



Sharif University of Technology
Scientia Iranica
Transactions A: Civil Engineering
<http://scientiairanica.sharif.edu>



Development of predictive models for shear strength of HSC slender beams without web reinforcement using machine-learning based techniques

A. Kaveh^{a,*}, T. Bakhshpoori^b, and S.M. Hamze-Ziabari^c

a. *Centre of Excellence for Fundamental Studies in Structural Engineering, Iran University of Science and Technology, Narmak, Tehran, P.O. Box 16846-13114, Iran.*

b. *Department of Civil Engineering, Faculty of Technology and Engineering, University of Guilan, Rudsar-Vajargah, Iran.*

c. *School of Civil Engineering, Iran University of Science and Technology, Narmak, Tehran, P.O. Box 16846-13114, Iran.*

Received 30 March 2017; received in revised form 20 July 2017; accepted 11 September 2017

KEYWORDS

High Strength
Concrete (HSC);
Slender beams;
Shear strength;
Machine learning;
GMDH;
MARS.

Abstract. Shear failure of slender beams made of High Strength Concrete (HSC) is one of the most crucial failures in designing reinforced concrete members. The accuracy of the existing design codes for HSC, unlike the Normal Strength Concrete (NSC) beams, seems to be limited in predicting shear capacity. This paper proposes a new set of shear strength models for HSC slender beams without web reinforcement using conventional multiple linear regression, advanced machine learning methods of Multivariate Adaptive Regression Splines (MARS), and Group Method of Data Handling (GMDH) network. In order to achieve high-fidelity and robust regression models, this study employs a comprehensive database including 250 experimental tests. Various influencing parameters, including the longitudinal steel ratio, shear span-to-depth ratio, compressive strength of concrete, size of the beam specimens, and size of coarse aggregate, are considered. The results indicate that the MARS approach has the best estimation in terms of both accuracy and safety aspects in comparison with regression methods and GMDH approach. Moreover, the accuracy and safety of predictions of MARS model is also remarkably more than the most common design equations. Furthermore, the robustness of the proposed models is confirmed through sensitivity and parametric analyses.

© 2019 Sharif University of Technology. All rights reserved.

1. Introduction

In the last 30 years, the application of High Strength Concrete (HSC) has remarkably increased due to significant development in building and materials technology. Several advantages, including better mechanical

properties, structural efficiency, and economic benefits, in terms of both structural and aesthetic aspects have made the HSC a widely used material in the construction of high-rise buildings and long-span bridges in all over the world. The definition of HSC has changed parallel to the development of its properties in recent years. HSC can be defined as a concrete with compressive strength which is significantly beyond and more than double from what is used in normal practice (normal strength concrete, NSC). Different design codes and researchers proposed different limits for compressive strength of concrete to set a demarcation line between NSC and HSC materials. The present study follows

*. *Corresponding author. Tel.: +98-21-77240104;
Fax: +98-21-77240398
E-mail addresses: alikaveh@iust.ac.ir (A. Kaveh);
tbakhshpoori@guilan.ac.ir (T. Bakhshpoori);
ziabari@civileng.iust.ac.ir (S.M. Hamze-Ziabari)*

the definition of HSC by ACI 363R-10 [1], which is defined as a concrete with compressive strength more than 40 MPa.

One of the most crucial failures in reinforced structures made of HSC or NSC materials is the shear failure. Failure due to the shear in beams with HSC is brittle and occurs suddenly without warning. Several researchers have investigated experimentally and analytically to evaluate shear carrying capacity of reinforced concrete members. It has been found that several mechanisms, including the shear transfer in the compression zone, aggregate interlock across crack face, stirrups crossing the shear crack, and dowel action of longitudinal reinforcing bars crossing the crack, can be involved in providing shear resistance of Reinforced Concrete (RC) beams [2]. In beams without web reinforcing stirrups, crossing the shear crack mechanism cannot be considered as an internal force that contributes to shear resistance. It is understood that the shear failure mechanism in concrete members reinforced longitudinally, yet without transverse reinforcement under bending, varies significantly, and the size of member and shear span-to-depth ratio are the two main parameters that influence the shear failure mechanism. In fact, the shear span-to-depth ratio (a/d) determines the failure modes of RC beams and divides them into deep beams (< 2.5) and slender beams (> 2.5). For deep beams with a/d approximately less than 2.5, arch action occurs [3]. For slender beams with a/d greater than 2.5, the shear strength of reinforced member with longitudinal steel bars is a function of a/d , too. For these members, a/d represents the interacting effect of the moment and shear at a section on the shear strength. Accurate prediction of the shear behavior of the RC beams, unlike flexural behavior that can generally be predicted well, is a challenging work due to the complexity of the shear transfer mechanism.

Furthermore, there are remarkable differences between NSC and HSC beams without web reinforcement [4-10]. According to reports, the fracture surface in HSC beams is smooth and develops along the transition zone between the matrix and aggregates, whereas it is rough in NSC beams. In addition, internal shear transferring mechanisms contribute differently to different concrete strengths. At higher concrete strength, the aggregate interlocking does not contribute greatly because of the smooth fracture planes and the straight cracks, which do not go around the aggregate particles [11,12].

Shear strength prediction of HSC slender beams without web reinforcement is still contentious. There is no unified rational theory explaining the interaction of the three internal forces contributing to shear resistance, especially for HSC. Various design equations are available to estimate shear strength of reinforced

concrete beams without stirrups. Their accuracy, however, seems limited for slender reinforced HSC beams, as these equations are empirically developed using predefined forms with experimental data, generated mainly for a limited number of influencing parameters.

Recently, machine-learning approaches have been successfully employed to overcome these limitations in many different problems of civil engineering (e.g., [13-25]). In this regard, particularly, Elsanadedy et al. [12] very recently applied regression models and Artificial Neural Networks (ANN) to predict the shear strength of HSC slender RC beams without stirrups. Their results showed that ANN approach generally provided better predictions than regression approach did, in terms of accuracy. However, ANN implementation suffers from several drawbacks. The major disadvantage of this approach is that the ANN training process is achieved through a gradient descent algorithm on the error space, which can be very complex and may contain many local minima. Furthermore, trial-and-error processes are required to determine the network structure. In addition, ANN method does not give enough insight into the generated models and is not as easy to be used as the empirical formulas.

Considering these drawbacks and the existing unsatisfactory equations, Kaveh et al. [26] very recently successfully used M5' algorithm as one of the model-tree based algorithms used in developing predictive and simple formulas to estimate the shear strength of HSC slender beams without stirrups. This paper illustrates the potential of two more efficient alternative machine-learning-based approaches to this problem. Among soft computing approaches, the Group Method of Data Handling (GMDH) network and (multivariate adaptive regression splines) MARS algorithms are known as self-organized and non-parametric methods to model and discover the behaviors of unknown or complicated systems based on given input-output data points [27-29]. The main advantage of GMDH and MARS methods in comparison with ANN method is that the dependencies between input parameters and output parameter are represented in parametric form as an equation, while these dependencies are hidden within neural network structures in ANN method.

To develop new predictive models based on GMDH and MARS, a comprehensive existing database was employed very recently, used by Elsanadedy et al. [12], that contained 250 experimental tests. The developed GMDH model related shear capacity to the longitudinal steel ratio, the shear span-to-depth ratio, compressive strength of concrete, the size of the beam specimens, and the size of coarse aggregate. The results of the developed GMDH and MARS were compared to the most common existing equations and to multiple nonlinear regression (MNLR) approach developed in this study through statistical error indicators. Results

confirmed that the developed GMDH and MARS models outperformed the existing equations and the developed regression-based equations in terms of both accuracy and safety aspects. The relative importance of significant parameters dealing with shear strength was also investigated through sensitivity analysis. The robustness of the proposed models was also verified through a parametric analysis.

This paper is outlined as follows: The existing design equations and their limitations are given in the following section. The GMDH and MARS algorithms are described next in Section 3. The applied dataset and modeling process are presented in Section 4 followed by the results and discussion section. Finally, the summary and conclusion are presented in Section 6.

2. Design equations for shear strength

2.1. Design equations

The existing equations for estimating shear capacity of RC beams (V) are presented in Eqs. (1) to (8). A significant gap exists in the process of selecting the main parameters and their effects on V , because these equations have been empirically derived. In the following equations b is beam width; d is the effective depth of beam; h beam depth; d_v effective shear depth; f'_c specified compressive strength of concrete; f_{ck} characteristic compressive cylinder strength of concrete; a shear span; a_g specified nominal maximum size of coarse aggregate; E_s modulus of elasticity of steel; A_s area of tension steel; ρ longitudinal steel ratio; M_u and V_u factored moment and shear force, respectively; M_f and V_f factored moment and shear force, respectively; M_{Ed} and V_{Ed} design moment and shear force; λ factor accounting for concrete density (units are N and mm).

