Journal of Soft Computing in Civil Engineering 3-2 (2019) 43-53

Contents lists available at **SCCE**

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Modelling of Concrete Compressive Strength Admixed with GGBFS Using Gene Expression Programming

O.O. Akin[1], O.S. Abejide[1*]

1. Postgraduate Student, Department of Civil Engineering, Ahmadu Bello University, Zaria, Nigeria
2. Professor, Department of Civil Engineering, Ahmadu Bello University, Zaria, Nigeria

Corresponding author: *abejideos@yahoo.com*

https://doi.org/10.22115/SCCE.2019.178214.1103

ARTICLE INFO

Article history:
Received: 09 April 2019
Revised: 11 July 2019
Accepted: 04 August 2019

Keywords:
Concrete;
Strength;
GGBFS;
GEP.

ABSTRACT

Several studies have established that strength development in concrete is not only determined by the water/binder ratio, but it is also affected by the presence of other ingredients. With the increase in the number of concrete ingredients from the conventional four materials by addition of various types of admixtures (agricultural wastes, chemical, mineral and biological) to achieve a desired property, modelling its behavior has become more complex and challenging. Presented in this work is the possibility of adopting the Gene Expression Programming (GEP) algorithm to predict the compressive strength of concrete admixed with Ground Granulated Blast Furnace Slag (GGBFS) as Supplementary Cementitious Materials (SCMs). A set of data with satisfactory experimental results were obtained from literatures for the study. Result from the GEP algorithm was compared with that from stepwise regression analysis in order to appreciate the accuracy of GEP algorithm as compared to other data analysis program. With R-Square value and MSE of -0.94 and 5.15 respectively, The GEP algorithm proves to be more accurate in the modelling of concrete compressive strength.

1. Introduction

There has been a huge rise in the production of Portland cement over time, Just in the year 2011, production was reported to be about 3.6 billion tonnes [1]. A major problem associated with the

How to cite this article: Akin OO, Abejide OS. Modelling of concrete compressive strength admixed with GGBFS using gene expression programming. J Soft Comput Civ Eng 2019;3(2):43–53. https://doi.org/10.22115/scce.2019.178214.1103.

production of Portland cement of is the emission of carbon dioxide (CO_2) during the processes. It is noted that for every one tonne of Portland cement clinker produced, there is an approximate release of one tonne of CO_2 into the atmosphere [2]. This fact is thought provoking when the huge tonnes of cement produced annually worldwide is considered. As a result of the rising need for cement, engineers and researchers are searching for possible means to reduce the quantity of Portland cement needed for concrete production.

The huge quantity of natural materials expended during the production of concrete, has necessitated the search for new solutions for sustainable concrete production and development of infrastructures. A major means to reduce the effects on the environment is the use of admixtures such as GGBFS, rice husk ash, fly ash and natural pozzolans as supplementary cementitious materials. The application of these materials in production of concrete increases the compressive strength and pore structure of the mortars.[3,4]. The use of these alternative materials have the ability to accomplish substantial reductions in the embodied energy and greenhouse gas emissions inherent in Portland cement production. And by extension improve the overall sustainability of concrete. In the future, the use of alternative materials to Portland cement will only increase, so there is need to come up with appropriate means to model its properties so as to fully understand its behavior under different conditions.[5].

Since concrete is expected to show resistance under austere conditions such as acidic attack, its properties need to be improved. Other than just resistance against extreme conditions, concrete is expected to always exhibit good workability, strength and durability. As a result of the advancement in technology, concrete that meets these requirements can now be produced. However, there seem not to be a clearly defined method in which concrete mix can be optimized according to the required properties. Only a few attempts have so far been made at that problem. The major reason behind this is the possibility of different mix proportions and the way to optimize the problem under different variables and properties (designated by single or multi objective functions) mathematically appears quite challenging.[6]. Therefore this paper presents the use of GEP (a soft computing technique) as compared to the conventional regression analysis in the modeling of concrete compressive strength when admixed with GGBFS.

