Challenge Journal of Structural Mechanics

Research Article

A comparative study of concrete strength prediction using artificial neural network, multigene programming and model tree

Preeti Kulkarni *, Shreenivas N. Londhe, Pradnya R. Dixit

Department of Civil Engineering, Vishwakarma Institute of Information Technology, Pune 411048, India

ABSTRACT

In the current study 28 day strength of Recycled Aggregate Concrete (RAC) and Fly ash (class F) based concrete is predicted using Artificial Neural Network (ANN), Multigene Genetic Programming (MGGP) and Model Tree (MT). Four sets of models were designed for per cubic proportions of materials, Properties of materials and non-dimensional parameters as input parameters. The study shows that the predicted 28 day strength is in good agreement with the observed data and also generalize well to untrained data. ANN outperforms MGGP and MT in terms of model performance. Output of the developed models can be presented in terms of trained weights and biases in ANN, equations in MGGP and in the form of series of equations in MT. ANN, MGGP and MT can grasp the influence of input parameters which can be seen through Hinton diagrams in ANN, input frequency distribution in MGGP and coefficients of input parameters in MT. The study shows that these data driven techniques can be used for developing model/s to predict strength of concrete with an acceptable performance.

ARTICLE INFO

Article history:
Received 23 February 2019
Revised 26 April 2019
Accepted 15 May 2019

Keywords:
Recycled aggregate concrete
Fly ash concrete
Artificial neural network
Multigene genetic programming
Model tree

1. Introduction

Recycled Aggregates and fly ash are the alternative materials used in concrete which can be termed as a step towards use of waste materials in concrete. Ascertaining strength of such concrete is a tedious and difficult task owing to the different properties of Recycled aggregates and fly ash (Hansen and Narud, 1983; Yueh and Hwang, 2006; Ryu, 2002). Determination of compressive strength of concrete has great importance as it offers an option to do the essential modification on the mix proportion to avoid circumstances where concrete does not attain the design strength and also for more economic use of raw material and fewer construction failures, thus reducing construction cost. Traditional determination of compressive strength of concrete needs actual testing which requires time and materials, which can be reduced by using data driven techniques like Artificial Neural Network (ANN), Genetic Programming (GP), and Model Tree (MT) etc. Prediction of compressive strength of concrete has been an active area of research in last two decades or so (Dias and Pooliyadda, 2001; I-Cheng, 2007; Ni and Wang, 2000; Ahmet et al., 2006; Adriana et al., 2013; Duan et al., 2013; Deshpande et al., 2014; Gorphade et al., 2014; Sarıdemir, 2010; Bayazidi et al., 2014). Relatively new techniques of GP and MT have been used sparingly for modeling the compressive strength of concrete. ANN models were developed to predict the strength and slump of ready mix concrete with admixtures in which the input parameters were non-dimensional ratios transformed from the material weights per unit volume. Neural network was also developed with the natural logarithms of both inputs and outputs (Dias and Pooliyadda, 2001). Weights of mixes per unit volume were considered as input parameters to predict slump of High Performance Concrete (HPC) using ANN (I-Cheng, 2007). A Three layered neural network model was built to implement the complex nonlinear relationship between the inputs (11 factors that influence concrete strength) and the output (concrete strength). The neural network models give high prediction accuracy, and the research results conform to some rules of mix proportion of concrete (Ni and Wang, 2000). ANN models were developed to predict the Compressive

* *Corresponding author.* Tel.: +91-955-252-5907 ; E-mail address: preeti.kulkarni@viit.ac.in (P. Kulkarni)
ISSN: 2149-8024 / DOI: https://doi.org/10.20528/cjsmec.2019.02.002

Strength and slump of High Strength Concrete (HSC) with input parameters such as water to binder ratio, water content, fine aggregate ratio, fly ash content etc. (Ahmet et al., 2006). ANN is used as an attempt to obtain more accurate concrete strength prediction based on parameters like concrete mix design, size and shape of specimen, curing technique and period, environmental conditions, etc. (Gupta et al., 2006). ANN with mix proportions as input parameters was used to predict strength of concrete from ready-mixed concrete companies (Jong-In Kim et al., 2004). Particularly in the field of RAC, ANN was used to predict strength of RAC (Adriana et al., 2013; Duan et al., 2013; Deshpande et al., 2014). Application of Genetic Algorithm based neural network models for predicting the Compaction factor, VB time and Compressive strength, Tensile strength, Flexural strength and Young's modulus of High performance concrete showed a prediction accuracy of 95% (Gorphade et al., 2014). In a study, two models using gene expression programming (GEP) approach were developed for predicting compressive strength of concrete containing rice husk ash at the age of 1, 3, 7, 14, 28, 56 and 90 days (Sarıdemir, 2010). MGGP as a technique was utilized to predict modulus of elasticity of concrete. A general model proposed for Normal strength concrete and High strength concrete using the 28 day strength data (Bayazidi et al., 2014). Model Tree (MT) was used to predict strength of conventional and Recycled aggregate concrete (Deshpande et al., 2014; Deepa et al., 2010).

The study concluded that ANN facilitates a better correlation among inputs and output and displays a good performance. Very few applications of GP (specifically MGGP) have been reported in recent literature focused on predicting strength of concrete. Similarly very few works can be seen which used MT to predict the concrete strength (CS). The study mentioned earlier focused mainly on performance of tool used rather than discussing the influence of input parameters on output which is necessary, for the tag of 'Black box' on these techniques to be removed. In the present work, three techniques viz. Artificial Neural Network (ANN), MultiGene Genetic Programming (MGGP) and Model Tree (MT) are used separately to develop models to predict strength of Recycled aggregate concrete and Fly ash based concrete respectively. Secondly, in total 8 models for each technique were developed with Mix proportions of materials, properties of materials and non-dimensional parameters as input parameters for developing different models. The data sets were designed in the said way so that the study is not limited to only type of input parameter/s for a similar output. Third, the influence of parameters affecting the strength of concrete are shown in the form of Hinton diagram in ANN, in the form of coefficients and input frequency in MGGP and coefficients of parameters in MT. Fourth, the comparative analysis of the modeling approaches (ANN, MGGP and MT) are validated with the observed values and best approach is suggested for predicting the 28 day compressive strength of concrete.

The paper is further organized as follows: The next section gives an overview of ANN, GP-MGGP and MT techniques. The information about data used is provided in following section followed by the methodology adopted. The results of models developed and influences of parameters are discussed in the next section followed by concluding remarks.

2. Modeling Techniques

In the current study, prediction of 28 day concrete strength (CS) for RAC and Fly ash based concrete is done using Artificial Neural Networks, Genetic Programming-Multi Gene Genetic Programming, and Model Tree with M5 algorithm. These approaches are described in brief below.

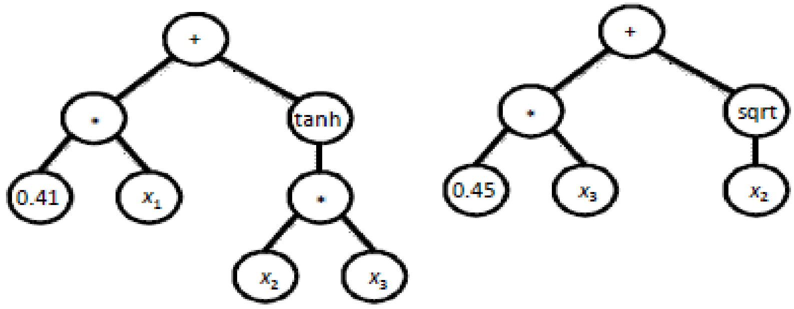
2.1. Artificial Neural Network (ANN)

ANN is a soft computing technique involving an input layer, one or more hidden layer (s) and an output layer. The hidden layer is connected to the other layers by weights, biases and transfer functions. An error function is determined by the difference between network output and the target. The error is propagated back and the weight and biases are adjusted using some optimization technique which minimizes the error. The entire process called training is repeated for number of epochs (iterations) till the desired accuracy in output is achieved. Once the network is trained it can be used to validate against unseen data using trained weights and biases (The ASCE Task Committee, 2000; Maier and Dandy, 2000).

2.2. Genetic Programming (GP)

Genetic programming (GP) is a biologically inspired machine learning method that evolves computer programs to perform a task (usually represented by tree structures) and then breeding together the best performing trees to create a new population. The three genetic operations are as follows: Reproduction, Cross over and Mutation (Londhe and Dixit, 2012). In MGGP, multigene individual consists of one or more genes, each of which is a "traditional" GP tree (Searson et al., 2010). Genes are acquired incrementally by individuals in order to improve fitness (e.g. to reduce a model's sum of squared errors on a data set). The overall model is a weighted linear combination of each gene. The resulting pseudo-linear model can capture non-linear behavior. When the transformations are forced to be low order (by restricting the GP tree depth), allows the evolution of accurate, relatively compact mathematical models of predictor – response (input – output) data sets, even when there are a large number of input variables. For example, the multigene model shown in Fig. 1 predicts an output variable using input variables x1, x2 and x3.

This model structure contains non-linear terms (e.g. the hyperbolic tangent) but is linear in the parameters with respect to the coefficients d0, d1 & d2. In practice, the user specifies the maximum number of genes Gmax a model is allowed to have and the maximum tree depth Dmax any gene may have and therefore can exert control over the maximum complexity of the evolved models. In particular, we have found that enforcing stringent tree depth restrictions (i.e. maximum depths of 4 or 5 nodes) often allows the evolution of relatively compact models that are linear combinations of low order non-linear transformations of the input variables. Multigene GP combines the power of classical linear regression with the ability to capture non-linear behavior without needing to pre-specify the structure of the non-linear model (Searson et al., 2010; Searson et al., 2007). The uniqueness of the multi-gene genetic programming based model is that it automatically evolves a mathematical expression in a symbolic form which can be analyzed further to find which variables impact the final prediction and in what fashion (Pandey et al., 2015).



$$y = d_0 + d_1(0.41x_1 + \tanh(x_2x_3) + d_2(0.45x_3 + \mathrm{sqrt}(x_2))$$

Fig. 1. Example of a Multigene symbolic model.

2.3. Model Tree (MT)

MT utilizes divide-and-conquer approach and provides rules for reaching the models at the leaf nodes. The linear models are then used to quantify the contribution of each attribute to the overall predicted value. M5P, a reconstruction of Quinlan's M5 algorithm is used for inducing trees of regression models and combines a conventional decision tree with the possibility of linear regression functions at the nodes. First, a decision-tree induction algorithm is used to build a tree and a splitting criterion is then used that minimizes the intra-subset variation in the class values down each branch. The splitting procedure in M5 stops if the class values of all instances that reach a node vary very slightly, or only a few instances remain. Second, the tree is pruned back from each leaf. When pruning an inner node is turned into a leaf with a regression plane. In comparison with classical regression trees, Model Trees deliver better compactness and prediction accuracy (Deepa et al., 2010; Quinlan, 1992).

