AUT Journal of Civil Engineering

AUT J. Civil Eng., 4(4) (2020) 423-436
DOI: 10.22060/ajce.2019.16124.5569

Developing a Novel Machine Learning Method to Predict the Compressive Strength of Fly Ash Concrete in Different Ages

H. Naseri[1], H. Jahanbakhsh[1,2], F. Moghadas Nejad[1,*], A. Golroo[1]

[1] Department of Civil and Environmental Engineering, Amirkabir University of Technology, Tehran, Iran.
[2] Asphalt Technology Laboratory, University of Science and Culture, Tehran, Iran.

ABSTRACT: Estimating the compressive strength of concrete before fabricating, has been one of the most important challenges because designing a mixture proportion by experimental methods needs expert workers, consumes energy, and wastes materials. Therefore, in this study, the influences of materials and the age of samples on the compressive strength of fly ash concrete are investigated, and a novel method for predicting the compressive strength is presented. To this end, the water cycle algorithm and genetic algorithm are utilized, and their outcomes are compared with the classical regression models. Various performance indicators are used to gauge the accuracy of the models. By analyzing the results, it is comprehended that the water cycle algorithm is the most accurate model according to all performance indicators. Besides, the outcomes of the water cycle algorithm and genetic algorithm are by far better than those of classical methods. The mean absolute error of water cycle algorithm, genetic algorithm, linear regression, partial-fractional regression, and fractional regression are 3.01, 3.12, 5.47, 9.70, and 5.37 MPa for training data and 2.90, 3.44, 5.47, 9.70, and 5.37 MPa for testing data respectively. Furthermore, the water cycle algorithm is the only algorithm whose mean absolute error of testing data is less than that of training data. At last, it was concluded that the mixture with less than 35% fly ash (weight of the binder) had maximum amounts of compressive strength. Also, the compressive strength of concrete decreased significantly as the amount of fly ash increased more than this definite level.

Review History:

Received: Apr. 14, 2019
Revised: Jul. 14, 2019
Accepted: Jul. 22, 2019
Available Online: Jul. 24, 2019

Keywords:

Water Cycle Algorithm
Genetic Algorithm
Regression
Fly Ash Concrete
Compressive Strength

1- Introduction

The cement considered as the global anthropogenic e-CO2 emissions through the more than 5-7% of CO2 emission and utilizing a significant amount of energy by the cement industry [1, 2]. Furthermore, about 1.5 tons of Raw materials along with 3000-4300 MJ of fuel energy and 120-160 kWh of electrical energy are needed for producing each ton of cement [2]. Based on the aforementioned concepts, utilizing some strategies to reduce cement consumption should be of concern. This can lead to preserving the environment, Raw materials, fuel, and energy with the reduction of pollution emitted during the process of cement production. To this end, the researchers have been interested in utilizing the waste materials and by-products of industries as cement replacement [3, 4].

Based on the aforementioned concepts, supplementary cementitious materials (SCM) have been extensively used in the concrete industry to improve the mechanical properties and durability of the concrete. By these materials, the more economical profit can be gained by replacing a substantial part of the Portland cement by cheap natural pozzolans or industrial by-products. SCMs are known as eco-friendly materials because the CO2 emissions are considerably reduced by decreasing the amount of cement in mixture proportions. Moreover, industrial and societal by-products and wastes can be utilized as construction materials [5]. As such, Fly ash as the most valuable SCMs is commonly used cement replacement material in recent years. Up to 50% of cement can be replaced with fly ash, and it is competent to use in precast elements and reinforced cement concrete construction. The appropriate use of fly ash can prevent expansion due to alkali-silica reaction (ASR) in concrete. The use of fly ash will, by far increase the service life of structures exposed to chloride environments. It is also fruitful to enhance the long term strength of concrete structures [6, 7].

One of the most important issues regarding the fabrication of a suitable concrete mixture containing SCMs is finding a suitable mixture design. As reported by some researchers, a direct correlation exists between the amounts of materials in mixture proportion and properties of concrete [8, 9]. In previous decades, experimental tests were conducted to find the appropriate mixture design. Nonetheless, considering the required several expert workers takes a long time, and wastes irreproducible materials, an experimental method would not be an appropriate procedure designing the convenient environment-friendly concrete. Moreover, the appropriate mixture design may not be found by this technique. Therefore, a wide range of computational approaches has been taken to recognize the optimal mixture design of different sorts of

*Corresponding author's email: moghadas@aut.ac.ir.

concretes and to predict their characteristics. To this end, Adaptive neuro-fuzzy inference system (ANFIS), extreme learning machine model, artificial neural network (ANN), hybrid artificial neural network, various kinds of regressions, deep learning theory, M5P model are the commonly used methods to find the relation between inputs and output(s) in modeling the systems in many engineering applications especially concrete mix design [8-12].

Various kinds of regressions and multiple regressions have been used to predict the compressive strength of concrete as one of the most important mechanical properties and the overall quality of concrete to design optimal mixture characteristics [9-14]. However, their accuracy may not be ideal. ANNs are also the conventional tools that help the researchers to predict the mechanical properties, workability, and durability properties of concrete before casting it. Despite they are powerful and trustable methods, they are considered black-box tools. That is to say, generating practical equations are usually infeasible. On the other hand, the number of hidden layer nodes and neurons along with biases and weights are calculated by trial and error to achieve suitable performance [15, 16].

Fly ash concretes known as high long-term compressive strength mixtures. Therefore, several efforts have been made to predict the compressive strength of this type of concrete with new machine learning methods. Regarding the Prediction of the characteristics of fly ash concretes, Taehwan Kim et al. utilized a particle Model and a new classification approach to analyze fly ash particle characterization for predicting concrete compressive strength [17]. Moreover, the prediction of compressive strength of concrete with fly ash as sand replacement material was considered in the Rajamane investigation [18]. With the help of ANN and ANFIS, the compressive strength of fly ash concretes was predicted in previous studies [19, 20].

Regarding all the above concepts, although common techniques are capable of predicting the compressive strength of fly ash concretes, their accuracy may not be perfect moreover, because the concrete compressive strength test is generally conducted at 7 or 28 days of curing. Ergo, most research studies applied machine learning methods to predict the strength of concrete after 28-days of curing. On the flip side, as the fly ash substantially affects the compressive strength of concrete over long ages, previously developed prediction methods cannot be efficient for the design of fly ash concrete. To the best of the authors' knowledge, there is no machine learning prediction approach of concrete's compressive strength considering the age of casted concrete. Therefore, this research aimed to propose a predictive relationship between concrete strength and mix design parameters considering the age of specimens employing a new machine learning method. To this end, several techniques containing linear regression (LR), fractional regression (FR), partial fractional regression (PFR), Genetic algorithm (GA), and Water cycle algorithm (WCA) were used in this study. Comparative study of the application of these methods to the design of fly ash concrete has also been carried out.

2- Methodology

This study aimed to find the relation between the compressive strength of fly ash concrete and the amounts of materials used in the mixture design. This relation helps to find the appropriate mixture proportion before casting the concrete. To reach the objective of this research, 239 experimental data are used which are extracted from international published articles [21, 22]. The age of samples (day), their components (kg/m^3) including cement, fly ash, water, superplasticizer, coarse aggregate, and fine aggregate, and the ratios of water to the binder, fly ash to the binder, superplasticizer to binder, fine aggregate to total aggregate, coarse aggregate to binder are considered as the inputs of model and compressive strength is the output of the model. The inputs and output of the problem have various ranges. Thus, they ought to be scaled in the same range. In these situations, all the data are usually scaled between 0 and 1. However, in this study, logarithmic equations can be selected by algorithms, and the amount of 0 is unable to use in these modes. Ergo, drawing on Eq. (1) all the data are scaled between 0.1 and 0.9 [23].

$$S_i = 0.1 + \left(0.9 - 0.1\right) \times \frac{i - i_{min}}{i_{max} - i_{min}} \quad (1)$$

where i is the rough data, S_i is the scale value of data, i_{max}, and i_{min} are the maximum and minimum rough data values, respectively. The standard deviation, maximum, minimum, and average amount of rough data used in this study are presented in Table 1.