Soft computing techniques have been suggested and reported in some studies to have faster processing time to achieve a better results.[7,8] In this work GEP and SPSS would be used and their results compared. With the growing numbers of soft computing techniques, there is not a superior algorithm, because all have their merit and demerits. The problem type is what determines which technique is appropriate. [9]

2. Background of study

2.1. Use of GGBFS as supplementary cementitious material (SCM)

GGBFS, also referred to as slag cement, is produced from blast iron, it is a non-metallic hydraulic cement comprising basically aluminosilicates of calcium that is established in a molten state simultaneously with iron in a blast furnace. The molten slag at high temperature of about 1500^0C, is cooled quickly in water to form a granular material that is sand-like. The specific gravity of GGBFS is found to be in the range of 2.85 to 2.95. In the presence of water and an alkali activator CaOH or NaOH which is gotten from Portland cement, granulated slag undergoes hydration and sets just like Portland cement. [10].

The use of GGBFS as a SCM in concrete has a number of benefits, some of which are: improved workability and durability and also economic benefits.[11]. One of the noticeable improvements when slag is introduced in concrete, is the compressive strength which is as a result of the very fine state of the GGBFS and the hydration process.[12].

2.2. Overview of gene expression programming (GEP)

Gene Expression Programming (GEP) is a computational algorithm that generate computer programs or models. These programs are complex tree structures that is trained by changing their sizes, shapes, and composition, much like a living organism. GEP was first developed by Ferreira [13], with the assumption of it being an hybrid of both genetic algorithm(GA) and genetic programming(GP) [14], [15]. GEP is an advance data analysis algorithm which has been used widely across many discipline. The shortcomings of other data analysis tools are addressed by the use of gene expression programming,[16].

GEP algorithm make use of linear chromosomes character that are made up of genes that are organized structurally in a head and a tail. These chromosomes are programmed to act as a genome and can be modified by means of many phenomena such as transmutation, transposition, root transposition, gene transposition, gene recombination, and one- and two-point recombination. The chromosomes encode expression trees, which are the object of selection. [10].

The creation of the genetic varieties in the GEP algorithm is quite easy simply because of the genetic mechanism of this technique at the chromosome level. More so, due to the multigenetic nature of GEP, complex programmes and nonlinear programmes can be developed with various subprograms [17]. GEP algorithm make use of a fixed length of character strings to denote the problem solutions, they are eventually expressed as parse-trees referred to as ‘‘expression tree” in GEP of various shapes and sizes during evaluation of fitness [18]. The fixed length of the GEP is usually predefined for a any problem. So, what changes in GEP is not the genes length, but rather the size of the resulting Expression Trees (ETs). [19].

2.2.1. GEP genes and expression tree

The structures of GEP genes is best understood in form of Open Reading Frames (ORFs). Each GEP gene is made up of a set of symbols with a fixed length that can be any element from a list of functions like +,log, *,tan, -, /, √ and the variable or terminal set like (x_1, x_2, x_3…x_n). An example of GEP gene with the selected function and terminal is given in equation (1):

$$+, -, sin, x_1, x_2, +, c_1, x_1 \tag{1}$$

Where x_1 *and* x_2 are variables and C_1 a constant say 3. Equation (1) is referred to as Karva-notation or K-expression. This expression is may be diagramed into an expression tree (ET) through a width-first fashion. The sample gene in Equation (1) is shown in Figure 1. The conversion begins from the first point in the Karva expression, which is corresponding to the root of the expression tree, and it is read from the string one after the other. The GEP gene in Equation (1) can also be written mathematically as:

$$(x_1 - x_2) + \sin(3 + x_1) \tag{2}$$

ETs can be converted inversely into a K-expression by accurately recording from left to right nodes in each layer of a particular gene in the ET, from root layer down to the deepest one to form the string. Figure 1 shows the gene expression tree from which Equation (2) is decoded.

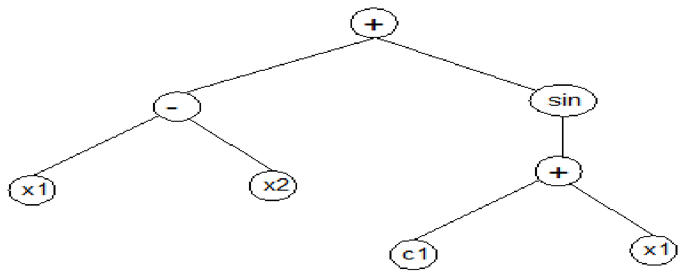

Fig. 1 Expression Tree (ET).