3. Data and Model Development

For predicting concrete strength using ANN, MGGP and MT, experimentation work was carried out by the authors and few data was also collected from literature (Hansen and Narud, 1983; Yueh and Hwang, 2006; Ryu, 2002; Khatib, 2005; Padmini et al., 2003; Dapena et al., 2011; Corinaldesi, 2010; Fathifazl et al., 2009; Yong and Teo, 2009; Yaprak et al., 2011; ChakradharaRao et al., 2010; Schoppe, 2011; Kumutha and Vijai, 2010; Evangelista and Brito, 2010; Zega and Maio, 2003; Kou, 2006; Konin and Kouadio, 2011; Poon et al., 2004; Kotrayothar, 2012; Adnan et al., 2011; Poon et al., 2007; Domingo-Cabo et al., 2009; Pereira et al., 2012; Pelufo et al., 2009; Agarwal et al., 2011; Evangelista and Brito, 2004; Goncalves et al., 2004; Nikoo et al., 2015). The data used in the current work is divided into four sets i.e Set 1, Set 2, Set 3 and Set 4 and 2 models each. Set 1 is designed with process parameters related to Recycled aggregate concrete (RAC), Set 2 with Fly ash based concrete, Set3 with non-dimensional parameters for RAC and Set 4 with non-dimensional parameters of Fly ash based concrete. The data sets were designee in the said way so that the study is not limited to only one type of input parameters and number of data sets for a similar output.

The process parameters that have been used as model input parameters for models in Set1 are: Content of materials in kg/m^3 for Cement (RC, kg/m^3), Natural fine aggregate (RNFA, kg/m^3), Natural coarse aggregate-20mm (RNCA-20, kg/m^3), Natural coarse aggregate-10mm (RNCA-10, kg/m^3), Recycled coarse aggregate-20mm (RCA-20, kg/m^3), Recycled coarse aggregate-10mm (RCA-10, kg/m^3), Admixture(RA, kg/m^3) and water (RW, kg/m^3). Water absorption of conventional coarse aggregates (WA-NA, %) and water absorption of Recycled aggregates (WA-RA, %) were used as additional input parameters in Set 1: model 2. The input parameters for models in Set 2 were: Cement (FC, kg/m^3), Fly ash – Class F (F, kg/m^3), Fine aggregate (FNFA, kg/m^3), Natural coarse aggregate-20mm-1 (FNCA-20, kg/m^3), Natural coarse aggregate-10mm (FNCA-10, kg/m^3), water (FW, kg/m^3), and Admixture (FA, kg/m^3). Specific gravity of FNFA (FSP-NFA), Specific gravity of NCA-20 (FSP-20) and Specific gravity of NCA-10 (FSP-10) were additional input parameters in Set 2: model1. The input parameters for Set 3 were dimensionless parameters such as ratio of Water to cement ratio (RW/C), natural fine aggregate to total aggregate ratio (RNFA/A), Natural coarse aggregate-20mm to cement content (RNC20/A), Natural coarse aggregate-10mm to cement content (RNC10/A), Recycled coarse aggregate-20mm (RCA-20) to cement (RC20/C), Recycled coarse aggregate-10mm (RCA-10) to cement (RC10/C), water to total materials (RW/T). Replacement ratio (R-RR) was used as an additional parameter in Set 3:model2. The input parameters for Set 4 were non-dimensional parameters for Fly Ash based concrete namely Water to binder ratio (FW/B), machine made sand by aggregate ratio (FNFA/A), Natural coarse aggregate-20mm to cement ratio (FNCA20/C), Natural coarse aggregate-10mm to cement ratio (FNCA10/C) and water to total materials ratio (FW/T). Replacement ratio (F-RR) was used as an additional input parameter in Set 4:model2.

The output for each model is 28 day compressive strength of respective type of concrete (CS). The details of data used in developing the models are shown in Tables 1-4. The detail of models developed in each set is shown in Table 5.

Three layered Feed Forward Back-Propagation ANN models were developed using MATLAB 2016, to predict the 28 day CS and trained till a very low performance error (mean squared error) was achieved. All the networks were trained using Levernberg-Marquardt algorithm with 'log-sigmoid 'transfer functions in between the first (input) and second (hidden) layer and 'linear' transfer function between the second and third layer (output). Trial and error method was utilized to determine the optimal number of hidden neurons. MGGP models were developed using GPTIPS-2. Readers are referred for features of GPTIPS (Searson et al., 2007; Searson et al., 2010). The RMSE function was adapted for error minimization during runs (Searson et al., 2007; Searson et al., 2010). The adopted function set to develop the GP model are as shown in table 6 for each model. The parameters were selected which yielded best performance of the models. These settings were based on experience with the predictive modeling of other data sets of similar size, and so they may not be optimal. A fairly large number of population and generations were tested to find models with minimum error. The programs run until the number of generations were reached as in Table 6.

Table 1. Details of data in Set 1.

Sr. No	Parameters	Values (min-max)	Correlation with Output
1	RC (kg/m^3)	235-645	0.477
2	RNFA (kg/m3)	217-1050	0.004
3	RNCA-20mm (kg/m^3)	0-1508.640	0.118
4	RNCA-10mm (kg/m^3)	0-553	0.281
5	RCA-20mm (kg/m^3)	0-1508.640	-0.3011
6	RCA-10mm (kg/m^3)	0-840	-0.0989
7	RW (kg/m^3)	120-271	0.0451
8	RA (kg/m^3)	0-41.600	-0.3186
9	WA-RA (%)	0-10.600	-0.109
10	WA-NA (%)	0-3.560	-0.207
11	S (N/mm^2)	10.319-100.500	

Table 2. Details of data in Set 2.

Sr. No	Parameters	Values (min-max)	Correlation with Output
1	FC (kg/m^3)	130-460	0.861
2	F (kg/m^3)	0-120	-0.247
3	FNFA-1 (kg/m^3)	398-1011	-0.421
4	FNCA-20mm (kg/m^3)	0-958	-0.163
5	FNCA-10mm (kg/m^3)	482-1242	0.039
6	FW (kg/m^3)	127-202	0.099
7	FA (kg/m^3)	0-5.520	0.519
8	FSP-NFA	2.700-2.980	0.199
9	FSP-20	0-3.050	-0.036
10	FSP-10	2.850-3.040	0.225
11	S (N/mm^2)	12-60.2	

Table 3. Details of data in Set 3.

Sr. No	Parameters	Values (min-max)	Correlation with Output
1	FW/B	0.315-0.980	-0.843
2	FNFA/A	0.200-0.489	-0.197
3	FNC20/C	1.175-5.393	-0.622
4	FNC10/C	0-4.354	-0.499
5	FW/T	0.051-0.0867	0.107
6	F-RR	0-48	-0.48
7	S (N/mm^2)	12-60.200	

Table 4. Details of data in Set 4.

Sr. No	Parameters	Values (min-max)	Correlation with Output
1	RWC	0.229-0.860	-0.584
2	RNFA/A	0.148-1.566	0.085
3	RNC20/C	0-4.726	0.029
4	RNC10/C	0-2.196	0.204
5	RC20/C	0-5.184	-0.359
6	RC10/C	0-2.333	-0.125
7	RW/T	0.054-0.139	-0.25
8	R-RR	0-100	-0.251
9	S (N/mm^2)	10.319-100.5	

Table 5. Model development.

Sr. No	Set No	Model No.	Input Parameters	No. of Data Sets
1	Set 1 (Recycled Aggregate Concrete)	1-1	RC, RNFA, RNC20, RNC10, RC20, RC-10,RA, RW	226
		1-2	RC, RNFA, RNC20, RNC10, RC20, RC-10, RA, RW, WA-RA, WA-NA	226
2	Set 2 (Fly Ash based Concrete)	2-1	FC, F, FNFA, FNC20, FNC10, FA, FW	113
		2-2	FC, F, FNFA, FNC20, FNC10, FA, FW, FSP-NFA, FSP-NC20, FSPNC10, FW,FA	113
3	Set 3 (Non-Dimensional Parameters for Recycled Aggregates)	3-1	RWC, RNFA/A, RNC20/C, RNC10/C, RC20/C, RC10/C, RW/T	226
		3-2	RWC, RNFA/A, RNC20/C, RNC10/C, RC20/C, RC10/C, RW/T, R-RR	226
4	Set 4 (Non Dimensional parameters related to Fly Ash based Concrete)	4-1	FW/B, FNFA/A, FNC20/C, FNC10/C, FW/T	113
		4-2	FW/B, FNFA/A, FNC20/C, FNC10/C, FW/T, F-RR	113

Table 6. Parameter settings for the MGGP.

GP Parameters	Parameter Settings
Population size	1000
Number of generation	200,500
Selection method	tournament
Tournament size	15
Crossover rate	0.84
Mutation rate	0.14
Termination criteria	500 generation or fitness value
	less than 0.00 whichever is earlier
Maximum number of genes	6,8
Maximum tree depth	4,5,6
Mathematical operations	+, -, x, /, sin, cos, exp, √, exp,{}

The maximum allowable number of genes in an individual and the maximum tree depth directly influence the size of the search space and the number of solutions explored within the search space (Searson et al., 2007; Searson et al., 2010; Pandey et al., 2015). The allowable number of genes and tree depth were, respectively, set to optimal values as tradeoffs between the running time and the complexity of the evolved solutions (Searson et al., 2007; Searson et al., 2010; Pandey et al., 2015). The best MGGP models were chosen on the basis of providing the best fitness value on the training and testing data as well as the simplicity of the models (Bayazidi et al., 2014). All these parameter combinations were tested and 2 replications for each were carried out. Multiple individual runs are suggested where the populations are automatically merged after the completion of the runs. This approach mitigates problems with the possible loss of model diversity over a run and with the GP algorithm getting stuck in local minima (Searson, 2015). The overall number of optimal individual runs equals to 12x8x2=192 (6 group of models for each generations 200 and 500. Each generation with genes 6 and 8 and further each generation and gene with tree depth 4, 5 and 6. Thus 12 group of models each for generations 200 and 500 and 2 replications for each models. The methodology is adopted for set 1 with 4 models and set 2 with 4 models).

For Model Tree as a technique, M5P algorithm implemented in software WEKA was used for calibrating the model (Frank et al., 2016; Deepa, 2010). To check the accuracies and robustness of the model, the dataset was divided for training and testing purposes. From the available data, 70% was selected to be used for training purposes and the remaining 30% was used for model validation. The performance of the model was assessed by statistical measures like correlation coefficient (r) (Eq. (1)), Root mean squared error ($RMSE$) (Eq. (2)), Average absolute error ($AARE$) (Eq. (3)), Mean absolute error (MAE) (Eq. (4)), and Nash-Sutcliffe Efficiency (E) (Eq. (5)) (David and Gregory, 1999; Jain et al., 2008; Londhe, 2008). Lower $RMSE$ indicates good prediction, but this statistic is biased towards to high error values. Coefficient of correlation (r) measures the degree of association between the observed and predicted values and r closer to 1 indicates an almost perfect linear relationship between them. The value of zero for the coefficient of efficiency (E) indicates that the observed mean is as good a predictor as the model, while negative values indicate that the observed mean is a better predictor than the model. E is sensitive to outliers (David and Gregory, 1999). The degree to which $RMSE$ exceeds MAE is an indicator of the extent to which outliers (or variance in the differences between the modeled and observed values) exist in the data (David and Gregory, 1999; Jain et al., 2008; Londhe, 2008).

$$r = \frac{\sum(S_{obs} - \bar{S}_{obs})(S_{cal} - \bar{S}_{cal})}{\sqrt{\sum(S_{obs} - \bar{S}_{obs})^2 (S_{cal} - \bar{S}_{cal})^2}} \tag{1}$$

$$RMSE = \sqrt{\sum_{i=1}^{n}(S_{obs} - \bar{S}_{cal})^2/n} \tag{2}$$

$$AARE = \frac{1}{N}\sum\left|\frac{(S_{cal} - S_{obs})}{S_{obs}}\right| \times 100 \tag{3}$$

$$MAE = \frac{\sum_{i=1}^{N}|S_{cal} - \bar{S}_{cal}|}{N} \tag{4}$$

$$E = 1 - \frac{\sum(S_{cal} - S_{obs})^2}{\sum(S_{obs} - \bar{S}_{obs})^2} \tag{5}$$

where S_{obs}=observed strength; S_{cal}=strength calculated from a model; $\bar{S}_{obs}$=average observed strength; $\bar{S}_{cal}$=average calculated strength; n=total number of data points predicted and all the summations run from 1 to N.