In this research, by the advantages of heuristic and meta-heuristic algorithms, a novel regression is developed, finding the correlation between inputs and output of the model with a high level of accuracy. Water cycle algorithm (WCA) and genetic algorithm (GA) are opted for introducing the equation of compressive strength of fly ash concrete because they are one of the most powerful algorithms that have been suitably qualified for solving integer problems. To make the model more authentic, 8000 different modes are considered for each input, and 400 various functions can be selected as the final equation. Accordingly, there are more than 4.39×10^{54} feasible solutions in the feasible region.

To assess the accuracy of the model, 6 parameters including mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), percentage of data which their error are fewer than 5 Mega-Pascal (PE5), coefficient of determination (R^2), and correlation coefficient (R) are used as performance indicators. The equations of these performance indicators are as follows:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \left| EXP_i - PRE_i \right|}{n} \quad (2)$$

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n} (EXP_i - PRE_i)^2}{n} \quad (3)$$

Table 1. Standard deviation, maximum, minimum, and average amount of data

Variables		Abbreviation	Minimum	Maximum	Average	Standard deviation
Input Variable						
Cement	(kg/m^3)	CE	134.7	505	232.5	55.235
Fly ash	(kg/m^3)	FL	59	200.1	123.3	28.088
Water	(kg/m^3)	WA	142	221.4	174.5	17.515
Superplasticizer	(kg/m^3)	SU	0	20	8.8	3.315
Coarse aggregate	(kg/m^3)	CO	801	1098	997	70.209
Fine aggregate	(kg/m^3)	FI	630	905.9	811.4	59.603
Age	(days)	AG	3	100	37.2	30.694
Water to binder ratio		WAB	0.27	0.7	0.5	0.005
Fly ash to binder ratio		FLB	0.1	0.55	0.35	0.008
Superplasticizer to binder ratio		SUB	0	0.06	0.03	0.000
Fine to total aggregate ratio		FIT	0.37	0.51	0.45	0.001
Coarse aggregate to binder ratio		COB	1.6	3.76	2.85	0.182
Output variable						
Compressive Strength (MPa)		CS	8.49	66.42	31.2	13.394

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(EXP_i - PRE_i)^2}{n}} \tag{4}$$

$$\mathrm{PE5} = \frac{\begin{pmatrix}\text{number of data which their} \\ \text{error are fewer than 5MPa}\end{pmatrix} \times 100}{\mathrm{n}} \tag{5}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}\left(EXP_i - \overline{EXP_i}\right) \times \left(PRE_i - \overline{PRE_i}\right)}{\sqrt{\sum_{i=1}^{n}(EXP_i - \overline{EXP_i})^2 \times \sum_{i=1}^{n}(PRE_i - \overline{PRE_i})^2}}\right)^2 \tag{6}$$

$$R = \frac{\sum_{i=1}^{n}(EXP_i - \overline{EXP_i}) \times (PRE_i - \overline{PRE_i})}{\sqrt{\sum_{i=1}^{n}(EXP_i - \overline{EXP_i})^2 \times \sum_{i=1}^{n}(PRE_i - \overline{PRE_i})^2}} \tag{7}$$

where EXP_i $\overline{EXP_i}$ PRE_i $\overline{PRE_i}$ and are the results of experimental data, the average of results of experimental data, predicted data, the average of predicted data, and the number of samples in the order given.

The data are divided into two groups. 200 data randomly are considered as training data utilizing to make the model. Moreover, the remaining data (39 samples) are used to gauge the reliability of the results and validation of the model. On the other hand, the three different kinds of traditional regressions are applied, and the outcomes of algorithms and traditional regressions are compared for both training and testing data. The flowchart of the methodology is demonstrated in Fig.1.

3- Modeling the compressive strength of fly ash concrete

As previously mentioned, conventional methods may be unable to consider various modes and several functions to predict the compressive strength of fly ash concrete. In contrast, heuristic and meta-heuristic algorithms can be qualified for deliberating several modes for each input and complicated functions. In this study, 12 inputs are available, and their relation with the compressive strength of fly ash concrete is investigated. To improve the accuracy of the model and to reduce the error, 20 different functions and 400 positive and negative constant values are chosen for each input, and the modes of each input are the multiplication of the functions and the numbers. That is to say, 20 different functions including logarithmic functions with different basics, trigonometric functions (sin, cos, tan, and cot), different types of radical functions, exponential functions, and the combination of these functions are chosen as the forms of each variable, and these functions are assigned to each input. Moreover, for each variable, a coefficient is selected among 400 alternatives, and it is multiplied by the selected function of the input. Hence, 8000 possible modes are available for each input, and the algorithms opt for one of them to find the ideal correlation. Alternatively, 400 functions are considered

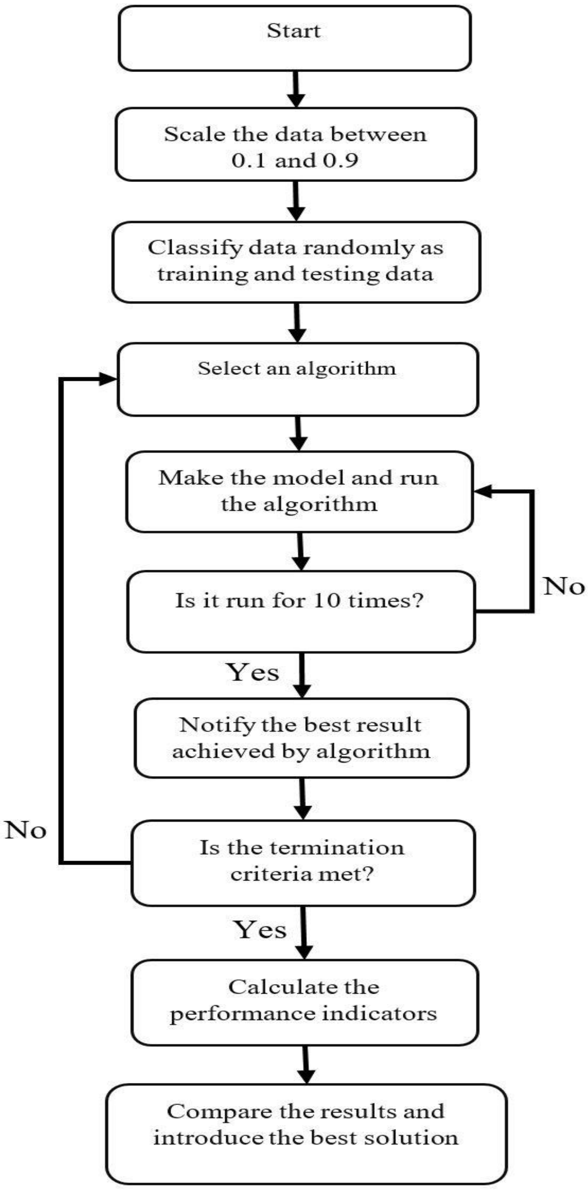


Fig. 1. Flowchart of process.

to find the best combination of inputs. Moreover, several individual constants belong to functions that can be used in their equations.

3.1. Genetic algorithm

A genetic algorithm (GA) is a heuristic algorithm that performs optimization searches in the feasible region to find better solutions. It was inspired by Darwinian principles of natural selection, and it was developed by John H. Holland [24, 25]. In the genetic algorithm, each data is assigned to a chromosome that contains given genes, and each gene represents a feature of data. The population is constituted by chromosomes, and in each iteration, new populations are generated by some process operators including selection operators, mutation operators, and crossover operators. Afterward, novel and previous populations are compared based on fitness value (objective function), and the best chromosomes are survived. Ultimately, the best solution is reported as a genetic algorithm solution [26, 27]. In this study, tournament selection, random selection, and roulette wheel selection operators are chosen as selection operators, and in each selection, one of them is chosen randomly. On the other hand, a two-point (double point) crossover is utilized to cover more data. After running the algorithm, the genetic parameters are adjusted and calibrated. The number of population, number of iterations, crossover percentage, mutation rate, and mutation percentage are considered 500, 6000, 0.8, 0.1, and 0.2, respectively. The algorithm is coded in MATLAB, it is run 10 times and the best result is notified as to the solution of the genetic algorithm. The flow chart of the genetic algorithm used in this study is illustrated in Fig. 2.

