Computers and Concrete, *Vol. 24, No. 2 (2019) 137-150*
DOI: https://doi.org/10.12989/cac.2019.24.2.137

Prediction of the compressive strength of self-compacting concrete using surrogate models

Panagiotis G. Asteris[*1], Ali Ashrafian[2] and Mohammad Rezaie-Balf[3]

[1]Computational Mechanics Laboratory, School of Pedagogical and Technological Education, Heraklion, GR 14121, Athens, Greece
[2]Department of Civil Engineering, Tabari University of Babol, P.O. Box 47139-75689, Babol, Iran
[3]Department of Civil Engineering, Graduate University of Advanced Technology-Kerman, P.O. Box 76315-116, Kerman, Iran

(Received April 12, 2019, Revised June 5, 2019, Accepted June 12, 2019)

Abstract. In this paper, surrogate models such as multivariate adaptive regression splines (MARS) and M5P model tree (M5P MT) methods have been investigated in order to propose a new formulation for the 28-days compressive strength of self-compacting concrete (SCC) incorporating metakaolin as a supplementary cementitious materials. A database comprising experimental data has been assembled from several published papers in the literature and the data have been used for training and testing. In particular, the data are arranged in a format of seven input parameters covering contents of cement, coarse aggregate to fine aggregate ratio, water, metakaolin, super plasticizer, largest maximum size and binder as well as one output parameter, which is the 28-days compressive strength. The efficiency of the proposed techniques has been demonstrated by means of certain statistical criteria. The findings have been compared to experimental results and their comparisons shows that the MARS and M5P MT approaches predict the compressive strength of SCC incorporating metakaolin with great precision. The performed sensitivity analysis to assign effective parameters on 28-days compressive strength indicates that cementitious binder content is the most effective variable in the mixture.

Keywords: artificial intelligence models; compressive strength; multivariate adaptive regression splines (MARS); M5P model tree; self-compacting concrete; surrogate models

1. Introduction

Concrete is among the most fundamental and most widely used materials in the construction industry and it is, therefore, necessary to continuously investigate further improvements to it such as concrete structures with improved durability and mechanical properties (Sabur *et al*. 2001). In the last years, the introduction of self-compacting concrete (SCC) has brought huge technological advances. The use of SCC has facilitated the placing of concrete between the rebar without need of external vibration, just by means of the weight of concrete itself. Utilizing SCC leads to the reduction of construction time and cost as well as the reduction of noise in construction sites Khatib (2008). Concrete workability is an important factor for proper execution, which after widespread application of reinforcing bars in concrete in the beginning of the 20th century; it has been maintained for a long time by adding water to the cement. However, in recent research works it has been found that the use of large amounts of the water in the cement can lead to negative results Alabi *et al*. (2012). In SCC, superplasticizer and binder materials are important to achieve high workability and proper viscosity while eliminating the separation, and to provide solutions for the design of the optimal mix of concrete toward reducing the aggregate to cement materials ratio, increasing the amount of cement- paste with a certain water to cement ratio, and control of the largest coarse aggregate size (Aggarwal *et al*. 2008). The volume of binder materials used in SCC, in comparison to conventional concretes, is higher and this indicates the importance of utilizing proper type of the material and weight combination of these materials to provide higher durability and strength of concrete and also its corresponding effects such as reduced generation of pollutant gazes during cement production and participation in the sustainable development (Mehta 1978). With respect to this issue, i.e., that consumption of high amounts of cement and superplasticizer requires huge expenses, the utilization of certain alternative supplementary cementitious materials (SCMs) such as metakaolin as a replacement for Portland cement has been in consideration. The environmental concerns over extraction of raw materials and emission of CO2 during cement production have urged us to reduce the amount of consumed cement by application of additives (AzariJafari *et al*. 2019). Utilizing metakaolin increases the concrete strength and durability against chemical attacks, alkali silica reaction and freeze-thaw cycles. Metakaolin is also effective in certain mechanical properties of concrete including compressive strength, early age and flexural strength (Poon *et al*. 2006, Wild *et al*. 1996, Coleman and Page 1997, Frias and Cabera 2000, Ramzanianpour and Bahrami Juvein 2012, Sfikas *et al*. 2014, Hassan *et al*. 2012a). The wide range of materials and substances used in this type of concrete and the complexity of its corresponding mix design, which is affected by

∗Corresponding author, Professor
E-mail: asteris@aspete.gr; panagiotisasteris@gmail.com

http://www.techno-press.org/?journal=cac&subpage=8
ISSN: 1598-8198 (Print), 1598-818X (Online)

various parameters, as well as the difficulty in formulating relationships between these parameters have made it necessary to propose and present a model for mix design of the SCC incorporated metakaolin.

The lack of a reliable and robust prediction method for the mechanical properties, such as the 28-day compressive strength, is mainly due to the large number of parameters affecting its nonlinear behavior. As the deterministic methods have not been able to offer reliable predictions in the last two decades, soft computing techniques, such as surrogate models, have started to contribute to the problems' solutions in a significant way. In the last decades, the use of artificial intelligence methods for modeling and prediction in civil engineering has become widespread due to their advantages these methods offer. Soft computing techniques have been used for the modeling of mechanical characteristics of concrete materials. Namely, artificial neural networks (ANNs) (Wang *et al*. 2015, Šipoš *et al*. 2017, Asteris and Kolovos 2019, Asteris *et al*. 2016, 2017, 2018, Sonebi *et al*. 2016a, Mansouri *et al*. 2016), adaptive neuro-fuzzy inference systems (ANFIS) (Wang *et al*. 2015, Mansouri *et al*. 2016), support vector machines (SVM) (Sonebi *et al*. 2016b, Gilan *et al*. 2012), genetic expression programming (GEP) (Kiani *et al*. 2016, Gholampour *et al*. 2017), model trees (MT), and multivariate adaptive regression splines (MARS) (Mansouri *et al*. 2016, Kaveh *et al*. 2017, Šipoš *et al*. 2017, Sonebi *et al*. 2016a) have been proposed for the prediction of the mechanical and physical properties of concrete materials. The ANN models from these studies reduce, significantly, the mixture cost and time. Mansouri *et al*. (2016) used a comprehensive dataset, which had been utilized in evolutionary algorithm models such as ANFIS, ANN, MT and MARS for estimating the properties of fiber reinforced –polymer (FRP) confined concrete. Kiani *et al*. (2016) proposed new equation for the compressive strength of foam cellular concrete using GEP, based on comprehensive laboratory data, in a wide range of mixture components. Ashrafian *et al*. (2018) developed heuristic models for the prediction of compressive strength characteristics of fiber reinforced SCC containing nano silica using artificial intelligence method; they resulted MARS techniques was superior to other proposed models.

ANN, SVM and ANFIS provide robust modeling technologies, which are alternative to regression analysis approaches; however they are not formula-based heuristic regression methods. Surrogate models such as MARS and M5P MT are capable of generating linear relationships between input and output variables and provide ‘transparent’ results as regression-based formula. In this paper, we present a new formulation for SCC incorporating metakaolin using two predictive methods namely (a) a MARS method based on nonlinear regression, and (b) an MT expanded version of M5 Prime algorithm (M5P MT). The structure of this paper is as follows: Section 2 describes the experimental dataset used in this study as well as a statistical analysis of each one of them. Section 3 describes the proposed methodological approaches, as well as MARS and M5P MT specific methods. The comparative results and the modeling approach are presented in Section 4. Analysis of results and discussion are presented in Section 5. Finally, the conclusions are described in last section.

2. Background of experimental dataset

Improvement of the performance specifications of SCC are usually acquired by the addition of different supplementary cementitious based materials, admixtures and water to the ordinary concrete mixture designs. (Wild *et al*. 1996). Recent advances have been assisted by the utilization and comprehension of chemical combination (e.g., super plasticizers and SCMs for example metakaolin). Metakaolin has essential efficacy on the workability and early age and long-term compressive strength of SCC (Ramzanianpour and Bahrami Juvein 2012). In many cases, there is the economic benefit of price differential between cement and SCMs. Furthermore, partial substitution of cement allows for a considerable decrease in the quantity of the chemical admixture (Hassan *et al*. 2012a). However, the behavior simulation of the compressive strength of SCC is more difficult than that of ordinary concrete.

Recently, Alyhya (2016) performed a regression analysis to formulate the compressive strength of a SCC mixture by the ratio of water to binder (W/B). In addition, he collected experimental data from several studies available in literature (Alyhya 2016). It was found that the compressive strength of SCC cube specimens at the age of 28-days (CS28) could be calculated by following equation

$$\mathrm{CS}_{28} = \frac{195}{12.65^{w/b}} \tag{1}$$

In this paper, the MARS and M5P MT methods generate new predictive formulas to estimate CS28 of SCC and to analyze relationships between this key-property and the influencing parameters as follows

$$\mathrm{CS}_{28} = \mathrm{f}\left(\mathrm{C}, \mathrm{MK}, \mathrm{B}, \mathrm{SP}, \frac{\mathrm{CA}}{\mathrm{FA}}, \frac{\mathrm{W}}{\mathrm{B}}, \mathrm{D}_{\max}\right) \tag{2}$$

where C (Kg/m^3), MK (Kg/m^3), B (Kg/m^3), SP (Kg/m^3), CA/FA, W/B and Dmax (mm) are the cement, metakaolin, binder, super plasticizer, coarse aggregate to fine aggregate ratio, water to binder ratio and maximum size of aggregate in the mix design, respectively. The above parameters have been chosen as the input variables on this study and literature review.

