Periodica Polytechnica
Civil Engineering

62(2), pp. 281–294, 2018
https://doi.org/10.3311/PPci.10799

RESEARCH ARTICLE

M5' and Mars Based Prediction Models for Properties of Self-Compacting Concrete Containing Fly Ash

Ali Kaveh[1*], Taha Bakhshpoori[2], Seyed Mahmood Hamze-Ziabari[1]

Received 24 March 2017; Revised 31 May 2017; Accepted 10 July 2017

Abstract
The main purpose of this paper is to predict the properties (mechanical and rheological) of the self-compacting concrete (SCC) containing fly ash as cement replacement by using two decision tree algorithms: M5' and Multivariate adaptive regression splines (Mars). The M5' algorithm as a rule based method is used to develop new practical equations while the MARS algorithm besides its high predictive ability is used to determine the most important parameters. To achieve this purpose, a data set containing 114 data points related to effective parameters affect on SSC properties is used. A gamma test is employed to determine the most effective parameters in prediction of the compressive strength at 28 days, the V-funnel time, the slump flow, and the L-box ratio of SCC. The results from this study suggests that tree based models perform remarkably well in predicting the properties of the self-compacting concrete containing fly ash as cement replacement.

Keywords
self-compacting concrete (SCC), fly ash, M5', MARS

[1] Centre of Excellence for Fundamental Studies in Structural Engineering, Iran University of Science and Technology, Tehran-16, Iran
[2] Faculty of Technology and Engineering, Department of Civil Engineering, East of Guilan, University of Guilan, Rudsar-Vajargah, Iran
[*] Corresponding author, email: alikaveh@iust.ac.ir

1 Introduction

Concrete as a vital material in civil engineering structures and infrastructures along with its growing usage (annually more than 30 billion tons as an estimation) has experienced a significant development in the recent decades. Such a tremendous development is due to its superior physical and mechanical properties and long service life. One of the most revolutionary developments in concrete construction can be the Self-Compacting Concrete (SCC) [1]. SCC was originally developed in Japan to overcome the unsatisfactory compaction because of the inefficient workforce, and complex design and reinforcement details of modern structures [2]. SCC flows under its own weight, maintains homogeneity, fills any form works and passes around dense reinforcement in its plastic state [1, 3]. It equals or excels standard concrete with respect to strength and durability but with the potential of segregation, creep and shrinkage in the hardened state [3, 4]. Nowadays, SCC is used in both precast and in situ construction forms all over the world because of its several economically beneficial factors. In this regards characterizing the properties of SCC are of great importance.

The properties of SCC depend on several factors. From its mechanical properties, the compressive strength is indispensable to model, analyze and design the structures or members. The rheological properties of SCC are also of great importance like its mechanical properties. A concrete mix can only be classified as Self-Compacting Concrete if it satisfies certain requirements in filling ability, passing ability, and segregation resistance [1]. Consequently, development of methods to determine the properties of SCC is critical. The properties of SCC are complex nonlinear regression problems and highly difficult to predict due to the nonlinearity. This aim can be achieved by many different test methods in its plastic and hardened states. On the other hand, data-driven techniques provide the opportunity to tackle such highly nonlinear prediction problems. These techniques have been interested in many fields and as well applied to civil engineering problems in general engineering such as [5–23], and especially in the concrete engineering such as [13, 24–29]. Data-driven techniques were also interested for predicting the properties of SCC.

Beside extensive empirical research on SCC [30], deficiency of theoretical relationship between mixture proportioning and measured engineering properties, push researchers to this subject. This deficiency is overcome mostly by statistical models [31–34], subjectively assuming certain empirical relationships using regression functions [35] and also artificial neural networks (ANN) method based on limited experimental data. Many of them are suitable for the studied experimental domain and consider mostly one or a limited number of properties of SCC for prediction. In 2001, Nehdi et al. [36] for the first time, developed four ANN-based models with the same network architecture to predict slump flow, filling capacity, segregation, and compressive strength of SCC with 87, 75, 52 and 29 experimental data sets containing 10 input variables (cement, water, fly ash, slag, silica fume, limestone filler, sand, gravel, VEA, and high range water reducing admixture), respectively. Siddique et al. [37] used Support Vector Machine (SVM) and ANN for predicting compressive strength and slump flow of SCC using 80 data sets with 6 input variables. Guneyisi et al. [38] used ANN for predicting the initial and final setting times of SCC with mineral admixtures based on 65 SCC mixtures containing 10 input variables. Prediction of compressive strength of SCC containing different types of additions is studied by ANN [39-43], Fuzzy Logic [43], and Least Square Support Vector Machine (LSSVM) and Relevance Vector Machine (RVM) [44]. Very recently Sonebi et al. used SVMs for predicting of six fresh properties of SCC, separately. For this aim they tested 20 mixtures of SCC which results in 60 pairs of datasets with the dosages of cement, limestone powder, water, sand, coarse aggregate and the testing time as input parameters, and slump flow, T50, T60, V-funnel, Orimet, and L-box ratio as output predictions.

Superplasticizers are essential admixtures of SCC to provide the necessary workability. Fly ash, silica fume, and ground granulated blast furnace slag are common additions to improve and maintain the workability of SCC, as well as to regulate the cement content and so reduce the heat of hydration without increasing its cost. Fly ash is a fine inorganic material with pozzolanic properties [1]. Therefore, SCC with the fly ash as an addition is a common form of SCC used in the practice. Many studies have been carried out on the replacement of fly ash as a fine aggregate r in cement concretes. Use of municipal solid waste incineration fly ash in concrete has been studied and from mechanical and durable points of view, the ash incorporated in the concrete behaves like ordinary sand [45]. Investigation on FA as sand (fine aggregate) replacement material and a formula to predict the compressive strength of concrete at 28 days can be found in the work of [46].

Very recently Douma et al. [47] used ANN for prediction of properties of self-compacting concrete containing fly ash. To construct the model, they have collected a total number of 114 different experimental data from the specialized literature. Each data set contains 6 input parameters (Binder content, fly ash percentage, water–binder ratio, fine aggregates, coarse aggregates and superplasticizer) and four output parameters (the compressive strength at 28 days, the V-funnel time, slump flow diameter, and the L-box ratio).

As it is clear in the recent decade remarkable studies were conducted to predict the properties of SCC as a complex nonlinear regression problem using soft computing based techniques. In these studies, which are almost entirely reviewed in the previous paragraph, three issues can be raised: (i) the considered data set; (ii) considered input and output predictive parameters; and (iii) selecting suitable soft computing based technique. It is well known that efficiency of the predictive models significantly depends on how comprehensive the training data is. In addition to mechanical properties, the rheological properties are interested in the case of SCC for guarantee its workability. Some inevitable subjective selection of data driven based techniques is needed.

The main objective of this study is to investigate the potential of two decision tree algorithms: M5′ and Multivariate adaptive regression splines (MARS) for predicting the properties of SCC with fly ash as a common addition. The same database used in the study by Douma et al. [47] is considered here. It should be noted that gathering the database from the preexisting experimental studies can be limited significantly because of exclusion of one or more of SCC properties in some studies and the ambiguity of mixture proportions and testing methods in others [36]. In this regard, 114 data sets can be considered comprehensive to train and test the selected data driven based techniques. Predicting the most important mechanical property (compressive strength at 28 days) and three rheological properties (the slump flow diameter, the L-box ratio, and the V-funnel time) which are required to characterize filling ability, passing ability and segregation resistance and as a sequence to guaranty workability of SCC from dosages of its six main contents (binder content, fly ash percentage, water–binder ratio, fine aggregates, coarse aggregates and superplasticizer) as input parameters makes the models as much as possible to be successful and applicable in the field [47]. The M5′ as one of the model tree algorithms is used for developing predictive and simple formulas for estimation of SCC properties. Unlike most of the data driven based algorithms such as ANN, SVM, and ANFIS, the M5′ algorithm can provide transparent formulas that are physically sound and interpretable. Furthermore, the MARS algorithm besides considering as an efficient algorithm to investigate the efficiency of the M5′ algorithm because of its high predictive ability is used to discover the most significant parameters dealing with the prediction of SCC properties. A gamma test is employed also to determine the most effective parameters in the prediction of the slump flow, the L-box ratio, the V-funnel time and the compressive strength at 28 days of SCC. The results from this study suggest that tree based models perform remarkably well in predicting the properties of the self-compacting concrete containing fly ash as cement replacement.

The remaining sections of this paper are organized as follows. In Section 2 the M5′ and MARS algorithms are outlined and also the performance measures for evaluating the algorithms are presented. Section 3 describes the database used and presents the derived models by algorithms. In the penult section after performance analyses of the algorithms, and the parametric study of the prediction problem, the sensitivity analysis is performed. At the end, the paper is concluded in Section 5.

2 Predictive data mining techniques used in this research

Two different data mining methods consisting of the Multivariate Adaptive Regression Splines (MARS) and M5′ algorithms are used to develop robust models for predicting properties of SCC. The algorithms are outlined at the following [13].