ACI 318-11 [30]:

$$V = \left(0.16\sqrt{f'_c} + 17\rho\frac{V_u d}{M_u} \right) bd \leq 0.29\sqrt{f'_c}bd. \quad (1)$$

CSA A23.3-04 [31]:

$$V = \lambda\beta\sqrt{f'_c}bd_v,$$

where:

$$\begin{aligned} \beta &= \frac{0.4}{(1 + 500\varepsilon_x)} \times \frac{1300}{(1000 + s_{ze})}, \\ \varepsilon_x &= \frac{M_f/d_v + V_f}{2E_s A_s} \leq 0.003, \\ s_{ze} &= \frac{35d_v}{15 + a_g} \geq 0.85d_v. \end{aligned} \quad (2)$$

Fib Model Code [32]:

$$V = k_v\sqrt{f_{ck}}bz$$

where:

$$k_v = \frac{0.4}{1 + 1500\varepsilon_x} \times \frac{1300}{1000 + k_{dg}z},$$

$$\varepsilon_x = \frac{M_{Ed}/z + V_{Ed}}{2E_s A_s} \leq 0.003,$$

$$k_{dg} = \frac{32}{16 + a_g} \geq 0.75,$$

$$z = 0.9d. \quad (3)$$

Eurocode-2 [33]:

$$V = 0.18k\sqrt[3]{100\rho f_{ck}}bd \geq 0.035\sqrt{k^3 f_{ck}}bd,$$

where:

$$k = 1 + \sqrt{\frac{200}{d}} \leq 2. \quad (4)$$

CEB-FIP Model Code [34]:

$$V = 0.15\sqrt[3]{\frac{3}{a/d}}\zeta\sqrt[3]{100\rho f_{ck}}bd,$$

where:

$$\zeta = 1 + \sqrt{\frac{200}{d}} \quad (5)$$

AS 3600-2009 [35]:

$$V = \beta_1\beta_2\beta_3\sqrt[3]{\rho}f_{cv}bd,$$

where:

$$\beta_1 = 1.1 \left(1.6 - \frac{d}{1000} \right),$$

$$\beta_2 = 1.0 \text{ for beams without axial force,}$$

$$1 \leq \beta_3 = \frac{2}{a/d} \leq 2,$$

$$f_{cv} = \sqrt[3]{f_{ck}} \leq 4 \text{ MPa.} \quad (6)$$

JSCE Guidelines [36]:

$$V = \beta_d\beta_p f_{vcd}bd,$$

$$\beta_d = \sqrt[4]{\frac{1000}{d}} \leq 1.5, \quad \beta_p = \sqrt[3]{100\rho} \leq 1.5,$$

$$f_{vcd} = 0.2\sqrt[3]{f_{ck}} \leq 0.72 \text{ MPa.} \quad (7)$$

Cladera and Mari [37]:

$$V = 0.225\zeta\sqrt{100\rho}f'_c{}^{0.2}bd,$$

where:

$$\zeta = 1 + \sqrt{\frac{200}{0.9d}} \leq 2.75,$$

$$\rho \leq 0.02 \left(1 + \frac{f'_c}{100} \right). \quad (8)$$

Table 1. Details of existing design equations and their performances in predicting shear strength of slender HSC beams.

Equations or models	Eqs.	Design parameters								Statistical parameters on $V_{\text{measured}}/V_{\text{predicted}}$		
		b	d	A	a_g	f'_c		ρ (%)		M	STD	COV
						Limitation	Limitation					
ACI 318-11	(1)	O	O	—	—	O	≤ 65 MPa	O	—	1.28	0.48	0.23
CSA A23.3-04	(2)	O	O	—	O	O	≤ 64 MPa	—	—	1.62	0.61	0.38
fib Model Code	(3)	O	O	—	O	O	≤ 64 MPa	—	—	1.59	0.60	0.36
Eurocode-2	(4)	O	O	—	—	O	≤ 90 MPa	O	≤ 0.02	1.09	0.35	0.12
CEB-FIP Model	(5)	O	O	O	—	O	—	O	—	1.24	0.29	0.08
AS 3600-2009	(6)	O	O	O	—	O	≤ 64 MPa	O	—	1.12	0.28	0.08
JSCE Guidelines	(7)	O	O	—	—	O	≤ 47 MPa	O	≤ 0.03	1.35	0.38	0.14
Cladera and Mari [37]	(8)	O	O	—	—	O	≤ 60 MPa	O	$\leq g(f'_c)$	1.27	0.31	0.10
GMDH (proposed)	(19)	O	O	O	O	O	—	O	—	1.01	0.17	0.03
MARS (proposed)	(20)	O	O	O	O	O	—	O	—	1.02	0.10	0.01

2.2. Limitations

Table 1 summarizes the main design parameters included in Eqs. (1) to (8). According to this table, all design equations put maximum limitation on f'_c except CEB-FIP [34] model. Most of these maximum limitations are around 65 MPa, while f'_c for HSC slender beams without web reinforcement has been recorded notably more than the aforementioned value in field or laboratory. The statistical error parameters in terms of average (M), Standard Deviation (STD), and Coefficient Of Variation (COV) between the standardized i th predicted and measured values of V are presented. It can be expected that the performance of these equations for slender beams made of HSC be limited, as is clear in Table 1.

In addition, there is inconsistency between the existing equations for predicting shear strength. For example, the portions of longitudinal steel reinforcement contributing to V are given by $\rho^{1/3}$ in equations of Eurocode-2 [33], CEB-FIP [34] model code, AS 3600 [35], and JSCE [36], whereas its contributions to study of Cladera and Mari [37] and ACI 318-11 [30] are proportional to $\rho^{1/2}$ and ρ , respectively. Furthermore, some codes (Eurocode-2 [33], CEB-FIP [34] model, AS 3600 [35], and JSCE [36]) consider the shear strength of HSC beams proportional to $f'_c{}^{1/3}$, while other codes (ACI 318-11 [30], CSA A23.3-04 [31], and fib model [32]) consider it to be proportional to $f'_c{}^{1/2}$. This proportion was also reported as $f'_c k k k_c^{0.2}$ by Cladera and Mari [37].

Furthermore, according to Table 1, some important parameters were not incorporated in most of these design codes. For example, the influence of the nominal

maximum size of coarse aggregate only was considered in CSA A23.3-04 [31] and level II approximation of the fib model [32] code, whereas dependency of HSC shear strength on this parameter was reported by Muttoni [38]. Only CEB-FIP [34] model and AS 3600 [35] consider the effect of shear span-to-depth ratio on shear strength of slender HSC beams. In addition, ACI 318-11 [30] does not consider the effect of depth, d , whereas other design equations show notable sensitivity to the change of parameter d .

To take full advantage of the HSC RC beam, generating new models to predict the shear capacity to obtain both reliability and accuracy is indispensable. Recently, machine-learning approaches have been widely applied to engineering problems. The main advantage of these approaches can be mentioned as they considered all possible relations between input and output parameters and checked different combinations of input parameters for estimating output parameter, unlike the regression approaches. In the following section, Multivariate Adaptive Regression Splines (MARS), Group Method of Data Handling (GMDH) network, and Multiple Nonlinear Regression methods (MNL) are introduced in this study to predict the shear strength of HSC slender beams without web reinforcement.

3. Soft computing approaches

Recently, the soft computing approaches, such as Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS), have been employed as useful tools for modeling and forecasting complex structural engineering problems. Among these methods, the Group Method of Data Handling (GMDH)

network and Multivariate Adaptive Regression Splines (MARS) have not yet been applied widely in structural engineering. The main advantage of GMDH and MARS methods in comparison with ANN method is that the dependency between input parameters and responses is represented in parametric form as an equation, while these dependencies are hidden within neural network structures in ANN method. Besides, ANN methods need an essential time for learning; therefore, it is difficult to apply it for modeling and forecasting under real-time systems. The descriptions of the GMDH network, the MARS, and Multiple Nonlinear Regression methods (MNLRL) are discussed as follows.

3.1. Multivariate adaptive regression splines

Multivariate Adaptive Regression Splines (MARS) is a nonlinear and nonparametric intelligent computing regression algorithm that models the nonlinear responses between the inputs and outputs of a system using a series of piecewise linear or cubic segments (splines). The resulting piecewise equation is known as Basis Functions (BFs). The slope of regression function is allowed to change from one segment to the next. The end points of each segment are called knots. A knot marks the end of one region of data and the beginning of another. Unlike the well-known parametric linear regression analysis, MARS provides greater flexibility to explore nonlinear relation between a response variable and predictor variables. In addition, MARS also searches for possible interactions between variables by checking all degrees of interactions. Because it allows for all functional forms and interactions, MARS is able to track the complex data structures from high-dimensional datasets. The general MARS function can be expressed using the following equation:

$$\tilde{f}(x) = \beta_0 + \sum_{m=1}^M \beta_m \lambda_m(x), \quad (9)$$

where $\tilde{f}(x)$ is the predicted response, β_0 and β_m are constants, which are estimated to yield the best data fit, and M is the number of basis functions included into the model. The basis function in MARS model can be either one single spline function, or a product of two or more spline functions for different predictor variables. The spline basis function, $\lambda_m(x)$, can be specified as follows:

$$\lambda_m(x) = \prod_{k=1}^{k_m} [s_{km}(x_{v(k,m)} - t_{k,m})], \quad (10)$$

where k_m is the number of knots, s_{km} takes either 1 or -1 and indicates the right/left regions of the associated step function, $v(k, m)$ is the label of the predictor variable, and $t_{k,m}$ is the knot location. MARS generates

basis functions by searching in a stepwise way. An adaptive regression algorithm is used to select the knot locations. An optimal MARS is developed through a two-stage forward and backward procedure. In the forward stage, MARS overfits data by considering a great number of basis functions. In the backward stage, to avoid overfitting, redundant basis functions are deleted from Eq. (9). MARS adopts Generalized Cross-Validation (GCV) to delete the redundant basis functions. The expression of GCV is given as follows:

$$\text{GCV} = \frac{\frac{1}{N} \sum_{i=1}^N [y_i - \hat{f}(x_i)]^2}{\left[1 - \frac{C(B)}{N}\right]^2}, \quad (11)$$

in which N is the number of data, and $C(B)$ is a complexity penalty that increases with the number of Basis Function (BF) in the model. It is defined as follows:

$$C(B) = (B + 1) + dB, \quad (12)$$

where d is a penalty for each BF included into the model, and B is the number of basis functions [24].

