Structural Engineering and Mechanics, *Vol. 70, No. 6 (2019) 671-681*
DOI: https://doi.org/10.12989/sem.2019.70.6.671

Prediction of compressive strength of bacteria incorporated geopolymer concrete by using ANN and MARS

John Britto X.[a] and Muthuraj M.P.*

Department of civil Engineering, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India 641014

(Received February 2, 2019, Revised March 9, 2019, Accepted March 16, 2019)

Abstract. This paper examines the applicability of artificial neural network (ANN) and multivariate adaptive regression splines (MARS) to predict the compressive strength of bacteria incorporated geopolymer concrete (GPC). The mix is composed of new bacterial strain, manufactured sand, ground granulated blast furnace slag, silica fume, metakaolin and fly ash. The concentration of sodium hydroxide (NaOH) is maintained at 8 Molar, sodium silicate (Na_2SiO_3) to NaOH weight ratio is 2.33 and the alkaline liquid to binder ratio of 0.35 and ambient curing temperature (28°C) is maintained for all the mixtures. In ANN, back-propagation training technique was employed for updating the weights of each layer based on the error in the network output. Levenberg-Marquardt algorithm was used for feed-forward back-propagation. MARS model was developed by establishing a relationship between a set of predictors and dependent variables. MARS is based on a divide and conquers strategy partitioning the training data sets into separate regions; each gets its own regression line. Six models based on ANN and MARS were developed to predict the compressive strength of bacteria incorporated GPC for 1, 3, 7, 28, 56 and 90 days. About 70% of the total 84 data sets obtained from experiments were used for development of the models and remaining 30% data was utilized for testing. From the study, it is observed that the predicted values from the models are found to be in good agreement with the corresponding experimental values and the developed models are robust and reliable.

Keywords: geopolymer concrete; bacteria; compressive strength; artificial neural network; multivariate adaptive regression splines

1. Introduction

Due to the continuous depletion of the ozone layer and global warming, awareness of the eco-friendlier construction materials has been increasing. In view of this, GPC has gained significant attention by the research community, construction practitioners, and design engineers.

It is very well known that GPC can be produced by utilizing several industrial by-products such as fly ash (FA), metakaolin (MK), slag (SG), rice husk ash (RHA), and high calcium wood ash (HCWA) through polymerization and using alkaline solution. The quality and efficiency of GPC will depend upon activators and type and source of aluminosilicates. It possesses superior mechanical and durability properties against conventional concrete and several studies were reported on GPC (Guo *et al*. 2010, Pacheco-Torgal *et al*. 2011, Shi *et al*. 2012, Hardjito *et al*. 2004, Sumajouw *et al*. 2007, Fernandez-Jimenez *et al*. 2006).

In view of difficulty in conducting experiments numerous times and to reduce time and effort, analytical models/numerical models will be useful to predict the mechanical properties of several concretes. Several advanced statistical models, namely, Artificial Neural Network, Gaussian regression process, relevance vector machine, least squares support vector machine, extreme learning machine and multivariate adaptive regression splines are available in the literature to train the data and to predict the response of the structural components or mechanical and durability properties of concrete mixes (Yuvaraj *et al*. 2013a, Yuvaraj *et al*. 2013b, Yuvaraj *et al*. 2014a, Yuvaraj *et al*. 2014b, Shantaram *et al*. 2014, Vishal *et al*. 2014, Susom Dutta *et al*. 2017, Jaideep Kaur *et al*. 2017, Erdem 2017, Engin *et al*. 2015, Mansouri *et al*. 2016, Prasanna *et al*. 2018).

An artificial neural network (ANN) is a mathematical model developed from the inspiration of biological neural networks. ANN consists of group of artificial neurons interconnected to each other, and it processes information using a connectionist approach to computation. ANN model for predicting fracture toughness and tensile strength was successfully demonstrated by Mohammed and Sudhakar (2002). Ince (2004) predicted facture parameters of concrete by artificial neural networks and compared the results of two-parameter model (TPM) and ANN. ANN was proven to be successful in civil engineering (Hakan 2009, Siddique *et al*. 2011, Loizos and Karlaftis 2006).

MARS is relatively a new technique used for modeling data depicting non-linear relationship (Friedman, 1991). It establishes the relationship in non-linear form between the response and predictor variables and identifies the interactions and conditional relationships among the predicator variables. MARS was successfully applied in

*Corresponding author, Assistant Professor
E-mail: muthuraj@cit.edu.in
[a] Ph.D. Student

ISSN: 1225-4568 (Print), 1598-6217 (Online)

various fields including biological sciences, cancer research (Mallick *et al*. 1997), communication (Ekman and Kubin 1999), Engineering (Jin *et al*. 2000, Nii *et al*. 2009) and genetics (York and Eaves 2001). From the available literature, it is observed that MARS is employed in various fields and to the best of authors' knowledge, no research investigations are reported in the field of GPC.

In the present investigation, it is proposed to employ artificial neural network and MARS to predict the compressive strength of various bacterial based GPC mixes.

2. Compressive strength of various GPC mixes

Materials used for preparation of GPC are manufactured sand (M-sand), ground granulated blast furnace slag (GGBS), silica fume (SF), metakaolin (MK) and low calcium fly ash (ASTM class F). Physical properties of fly ash (FA), GGBS, SF and MK are presented in Table 1.

Combination of NaOH solution and Na_2SiO_3 solution was used as alkaline activator solution (AAS). SP used was Master Glenium Sky 8233. Several trial mixes (range of molarity: 4 to 16; the ratio of alkaline liquids to GGBS: 0.3 to 0.45; the ratio of sodium silicate to sodium hydroxide: 2 to 2.5) were carried out and the desired mechanical properties were obtained for the combination of 8 Molar, the ratio of sodium silicate (Na2SiO3) to NaOH weight ratio as 2.33 and the ratio of alkaline liquid to binder ratio as 0.35. Ambient curing temperature (28^{o}C) is maintained for all the mixtures. Two new *Bacillus strains such as Bacillus pumilus* (B1) and *Bacillus velezensis* (B2) isolated from manufacturing sand were used for preparation of GPC. The cell concentration was 10^6 cells/ ml. Compressive strength is determined for four broader mixes with and without bacteria. The size of the cube is 150mm. For all the mixes, GGBS is fixed as 70%. Molarity is 8M and SP is 2.5% of total cementitious materials. There are total four control mixes (without bacteria), namely, 8M1C (GGBS = 70% and FA = 30%), 8M2C (GGBS = 70%, FA=15%, SF=10%, MK=5%), 8M3C (GGBS=70%, FA=17%, SF=5%, MK=8%) and 84M4C (GGBS=70%, FA=12%, SF=10%, MK=8%). Each control mix is incorporated with bacteria (B1 or B2) in varied proportions i.e. 12.5, 25, 37.5, 50, 62.5, 75, 87.5 and 100 ml/litre. Detailed input and output data is shown in Table 3.

Table 2 shows the physical properties coarse and fine aggregates.