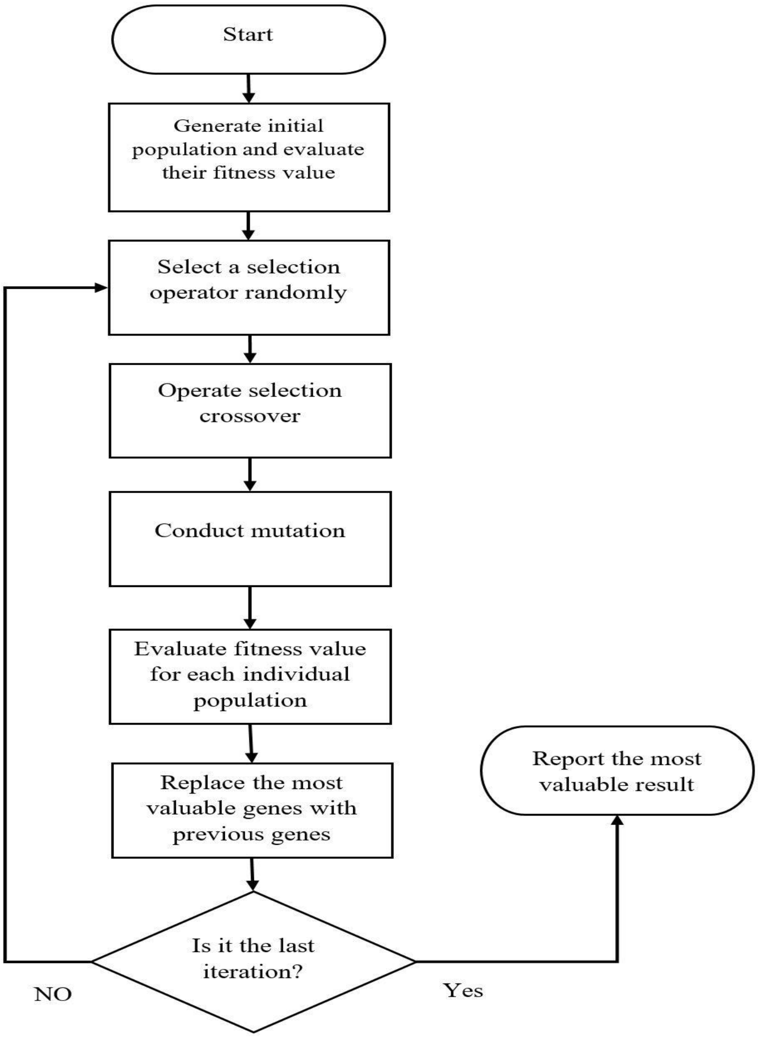


Fig. 2. Flowchart of genetic algorithm.

3.2. Water cycle algorithm

The water cycle algorithm (WCA) is a meta-heuristic algorithm that intelligently looks into the feasible solutions to spot optimal or near-optimal solutions. It was inspired by the water cycle in the environment and how water flows from streams and rivers to the seas [28-30]. In this method, all of the data are considered raindrops. The most valuable data is associated with the sea, and good data are considered rivers. Besides, the remaining data are considered streams that flow to rivers and sea. When the distances of streams and rivers or rivers and sea become less than a given amount, it is rained, and new raindrops are generated [28-30]. In this paper, the water cycle algorithm is coded in MATLAB, and tuning is performed for the parameters of the algorithm. The number of raindrops, number of iterations, number of rivers and sea, and the evaporation condition constant are considered 500, 6000, 8, and a random number fewer than 0.1 which is changed in every iteration. The algorithm is run 10 times, and the best solution is reported. The flow chart of the water cycle algorithm adjusted to solve this problem is displayed in Fig.3.

3.3. Regression

For evaluation of heuristics algorithms, the accuracy of the GA and WCA are compared with the classical regression models. Regressions are statistical tools that find the relationships between the inputs and output(s) variables. The structure of regression models should be defined in advance. Thus, these predetermined structures may reduce the accuracy of the model [31]. Linear regression (LR), fractional regression (FR), and partial fractional regression (PFR) are the most advantageous regressions which have proven useful for many engineering applications [50-54]. Hence, these regressions are utilized to achieve the aims of this study. Similar to GA and WCA models, the data are scaled between 0.1 and 0.9 to enhance the exactness and accuracy of the model. The structure of linear regression, fractional regression, and partial fractional regression used in this study are shown in Eqs. (8), (9), and (10), respectively. In this investigation, Microsoft Office Excel and MATLAB are applied to predict the compressive strength of fly ash concrete by classical regressions.

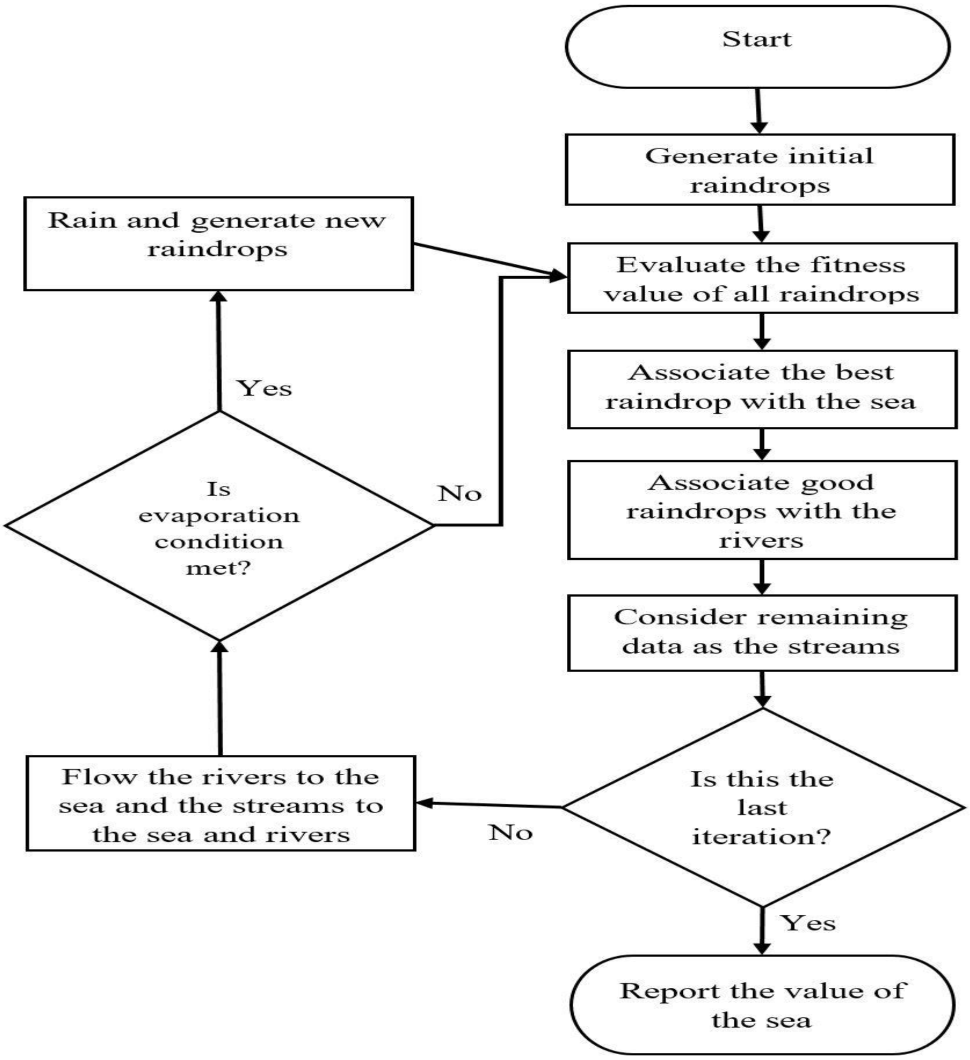


Fig. 3. Flowchart of water cycle algorithms.

$$y = \sum_{i=1}^{n} \left(a_i \times x_i \right) + c \tag{8}$$

$$y = \sum_{j=1}^{m} \left(b_j \times z_j \right) + c \tag{9}$$

$$y = \sum_{i=1}^{n} \left(a_i \times x_i \right) + \sum_{j=1}^{m} \left(b_j \times z_j \right) + c \tag{10}$$

where x_i and z_j represent the value of inputs and the amounts of fractional variables. a_i, b_j, and are the amounts of constants achieved by models. Also, n and m are the numbers of input variables and the fractional variables in the order given.