It is clear that the techniques derived utilizing the MARS, M5P MT or other data driven methods, in most cases, have a predictive capability which lies within the data range used for their development. The amount of data used for the training process of the MARS and M5P MT methods bears heavily on the reliability of the final models. Optimal yield of the proposed models to predict the compressive strength of SCC considerably depends on the integrity of the data. It is important that a large variety of experimental data needed to develop concrete relationships between variables must be accessible. To develop these proposed models, a collection of 204 particular data records from experiments has been obtained from published papers (Madandoust and Mousavi 2012, Dinakar and Manu 2014, Hassan *et al*. 2012b, Güneyisi *et al*. 2009, Hassan *et al*. 2015, Kavitha *et al*. 2015, Kannan and Ganesan 2015, Abouhussien and Hassan 2015, Seyd 2015, Gilan *et al*. 2012, Ahari *et al*. 2015, Megat Johari 2011, Khatib 2008, Justice and Kurtis 2007, Badogiannis *et al*. 2015, Dadsetan

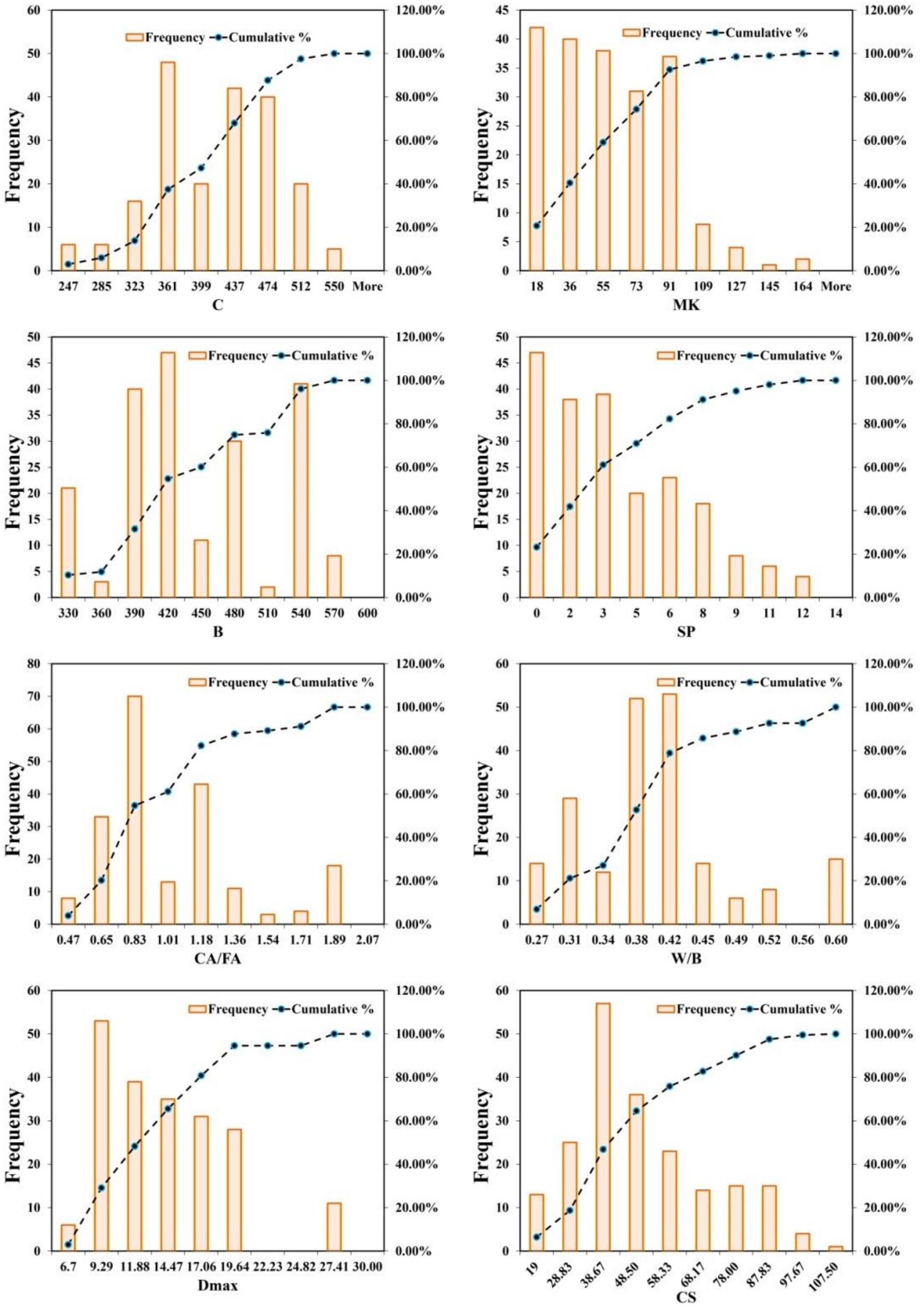


Fig. 1 Histograms of the parameters

and Bai 2017, Gill and Siddique 2018, Lenka and Panda 2017, Wild *et al*. 1996, Ferreira *et al*. 2015, Tirumala *et al*. 2018, Khotbehsara *et al*. 2017, Joseph *et al*. 2017) and is presented in Table 1. The dataset includes the values of C, MK, B, SP, CA/FA, W/B, Dmax, and CS28 of SCC mixtures. Moreover, Fig. 1 demonstrates the frequency

Table 1 Data from experiments published in literature

No.	Reference	No. of Datasets	Compressive Strength [MPa]
1	Madandoust and Mousavi (2012)	15	31-54
2	Dinakar and Manu (2014)	3	94.1-107.5
3	Hassan *et al*. (2012)	8	39.4-48.9
4	Guneyisi *et al*. (2009)	22	60.7-98.6
5	Hassan *et al*. (2015)	18	29.6-55.22
6	Kavitha *et al*. (2015)	4	44-71
7	Kannan and Ganesan (2015)	8	40.77-55.7
8	Abouhussien *et al*. (2015)	4	48.66-55.22
9	Syed Ahmed (2009)	7	39-46
10	Gilan *et al*. (2012)	25	19-50
11	Ahari *et al*. (2015)	4	45.5-64.5
12	Megat Johari *et al*. (2011)	3	91.5-103.5
13	Khatib (2008)	6	66.1-89.7
14	Justice and Kurtis (2007)	6	45-75
15	Sfikas *et al*. (2014), Badogiannis *et al*. (2015)	9	62.3-91
16	Dadsetan and Bai (2017)	4	82-89
17	Singh Gill and Siddique (2018)	4	41.4-52.4
18	Lenka and Panda (2017)	10	37-68
19	Wild *et al*. (1996)	7	62.6-82.47
20	Ferreira *et al*. (2015)	11	23.4-48
21	Chitroju and Yerikenaboina (2018)	14	30-46
22	Khotbehsara *et al*.(2017)	6	42.5-46
23	Joseph *et al*. (2017)	6	64.1-71.5
Total		204	19-107.5

Table 2 Correlation matrix of the input variables

Variable	C	MK	B	SP	CA/FA	W/B	D_{max}
C	1.00	-0.45	0.53	-0.02	-0.07	-0.42	-0.3
MK	-0.45	1.00	0.11	0.26	-0.02	-0.09	-0.06
B	0.53	0.11	1.00	0.31	-0.01	-0.74	-0.38
SP	-0.02	0.26	0.31	1.00	-0.04	-0.31	-0.34
CA/FA	-0.07	-0.02	-0.01	-0.04	1.00	-0.2	0.61
W/B	-0.42	-0.09	-0.74	-0.31	-0.2	1.00	0.26
D_{max}	-0.3	-0.06	-0.38	-0.34	0.61	0.26	1.00

histograms of the parameters. Basically, some of the SCC variables could be dependent on each other. High negative or positive values of the correlation coefficient between the variables may result to poor efficiency of the methods and to the difficulty in construing the effects of the expository variables on the respond. Subsequently, the correlation coefficients between all possible variables have been specified and presented in Table 2. As can be seen in the table, there are not significant correlations between the independent variables.

For the MARS and M5P MT analyses, the database sets were randomly divided into two phases: approximately 75% (153 data points) of the datasets were used for the training phase while the remained datasets (51 data points) were kept for testing the generalization capability of the approaches.