2.1 MARS

MARS is a well-known nonlinear and nonparametric data mining approach that discovers the nonlinear responses between the inputs and outputs of a system using a series of piecewise linear or cubic segments, which are known as splines. The resulting piecewise equation is known as the basis functions (BFs). The slope of regression function is allowed to vary from one segment to the next. The end points of each segment are called knots that mark the end of one region of data and the beginning of another. In contrast to well-known parametric linear regression analysis, MARS provides greater flexibility to explore the 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. MARS algorithm can track and discover the complex structures existing in high-dimensional datasets because it considers all functional forms and also interactions between input variables. The general MARS function can be expressed using the following equation:

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

wher $\tilde{f}(x)$ is the predicted response, β_0 and β_m are constants, which are estimated to yield the best data fit, M is the number of BFs included in the model, and x is the input variable. 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:

$$\lambda_m(x) = \prod_{k=1}^{k_m} [s_{km}(x_{v(k,m)} - t_{k,m})] \tag{2}$$

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 produces BFs by searching in a stepwise manner. An adaptive regression algorithm is used for selecting the knot locations. An optimal MARS is selected through a two-stage forward and backward procedure. In the forward stage, MARS over fits data points by considering a great number of BFs. In the backward stage, redundant BFs are deleted from Eq. (2) to avoid overfitting problem. MARS adopts Generalized Cross-Validation (GCV) as a criterion to delete the redundant basis functions. The expression of GCV is given as:

$$GCV = \frac{\frac{1}{N}\sum_{i=1}^{N}\left[y_i - \hat{f}(\mathrm{x}_i)\right]^2}{\left[1 - \frac{\mathrm{C(B)}}{N}\right]^2} \tag{3}$$

in which N is the number of data, y_i is the response values, and C(B) is a complexity penalty that increases with the number of BFs in the model. It is defined as:

$$C(\mathrm{B}) = (\mathrm{B} + 1) + \mathrm{dB} \tag{4}$$

where d is a penalty for each BF included in the model and B is the number of BFs [48].

2.2 M5′ algorithm

One of the most efficient techniques in data mining approaches is the so-called M5 model tree. M5 model was originally developed by Quinlan [49] and improved later in 1997 in a system called M5′ by Wang and Witten [50]. M5 model trees are more accurate and understandable than regression trees and the ANNs. This model can handle a large number of attributes and high dimensions [51, 52].

The algorithm constitutes of three main steps: Building a tree, pruning the tree and smoothing. In the first step, the basic tree is formed based on a splitting criterion. It uses the standard deviation of the class values for each node as a measure of the error at that node and then calculates the expected error reduction as a result of testing each attribute at that node. Then, the attribute that maximizes the expected error reduction is selected to split the data at that node. The Standard Deviation Reduction (SDR) for M5 is calculated using the following formula:

$$SDR = sd(\mathrm{T}) - \sum_i \frac{|T_i|}{|T|} \times sd(\mathrm{T}_i) \tag{5}$$

here T is the set of examples that reached the node, T_i is the resulted set from splitting the node (leaf) according to the selected attribute and sd is the standard variation [50]. The splitting procedure is ceased when the class values of all instances that reach a node vary by less than 5% of the standard deviation of the original instance set, or when only a few instances remain.

The overfitting problem can occur in the model tree construction process using training data. To alleviate this problem a method is termed “pruning” should be used [53]. The pruning procedure uses an estimate of the expected error that experienced at each node in the test data. First, the absolute difference between the predicted and the actual output values is averaged for each one of the training examples that reach the node. Because the trees have been built expressly for this dataset, this average amount will underestimate the expected error for new cases. In this regards the output value is multiplied by the factor $(n + v)/(n - v)$. n is the number of training examples that reach the node and v is the number of attributes in the model that represents the output value at that node. Therefore, this multiplication is benefited to avoid error underestimating for new data, rather than the data against which the model was trained. If the estimated error is lower than that of the parent (previous splitting attribute), the leaf node would be dropped.

Another problem is the sharp discontinuousness at the leaves of the pruned tree. Smoothing procedure described by Quinlan [49] can reduce and solve this problem which uses the leaf model to compute the predicted value. In the smoothing process, estimated value of each leaf is filtered along the path back to the root. The value at each node that is joining with the predicted value of the linear model for that node can be calculated as follows:

$$P' = \frac{np + kq}{n + k} \tag{6}$$

where P' is prediction which exceeded to a higher node, p is prediction which passes to current node from the below, q is the predicted value by the model at the node, n is the number of training instances reach to the previous node, and k is the Wang & Witten constant.

2.3 Performance measures

The developed model's ability in the prediction of SCC properties will be quantified using four frequently used performance measurements: MAE, RMSE, R, and R2 presented mathematically at the following:

$$MAE = \frac{1}{N}\sum_{1}^{N}\left|P_i - O_i\right| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{1}^{N}\left(P_i - O_i\right)^2} \tag{8}$$

$$R = \frac{\sum_{1}^{N}\left(P_i - P_m\right)\left(O_i - O_m\right)}{\left(\sum_{N}^{1}\left(P - P_m\right)^2\right)^{0.5}\left(\sum_{1}^{N}\left(P_i - O_m\right)^2\right)^{0.5}} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}(O_i - O_m)^2} \tag{10}$$

where O_i is the measured value, P_i stands for prediction values; N is the number of observation data, O_m is the mean value for observation and Pm is the mean value of prediction. The correlation coefficient (R) is a measure of linear relationships between measured and predicted values. The R value varies between -1 and 1. If R is zero, it indicates that there is no linear relationship between variables. If there is a direct linear relationship between the variables, the R value is 1 and in the case of inverse linear relationship is -1. Note that even if the R-value is close to 1, it doesn´t mean that predicted and measured values match; they only tend to vary similarly. However, R does not necessarily indicate the goodness of the model performance, particularly when data range is very wide and the data points distributed about their mean. Therefore, the coefficient of determination, R^2, can be used as an unbiased estimate and can be a better measure for evaluating model performance. The MAE and RMSE measure the difference between predicted and measured values. The values near to zero indicate a close match.

3 Dataset and model development

The study is performed in two steps: (i) the MARS algorithm besides its high predictive ability, is used to discover the most significant parameters dealing with the prediction of SCC properties; (ii) The M5′ is used for developing predictive and simple formulas for estimation of SCC properties. At the following subsections after the description of the dataset used, the developed models based on the MARS and M5′ algorithms are presented. To develop the models, all the data set (114) randomly is divided into two sets considering 80% (91) as training dataset and 20% (23) for testing dataset. The training data is used for the learning of the algorithms. The testing dataset is used to specify the generalization capability of the models to new data they are not trained with.

3.1 Dataset used and considered influential parameters

The built data set and the considered influential input parameters are two major issues in developing new models for the prediction problem at hand. It is well known that efficiency and reliability of the predictive models significantly depend on how comprehensive the training dataset is. In addition to mechanical properties, the rheological properties are interested in the case of SCC for guaranteeing its workability.

Very recently Douma et al. [47] collected and used a total number of 114 different experimental data to develop an ANN model to predict rheological and mechanical properties of self-compacting concrete with fly ash. It should be noted that gathering the database from the preexisting experimental studies can be limited significantly because of exclusion of one or more of SCC properties in some studies and the ambiguity of mixture proportions and testing methods in others. In this regard, 114 data sets can be considered comprehensive to train and test the selected data driven based techniques. Predicting the most important mechanical property: compressive strength at 28 days ($Fc28$, Mpa) and three rheological properties: the slump flow diameter (D, mm), the L-box ratio (L_{box}), and the V-funnel time (V_{funnel}, s) which are required to characterize filling ability, passing ability, and segregation resistance and as a sequence to guaranty workability of SCC is desirable. The following dosages of its six main contents: binder content (B, Kg/m^3), fly ash percentage (P, %), water–binder ratio (W/B), fine aggregates (F, Kg/m^3), coarse aggregates (C, Kg/m^3) and superplasticizer (SP, Kg/m^3), are well-known to be the most effective parameters and make the models as much as possible to be successful and applicable in the field [47]. The histograms of input parameters are shown in Figure 1. For example, the binder content (B) varies between 370 and 733. The output parameters $Fc28$ (Mpa), D (mm), L_{BOX}, and the V_{funnel} (s) varies between [10.2, 86.8], [1.92, 19.2], [0.6, 1], and [480, 880] intervals, respectively. It should be noted that the reliability of developed models in predictions of properties of SCC with fly ash is more in the ranges with more concentrated data points. The prediction problem can be stated mathematically as follows:

$$(Fc28, D, L_{BOX}, V_{funnel}) = f(B, P, W/B, F, C, SP) \quad (11)$$

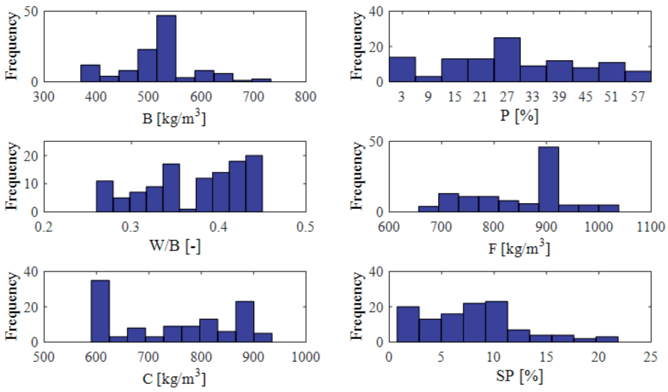

Fig. 1 The histograms of input parameters

3.2 Model development using MARS algorithm

As stated in the Subsection (2.1), the segments (splines) of MARS algorithm can be either piecewise linear or cubic segments. Using both segments and comparing the obtained results, leads us to benefit the piecewise linear segment for model development, which has the better performance. After presenting training data set to the MARS algorithm, the following equations for compressive strength at 28 days ($Fc28$, Mpa) and three rheological properties: the slump flow diameter (D, mm), the L-box ratio (L_{BOX}%), and the V-funnel time (V_{funnel}, sec) are derived:

$$\begin{aligned} Fc28 = {} & 52 - 1.2BF_1 - 0.0058BF_2 + 0.013BF_3 \\ & - 0.0049BF_4 - 0.59BF_5 + 81BF_6 - 0.042BF_7 \\ & - 0.011BF_8 - 0.28BF_9 + 0.045BF_{10} + 0.00063BF_{11} \end{aligned} \quad (12)$$

$$\begin{aligned} V_{funnel} = {} & 3.3 + 61BF_1 - 0.0018BF_2 - 1.3BF_3 - 0.011BF_4 \\ & - 0.035BF_5 + 0.0006BF_6 + 0.0015BF_7 - 5.2BF_8 \\ & + 0.00016BF_9 - 0.011BF_{10} + 0.0022BF_{11} \\ & + 7.5\times10^{-5}BF_{12} \end{aligned} \quad (13)$$

$$\begin{aligned} D = {} & 730 - 150BF_1 + 3.1BF_2 - 0.023BF_3 - 0.002BF_4 \\ & - 5.3BF_5 + 0.0055BF_6 \end{aligned} \quad (14)$$

$$\begin{aligned} L_{BOX} = {} & 0.93 - 0.27BF_1 + 0.0084BF_2 - 0.0025BF_3 \\ & - 0.0018BF_4 + 0.06BF_5 + 1.6\times10^{-5}BF_6 \\ & - 0.026BF_7 + 1.1\times10^{-5}BF_8 \end{aligned} \quad (15)$$

Table 1 Basis functions of the developed MARS models for prediction the properties of SCC with fly ash.

Basis function	Equation			
	$Fc28$ (Mpa)	V_{funnel} (s)	D (mm)	L_{BOX} (%)
BF_1	max (0, 730 - C) × max (0, W/B - 0.35)	max(0, 0.43 –W/B)	max (0, 2.1 - SP)	max (0, 1.4 - SP)
BF_2	max (0, P - 10) × max (0, 670 - C)	max (0, F - 710) × max (0, 860 - C) × max (0, W/B - 0.35)	BF1 × max (0, 400 - B)	max (0, W/B - 0.36)× max (0, 4.5 - SP) × max (0, 450 - B)
BF_3	max (0, B - 480) × max (0, SP -4.5)	BF1 × max (0, P - 25)	max (0, C - 610) × max (0, SP -1.4)	max (0, 54 - P)
BF_4	max (0, 730 - C) × max (0, 37 - P)	max (0, F - 710) × max (0, 0.43 – W/B) × max (0, 54 - P)	max (0, B -450)× max (0, 880 - C)	max (0, W/B - 0.36) × max (0, SP - 4.5) × max (0, B - 500) × max (0, 30 - P)
BF_5	max (0, 0.42 – W/B) × max(0, F -790) × max (0, 7.5 - SP)	max (0, F - 710) × max (0, 0.43 – W/B) × max(0, P - 54)	max (0, F -830)× max (0, W/B - 0.33)	max (0, B - 430) × max (0, 8.4 - SP) × max (0, W/B - 0.42)
BF_6	max (0, 0.42 – W/B) × max (0, 9.7 - SP)	max (0, 0.43 –W/B) × max(0, P-15) ×max(0, F-770) × max(0, B-480)	BF3 × max (0, 450 - B)	max (0, B - 430) × max (0, 610 - C) × max (0, 25 - P)
BF_7	max (0, P - 10) × max (0, 11- SP)	max (0, 900 - C) × max (0, 40 - P)	-	max (0, B - 430) × max (0, 8.4 - SP) × max (0, W/B - 0.39)
BF_8	max (0, P -10) × max (0, SP - 10) × max (0, 570 - B)	BF1 × max (0, F - 970)	-	BF6 × max (0, SP - 9.9)
BF_9	max (0, 0.42 – W/B) × max (0, 920 - F) × max (0, 9.7 - SP)	BF2 × max (0, P - 20)	-	-
BF_{10}	max (0, 0.42 – W/B) × max (0, F - 790) × max (0, 30 - P)	BF7 × max (0, 0.44 – W/B)	-	-
BF_{11}	max (0, 0.42 – W/B) × max (0, 920 - F) × max (0, B – 500) × max (0, C- 900)	max (0, 900 - C) × max (0, F - 970)	-	-
BF_{12}	-	BF4 × max (0, C - 610)	-	-

Table 1 lists the BFs and their corresponding equations. Derived models for $Fc28$, V_{funnel} time, D, and L_{BOX} ratio contain 11, 12, 6 and 8 integrated BFs with interaction terms, respectively. It can be seen that the derived models are not simply additive and those interaction terms play a significantly important role. The final models of MARS (Eqs. (12) to (15)) are achieved via GCV based on forward selection and backward deletion process. As observed, one of the advantages of MARS algorithm is that not only captures complex relationships between independent and dependent variables but also does not require additional effort to verify a priori assumption for the relationship between them. The latter feature becomes more important as the dimension of the problem increases [13].

3.3 Model development using M5′ algorithm

The M5′ algorithm only presents a linear multivariate equation in each class. To compensate this limitation, the data points were transformed into the logarithm space. Then, the final developed model can be rewritten as:

$$(Fc28, D, L_{BOX}, V_{funnel}) = a_1 B^{a_2}\, P^{a_3}\, (W/B)^{a_4}\, F^{a_5}\, C^{a_6}\, SP^{a_7} \quad (16)$$

where a_1, a_2, …, and a_7 are constants. The provided models by the M5′ algorithm are in the form of rules. These rules are simple and interpretable that can be easily used in calculating the mechanical and rheological properties of SCC with fly ash. The developed rules are presented at the following:

Compressive strength at 28 days ($Fc28$, Mpa):

$$Fc_{28} = 0.0022\, B^{0.1668} (1-p)^{0.428} (W/B)^{-0.4676} \times F^{0.2818}\, C^{0.5033}\, SP^{0.4588} \qquad C \le 601.24 \quad (17a)$$

$$Fc_{28} = 2.076\times 10^{-6}\, B^{0.5868} (1-p)^{0.3849} \times (W/B)^{-1.0361} F^{0.7389}\, C^{0.8583} \qquad C > 601.24 \quad (17b)$$

The V-funnel time (V_{funnel}, s):

$$V_{funnel} = 0.517 B^{0.0617} (1-p)^{-0.4067} (W/B)^{-0.4918} \times F^{0.3228}\, C^{0.2019} \qquad C \le 601.24 \quad (18a)$$

$$V_{funnel} = 4.38\times 10^{-6} B^{0.9395} (1-p)^{0.0824} \times (W/B)^{-0.3919} F^{1.0446}\, C^{0.1357} \qquad 601.24 < C \le 821.39 \quad (18b)$$

$$V_{funnel} = 6.37\times 10^{-17} B^{3.3835} (1-p)^{0.1458} \times (W/B)^{-0.3311} F^{2.538} \qquad C > 821.39 \quad (18c)$$

The slump flow diameter (D, mm):

$$D = 1341.576\, B^{-0.1178} (1-p)^{-0.228} \times (W/B)^{0.0203}\, C^{0.118}\, SP^{0.2275} \qquad SP \le 9.299 \ \&\ B \le 440.09 \quad (19a)$$

$$D = 422.7158\, B^{-0.0243} (1-p)^{-0.0789} \times (W/B)^{-0.0973} F^{0.0314}\, C^{0.0898}\, SP^{0.0419} \qquad SP \le 9.299 \ \&\ B > 440.09 \ \&\ P \ge 38.5 \quad (19b)$$

$$D = 295.9528\, B^{-0.1293} (1-p)^{-0.031} \times (W/B)^{-0.0216} F^{0.0768}\, C^{0.1703}\, SP^{0.0419} \qquad SP \le 9.299 \ \&\ B > 440.09 \ \&\ P < 38.5 \ \&\ C \le 604.86 \quad (19c)$$

$$D = 383.4465\, B^{-0.1293} (1-p)^{-0.0731} \times (W/B)^{-0.0216} F^{0.1037}\, C^{0.1343}\, SP^{0.0419} \qquad SP \le 9.299 \ \&\ B > 440.09 \ \&\ P < 38.5 \ \&\ C > 604.86 \quad (19d)$$

$$D = 1864.224\, B^{-0.1807} (1-p)^{-0.0484} \times (W/B)^{0.031}\, C^{0.0413}\, SP^{0.0256} \qquad SP > 9.299 \quad (19e)$$

The L-box ratio (L_{box}%):

$$L_{BOX} = 1955.89 B^{-0.5627} (1-p)^{-0.1111} (W/B)^{-0.1733} \times F^{-0.3046} C^{-0.2823} \quad (20)$$

4 Analytical results and discussion

At the following, the developed models are evaluated through performance analysis, and at the end sensitivity analysis is performed using the Gamma Test after presenting the parametric study for the problem.