4- Results and discussion

In this study, the prediction of compressive strength of fly ash concrete is investigated. The age of samples, their ingredients including cement, fly ash, water, superplasticizer, coarse aggregate, and fine aggregate, and five ratios including water to binder, fly ash to the binder, superplasticizer to the binder, fine aggregate to total aggregate, coarse aggregate to binder are considered as the input variables and compressive strength is the output of the model. The model is solved by the genetic algorithm and water cycle algorithm, and their outcomes are compared with linear regression, fractional regression, and partial fractional regression.

Even though considering the age of the samples as the input variable may deteriorate the accuracy of the model, it should be scrutinized in predicting the compressive strength of fly ash concretes. Therefore, 239 data which are the results of compressive strength experimental test 3, 14, 28, 56, and 100 days after making the samples are used in this investigation. These data are randomly categorized into two groups, including training data and testing data. MAE is considered as the objective function, and it is tried to reduce the error of prediction models. After solving the models, they are compared with classical regressions based on performance indicators. Mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), percentage of data which their error are fewer than 5 Mega-Pascal (PE5), coefficient of determination (R^2), and correlation coefficient (R) are the performance indicators used to evaluate and to compare the models.

The genetic algorithm and water cycle algorithm are compared in the same condition. That is to say, the number of populations and the number of iterations are considered the same, they are run 10 times, and the best-achieved solutions are compared with each other. Drawing on results, 3.12, and 3.01 are the best solutions introduced by GA and WCA, respectively. Hence, it can be comprehended that the performance of GA and WCA for training data is approximately the same. However, the performance of WCA is a bit (0.11 MPa) better than that of GA.

The resulted equations of water cycle algorithm, genetic algorithm, linear regression, fractional regression, and partial fractional regression which predict the compressive strength of fly ash concrete trained by training data are displayed in Eqs. (11) to (15) in the order named.

$$\begin{aligned} f'_{c(WCA)} &= \log_{10}((-0.378\times\sqrt{CE})+(1.299\times FL)+(1.25\times 2^{WA}) \\ &+(-0.519\times SU)+(0.4\times 4^{co})+(1.565\times FL)+ \\ &(0.671\times\log_2(AG))-0.35)+[\sin((-0.433\times\log_2(WAB)) \\ &-(0,525\times\tan(FLB)))\times\cos(((0.671\times\sqrt[3]{SUB}) \\ &+(0.3\times\sqrt[3]{FIT}))((.671\times\sqrt[3]{SUB})+(0.3\times\sqrt[3]{FIT} \end{aligned} \tag{11}$$

$$\begin{aligned} f'_{c(GA)} &= \frac{\begin{gathered}0.2\times 3^{CE}+(1.25\times\cos(FL))-(0,1\times\cos(SU)) \\ +(0.075\times\tan(CO)\end{gathered}}{(0.3\times\sin(WA)+(0.05\times\cot(AG))+0.3} \\ &+0.4\times\cos(WAB)+(0.05\times\cos(FLB))-(0.15\times \\ &\tan(SUB))+(0.15\times\sin(FIT)+0.02\times\cos(COB))-0.756 \end{aligned} \tag{12}$$

$$\begin{aligned} f'_{c(LR)} &= (1.256\times CE)+(0.243\times FL)-(0.204\times WA) \\ &+(0.075\times SU)+(0.184\times CO)+(0.241\times FI)+ \\ &(0.464\times AG)-0.472 \end{aligned} \tag{13}$$

$$\begin{aligned} f'_{c(FR)} &= -(0.894\times WAB)-(0.374\times FLB) \\ &-(0.080\times SUB)+(0.107\times FIT)+(0.212\times COB)+0.954 \end{aligned} \tag{14}$$

$$\begin{aligned} f'_{c(PFR)} &= (0.638\times CE)+(0.508\times FL)-(0.297\times WA) \\ &-(0.156\times SU)-(0.036\times CO)+(0.191\times FI)+ \\ &(0.462\times AG)-(0.087\times WAB)-(0.602\times FLB) \\ &+(0.145\times SUB)-(0.162\times FIT) \\ &-(0.001\times COB)+0.305 \end{aligned} \tag{15}$$

where CE, FL, WA, SU, CO, FI, AG, WAB, FLB, SUB, FIT, COB, and f_c' are the scaled values of cement, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, age, water to binder ratio, fly ash to binder ratio, superplasticizer to binder ratio, fine aggregate to total aggregate ratio, coarse aggregate to binder ratio, and compressive strength of fly ash concrete, respectively.

As can be perceived from Eqs. (11) and (12), the water cycle model does not include the scaled value of coarse aggregate to binder ratio. Moreover, the scaled value of the fine aggregate does not exist in the genetic algorithm model. These are because of the flexibility of the models. The heuristic algorithms can select between several alternatives, and it helps to find better answers. Nevertheless, the classical models can find the relation between input variables and output(s) variables by utilizing predetermined structures, and it may decrease the accuracy of these models.

Table 2 provides information about the accuracy of the models. As can be seen, data are sorted into training and testing data and the amounts of mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), percentage of data which their error are fewer than 5 Mega-Pascal (PE5), coefficient of determination (R^2), and correlation coefficient (R) are calculated for them.

The mean absolute error of models is illustrated in Fig. 4. Regarding this figure and Table 2, the water cycle algorithm is the most reliable and trustful model followed by genetic algorithm, linear regression, partial-fractional regression, and fractional regression. The MAE of the water cycle algorithm is 3.01 and 2.90 for training and testing data respectively. In other words, the average error of this algorithm for predicting the compressive strength of fly ash concrete is less than 3 MPa (2.90), which can be overlooked. In addition, the performance of the genetic algorithm can be acceptable, and its MAE for testing data is 3.44 MPa. Conversely, the average absolute error of all regression models is more than 5.3 MPa for testing data they may not be allowable. Needless to say, fractional regression is the worst model, and its MAE for testing data is equal to 9.70 MPa. Thus, it is not logical to apply this model for the prediction of the compressive strength of concrete.

Fig.5 gives details about the root mean square error of the model for training and testing data. Whatever the amount of RMSE is closer to 0, it is comprehended that the accuracy of the model is better. Concerning Fig.5 and Table 2, the most reliable model is relevant to the water cycle algorithm with the RMSE of 3.76 and 3.88 for training and testing data in the order given. Furthermore, the next valuable model is the genetic algorithm, and its RMSE is equal to 4.29 and 5.01 for training and testing data, respectively. In addition, linear regression and partial-fractional regression are the next authentic models, and their performance based on RMSE is approximately the same. The RMSE of linear regression and partial-fractional regression are nearly 5.4 and 6.7 for training and testing data in the order named. Concerning the results, fractional regression is the inaccurate model, and it may not be capable of finding the relation between the compressive

Table 2. The amounts of performance indicators of models for training and testing data.

Method	Data	MAE (MPa)	PE5 (%)	MSE (MPa^2)	RMSE (MPa)	R	R^2
WCA	Training data	3.01	83.5	14.11	3.76	0.96	0.92
	Testing data	2.90	79.49	15.03	3.88	0.94	0.89
GA	Training data	3.12	81.5	18.43	4.29	0.95	0.90
	Testing data	3.44	71.79	25.06	5.01	0.92	0.85
Linear regression	Training data	4.36	64	29.50	5.43	0.92	0.84
	Testing data	5.47	56.41	46.12	6.79	0.87	0.75
Fractional regression	Training data	7.90	39	94.54	9.72	0.70	0.49
	Testing data	9.70	25.64	141.91	11.91	0.28	0.08
Partial-fractional regression	Training data	4.34	63	28.76	5.36	0.92	0.84
	Testing data	5.37	56.41	44.96	6.70	0.87	0.75