Figure 2 gives the schematic representation of the GEP algorithm. The algorithm begins by randomly creating the chromosome with the fixed length of every evolving individuals. Afterwards, the chromosomes are confirmed, and the fitness of every individual is assessed. Next, the individuals are selected based on their fitness level in order to apply the reproduction. This process is repeated over and over with the new individual for a many of generations until the best solution is found.

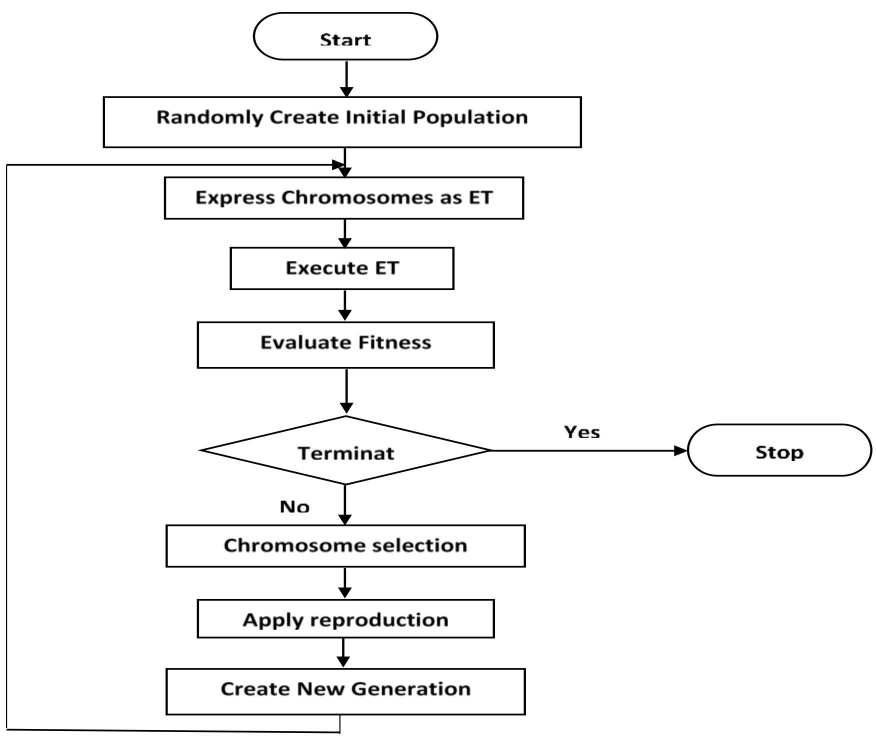

Fig. 2. Schematic representation of GEP Algorithm.

3. Methodology

3.1. Experimental dataset

The dataset used for the formulation are gathered from work by the following authors, Siddique and Kaur [20] and Oner and Akyuz [21]. The dataset was critically analysed and the reported experimental procedures were fully considered. The summary of the dataset is given in Table 1.

Table 1
Statistical view of concrete mix with GGBFS as SCM.

Parameters	Minimum	Maximum	Mean	Standard Deviation
Cement (kg/m^3)	175.00	450.00	248.0556	65.3155
Fly ash (kg/m^3)	0.00	440.00	146.8056	129.5123
Fine Aggregate (kg/m^3)	477.00	768.00	640.2500	88.9656
Coarse aggregate (kg/m^3)	723.00	1166.00	1005.6944	104.6529
Water (kg/m^3)	203.00	295.00	234.0556	22.5755
Strength (N/mm^2)	13.00	48.40	31.9197	8.8058

3.2. Model construction using gene expression programming

The dataset described in Tables 1 is used for the modelling compressive strength of concrete. What is necessary here is for the linking functions between the input variables x_1, x_2...x_5 and the output or target value y to be clearly defined. Equation (3) gives the typical expression of the function:

$$y_i = f(x_1, x_2, \ldots, x_5)\forall_i \tag{3}$$

The models are developed for the 28day compressive strength of concrete with GGBFS as SCM. The variables for the modelling is given in Table 2.

Table 2
Design variables for the concrete mix.