4.1 Performance analysis

Performances of the developed MARS and M5′ models are shown in Figure 2. Comparison between measured and predicted values for mechanical and rheological properties of SCC for whole test data in this figure demonstrates that there are little scatters around the line of equality between measured and predicted values. As shown, the proposed model for compressive strength of SCC has the highest accuracy and the least scatter in comparison with other proposed models.

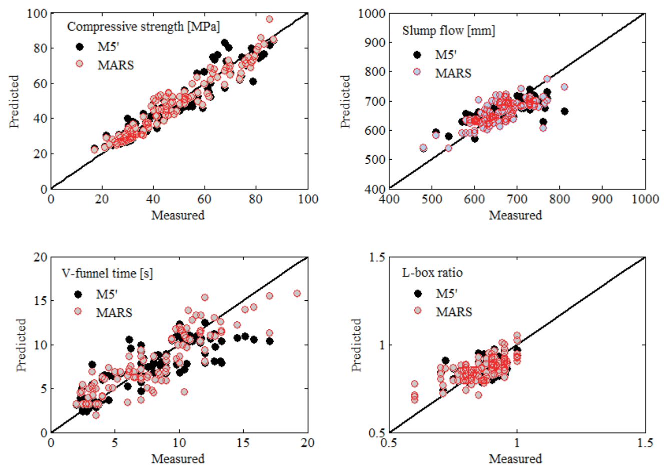

Fig. 2 Comparison between measured and predicted values for properties of SCC with fly ash based on the developed models for whole dataset

To further evaluation of proposed models, the analytical performance measures are presented in Table 2 with respect to training, testing and total datasets. It should be noted that even if R will be close to 1, the predicted and observed values may not match each other; they only tend to vary similarly. To compensate this limitation, the coefficient of determination R^2 can be used. To have precise results, R^2 values should be close to 1 also. It should be noted also that the RMSE and MAE values should be close to zero to have a precise performance. MARS algorithm performs almost better than M5′ algorithm but the differences are negligible. However, models based on the MARS algorithm are recommended theoretically to use for the first two outputs: *Fc*28 and V_{funnel}. It should be noted that the models developed based on the MARS algorithm are more complicated.

Table 2 Performance measurements calculated for developed models according to training, testing and total data sets.

Model		MAE	RMSE	R	R^2	Model
*Fc*28 (MPa)	M5′	Train	4.11	5.23	0.95	0.91
	MARS		3.45	4.45	0.96	0.93
	M5′	Test	5.50	6.84	0.94	0.83
	MARS		4.51	5.56	0.94	0.89
	M5′	Total	4.39	5.59	0.94	0.89
	MARS		3.66	4.69	0.96	0.92
V_{funnel} (s)	M5′	Train	1.64	2.38	0.77	0.57
	MARS		1.64	1.97	0.84	0.71
	M5′	Test	1.53	2.38	0.83	0.66
	MARS		1.11	1.46	0.95	0.87
	M5′	Total	1.62	2.38	0.79	0.61
	MARS		1.53	1.88	0.87	0.75
D [mm]	M5′	Train	27.96	37.12	0.76	0.56
	MARS		28.06	36.29	0.76	0.57
	M5′	Test	27.66	37.04	0.82	0.56
	MARS		35.88	42.15	0.65	0.41
	M5′	Total	28.15	37.27	0.77	0.55
	MARS		29.64	37.55	0.74	0.54
L_{BOX} (%)	M5′	Train	0.05	0.07	0.57	0.32
	MARS		0.05	0.06	0.70	0.49
	M5′	Test	0.07	0.09	0.38	0.13
	MARS		0.059	0.069	0.76	0.56
	M5′	Total	0.06	0.07	0.52	0.27
	MARS		0.05	0.06	0.71	0.51

The ratios between predicted and experimental values of *Fc*28, *D*, V_{funnel}, and L_{BOX} with respect to three main influential parameters: percentage of fly ash (*P* %), water–binder ratio (*W*/*B*), and percentage of superplasticizer (*SP*, Kg/m³) are shown in figures 3(a) to (c), respectively for the derived M5′ based models. As the scattering increases in these figures, the model accuracy will consequently decrease. It can be observed from these figures that the predictions obtained by the proposed models have a very good accuracy with no significant trend with respect to the main parameters. The errors of a good prediction model should be independent of physical parameters involved in that problem. Otherwise, it can be concluded that those physical parameters should be added to that prediction model or they didn't consider correctly in that model. It should be mentioned that the errors of developed models for V_{funnel} slightly decreases with increasing of superplasticizer.

Figure 4 depicts the histogram diagrams of Discrepancy Ratio (DR) of Predicted/Measured values based on the M5′ algorithm for the four output variables and their fitted normal distribution functions. For having a precise and accurate predictive model, the error distribution of Measured/Predicted values should be symmetrical around their mean value and close to 1. Wider distribution generally leads to more uncertainty. As it is clear error distribution of the developed models based on the M5′ algorithm conform the normal distribution. This observation indicates that the uncertainty of the developed models has a deterministic behavior and can be easily modeled.

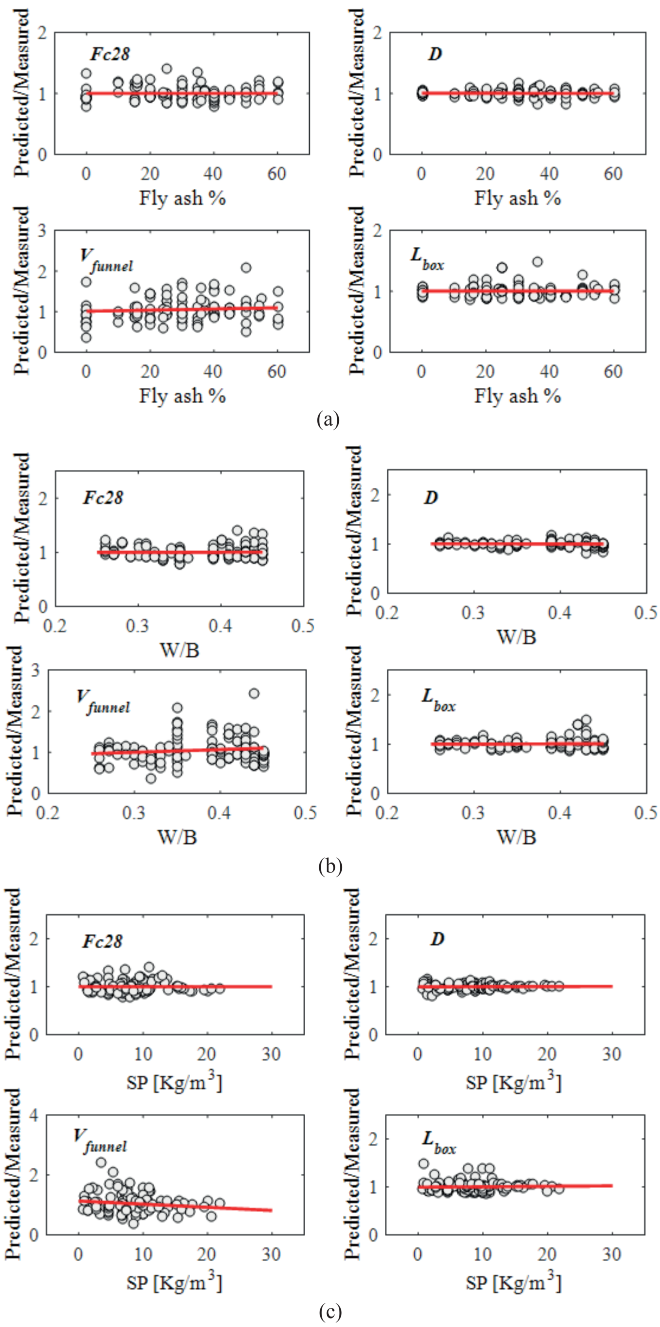

Fig. 3 The ratios between predicted and experimental values of output variables with respect to: (a) *P*, (b) *W*/*B*, and (c) *SP*

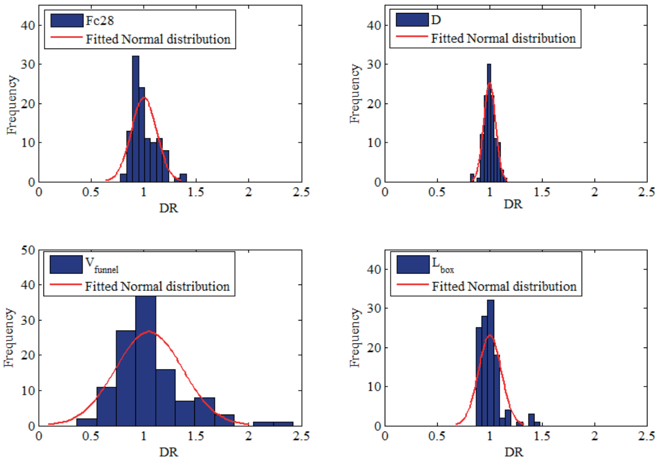

Fig. 4 Histograms of Discrepancy Ratio (DR) of Predicted/Measured values by M5′ algorithm for four output parameters and their fitted normal distribution functions

4.2 Parametric study

To evaluate the robustness of the developed models, a parametric study should be done to ensure that the results of developed models are in line with physical concepts and also previous experimental results. To achieve this, the variations of the predicted compressive strength at 28 days (*Fc*28, MPa) as the main mechanical property of SCC and the predicted slump flow diameter (*D*, mm) as one of the main rheological properties of SCC by the derived M5′and MARS based models are investigated with respect to change of each input parameter. For monitoring parametric behavior of each influential parameter, other ones are fixed to their mean values while the considered parameter increased incrementally from its lower bound to the upper bound based on the experimental database at hand.