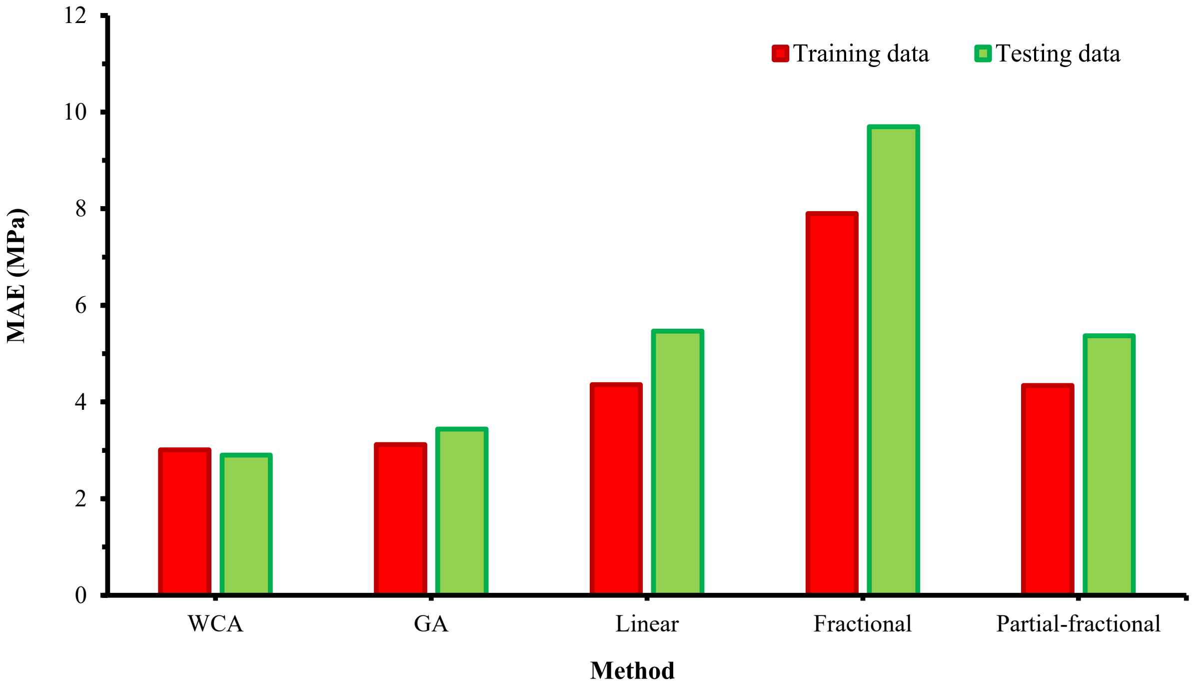


Fig. 4. Mean absolute error of the models.

strength of fly ash concrete and its proportions. According to RMSE, the accuracy of the water cycle algorithm is much better than classical regressions.

The coefficient of determination of the models (R^2) is the other performance indicator used in this study to compare the results of the various model. If the coefficient of determination of a model is greater than the others, it is concluded that it provides the best prediction model among all alternatives. The coefficient of determination of the models is indicated in Fig.6 and Table 2. As can be seen, the most amount of coefficient of determination is related to water cycle algorithm, followed by genetic

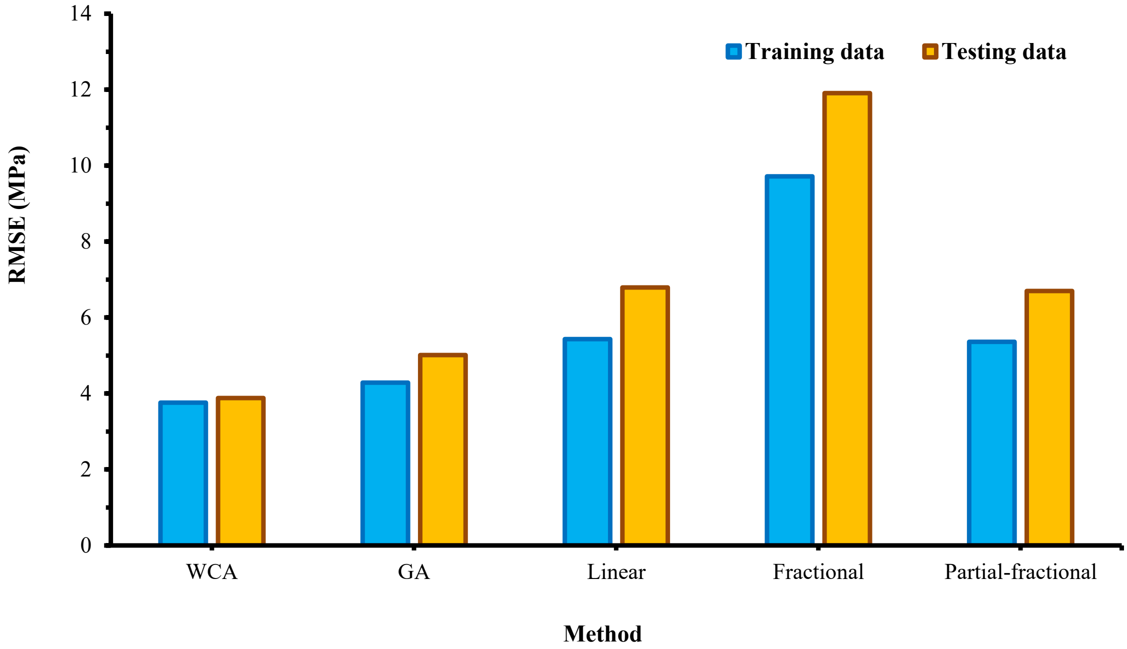


Fig. 5. Root mean square error of the models.

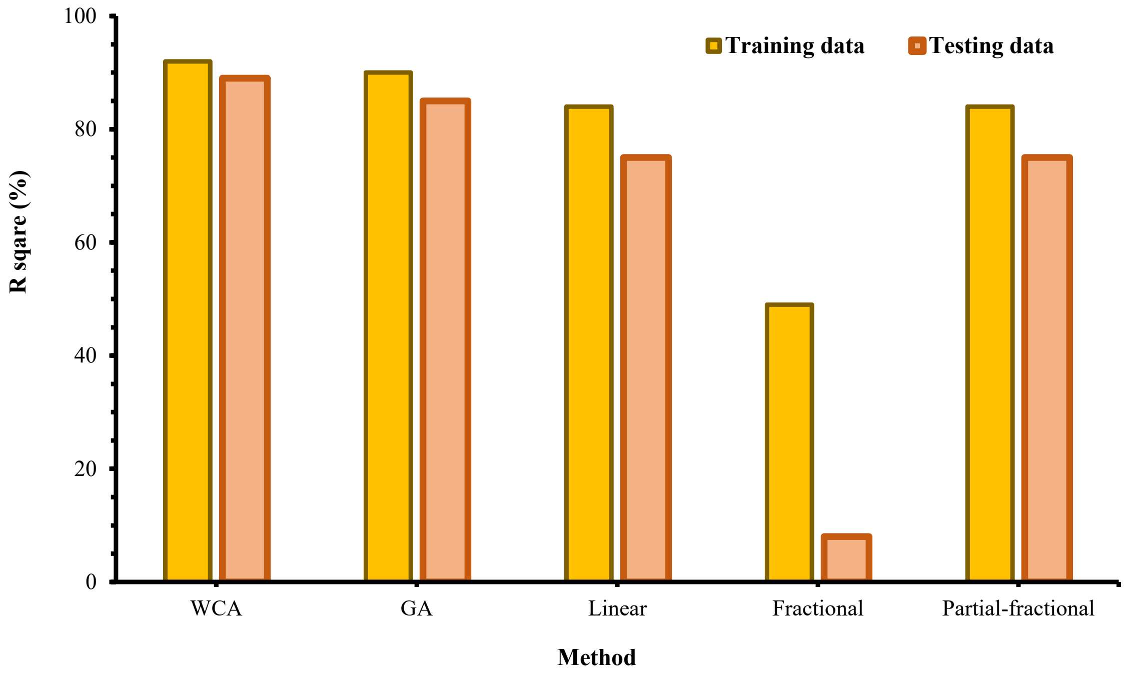


Fig. 6. Coefficient of determination of the models.

algorithm, linear regression, partial-fractional regression, and fractional regression, respectively and their values are equal to 0.92, 0.90, 0.84, 0.84, and 0.49 for training data and 0.89, 0.85, 0.75, 0.75, and 0.08 for testing data in the order given. Hence, the result of this part is in line with the other parts of this investigation, and the accuracy of the water cycle algorithm is considerably better than other models.