3.2. Group method of data handling

Group Method of Data Handling (GMDH) is a learning machine based on the polynomial theory of complex systems [22]. From this network, the most significant input parameters, the number of layers, the number of neurons in middle layers, and optimal topology design of the network are defined automatically. The structure of the GMDH network is configured through the training stage with polynomial model, which produces the minimum error between the predicted value and observed output. The formal definition of the system identification problem is to find an approximate function, \hat{f} , that can be used to predict actual output, \hat{y} , or a given input vector, $X = (x_1, x_2, \dots, x_n)$, as close as possible to actual output, y . Therefore, n observations of multi-input-single-output data pairs are considered as follows:

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}), \quad i = 1, 2, \dots, M. \quad (13)$$

The general relationship between input and output variables can be expressed by a complicated discrete form of the Volterra function as a series in the form of:

$$y = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=1}^n w_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n w_{ijk} x_i x_j x_k + \dots, \quad (14)$$

which is known as the Kolmogorov-Gabor polynomial [39]. In the present study, a quadratic polynomial of the GMDH network is used that is written as:

$$\text{Quadratic: } \hat{y} = w_0 + w_1x_i + w_2x_j + w_3x_ix_j + w_4x_i^2 + w_5x_j^2. \quad (15)$$

The weighting coefficients of Eq. (14) are calculated using regression techniques such that the difference between actual output, y , and the calculated value, \hat{y} , for each pair of x_i and x_j as input variables is minimized. In this way, the weighting coefficients of quadratic function G_i are obtained to optimally fit the output to the whole set of input-output data pairs, defined as follows:

$$E = \frac{\sum_{i=1}^M [y_i - G_i()]^2}{M} \rightarrow \min. \quad (16)$$

In this study, the GMDH network is improved using a back propagation algorithm. This method includes two main steps:

1. Weighting coefficients of the quadratic polynomial are determined using the least squares method from the input layer to output layer in the form of a forward path;
2. Weighting coefficients are updated using a back-propagation algorithm in a backward path. This procedure may continue until the error of the training network (E) is minimized.

3.3. Multiple nonlinear regression method

Let y be a dependent variable and have a nonlinear relation with n independent variables as x_1, x_2, \dots, x_n . The nonlinear relation between them can be expressed as:

$$y = a_0x_1^{a_1}x_2^{a_2}\dots x_n^{a_n}. \quad (17)$$

By applying logarithmic transformation, the following equation can be obtained as follows:

$$\log y = \log a_0 + a_1 \log x_1 + a_2 \log x_2 + \dots + a_n \log x_n, \quad (18)$$

where coefficients a_0, a_1, \dots, a_n can be determined by applying the least squares method, similar to Multi-Linear Regression (MLR) method.

4. Model development

4.1. Model inputs and outputs

Five variables were presented to the MARS, GMDH, and MNLR as model inputs including the effective depth (d), the shear span-to-depth ratio (a/d), the compressive strength of concrete (f'_c), the aggregate size-to-depth ratio (a_g/d), and the longitudinal steel ratio (ρ). The single model output is concrete shear capacity, v_u .

4.2. Data division and pre-processing

The data used to calibrate and validate MARS, GMDH, and MNLR models obtained from the literature include a series of 250 concrete shear capacity tests, compiled by Elsanadey et al. [12] from 33 experimental studies performed between 1957 and 2013. All selected beams are without stirrups, and they were longitudinally reinforced with non-prestressed steel rebars. Only slender beams with $a/d \geq 2.5$, whose failures primarily occur due to shear, were considered. The specimens were monotonically loaded using either one or two concentrated loads until failure. Details of the data used can be found in [12].

The available data were randomly divided into two sets: a training set for model calibration and an independent validation set for model verification. The validation dataset was used to specify the generalization capability of the models to new data with which they had not been trained. In other words, the testing data were applied to measure the performance of the models obtained by the proposed algorithms when applied to dataset, which played no role in building the models. The statistics of the data used in the training and validation sets are given in Table 2, which include the mean, standard deviation (STD), minimum, and maximum. For more visualization, a matrix-plot of input and output parameters is represented in Figure 1. This plot presents all possible scatter plots of input and output parameters one by one. The plots in the diagonal of this matrix are the histograms of input and output parameters for the whole data points. As shown, most of data points are concentrated in the ranges of $d \leq 400$ mm, $f'_c \leq$

Table 2. Ranges of input and output parameters for training and testing sets.

Variable	Subset	Min	Max	Mean	STD
d (mm)	Training	133	925	284.88	158.94
	Testing	135	718	255.98	96.99
a/d	Training	2.46	6.1	3.59	0.86
	Testing	2.47	6.1	3.66	0.89
a_g/d	Training	0.01	0.17	0.06	0.02
	Testing	0.02	0.17	0.06	0.02
f'_c (MPa)	Training	42.5	183	65.40	21.44
	Testing	45.3	155	64.56	20.75
ρ (%)	Training	0.33	6.64	2.17	1.31
	Testing	0.33	6.64	2.29	1.26
V (kN)	Training	0.48	4.08	1.72	0.78
	Testing	0.57	3.85	1.8	0.70

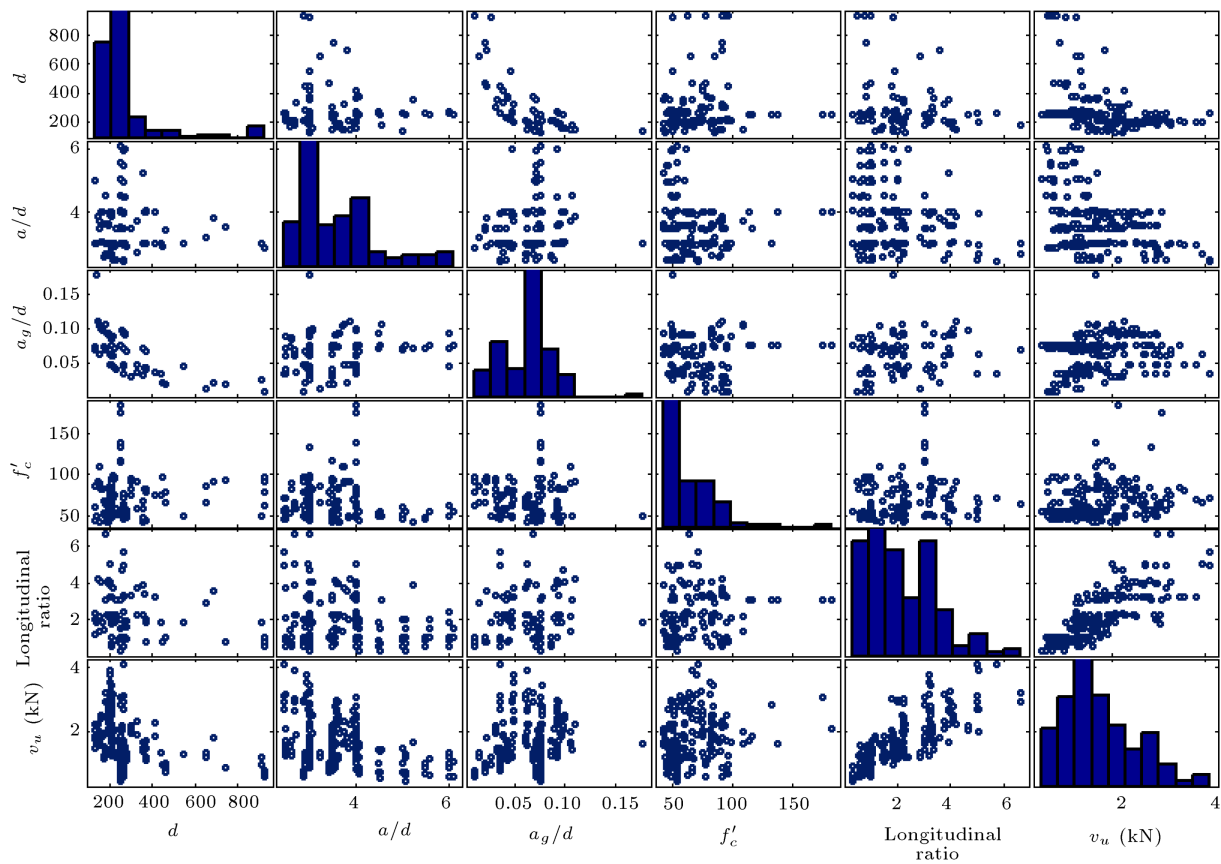


Figure 1. Matrix plot of input and output parameters.

100 MPa, and $\rho \leq 4$ (%). It should be noted that the derived model could be more reliable in these ranges.

Out of the 250 tests, 200 data vectors (80%) were taken for the training process. The remaining 50 data vectors (20%) were used to test the models. It should be noted that, like other empirical models, the proposed models are only valid for the ranges in which they are trained and show better performance in those ranges in which data are denser.