Figure 5 shows the results of the parametric study for the Fc28 as one of the main mechanical properties for prediction considered in this study. Based on this figure the following observations can be made. Increasing the dosages of binder content (*B*) and coarse aggregate (*C*) has an increasing influence on the Fc28. On the other hand, the dosage of fly ash (*P*) and water–binder ratio (*W*/*B*) influence in the contrary. It should be noted that the variation of *Fc*28 with respect to 100-P is presented here for more clarity. Other influential parameters (*F* and *SP*) are almost affectless on the *Fc*28. The parametric behavior of B and *W*/*B* are completely sound with the concrete engineering scenes. Experimental studies on the effect of fly ash on compressive strength of self-compacting concrete under different curing conditions resulted that high dosages of *P* decrease the *Fc*28 [54]. Coarse aggregates can have an influence on the compressive strength due to their shape, nominal maximum size, surface texture, and origin [55]. Experimental studies show that dosage of higher strength course aggregates like basalt can has an increasing effect on the *Fc*28 of SCC [56]. Based on this figure and the experimental database used in this study both Mars and M5′ based models can capture the same trends.

In Figure 6 the variations of *D* predicted by M5′ and MARS models to change of *B*, *P*, *W*/*B*, *F*, *C*, and *SP* are illustrated. The main observations from this figure are outlined at the

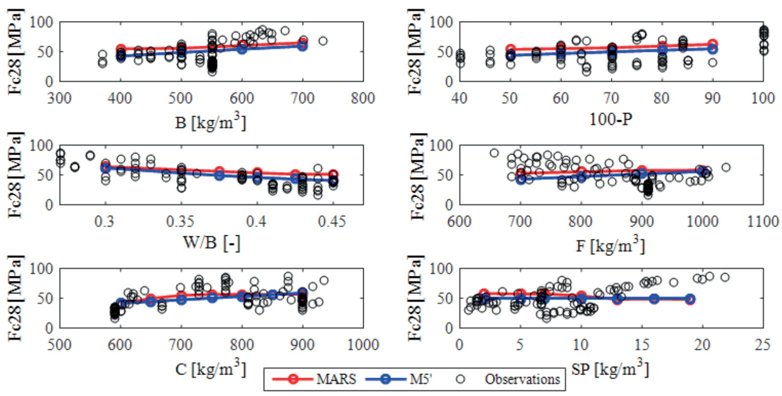

Fig. 5 Parametric study of the compressive strength at 28 days (*Fc*28, Mpa)

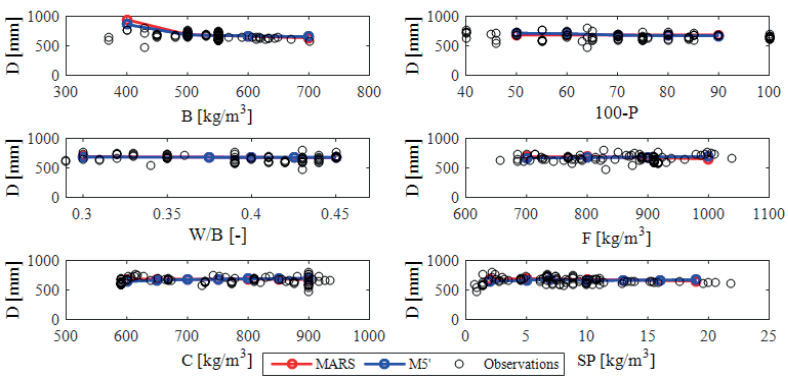

Fig. 6 Parametric study of the slump flow diameter (*D*, mm)

following. (i) Both M5′ and MARS models show a similar trend in variation of B. The slump flow diameter decreases as binder content increases. This trend is also observed in previous experimental studies. The main reason for this trend can be attributed to this fact that the plasticity and cohesiveness of the paste can be improved by decreasing the binder content. (ii) The D parameter is slightly increased by increasing the fly ash replacement level. It should be noted that in the related subfigure the horizontal axis is captured as 100-P for more clarity. It was reported that the increase of fly ash can help to improve the plasticity and cohesiveness of the paste. (iii) The D parameter is not sensitive to change of the ratio between water and binder (W/B). It can be seen from this figure that this behavior is in agreement with the results of other published works. The similar trends were also observed for parameters of fine aggregate and coarse aggregate contents as shown. (iv) According to predictions of developed models, D increases as superplasticizer dosage increases up to a specific dosage (1% to 2%) and after that, it is not sensitive to change of SP dosages. It was reported in the literature that the SP can help to improve the flowability of SCC by liquefying and dispersing actions. Furthermore, the captured behavior of SP by M5′ and MARS models are inconsistent with experimental results of previously published works as shown.

4.3 Sensitivity analysis

To evaluate the most effective influential parameters the Gamma Test (GT) can be used. The GT can examine the relationship between the input and output variables without the construction of a new prediction model. It estimates the minimum Mean Square Error (MSE) that should be obtained by any smooth nonlinear function [14]. Suppose a set of input and output parameters that are given as:

$$\{(x_i, y_i), 1 \leq i \leq M\} = (X, Y) \tag{21}$$

where vector X denote the input variables that are confined to some closed bounded set $c \in R^m$ and y is a response variable. The relationship between X and y can be stated as:

$$y = f(X) + r \tag{22}$$

where f is smooth function and r is a noise variable with random nature. The GT calculate the variance of the random variable ($Var(r)$) without knowing the function of f. To achieve this, the GT estimates the following equations:

$$\delta_M(\mathrm{k}) = \frac{1}{M}\sum_{i=1}^{M}\left|x_{N(\mathrm{i,k})} - x_i\right|^2 \tag{23}$$

$$\gamma_M(\mathrm{k}) = \frac{1}{2M}\sum_{i=1}^{M}\left|y_{N(\mathrm{i,k})} - y_i\right|^2 \tag{24}$$

where $\delta_M(k)$ is the mean square distance to the kth nearest neighbor, $Y_M(k)$ is the corresponding gamma function of the output variable, $x_{N(i,k)}$ is the index of kth nearest neighbor to x_i, $y_{N(i,k)}$ is the corresponding output of $x_{N(i,k)}$ and |…| denotes Euclidean distance.

In general, the GT calculates mean-squared k^{th} nearest neighbor distances $\delta(k)$, ($1 \leq k \leq k_{Max}$) and corresponding $Y(p)^2$. To estimate $Var(r)$, the GT computes the best the intercept of the linear regression line of $Y_M(k)$ versus $\delta_M(k)$, which is often shown by Γ. This parameter is also an indicator for the complexity of function f. The results of GT can also be standardized by considering the V_{ratio} term which returns scale-invariant noise between zero and one. It is defined as:

$$V_{ratio} = \frac{\Gamma}{Var(y)} \tag{25}$$

where $Var(y)$ is the variance of y.

To determine the most important parameters in prediction of the SCC mechanical and rheological properties, seven scenarios are considered. The GT analyses were employed to evaluate the effectiveness of each parameter in construction of a function for estimation of SCC properties. In the first scenario, all input parameters were considered in GT analysis. In the remaining scenarios the input variables are excluded one by one from the dataset and then a new GT analysis is done. The results of GT analysis for each output parameter are presented in Table 3. The most important GT parameter, V_{ratio} is used to analysis the performance of each scenario. Excluding each parameter from analysis leads to change of GT parameters, which can be used to evaluate the importance of that excluded parameter. More changes in GT values indicate that the corresponding excluded parameter has more contribution in the prediction of the related output parameter. As stated, the V_{ratio} can be varied between zero and 1. The value of V_{ratio} close to zero indicates that there is a high degree of accuracy. The first two scenarios with more difference with respect to the total scenario are highlighted in Gray. According to this table excluding (P, F), (W/B, F), (P, SP), and (P, F) causes a significant increase of GT parameters in predicting $Fc28$, V_{funnel}, L_{BOX}, and D, respectively. Therefore, it can be concluded that these parameters are the most important parameters in the prediction of mechanical and rheological parameters of SCC.

Table 3 Results of Gamma test analysis for M5′ in predicting the mechanical and rheological properties of SCC.