Table 1 Physical properties of GGBS, SF, MK and FA

Physical Properties	GGBS	SF	MK	FA
Specific gravity	2.25	2.95	2.5	2.05
Mean Particle size (MPS) (μm)	12.56	79.38	6.25	14.06
Multipoint BET Fineness (m²/kg)	3,700	21,410	14,970	2,270
Density (Kg/m^3)	1275	410	840	800

Table 2 Physical properties of coarse aggregates and Fine Aggregates

Physical Properties	20mm	12.5mm	M-sand
Specific gravity	2.84	2.84	2.78
Water Absorption (%)	0.48	0.51	2.35
Dry Loose Bulk Density (Kg/m^3)	1499	1472	1627
Material finer than 75-μ IS Sieve (%)	0.12	0.16	4.40
Fineness Modulus	6.5	6.3	2.68
Impact Value (%)	18	17	-
Crushing Value (%)	21	22	-
Flakiness Index (%)	7	5	-
Elongation Index (%)	8	9	-
Clay Lumps (%)	Nil	Nil	Nil

3. Artificial neural network

ANNs learn through the example problems rather than programming. Although detailed methodology of ANN has been reported in literature (Mohammed *et al*. 2002, Ince 2004, Kamarthi and Pittner 1999), a brief description about the development of model is described below. Neural networks can be classified into two different categories, feed-forward and feed-back networks. The feedback networks contain nodes that can be connected to them, enabling a node to influence other nodes as well as it. In Feed-Forward Networks (FFNs), the signals from the input neurons to the output neurons flow only in one direction. Feed-forward ANNs are straight forward networks (no loops) that associate inputs with outputs. They are extensively used in pattern recognition. This type of organisation is also referred to as top-down. The information distribution is parallel for all the nodes of the succeeding layer. Back-propagation neural networks are adopted in the present study, as they have a high capability of data mapping (Hecht 1990).

The Back-propagation learning is based on the gradient descent along the error surface where in the weight adjustment is proportional to the negative gradient of the error with respect to the weight.

In mathematical form

$$W_{k+1} = W + nd_k \quad (1)$$

where, W_k is the weight matrix at η epoch k. The direction vector d_k is negative of the gradient of the output error function 'E' and is given by the equation 2.

$$d_k = - \nabla E(w_k) \quad (2)$$

There are two standard learning schemes for the back propagation algorithm, namely, on-line learning and batch learning. In on-line learning, the weights of the network are updated immediately after the presentation of each pair of input and target patterns. In batch learning, all the pairs of patterns in the training sets will be treated as a batch and the network is updated after processing of all training patterns in the batch. In either case the vector w_k contains the weights computed during kth iteration, and the output error function E is a multivariate function of the weights in the network (Kamarthi *et al*. 1999) and is as follows (equation 3).

Table 3 Compressive strength of bacteria incorporated geo polymer concrete

Mix ID	FA	SF	MK	Bacillus bacteria ml/litre	Average compressive strength, MPa					
					1 Day	3 day	7 day	28 day	56 day	90 day
8M1C	30	0	0	0	18.62	30.48	39.57	44.87	48.31	50.08
8M1B112.5	30	0	0	12.5	19.24	31.97	42.37	48.39	51.09	53.17
8M1B125	30	0	0	25	19.31	32.35	42.66	49.44	52.84	54.81
8M1B137.5	30	0	0	37.5	19.38	32.63	42.97	50.46	53.43	55.96
8M1B150	30	0	0	50	19.68	33.15	43.68	51.48	55.04	57.11
8M1B162.5	30	0	0	62.5	19.72	33.93	46.17	53.71	57.49	59.17
8M1B175	30	0	0	75	20.14	34.28	46.67	54.82	58.59	60.82
8M1B187.5	30	0	0	87.5	20.21	34.39	46.78	54.96	58.74	60.98
8M1B1100	30	0	0	100	20.24	34.5	46.89	55.1	58.89	61.13
8M1B1112.5	30	0	0	112.5	20.17	34.29	46.66	54.91	58.77	60.95
8M1B1125	30	0	0	125	19.98	34.06	46.58	54.73	58.47	60.78
8M1B212.5	30	0	0	12.5	19.38	32.43	42.19	48.91	52.09	55.07
8M1B225	30	0	0	23	19.41	32.73	43.26	50.65	54.17	56.32
8M1B237.5	30	0	0	37.5	19.46	33.14	43.94	51.99	55.68	57.37
8M1B250	30	0	0	50	19.94	33.55	45.1	53.42	57.18	59.35
8M1B262.5	30	0	0	62.5	20.39	34.88	47.27	56.19	60.08	62.67
8M1B275	30	0	0	75	20.44	35.64	47.66	57.06	61	63.31
8M1B287.5	30	0	0	87.5	20.46	35.76	47.96	57.20	61.16	63.49
8M1B2100	30	0	0	100	20.65	35.87	48.25	57.33	61.31	63.66
8M1B2112.5	30	0	0	112.5	20.49	35.71	47.99	57.27	61.13	63.46
8M1B2125	30	0	0	125	20.41	35.59	47.52	57.13	61.03	63.33
8M2C	15	10	5	0	30.47	50.17	65.14	74.02	79.97	83.07
8M2B112.5	15	10	5	12.5	31.68	52.87	68.98	80.07	85.67	89.17
8M2B125	15	10	5	25	31.72	53.63	70.81	82.46	88.37	91.94
8M2B137.5	15	10	5	37.5	31.87	53.98	71.36	83.53	89.07	92.08
8M2B150	15	10	5	50	32.32	54.53	71.91	84.6	90.69	94.47
8M2B162.5	15	10	5	62.5	32.89	55.58	75.19	88.43	94.17	98.07
8M2B175	15	10	5	75	33.02	56.62	76.98	89.7	96.26	100.16
8M2B187.5	15	10	5	87.5	33.09	56.73	77.15	90.26	96.90	100.75
8M2B1100	15	10	5	100	33.15	56.84	77.32	90.82	97.54	101.34
8M2B1112.5	15	10	5	112.5	33.11	56.75	77.13	90.61	96.91	100.73
8M2B1125	15	10	5	125	33.07	56.6	77.29	90.09	96.37	100.59
8M2B212.5	15	10	5	12.5	31.69	52.67	70.09	80.11	86.73	90.07
8M2B225	15	10	5	25	31.75	53.89	71.19	83.57	89.63	93.26
8M2B237.5	15	10	5	37.5	31.92	54.09	72.11	85.12	91.87	97.08
8M2B250	15	10	5	50	32.54	55.14	74.06	88.45	94.9	98.75
8M2B262.5	15	10	5	62.5	32.97	57.29	76.67	91.39	98.07	102.17
8M2B275	15	10	5	75	33.18	57.72	77.99	92.23	99.15	103.06
8M2B287.5	15	10	5	87.5	33.21	58.07	78.21	93.97	100.39	104.34
8M2B2100	15	10	5	100	33.76	59.15	79.21	94.67	101.62	105.62
8M2B2112.5	15	10	5	112.5	33.45	58.53	78.54	93.48	100.48	104.12
8M2B2125	15	10	5	125	33.29	58.08	78.08	92.97	99.81	103.87
8M3C	17	5	8	0	29.58	50.02	64.96	74.07	80.09	83.36
8M3B112.5	17	5	8	12.5	30.69	52.97	68.99	81.19	86.09	89.39
8M3B125	17	5	8	25	30.81	53.47	70.88	83.01	89.15	92.93
8M3B137.5	17	5	8	37.5	31.07	54.01	71.52	83.97	90.34	93.97