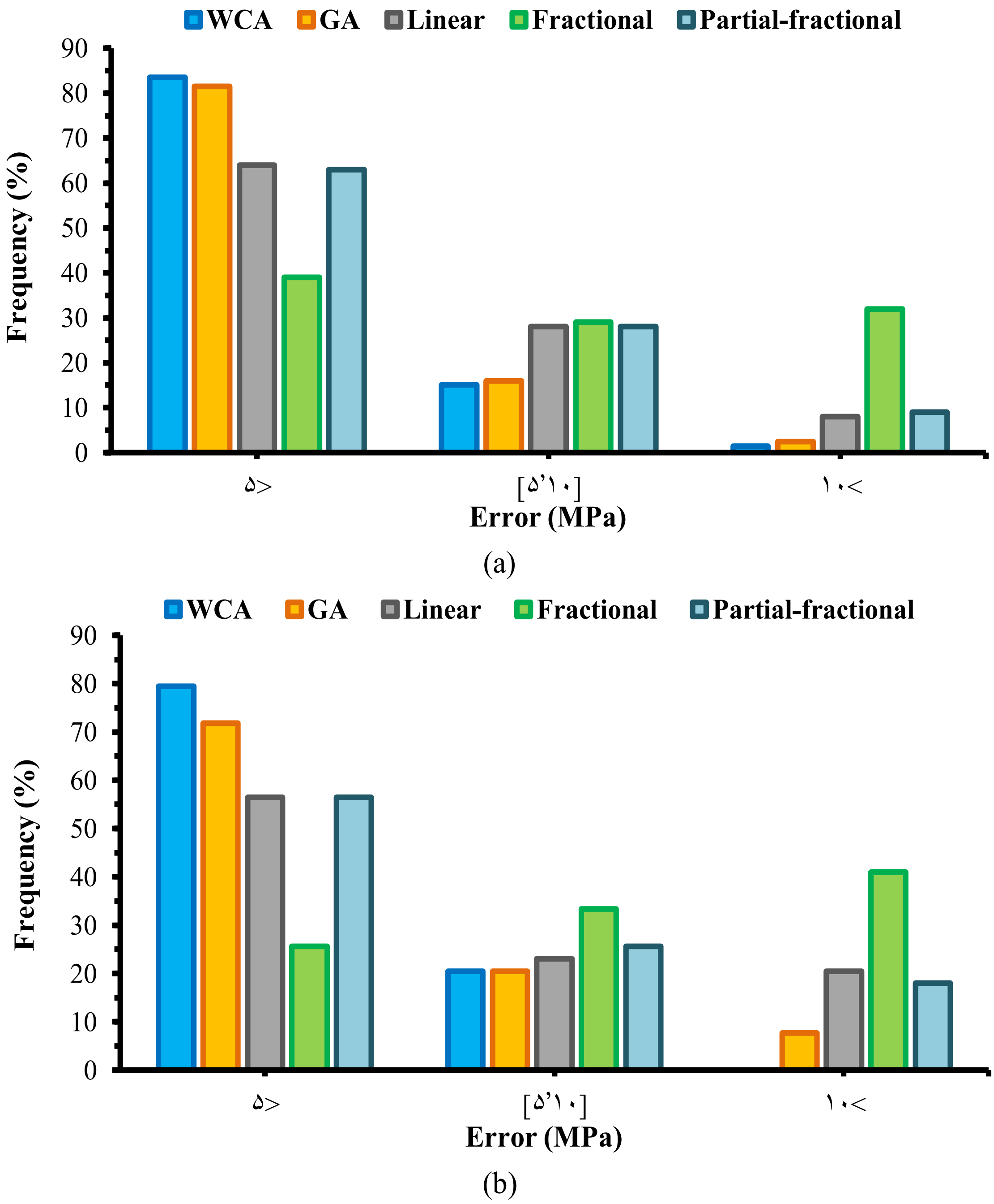
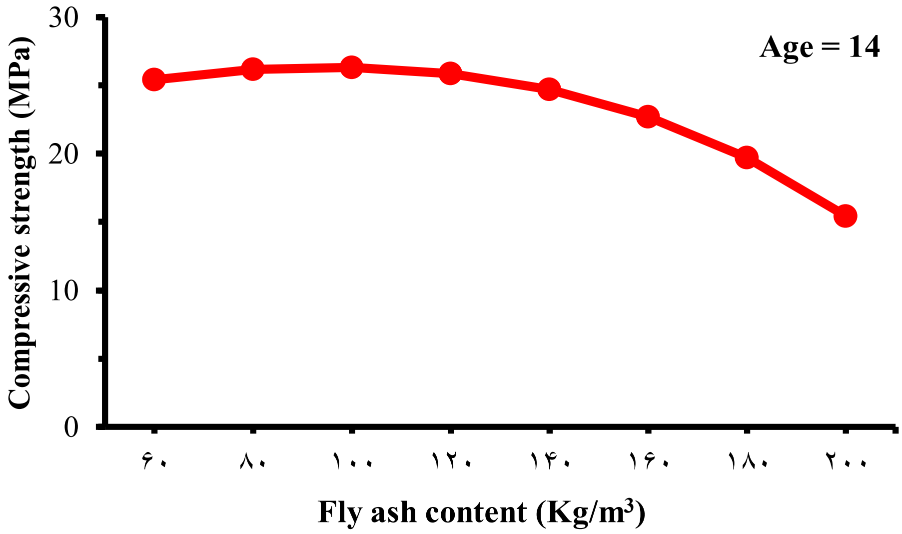


Fig. 7. Error histogram of the models, (a): training data, (b): testing data.

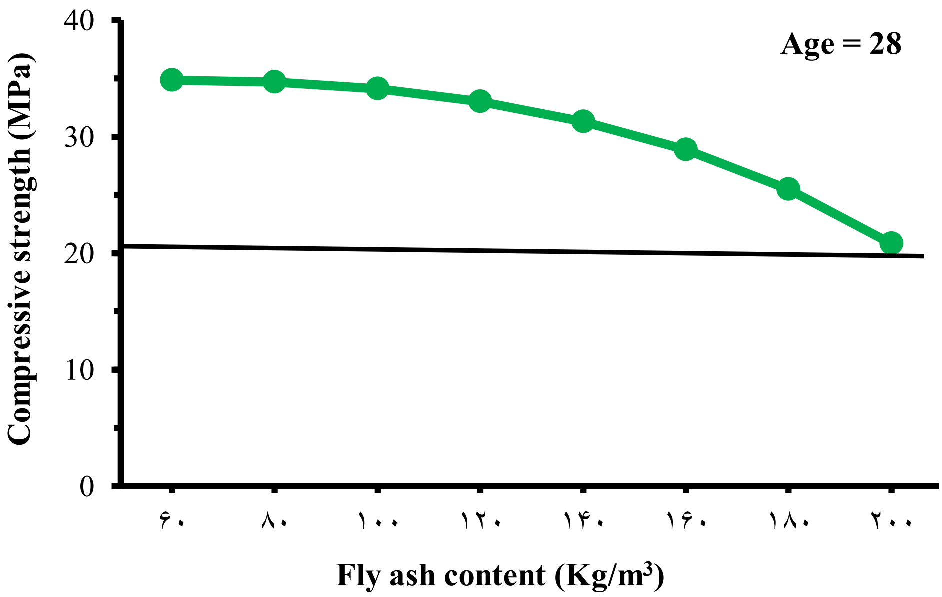
The error histogram of the models for testing and training data is presented in Fig.7. A more detailed look at this figure reveals that 83.5% and 79.5% of the compressive strength values for training and testing data predicted by the water cycle algorithm have an error of less than 5 MPa. Moreover, the errors of all testing data estimated by the water cycle algorithm are less than 10 MPa. In contrast, other models are not competent to predict all the training data with an error of less than 10 MPa. Drawing on PE5, the water cycle algorithm is the most trustworthy model for predicting the compressive strength of fly ash concrete followed by genetic algorithm, linear regression, partial-fractional regression, and fractional regression with the PE5 of 83.5, 81.5, 64, 63, and 39 for training data and 79.5, 71.8, 56.4, 56.4, and 25.6 respectively. Likewise, the water cycle performs better than other models according to the percentage of data whose errors are more than 10 MPa. The genetic algorithm, partial-fractional regression, linear regression, fractional regression are the next fruitful model based on having the error of less than 10 MPa and they include 7.7, 17.9, 20.5, and 41 percent testing data which their error are more than 10 MPa. The result of PE5 and error histogram are consistent with the results of other performance indicators, and it is conceived that the performance of the water cycle algorithm is by far better than other models, especially classical regressions. Furthermore, the genetic algorithm is outweighed the common regression models.