Scenario	Input variables	V-ratio			
		$Fc28$	V_{funnel}	L_{BOX}	D
1	Total	0.0337	0.1196	0.3703	0.7397
2	Lack of B	0.0029	0.1486	0.3633	0.7616
3	Lack of P	0.1060	0.1020	0.5173	0.5150
4	Lack of W/B	0.0564	0.1646	0.3847	0.5888
5	Lack of F	0.0838	0.2363	0.4982	0.6506
6	Lack of **C**	0.0500	0.1310	0.4686	0.7639
7	Lack of SP	0.0153	0.1089	0.6694	0.6373

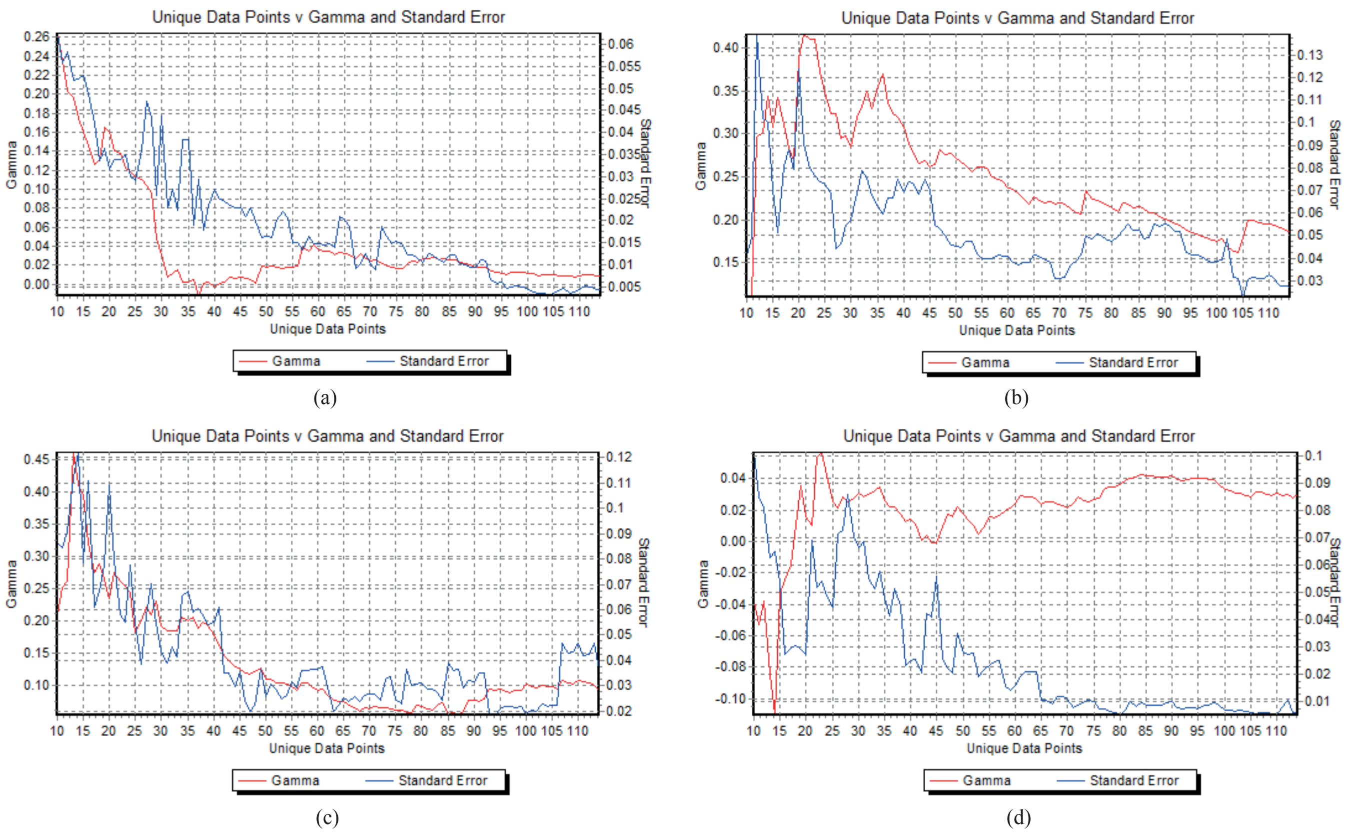

Fig. 7 Variation of gamma parameter besides the standard error values for all data sets for developed gamma models: (a) $Fc28$, (b) V_{funnel}, (c) D, and (d) L_{BOX}

The amount of data used for developing a new model using machine learning approaches plays a crucial role in modeling procedure. To have a reliable model it is suggested that the minimum ratio of samples over the number of involved variables should be 3 [57]. A safer value of 5 can be more reasonable. In the present study, this ratio is much higher and is equal to 114/6 = 19. To more verification of this statement, the gamma test can also be employed. To achieve this purpose, the variations of gamma parameter besides the standard error values for all data sets for developed gamma models of V_{funnel}, $Fc28$, D, and L_{BOX} are shown in figures 7 (a) to (d). According to the Figure 7 (a), the curves of the standard error and gamma are almost flat after point 95. From this observation, it can be concluded that the number of data points required for modeling compressive strength at 28 days of SCC is sufficient i.e. it just needs 95 data points to develop a predictive model. Similarly, the curves of the standard error and gamma are flat after points 105, 106, and 81 for modeling the D, L_{BOX}, and V_{funnel} parameters, respectively.

One of the main issues in mix design of the SCCs is to measure the capacity of concrete to tolerate changes in materials and procedures, which is also known as robustness of the SCC. These variations are usually inevitable during production time on any significant scale in practice. The robustness of the SCC depends on the mix design, the mixing procedure, and the application of the mixture. In this regard, the weighing tolerances due to inaccuracies in measuring process are also inevitable. Hence, the standards ACI 117-90 [58] and EN 206-1 [59] provide restrictions as demonstrated in Table 4 to allow variations in mixture proportions. Therefore, in this study, four different SCCs are considered to evaluate the sensitivity of developed models to small tolerances in materials based on Table 4. Details related to properties of considered concretes are provided in Table 5.

Table 4 Tolerances on material proportions according to ACI and EN.

Components	Limits ACI 117-90 [58]	Limits EN 206-1 [59]
Cementitious materials	±1%	±3%
Sand	±2%	±3%
Gravel	±2%	±3%
Water	±3%	±3%
Admixture	±3%	±5%

Table 5 Details on properties of different considered SCCs.

Dosage	SCC1	SCC2	SCC3	SCC4
B	480	530	607	500
P	10	20	25	40
W/B	0.4	0.45	0.27	0.35
F	890	768	774	923
C	810	668	772	663
SP	9.9	4.55	15.12	7.5
D	665	680	640	680
L_{BOX}	0.85	0.95	0.83	0.88
V_{Funnel}	9	9.8	10.8	6.2
Fc28	46	37.9	74.5	55

The errors in weighing the constituent are altered to analyze how the compressive strength, L_{BOX}, D, and V_{funnel} parameters are changed. Errors in weighing water, binder, fine aggregate, coarse aggregate, fly ash, superplasticizer are the factors considered. The levels of variation for each constituent are presented in Table 6. These variations for each constituent are chosen to cover the usual tolerances occurrence during weighing process in continuous production. According to the study by Rigueira et al. [60], 18 different mixtures based on orthogonal arrays and derived factorial plans for each SCC presented in Table 5 are considered. All these mixtures are shown in Table 7. The properties of these mixtures are presented to developed predictive models and the fresh and hardened properties of them are calculated based on these models.

Table 6 Level of variation for each factor (errors in percentage).

Constituent	Level of variation
Binder	-3; 0; +3
Fly ash	-6; 0; +6
Water	-3; 0; +3
Fine aggregate	-3; 0; +3
Coarse aggregate	-3; 0; +3
Superplasticizer	-5; 0; 5

The effect of weighing errors on the properties of each four SCC base mixtures are evaluated based on analysis of variance (ANOVA). ANOVA determines the most influential factors that effect on response variables considered [60]. The results of ANOVA for fresh and hardened properties of four SCCs

Table 7 Variations applied to each SCC base dosage, in percentage.

Mixture	Binder	Fly ash	Water	Fine aggregate	Coarse aggregate	Superplasticizer
1	0	0	0	0	0	0
2	3	0	3	3	3	5
3	-3	0	-3	-3	-3	-5
4	0	6	3	0	3	-5
5	3	6	-3	3	-3	0
6	-3	6	0	-3	0	5
7	0	-6	-3	3	0	5
8	3	-6	0	-3	3	-5
9	-3	-6	3	0	-3	0
10	0	0	3	-3	-3	5
11	3	0	-3	0	0	-5
12	-3	0	0	3	3	0
13	0	6	0	3	-3	-5
14	3	6	3	-3	0	0
15	-3	6	-3	0	3	5
16	0	-6	-3	-3	3	0
17	3	-6	0	0	-3	5
18	-3	-6	3	3	0	-5

Table 8 ANOVA results calculated based on weighing errors data.