Table 3 (continued)

8M3B150	17	5	8	50	31.38	54.54	72.15	85.31	91.52	95.36
8M3B162.5	17	5	8	62.5	31.97	56.07	75.57	86.83	95.71	99.17
8M3B175	17	5	8	75	32.09	56.66	76.83	90.51	97.21	101.29
8M3B187.5	17	5	8	87.5	32.18	56.78	77.12	90.73	97.43	101.53
8M3B1100	17	5	8	100	32.27	56.9	77.41	90.95	97.64	101.76
8M3B1112.5	17	5	8	112.5	32.23	56.75	77.24	90.86	97.55	101.56
8M3B1125	17	5	8	125	32.14	56.69	75.48	90.44	97.28	101.3
8M3B212.5	17	5	8	12.5	30.62	52.81	68.99	80.37	86.97	90.08
8M3B225	17	5	8	25	30.83	53.78	71	83.76	89.84	93.59
8M3B237.5	17	5	8	37.5	31.01	54.11	71.97	85.39	91.32	97.09
8M3B250	17	5	8	50	31.66	55.07	74.07	88.27	94.8	98.75
8M3B262.5	17	5	8	62.5	32.10	57.09	76.31	91.17	97.19	101.73
8M3B275	17	5	8	75	32.23	57.57	77.93	92.05	98.82	103.01
8M3B287.5	17	5	8	87.5	32.53	58.28	78.11	93.07	100.39	104.62
8M3B2100	17	5	8	100	32.83	58.98	79.23	94.89	101.96	106.23
8M3B2112.5	17	5	8	112.5	32.5	58.19	78.43	93.11	100.28	105.17
8M3B2125	17	5	8	125	32.08	57.97	77.99	92.63	99.12	104.11
8M4C	12	10	8	0	30.32	51.73	67.88	77.97	84.39	88.11
8M4B112.5	12	10	8	12.5	31.20	54.17	72.91	84.37	91.97	95.01
8M4B125	12	10	8	25	31.47	55.34	74.04	87.4	93.93	98.27
8M4B137.5	12	10	8	37.5	31.81	55.84	74.73	88.62	94.97	99.64
8M4B150	12	10	8	50	32.15	56.33	75.41	89.83	96.56	101
8M4B162.5	12	10	8	62.5	32.34	56.75	78.19	93.09	100.11	105.04
8M4B175	12	10	8	75	32.53	57.16	79.49	94.45	101.5	106.09
8M4B187.5	12	10	8	87.5	32.81	58.15	80.24	95.26	102.36	107.00
8M4B1100	12	10	8	100	33.09	59.13	80.98	96.06	103.22	107.91
8M4B1112.5	12	10	8	112.5	32.96	58.31	80.07	95.79	102.96	106.87
8M4B1125	12	10	8	125	32.71	57.98	79.81	94.97	101.97	106.13
8M4B212.5	12	10	8	12.5	31.40	54.17	72.67	85.07	92.19	95.17
8M4B225	12	10	8	25	31.62	55.71	74.4	88.24	94.74	99.09
8M4B237.5	12	10	8	37.5	31.92	56.39	75.17	89.99	96.37	101.08
8M4B250	12	10	8	50	32.47	57.06	77.53	93.74	100.82	105.31
8M4B262.5	12	10	8	62.5	33.20	59.87	80.37	97.53	104.81	109.81
8M4B275	12	10	8	75	33.45	60.93	82.26	98.09	106.74	111.64
8M4B287.5	12	10	8	87.5	33.56	61.01	82.58	99.59	107.31	111.98
8M4B2100	12	10	8	100	33.66	61.09	82.9	99.87	107.44	112.31
8M4B2112.5	12	10	8	112.5	33.47	60.98	82.47	99.47	107.13	111.17
8M4B2125	12	10	8	125	33.11	59.18	82.07	99.03	106.79	109.57

$$E(w_k) = \begin{cases} E_p(w_k)[\text{on} - \text{line}] \\ \sum_p E_p(w_k)[\text{batch}] \end{cases} \quad (3)$$

Where, $E_p(W_k)$ signifies the half - sum – of – squares error functions of the network output for a certain input pattern 'p'. The purpose of the supervised learning (or training) is to ascertain a set of weight that can minimize the error E over the complete set of training pair. Every cycle in which each one of the training patterns is presented once to the neural network is called an epoch. The direction vector d_k, expressed in terms of error gradient depends upon the choice of activation function. When the sigmoid function i.e. $f(x) = (1+\exp(-x))^{-1}$ is adopted, the back propagation algorithm turns as 'Back propagation for the Sigmoid Adaline'(Widrow *et al*. 1990). Generally, the back-propagation learning apply the weight change proportional to the negative gradient of the instantaneous error. It considers the only the first derivative of the instantaneous

error with respect to the weight. A more effective method can be derived by using Taylor series expansion of the error as a function of the weight vector

$$E(w+\Delta w) = E(w) + g^T \Delta w + \frac{1}{2}\Delta w + \ldots \quad (4)$$

Where, $g = \frac{\Delta E}{\Delta w}$ is the gradient vector, and $H = \frac{\partial^2 w}{\partial w^2}$ the hessian matrix. Newton's method of Back propagation learning uses this Hessian matrix as a component of weight update.

4. Multivariate adaptive regression splines

The ultimate aim of the Multivariate adaptive regression splines (MARS) model is to capture the relationship between the dependent variable and the independent variable from the data. In general, MARS function can be represented by using the following equation (Friedman, 1991)

$$y = \hat{f}(x) = a_0 + \sum_{m=1}^{M} a_m B_m^{(q)}(x) \quad (5)$$

where,

a_0 = coefficient of the constant basis function, or the constant term

$\{a_m\}_1^M$ = vector of coefficients of the non-constant basis functions, m = 1, 2, …, M

$B_m^{(q)}$ are the basis functions that are selected for inclusion in the model of q^{th} order

$$B_m^{(q)}(x) = \prod_{k-1}^{k_m} \left[s_{km} \cdot \left(x_{v(k,m)} - t_{km} \right) \right]_+^q \quad (6)$$

where

$B_m^{(q)}(x)$ = vector of non-constant (truncated) basis functions, or the tensor product spline basis

m = number of non-constant basis functions (1, 2… M)

q = the power to which the spline is raised in order to control the degree of smoothness of the resultant function estimate, which in this case is equal to 1'+' = denotes that only positive results of the right-hand side of the equation are considered; otherwise, the functions evaluate to 0.