4.3. Derived models

Following data division, training dataset was presented to GMDH, MARS, and MNLR algorithms for model training. The MARS model returned the following equation for concrete shear capacity:

$$\begin{aligned}
 v_u \text{ (kN)} = & 5.4 - 3.5 \times \text{BF}_1 + 22 \times \text{BF}_2 + 41 \\
 & \times \text{BF}_3 - 0.079 \times \text{BF}_4 + 0.078 \times \text{BF}_5 - 0.049 \\
 & \times \text{BF}_6 + 0.054 \times \text{BF}_7 - 0.54 \times \text{BF}_8 + 38 \\
 & \times \text{BF}_9 + 0.096 \times \text{BF}_{10} - 0.47 \times \text{BF}_{11} - 3.6 \\
 & \times \text{BF}_{12} + 31 \times \text{BF}_{13} + 700 \times \text{BF}_{14} - 4100
 \end{aligned}$$

$$\times \text{BF}_{15} - 16 \times \text{BF}_{16} - 15 \times \text{BF}_{17} + 1.7$$

$$\times \text{BF}_{18} + 4.3 \times \text{BF}_{19} - 0.09 \times \text{BF}_{20} - 0.79$$

$$\times \text{BF}_{21} - 23 \times \text{BF}_{22} - 0.15 \times \text{BF}_{23} + 0.48$$

$$\times \text{BF}_{24} + 17 \times \text{BF}_{25} + 0.093 \times \text{BF}_{26} + 6.8 \times 10^{-5}$$

$$\times \text{BF}_{27} + 2.4 \times \text{BF}_{28}. \quad (19)$$

Table 3 lists the BFs and their corresponding equations. It should be noted that, based on Table 3, 28 BFs and 19 BFs with interaction terms are integrated in this model, indicating that the model is not simply additive and that interaction terms play a significantly important role. The final model of MARS (Eq. (19)) is achieved via GCV obtained from forward selection and backward deletion process. As observed, one of the advantages of MARS algorithm is that it not only captures complex relationships between independent and dependent variables, but also does not require additional effort to verify a priori assumption about the relationship between the set of independent variables and dependent response variable. The latter feature becomes more important as the dimension of the problem increases.

Table 3. Basis functions of the developed MARS model.

Basis functions	Equations
BF ₁	$\max(0, \rho - 3.2)$
BF ₂	$\max(0, 3.2 - \rho)$
BF ₃	$\max(0, 2.5 - a/d)$
BF ₄	$\max(0, d - 190)$
BF ₅	$\max(0, d - 230)$
BF ₆	$\max(0, 230 - d)$
BF ₇	$\max(0, 170 - d)$
BF ₈	$\max(0, a/d - 2.5) \times \max(0, 1.4 - \rho)$
BF ₉	$\max(0, 230 - d) \times \max(0, 0.061 - a_g/d) \times \max(0, 3.9 - a/d)$
BF ₁₀	$\max(0, 360 - d) \times \max(0, 0.077 - ag/d)$
BF ₁₁	$\text{BF}_{10} \times \max(0, \rho - 4)$
BF ₁₂	$\text{BF}_9 \times \max(0, 60 - f'_c)$
BF ₁₃	$\max(0, 230 - d) \times \max(0, 0.061 - a_g/d) \times \max(0, a/d - 3)$
BF ₁₄	$\max(0, 230 - d) \times \max(0, 0.061 - a_g/d) \times \max(0, 3 - a/d)$
BF ₁₅	$\max(0, 230 - d) \times \max(0, 0.061 - a_g/d) \times \max(0, \rho - 3.2)$
BF ₁₆	$\max(0, 230 - d) \times \max(0, 0.061 - a_g/d) \times \max(0, 3.2 - \rho)$
BF ₁₇	$\max(0, 230 - d) \times \max(0, \rho - 1.4) \times \max(0, 0.063 - x3)$
BF ₁₈	$\text{BF}_5 \times \max(0, a_g/d - 0.068)$
BF ₁₉	$\max(0, \rho - 3.4)$
BF ₂₀	$\max(0, 310 - d) \times \max(0, 0.076 - a_g/d) \times \max(0, 6 - a/d)$
BF ₂₁	$\max(0, d - 170) \times \max(0, a_g/d - 0.071)$
BF ₂₂	$\max(0, 3.2 - \rho)$
BF ₂₃	$\max(0, a/d - 2.5) \times \max(0, \rho - 1)$
BF ₂₄	$\max(0, a/d - 2.5) \times \max(0, 1 - \rho)$
BF ₂₅	$\max(0, 230 - d) \times \max(0, a_g/d - 0.061) \times \max(0, \rho - 3.2) \times \max(0, 3 - a/d) \times \max(0, f'_c - 64)$
BF ₂₆	$\max(0, d - 170) \times \max(0, 0.071 - a_g/d) \times \max(0, \rho - 3.3)$
BF ₂₇	$\max(0, 360 - d) \times \max(0, f'_c - 70)$
BF ₂₈	$\text{BF}_3 \times \max(0, f'_c - 54)$

GMDH algorithm returned the following selective polynomials to predict concrete shear capacity of HSC slender beams:

Layer 1:

$$\begin{aligned}
 L_1 = & -1.7 + 1.3\rho + 1.1 \left(\frac{a}{d}\right) + 0.02d \\
 & + 0.27 \left(\frac{a}{d}\right) \rho - 0.004d\rho - 0.009d \left(\frac{a}{d}\right) - 0.18\rho^2 \\
 & - 0.18 \left(\frac{a}{d}\right)^2 - 1.5 \times 10^{-5} d^2 - 0.00041d \left(\frac{a}{d}\right) \rho \\
 & - 0.026 \left(\frac{a}{d}\right) \rho^2 - 0.016 \left(\frac{a}{d}\right)^2 \rho + 0.00059d\rho^2 \\
 & + 0.0014d \left(\frac{a}{d}\right)^2 + 2.9 \times 10^{-6} d^2 \rho
 \end{aligned}$$

$$\begin{aligned}
 & + 2 \times 10^{-6} d^2 \left(\frac{a}{d}\right) + 0.0067\rho^3 - 0.0094 \left(\frac{a}{d}\right)^3 \\
 & + 4.3 \times 10^{-9} d^3.
 \end{aligned}
 \tag{20a}$$

Layer 2:

$$\begin{aligned}
 v_u \text{ (kN)} = & 2.5 - 1.4L_1 - 0.03f'_c - 25 \left(\frac{a_g}{d}\right) \\
 & + 0.0017f'_c L_1 + 42 \left(\frac{a_g}{d}\right) L_1 + 0.7 \left(\frac{a_g}{d}\right) f'_c \\
 & + 0.42L_1^2 + 7.9 \times 10^{-5} f'^2_c - 4.1 \times 10^2 \left(\frac{a_g}{d}\right)^2 \\
 & - 0.3 \left(\frac{a_g}{d}\right) f'_c L_1 + 0.0095f'_c L_1^2 \\
 & - 6.7 \times 10^{-5} f'^2_c L_1 - 7.1 \left(\frac{a_g}{d}\right) L_1^2
 \end{aligned}$$

$$-0.0039 \left(\frac{a_g}{d}\right) f_c'^2 + 6 \left(\frac{a_g}{d}\right)^2 L_1 + 2.6 \left(\frac{a_g}{d}\right)^2 f_c' \\ - 0.092 L_1^3 + 1.2 \times 10^{-6} f_c'^3 + 7200 \left(\frac{a_g}{d}\right)^3. \quad (20b)$$

To further evaluate the proposed MARS and GMDH models against the most common regression approaches, a new regression equation was developed using MNLR as follows:

$$v_u = 14.50d^{-0.39} \left(\frac{a}{d}\right)^{-0.62} f_c'^{0.12} \rho^{0.45}. \quad (21)$$

To analytically evaluate the performances of the developed models, the following statistical error parameters were applied: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), correlation coefficient (R), and coefficient of determination (R^2).

$$MAE = \frac{\sum_{i=1}^N |P_i - O_i|}{N}, \quad (22)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2}, \quad (23)$$

$$R = \frac{\sum_{i=1}^N (P_i - P_m)(O_i - O_m)}{\sqrt{(\sum_{i=1}^N (P_i - P_m)^2) \sum_{i=1}^N (O_i - O_m)^2}}, \quad (24)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - O_m)^2}, \quad (25)$$

where O_i is the measured value, P_i stands for the prediction values, N is the number of data points, O_m is the mean value for observation, and P_m is the mean value of prediction. The correlation coefficient (R) is a measure of the relative correlation between the predicted and measured values. R values ranged between -1 and 1 . If R value is close to 1 , it indicates that there is a direct linear relationship between measured and predicted values. However, R does not necessarily indicate the goodness of the model's performance, particularly when the range of data is very wide and the data points are distributed around their mean. Therefore, the coefficient of determination, R^2 , can be used as an unbiased estimate and, also, a better measure for evaluating the model's performance. The MAE and RMSE measure the difference between predicted and measured values, and values near to zero indicate a close match.

5. Results and discussion

The results of developed models and the most common

design equations to predict shear strength of HSC slender beams without web reinforcement were compared in this section. Moreover, the influence of important parameters on shear strength was investigated through parametric and sensitivity analyses. The most important predictive parameters were determined through sensitivity analysis in predicting shear strength. Finally, safety analysis was done by using demerit point's classifications of Collins.