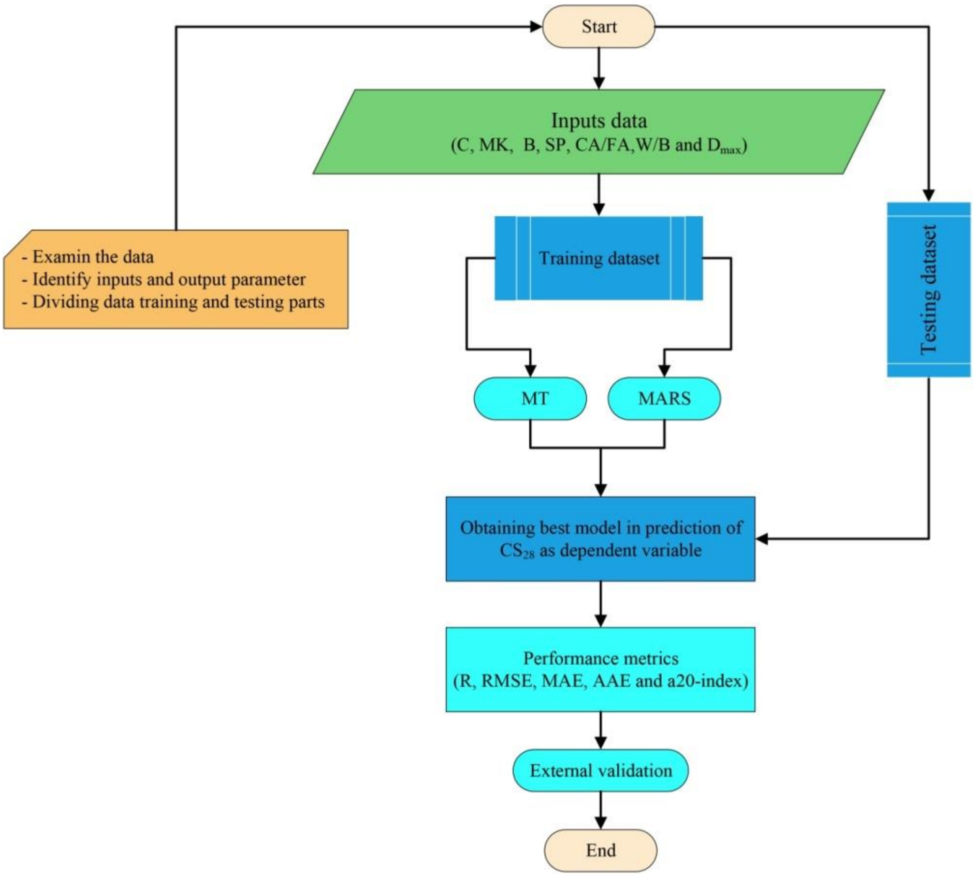
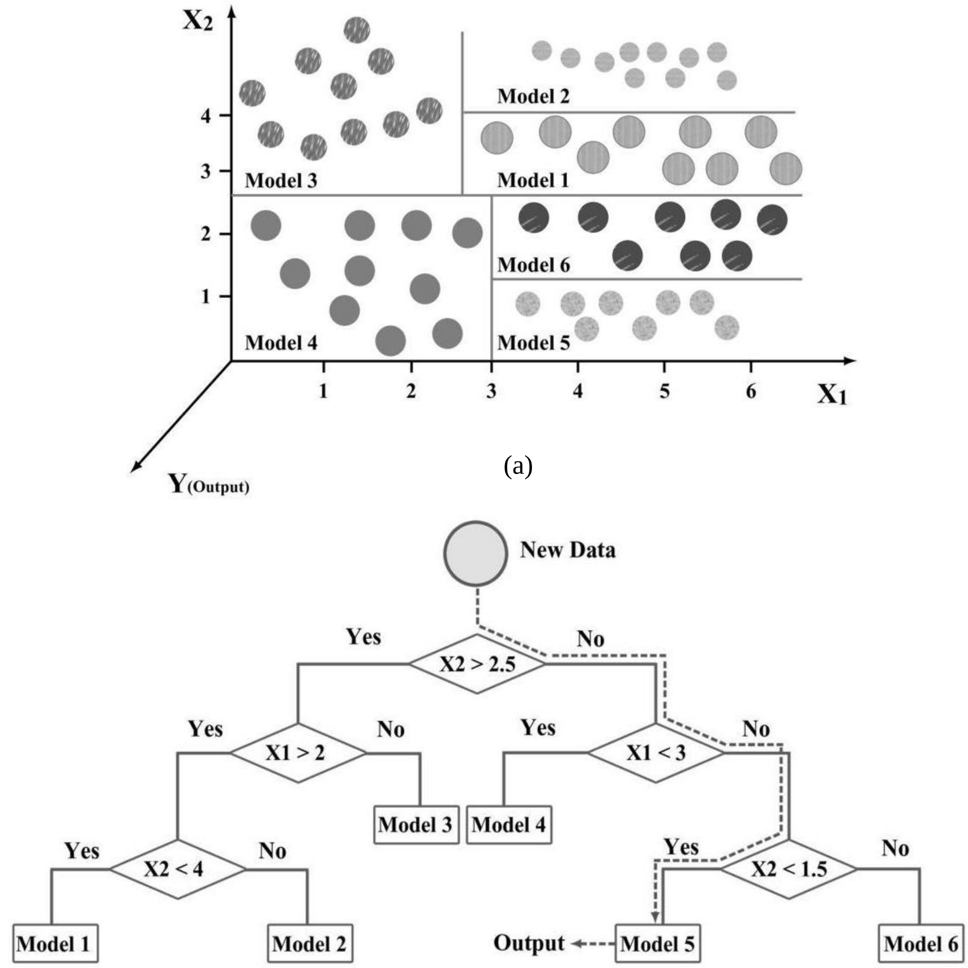


Fig. 2 Schematic flow diagram for the proposed data driven approaches



(b)

Fig. 3 Splitting the input space and prediction by the model tree for a new data record: (a) splitting of the input space (X1×X2) by the M5P MT; (b) predicting a new data by the M5P MT (Rezaie-Balf *et al*. 2017)

3. Predictive data driven approaches

Traditional modeling methods are on the basis of empirical relations obtained from experimental results and observations. Therefore, the application of artificial intelligence methods for modeling and problems prediction is widespread in civil engineering due to its considerable benefits (Kiani *et al*. 2016). In the past, the complexity of the hardened properties of SCC resulted in attempts mostly empirical in nature. Variety of materials used in self-compacting concrete and the complexity of its mixture proportions show that in order to propose improvements fast and accurate models are necessary to predict the properties. Thus, 28-days compressive strength is examined, in this paper, as hardened properties of SCC via new methods of MARS and M5P MT. A schematic flow diagram for the proposed data driven approaches is presented in Fig. 2.

3.1 M5 prime model tree

M5P MT is a supervised learning method which have been widely used to numeric attributes. This technique was first introduced by Quinlan (1992), Wang and Witten (1997) enhanced the method by means of an algorithm called M5P. M5P MT is a tree that contains a root node and leaves with linear regression functions at the top and bottom of the tree. The main aim of this model is to determine the relevancy of independent and dependent variables (Witten and Frank 2005). One of the advantages of model tree is to distribute space of input variables into some disparate regions to create a linear regression in every region. This advantage improves the accuracy of model. In fact it divides problems to sub-problems and combines the results of these sub-problems in order to solve those (Rezaie-Balf *et al*. 2017).

This algorithm comprises two distinct steps: the growth of tree and the tree step. Initially, the instance space recursively splits to construct a regression tree in M5P algorithm. A splitting criterion is used to minimize the intra-subset variability in the values down from the root through the branch to the node. By means of testing each attribute at a specific node, the variability can be calculated by the standard deviation of the values that reach the node from the root through the branch (Witten and Frank 2005). Fig. 3(a) schematically illustrates the splitting phase of input space and the general structure along with the dependent leaves is presented in Fig. 3(b).

In order to organize the basic tree, Standard Deviation Reduction (SDR) is applied as the splitting criterion in M5P MT. This criterion can be calculated as

$$\mathrm{SDR} = \mathrm{sd}(\mathrm{K}) - \sum_{\mathrm{i}} \left|\frac{\mathrm{K_i}}{\mathrm{K}}\right| \times \mathrm{sd}(\mathrm{K_i}) \tag{3}$$

where K indicate a set of data that reach the node; K_i denotes the subsets of data that have the *i*th outcome of the potential set; and *sd* is abbreviation of the standard deviation (Witten and Frank 2005). This splitting process forces the child node to have smaller value of standard deviation as compared to parent node thus making them more pure (Quinlan 1992). The implementation of M5P MT chooses the split that maximizes the expected error reduction after testing all the possible splits. This data division during M5 algorithm implementation produces a large tree which may be the cause of over-fitting. Furthermore, a certain pruning method was offered by Quinlan (1992) to overcome the problem of over-fitting in this paradigm. In general, the pruning is achieved by replacing a sub tree with a linear regression function. More details in this respect can be found in Quinlan (1992), Wang and Witten (1997).

3.2 Multivariate adaptive regression splines

MARS is a non-linear and non-parametric regression method which was presented by Friedman (2010). It is implemented by models of non-linear responses between a system input and output using a set of splines (piecewise polynomials) with different gradients. There is no need for a permanent assumption about basic functional relationship between input and output variables. Endpoints of the segments are called nodes. A node defines endpoint of an area of data and beginning of another area of data. The resulted splines (known as base functions) provide more flexibility for the model and curvatures, thresholds and other deviations of linear functions. The MARS method creates basis functions by step searching. Adaptive regression algorithm is used to select nodes position. MARS models are implemented via a two-step method. In the first step, functions are added up and probabilistic nodes are found for performance improvement leading to a model with a precise curve fitting (primary phase). The second step involves the removal of minimum real terms (secondary phase). Suppose that y is a deterministic output and $X=(X1, ..., X_p)$ is input variable matrix, P. Thus, it is considered that data are obtained from an unknown “real” model. Consequently, the response is as follows

$$\mathrm{y} = \mathrm{f}(\mathrm{X1}, \dots, \mathrm{Xp}) + \mathrm{e} = \mathrm{f}(\mathrm{x}) + \mathrm{e} \tag{4}$$

where, e is the error. MARS is used to approximate the function f by employing basis functions (BFs). Basis functions are referred to splines (smooth polynomials) comprising piecewise-linear functions and piecewise-cubic functions. In this study, piecewise-linear functions are employed and these functions are explained in the following.