To analyze the effect of various amounts of fly ash on compressive strength, the average amount of materials used in samples are considered as the weights of materials in mixture proportion. Furthermore, the average amount of binder (355.8 kg/m^3) is taken into account as the binder's amount in mixture design. Therefore, by increasing the weight of fly ash, the weight of cement is reduced. Different amounts in the allowable range (between the minimum and maximum range in Table 1) are chosen as the amount of fly ash and the compressive strength in four different ages (14, 28, 56, and 90 days) are predicted. The water cycle algorithm is applied to estimate the compressive strength of fly ash concrete because it has the highest accuracy among all the models. The relation between the amount of fly ash and the compressive strength of fly ash concrete in different ages is shown in Fig.8.

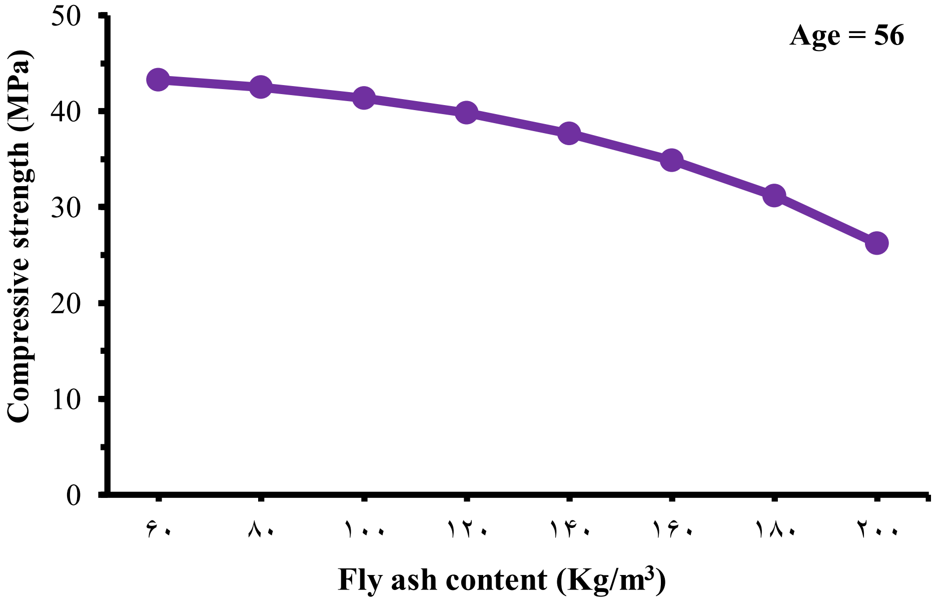
Based on the results depicted in Fig.8, if the average amount of materials were considered as their weight in mixture proportioning, the maximum amount of compressive strength would be relevant to the mixtures whose amount of fly ash is less than or equal to 120 kg/m^3. That is to say, using fly ash more than 35% of binder weight is not optimal, and it makes the compressive strength reduced. Nevertheless, according to ACI 318 [55], the minimum 28-day compressive strength of structural concretes is 2500 psi (17.23 MPa) for general elements and 3000 psi (20.68 MPa) for special moment



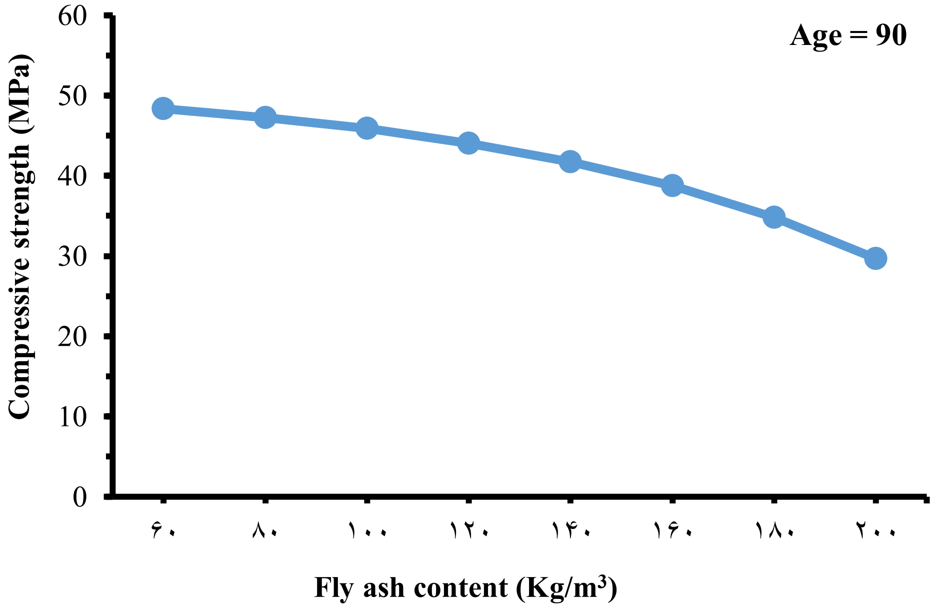
(a)



(b)



(c)



(d)

Fig. 8. The impacts of the amount of fly ash weight on compressive strength of concrete on the age of (a): 14, (b): 28, (c): 56, and (d): 90.

frames and special structural walls and all the amounts of fly ash in the allowable range can be used in mixture design because their 28-days compressive strength of all of them is more than 3000 psi.

Table 3 summarizes the recent works in concrete compressive strength prediction. According to this Table, the range of R for testing data varies from 0.73 to 0.98, with an average value of 0.88. Therefore, the coefficient of determination of the models presented in this study is more than 0.88, and it is acceptable. Furthermore, the mean absolute error and root mean square of WCA and GA are by far less than those of the other investigations. Hence, the accuracy of the model developed in this study is high, and the proposed models are trustable.

Table 3. Comparison of the obtained results with recent works.

Concrete type	Model	Training data			Testing data			Reference
		R	MAE	RMSE	R	MAE	RMSE	
Fly ash	WCA	0.96	3.01	3.76	0.94	2.90	3.88	This research
Fly ash	GA	0.95	3.12	4.29	0.92	3.44	5.01	This research
Glass cullet	Genetic programming I	0.81	4.78	5.78	0.81	7.27	9.42	[36]
Glass cullet	Genetic programming II	0.81	6.62	7.80	0.73	8.31	10.40	[36]
High performance	M5P model tree algorithm	0.95	3.92	5.48	0.94	4.48	6.17	[8]
Silica fume	Grey wolves ANN I	0.96	5.38	7.12	0.96	5.55	7.70	[10]
Silica fume	Grey wolves ANN II	0.99	2.38	3.43	0.98	3.47	5.43	[10]
Silica fume	biogeography	0.93	6.66	9.36	0.91	7.61	10.69	[31]

5- Summary and Conclusions

In this study, the impacts of fly ash concrete materials and the age of samples on their compressive strength are scrutinized, and a novel method for predicting the compressive strength of fly ash concrete is introduced. Water cycle algorithm and genetic algorithm are used to pursue the aims of this investigation, and their results are compared with the classical regression models, including linear regression, fractional regression, and partial-fractional regression. The obtained results can be summarized as follow:

- Heuristic and meta-heuristic algorithms are competent to create complicated nonlinear equations with high accuracy to find the relation of inputs and output of this model. Nonetheless, classical methods like regressions may not be trustable because of their predetermined structures.
- The mean absolute errors of the water cycle algorithm and genetic algorithm are 3.01 and 3.12 MPa for training data and 2.90 and 3.44 MPa for testing data in the order given. These errors are acceptable and can be overlooked. Conversely, the mean absolute errors of linear regression, fractional regression, and partial-fractional regression are 4.36, 7.90, and 4.34 for training data and 5.47, 9.70, and 5.37 MPa for testing data, respectively. Therefore, the water cycle algorithm is the only algorithm whose mean absolute error of testing data is less than that of training data.
- The water cycle algorithm outweighs other models based on PE5. According to PE5, the most accurate model is the water cycle algorithm, followed by genetic algorithm, linear regression, partial-fractional regression, and fractional regression. Besides, the water cycle algorithm is the only algorithm whose error of prediction for all testing data is less than 10 MPa.
- The outcomes of the coefficient of determination criteria and root mean square error validate that the water cycle algorithm is the most accurate model, and it is consistent with the results of other performance indicators.
- The influences of the age of samples can be taken into account in heuristic algorithm models, and the long-term features of concrete can be evaluated in advance.
- The maximum amounts of compressive strength are related to the mixtures which the amount of fly ash is less than 35% of the weight of the binder. Increasing the amount of fly ash more than this definite level reduces the compressive strength of concrete considerably.