Parameter	$Fc28$		V_{funnel}		D		L_{BOX}	
Mean	0.003		0.005		0.002		-0.0001	
Min	-0.094		-0.075		-0.016		-0.036	
Max	0.106		0.193		0.035		0.033	
Range	0.201		0.268		0.051		0.069	
Variance	0.003		0.002		0.0001		0.0001	
Standard deviation	0.058		0.048		0.011		0.017	
Standard error of mean	0.007		0.006		0.001		0.002	
Constituent	Pearson correlation	Strength	Pearson correlation	strength	Pearson correlation	strength	Pearson correlation	strength
Binder	-0.618	Strong	-0.558	Strong	0.311	Strong	0.596	Strong
Fly ash	0.371	Strong	0.076	Weak	-0.610	Strong	-0.466	Strong
Water	0.447	Strong	0.298	Strong	0.191	Weak	0. 441	Strong
Fine aggregate	-0.384	Strong	-0.478	Strong	-0.107	Weak	0.396	Strong
Coarse aggregate	-0.287	Strong	0.064	Weak	-0.166	Weak	0.247	Strong
Superplasticizer	-0.084	Weak	0.051	Weak	-0.113	Weak	-0.042	Weak

considered in this study are presented in Table 8. According to this table, the compressive strengths of the considered SCCs are mostly affected by changing in water and binder contents. It should be noted that tolerances in fine aggregate, fly ash and coarse aggregate, respectively, can also have significant influence on compressive strength. For V_{funnel} parameter, the tolerances in cement, water, and fine aggregate have the most influences in predictive ability of developed models. In total, the fresh properties of the considered SCCs are mostly affected by the corresponding variations in the cement, water, and fly ash contents. These observations are also in line with experimental results of Rigueira et al. [60].

5 Conclusions

Self-Compacting Concrete (SCC) with the fly ash as an addition is a common form of SCC used in the practice. The objective of this study is to investigate the capability of decision tree based algorithms for predicting the properties of SCC with fly ash as an addition. Selected algorithms are the M5′ and Multivariate Adaptive Regression Splines (MARS). The M5′ algorithm as a rule based method is used to develop new practical equations while the MARS algorithm besides its high predictive ability is used to determine the most important parameters.

To construct the models, very recently collected and available dataset with a total number of 114 different experimental data is used. Each data set contains 6 input influential parameters (Binder content, fly ash percentage, water–binder ratio, fine aggregates, coarse aggregates and superplasticizer) and four output parameters (the compressive strength at 28 days, the V-funnel time, slump flow diameter, and the L-box ratio). The compressive strength is indispensable to model, analyze and design the structures or members as one of the main mechanical properties of SCC. The rheological properties of SCC are also of great importance like its mechanical properties. A concrete mix can only be classified as Self-Compacting Concrete if it satisfies certain requirements in filling ability, passing ability, and segregation resistance. Therefore, such developed models are applicable and desired in the field.

Performance analysis is done based on the frequently used performance measurement to evaluate the effectiveness of the developed model's ability in the prediction of SCC properties. Also to evaluate the robustness of the developed models, a parametric study has been done to ensure that the results of developed models are in line with the physical concepts and also with the previous experimental results. Furthermore, sensitivity analysis based on the gamma test is employed to determine the most effective influential parameters. Results show that tree based models which yield to closed form prediction equations, perform remarkably well in predicting the properties of the self-compacting concrete containing fly ash as cement replacement. Although the models based on the MARS algorithm are somehow more precise, but the models developed based on the M5′ algorithm which are also precise are recommended to use because of the simplicity.

References

[1] EFNARC, F., "Specification and guidelines for self-compacting concrete. European 29 Federation of National Associations Representing producers and applicators of specialist 30 building products for Concrete (EFNARC)". 32. pp. 31. 2002. http://www.efnarc.org/pdf/SandGforSCC.PDF

[2] Nagamoto, N., Ozawa, K". Mixture properties of self-compacting, high-performance concrete. Special Publication. 172, pp. 623–636. 1999.

[3] Felekoğlu, B., Türkel, S., Baradan, B. "Effect of water/cement ratio on the fresh and hardened properties of self-compacting concrete". *Building and Environment* 42(4), pp. 1795–1802. 2007. https://doi.org/10.1016/j.buildenv.2006.01.012

[4] Ravindrarajah, R. S., Siladyi, D., Adamopoulos, B. "Development of high strength self compacting concrete with reduced segregation potential". In: *International RILEM Symposium on Self-Compacting Concrete*. (Wallevik, O., Nielsson, I. (Eds.)), pp. 530–532. RILEM Publications SARL. 2003.

[5] Alavi, A. H., Gandomi, A. H., Gandomi, M., Sadat Hosseini, S. S. "Prediction of maximum dry density and optimum moisture content of stabilised soil using RBF neural networks". *The IES Journal Part A: Civil & Structural Engineering*, 2(2), pp. 98–106, 2009. https://doi.org/10.1080/19373260802659226

[6] Cong, M., Yuehu, T., Erbing, L., Meng, Y., Jianli, D. "Allowable Deformation Prediction for Surrounding Rock of Underground Caverns Based on Support Vector Machine". *Periodica Polytechnica Civil Engineering*, 60(3). pp. 361–369. 2016. https://doi.org/10.3311/PPci.7732

[7] Dabiri, M., Ghafouri, M., Rohani Raftar, H. R., Björk, T. "Neural network-based assessment of the stress concentration factor in a T-welded joint". *Journal of Constructional Steel Research*, 128, pp. 567–578. 2017. https://doi.org/10.1016/j.jcsr.2016.09.024

[8] Facchini, L., Betti, M., Biagini, P. "Neural network based modal identification of structural systems through output-only measurement". *Computers & Structures*, 138, pp. 183–194. 2014. https://doi.org/10.1016/j.compstruc.2014.01.013

[9] Gandomi, A. H., Sahab, M. G., Heshmati, A. A. R., Alavi, A. H., Gandomi, M., Arjmandi, p. "Application of a coupled simulated annealing-genetic programming algorithm to the prediction of bolted joints behavior". *American-Eurasian Journal of Scientific Research*, 3(2), pp. 153–162. 2008. https://pdfs.semanticscholar.org/69ff/2bb619dd958b653f204b58b23aa60023866c.pdf

[10] Gandomi, A.H., Alavi, A. H., Mirzahosseini, M. R., Nejad, F. M. "Nonlinear genetic-based models for prediction of flow number of asphalt mixtures". *Journal of Materials in Civil Engineering*, 23(3), pp. 248–263. 2010. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000154

[11] Gholizadeh, S., Pirmoz, A., Attarnejad, R. "Assessment of load carrying capacity of castellated steel beams by neural networks". *Journal of Constructional Steel Research*, 67(5), pp. 770–779. 2011. https://doi.org/10.1016/j.jcsr.2011.01.001

[12] Kaveh, A., Bakhshpoori, T., 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. https://doi.org/10.1080/13632469.2015.1104758

[13] Kaveh, A., T. Bakhshpoori, and S.M. Hamze-Ziabari, "New model derivation for the bond behavior of NSM FRP systems in concrete". *Iranian Journal Of Science and Technology, Transactions of Civil Emgineering*, 2017. https://doi.org/10.1007/s40996-017-0058-z

[14] Kaveh, A., Hamze-Ziabari, S. M., Bakhshpoori, T. "Patient rule-induction method for liquefaction potential assessment based on CPT data". *Bulletin of Engineering Geology and the Environment*, pp. 1–17. 2016. https://doi.org/10.1007/s10064-016-0990-3

[15] Li, Z.-X., Yang, X.-M. "Damage identification for beams using ANN based on statistical property of structural responses". *Computers & Structures*, 86(1–2), pp. 64–71. 2008. https://doi.org/10.1016/j.compstruc.2007.05.034

[16] Tadesse, Z., Patel, K. A., Chaudhary, S., Nagpal, A. K. "Neural networks for prediction of deflection in composite bridges". *Journal of Constructional Steel Research*, 68(1), pp. 138–149. 2012. https://doi.org/10.1016/j.jcsr.2011.08.003

[17] Tarawneh, B., Nazzal, M. D. "Optimization of resilient modulus prediction from FWD results using artificial neural network". *Periodica Polytechnica Civil Engineering*, 58(2), pp. 143–154. 2014. https://doi.org/10.3311/PPci.2201

[18] Ün"es, F., Demirci, M., Kisi, Ö. "Prediction of millers ferry dam reservoir level in USA using artificial neural network". Periodica Polytechnica Civil Engineering, 59(3), pp. 309–318. 2015. https://doi.org/10.3311/PPci.7379

[19] Zargari, S. A., et al., "A computational intelligence-based approach for short-term traffic flow prediction". *Expert Systems*, 29(2), pp. 124–142. 2012. https://doi.org/10.1111/j.1468-0394.2010.00567.x

[20] Kaveh, A., Hamze-Ziabari, S. M., 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. http://www.ijmo.org/vol7/557-MO004.pdf

[21] Kaveh, A., Ghazaan, M. I., Bakhshpoori, T. "An improved ray optimization algorithm for design of truss structures". *Periodica Polytechnica Civil Engineering*, 57(2), pp. 97–112. 2013. https://doi.org/10.3311/PPci.7166

[22] Kaveh, A., Bakhshpoori, T. "Water evaporation optimization: a novel physically inspired optimization algorithm". *Computers & Structures*, 167, pp. 69–85. 2016. https://doi.org/10.1016/j.compstruc.2016.01.008

[23] Kaveh, A., Bakhshpoori, T. "An accelerated water evaporation optimization formulation for discrete optimization of skeletal structures". *Computers & Structures*, 177, pp. 218–228. 2016. https://doi.org/10.1016/j.compstruc.2016.08.006