S_{km} = indicates the (left/right) sense of truncation, which assumes only 2 values (± 1), representing the standard basis function and its mirror image. For S_{km} equal to +1, the basis function will have a value x-t if x>t and 0 if x≤ t. If it is -1, the basis function will have a value t-x when x<t, while 0 if x ≥ t

$X_{v(k,m)}$ = value of the predictor

v(k,m) = label of the predictor ($1 \leq v(k,m) \leq n$

n = number of predictors

t_{km} = "knot" location on the corresponding predictor space or region, or value that defines an inflection point along the range of the predictor

K= maximum level or order of interaction, or the number of factors, in the m^{th} basis function (1, 2,…, K_m) Basis functions are a set of functions used to represent the information contained in one or more variables. Like principal components, basis functions re-express the relationship of the predictors with the dependent variable.

Parameters of Mars can be estimated by the Penalized Least Squares (PLS) with the form:

$$P(x) = \min \sum \left(y_i - \hat{f}(x_i) \right)^2 + \lambda \int f^n(x_i) dx_i \quad (7)$$

In the above equation the first term represents the residual sum of squares and the second term represents the roughness penalty term, which is weighted by λ (known as the smoothing constant).

The penalty term is large when the integrated second derivative of the regression function $f''(x)$ is large – that is, when f (x) is 'rough' (with rapidly changing slope). At one extreme, when the λ is set to zero (and if all the values of x are distinct), the objective function simply interpolates the data. At the other extreme, if λ is very large, then the objective function will be selected so that its second derivative is everywhere zero, implying a globally linear least-squares fit to the data.

When fitting a MARS model, knots are chosen in an iterative (recursive, i.e., from low to high interaction order) forward stepwise procedure. After over-fitting the model with so many basis functions, a backward spurning or snubbing procedure is applied in which those basis functions that contribute least to model fit are progressively removed. At this stage, a predictor variable can be dropped from the model completely if none of its basis functions contribute meaningfully to predictive performance. The sequence of models generated from this process is then evaluated using the Generalized Cross-Validation (GCV), and the model with the best predictive fit is finally selected. The GCV can be expressed as follows

$$GCV(M) = \frac{\frac{1}{N}\sum_{i=1}^{N}\left[y_i - \hat{f}_M(x_i) \right]^2}{\left[1 - \frac{C(M)}{N} \right]^2} \quad (8)$$

Where, the numerator denotes lack-of-fit on the training data (sort of "bias") and the denominator accounts the (inverse) penalty for increasing model complexity C(M) (sort of "variance")

N = observations

C(M) = Cost penalty measures of a model

M = basis functions

Fm(xi)= basis function model

MARS minimizes GCV (M), which reduces the bias of the model estimates but at the same time increases the variance due to additional parameters included to improve the fit of the model.

In order to reveal considerable information about the predictive relationship between the dependent variable and a set of predictors, eq. (5) can be recast into the following form:

$$\hat{f}(x) = a_0 + \sum_{K_m=1} f_i(x_i) + \sum_{K_m=2} f_{ij}(x_i, x_j) + \sum_{K_m=3} f_{ijk}(x_i, x_j, x_k) + \ldots \quad (9)$$

This is referred to as the ANOVA decomposition of the

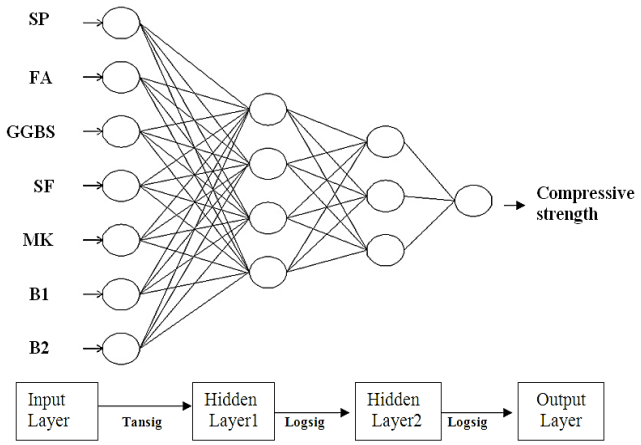


Fig. 1 Typical Architecture of ANN

MARS model. The first sum is over all basis functions that involve only a single variable. The second sum is over all basis functions that involve exactly two variables, representing (if present) two-variable interactions. Similarly, the third sum represents (if present) the contributions from three-variable interactions and so on. MARS can also handle the missing value problems using dummy variable skills. By allowing any arbitrary shape for the function and interactions, MARS is capable of tracking very complex data structures that hide in high dimensional frequently data. More details regarding the model building process can be found in Friedman (1991).

5. Development of models

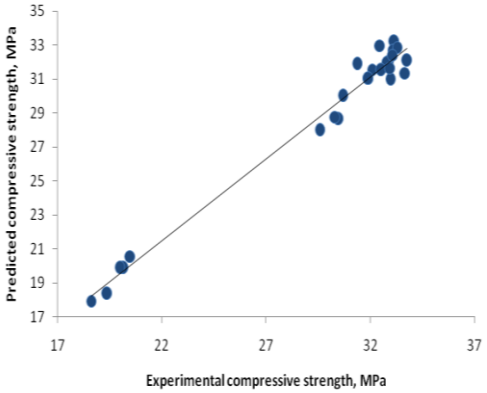
Six individual ANN models and six individual MARS models were developed for the prediction of compressive strength of various GPC mixes. The compressive is primarily dependent on fly ash, GGBS, SF, MK, Bacteria 1 or Bacteria 2 and SP. The compressive strength was predicted for 1, 3, 7, 28, 56 and 90 days. MATLAB software was used to develop models. The data that forms an input vector has different quantitative limits as shown in Table 3. Normalization of the data is to be carried out before presenting the input patterns to ANN and MARS. Eq. (7) is used for the linear normalization of the data to the data values between 0 and 1.

$$x_i^n = \frac{x_i^a - x_i^{\min}}{x_i^{\max} - x_i^{\min}} \tag{10}$$

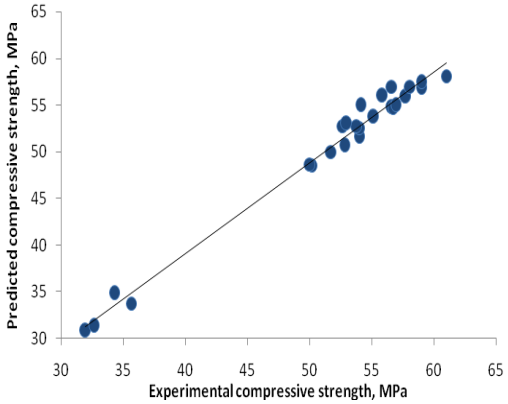
Where, x_i^a and x_i^n are the i[th] components of the input vector before and after normalization, respectively, and x_i^{max} and x_i are the maximum and minimum values of all the components of the input vector before the normalization.