References

[1] A.M.J.J.O.C.P. Rashad, An exploratory study on high-volume fly ash concrete incorporating silica fume subjected to thermal loads, 87 (2015) 735-744.

[2] R. Feiz, J. Ammenberg, L. Baas, M. Eklund, A. Helgstrand, R.J.J.O.C.P. Marshall, Improving the CO2 performance of cement, part I: utilizing life-cycle assessment and key performance indicators to assess development within the cement industry, 98 (2015) 272-281.

[3] S. Ghavami, B. Farahani, H. Jahanbakhsh, F.J.A.J.O.C.E. Moghadas Nejad, Effects of Silica Fume and Nano-silica on the Engineering Properties of Kaolinite Clay, 2(2) (2018) 135-142.

[4] F.M. Nejad, M. Habibi, P. Hosseini, H.J.J.o.c.p. Jahanbakhsh, Investigating the mechanical and fatigue properties of sustainable cement emulsified asphalt mortar, 156 (2017) 717-728.

[5] B. Lothenbach, K. Scrivener, R.J.C. Hooton, c. research, Supplementary cementitious materials, 41(12) (2011) 1244-1256.

[6] M.H. Shehata, M.D.J.C. Thomas, C. Research, The effect of fly ash composition on the expansion of concrete due to alkali-silica reaction, 30(7) (2000) 1063-1072.

[7] R.J.C. Siddique, C. Research, Performance characteristics of high-volume Class F fly ash concrete, 34(3) (2004) 487-493.

[8] A. Behnood, V. Behnood, M.M. Gharehveran, K.E.J.C. Alyamac, B. Materials, Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm, 142 (2017) 199-207.

[9] F. Deng, Y. He, S. Zhou, Y. Yu, H. Cheng, X.J.C. Wu, B. Materials, Compressive strength prediction of recycled concrete based on deep learning, 175 (2018) 562-569.

[10] A. Behnood, E.M.J.J.o.c.p. Golafshani, Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves, 202 (2018) 54-64.

[11] B. Vakhshouri, S.J.N. Nejadi, Prediction of compressive strength of self-compacting concrete by ANFIS models, 280 (2018) 13-22.

[12] Z.M. Yaseen, R.C. Deo, A. Hilal, A.M. Abd, L.C. Bueno, S. Salcedo-Sanz, M.L.J.A.i.E.S. Nehdi, Predicting compressive strength of lightweight foamed concrete using extreme learning machine model, 115 (2018) 112-125.

[13] H.J.I.J.O.I. Naseri, Management, Technology, Cost Optimization of No-Slump Concrete Using Genetic Algorithm and Particle Swarm Optimization, 10(1) (2019).

[14] J.-S. Chou, C.-F.J.A.i.C. Tsai, Concrete compressive strength analysis using a combined classification and regression technique, 24 (2012) 52-60.

[15] D.-K. Bui, T. Nguyen, J.-S. Chou, H. Nguyen-Xuan, T.D.J.C. Ngo, B. Materials, A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete, 180 (2018) 320-333.

[16] H. Naderpour, A.H. Rafiean, P.J.J.O.B.E. Fakharian, Compressive strength prediction of environmentally friendly concrete using artificial neural networks, 16 (2018) 213-219.

[17] T. Kim, J.M. Davis, M.T. Ley, S. Kang, P.J.C. Amrollahi, B. Materials, Fly ash particle characterization for predicting concrete compressive strength, 165 (2018) 560-571.

[18] N. Rajamane, J.A. Peter, P.J.C. Ambily, c. composites, Prediction of compressive strength of concrete with fly ash as sand replacement material, 29(3) (2007) 218-223.

[19] I.B. Topcu, M.J.C.M.S. Sarıdemir, Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic, 41(3) (2008) 305-311.

[20] M.M. Khotbehsara, B.M. Miyandehi, F. Naseri, T. Ozbakkaloglu, F. Jafari, E.J.C. Mohseni, B. Materials, Effect of SnO2, ZrO2, and CaCO3 nanoparticles on water transport and durability properties of self-compacting mortar containing fly ash: Experimental observations and ANFIS predictions, 158 (2018) 823-834.

[21] I.-C.J.C. Yeh, C. research, Modeling of strength of high-performance concrete using artificial neural networks, 28(12) (1998) 1797-1808.

[22] I.-C.J.J.C.I.C.H.E. Yeh, Prediction of strength of fly ash and slag concrete by the use of artificial neural networks, 15(4) (2003) 659-663.

[23] J. Sobhani, M. Najimi, A.R. Pourkhorshidi, T.J.C. Parhizkar, B. Materials, Prediction of the compressive strength of no-slump concrete: A comparative study of regression, neural network and ANFIS models, 24(5) (2010) 709-718.

[24] J.R. Sampson, Adaptation in natural and artificial systems (John H. Holland), in, Society for Industrial and Applied Mathematics, 1976.

[25] H. Naseri, M.A.E.J.I.J.o.I. Ghasbeh, Management, Technology, Time-Cost Trade off to Compensate Delay of Project Using Genetic Algorithm and Linear Programming, 9(6) (2018).

[26] D. Neeraja, T. Kamireddy, P. Santosh Kumar, V. Simha Reddy, Weight optimization of plane truss using genetic algorithm, in: Materials Science and Engineering Conference Series, 2017, pp. 032015.

[27] Z. Aydın, Y.J.K.J.O.C.E. Ayvaz, Overall cost optimization of prestressed concrete bridge using genetic algorithm, 17(4) (2013) 769-776.

[28] H. Eskandar, A. Sadollah, A. Bahreininejad, M.J.C. Hamdi, Structures, Water cycle algorithm–A novel metaheuristic optimization method for solving constrained engineering optimization problems, 110 (2012) 151-166.

[29] A. Sadollah, H. Eskandar, A. Bahreininejad, J.H.J.S.C. Kim, Water cycle algorithm for solving multi-objective optimization problems, 19(9) (2015) 2587-2603.

[30] A. Sadollah, H. Eskandar, J.H.J.A.S.C. Kim, Water cycle algorithm for solving constrained multi-objective optimization problems, 27 (2015) 279-298.

[31] E.M. Golafshani, A.J.C. Behnood, C. Composites, Estimating the optimal mix design of silica fume concrete using biogeography-based programming, 96 (2019) 95-105.

[32] E.A. Ramalho, J.J. Ramalho, P.D.J.J.o.P.A. Henriques, Fractional regression models for second stage DEA efficiency analyses, 34(3) (2010) 239-255.

[33] E.A. Ramalho, J.J. Ramalho, J.M.J.J.o.E.S. Murteira, Alternative estimating and testing empirical strategies for fractional regression models, 25(1) (2011) 19-68.

[34] M. Mokhtari, s. Abedian, S.A. Almasi, Rockfall Susceptibility Mapping Using Artificial Neural Network, Frequency Ratio, and Logistic Regression: A Case Study in Central Iran, Taft County %J AUT Journal of Civil Engineering, (2019) -.

[35] A. Committee, I.O.f. Standardization, Building code requirements for structural concrete (ACI 318-08) and commentary, in, American Concrete Institute, 2008.

[36] M. Mirzahosseini, P. Jiao, K. Barri, K.A. Riding, A.H.J.E.C. Alavi, New machine learning prediction models for compressive strength of concrete modified with glass cullet, 36(3) (2019) 876-898.

HOW TO CITE THIS ARTICLE

H. Naseri, H. Jahanbakhsh, F. Moghadas Nejad, A. Golroo, Developing a Novel Machine Learning Method to Predict the Compressive Strength of Fly Ash Concrete in Different Ages, AUT J. Civil Eng., 4(4) (2020) 423-436.

DOI: 10.22060/ajce.2019.16124.5569