Input Variables	Code	Output Variable	Code
Portland cement	x_1		
GGBFS	x_2		
Fine Aggregate	x_3	28day Compressive strength	y
Coarse aggregate	x_4		
Water	x_5		

The R-square (R^2), Mean Square Error (MSE) and Root Mean Squared Error (RMSE) are the results and the statistical criteria for evaluating the performance of the model obtained from the GEP algorithm. These criteria are defined in Equations (4), (5) and (6)

$$R^2 = 1 - \left(\frac{\sum_i (t_i - o_i)^2}{\sum_i (o_i)^2}\right) \quad (4)$$

$$MSE = \frac{1}{n}\sum (t_i - o_i)^2 \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum (t_i - o_i)^2} \quad (6)$$

Where t = target

o = output

n = number of dataset

For an ideal or perfect fit $t_i = o_i$, and $MSE_i = 0$. Therefore, the range of MSE_i index is from 0 to infinity, with the value of 0 representing idea and absolute prediction. That is, the lower the MSE value the better the model.

The fitness f_i of an individual program according to Sarıdemir [22], is given by Equation(7)

$$f_i = \sum_{j=1}^{c_t} \left(M - \left|C_{ij} - T_j\right|\right) \quad (7)$$

Where M = range of selection

C_{ij} = value returned by the individual chromosome i for fitness case j

T_j = target value for fitness case

The results obtained from literature are compared with the results derived from the GEP-based model and the regression-based formulation from Statistical Package for Social Science (SPSS). The R^2 values are tabulated in the results section. The GEP algorithm parameters used in this analysis are given in Table 3

Table 3
The GEP algorithm Parameters.

Parameter definition	Values/designations
Number of generation	100000–500000
Number of chromosomes	30
Function set	+, -, *, /, x^2, sin, cos, arc tan, e
Number of genes	3
Head size	8
Linking function	+
Mutation rate	0.00138
One-point recombination rate	0.00277
Two-point recombination rate	0.00277
Gene recombination rate	0.00277
Lower bound	-10
Upper bound	10

4. Results and discussion

4.1. Results of prediction models

The equation for the prediction of the compressive strength of concrete having GGBFS as SCM is generated using two methods namely: Gene Expression Programming and the stepwise regression analysis using SPSS.

(a) Models generated using the GEP algorithm

Running the GEP algorithm for the concrete mix dataset needs huge computational time depending of the computer processing ability.

The expression tree result from which the model was formulated is given in Figure 3.