Piecewise-linear functions are a type of max (0, x-t), where a node is located on t value, max (.) denotes that only positive part of (.) is used; otherwise, it is zero.

$$\max(0, \mathrm{x} - \mathrm{t}) \begin{cases} \mathrm{x} - \mathrm{t} & \text{if } \mathrm{x} \geq \mathrm{t} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

The MARS method is a linear combination of BFs expressed as follows

$$\mathrm{f}(\mathrm{x}) = \beta_0 + \sum_{\mathrm{m}=1}^{\mathrm{M}} \beta_{\mathrm{m}} \lambda_{\mathrm{m}}(\mathrm{x}) \tag{6}$$

where, λ_m is smoothing parameter. Each $\lambda_m(x)$ is a basis function which comprises one spline function or the product of two or more spline functions (although data might impose the use higher order; here, a maximum of a second degree order is considered). The coefficients β are constant and can be estimated using least squares method. The MARS modeling stems from data. First, the primary method is applied to the training data for fitting the model. A model is constructed with the intercept, β_0, and the basis pair that generates the largest reduction in the training error. Next model is added to the model, based on present model of basic function

$$\hat{\beta}_{\mathrm{M}+1}\lambda_1(\mathrm{X}) \max\left(0, \mathrm{X_j} - \mathrm{t}\right) + \hat{\beta}_{\mathrm{M}+2}\lambda_1(\mathrm{X}) \max\left(0, \mathrm{t} - \mathrm{X_j}\right) \tag{7}$$

where the least squares method is used for the estimation. Mutual effects between BFs which are present in the model are also considered, since the basis function is added to the model space. Then BFs are added to the model to obtain the maximum number of terms leading to a purposely over fit model. Then a secondary removal approach is employed to reduce the number of terms. This removal method is applied to find a model which is closest to optimal range by eliminating extraneous variables. In this method, BFs with minimum contribution to the model are eliminated in order to find the best sub-model. Therefore, BFs selected from set of all BFs which were used in primary selection step, comprise the final optimized model. Generalized cross validation (GCV) method is used to compare subsets of the model due to its low computational cost. The GCV equation which is an adaptive amount that serves to decrease the chance of over fitting and approximates high dimensional BFs for decreasing goodness of fit probability. N observations are used to calculate GCV of the training data model (Friedman 2010, Jeckabsons 2010)

$$\mathrm{GCV} = \frac{\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}[\mathrm{y_i} - \mathrm{f}(\mathrm{x_i})]^2}{\left[1 - \frac{\mathrm{M} + \mathrm{d} \times (\mathrm{M}-1)/2}{\mathrm{N}}\right]^2} \tag{8}$$

where M is the number of BFs, N is the number of observations, d is the estimation parameter, and f (x_i) represents values predicted by the MARS model. An average squared error of the evaluated model in training data is the numerator which is estimated by the penalizing denominator. The denominator increases the complexity of the model by assuming an ascending variance. It is worth mentioning that (M-1)/2 is the number of nodes of the basis function. GCV not only estimates the number of BFs of a model but also estimates the number of nodes (Hastie *et al*. 2009, Zhang and Goh 2013). In order to minimize Eq. (6), one BF is eliminated in each removal step such that the presented model is fitted sufficiently (Friedman 2010). MARS is an adaptive technique, since BFs and positions of variable node are selected by data-driving and are specific for each problem.

3.3 Performance evaluation criteria

To compare the statistical criteria of the developed techniques, different performance measures (Eqs. (9)-(13)) are calculated. The indices include correlation coefficient (R), root mean square error (RMSE), mean absolute error (MAE), average absolute error (AAE) and engineering index (a20-index).

$$R = \frac{\sum_{i=1}^{M}(O_i - \bar{O}).(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{M}(O_i - \bar{O})^2 \sum_{i=1}^{M}(P_i - \bar{P})^2}} \quad (9)$$

$$RMSE = \frac{\sum_{i=1}^{M}\ (P_i - O_i)^2}{M} \quad (10)$$

$$MAE = \frac{\sum_{i=1}^{M}\ |Pi - Oi|}{M} \quad (11)$$

$$AAE = \frac{1}{M}\sum_{i=1}^{n}\frac{|O_i - P_i|}{O_i} \times 100 \quad (12)$$

$$\boldsymbol{a20 - index} = \frac{m20}{M} \quad (13)$$

where $\bar{O}$ is the mean of O (observed target), $\bar{P}$ is the mean of P (predicted target), M is the number of dataset sample and $m20$ is the number of samples with value of rate O_i/P_i between 0.80 and 1.20 ($0.80 \leq O_i/P_i \leq 1.20$). Note that for a robust predictive model, the values of RMSE, MAE and AAE are expected to be zero, while the R and a20-index values are expected to be equal to 1.00. The recently proposed a20-index has the advantage that their value has a physical engineering meaning. It declares the amount of the samples that satisfies predicted values with a deviation ±20% compared to experimental values (Apostolopoulou *et al*. 2018, Asteris *et al*. 2018, Asteris and Nikoo 2019, Chen *et al*. 2019).

4. Modeling and results

Evaluation of hardened properties of SCC containing metakaolin will be accomplished using 28-days compressive strength. The effects of curing and environmental conditions are not investigated in this research work though their influence on compressive strength at age of 28-day has been reported in previous experimental studies. The data used in the proposed MARS and M5P MT are arranged in a format of seven input parameters including cement, metakaolin, binder (cement+ supplementary cementitious materials), super plasticizer, coarse aggregate to fine aggregate ratio, water to binder ratio and maximum size of aggregate. The dataset presented in this study comprises 204 samples collected from the literature (Table 1).

Table 3 The range of experimental variables

Components	Minimum	Maximum	Average	Standard deviation
Input variables				
Cement (kg/m^3)	209.5	550	394.5	69.56
Metakaolin (kg/m^3)	0	163.5	47.17	34.81
Binder (kg/m^3)	330	600	461.30	70.79
Super plasticizer (kg/m^3)	0	14	4.40	3.24
Coarse aggregate to fine aggregate ratio	0.47	2.6	1.11	0.393
Water to binder ratio	0.27	0.6	0.41	0.082
maximum size of aggregate (mm)	6.7	30	15.19	5.30
Output variable				
Compressive strength (MPa)	19	107.5	54.96	20.09

Based on the above database, each input training vector is of dimension 1×7 and consists of the values of C, MK, B, SP, CA/FA, W/B and Dmax. The corresponding output training vectors are of dimension 1×1 and consist of the value of the 28-days compressive strength of the SC concrete specimen. Their mean values together with the minimum, maximum values as well standard deviation (STD) values are listed in Table 3.

4.1 Development of M5P MT

The M5P MT procedure for prediction of the 28-days compressive strength (CS28) utilized WEKA 3.7 software. In this section, the capability of M5P MT is evaluated toward finding the mathematical formulation of linear equations for CS28 of self-compacting concrete containing metakaolin. The initial parameters of M5P MT technique were taken to their default values; pruning factor 5.0 and smoothing option. After classifying, the model tree method including seven inputs and one output parameter was implemented for prediction of CS28 of SCC using 6 rules. These rules based on conditional sentences are presented in Table 4.

A schematic diagram of tree-building of the M5P MT modeling technique for the estimation of compressive strength of SCC containing metakaolin is presented in Fig. 4. All input variables were taken into account in the estimation of the compressive strength of SCC containing MK; they are significant in development of the proposed linear models (LMs).

4.2 Development of MARS

An open source code of MARS which implements the main functionality of the MARS method for regression proposed in is used to accomplish the analyses presented in this study. Table 5 presented analytical details of MARS, including basis functions (BFs) type, numbers and maximum interactions in the final model and GCV value. Also, this study used a 10-fold cross-validation approach to avoid model performance assessment bias. Analysis of variance (ANOVA) test which is statistical methodology to

Table 4 General features of the proposed model tree method

<table>
<tr><th>Rules of MT approach</th><th>LM</th><th>LM No.</th><th>Equation No.</th></tr>
<tr><td rowspan="6">SP <= 4.316
| SP <= 1.26
| | W/B <= 0.44 : LM1
| | W/B > 0.44 : LM2
| SP > 1.26
| | CA/FA <= 0.971 : LM3
| | CA/FA > 0.971 : LM4
SP > 4.316
| CA/FA <= 1.03 : LM5
| CA/FA > 1.03 : LM6</td><td>-0.009 * B + 1.95* SP + 5.99 * CA/FA - 5.045 * W/B + 38.63</td><td>(1)</td><td>(14)</td></tr>
<tr><td>-0.009 * B + 1.95 * SP + 8.38 * CA/FA - 72.64 * W/B + 60.99</td><td>(2)</td><td>(15)</td></tr>
<tr><td>0.06 * C + 0.10 * MK - 0.051 * B + 3.53 * SP - 5.56 * CA/FA - 76.13 * W/B + 63.81</td><td>(3)</td><td>(16)</td></tr>
<tr><td>0.031* MK - 0.025 * B + 2.97 * SP + 9.94 * CA/FA - 12.15 * W/B + 46.79</td><td>(4)</td><td>(17)</td></tr>
<tr><td>-0.032 * MK - 0.02 * B + 1.18 * SP - 15.82 * CA/FA- 3.99 * W/B + 74.96</td><td>(5)</td><td>(18)</td></tr>
<tr><td>0.035 * C - 0.027 * MK + 2.74 * SP + 6.94 * CA/FA + 32.68 * W/B + 25.2</td><td>(6)</td><td>(19)</td></tr>
</table>