5.1. Performance analysis

The number of data used to develop a new predictor model based on machine-learning approaches plays a crucial role in modeling process. For developing a reliable model based on data mining approaches, Frank and Todeschini [40] suggested that the minimum ratio between the number of data used and the number of involved variables should be 3 . A safer value of 5 can be more conservative. In the present study, this ratio is remarkably higher and is equal to $250/6 = 41.66$. The performances of the developed MARS and GMDH models shown in Figure 2(a) and (b) demonstrate that there are little scatter around optimal line between measured and predicted shear strengths for both training and testing datasets. For further verification of the developed models, analytical analysis of statistical error parameters for training and testing datasets is presented in Table 4. In addition, the results of the developed MNLR method in this study were compared with those of the other proposed ones.

Smith [41] suggested that if $|R| > 0.8$, it can be expected that there is a strong correlation between observed and predicted values. It should be recognized that even if R is close to 1 , the predicted and observed values may not match each other; they only tend to vary similarly. To compensate for this limitation, coefficient of determination R^2 can be used. To achieve precise results, R^2 values should be close to 1 . In

Table 4. Performance of GMDH and regression models in predicting bond strength for training and testing datasets.

Subset		MAE	RMSE	R	R^2
Training	MARS	0.1405	0.1849	0.9715	0.9438
	GMDH	0.2223	0.2983	0.9241	0.8537
	MNLR	0.2741	0.3626	0.8866	0.7839
Testing	MARS	0.1395	0.1902	0.9630	0.9253
	GMDH	0.2592	0.3670	0.8519	0.7217
	MNLR	0.2684	0.3831	0.8414	0.6966
Total	MARS	0.1403	0.1860	0.9699	0.9408
	GMDH	0.2297	0.3132	0.9122	0.8321
	MNLR	0.2729	0.3668	0.8789	0.7697

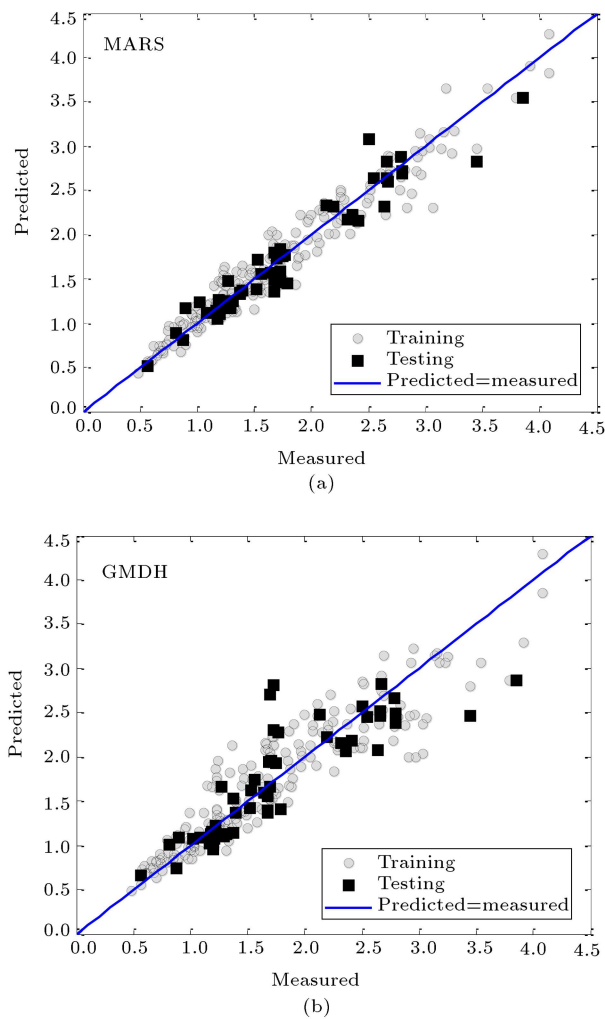


Figure 2. Comparison between measured and predicted shear strengths (V (kN)) by (a) MARS and (b) GMDH models for training and testing datasets.

all cases, RMSE and MAE should be close to zero. As shown in Table 4, the MARS model outperformed GMDH and MNLR models in terms of accuracy for

both training and testing datasets. For example, the MARS model decreased RMSE value by 40.6% and 49.2% and increased R^2 values by 13% and 22.2% with respect to GMDH and MNLR models, respectively. Of note, the performances of the developed GMDH and MNLR models were acceptable, and GMDH outperformed the MNLR model. Lower accuracy of MNLR can be attributed to its limitations in predicting the phenomenon with a highly nonlinear relationship or with multiple criteria.

The results of the developed models are also compared with those obtained from the widely available design codes and standards including ACI 318-11 [30], CSA A23.3-04 [31], fib model code [32], Eurocode-2 [33], CEB-FIP [34] model code, AS 3600-2009 [35], JSCE [36] guidelines, and Cladera and Mari [37] in Table 5. In general, the accuracy of design codes is remarkably limited in predicting the shear strength of HSC slender beams without web reinforcement, especially ACI 318-11, CSA A23.3-04, fib model code, and JSCE guidelines. In fact, these design codes, except CEB-FIP models, were not exclusively suggested for HSC slender beams without web reinforcement and had limitations on the maximum value of f'_c (ACI 318-11 [30], CSA A23.3-04 [31], fib model code [32], Eurocode-2 [33], model code, AS 3600-2009 [35], JSCE [36] guidelines, and Cladera and Mari [37]) or on the percentage of tension steel (Eurocode-2 [33] and Cladera and Mari [37]). Therefore, most of these equations cannot correctly be employed to predict the shear strength of HSC slender beams. Although AS 3600-2009 [35] model had the best performance amongst the other design codes, which had R^2 of 0.50 (highest) and RMSE of 0.54 MPa (lowest), this model performed significantly lower than the three developed models of GMDH, MARS, and MNLR did. For example, MARS model improved RMSE and R^2 by 65.5% and 88%, respectively, in comparison with the best design code. Thus, the

Table 5. Performance of existing equations for prediction of shear strength of slender HSC beams.

Model	MAE	RMSE	R	R^2
ACI31811	0.5891	0.7939	0.6676	-0.0787
CSAA23304	0.6976	0.9358	0.5289	-0.4988
Fib	0.6895	0.9254	0.5252	-0.4656
Eurocode2	0.4341	0.6161	0.7311	0.3503
CEBFIP	0.4273	0.6059	0.8498	0.3718
AS36002009	0.3758	0.5402	0.8241	0.5005
JSCE	0.5430	0.7525	0.8188	0.0307
Cladera and Mari [37]	0.4540	0.6399	0.8104	0.2992
Elsanadedy et al. [12]	0.7043	0.8346	0.8777	-0.1921
MNLR (this study)	0.2729	0.3668	0.8789	0.7697
GMDH (this study)	0.2297	0.3132	0.9122	0.8321
MARS (this study)	0.2729	0.3668	0.8789	0.7697

developed models can predict the target values of the shear strength of HSC slender beams without web reinforcement with acceptable accuracy and less error than the available design codes over a wide range of input variables.

To have a deeper understanding of the errors in design codes, the Discrepancy Ratio (DR) between measured and predicted shear strengths as an error indicator was plotted against the most effective parameters, including longitudinal steel ratio (ρ), f'_c , d , and a/d , for testing the dataset in Figure 3. The results of Eurocode-2 [33], CEP-FIP [34], and AS 3600-2009 [35], which had higher acceptable accuracy, were compared with those of GMDH and MARS models in this aspect. The best linear regression also fits prediction values of the mentioned models. Errors of an appropriate model should be independent of, or less sensitive to, the variation of the input parameters involved in that phenomenon. Otherwise, it can be interpreted that those input parameters are neither correctly incorporated or included in that model [42]. As shown in Figure 3, the errors of three design codes are sensitive to changes of ρ as if their prediction varies from overestimation status to underestimation status as ρ increases. It can be interpreted that this parameter is not correctly involved in their equations. On the

other hand, the errors of GMDH and MARS models are completely independent of this parameter.

According to this figure, the errors of all models are almost independent of parameters f'_c and d . Thus, it can be expected that these parameters be probably well defined in the mentioned models. It should be noted that the errors of Eurocode-2 [33] in variation of parameter d are more than those of other models. In addition, sensitivity of the errors of GMDH and MARS models to variations of d and f'_c is less than that of the three design codes. Figure 3 demonstrates the errors of the mentioned models as a function of parameter a/d . Of note, the predictions of design codes vary from underestimation status to overestimation status as a/d increases, and this inclination in CEB-FIP [34] model is less than the other two equations. As a result, these figures show that all input parameters are successfully included into the newly proposed models, and their errors are completely independent of these parameters.

5.2. Parametric and sensitivity analyses

To further examine the robustness of the developed GMDH and MARS models, a parametric analysis can be carried out. Parametric analysis investigates how closely the model's predictions agree with available structural knowledge and experimental data and with one's expectation. Therefore, parametric analysis was carried out with the aim of gaining a deeper understanding of concrete shear strength of HSC slender beams without web reinforcement. The parametric analysis investigates the response of the predicted shear strength from the proposed models to a set of hypothetical input data generated over the ranges of the minimum and maximum data used for the model training. The methodology was presented based on the change of only one input variable varied at a time, while the other variables were kept constant on their average in the applied datasets. A set of synthetic data for the single parameter was generated by increasing it incrementally. These inputs were presented for the prediction equation, and the shear strength was calculated. This procedure was repeated using another variable until the model's response was tested for all input variables.

