Ian H. and EF. Data mining: practical machine learning tools and techniques with Java implementations. Acm Sigmod Rec 311 2002:76–7.

[42] Ahmadianfar I, Jamei M, Karbasi M, Sharafati A, Gharabaghi B. A novel boosting ensemble committee-based model for local scour depth around non-uniformly spaced pile groups. Eng Comput 2022;38:3439–61. https://doi.org/10.1007/s00366-021-01370-2.

[43] Friedman JH. Multivariate Adaptive Regression Splines. Ann Stat 1991;19:1–67. https://doi.org/10.1214/aos/1176347963.

[44] Deo RC, Kisi O, Singh VP. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. Atmos Res 2017;184:149–75. https://doi.org/10.1016/j.atmosres.2016.10.004.

[45] Zhang W, Wu C, Li Y, Wang L, Samui P. Assessment of pile drivability using random forest regression and multivariate adaptive regression splines. Georisk Assess Manag Risk Eng Syst Geohazards 2021;15:27–40. https://doi.org/10.1080/17499518.2019.1674340.

[46] Jang J-SR. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 1993;23:665–85. https://doi.org/10.1109/21.256541.

[47] Tien Bui D, Pradhan B, Lofman O, Revhaug I, Dick OB. Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. Comput Geosci 2012;45:199–211. https://doi.org/10.1016/j.cageo.2011.10.031.

[48] Firat M. Scour depth prediction at bridge piers by Anfis approach. Proc Inst Civ Eng - Water Manag 2009;162:279–88. https://doi.org/10.1680/wama.2009.00061.

[49] Larras J. Profondeurs Maximales d'Erosion des Fonds Mobiles Autour des Piles en Rivere. Ann Ponts Chaussees 1963;133:411–24.

[50] Breusers HNC. Scouring around drilling platforms. J Hydraul Res IAHR, Bull 1965;19:276.

[51] Hancu S. Sur le calcul des affouillements locaux dams la zone des piles des ponts. Proc. 14th IAHR Congr. Paris, Fr., vol. 3, 1971, p. 299–313.

[52] Melville BW. Pier and Abutment Scour: Integrated Approach. J Hydraul Eng 1997;123:125–36. https://doi.org/10.1061/(ASCE)0733-9429(1997)123:2(125).

[53] Chabert J. Etude des affouillements autour des piles de ponts. Rep Natl Hydraul Lab, Chatou 1956.

[54] Norman VW. Scour at selected bridge sites in Alaska. vol. 32. US Geological Survey, Water Resources Division; 1975.

[55] Jain SC, Fischer EE. Scour around bridge piers at high Froude numbers,‖ Federal Highway Administration. US Dep Transp Washington, DC 1979.

[56] Chee RKW. Live-bed scour at bridge piers. Publ Auckl Univ New Zeal 1982.

[57] Chiew YM. Local scour at bridge piers 1984.

[58] Butch GK. Measurement of bridge scour at selected sites in New York, excluding Long Island. vol. 91. Department of the Interior, US Geological Survey; 1991.

[59] Wilson Jr K V. Scour at selected bridge sites in Mississippi. No. 94-4241. US Geological Survey; Earth Science Information Center, Open-File Reports; 1995.

[60] US Geological Survey. National bridge scour database, accessed April 15, 2014 2001. https://water.usgs.gov/osw/techniques/bs/BSDMS/index.htm.

[61] Sheppard DM, Miller W. Live-Bed Local Pier Scour Experiments. J Hydraul Eng 2006;132:635–42. https://doi.org/10.1061/(ASCE)0733-9429(2006)132:7(635).

[62] Holnbeck SR. Investigation of Pier Scour in Coarse-Bed Streams in Montana, 2001 through 2007 2011.

[63] Kothyari UC. Scour around bridge piers. University of Roorkee; 1989.

[64] Daneshfaraz R, Bagherzadeh M, Esmaeeli R, Norouzi R, Abraham J. Study of the performance of support vector machine for predicting vertical drop hydraulic parameters in the presence of dual horizontal screens. Water Supply 2020;21:217–31. https://doi.org/10.2166/ws.2020.279.

[65] Dasineh M, Ghaderi A, Bagherzadeh M, Ahmadi M, Kuriqi A. Prediction of Hydraulic Jumps on a Triangular Bed Roughness Using Numerical Modeling and Soft Computing Methods. Mathematics 2021;9. https://doi.org/10.3390/math9233135.

[66] Willmott CJ. Some Comments on the Evaluation of Model Performance. Bull Am Meteorol Soc 1982;63:1309–13. https://doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2.

[67] Nash JE. River flow forecasting through conceptual models, I: A discussion of principles. J Hydrol 1970;10:398–409.

[68] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software. ACM SIGKDD Explor Newsl 2009;11:10–8. https://doi.org/10.1145/1656274.1656278.

[69] Butte NF, Wong WW, Adolph AL, Puyau MR, Vohra FA, Zakeri IF. Validation of Cross-Sectional Time Series and Multivariate Adaptive Regression Splines Models for the Prediction of Energy Expenditure in Children and Adolescents Using Doubly Labeled Water. J Nutr 2010;140:1516–23. https://doi.org/10.3945/jn.109.120162.

[70] Friedman JH, Roosen CB. An introduction to multivariate adaptive regression splines. Stat Methods Med Res 1995;4:197–217. https://doi.org/10.1177/096228029500400303.