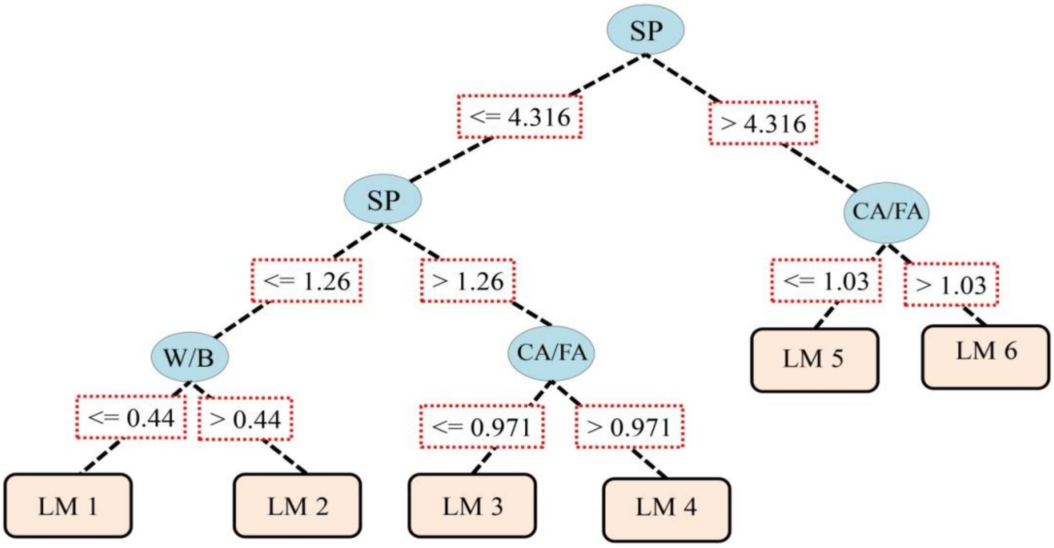


Fig. 4 Proposed pruned tree generation of the CS28 of SCC based on M5P algorithm

Table 5 MARS details to predict CS28

Type of BFs	Piecewise-linear
Number of BFs	24
Max interaction	2
GCV	65.279

determine important variables and important interactions between predictor variables in high-dimensional models was implemented for the model utilizing the training data. In Table 6, the GCV and associated input variables show the importance of that specific ANOVA function which were investigated via the increase in the GCV value caused by the removal of the considered predictor variables from the implemented MARS method. As can be seen in Table 7, the ANOVA decomposition identified CA/FA and B as the most influential variables in the prediction of CS28, respectively. The other effective variables on compressive strength include W/B, Dmax, SP, and MK which were ranked from higher to lower values.

The MARS model determined the basis functions and their related equations. This listed functions and equations in order to choose a best model and optimal equation of desire output from MARS is provided (Zhang *et al*. 2016). The details of BFs for CS28 are shown in Table 7. The excellence of the processing speed of MARS is obvious.

The interpretable MARS method to estimate CS28 of SCC containing metakaolin is given by

$$\begin{aligned} CS_{28} = & -10.526 + 1.431 * BF1 - 2.961 * BF2 \\ & +183.402 * BF3 - 532575.083 * BF4 + 55078.985 \\ & * BF5 - 54916.556 * BF6 - 0.353 * BF7 - 0.272 \\ & * BF8 + 1168926.533 * BF9 - 4.864 * BF10 \\ & +0.023 * BF11 + 0.421 * BF12 - 1.247 * BF13 \\ & -0.434 * BF14 - 364.686 * BF15 - 0.535 * BF16 \\ & +532451.940 * BF17 - 28466.155 * BF18 \\ & -1169072.366 * BF19 - 55085.856 * BF20 \\ & +55003.452 * BF21 - 2.050 * BF22 + 1.143 \\ & * BF23 - 0.026 * BF24 \end{aligned} \tag{20}$$

4.3 External validation of proposed models

Tropsha *et al*. (2003) have introduced some new criteria for the evaluation of the methods according to their performance using datasets incorporated for the tests. It is suggested that at least one slope of regression lines passing through the origin for estimated values against actual values or vice versa should be close to unity (Sattar and Gharabaghi 2015).

Table 6 ANOVA decomposition based on error

No. of BFS	GCV	STD	variable(s)
1	76.528	4.404	MK
1	84.813	14.714	B
2	71.984	7.863	SP
4	164.553	6640.332	CA/FA
1	90.541	10.798	D_{max}
1	68.989	2.098	C and B
1	73.334	2.393	C and CA/FA
1	68.090	2.406	C and W/B
1	65.728	3.210	MK and SP
3	88.945	23.528	B and CA/FA
4	84.465	6629.914	SP and CA/FA
1	92.105	13.965	CA/FA and W/B
3	120.737	10.711	CA/FA and D_{max}

Table 7 Basis functions and related equations of MARS model for CS28

Basis function	Equation
BF1	BF1=max(0, SP - 6)
BF2	BF2=max(0,6 -SP)
BF3	BF3=max(0, CA/FA - 0.807)
BF4	BF4=max(0, CA/FA -1.284)
BF5	BF5=max(0,1.284 – CA/FA) * max(0, SP -1)
BF6	BF6=max(0,1.284 – CA/FA) * max(0,1 -SP)
BF7	BF7=BF3 * max(0,296.300 - C)
BF8	BF8 = max(0,41.250 -MK)
BF9	BF9=max(0,1.284 – CA/FA) * max(0, Dmax -19)
BF10	BF10=max(0,1.284 – CA/FA) * max(0,19-D_{max})
BF11	BF11=BF2 * max(0,80 - MK)
BF12	BF12=max(0,450 -B)
BF13	BF13=BF3 * max(0, B -550)
BF14	BF14=BF3 * max(0,550 -B)
BF15	BF15=max(0,1.284 –CA/FA) * max(0, W/B-0.380)
BF16	BF16=max(0, B -450) * max(0, CA/FA -1.284)
BF17	BF17=max(0, CA/FA -1.284)
BF18	BF18=max(0,1.284 – CA/FA)
BF19	BF19=BF18 * max(0, Dmax -19)
BF20	BF20=BF18 * max(0, SP -1.52)
BF21	BF21=BF18 * max(0,1.52 - SP)
BF22	BF22=max(0, Dmax -10)
BF23	BF23=max(0, C -340) * max(0, W/B -0.4)
BF24	BF24= F12 * max(0, C -403)

$$\mathrm{k} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\mathrm{T_i} \times \mathrm{P_i}}{\mathrm{P_i^2}} \quad \text{or} \quad \mathrm{k}' = \frac{\mathrm{T_i} \times \mathrm{P_i}}{\mathrm{T_i^2}} \tag{21}$$

Furthermore, the coefficient of determination calculated for the regression line passing the origin should be smaller than 0.1.

$$\mathrm{m} = \frac{\mathrm{R^2} - \mathrm{R_0^2}}{\mathrm{R^2}} \tag{22}$$

$$\mathrm{n} = \frac{\mathrm{R^2} - \mathrm{R_0'^2}}{\mathrm{R^2}} \tag{23}$$

Table 8 Validation statistical criterions for proposed approaches

Model	*R*	*K*	*K'*	*m*	*n*	*Rm*
	(*R*>0.8)	(0.85<*K*, *K'*<1.15)		(*m*, *n*<0.1)		(*Rm*>0.5)
MT	0.875	0.964	1.00	- 0.385	- 0.402	0.511
MARS	0.930	0.992	0.993	- 0.154	- 0.154	0.548
Alyhya	-0.030	0.926	0.916	-1.00e+3	-940.26	2.76e-5

Table 9 Comparison of results obtained from proposed models

Performance	Model	Statistical criteria*				
		R	RMSE (MPa)	MAE (MPa)	AAE (%)	a20-index (%)
Training	MT	0.928	77.10	6.59	12.7	84.31
	MARS	0.972	23.60	3.91	8.3	90.85
	Alyhya	0.130	602.32	19.857	43.2	36.60
Testing	MT	0.875	93.92	7.76	14.9	84.31
	MARS	0.930	46.46	5.30	10.3	84.31
	Alyhya	-0.030	543.22	18.25	36.71	31.37

*Bold text refers to best performance

In addition, it is necessary that cross validation satisfy the following condition

$$\mathrm{R_m} = \mathrm{R^2} \times \left(1 - \sqrt{|\mathrm{R^2} - \mathrm{R_0^2}|}\right) > 0.5 \tag{24}$$

The following relationships give the squared correlation factors through the origin between the forecasted and actual values R_0^2, and between the actual and forecasted values $R_0'^2$ as follows

$$\mathrm{R_0^2} = 1 - \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{P_i^2}(1-\mathrm{k})^2 / \sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{P_i} - \overline{\mathrm{P}})^2 \tag{25}$$

$$R_0'^2 = 1 - \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{T_i^2}(1-\mathrm{k}')^2 / \sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{T_i} - \overline{\mathrm{T}})^2 \tag{26}$$

Table 8 shows the validation criteria and associated performance measures of proposed AI techniques in the developed form. If some or all the necessary conditions are satisfied by these techniques they would be regarded as valid. As seen the required criteria are all satisfied implying that they have the predictive potential and are not regarded as just accidental correlations.