The architecture of ANN models and no. of equations developed for each model by MT and the parameters not considered in the equation developed by MGGP is shown in Table 7.

4. Results and Discussion

4.1. Models developed using ANN

Performance of each of the models developed in testing using the ANN technique is shown in Table 8.

Table 7. Details of models developed.

Set. No	Model No.	ANN Architecture	No. of equation in MT	MGGP parameters not considered in the equation
1	1-1	8:25:01	2	NIL
	1-2	10:23:01	11	NIL
2	2-1	7:13:01	3	NIL
	2-2	10:08:01	5	SP-NC20
3	3-1	7:24:01	2	NIL
	3-2	8:25:01	2	RNC20/C
4	4-1	5:16:01	1	NIL
	4-2	6:10:01	1	NIL

Table 8. Performance of models developed using ANN.

Set. No	Model No.	*RMSE* (m/s)	*MAE* (m/s)	*E*	*AARE*	*r*
1	ANN 1-1	5.759	4.3375	0.87934	11.4091	0.9440
	ANN 1-2	7.8139	5.9043	0.78034	14.6269	0.8890
2	ANN 2-1	3.6983	2.5631	0.85917	8.3445	0.9388
	ANN 2-2	4.0241	2.7790	0.83327	8.7192	0.9293
3	ANN 3-1	6.5255	4.6132	0.84508	11.9555	0.9219
	ANN 3-2	6.2345	4.4092	0.85859	10.8077	0.9268
4	ANN 4-1	3.2909	2.4703	0.87941	7.7395	0.9438
	ANN 4-2	3.5056	2.7136	0.88577	10.1742	0.9452

Fig. 2 shows the scatter plot for ANN 1-1. Comparison between measured & predicted values for strength characteristics of RAC for whole test data in Fig. 2 demonstrates that there are few scatters away from the line of equality between measured and predicted values. As shown, the proposed model for compressive strength of RAC has a reasonable accuracy with less scatter and a high value of correlation coefficient (r=0.94). Fig. 3 shows the Hinton diagram which depicts the influence of various parameters on the compressive strength of concrete in ANN.

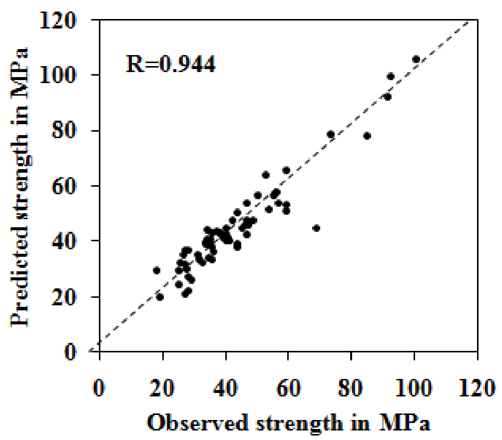


Fig. 2. Scatter plot for ANN1-1.

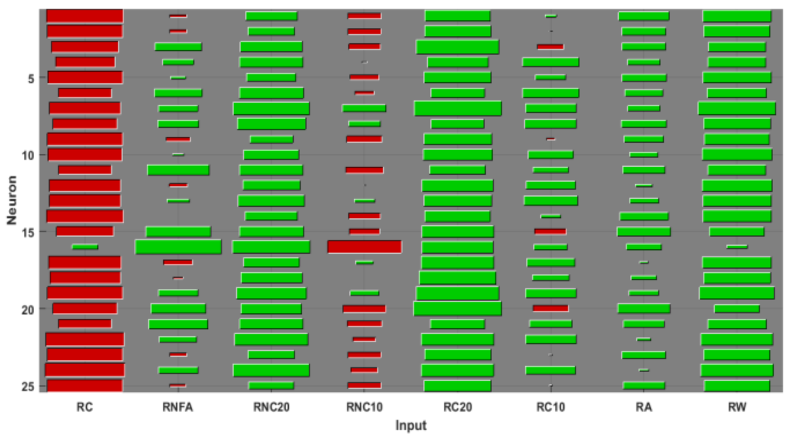


Fig. 3. Hinton diagram for ANN1-1.

A Hinton diagram is plot of weight matrix of a neural network, where the size of the square represents the magnitude, and the color represents the polarity (red=positive, green=negative). A Hinton diagram thus at a glance shows the units which are strongly active, which input parameters are off and which input parameters are intermediate in influence towards predicting the 28 day CS of RAC (Ahmet et al., 2006). A Hinton diagram for ANN1-1, shows maximum influence of cement content, water content followed by recycled aggregate content in ANN1-1 on the CS of RAC. Hinton for ANN2-2 in Fig. 4 shows cement content, aggregate content and fly ash content as influential parameters in decreasing order followed by other input parameters. The Hinton diagram can thus eliminate the need for the sensitivity analysis. The scatter plot for ANN3-1 is shown in Fig. 5, shows over prediction of strength. Hinton diagram for ANN4-2 is as shown in Fig. 6 which shows aggregate to cement ratio and water to binder ratio as the highly influential parameters and with F-RR as the least influential parameter.

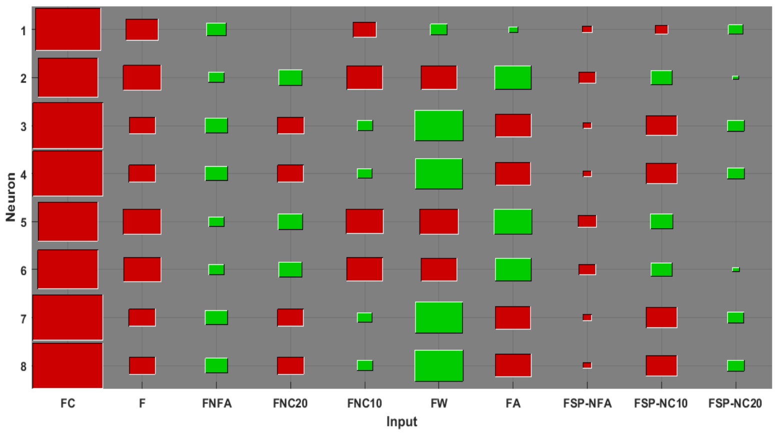


Fig. 4. Hinton diagram for ANN2-2.

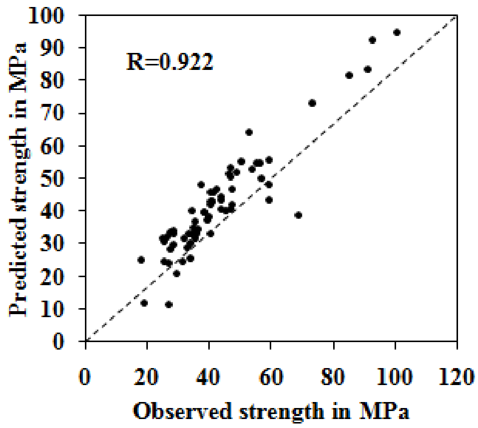


Fig. 5. Scatter plot for ANN2-2.

28 day strength of concrete is affected due to the water binder ratio (W/B) & increase in W/B can decrease the strength. Also replacement ratio of RA or Fly ash in concrete can decrease the strength of concrete (Shetty, 2005; Neville, 2012). Specifically in RAC, in a given mix when Aggregate to cement ratio increases, the strength decreases. Also increase in the CS can be seen with increase in water to total materials ratio and fine aggregate to total aggregate content up to a certain limit and further it shows a decrease in strength (Deshpande, 2016).

4.2. Model formulation using MGGP

With input parameters as mix proportions of concrete, properties of materials and non-dimensional parameters, models were calibrated using MGGP as shown in Table 1 and Table 2. Performance of each of the model in testing is shown in Table 9.

To find the optimal model, the MGGP algorithm was run several times with different combinations of the parameters as shown in Table 6. Results of models with Population:1000, generation:500, tree depth:4 and no. of genes:6 were found to be satisfactory and recorded here. Fig. 7 and Table 10 show the individual genes/model terms for the best models of MGGP1-1 and MGGP2-1 that were obtained during the conducted runs.

Each gene includes its weighing coefficient. It is seen that the weight of the genes (sub-programs) 1, 3, 5 and the bias terms are higher than the other genes for MGGP2-1 and high importance of gene 6, 4, 3 and bias term in MGGP1-1. This means that they have higher contribution to the strength prediction of concrete. Fig. 8 shows the expressional trees for the best models that were obtained during the conducted runs for MGGP3-1. Each gene includes its weighing coefficient. As can be observed from Fig. 8, the derived model is composed of complicated array of operators, variables, and constants to estimate the 28 CS of concrete.

Table 9. Performance of models developed using MGGP.

Set. No	Model No.	*RMSE* (m/s)	*MAE* (m/s)	*E*	*AARE*	*r*
1	MGGP 1-1	6.9110	5.5560	0.8262	14.3747	0.9090
	MGGP 1-2	8.5412	6.3821	0.7375	16.4551	0.8588
2	MGGP 2-1	4.0559	2.8832	0.8222	9.0642	0.9283
	MGGP 2-2	3.6216	2.8611	0.8649	9.6019	0.9414
3	MGGP 3-1	6.9458	5.3300	0.8244	13.3762	0.9095
	MGGP 3-2	8.4083	6.4100	0.7427	16.4922	0.8647
4	MGGP 4-1	4.2349	2.8108	0.8003	9.7842	0.9177
	MGGP 4-2	3.5550	2.7290	0.8824	10.3964	0.9518

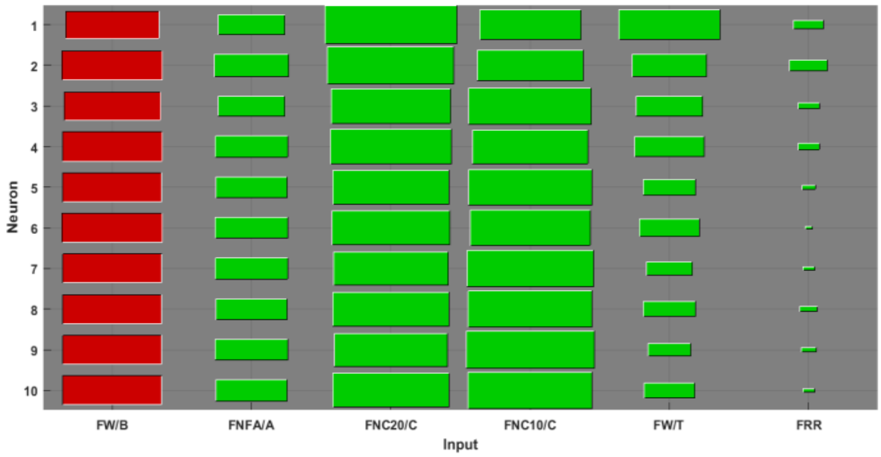


Fig. 6. Hinton diagram for ANN4-2.