The results of the parametric analysis and, also, the used experimental data are shown in Figure 4. It can be seen that the prediction behavior of concrete shear strength from GMDH and MARS models agrees well with the experimental results. Figure 4(a) presents the variations of concrete shear strength as a function of effective depth (d). As shown, the failure shear stress decreases with the increase of effective depth. This behavior is well documented in the literature for RC slender beams without web reinforcement. There are two common approaches and hypotheses to explain this behavior: one of these approaches highlights the

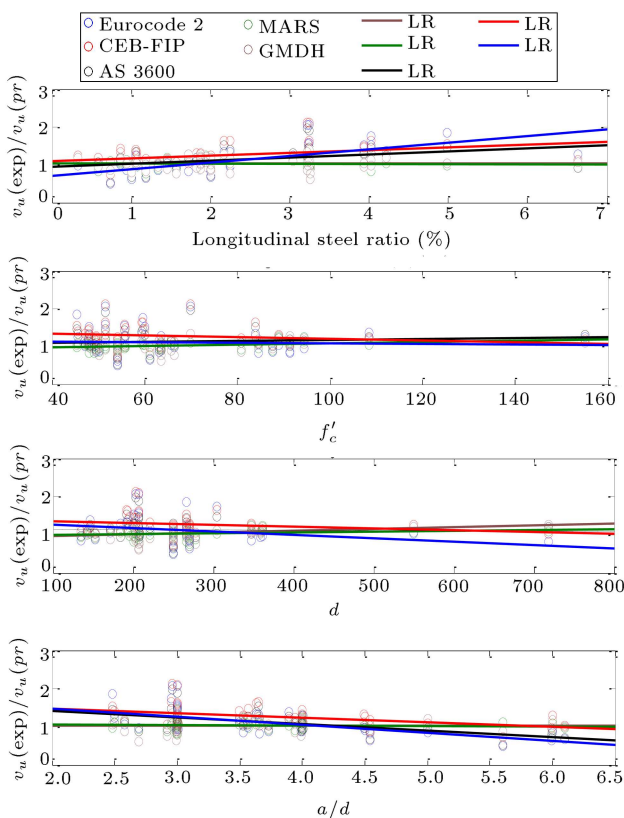


Figure 3. Variation of Discrepancy Ratio (DR) between measured and predicted shear strengths as a function of ρ , f'_c , d , and a/d .

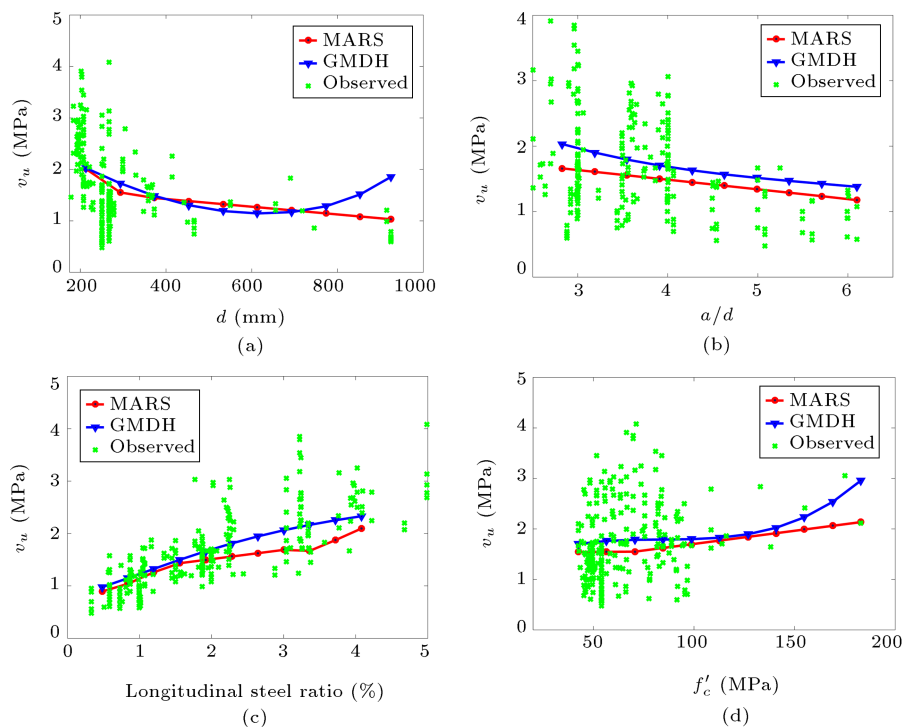


Figure 4. Shear strength parametric analysis in GMDH- and MARS-based models for different ranges of (a) d , (b) a/d , (c) ρ , and (d) f'_c .

role of aggregate interlock, which is reduced in larger members due to the development of wider cracks. The other common approach is based on fracture mechanic concepts, indicating that larger members release more energy, resulting in wider cracks.

The variations of predicted shear strength by GMDH and MARS models with a/d ratio are shown in Figure 4(b). As expected, the shear strength of RC beams decreases as a/d ratio increases. Furthermore, the prediction of shear strength almost coincides with the experimental test results obtained from literature, according to this figure. The ratio of a/d determines the behavior of deep and slender beams and also their failure modes. In deep beams, it can be expected that marginal enhancement of resistance occurs by the beam beyond the first diagonal crack due to arch action with redistribution of stresses. On the other hand, failure occurs after the first diagonal crack in slender beams. It is indicated that there is possible meager reserve strength beyond diagonal cracking. The similar effect of a/d ratio on shear strength of RC beams is also reported in the literature.

The effect of longitudinal steel ratio (ρ) on shear strength is presented in Figure 4(c). It can be seen that shear strength increases as the percentage of longitudinal reinforcement increases. This contribution of longitudinal reinforcement in increasing shear resistance can be attributed to dowel action and can be about 30% in slender member, such as slabs. It should be noted that, in HSC beams, it demands higher

percentage of longitudinal reinforcement to enhance the dowel force than beams made of Normal Strength Concrete (NSC).

Based on Figure 4(d), it can be seen that the compressive strength of concrete increases the shear strength when all other parameters are constant. The shear strength in RC slender beams significantly depends on the diagonal cracking strength. The diagonal cracking strength of a NSC beam without web reinforcement is often expressed as a function of square root of compressive strength of concrete [30]. However, the shear strength of RC beams made of HSC tends to vary differently from the square root of compressive strength of concrete. From these observations, it can be concluded that the developed GMDH and MARS models are in good agreement with the physical characteristics of shear strength phenomenon and previous findings.

The data used in parametric analysis were also employed to explore and quantify the relative importance of model inputs to its output by measuring the effects on the output when the inputs vary between their minimum and maximum in the dataset used. The quantification of this process was determined using the data obtained from holding all input variables at their average values except one output that varied between its ranges ($x_i \in \{x_1, x_2, \dots, x_n\}$). The output y_i for n levels of particular input, x_i , was used to identify which variables have more influence using the sensitivity measure, S_g , of the average gradient over all

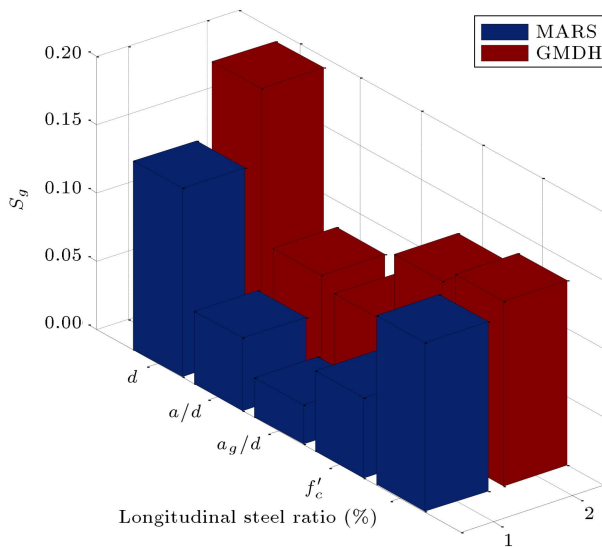


Figure 5. Sensitivity analysis of important parameters based on GMDH and MARS models.

intervals as follows [43]:

$$S_g = \sum_{i=1}^n |y_i - y_{i+1}| / (n - 1). \quad (26)$$

In fact, this approach investigates how output may change if the variables in the trained MARS and GMDH models are perturbed in their ranges. Figure 5 presents the calculated S_g values of GMDH and MARS models for the effective depth, d , shear span-to-depth ratio, a/d , the aggregate size-to-depth ratio, a_g/d , the compressive strength of concrete, f'_c , and the longitudinal steel ratio, ρ . According to this figure, the effective depth provides greater importance and is considered the most significant factor affecting the shear strength of HSC slender beams. On the other hand, the results demonstrate that the aggregate size-to-depth ratio holds the least importance. It can be also observed from this figure that the longitudinal steel ratio is the second important parameter and is followed by the compressive strength of concrete and the shear span-to-depth ratio.

5.3. Safety analysis

A reliable prediction of concrete shear capacity is vital for engineers to have a safe, technically correct

and cost-effective design. According to the previous sections, it can be expected that the uncertainty of the design codes is greater than that of the proposed ones due to their low accuracy. However, it is impossible to know the uncertainty or safety factor incorporated in them. To evaluate the reliability and uncertainty of the existing models, box plot of DR values of different approaches can be used. To generate each box, the DR values for entire database are sorted from the largest to smallest values; then, their median values specify the central mark of the box. The edges of the box represent the 25th and 75th percentiles. The generated box plot can be used as a tool to graphically demonstrate the uncertainty in different models. The spacing between different parts of the box indicates the degree of dispersion (spread) or skewness.