5. Discussion

The admission or rejection of the approaches was determined by their capability to estimate the 28-days compressive strength (CS28) of SCC. To test the precision of the model, a comparative study has been performed in terms of R, RMSE, MAE, AAE and a20-index statistical metrics. Figs. 5 and 6 delineated the fitting sufficiently and a20-index of the proposed models for the prediction of CS28. It can be observed that the points are perceptibly

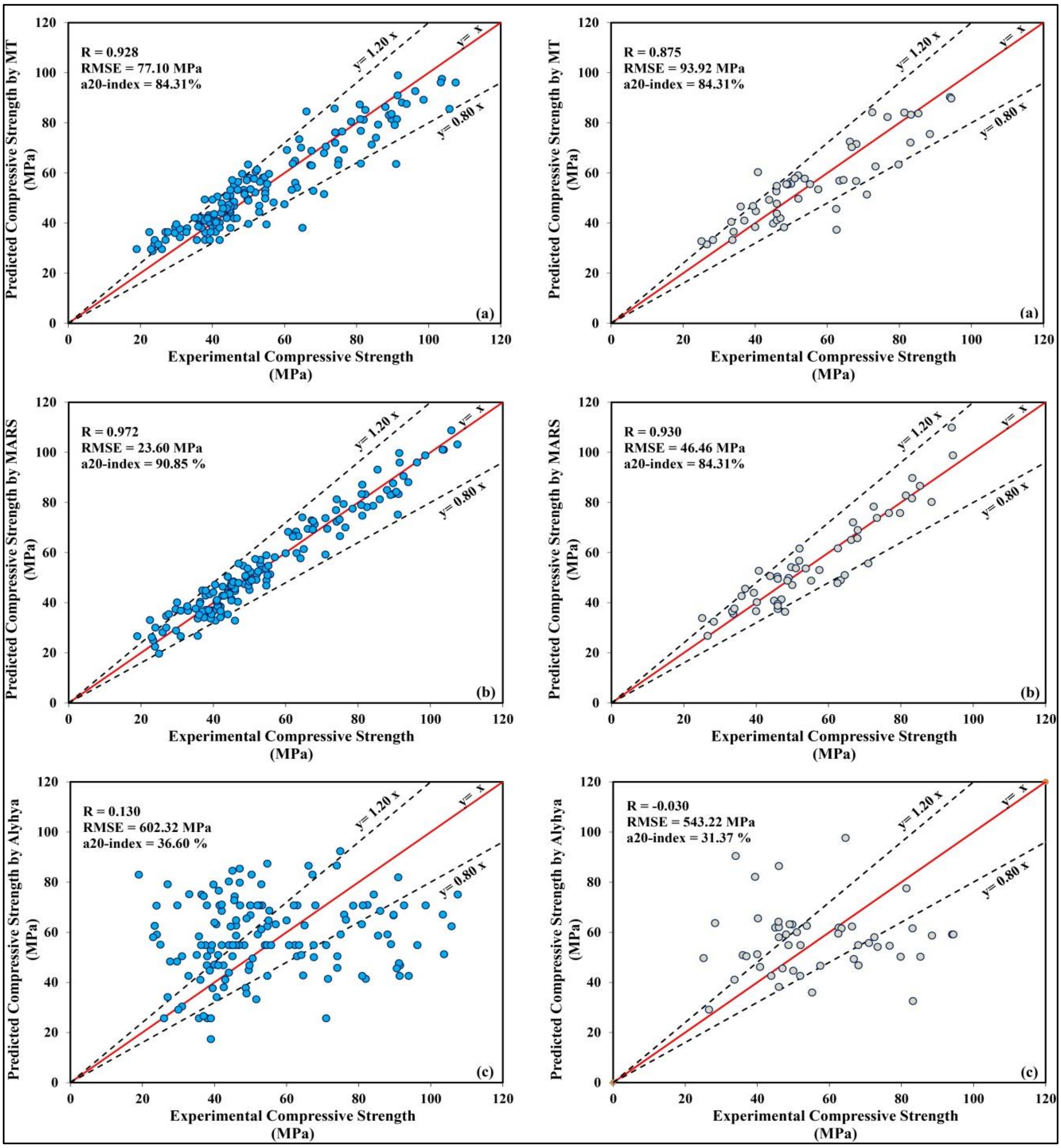


Fig. 5 Scatter plots of observed and predicted CS28 for training (dark color) and testing (light color) performances of the proposed models; (a): MT, (b): MARS, (c): Alyhya model

closed and around the ideal line (y=x). The comparison of the derived results with the experimental findings demonstrates the ability of the proposed models to approximate the compressive strength of self-compacting concrete. Specifically, the developed MARS model can predict the compressive strength in a reliable and robust manner. Furthermore, the developed MARS model outperforms the proposed MT and Alyhya regression model in the training and testing performances, with 12.7% and 14.9 % error, respectively.

Fig. 5 illustrated observed to predicted CS28 ratio. As it can be seen, the closer the ratio to unity, the more precise the models. Distribution plot of observed to predicted ratio for the investigated models indicated in Fig. 6. Values of the AI and Alyhya models indicated outside the range data points are almost of over predicted or under predicted distribution in the prediction procedure.

As shown in Fig. 7 the comparison of values observed and computed of the M5P MT, MARS and alyhya developed model is presented. The results of plots of Fig.7 indicated that the MARS technique to predict local maximum and minimum of data point has better performance than the other models.

Also M5P MT and MARS in the estimation of

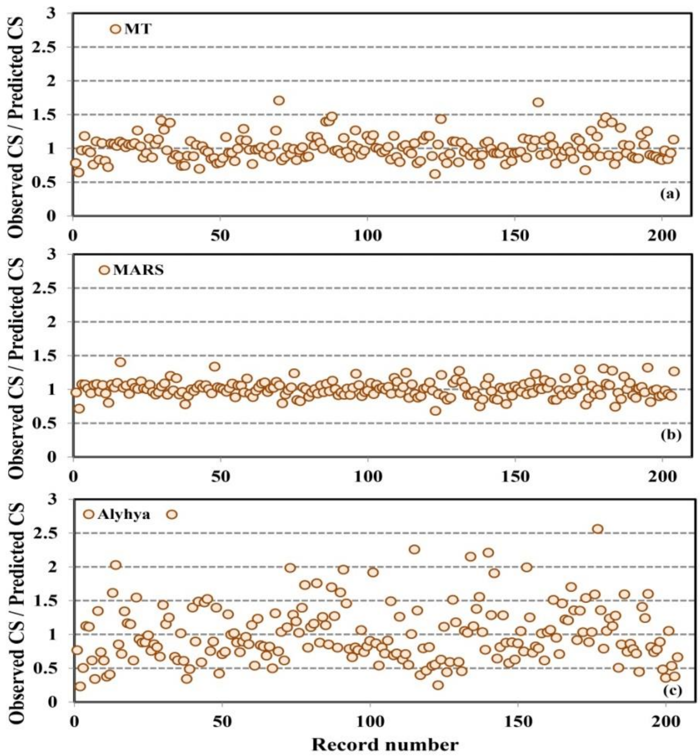


Fig. 6 Ratio between the observed and predicted CS28; (a) MT; (b) MARS; (c) Alyhya model

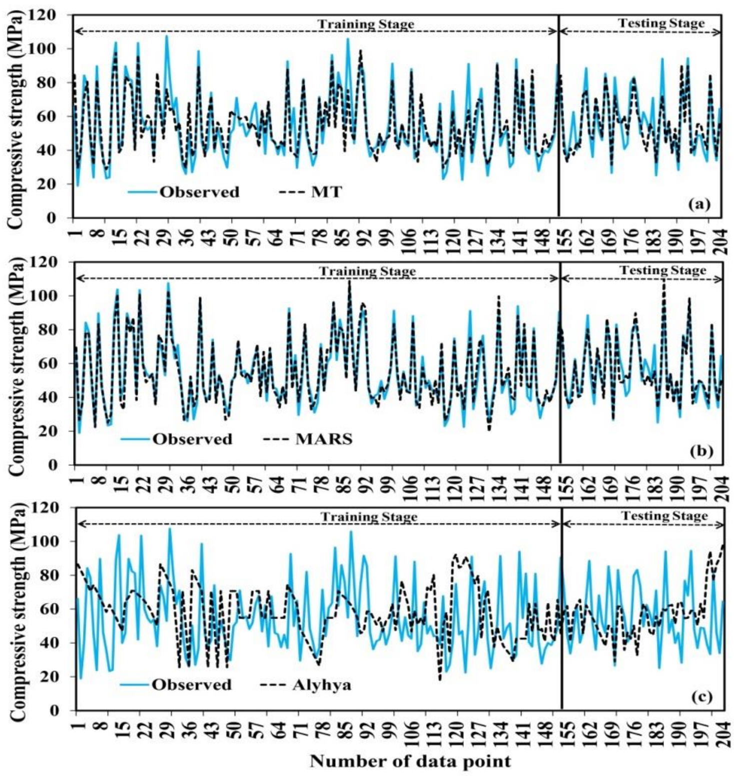


Fig. 7 Time series plot of observed and predicted compressive strength for proposed models: (a) MT; (b) MARS; (c) Alyhya model

compressive strength in a range of 60-100 (MPa) respectively have accurate results reported. Comparisons of R, RMSE, MAE and AAE, in Table 9, indicate that MARS gives only marginally better predictions than M5P MT. The predicted CS28 values based on the MARS model indicate high degree of dependency with the experimental results considered than M5P MT approach both for training and testing process. The statistical criteria have shown obviously this condition. In this study, the maximum value of correlation of coefficient is 0.972 for training stage in the MARS method and the minimum value is 0.875 for testing stage in the M5P MT methods. Therefore, both MARS and M5P MT approaches could be used as trustworthy tools to estimate compressive strength of the SCC containing metakaolin.