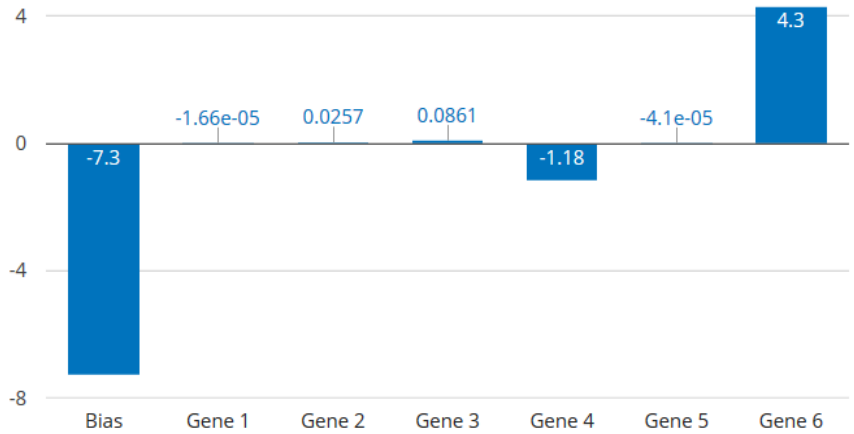


Fig. 7. Weights of the genes (sub-programs) of MGGP1-1.

Table 10. Individual genes/model terms for the prediction of CS for MGGP2-1.

Term	Value
Bias	14.5
Gene 1	$(1.2\ x_1^2)\ /\ (x_6 - 1.0\ x_7)$
Gene 2	$-(0.0139\ x_1^2\ x_7)\ /\ (x_6 + x_7)$
Gene 3	$-(1.16\ x_1^2)\ /\ (x_6 + x_7)$
Gene 4	$(0.00531\ x_2\ x_5)\ /\ (\text{psqroot}(x_6) + \text{psqroot}(x_7))$
Gene 5	$-(9.19\ x_2\ x_5)\ /\ x_6 + 5.08)^2$
Gene 6	$-\big(0.00149\ (x_3 + x_7)\ (x_3 + x_6 + \text{psqroot}(x_6))\big)\ /\ (2.0\ x_2 + x_6)$

Each gene includes its weighing coefficient. It is seen that the weight of the genes (sub-programs) 1, 3, 5 and the bias terms are higher than the other genes for MGGP2-1 and high importance of gene 6, 4, 3 and bias term in MGGP1-1. This means that they have higher contribution to the strength prediction of concrete. Fig. 8 shows the expressional trees for the best models that were obtained during the conducted runs for MGGP3-1. Each gene includes its weighing coefficient. As can be observed from Fig. 8, the derived model is composed of complicated array of operators, variables, and constants to estimate the 28 CS of concrete.

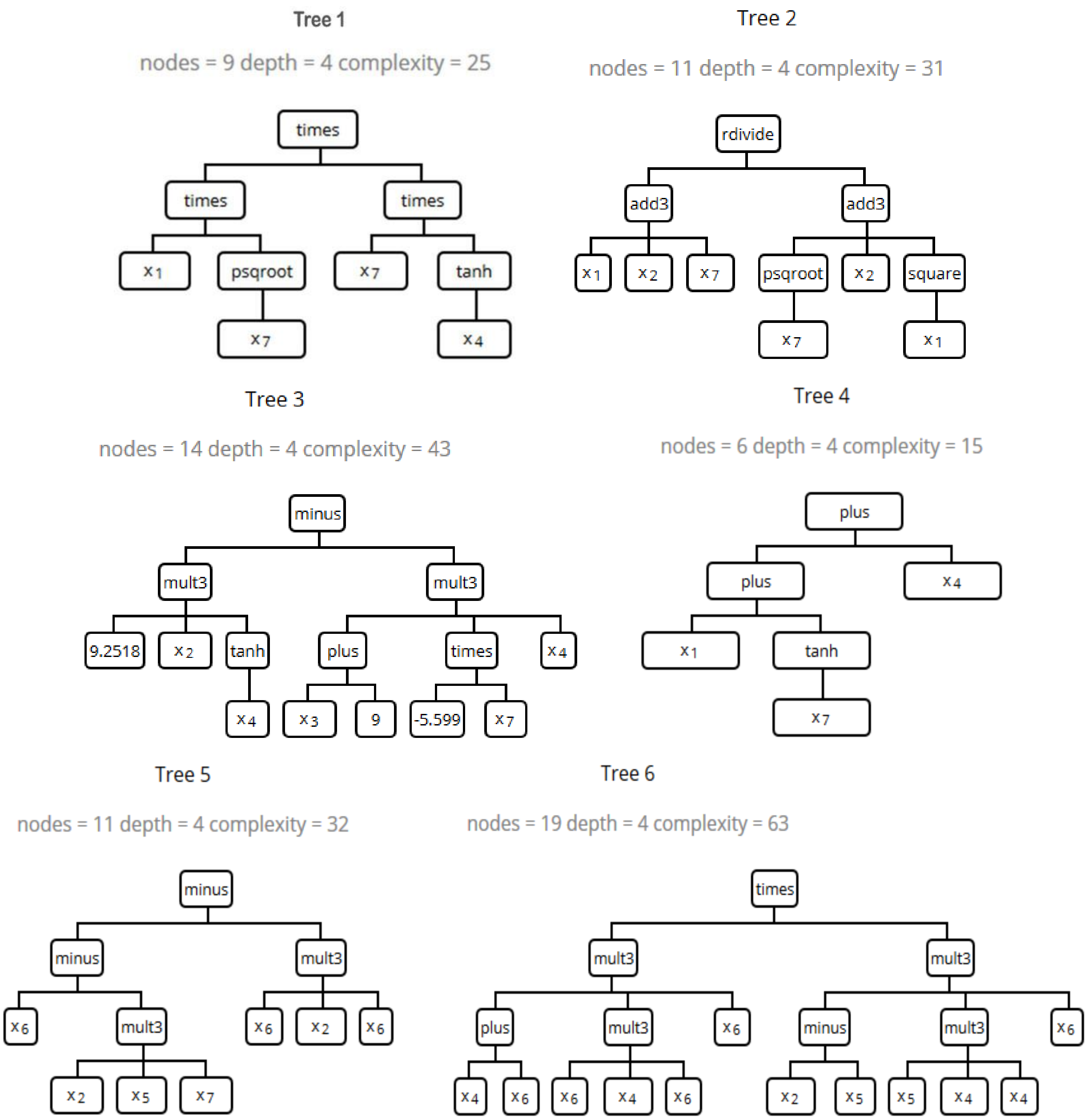


Fig. 8. Expression trees of the best models for the prediction of CS of concrete for MGG 3-1.

To facilitate the use of the developed model, model MGGP3-1 was transformed into a simplified functional form (Eq. (6)):

$$y = 16.9\, x_6 - 94.0\, x_4 - 94.0\, x_1 - 94.0 \tanh(x_7) - (1.0\,(7.79e^{15}\, x_1 + 7.79e^{15}\, x_2 + 7.79e15\, x_7)) / (1.76e^{13}\, x_2 + 1.76e^{13} \cdot \mathrm{sqrt}(x_7) + 1.76e^{13}\, x_1^2) - 16.9\, x_2\, x_6^2 + 153.0\, x_2 \tanh(x_4) + 92.9\, x4\, x7\, (x_3 + 9.0) - 16.9\, x_2\, x_5\, x_7 - 4988.0\, x_1\, x_7 \cdot \mathrm{sqrt}(x_7) \tanh(x_4) + 995.0\, x_4^3\, x_5\, x_6^4\, (x_4 + x_6)\,(x_2 - 1.0\, x_5) + 550.0 \qquad (6)$$

where x_1=RWC, x_2=RNFA/A, x_3=RNC20/C, x_4= RNC10/C, x_5=RC20/C, x_6=RC10/C, x_7=RW/T.

Fig. 9 presents the accuracy against the complexity of the evolved models. Green dots represent the Pareto front of models in terms of model performance (1 – R^2) and model complexity. Blue dots represent non-Pareto models. The red circled dot represents the best model in the population in terms of $R2$ on the training data (Searson, 2015). The red circle in Fig. 10 for MGGP3-1 designates the best model presented herein that is not outperformed by any other model in terms of complexity and fitness. A less complex model for MGGP2-1 can be seen in Fig. 9.

From the Pareto front (Figs. 9 and 10), user can decide whether the incremental gain in performance is worth with associated model complexity. Concisely, the MGGP paradigm evolves multiple models which provide more number of choices to the designer. A single model can be selected based on the application requirements (Searson et al., 2007; Pandey, 2015). Figs. 11 and 12 also depict the convergence characteristics of the genetic programming algorithm.

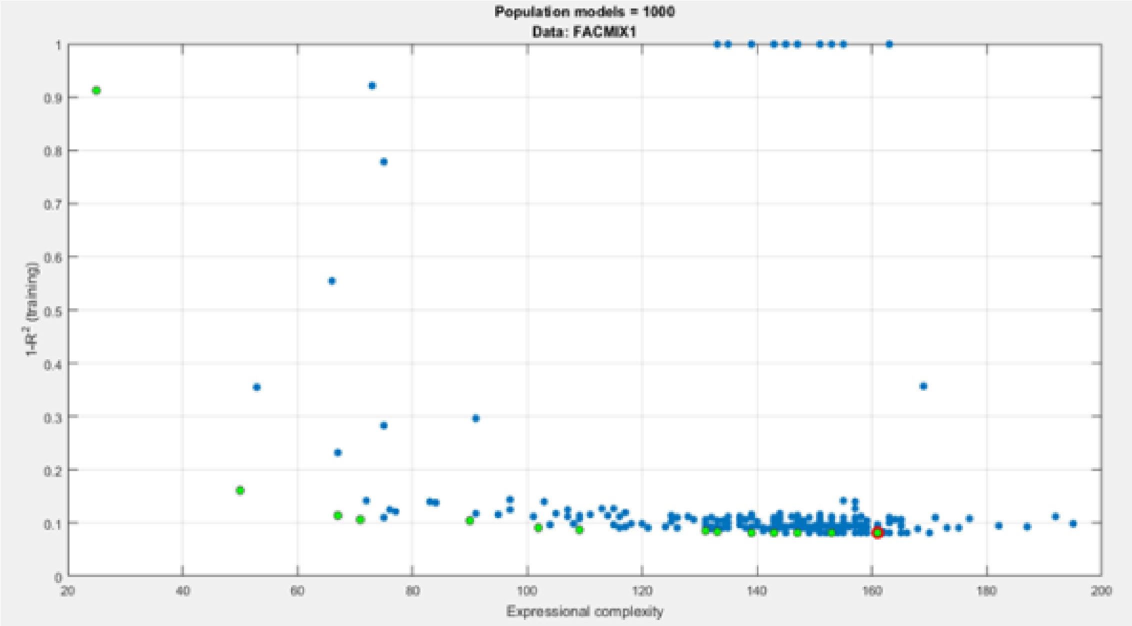


Fig. 9. Pareto front report for MGGP2-1.