Furthermore, Collins [44] presented a new scale to evaluate and classify the reliability of design codes. This scale is known as Demerit Points Classification (DPC), which considers the safety, accuracy, and scattering of design codes as a function of the ratio between the ultimate resistance in experimental tests and the theoretically estimated shear capacity. In the present study, this ratio is represented by parameter DR. Table 6 presents an adaption in the present study derived from the original values proposed by Collins.

Figure 6 presents the box plots and DR values of different approaches. Furthermore, five classified areas based on Collins criteria, including extremely dangerous, dangerous, appropriate and safe, conservative, and extremely conservative, are specified in Figure 6. As shown, ACI 318-11 [30], CSA A23.3-04 [31], and fib model [32] equations are remarkably conservative and have the highest uncertainty (longer box) as if their predictions vary from extremely conservative to dangerous areas and even extremely dangerous area for ACI 318-11 [30]. The other design codes are also conservative and their uncertainties are remarkable; however, their amounts are low and more reasonable than the three mentioned design codes. However, Figure 6 shows that box plots of the proposed MARS and GMDH models are narrower than those of others are, which is an indicator of a higher confidence level. In addition, the proposed methods show good predictive capability in terms of accuracy with median values close to 1,

Table 6. Classification by demerit points.

$v_{u,exp}/v_{u,predicted}$	Classification	Demerit points
< 0.50	Extremely dangerous	10
$[0.50 - 0.85)$	Dangerous	5
$[0.85 - 1.15)$	Appropriate and safe	0
$[1.15 - 2.00)$	Conservative	1
≥ 2.00	Extremely conservative	2

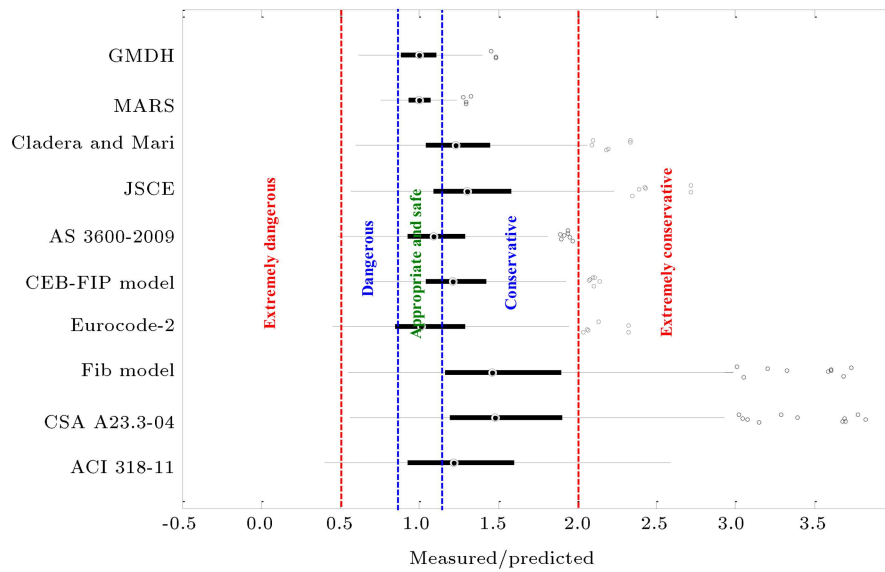


Figure 6. Box plot of different equations.

unlike the design codes. It should be noted that the MARS model outperforms the GMDH model in this aspect.

In order to evaluate quantitatively the safety of the proposed models and existing design codes, first, a demerit point is attributed to each prediction of these equations for total 250 data points based on Table 6. Then, the total values of demerit of each formula are calculated by the sum of the products of the number of specimens at each interval and their corresponding demerit penalty. The lower value of the total sum indicates that the considered formula is safer. Table 7 presents the evaluation of GMDH, MARS and also design equations as a function of the adapted criteria from Ramirez et al. [3]. According to this table, the equations of Eurocode-2 [33], ACI 318-11 [30], and CSA A23.3-04 [31] with respective 427, 413, and 315 total demerit points, as expected, have the weakest performance among design codes. The CEP-FIP [34] model presents the lowest total demerit points amongst the other design equations. However, 60% of its prediction values are either in the fourth classification range (between 1.15 and 2) or the fifth classification range (greater than 2), which is classified as conservative and extremely conservative in terms of safety and leads to an unprofitable design. The total demerit point of the proposed GMDH model is nearly equal to CEP-FIP equation. Unlike the CEP-FIP, 68.4% of GMDH predictions are classified in the third region, which is appropriate and safe. In general, the developed MARS model outperformed other design codes in terms of safety by improving the total demerit point by 69.6% with respect to CEP-FIP equation as the safest design code and GMDH model.

6. Concluding remarks

The robustness of Group Method of Data Handling (GMDH) and Multivariate Adaptive Regression Splines (MARS) as alternative approaches was investigated and assessed in prediction of the shear strength of slender HSC reinforced concrete beams. An available database, including 250 experimental tests of shear strength, was employed to develop the MARS and GMDH models. The proposed models related the shear strength to longitudinal steel ratio, the shear span-to-depth ratio, compressive strength of concrete, the size of the beam specimens, and the size of coarse aggregate. The predictive abilities of GMDH and MARS models were examined by comparing their predictions with those obtained from the most common formulae and MNLR regression approach. The most important outcomes of this study can be summarized as follows:

- Among the existing equations in the literature, there was good agreement between measured and predicted maximum shear strengths using AS 3600-2009 equation. However, a notable number of its predictions were categorized in unsafe region based on Collins criteria. In contrast, the CEB-FIP model represented the safest predictions according to this criterion while its accuracy was remarkably limited;
- The newly proposed GMDH and MARS models outperformed the existing design equations in terms of both accuracy and safety. The errors of the developed GMDH and MARS models also showed a symmetrical and predictable behavior. The MARS model also outperformed GMDH and MNLR models in both accuracy and safety aspects;
- The results of statistical measures showed that

Table 7. Classification of different equations according to the criteria of Collins.

Model	$DR < 0.5$	$0.5 \leq DR < 0.85$	$0.85 \leq DR < 1.15$	$1.15 \leq DR < 2$	$DR > 2$	Demerit points
MARS	—	23	207	20	—	135
	—	23×5	207×0	20×1	—	
GMDH	—	37	171	42	1	227
	—	37×5	171×0	42×1	1×2	
ACI 318-11	6	38	64	121	21	413
	6×10	38×5	64×0	121×1	21×2	
CSA A23.3-04	—	12	40	141	57	315
	—	12×5	40×0	141×1	57×2	
Fib model code	—	14	45	138	53	314
	—	14×5	45×0	138×1	53×2	
Eurocode-2	2	63	99	80	6	427
	2×10	63×5	99×0	80×1	6×2	
CEB-FIP model	—	15	87	142	6	229
	—	15×5	87×0	142×1	6×2	
AS 3600-2009	1	40	108	101	—	311
	1×10	40×5	108×0	101×1	—	
JSCE	—	14	65	154	17	258
	—	14×5	65×0	154×1	17×2	
Cladera and Mari [37]	—	16	76	149	9	247
	—	16×5	76×0	149×1	9×2	

the proposed MARS model outperformed the other design equations in the literature. The proposed models yielded $R^2 = 0.94$ and $RMSE = 0.18$ kN, and this represented a 65.5% improvement in RMSE and a 88% improvement in terms of R^2 with respect to AS 3600-2009 equation as the best equation among the other design equations;

- The sensitivity analysis based on GMDH and MARS models indicated that shear capacity was mostly affected by effective depth, whereas the aggregate size-to-depth ratio were of the least importance. The longitudinal steel ratio was the second important parameter and was followed by the compressive strength of concrete and the shear span-to-depth ratio;
- The robustness of GMDH and MARS models in capturing the underlying physical behaviors of shear strength was verified through parametric study. Results of parametric analysis were also confirmed by the outcomes of sensitivity analysis.

In general, the results indicated that GMDH and MARS algorithms could be successfully used as reliable alternative approaches to predict shear strength of HSC slender beams without web reinforcement.