6. Conclusions

In this study, data driven approaches such as the MARS and M5P algorithm based model trees have been used to a novel formulation of the 28-days compressive strength of SCC incorporated metakaolin. The application of MARS and M5P MT to generate linear contribution of input and output variables and regression based formula has been investigated. The models were developed by using 204 data sets including mixture proportions specification such as C, MK, B, SP, CA/FA, W/B and Dmax as inputs. Statistical metrics were utilized to validate the efficiency of the proposed data driven models. The performance criteria indicated that the MARS and M5P MT were able to produce accurate estimation of CS28 based on mixture content. The reliability of the developed predictive models was investigated by external validation. The results indicate that the proposed models are robust and provide more accurate predictions than previous developed model. The methods presented can be employed as alternative techniques for simulating the compressive strength of self-compacting concrete materials.

Acknowledgments

The authors would like to thank Dr. Liborio Cavaleri, Professor of Structural Engineering and Seismic Design at Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, dei Materiali, University of Palermo, Italy, and Dr. Danial Jahed Armaghani, PostDoc. Researcher at Institute of Research and Development, Duy Tan University, Vietnam for their valuable comments and discussions. The authorswould also like to thank the anonymous reviewers for their valuable comments and suggestions which improved the quality of the paper.

References

Abouhussien, A.A., Hassan, A.A.A. and Ismail, M.K. (2015), “Properties of semi-lightweight self-consolidating concrete containing lightweight slag aggregate”, *Constr. Build. Mater.*, **75**, 63-73. https://doi.org/10.1016/j.conbuildmat.2014.10.028.

Aggarwal, P., Siiddique, R., Aggarwal, Y. and Gupta, S.M. (2008), “Self-compacting concrete-procedure for mix design”, *Leonardo Elec. J. Pract. Technol.*, **7**(12), 15-24.

Ahari, R.S., Erdem, T.K. and Ramyar, K. (2015), “Time-dependent rheological characteristics of self-consolidating concrete containing various mineral admixtures”, *Constr. Build. Mater.*, **88**, 134-142. https://doi.org/10.1016/j.conbuildmat.2015.04.015.

Ahmed, S. (2009), “Fresh and mechanical properties of self-consolodating concrete incorporationg silica fume and metakaolin”, Theses and dissertation of Master of Civil Engineering, Reyson University.

Alabi, S.A., Olanitori, L.M. and Afolayan, J.O. (2012), “Optimum mix design for minimum concrete strength requirement using akure pit-sand as fine aggregate”, *J. Emerg. Trend. Eng. Appl. Sci.*, **3**(4), 718-724.

Alyhya, W.S. (2016), “Self-compacting concrete: mix proportioning, properties and its flow simulation in the V-funnel”, Doctoral Dissertation, Cardiff University.

Apostolopoulour, M., Douvika, M.G., Kanellopoulos, I.N., Moropoulou, A. and Asteris, P.G. (2018), “Prediction of compressive strength of mortars using artificial neural networks”, *1st International Conference TMM_CH, Transdisciplinary Multispectral Modelling and Cooperation for the Preservation of Cultural Heritage*, Athens, Greece, October.

Ashrafian, A., Amiri, M.J.T., Rezaie-Balf, M., Ozbakkaloglu, T. and Lotfi-Omran, O. (2018), “Prediction of compressive strength and ultrasonic pulse velocity of fiber reinforced concrete incorporating nano silica using heuristic regression methods”, *Constr. Build. Mater.*, **190**, 479-494, https://doi.org/10.1016/j.conbuildmat.2018.09.047.

Asteris, P.G. and Kolovos, K.G. (2019), “Self-compacting concrete strength prediction using surrogate models”, *Neur. Comput. Appl.*, **31**, 409-424, https://doi.org/10.1007/s00521-017-3007-7.

Asteris, P.G. and Nikoo, M. (2019), “Artificial bee colony-based neural network for the prediction of the fundamental period of infilled frame structures”, *Neur. Comput. Appl.*, 1-11. https://doi.org/10.1007/s00521-018-03965-1.

Asteris, P.G., Argyropoulos, I., Cavaleri, L., Rodrigues, H., Varum, H., Thomas, J., Paulo B. and Lourenço, P.B. (2018), “Masonry compressive strength prediction using artificial neural networks”, *1st International Conference TMM_CH, Transdisciplinary Multispectral Modelling and Cooperation for the Preservation of Cultural Heritage*, Athens, Greece, October.

Asteris, P.G., Kolovos, K.G., Douvika, M.G. and Roinos, K. (2016), “Prediction of self-compacting concrete strength using artificial neural networks”, *Eur. J. Environ. Civil Eng.*, **20**(sup1), s102-s122, https://doi.org/10.1080/19648189.2016.1246693.

Asteris, P.G., Nozhati, S., Nikoo, M., Cavaleri, L. and Nikoo, M. (2018), “Krill herd algorithm-based neural network in structural seismic reliability evaluation”, *Mech. Adv. Mater. Struct.*, **26**(13), 1146-1153. https://doi.org/10.1080/15376494.2018.1430874.

Asteris, P.G., Roussis, P.C. and Douvika, M.G. (2017), “Feed-forward neural network prediction of the mechanical properties of sandcrete materials”, *Sens., (Switzerland)*, **17**(6), 1344.

Asteris, P.G., Tsaris, A.K., Cavaleri, L., Repapis, C.C., Papalou, A., Di Trapani, F. and Karypidis, D.F. (2016), “Prediction of the fundamental period of infilled RC frame structures using artificial neural networks”, *Comput. Intel. Neurosci.*, 5104907. https://doi.org/10.1155/2016/5104907.

AzariJafari, H., Amiri, M.J.T., Ashrafian, A., Rasekh, H., Barforooshi, M.J. and Berenjian, J. (2019), “Ternary blended cement: an eco-friendly alternative to improve resistivity of high-performance self-consolidating concrete against elevated temperature”, *J. Clean. Prod.*, **223**, 575-586. https://doi.org/10.1016/j.jclepro.2019.03.054.

Badogiannis, E.G., Sfikas, I.P., Voukia, D.V., Trezos, K.G. and Tsivilis, S.G. (2015), “Durability of metakaolin self-compacting concrete”, *Constr. Build. Mater.*, **82**, 133-141. https://doi.org/10.1016/j.conbuildmat.2015.02.023.

Chen, H., Asteris, P.G., Armaghani, D.J., Gordan, B. and Pham, B.T. (2019), “Assessing dynamic conditions of the retaining wall using two hybrid intelligent models”, *Appl. Sci*., **9**, 1042.