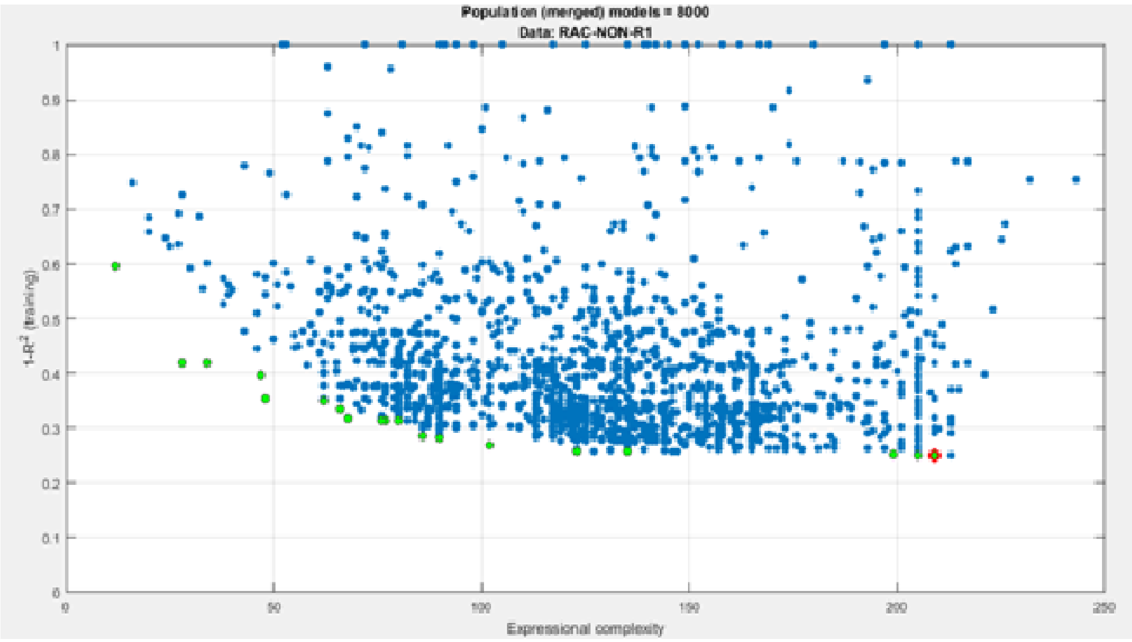


Fig. 10. Pareto front report for MGGP3-1.

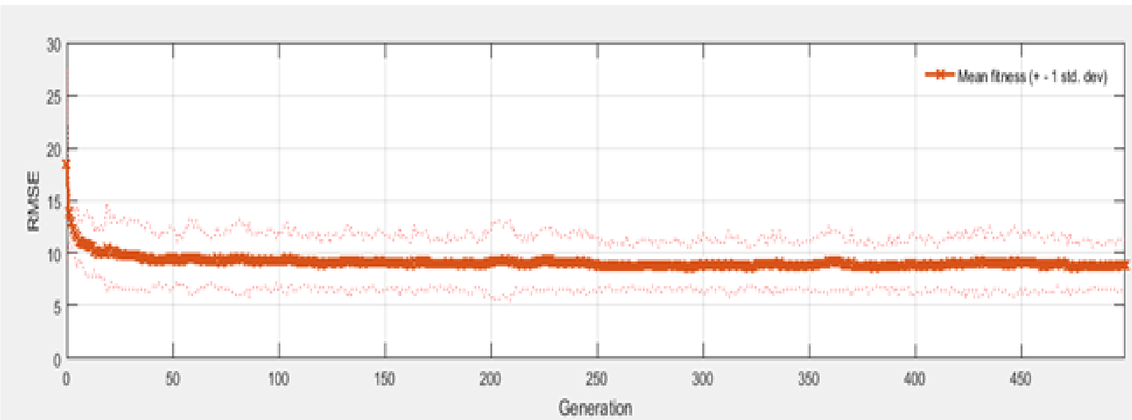


Fig. 11. Convergence of the MGGP solutions for MGGP1-2.

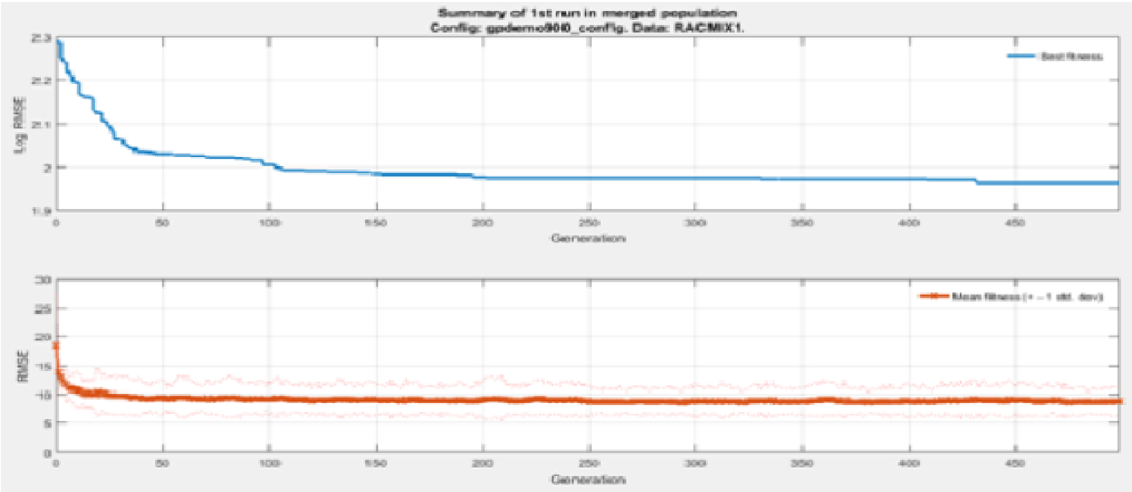


Fig. 12. Convergence of the MGGP solutions for MGGP3-2.

It is evident that the mean fitness of the curve becomes smoother after 100 (in both the figures) generations and that the change in objective function is not significant near the end of the genetic programming run. It indicates that running the genetic programming for more generations does not result in a more favorable outcome. However, as the best fitness is reported at 450 (RMSE-7.14 for MGGP1-2) in this particular case, it suggests that the genetic programming algorithm should have to run for at least 500 generations for all the models developed here. Figs. 13-16 show the frequency of input data for models in GP with coefficient of determination $R^2 \geq 0.6$ (Searson, 2010; Searson et al., 2007; Pandey, 2015). Input frequency of the graphical input frequency analysis of single model or of a user specified fraction of the population is used to provide the identification of Input variables that are significant to the output (Searson et al., 2007; Singh, 2014).

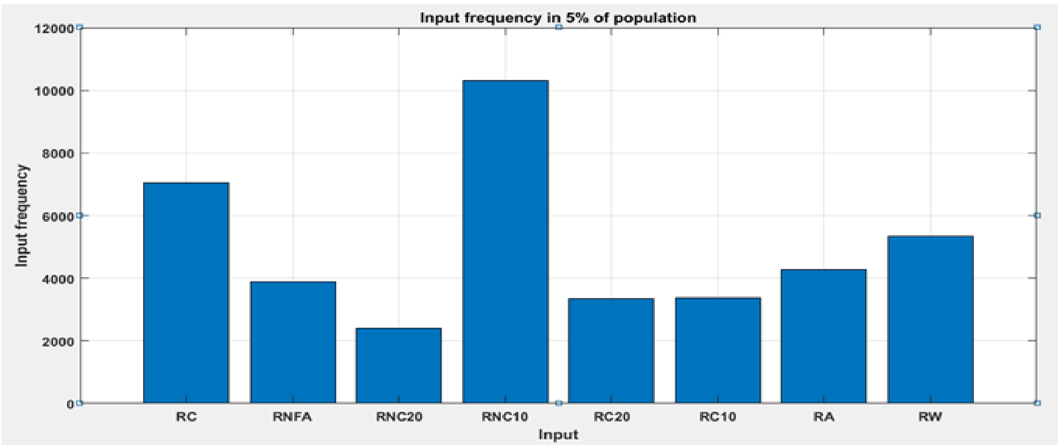


Fig. 13. Input frequency for MGGP1-1.

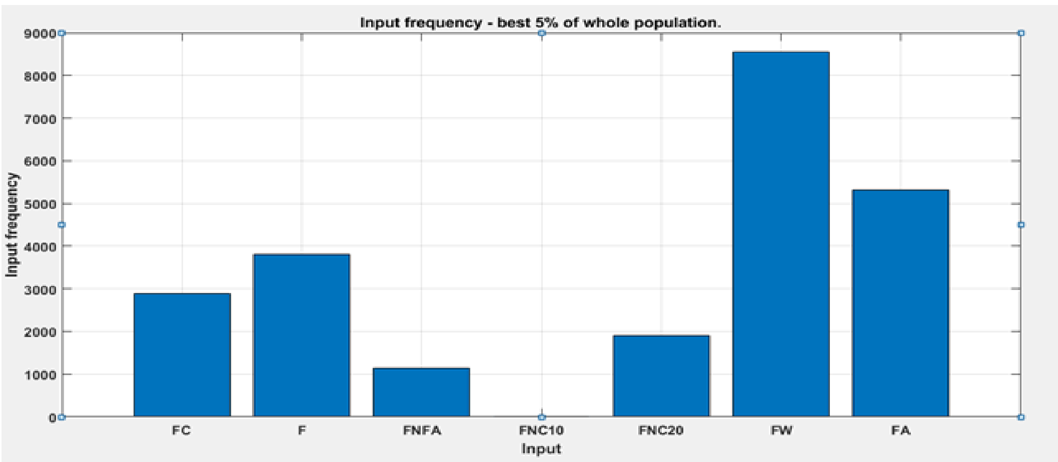


Fig. 14. Input frequency for MGGP2-1.

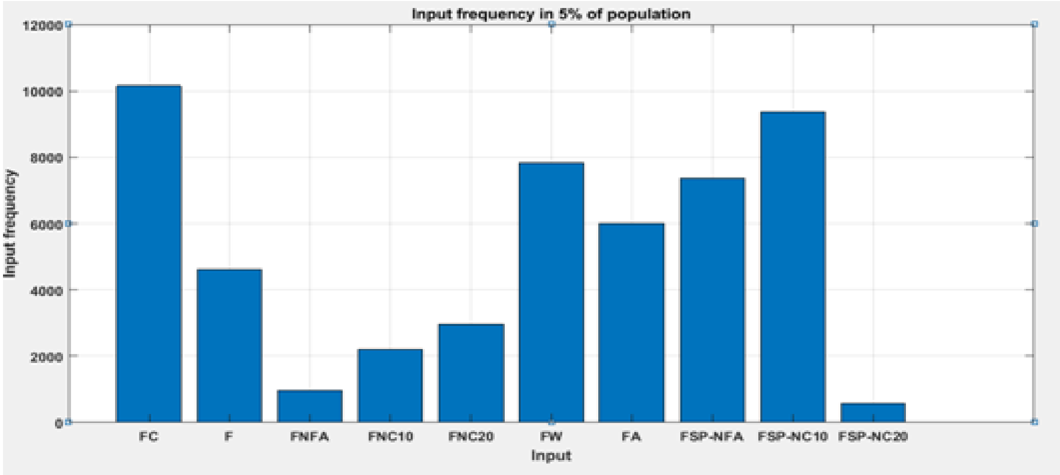


Fig. 15. Input frequency for MGGP2-2.

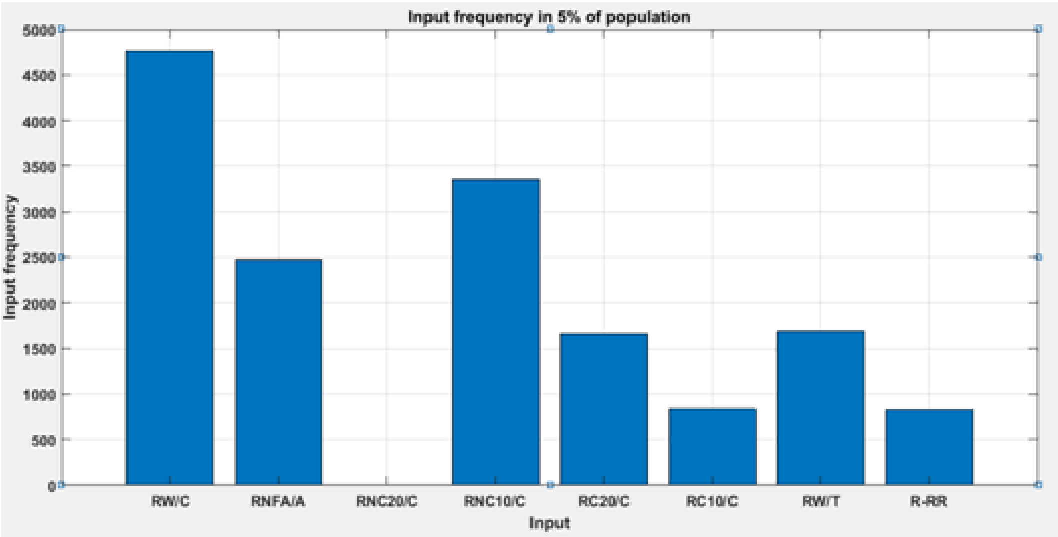


Fig. 16. Input frequency for MGGP3-2.