5.1 ANN based analysis

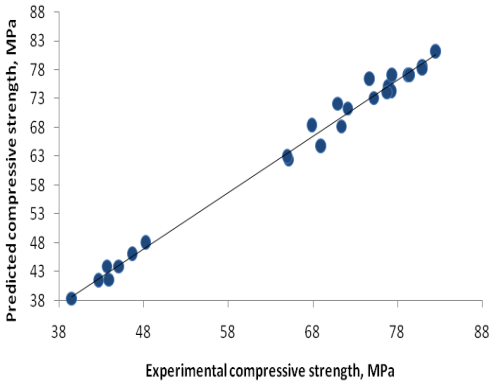
On successful completion of ANN training with 60 dataset, the model is verified with remaining 24 dataset. The output vector obtained from the ANN model is a normalized data and hence, the normalized data is reverted to its actual value by using eq. (11).



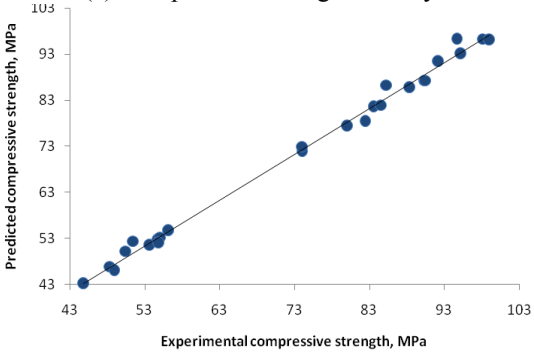
(a) Compressive strength – 1 day



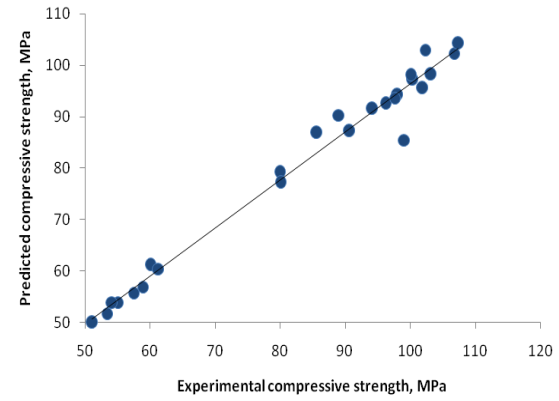
(b) Compressive strength – 3 days



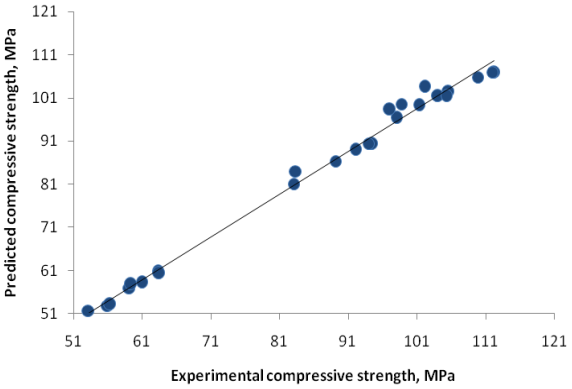
(c) Compressive strength – 7 days



(d) Compressive strength – 28 days



(e) Compressive strength – 56 days



(f) Compressive strength – 90 days

Fig. 2 Predicted through ANN vs experimental compressive strength

The training phase of the ANN converged at 900 - 1000 iterations or epochs.

$$X_i^a = X_i^n\left(X_i^{\max} - X_i^{\min}\right) + X_i^{\min} \tag{11}$$

where, x_i^n is the normalized result obtained after the test for the ith component. x_i^a is the actual result obtained for ith component, and x_i^{max} and x_i^{min} are the maximum and minimum values of all the components of the corresponding input vector before the normalization. The result is converged at 950 - 1050 iterations or epochs and the average error value is 1.8e-5. The coefficient determination (R^2) is 0.98765, 0.98965, 0.99342, 0.9812, 0.9801 and 0.9721 respectively for 1, 3, 7, 28, 56 and 90 days compressive strength. Fig. 2 presents ANN output vs corresponding experimental values. From Fig. 2, it can be observed that the ANN prediction is quite closer to corresponding experimental values.

5.2 MARSs based analysis

The MARS equation for the prediction of compressive strength for one day is given in eq. (9). It can be directly identified that the number of interaction effects and, in particular, interactions between efforts (as captured by basis function $B_4(x)$). Such interactions can be seen in eq. (12), when basis functions are part of the definition of other basis functions, e.g., $B_2(x)$ in $B_7(x)$, $B_8(x)$ etc. The presence of many such interactions suggests that the model is far from being additive and those interactions will play an important role in building an accurate model for code inspections. The user defined basis functions for over fitting the model was limited to 8 basis functions and the allowable highest degree of interaction was set to 2. The final model had 8 Basis functions as listed below.

Model I- Compressive strength (1 day)

The predicted model for compressive strength (one day) is given below

$$y = Comp.Str. = 0.356 + \sum_{m=1}^{8} a_m B_m^{(1)}(x) \tag{12}$$

Where

Basis function $B_m^{(1)}(x)$	Equation	Co-efficient (a_m)
$B_1(x)$	max(0, GGBS -0.3)	-0.501
$B_2(x)$	max(0, 0.7-FA) * max(0, FA -0.12)	-7.782
$B_3(x)$	max(0, 0.7-FA) * max(0, 0.12- SF)	3.746
$B_4(x)$	max(0, 0.7-FA) * max(0, MK-0.1)	0.182
$B_5(x)$	max(0, 0.7-FA) * max(0, 0.1 -MK)	-3.54
$B_6(x)$	$B_4(x)$ * max(0, FA -0.12)	10.12
$B_7(x)$	$B_4(x)$ * max(0, 12- FA)	3.65
$B_8(x)$	max (0, 0.7-FA) * max(0, FA-0.1)	1.213

Model II – Compressive strength (3 days)

The predicted model for compressive strength (3 days) is given below.

$$y = Comp.Str. = 0.478 + \sum_{m=1}^{8} a_m B_m^{(1)}(x) \tag{13}$$

Where,

Basis function $B_m^{(1)}(x)$	Equation	Co-efficient (a_m)
$B_1(x)$	max(0, GGBS -0.3)	-0.342
$B_2(x)$	max(0, 0.7-FA)	0.635
$B_3(x)$	$B_2(x)$ * max(0, FA -0.12)	-1.079
$B_4(x)$	$B_2(x)$* max(0, 0.12-SF)	-2.055
$B_5(x)$	$B_2(x)$* max(0, SF-0.12)	0.634
$B_6(x)$	$B_2(x)$* max(0, FA-0.12)	-0.421
$B_7(x)$	max(0, MK-0.1)	0.321
$B_8(x)$	max(0, 0.1-MK)	0.231

Model III – Compressive strength (7 days)

The predicted model for compressive strength (7 days) is given below.

$$y = Comp.Str. = 0.492 + \sum_{m=1}^{8} a_m B_m^{(1)}(x) \tag{14}$$

Where,

Basis function $B_m^{(1)}(x)$	Equation	Co-efficient (a_m)