https://doi.org/10.3390/app9061042.
Chitroju, S.T.D. and Yerikenaboina, A. (2018), "Study the influence of metakaolin and foundry sand on self-compacting concrete properites", *Int. Res. J. Eng. Technol. (IRJET)*, **5**(4), 4027-4033.
Coleman, N.J. and Page, C.L. (1997), "Aspects of the pore solution chemistry of hydrated cement pastes containing metakaolin", *Cement Concrete Res.*, **27**(1), 147-154. https://doi.org/10.1016/S0008-8846(96)00184-6.
Dadsetan, S. and Bai, J. (2017), "Mechanical and microstructural properties of self-compacting concrete blended with metakaolin, ground granulated blast-furnace slag and fly ash", *Constr. Build. Mater.*, **146**, 658-667. https://doi.org/10.1016/j.conbuildmat.2017.04.158.
Dinakar, P. and Manu, S.N. (2014), "Concrete mix design for high strength self-compacting concrete using metakaolin", *Mater. Des.*, **60**, 661-668. https://doi.org/10.1016/j.matdes.2014.03.053.
Ferreira, R.M., Castro-Gomes, J.P., Costa, P. and Malheiro, R. (2016), "Effect of metakaolin on the chloride ingress properties of concrete", *KSCE J. Civil Eng.*, **20**(4), 1375-1384. https://doi.org/10.1007/s12205-015-0131-8.
Frías, M. and Cabrera, J. (2000), "Pore size distribution and degree of hydration of metakaolin–cement pastes", *Cement Concrete Res.*, **30**(4), 561-569. https://doi.org/10.1016/S0008-8846(00)00203-9.
Friedman, J.H. (1991), "Multivariate adaptive regression splines", *Ann. Statist.*, **19**(1), 1-67.
Gholampour, A.A., Gandomi, A.H. and Ozbakkaloglu, T. (2017), "New formulations for mechanical properties of recycled aggregate concrete using gene expression programming", *Constr. Build. Mater.*, **130**, 122-145. https://doi.org/10.1016/j.conbuildmat.2016.10.114.
Gilan, S.S., Bahrami Jovein, H. and Ramezanianpour, A.A. (2012), "Hybrid support vector regression–Particle swarm optimization for prediction of compressive strength and RCPT of concretes containing metakaolin", *Constr. Build. Mater.*, **34**, 321-329. https://doi.org/10.1016/j.conbuildmat.2012.02.038.
Gilan, S.S., Jovein, H.B. and Ramezanianpour, A.A. (2012), "Hybrid support vector regression-particle swarm optimization for prediction of compressive strength and RCPT of concretes containing metakaolin", *Constr. Build. Mater.*, **34**, 321-329. https://doi.org/10.1016/j.conbuildmat.2012.02.038.
Gill, A.S. and Siddique, R. (2018), "Durability properties of self-compacting concrete incorporating metakaolin and rice husk ash", *Constr. Build. Mater.*, **176**, 323-332. https://doi.org/10.1016/j.conbuildmat.2018.05.054.
Güneyisi, E., Gesoglu, M. and Özbay, E. (2009), "Evaluating and forecasting the initial and final setting times of self-compacting concretes containing mineral admixtures by neural network", *Mater. Struct.*, **42**(4), 469-484. https://doi.org/10.1617/s11527-008-9395-5.
Hassan, A.A.A., Ismail, M.K. and Mayo, J. (2015), "Mechanical properties of self-consolidating concrete containing lightweight recycled aggregate in different mixture compositions", *J. Build. Eng.*, **4**, 113-126. https://doi.org/10.1016/j.jobe.2015.09.005.
Hassan, A.A.A., Lachemi, M. and Hossain, K.M.A. (2012a), "Effect of metakaolin and silica fume on the durability of self-consolidating concrete", *Cement Concrete Compos.*, **34**(6), 801-807. https://doi.org/10.1016/j.cemconcomp.2012.02.013.
Hassan, A.A.A., Lachemi, M. and Hossain, K.M.A. (2012b), "Effect of metakaolin and silica fume on the durability of self-consolidating concrete", *Cement Concrete Compos.*, **34**(6), 801-807. https://doi.org/10.1016/j.cemconcomp.2012.02.013.
Hastie, T., Tibshirani, R. and Friedman, J. (2009), *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd Edition, Springer.
Jekabsons, G. (2010), "VariReg: a software tool for regression modelling using various modeling methods", Riga Technical University, http://www.cs.rtu.lv/ jekabsons/.
Johari, M.M., Brooks, J.J., Kabir, S. and Rivard, P. (2011), "Influence of supplementary cementitious materials on engineering properties of high strength concrete", *Constr. Build. Mater.*, **25**(5), 2639-2648. https://doi.org/10.1016/j.conbuildmat.2010.12.013.
Joseph, A., Mathew, L.A. and John, R. (2017), "Performance of Metakaolin on high strength self compacting concrete", *Int. J. Sci. Technol. Eng.*, **3**(12), 110-114.
Justice, J.M. and Kurtis, K.E. (2007), "Influence of metakaolin surface area on properties of cement-based materials", *J. Mater. Civil Eng.*, **19**(9), 762-771. https://doi.org/10.1061/(ASCE)0899-1561(2007)19:9(762).
Kannan, V. and Ganesan, K. (2015), "Effect of Tricalcium aluminate on durability properties of self-compacting concrete incorporating rice husk ash and Metakaolin", *J. Mater. Civil Eng.*, **28**(1), 04015063. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001330.
Kavitha, O.R., Shanthi, V.M., Arulraj, G.P. and Sivakumar, P. (2015), "Fresh, micro-and macrolevel studies of metakaolin blended self-compacting concrete", *Appl. Clay Sci.*, **114**, 370-374. https://doi.org/10.1016/j.clay.2015.06.024.
Khatib, J.M. (2008), "Metakaolin concrete at a low water to binder ratio", *Constr. Build. Mater.*, **22**(8), 1691-1700. https://doi.org/10.1016/j.conbuildmat.2007.06.003.
Khatib, J.M. (2008), "Performance of self-compacting concrete containing fly ash", *Constr. Build. Mater.*, **22**(9), 1963-1971. https://doi.org/10.1016/j.conbuildmat.2007.07.011.
Kiani, B., Gandomi, A.H., Sajedi, S. and Liang, R.Y. (2016), "New formulation of compressive strength of preformed-foam cellular concrete: an evolutionary approach", *J. Mater. Civil Eng.*, **28**(10), 04016092. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001602.
Lenka, S. and Panda, K.C. (2017), "Effect of metakaolin on the properties of conventional and self compacting concrete", *Adv. Concrete Constr.*, **5**(1), 31-48. https://doi.org/10.12989/acc.2017.5.1.31.
Madandoust, R. and Mousavi, S.Y. (2012), "Fresh and hardened properties of self-compacting concrete containing metakaolin", *Constr. Build. Mater.*, **35**, 752-760. https://doi.org/10.1016/j.conbuildmat.2012.04.109.
Mansouri, I., Ozbakkaloglu, T., Kisi, O. and Xie, T. (2016), "Predicting behavior of FRP-confined concrete using neuro fuzzy, neural network, multivariate adaptive regression splines and M5 model tree techniques", *Mater. Struct.*, **49**(10), 4319-4334. https://doi.org/10.1617/s11527-015-0790-4.
Mehrinejad Khotbehsara, M., Mohseni, E., Ozbakkaloglu, T. and Ranjbar, M.M. (2017), "Durability characteristics of self-compacting concrete incorporating pumice and metakaolin", *J. Mater. Civil Eng.*, **29**(11), 04017218. https://doi.org/10.1061/(ASCE)MT.1943-5533.0002068.
Mehta, P.K. (1978), "Siliceous ashes and hydraulic cements prepared therefrom", U.S. Patent No. 4,105,459, The Regents of the University of California.
Poon, C.S., Kou, S.C. and Lam, L. (2006), "Compressive strength, chloride diffusivity and pore structure of high performance metakaolin and silica fume concrete", *Constr. Build. Mater.*, **20**(10), 858-865. https://doi.org/10.1016/j.conbuildmat.2005.07.001.
Quinlan, J.R. (1992), "Learning with continuous classes", Ed. Adams, S., *Proceedings of AI'92. World Scientific*, 343-348.
Ramezanianpour, A.A. and Jovein, H.B. (2012), "Influence of metakaolin as supplementary cementing material on strength and durability of concretes", *Constr. Build. Mater.*, **30**, 470-479. https://doi.org/10.1016/j.conbuildmat.2011.12.050.

Rezaie-Balf, M., Zahmatkesh, Z. and Kim, S. (2017), "Soft computing techniques for rainfall-runoff simulation: Local non-parametric paradigm vs. model classification methods", *Water Resour. Manage.*, **31**(12), 3843-3865. https://doi.org/10.1007/s11269-017-1711-9.

Sabir, B.B., Wild, S. and Bai, J. (2001), "Metakaolin and calcined clays as pozzolans for concrete: a review", *Cement Concrete Compos.*, **23**(6), 441-454. https://doi.org/10.1016/S0958-9465(00)00092-5.

Sattar, A.M.A. and Gharabaghi, B. (2015), "Gene expression models for prediction of longitudinal dispersion coefficient in streams", *J. Hydrol.*, **524**, 587-596. https://doi.org/10.1016/j.jhydrol.2015.03.016.

Sfikas, I.P., Badogiannis, E.G. and Trezos, K.G. (2014), "Rheology and mechanical characteristics of self-compacting concrete mixtures containing metakaolin", *Constr. Build. Mater.*, **64**, 121-129. https://doi.org/10.1016/j.conbuildmat.2014.04.048.

Šipoš, T.K., Miličević, I. and Siddique, R. (2017), "Model for mix design of brick aggregate concrete based on neural network modelling", *Constr. Build. Mater.*, **148**, 757-769. https://doi.org/10.1016/j.conbuildmat.2017.05.111.

Sonebi, M., Cevik, A., Grunewald, S. and Walravan, J. (2016b), "Modelling the fresh properties of self-compacting concrete using support vector machine approach", *Constr. Build. Mater.*, **106**, 55-64. https://doi.org/10.1016/j.conbuildmat.2015.12.035.

Sonebi, M., Grünewald, S., Cevik, A. and Walraven, J. (2016a), "Modelling fresh properties of self-compacting concrete using Neural network technique", *Comput. Concrete*, **18**(4), 903-921. https://doi.org/10.12989/cac.2016.18.4.903.

Tropsha, A., Gramatica, P. and Gombar, V.K. (2003), "The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models", *QSAR Combin. Sci.*, **22**(1), 69-77. https://doi.org/10.1002/qsar.200390007.

Wang, B., Man, T. and Jin, H. (2015), "Prediction of expansion behavior of self-stressing concrete by artificial neural networks and fuzzy inference systems", *Constr. Build. Mater.*, **84**, 184-191. https://doi.org/doi.org/10.1016/j.conbuildmat.2015.03.059.

Wang, Y. and Witten, I.H. (1997), "Induction of model trees for predicting continuous lasses", *Proceedings of the Poster Papers of the European Conference on Machine Learning*, University of Economics, Faculty of Informatics and Statistics, Prague.

Wild, S., Khatib, J.M. and Jones, A. (1996), "Relative strength, pozzolanic activity and cement hydration in superplasticised metakaolin concrete", *Cement Concrete Res.*, **26**(10), 1537-1544. https://doi.org/10.1016/0008-8846(96)00148-2.

Wild, S., Khatib, J.M. and Jones, A. (1996), "Relative strength, pozzolanic activity and cement hydration in superplasticised metakaolin concrete", *Cement Concrete Res.*, **26**(10), 1537-1544. https://doi.org/10.1016/0008-8846(96)00148-2.

Witten, I.H., Frank, E. and Hall, M.A. (2005), "Data mining: practical machine learning tools and techniques with Java implementations", Morgan Kaufmann Series in Data Management Systems, San Francisco.

Zhang, W.G. and Goh, A.T.C. (2013), "Multivariate adaptive regression splines for analysis of geotechnical engineering systems", *Comput. Geotech.*, **48**, 82-95. https://doi.org/10.1016/j.compgeo.2012.09.016.

Zhang, W.G. and Goh, A.T.C. (2016), "Multivariate adaptive regression splines and neural network models for prediction of pile drivability", *Geosci. Front.*, **7**(1), 45-52. https://doi.org/10.1016/j.gsf.2014.10.003.

CC