For MGGP1-1, out of 8 parameters, parameters: RNC10, RC and RW are influential followed by RNFA, RA, RC20, RC10 and RNC20 content. This finding is in tune with the fundamental knowledge of concrete technology (Shetty, 2005; Neville, 2012). For MGGP2-1, FW and F are seen as important parameters. With addition of properties of materials like Specific gravity of aggregate content, MGGP2-2 shows specific gravity as an influential parameter with cement content being the most influential parameter followed by water content and other parameters. In MGGP3-2 the frequency of input parameters is as shown in figure 16 which shows RC20/C as the least influential parameter.

4.3. Models developed using MT

Model tree is the third technique used to predict the 28 day strength of concrete. Fig. 17 below shows a typical Model Tree developed for Model 1-1. The linear regression equations developed by MT1-1 are shown in Fig. 18.

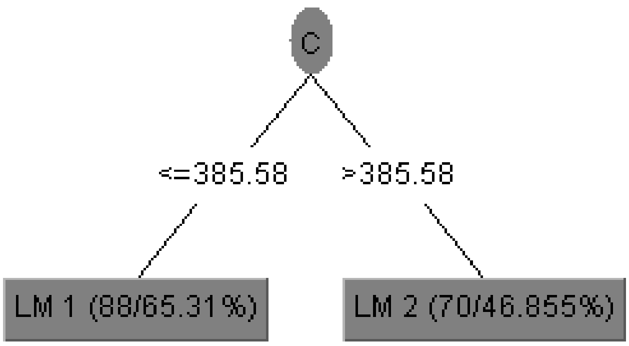


Fig. 17. Model Tree for MT1-1.
(The first number in the bracket is the number of samples in the subset sorted to this leave and the second one-root mean squared error (*RMSE*) of the corresponding linear model divided by the standard deviation of the samples subset for which it is built.) (expressed in percent)

```
Scheme:     weka.classifiers.trees.M5P -M 4.0
Relation:   RACData2-MT-TRAINING
Instances:  158
Attributes: 9
            C, NFA, NC-20, NC-10, RC-20, RC-10, A, W, S
=== Classifier model (full training set) ===
M5 pruned model tree: (using smoothed linear models)

C <= 385.58 : LM1 (88/65.31%)
C > 385.58 : LM2 (70/46.855%)

LM num: 1
S = 0.0142 * C - 0.002 * NFA - 0.0039 * NC-20 + 0.0172 * NC-10 - 0.0048 * RC-20
        - 0.0118 * RC-10 + 0.8318 * A - 0.0365 * W + 39.8858

LM num: 2
S = 0.0912 * C - 0.022 * NFA - 0.0047 * NC-20 + 0.0485 * NC-10 - 0.0058 * RC-20
        - 0.0044 * RC-10 - 0.3575 * W + 93.4442

Number of Rules : 2

Time taken to build model: 0.26 seconds
```

Fig. 18. Equations developed for MT1-1.

Similarly for other models, the number of equations developed are shown in Table 7. Equation developed for MT1-1 as shown in Fig. 17, shows positive coefficients to cement content, admixture content and RC-10 content.

Negative coefficients can be seen for other parameters specially water content, indicating that its increase in mix after a certain limit can decrease the strength of concrete which agrees with the domain knowledge (Shetty, 2005; Neville, 2012; Deshpande, 2016). This can also be seen in MT equation developed for Fly ash based concrete i.e Model MGGP2-1. The series of equations developed for MT2-1 areas shown in Fig. 17. Inclusion of replacement ratio in models for RAC and Fly ash based concrete in Model 3-2 and 4-2 are shown in Fig. 19 for RAC and Fig. 20 or Fly ash based concrete.

The performance of each model developed using MT in each set are as shown in Table 11.

```
Scheme:     weka.classifiers.trees.M5P -M 4.0
Relation:   RACR2-TRAINING
Instances:  159
Attributes: 9
            W/C, NFA/A, RNC20/C, RNC10/C, RC20/C, RC10/C, RW/T, R-RR, RCS
Test mode:  evaluate on training data
M5 pruned model tree: (using smoothed linear models)
W/C <= 0.452 : LM1 (62/57.194%) and W/C > 0.452 : LM2 (97/69.324%)
LM num: 1
RCS = -119.3856 * W/C + 2.2316 * NFA/A + 8.3196 * RNC10/C - 3.2418 * RC20/C + 27.2358
      * RW/T + 99.9119
LM num: 2
RCS = -12.915 * W/C + 1.5343 * NFA/A - 3.8624 * RNC20/C + 1.4344 * RNC10/C - 5.5796 *
      RC20/C - 6.1185 * RC10/C + 18.7246 * RW/T + 53.5031
Number of Rules : 2
```

Fig. 19. Equations developed for MT3-2.

```
Scheme:     weka.classifiers.trees.M5P -M 4.0
Relation:   NON-FAC-MD4-ANN-MT-TRAINING
Instances:  79
Attributes: 7
            FW/B , FNFA/A , FNC20/C, FNC10/C, FW/T, F-RR, FCS
Test mode:  user supplied test set: size unknown (reading incrementally)
=== Classifier model (full training set) ===
M5 pruned model tree:
(using smoothed linear models)
LM1 (79/39.428%)
LM num: 1
FCS = -66.8101 * FW/B - 50.6429 * FNFA/A - 3.9907 * FNC20/C - 2.9493 * FNC10/C
        - 0.167 *F RR + 103.5807
Number of Rules : 1
```

Fig. 20. Equations developed for MT4-2.

Table 11. Performance of models developed using MT.

Set. No	Model No.	*RMSE* (m/s)	*MAE* (m/s)	*E*	*AARE*	*r*
1	MT 1-1	10.66	7.213	0.586	17.123	0.767
	MT 1-2	9.883	6.176	0.706	14.011	0.843
2	MT 2-1	4.544	3.442	0.77	10.916	0.902
	MT 2-2	4.641	3.861	0.797	11.284	0.929
3	MT 3-1	10.66	7.212	0.586	17.122	0.767
	MT 3-2	10.66	7.212	0.586	17.122	0.767
4	MT 4-1	8.053	3.992	0.278	15.326	0.795
	MT 4-2	6.944	3.987	0.463	14.917	0.817

4.4. Comparison of models developed using ANN, MGGP and MT

The models were developed using same data division and their results were compared on testing data sets as shown in Tables 8, 9 and 11. Performance of each of the model was judged using 5 statistical error measures namely *RMSE*, *MAE*, *E*, *AARE* and *r*.

ANN outperformed the other data driven techniques as seen in the Tables 8, 9 and 11. ANN predicted the output of 28 day CS of concrete better, as compared to MGGP and MT. Figs. 21and 22 show the predictions of RAC and Fly ash based concrete in models 1 in Set 1 and Set 2.

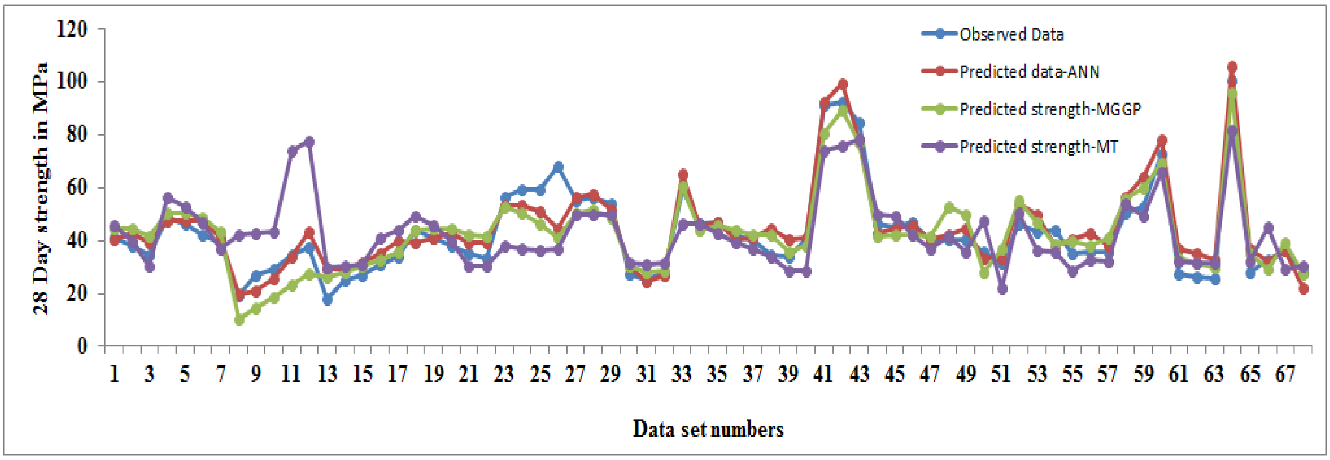


Fig. 21. Prediction trend for model 1-1.

Fig. 22. Prediction trend for model 2-1.

On the other hand, MGGP model performs better than MT models. Table 8 shows that the performance of model ANN1-1 is better as compared to ANN1-2. This can be also be seen in models developed using MGGP. Models developed using kg/m^3 proportions of materials predict strength better as compared to models developed with additional properties of materials i.e water absorption of aggregates in Set 1 in ANN. A similar performance can be seen with models developed using MGGP when Specific gravity of MNFA, FNCA20 and FNC10 become part of input parameters in Set 2. Presence of non-dimensional parameters as input parameters for development of models using ANN, MGGP and MT show a similar performance in RAC and fly ash based concrete. Scatter plots for ANN1-1 and MGGP1-1 are as shown in Figs. 23 and 24, respectively. Scatter plot for MGGP1-1 shows slight under prediction of RAC.

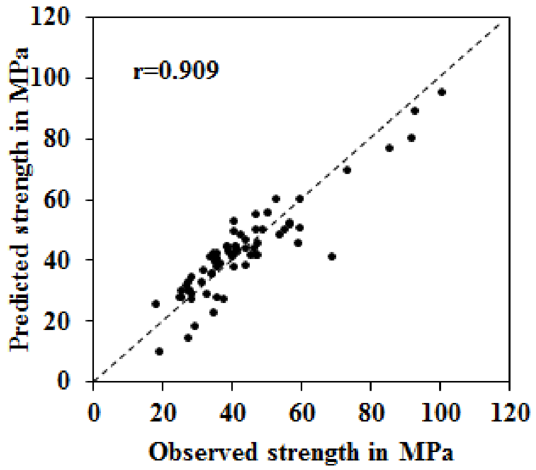


Fig. 23. Scatter plot for MGGP 1-1.