$B_1(x)$	max(0, GGBS -0.3)	-0.301
$B_2(x)$	max(0, 0.7-FA)	0.161
$B_3(x)$	$B_2(x)$ * max(0, FA-0.12)	-1.031
$B_4(x)$	$B_2(x)$ * max(0, 0.12-SF)	0.621
$B_5(x)$	$B_4(x)$ * max(0, SF-0.12)	-2.654
$B_6(x)$	$B_1(x)$ * max(0, FA-0.12)	0.854
$B_7(x)$	max (0, 0.7-FA) * max(0, SF-0.12) * max(0, FA-0.1)	6.231
$B_8(x)$	max (0, 0.7-FA) * max(0, SF-0.12) * max(0, 0.1-MK)	10.342

Model IV – Compressive strength (28 days)

The predicted model for compressive strength (28 days) is given below.

$$y = \text{Comp. } Str. = 0.523 + \sum_{m=1}^{8} a_m B_m^{(1)}(x) \quad (15)$$

Where,

Basis function $B_m^{(1)}(x)$	Equation	Co-efficient (a_m)
$B_1(x)$	max(0, GGBS -0.3)	-0.589
$B_2(x)$	max(0, 0.7-FA)	0.94
$B_3(x)$	$B_2(x)$ * max(0, FA-0.12)	-3.785
$B_4(x)$	$B_2(x)$ * max(0, 0.12-SF)	2.863
$B_5(x)$	$B_4(x)$ * max(0, SF-0.12)	-5.902
$B_6(x)$	$B_4(x)$ * max(0, FA-0.12)	1.692
$B_7(x)$	max (0, 0.7-FA) * max(0, SF-0.12) * max(0, MK-0.1)	6.451
$B_8(x)$	max (0, 0.7-FA) * max(0, SF-0.12) * max(0, 0.1-MK)	13.402

Model V – Compressive strength (56 days)

The predicted model for compressive strength (56 days) is given below.

$$y = \text{Comp. } Str. = 0.592 + \sum_{m=1}^{8} a_m B_m^{(1)}(x) \quad (15)$$

Where,

Basis function $B_m^{(1)}(x)$	Equation	Co-efficient (a_m)
$B_1(x)$	max(0, GGBS -0.3)	-0.602
$B_2(x)$	max(0, 0.7-FA)	0.163
$B_3(x)$	$B_2(x)$ * max(0, FA-0.12)	-4.200
$B_4(x)$	$B_2(x)$ * max(0, 0.12-SF)	2.904
$B_5(x)$	$B_4(x)$ * max(0, SF-0.12)	-5.432
$B_6(x)$	$B_4(x)$ * max(0, FA-0.12)	2.345
$B_7(x)$	max (0, 0.7-FA) * max(0, SF-0.12) * max(0, FA-0.1)	9.154
B8(x)	max (0, 0.7-FA) * max(0, SF-0.12) * max(0, 0.1-MK)	12.632

Model VI – Compressive strength (90 days)

The predicted model for compressive strength (90 days) is given below.

$$y = \text{Comp. } Str. = 0.621 + \sum_{m=1}^{8} a_m B_m^{(1)}(x) \quad (17)$$

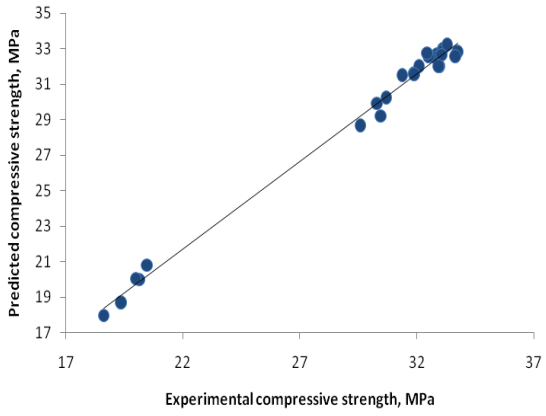
Where,

Basis function $B_m^{(1)}(x)$	Equation	Co-efficient (a_m)
$B_1(x)$	max(0, GGBS -0.3)	-0.176
$B_2(x)$	max(0, 0.7-FA)	0.689
$B_3(x)$	$B_2(x)$ * max(0, FA-0.12)	-1.098
$B_4(x)$	$B_2(x)$ * max(0, 0.12-SF)	2.076
$B_5(x)$	$B_4(x)$ * max(0, SF-0.12)	-3.905
$B_6(x)$	$B_4(x)$ * max(0, FA -0.12)	0.893
$B_7(x)$	max (0, 0.7-FA) * max(0, SF - 0.12) * max(0, FA-0.1)	10.876
B8(x)	$B_2(x)$ * max (0, MK-0.1)	14.703

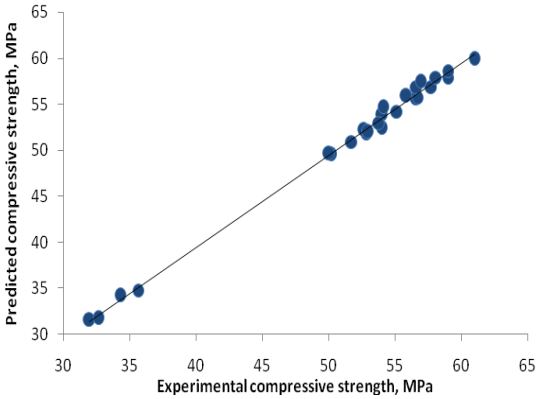
Table 4 shows the complete statistics for all the developed MARS models. The GCV was computed using the eq. (5) and the coefficient of correlation (R) was computed using eq. (18).

The value of coefficient of correlation (R) is determined by using the following formula

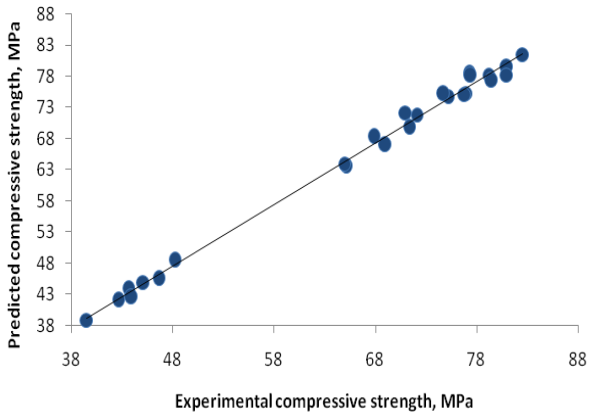
$$R = \frac{\sum_{i=1}^{n}\left(E_{ai} - \overline{E}_a\right)\left(E_{pi} - \overline{E}_p\right)}{\sqrt{\sum_{i=1}^{n}\left(E_{ai} - \overline{E}_a\right)}\sqrt{\sum_{i=1}^{n}\left(E_{pi} - \overline{E}_p\right)}} \quad (18)$$