[24] Mousavi, S. M., Aminian, P., Gandomi, Ah. H., Alavi, A. H., Bolandi, H. "A new predictive model for compressive strength of HPC using gene expression programming". *Advances in Engineering Software*, 45(1), pp. 105–114. 2012. https://doi.org/10.1016/j.advengsoft.2011.09.014

[25] Chou, J.-S., Pham, A.-D. "Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength". *Construction and Building Materials*, 49, pp. 554–563. 2013. https://doi.org/10.1016/j.conbuildmat.2013.08.078

[26] Chou, J.-S., Tsai, C. F. "Concrete compressive strength analysis using a combined classification and regression technique". *Automation in Construction*, 24, pp. 52–60. 2012. https://doi.org/10.1016/j.autcon.2012.02.001

[27] Yan, F., Lin, Z., Wang, X., Azarmi, F., Sobolev, K. "Evaluation and prediction of bond strength of GFRP-bar reinforced concrete using artificial neural network optimized with genetic algorithm". *Composite Structures*, 161, pp. 441–452. 2017. https://doi.org/10.1016/j.compstruct.2016.11.068

[28] Khademi, F., Akbari, M., Mohammadmehdi Jamal, S., Nikoo, M. "Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete". *Frontiers of Structural and Civil Engineering*, 11(1), pp. 90–99. 2017. https://doi.org/10.1007/s11709-016-0363-9

[29] Gandomi, A.H., Alavi, A. H., Gandomi, M., 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. https://doi.org/10.1016/j.measurement.2016.10.024

[30] Shi, C., Wu, Z., Lv, K., Wu, L. "A review on mixture design methods for self-compacting concrete". *Construction and Building Materials*, 84, pp. 387–398. 2015. https://doi.org/10.1016/j.conbuildmat.2015.03.079

[31] Khayat, K., Ghezal, A., Hadriche, M. "Utility of statistical models in proportioning self-consolidating concrete". *Materials and structures*, 33(5), pp. 338–344. 2000. https://doi.org/10.1007/BF02479705

[32] Sonebi, M., "Medium strength self-compacting concrete containing fly ash: Modelling using factorial experimental plans". *Cement and Concrete research*, 34(7), pp. 1199–1208. 2004. https://doi.org/10.1016/j.cemconres.2003.12.022

[33] Ghezal, A., Khayat, K. H. "Optimizing self-consolidating concrete with limestone filler by using statistical factorial design methods". *ACI Materials Journal*, 99(3), pp. 264–272. 2002. https://doi.org/10.14359/11972

[34] Patel, R., Hossain , K. M. A., Shehata , M., Bouzoubaa , N., Lachemi , M. "Development of statistical models for mixture design of high-volume fly ash self-consolidating concrete". *Materials Journal*, 101(4), pp. 294–302. 2004. https://doi.org/10.14359/13363

[35] Kasemchaisiri, R., Tangtermsirikul, S. "Deformability prediction model for self-compacting concrete". *Magazine of Concrete Research*, 60(2), pp. 93–108. 2008. https://doi.org/10.1680/macr.2008.60.2.93

[36] Nehdi, M., El Chabib, H., El Naggar, M. H. "Predicting performance of self-compacting concrete mixtures using artificial neural networks". *ACI Materials Journal*, 98(5), pp. 394–401. 2001. https://doi.org/10.14359/10729

[37] Siddique, R., Aggarwal, P., Aggarwal, Y., Gupta, S. M. "Modeling properties of self-compacting concrete: support vector machines approach". *Computers and Concrete*, 5(5), pp. 123–129. 2008. https://doi.org/10.12989/cac.2008.5.5.461

[38] Güneyisi, E., Gesoglu, M., Özbay, E. "Evaluating and forecasting the initial and final setting times of self-compacting concretes containing mineral admixtures by neural network". *Materials and Structures*, 42(4), pp. 469–484. 2009. https://doi.org/10.1617/s11527-008-9395-5

[39] Aggarwal, Y., Aggarwal, P. "Prediction of compressive strength of SCC containing bottom ash using artificial neural networks". *World Academy of Science, Engineering and Technology*, Vol. No. 5, pp. 762–767. 2011. waset.org/Publication/1097

[40] Prasad, B. R., Eskandari, H., Reddy, B. V. "Prediction of compressive strength of SCC and HPC with high volume fly ash using ANN". *Construction and Building Materials*, 23(1), pp. 117–128. 2009. https://doi.org/10.1016/j.conbuildmat.2008.01.014

[41] Siddique, R., Aggarwal, P., Aggarwal, Y. "Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks". *Advances in Engineering Software*, 42(10), pp. 780–786. 2011. https://doi.org/10.1016/j.advengsoft.2011.05.016

[42] Uysal, M., Tanyildizi, H. "Estimation of compressive strength of self compacting concrete containing polypropylene fiber and mineral additives exposed to high temperature using artificial neural network". *Construction and Building Materials*, 27(1), pp. 404–414. 2012. https://doi.org/10.1016/j.conbuildmat.2011.07.028

[43] Aggarwal, P. Aggarwal, Y. "Prediction of compressive strength of self-compacting concrete with fuzzy logic". *World Academy of Science, Engineering and Technology*, 53, pp. 745–752. 2011. http://waset.org/publications/3539/prediction-of-compressive-strength-of-self-compacting-concrete-with-fuzzy-logic

[44] Aiyer, B. G., Kim, D., Karingattikkal, N., Samui, P., Ramamohan Rao, P. "Prediction of compressive strength of self-compacting concrete using least square support vector machine and relevance vector machine". *KSCE Journal of Civil Engineering*, 18(6), pp. 1753–1758. 2014. https://doi.org/10.1007/s12205-014-0524-0
[45] Aubert, J. E., Husson, B., Vaquier, A. "Use of municipal solid waste incineration fly ash in concrete". *Cement and Concrete Research*, 34(6) pp. 957–963. 2004. https://doi.org/10.1016/j.cemconres.2003.11.002
[46] Rajamane, N. P., Annie Peter, J., Ambily, P. S. "Prediction of compressive strength of concrete with fly ash as sand replacement material". *Cement and Concrete Composites*, 29(3) pp. 218–223. 2007. https://doi.org/10.1016/j.cemconcomp.2006.10.001
[47] Douma, O. B., Boukhatem , B., Ghrici, M., Tagnit-Hamou, A. "Prediction of properties of self-compacting concrete containing fly ash using artificial neural network". *Neural Computing and Applications*, pp. 1–12. 2016. https://doi.org/10.1007/s00521-016-2368-7
[48] Friedman, J. H., "Multivariate adaptive regression splines". *The Annals of Statistics*, 19(1), pp. 1–67. 1991.
[49] Quinlan, J. R. "Learning with continuous classes". In: *Proceedings of Australian Joint Conference on Artificial Intelligence*. Hobart, 16–18 November 1992, pp. 343–348. Singapore. 1992.
[50] Wang, Y., Witten, I. H. "Induction of model trees for predicting continuous classes". 1996. https://pdfs.semanticscholar.org/3324/4b59fe506331926b3a33e348209ac456c532.pdf
[51] Solomatine, D. P., Xue, Y. "M5 Model Trees and Neural Networks: Application to Flood Forecasting in the Upper Reach of the Huai River in China". *Journal of Hydrologic Engineering*, 9(6), pp. 491–501. 2004. https://doi.org/10.1061/(ASCE)1084-0699(2004)9:6(491)
[52] Solomatine, D. P., Ostfeld, A. "Data-driven modelling: some past experiences and new approaches". *Journal of hydroinformatics*, 10(1), pp. 3–22. 2008. http://jh.iwaponline.com/content/ppiwajhydro/10/1/3.full.pdf
[53] Witten, I. H., Frank, E. "Data Mining: Practical machine learning tools and techniques". Morgan Kaufmann series in data management systems. 2005.
[54] Mohamed, H.A. "Effect of fly ash and silica fume on compressive strength of self-compacting concrete under different curing conditions". *Ain Shams Engineering Journal*, 2(2), pp. 79–86. 2011. https://doi.org/10.1016/j.asej.2011.06.001
[55] Khayat, K., De Schutter, G. "Mechanical properties of self-compacting concrete". Vol. 14. Springer. 2014. https://doi.org/10.1007/978-3-319-03245-0
[56] Kozul, R., Darwin, D. "Effects of aggregate type, size, and content on concrete strength and fracture energy". *University of Kansas Center for Research*, Inc. 1997. http://www2.ku.edu/~iri/publications/sm43.pdf
[57] Davies, A. M. C. "The data analysis handbook, IE Frank and R. Todeschini". (Frank, I. E., Todeschini, R.). *Book Review. Journal of Chemometrics*, 9(5), pp. 431–432. 1995. https://doi.org/10.1002/cem.1180090509
[58] ACI Committee 117, "Standard specification for tolerances for concrete construction and materials" (ACI 117-90). Farmington Hills, American Concrete Institute, 1990.
[59] European Committee for Standardization. "EN 206-1 Concrete - part 1: Specification, performance, production and conformity". Brussels, Belgium: CEN, 2000.
[60] Rigueira, J. W., García-Taengua, E., Serna-Ros, P. "Self-consolidating concrete robustness in continuous production regarding fresh and hardened state properties". *ACI Materials Journal*, 106(3), pp. 301–307, 2009.