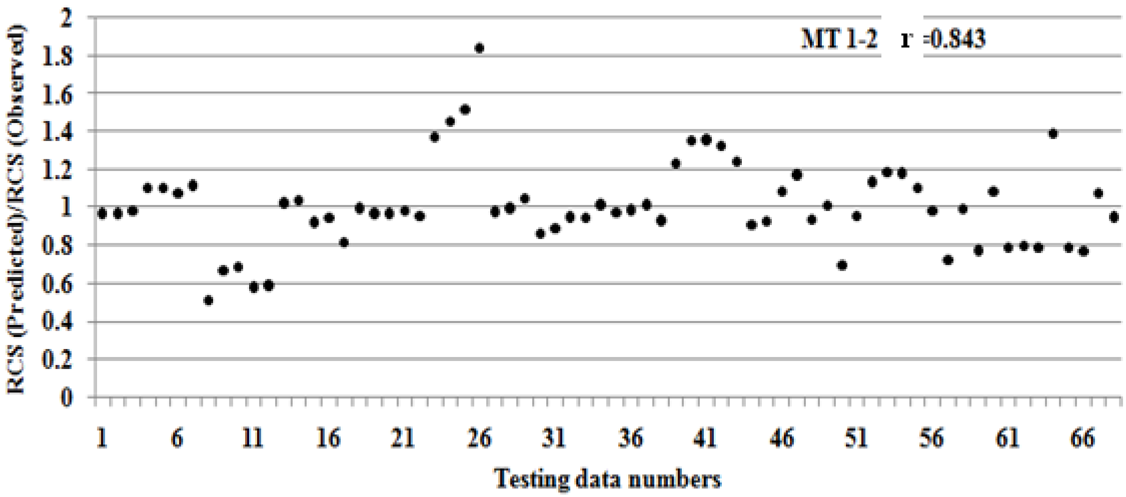


Fig. 24. Observed and Predicted values for MT 1-2.

However models in Set 1 and Set 2 developed using MT show performance of models with properties of materials better than relative proportions of materials i.e than MT1-1. This can be seen through lower *r*, *E* and higher *AARE* values for MT1-1. The same can be seen with models developed with fly ash based concrete too. The plot in Fig. 24 shows the trend of RAC prediction and observed for MT1-2. This figure shows the ratio of the predicted to observed RAC strength values. Apparently, a ratio closer to 1 indicates a more precise prediction.

Non-dimensional parameters have a greater significance in ascertaining strength characteristics of concrete. It has also been seen that instead of using proportions of materials as input parameters, if individual non-dimensional parameters are used, the performance of the models can be similar as that of the former or increases (Deshpande, 2014). Thus Set 3 and Set 4 were designed with non-dimensional parameters as input parameters. It was seen that with higher *r*, *AARE* and *E* values, models with non-dimensional parameters perform better when developed using ANN. A slight decrease in the performance of model ANN3-2 can be seen as compared to ANN3-1 in which *RR* was an additional parameter included. Thus it can be said that the influence of *R-RR* parameter is considered in the parameter: aggregate proportion. Similar performance can be seen in fly ash based concrete (model ANN4-1 and ANN4-2). MGGP3-1 shows a similar performance as MGGP1-1 and with *R-RR* as input parameter in MGGP3-2 shows *r* as 0.864. An increased (*r*=0.952) performance can be seen in MGGP4-2 as compared to MGGP4-1 (*r*=0.918).

Models developed using MT for RAC i.e MT3-1 and MT3-2 showed a similar performance. MT4-2 shows an increase in performance when *F-RR* was considered as an input parameter as compared to MT4-1.

Thus it can be said that models developed using ANN and MGGP displayed a good performance with relative proportions of constituents of materials in concrete; Non-dimensional parameters as input parameters performed similarly or slightly better than the model with proportions of materials as input parameters, however it is necessary to select non-dimensional parameters judiciously for a better representation of material proportions. Models developed using MT performed poor as compared to ANN and MGGP, however they have an advantage of series of equations which can be readily used. In the Hinton diagram for ANN1-1, RNC20, RW and RC are influential factors followed by other parameters. A similar trend of influential parameters can be seen in MGGP1-1 as well. A similar trend in terms of coefficients can be seen in model MT1-1 developed using MT (refer Fig. 19). Similarly for models ANN2-1, MGGP2-1 and MT2-1 with cement and fly ash content and specific gravity of fine aggregate are the most important parameters followed by other parameters. Parameter SP-NC20 was eliminated from MT models and seen as very low influence in ANN and MGGP models.

With non -dimensional parameters as input parameters, the influential parameters in ANN3-1, MGGP3-1 and MT3-1 are as shown in Figs. 25, 26 and 27, respectively. A slight difference in the influential parameters can be seen. ANN builds an approximate function that matches a list of inputs to the desired outputs. In the process, it adjusts the weights and biases to reach a predefined goal. This process makes ANN flexible and increases its performance as compared to GP.

GP, on the other hand, is based on evolutionary approach technique in which it does not involve any transfer function and evolves generations of 'offspring' based on the 'fitness criteria' and genetic operations. GP approach works with the concept of disregarding input parameters that do not contribute effectively in the model and thus based solely on 'fitness' criteria. In the process of building programs (through processes of mutation, crossover and reproduction), GP selects parameters which are useful in achieving the fitness criteria and deletes the remaining. This can reduce the - performance of GP as compared to ANN. This is in contrast to the work done by Londhe (2008). However the work in Londhe (2008), the problems were indeterministic in nature; whereas predicting strength of concrete is of deterministic nature. This may be the reason behind ANN working better than GP. However this needs to be explored further.

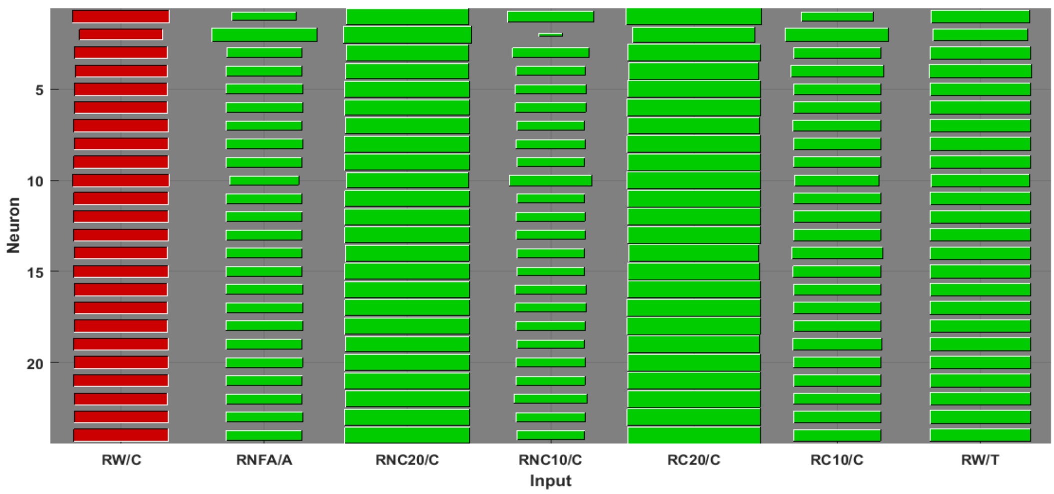


Fig. 25. Hinton diagram for ANN3-1.

In Model tree with M5P algorithm, the basic tree is formed based on a splitting criterion. It uses the standard deviation of the class values for each node as a measure of the error at that node and then calculates the expected error reduction as a result of testing each attribute at that node. Then, the attribute that maximizes the expected error reduction is selected to split the data at that node and the remaining are not considered in the developed equation. Thus RC10/C is included in Eq. (2) of MT3-2 and excluded in equation 1 of MT3-2 (refer Fig. 27). This can also be one of the reasons for poorer performance of MT as compared to ANN and MGGP. The errors of a good prediction model should be independent of physical parameters involved in that problem. Otherwise, it can be concluded that those physical parameters should be added to that prediction model or they weren't considered correctly in that model. It should be mentioned that the errors of developed models in set 3 and 4, show a similar or increased performance when RR is included as input parameter in ANN, MT and MGGP, except for model MGGP3-2. However the authors recommend the use of individual aggregate ratios as they display a better picture about contribution of each aggregate type on the strength of concrete. Thus it can be seen that models developed using ANN, MGGP and MT learn from the examples given and predict the strength of concrete with influential parameters which are in tune with the domain knowledge. The correlation of input parameters and output parameters seen in Tables 1, 2 and 3 also show significant parameters as cement content followed by water and aggregates for the current study.

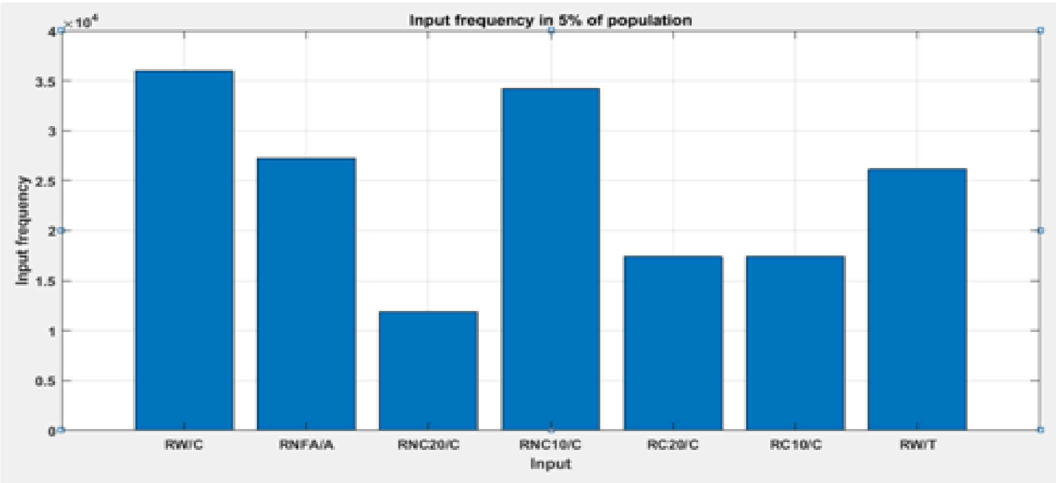


Fig. 26. Input frequency for MGGP3-1.

```
M5 pruned model tree:
(using smoothed linear models)

RW/C <= 0.452 : LM1 (62/57.016%)
RW/C >  0.452 : LM2 (96/69.464%)

LM num: 1
RCS =
        -119.4165 * RW/C
        + 2.2235 * RNFA/A
        + 8.3283 * RNC10/C
        - 3.2418 * RC20/C
        + 27.9404 * RW/T
        + 99.8744

LM num: 2
RCS =
        -13.0527 * RW/C
        + 1.5424 * RNFA/A
        - 3.8705 * RNC20/C
        + 1.4534 * RNC10/C
        - 5.587 * RC20/C
        - 6.1269 * RC10/C
        + 19.3821 * RW/T
        + 53.5492
```

Fig. 27. Equations for MT3-1.

Thus, models developed using ANN technique perform better as compared to MGGP and MT. MGGP on other hand perform better than MT. This can be seen in case of RAC and fly ash based concrete and different properties of materials. ANN though performs better; has a limitation of its ease of its use. MGGP and MT on other hand are easy to use with equation and series of equations developed respectively for ready use. Performance of MGGP and MT though less as compared to ANN, equations are developed by the techniques by understanding the basics of domain knowledge which can be seen through the equations.

