Construction and Building Materials 142 (2017) 199–207

Contents lists available at ScienceDirect

Construction and Building Materials

journal homepage: www.elsevier.com/locate/conbuildmat

# Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm

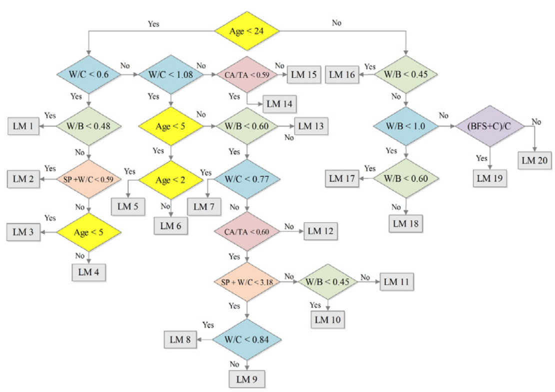
Ali Behnood [a,*], Venous Behnood [b], Mahsa Modiri Gharehveran [a], Kursat Esat Alyamac [c,a]

[a] Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall, West Lafayette, IN 47907-2051, USA
[b] University of Mohaghegh Ardabili, Daneshgah Street, 56199-11367 Ardabil, Iran
[c] Engineering Faculty, Civil Engineering Department, Firat University, 23119 Elazig, Turkey

## HIGHLIGHTS

- M5P model tree was used to predict the compressive strength of high performance concrete.
- The model is developed based using data collected from published literature.
- Log transformation of input and output parameters was used to account for non-linearity.
- The model developed shows accuracy above 80%.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 13 January 2017
Received in revised form 20 February 2017
Accepted 9 March 2017
Available online 17 March 2017

*Keywords:*
M5P model tree
High performance concrete
Estimation of compressive strength
Fly ash
Blast furnace slag

## ABSTRACT

Compressive strength of concrete is one the parameters required in many design codes. A reliable prediction of it can save in time and cost by quickly generating the needed design data. In addition, it can reduce the material waste by reducing the number of trial mixes. In this study, M5P model tree algorithm was used to predict the compressive strength of normal concrete (NC) and high performance concrete (HPC). Compared to other soft computing methods, model trees are able to offer two main advantages: (a) they are able to provide mathematical equations and offer more insight into the obtained equations and (b) they are more convenient to develop and implement. To develop the model tree, a total of 1912 distinctive data records were collected from internationally published literature. Overall, the results show that M5P model tree can be a better alternative approach for prediction of the compressive strength of NC and HPC using the amount of constituents of concrete as input parameters.



## 1. Introduction

Generally, normal concrete (NC) consists of three main components: Portland cement, coarse and fine aggregates, and water. However, making of the high-performance concrete (HPC) requires the addition of supplementary cementitious materials (SCM) such as fly ash (FA), blast furnace slag (BFS), silica fume and chemical admixtures (e.g., superplasticizer). One of the main differences

* Corresponding author.
*E-mail addresses:* abehnood@purdue.edu (A. Behnood), vbehnood@gmail.com (V. Behnood), mmodirig@purdue.edu (M. Modiri Gharehveran), kealyamac@firat.edu.tr, kalyamac@purdue.edu (K.E. Alyamac).

http://dx.doi.org/10.1016/j.conbuildmat.2017.03.061
0950-0618/

**Nomenclature**

| | |
|---|---|
| ANFIS | adaptive neuro-fuzzy inference systems |
| ANN | artificial neural network |
| BFS | blast furnace slag |
| C | ordinary Portland cement |
| CA | coarse aggregate |
| CA/B | coarse aggregate-binder ratio |
| CA/TA | coarse aggregate-total aggregate ratio |
| CC | correlation coefficient |
| CS | compressive strength |
| $\frac{C+FA}{C}$ | cement plus fly ash-cement ratio |
| $\frac{C+BFS}{C}$ | cement plus blast furnace slag-blast furnace slag ratio |
| DR | discrepancy ratio |
| FA | fly ash |
| FRP | fiber-reinforced polymer |
| GP | genetic programming |
| HCR | hierarchical classification and regression |
| HPC | high performance concrete |
| LM | linear model |
| MAE | mean absolute error |
| MAPE | mean absolute percentage error |
| NC | normal concrete |
| $R^2$ | linear correlation coefficient |
| RMSE | root mean square error |
| S | sand |
| SCM | supplementary cementitious material |
| SDR | standard deviation reduction |
| SP | superplasticizer |
| SP + W/C | summation of the amount of superplasticizer and W/C |
| SRL | slope of regression line |
| SVM | support vector machine |
| W | water |
| W/B | water-binder ratio |
| W/C | water-cement ratio |

between NC and HPC is the use of chemical and mineral admixtures to reduce the water content. Reduced water content can lead to reduced porosity and increased strength properties [1,2]. Due to the high strength properties of HPC, it has been widely used in many special structures such as tunnels, precast units and nuclear structures [3].