References

1. ACI “Report on high-strength concrete, ACI 363R-10”, American Concrete Institute, Detroit, MI (2010).
2. Taylor, H.P. “The fundamental behavior of reinforced concrete beams in bending and shear”, *Special Publication*, **42**, pp. 43-78 (1974).
3. Ramirez, J.A., French, C., Adebar, P., Bonacci, J., and Collins, M. “Recent approaches to shear design of structural concrete”, *Journal of Structural Engineering*, **124**(12), p. 1374 (1998).
4. Ashraf, H., Elzanaty, A.H.N., and Floyd, O.S. “Shear capacity of reinforced concrete beams using high-strength concrete”, *Journal Proceedings*, **83**(2), pp. 290-296 (1986).
5. Michael, P.C. and Daniel, K. “How safe are our

- large, lightly reinforced concrete beams, slabs, and footings?", *Structural Journal*, **96**(4), pp. 482-490 (1999).
6. Maekawa, K., Toongoenthong, K., Gebreyouhannes, E., and Kishi, T. "Direct path-integral scheme for fatigue simulation of reinforced concrete in shear", *Journal of Advanced Concrete Technology*, **4**(1), pp. 159-177 (2006).
 7. Hamrat, M., Boulekbache, B., Chemrouk, M., and Amziane, S. "Shear behaviour of RC beams without stirrups made of normal strength and high strength concretes", *Advances in Structural Engineering*, **13**(1), pp. 29-42 (2010).
 8. Gebreyouhannes, E. and Maekawa, K. "Numerical simulation on shear capacity and post-peak ductility of reinforced high-strength concrete coupled with autogenous shrinkage", *Journal of Advanced Concrete Technology*, **9**(1), pp. 73-88 (2011).
 9. Sagaseta, J. and Vollum, R. "Influence of aggregate fracture on shear transfer through cracks in reinforced concrete", *Magazine of Concrete Research*, **63**(2), pp. 119-137 (2011).
 10. Carrasquillo, R., Nilson, A., and Slate, F. "Properties of high strength concrete subject to short-term loads", *Journal Proceedings*, **78**(3), pp. 171-178 (2011).
 11. Sarkar, S., Adwan, O., and Bose, B. "Shear stress contributions and failure mechanisms of high strength reinforced concrete beams", *Materials and Structures*, **32**(2), pp. 112-116 (1999).
 12. Elsanadedy, H.M., Abbas, H., Al-Salloum, Y.A., and Almusallam, T.H. "Shear strength prediction of HSC slender beams without web reinforcement", *Materials and Structures*, **49**(9), pp. 3749-3772 (2015).
 13. Gandomi, A., Alavi, A., and Yun, G. "Nonlinear modeling of shear strength of SFRC beams using linear genetic programming", *Structural Engineering and Mechanics*, **38**(1), pp. 1-25 (2011).
 14. Gandomi, A.H., Alavi, A.H., Gandomi, M., and Kazemi, S. "Formulation of shear strength of slender RC beams using gene expression programming, part II: With shear reinforcement", *Measurement*, **95**, pp. 367-376 (2017).
 15. Hasanipanah, M., Faradonbeh, R.S., Amnieh, H.B., Armaghani, D.J., and Monjezi, M. "Forecasting blast-induced ground vibration developing a CART model", *Engineering with Computers*, **33**(2), pp. 307-316 (2017).
 16. Kaveh, A., Bakhshpoori, T., and Hamze-Ziabari, S.M. "Derivation of new equations for prediction of principal ground-motion parameters using M5' algorithm", *Journal of Earthquake Engineering*, **20**(6), pp. 910-930 (2016).
 17. Kaveh, A., Hamze-Ziabari, S.M., and Bakhshpoori, T. "Patient rule-induction method for liquefaction potential assessment based on CPT data", *Bulletin of Engineering Geology and the Environment*, pp. 1-17 (2017).
 18. Khandelwal, M., Armaghani, D.J., Faradonbeh, R.S., Yellishetty, M., Majid, M.Z.A., and Monjezi, M. "Classification and regression tree technique in estimating peak particle velocity caused by blasting", *Engineering with Computers*, **33**(1), pp. 45-53 (2017).
 19. Nasrollahzadeh, K. and Basiri, M.M. "Prediction of shear strength of FRP reinforced concrete beams using fuzzy inference system", *Expert Systems with Applications*, **41**(4), pp. 1006-1020 (2014).
 20. Shetty, N., Herbert, M.A., Shetty, R., Shetty, D.S., and Vijay, G. "Soft computing techniques during drilling of bi-directional carbon fiber reinforced composite", *Applied Soft Computing*, **41**, pp. 466-478 (2016).
 21. Kaveh, A. and Bakhshpoori, T. "A new metaheuristic for continuous structural optimization: water evaporation optimization", *Structural Multidisciplinary Optimization*, **54**(1), pp. 23-43 (2016).
 22. Kaveh, A. and Bakhshpoori, T. "An accelerated water evaporation optimization formulation for discrete optimization of skeletal structures", *Computers and Structures*, **177**, pp. 218-228 (2016).
 23. Kaveh, A., Bakhshpoori, T., and Hamze-Ziabari, S.M. "M5' and Mars based prediction models for properties of self-compacting concrete containing fly ash", *Periodica Polytechnica Civil Engineering*, **62**(2), pp. 1-17 (2017).
 24. Kaveh, A., Bakhshpoori, T., and Hamze-Ziabari S.M. "New model derivation for the bond behavior of NSM FRP systems in concrete", *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, **41**(3), pp. 249-262 (2017). DOI:10.1007/s40996-017-0058-z.
 25. Kaveh, A., Hamze-Ziabari, S.M., and Bakhshpoori, T. "Feasibility of pso-anfis-pso and ga-anfis-ga models in prediction of peak ground acceleration", *International Journal of Optimization in Civil Engineering*, **8**(1), pp. 1-14 (2018).
 26. Kaveh, A., Hamze-Ziabari, S.M., and Bakhshpoori, T. "M5' algorithm for shear strength prediction of HSC slender beams without web reinforcement", *International Journal of Modeling and Optimization*, **7**(1), pp. 48-53 (2017).
 27. Ivakhnenko, A.G. "Polynomial theory of complex systems", *IEEE Transactions on Systems, Man, and Cybernetics*, **1**(4), pp. 364-378 (1971).
 28. Ivakhnenko, A.G. and Ivakhnenko, G.A. "Problems of further development of the group method of data handling algorithms. Part I", *Pattern Recognition and Image Analysis c/c of Raspoznavaniye Obrazov i Analiz Izobrazhenii*, **10**(2), pp. 187-194 (2000).
 29. Friedman, J.H. "Multivariate adaptive regression splines", *The annals of statistics*, **19**(1), pp. 1-67 (1991).

30. ACI 318M-11, *Building Code Requirements for Structural Concrete and Commentary*, American Concrete Institute (2011).
31. Association CS, *Design of Concrete Structures*, Canadian Standard Association (2004).
32. Du Béton FI, *Model Code 2010: Final Draft*, International Federation for Structural Concrete (2012).
33. British Standards Institution, *Eurocode 2: Design of Concrete Structures: Part 1-1: General Rules and Rules for Buildings*, British Standards Institution (2004).
34. CEB-FIP, MC90 C, *Design of Concrete Structures*, British Standard Institution, London (1993).
35. *Concrete Structures*, Standards Australia. Sydney (2009).
36. Engineers JSOC, *Standard Specifications for Concrete Structures*, Japan Society of Civil Engineers, JSCE Guidelines for Concrete (2010).
37. Cladera, A. and Mari, A.R. “Shear design procedure for reinforced normal and high-strength concrete beams using artificial neural networks. Part I: beams without stirrups”, *Engineering Structures*, **26**(7), pp. 917-926 (2004).
38. Muttoni, A. “Punching shear strength of reinforced concrete slabs without transverse reinforcement”, *ACI Structural Journal*, **105**, Title no. 105-S42, pp. 440-450 (2008).
39. Amanifard, N., Nariman-Zadeh, N., Farahani, M., and Khalkhali, A. “Modelling of multiple short-length-scale stall cells in an axial compressor using evolved GMDH neural networks”, *Energy Conversion and Management*, **49**(10), pp. 2588-2594 (2008).
40. Frank, I.E. and Todeschini, R., *The Data Analysis Handbook*, Elsevier, **14** (1994).
41. Smith, G.N., *Probability and Statistics in Civil Engineering*, Collins London (1986).
42. Sahay, R.R. and Dutta, S. “Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm”, *Hydrology Research*, **40**(6), pp. 544-552 (2009).
43. Kewley, R.H., Embrechts, M.J., and Breneman, C. “Data strip mining for the virtual design of pharmaceuticals with neural networks”, *IEEE Transactions on Neural Networks*, **11**(3), pp. 668-679 (2000).
44. Collins, M. “Evaluation of shear design procedures for concrete structures”, A Report Prepared for the CSA Technical Committee on Reinforced Concrete Design (2001).

Biographies

Ali Kaveh was born in Tabriz, Iran in 1948. After graduation from the Department of Civil Engineering at the University of Tabriz in 1969, he continued his studies on structures at Imperial College of Science and Technology at London University and received his MSc, DIC, and PhD degrees in 1970 and 1974, respectively. He then joined the Iran University of Science and Technology. Professor Kaveh is the author of 550 papers published in international journals and 145 papers presented at national and international conferences. He has authored 23 books in Persian and 8 books in English published by Wiley, Research Studies Press, American Mechanical Society and Springer.

Taha Bakhshpoori was born in Astara, Iran. He earned his BSc in Civil Engineering from University of Guilan (2009), MSc (2011), and PhD (2015) both in Structural Engineering from Iran University of Science and Technology. He joined the University of Guilan, Department of Engineering in Eastern Guilan as an Assistant Professor in December 2015. His research interests include the following: artificial intelligence and its application to civil engineering problems, meta-heuristic modeling and optimization, multi-objective optimization, structural optimization, reliability, robust, and risk-based design optimization, structural health monitoring and damage detection.

Seyyed Mahmood Hamze-Ziabari was born in Ziabar, Iran. After earning his BSc from the Department of Civil Engineering at the University of Tehran in 2012 and his MSc in Marine Structures from Iran university of Science and Technology in 2014, he currently continued his PhD also in Marine Structures at IUST. His research interests include water wave modeling, nearshore hydrodynamics and sediment transport, ocean data analysis, and artificial intelligence and its application to civil engineering problems.