5. Conclusions

In the current study an attempt was made to predict 28 day compressive strength of Recycled Aggregate concrete and Fly ash based concrete with input parameters as kg/m^3 proportions of materials used in concrete, properties of materials used and non-dimensional parameters. The following outcomes can be noted from the current study:

- Models developed using ANN outperform MGGP and MT models with higher *R, AARE* and *E* values and lower RMSE and MAE values.
- ANN has an advantage of better performance; MGGP on other hand with acceptable accuracy can provide equations which can be readily used. Models developed using MT display performance less as compared to ANN and MGGP.
- Use of relative proportions of materials as input parameters predicts strength better than input parameters with properties of materials in ANN and MGGP. However MT shows performance with properties of materials better as compared to relative proportions of materials.
- Use of Non- dimensional parameters as input parameters can be encouraged for prediction of CS of concrete; however judicious selection of non-dimensional parameters needs to be done.

The study also shows that ANN, MGGP and MT learn from the examples given and display influential input parameters which are in tune with the domain knowledge of Concrete Technology specifically for ANN. GP and MT show a slight variation as compared to influential parameters observed in ANN and maybe due to the basic working concept of MGGP and MT.

REFERENCES

Adnan SH, Rahman I.A, Saman HM (2011). Recycled aggregate as coarse aggregate replacement in concrete mixes. *ASEAN Australian Engineering Congress 2011*, Kuching, Sarawak, Malaysia, 2011.

Adriana TAD, Monica BL, Koji de JN (2013). Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks. *Construction and Building materials*, 38, 717-722.

Agarwal SD, Karnnkal AB, Patil PN (2011). Feasibility study of use of RCA in concrete. *International Journal of Research in Engineering Sciences and Technology*, SP-II, 70-75.

Ahmet OZ, Murat PB, Erdogan O, Erdog K, Naci C, Bhatti A (2006). Predicting the compressive strength and slump of high strength concrete using neural network. *Construction and Building Materials*, 20, 769–775.

Bayazidi BM, Wang GG, Amir H, Alavi H, Gandomi AH (2014). Multigene genetic programming for estimation of elastic modulus of concrete. *Mathematical Problems in Engineering,* 2014, 1-10.

ChakradharaRao M, Bhattacharyya SK, Barai SV (2010). Influence of field recycled coarse aggregate on properties of concrete, *Materials and Structures*, 44(1), 205-220.

Corinaldesi V (2010). Mechanical and elastic behavior of concretes made of recycled-concrete coarse aggregates. *Construction and Building Materials*, 24, 1616–1620.

Dapena E, Alaejos P, Lobet A, Pérez D (2011). Effect of recycled sand content on characteristics of mortars and concretes. *Journal of Materials in Civil Engineering*, 23(4), 414-422.

David RL, Gregory JJr (1999). Evaluating the use of "goodness of fit" measures in hydrological and hydro climatic model validation. *Water Resources Research*, 35(1), 233-241.

Deepa C, SathiyaKumari K, Sudha KP (2010). Prediction of the compressive strength of high performance concrete mix using tree based modeling. *International Journal of Computer Applications*, 6(5), 18–24.

Deshpande N (2016). Study of Properties of Concrete Made with Recycled Aggregate and Industrial Wastes with Special Application to Housing. *Ph.D thesis*, Shivaji University.

Deshpande N, Londhe SN, Kulkarni SS (2014). Modeling compressive strength of recycled aggregate concrete by Artificial Neural Network, Model Tree and Non-linear Regression. *International Journal of Sustainable Built Environment*, 3, 187–198.

Dias WPS, Pooliyadda SP. (2001). Neural networks for predicting properties of concretes with admixtures. *Construction and Building Materials*, 15, 371-379.

Domingo-Cabo A, Lázaro C, López-Gayarre C, Serrano-López MA, Serna P, Castaño-Tabares JO (2009). Creep and shrinkage of recycled aggregate concrete. *Construction and Building Materials*, 23, 2545–2553.

Duan ZH, Kou SC, Poon CS. (2013). Prediction of compressive strength of recycled aggregate concrete using artificial neural networks. *Construction and Building Materials*, 40, 1200–1206.

Evangelista L, Brito JCde (2010). Durability performance of concrete made with fine recycled concrete aggregates. *Cement & Concrete Composites,* 32, 9–14.

Evangelista LR, Brito JCde (2004). Criteria for the use of fine recycled concrete aggregates in concrete production. *International RILEM Conference on Use of Recycled Building Materials in Building and Structures*, 8-11 Nov 2004, Barcelona, Spain.

Fathifazl G, Abbas A, Razaqpur AG, Isgor OB, Fournier B, Foo S (2009). New mixture proportioning method for concrete made with coarse recycled concrete aggregate. *Journal of Materials in Civil Engineering*, 21(10), 601-611.

Frank E, Mark A, Hall Ian, Witten H (2016). The WEKA Workbench. Fourth Edition. *Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques.* Morgan Kaufmann.

Goncalves A, Esteves A, Vieria M (2004). Influence of recycled concrete aggregates on concrete durability. *International RILEM Conference on Use of Recycled Building Materials in Building and Structures*, 8-11 Nov 2004, Barcelona, Spain.

Gorphade VH, Sudarsana Rao H, Beulah M (2014). Development of genetic algorithm based neural network model for predicting workability and strength of high performance concrete. *International Journal of Inventive Engineering and Sciences (IJIES),* 2(6), 1-8.

Gupta R, Kewalramani MA, Goel A (2004). Prediction of concrete strength using neural-expert system. *Journal of Materials in Civil Engineering*, 18(3), 462-466.

Hansen TC, Narud H (1983). Strength of recycled concrete made from crushed concrete coarse aggregate. *Concrete des Constr,* 5(1), 79-83.

I-Cheng Y (2007). Modeling slump flow of concrete using second-order regressions and artificial neural networks. *Cement & Concrete Composites*, 29, 474–480.

Jain A, KumarJha S, Misra S (2008). Modeling and Analysis of concrete slump using Artificial Neural Networks. *Journal of materials in Civil engineering*, 20(9), 628-633.

Jong-In KM, Kim DK, Maria MQ, Feng M, Yazdani F (2004). Application of Neural Networks for Estimation of Concrete Strength. *Journal of Materials in Civil Engineering*, May/June 2004, 257-264.

Khatib JM (2005). Properties of concrete incorporating fine recycled aggregate. *Cement Concrete Research,* 35(4), 763-769.

Konin A, Kouadio DM (2011). Influence of cement content on Recycled aggregates concrete properties. *Modern Applied Science*, 5(1), 23-31.

Kotrayothar D (2012). Recycled Aggregate Concrete for Structural Applications. *Ph.D thesis,* University of Western Sydney.

Kou S (2006). Reusing Recycled Aggregates in Structural Concrete. *Ph.D thesis,* Hong Kong Polytechnic University.

Kumutha R, Vijai K (2010). Strength of concrete incorporating aggregates recycled from demolition waste. *ARPN Journal of Engineering and Applied Sciences*, 5(5). 64-71.

Londhe SN (2008). Soft computing approach for real-time estimation of missing wave heights. *Ocean Engineering*, 35, 1080–1089.

Londhe SN, Dixit PR (2012). Genetic Programming: A Novel Computing Approach in Modeling Water Flows Genetic Programming – New Approaches and Successful Applications. Chapter 9, InTech.

Maier H, Dandy G (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modeling & Software*, 15(1), 101-124.

Neville AM (2012). Properties of Concrete. Pearson Education, USA and UK.

Ni HG, Wang JZ (2000). Prediction of compressive strength of concrete by neural networks. *Cement and Concrete Research*, 30, 1245-1250.

Nikoo M, Torabian Moghadam F, Sadowski A (2015). Prediction of concrete compressive strength by evolutionary artificial neural networks. *Advances in Materials Science and Engineering,* 2015, 1-8.

Padmini AK, Ramamurthy K, Matthews MS (2003). Relative moisture movement through recycled aggregate concrete. *Cement and Concrete Research*, 54(5), 377-384.

Pandey DS, Pan I, Das S, Leahy JJ, Kwapinski W (2015). Multi-gene genetic programming based predictive models for municipal solid waste gasification in a fluidized bed gasifier. *Bioresource Technology*, 179, 524–533.

Pelufo M J, Domingo A, Ulloa VA, Vergara NN (2009). Analysis of moisture state of recycled coarse aggregate and its influence on compression strength of the concrete. *Proceedings of the International Association for Shell and Spatial Structures (IASS) Symposium 2009, Valencia Evolution and Trends in Design, Analysis and Construction of Shell and Spatial Structures,* 28 September–2 October 2009, Universidad Politecnica de Valencia, Spain

Pereira P, Evangelista L, Brito Jde (2012). The effect of superplasticisers on the workability and compressive strength of concrete made with fine recycled concrete aggregates. *Construction and Building Materials*, 28, 722–729.

Poon CS, Kou SC, Lam L (2007). Influence of recycled aggregate on slump and bleeding of fresh concrete. *Materials and Structures,* 40, 981–988.

Poon CS, Shui ZH, Lam L, Fok H, Kou SC (2004). Influence of moisture states of natural and recycled aggregates on the slump and compressive strength of concrete. *Cement and Concrete Research*, 34, 31–36.

Quinlan JR (1992). Learning with Continuous Classes. *In: Adams and Sterling, (Eds.). Proceedings of 5th Australian Joint Conference on Artificial Intelligence*. World Scientific, Singapore.

Ryu JS (2002). An experimental study on the effect of recycled aggregate on concrete properties. *Magazine of Concrete Research*, 54(1), 7-12.

Sarıdemir M (2010). Genetic programming approach for prediction of compressive strength of concretes containing rice husk ash. *Construction and Building Materials*, 24, 1911–1919.

Schoppe BM (2011). Shrinkage and modulus of elasticity in concrete with recycled aggregates. *M.Sc. thesis*, California Polytechnic State University.

Searson DP (2015). GPTIPS 2: an open-source software platform for symbolic data mining. Chapter 22 in Handbook of Genetic Programming Applications, A.H. Gandomi et al., (Eds.), Springer, New York, NY.

Searson DP, Leahy DE, Willis MJ (2010). GPTIPS: An open source genetic programming toolbox for multigene symbolic regression. *Proceedings of the International Multi Conference of Engineers and Computer Scientists 2010*, I, IMECS 2010, March 17-19, 2010, Hong Kong.

Searson DP, Willis MJ, Montague GA (2007). Co-evolution of non-linear PLS model components. *Journal of Chemometrics*, 2, 592-603.

Searson PD (2015). GPTIPS 2: An open-source software platform for symbolic data mining. Cornell University Library. DOI: arXiv:1412.4690

Shetty MS (2005). Concrete Technology. 17th edition. S. Chand and Company, New Delhi.

Singh HK (2014). Prediction of shear strength of deep beam using Genetic Programming. *B.Tech thesis*, National Institute of Technology, Rourkela, India.

The ASCE Task Committee (2000). Artificial Neural Networks in Hydrology I: Preliminary concepts. *Journal of Hydrologic Engineering,* 5(2), 115-123.

Yaprak H, Aruntas HY, Demir I, Simsek O, Durmus G (2011). Effects of the fine recycled concrete aggregates on the concrete properties. *International Journal of the Physical Science,* 6(10), 2455-2461.

Yong PC, Teo DCL (2009). Utilization of Recycled aggregate as coarse aggregate in concrete. *UNIMAS-E Journal of Civil Engineering*, 1(1), 1-6.

Yueh TY, Chen YY, Hwang CL (2006). Properties of HPC with recycled aggregates. *Cement and Concrete research*, 36, 943-950.

Zega CJ, Di Maio AA (2003). Recycled concrete made with different natural coarse aggregates exposed to high temperature. *Construction and Building Materials*, 23, 2047–2052.