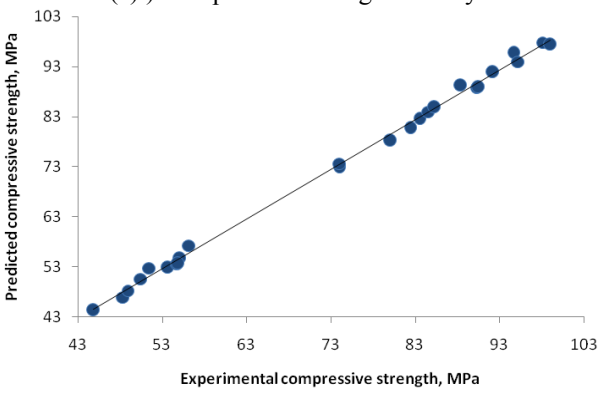
(a) Compressive strength – 1 day



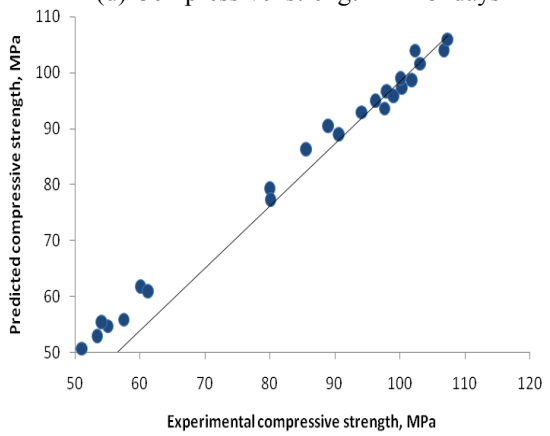
(b) Compressive strength – 3 days



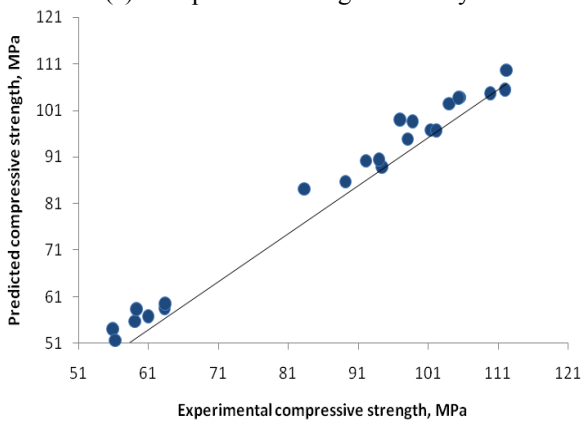
(c)) Compressive strength – 7 days



(d) Compressive strength – 28 days



(e) Compressive strength – 56 days



(f) Compressive strength – 90 days

Fig. 3 Predicted vs experimental compressive strength

Table 4 Testing of MARS

MARS results		Compressive strength					
		1 day	3 days	7 days	28 days	56 days	90 days
User defined max. no. of Basis functions		8	10	12	12	12	12
Interactions Ratio allowed		2	3	4	4	4	4
Final number of basis functions		8	8	8	8	8	8
Mean Square Error (MSE)	Train	4.32E-04	1.83E-04	3.79E-04	2.32E-04	1.76E-04	4.76E-04
	Test	0.001	0.0037	0.003	0.002	0.0032	0.004
Generalized Cross Validation		0.0010	0.0024	0.0032	0.0031	0.0014	0.0025
Coefficient of correlation (R)	Train	0.9991	0.9938	0.9939	0.9966	0.9927	0.9960
	Test	0.9977	0.9993	0.9959	0.9956	0.9910	0.9927

Table 5 R^2 values for training and testing datasets

Fracture characteristics	ANN		MARS	
	R^2 Train	R^2 Test	R^2 Train	R^2 Test
Comp. Str (1 day)	0.9876	0.9735	0.9982	0.9954
Comp. Str (3 days)	0.9896	0.9808	0.9876	0.9986
Comp. Str (7 days)	0.9934	0.9900	0.9879	0.9920
Comp. Str (28 days)	0.9812	0.9732	0.9932	0.9912
Comp. Str (56 days)	0.9801	0.9800	0.9854	0.9821
Comp. Str (90 days)	0.9721	0.9702	0.9921	0.9854

where, E_{ai} and E_{pi} are the actual and predicted values, respectively, Ea $\bar{E}_a$ and $\bar{E}_P$ are mean of actual and predicted E values corresponding to n patterns.

On successful development of MARS model with 60 dataset, the model is verified with remaining 24 dataset. The MARS output values versus Experimental values are plotted below (Fig. 3). From Fig. 3, it can be observed that the coefficient of determination R^2 is pretty much closer to the experimental values. Thus, developed MARS models are robust and reliable.

Table 5 presents the coefficient of determination (R^2) values for the various ANN models and MARS models for training dataset and testing dataset. Both the techniques scores R^2 values closer to 1 indicating robust to predict compressive strength of various GPC mixes. However, the R^2 values of MARS models scores bit higher values for testing datasets than that of the ANN models. In some cases the ANN prediction values show large variation from the experimental results compared to that of the MARS.

6. Summary and conclusions

Artificial neural network (ANN) and multivariate adaptive regression splines (MARS) based models have been developed to predict the compressive strength of bacteria incorporated GPC mixes. Compressive strength has

been predicted for 1 day, 3, 7, 28, 56 and 90 days of curing. GPC was made by using several ingredients like manufactured sand, GGBS, fly ash, silica fume, metakaolin, two new strains isolated from manufactured sand. For all the studies, the percentage of GGBS has been kept as fixed (70%) and molarity of NaOH is fixed at 8. From the results of ANN and MARS models, it is observed that the MARS models are preferred than the ANN models for the following reasons although both have good coefficient of determination:

- The computational resource required by MARS program is quite less than ANN.
- MARS models provide a relationship between the basis functions obtained from the dependent and the independent variables whereas ANN does not provide such a feature.
- In spite of ANN models prediction being fairly similar to that of the MARS models, MARS is found to be more stable for the dataset used in this study and ANNs are not stable (i.e. the predictions obtained using MARS model for a same parameter setting would remain same whereas in case of ANN this is unstable).

The developed ANN and MARS models can be used for prediction of compressive strength of bacteria incorporated GPC mixes. The predicted values will be useful for design of concrete structural components made up of bacteria incorporated GPC.

References

Dutta, S., Murthy, A.R., Kim, D. and Samui, P. (2017). "Prediction of compressive strength of self-compacting concrete using intelligent computational modelling", *Comput. Mater. Continua*, **53**(2), 157-174.

Ekman, T. and Kubin, G. (1999), "Nonlinear prediction of mobile radio channels: Measurements and Mars model designs", *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2667-2670, Arizona, U.S.A., March. https://doi.org/10.1109/ICASSP.1999.761246.

Engin, S., Ozturk, O. and Okay, F. (2015), "Estimation of ultimate torque capacity of the SFRC beams using ANN", *Struct. Eng. Mech.*, **53**(5), 939-956. https://doi.org/10.12989/sem.2015.53.5.939

Erdem, H. (2017), "Predicting the moment capacity of RC slabs with insulation materials exposed to fire by ANN", *Struct. Eng. Mech.*, **64**(3), 339-346. https://doi.org/10.12989/sem.2017.64.3.339.

Fernandez-Jimenez, A.M., Polomo, A. and Lopez-Hombrados, C. (2006), "Engineering properties of alkali-activated fly ash concrete", *ACI Mater. J.*, **103**(2), 106-112.