Predicting the mechanical properties of concretes such as compressive strength (CS) is an important task due to the fact that it is the required input parameter in many existing design practices [4,5]. Moreover, it can help to schedule operations such prestressing and removal of framework. A reliable and accurate model for prediction of the CS can save in time and cost by quickly generating the needed design data [6]. Many empirical equations presented in the codes and standards for such prediction are developed for NC. Therefore, the suitability and validity of these models for prediction of the CS of HPC, which has different properties compared to NC, might be questioned.

Machine learning techniques have been widely used in many fields of civil engineering for forecasting, planning, and management purposes. As an example, Sajjadi et al. used Extreme Learning Machine for prediction of heat load for individual consumers in district heating systems [7]. The predicted model enabled them to increase the fuel efficiency distribution by considering nine predictive models for different prediction horizons. Bilir et al. used Takagi and adaptive neuro-fuzzy inference systems (ANFIS) to predict the restrained drying shrinkage crack width of slag mortar composites [8]. The estimated predictive model for drying shrinkage allows designers to save in time and cost by taking precautions in designing process. In addition, this model would help to improve the durability of the concrete by reducing or even preventing the cracking. Lim et al. used genetic algorithm to determine the desired amount of different constituents of HPC in order to obtain a specific performance [3]. Reduced number of trial mixes could be one of the main advantages of using such model. To determine the longitudinal dispersion coefficient in rivers, Etemad-Shahidi and Taghipour utilized M5P model tree [9]. The reliable estimation of this coefficient could be helpful in determining the distribution of pollution concentration with high accuracy. The performance of the model developed by Etemad-Shahidi and Taghipour was compared with that of the previously existing formulas [9]. The findings from their study showed that the developed model outperforms the other previously existing formulas. In the area of pavement engineering, such techniques have been used to predict the properties of asphalt mixtures such as flow number [10,11]. Several researchers have used machine learning techniques to predict the properties of concretes such as ultimate conditions of fiber-reinforced polymer (FRP)-confined concrete [12], modulus elasticity of recycled aggregate concretes [4], slag concretes [13], rubberized concrete [14], high strength concretes [15] and self-compacting concretes [16], splitting tensile strength of steel fiber-reinforced concrete [5] and so on.

Some of the machine learning techniques that have been used to predict the CS of concretes include artificial neural network (ANN) [17–25], classification and regression trees [6], adaptive neuro-fuzzy inference systems (ANFIS) [17,26], support vector machine (SVM) [25,27], and genetic programming [28]. Öztaş et al. [18], Yeh [20], and Raghu Prasad et al. [29] studied the applicability of ANN, as one the most widely used forecasting technique, to predict the CS of HPC. These studies confirmed the suitability of ANN in prediction of the CS of HPC. The proposed ANN models can reduce the material waste and save in mix design time and cost. ANN has also been used for the prediction of the CS of HPC containing nano silica and copper slag [30]. Mousavi et al. [31] used a comprehensive dataset obtained from the literature to develop a predictive model for CS of HPC based on gene expression programming approach. In their study, a sensitivity analysis was also used to evaluate the contribution of the parameters affecting the CS. The performance of the developed model was reported to be superior than that of the regression models. Chou and Tsai [32] proposed a hierarchical classification and regression (HCR) approach in predicting the CS of HPC. By using a dataset that was obtained from laboratory tests, it was reported that HCR approach outperformed conventional flat prediction models such as linear regression, ANN and SVM.

Although ANN provides a more reliable and accurate alternative to statistical regression and numerical methods, it is not as transparent as regression-based methods and formulas. It is due to the fact that a mathematical expression (between the input and output variables of the system) is not developed as part of this modeling approach. In addition, ANN approach requires trial and error method to find the network parameters such as number of hidden layers and neurons, which makes it time consuming. Algorithms such as genetic programing (GP) or decision tree can be used to overcome these shortcomings [4]. GP is a technique that uses genetic algorithm to evolve the encoded set of genes. GP is a domain-independent method that generates prediction equation