Friedman, J.H. (1991), "Multivariate adaptive regression splines", *Annals Stat.*, **19**(1), 1-67.

Guo, X., Shi, H. and Dick, W.A. (2010), "Compressive strength and microstructural characteristics of class C fly ash geopolymer", *Cement Concrete Compos.*, **32**(2), 142-147. https://doi.org/10.1016/j.cemconcomp.2009.11.003.

Hakan Arslan, M. (2009), "Application of ANN to evaluate effective parameters affecting failure load and displacement of RC buildings", *Natural Hazards Earth Syst. Sci.*, **9**, 967-977. https://doi.org/10.5194/nhess-9-967-2009.

Hardjito D., Wallah S.E., Sumajouw D.M.J and Rangan B.V. (2004), "On the development of fly ash-based geopolymer concrete", *ACI Mater. J.*, **101**, 467-472.

Hecht-Nielsen, R. (1990), *Neurocomputing*, Addison-Wesely Publishing Company, Boston, U.S.A.

Ince, R. (2004), "Prediction of fracture parameters of concrete by Artificial Neural Networks", *Eng. Fracture Mech.*, **71**, 2143-2159. https://doi.org/10.1016/j.engfracmech.2003.12.004.

Jin, R., Chen, W. and Simpson, T.W. (2000), "Comparative studies of metamodelling techniques under multiple modelling criteria", *AIAA J.*, 2000-4801. https://doi.org/10.1007/s00158-001-0160-4.

Kamarthi, S.V. and Pittner, S. (1999), "Accelerating neural network training using weight extrapolation", *Neural Networks Soc.*, **12**(9), 1285-1299. https://doi.org/10.1016/S0893-6080(99)00072-6.

Kaur, J. and Kaur, K. (2017), "A fuzzy approach for an IoT-based automated employee performance appraisal", *CMC: Comput., Mater. Continua*, **53**(1), 23-36.

Loizos, A. and Karlaftis, M.G. (2006), "Neural networks and non-parametric statistical models: A comparative analysis in pavement condition assessment", *Adv. Appl. Stat.*, **6**(1), 87-110.

Mallick, B.K., Denison, D.G. and Smith, A.F. (1999), "Bayesian survival analysis using a MARS model", *Biometrics,* **55**(4), 1071-1077. https://doi.org/10.1111/j.0006-341X.1999.01071.x.

Mansouri, I., Safa, M., Ibrahim, Z., Kisi, O., Tahir, M.M., Baharom, S. and Azimi, M. (2016), "Strength prediction of rotary brace damper using MLR and MARS", *Struct. Eng. Mech.*, **60**(3), 471-488. https://doi.org/10.12989/sem.2016.60.3.471.

Mohammed, E.H. and Sudhakar, K.V. (2002), "ANN back-propagation prediction model for fracture toughness in microalloy steel", *J. Fatigue*, **24**(9), 1003-1010. https://doi.org/10.1016/S0142-1123(01)00207-9.

Nii, O.A., Cooger, K. and Mensah, S. (2009), "Multivariate adaptive regression (MARS) and hinged hyperplanes (HHP) for doweled pavement performance modeling", *Construct. Build. Mater.*, **23**(9), 3020-3023. https://doi.org/10.1016/j.conbuildmat.2009.04.010.

Pacheco-Torgal F., Moura, D., Ding, Y. and Jalali, S. (2011), "Composition, strength and workability of alkali-activated metakaolin based mortars", *Construct. Build. Mater.*, **25**(9), 3732-3745. https://doi.org/10.1016/j.conbuildmat.2011.04.017.

Parab, S., Srivastava, S., Samui, P. and Murthy, A.R. (2014), "Prediction of fracture parameters of high strength and ultra high strength concrete beams using Gaussian process regression and Least squares", *Comput. Model. Eng. Sci.*, **101**(2), 139-158.

Prasanna, P.K., Murthy, R.A. and Srinivasu, K. (2018), "Prediction of compressive strength of GGBS based concrete using RVM", *Struct. Eng. Mech.*, **68**(6), 691-700. https://doi.org/10.12989/sem.2018.68.6.691.

Shah, V.S., Shah, H.R., Samui, P. and Murthy, A.R. (2014), "Prediction of fracture parameters of high strength and ultra-high strength concrete beams using minimax probability machine regression and extreme learning machine", *Comput. Mater. Continua*, **44**(2), 73-84.

Shi, X., Collins, F., Zhao, X. and Wang, Q. (2012), "Mechanical properties and microstructure analysis of fly ash geopolymeric recycled concrete", *J. Hazard. Mater.*, **20**(9), 237-238. https://doi.org/10.1016/j.jhazmat.2012.07.070.

Siddique, R., Aggarwal, P. and Aggarwal, Y. (2011), "Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks", *Adv. Eng. Software*, **42**(10), 780-786. https://doi.org/10.1016/j.advengsoft.2011.05.016.

Sumajouw, D.M.J., Hardjito, D., Wallah, S.E. and Rangan, B.V. (2007), "Fly-ash based geopolymer concrete: study of slender reinforced columns", *J. Mater. Sci.*, **42**, 3124-3130. https://doi.org/10.1007/s10853-006-0523-8.

Widrow, B. and Lehr, M.A. (1990), "30 years of Adoptive Neural Netwoks; Perceptron, Madaline, and Back propagation", *Proceedings of the IEEE,* **78**, 1415-1442.

York, T.P and Eaves, L.J (2001), "Common disease analysis using multivariate adaptive splines (MARS): Genetic analysis workshop 12 simulated sequence data", *Genet Epidemiol*, **21**, 649-654. https://doi.org/10.1002/gepi.2001.21.s1.s649.

Yuvaraj, P., Murthy, A.R., Iyer, N.R., Samui, P. and Sekar, S.K. (2013b), "Multivariate adaptive regression splines model to predict fracture characteristics of high strength and ultra high strength concrete beams", *Comput*. *Mater*. *Continua*, **36**(1), 73-97.

Yuvaraj, P., Murthy, A.R., Iyer, N.R., Samui, P. and Sekar, S.K. (2014b), "Prediction of fracture characteristics of high strength and ultra high strength concrete beams based on relevance vector machine", *J. Damage Mech*., **23**(7), 979-1004. https://doi.org/10.1177/1056789514520796.

Yuvaraj, P., Murthy, A.R., Iyer, N.R., Sekar, S.K. and Samui, P. (2014a), "ANN model to predict fracture characteristics of high strength and ultra high strength concrete beams", *Comput*. *Mater*. *Continua*, **41**(3), 193-213. http://doi.org/10.3970/cmc.2014.041.193.

Yuvaraj, P., Murthy, A.R., Iyer, N.R., Sekar, S.K. and Samui, P. (2013a), "Support vector regression based models to predict fracture characteristics of high strength and ultra high strength concrete beams", *Eng*. *Fracture Mech*., **98**(1), 29-43. https://doi.org/10.1016/j.engfracmech.2012.11.014.

CC