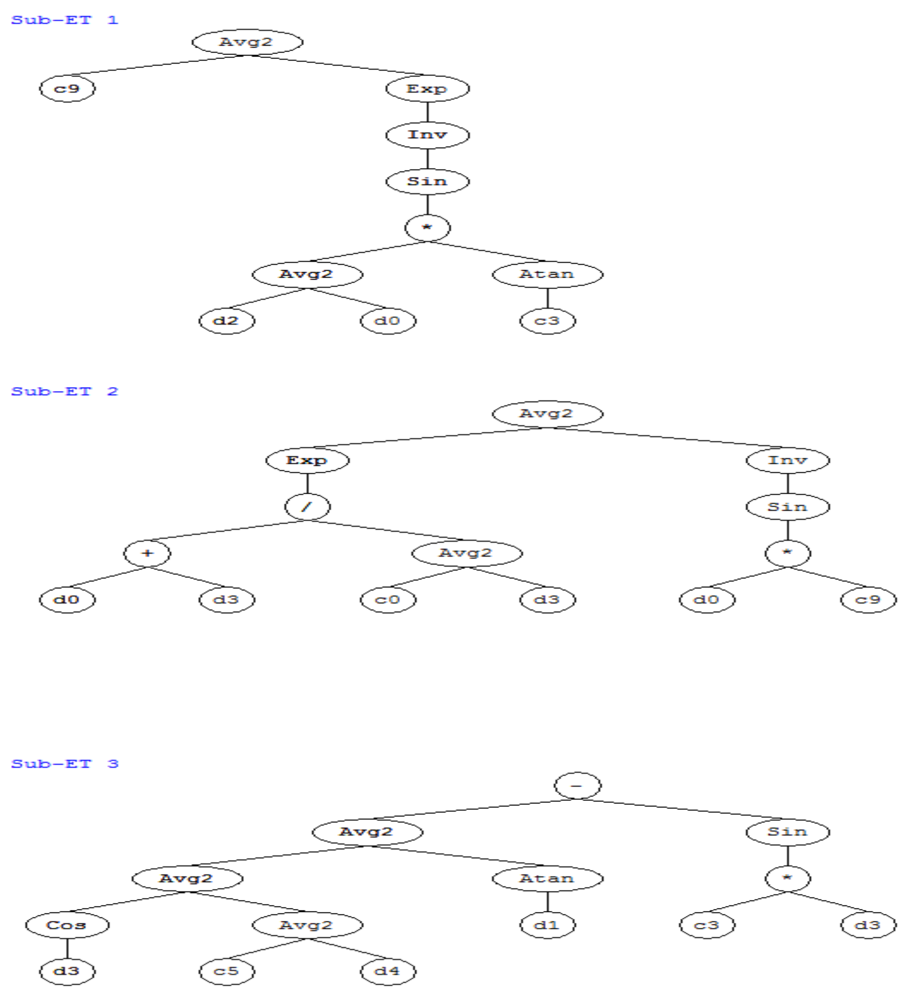

Fig. 3. Gene Expression Tree.

From the definition of the design variables and output in the previous section, the algorithm for the GEP was run and the best fitness function is given in equation (8).

$$y = \frac{1}{2}\left(-6.521 + e^{\frac{1}{\sin\left(\left(\frac{x_3+x_2}{2}\right)(\tan^{-1} 1.425)\right)}}\right) + \frac{1}{2}\left(\left(e^{\frac{2(x_1+x_4)}{-85.536+x_4}}\right) + \left(\frac{1}{\sin(-3.716x_1)}\right)\right) + \left(\cos x_4 + \frac{x_5-10.34}{2} + \frac{\tan^{-1} x_2}{2} - \sin(-8.475x_4)\right) \tag{8}$$

The test results for the model is presented in figure 4 and 5.

Fig. 4. Test result for GEP model (curve fitting).

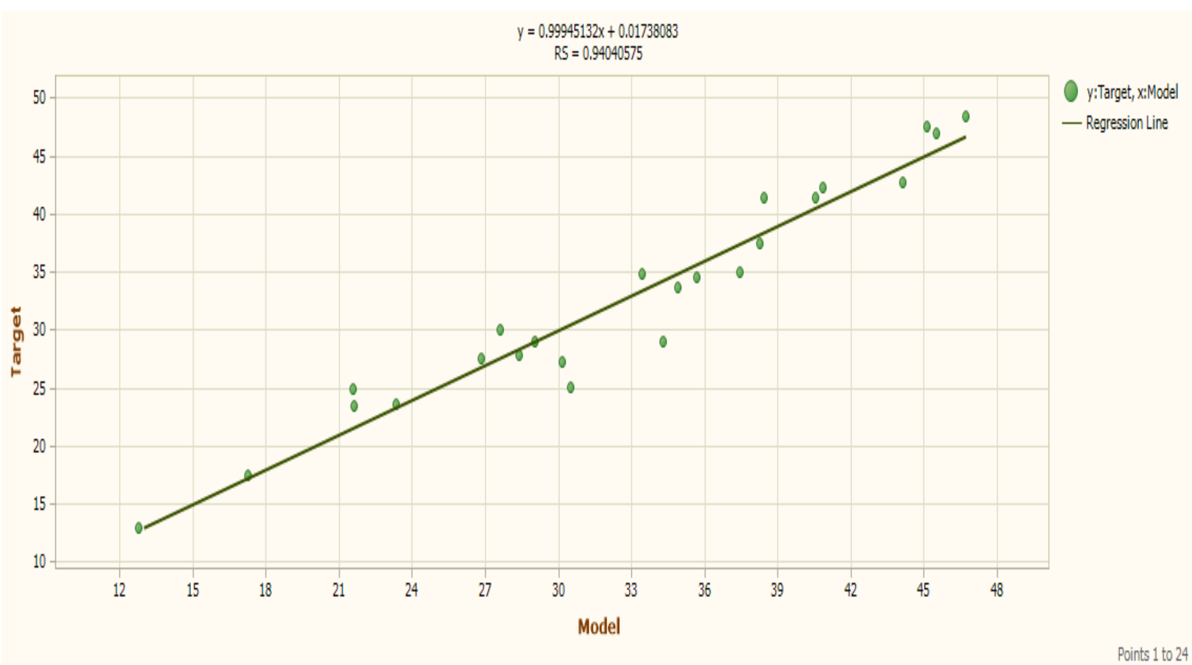

Fig. 5. Test result for GEP model (scatter plot).

Table 4
Performance Metrics for the GEP Algorithm Model.

Performance Metric	Values
R-square	0.940406
MSE	5.150249
RMSE	2.269416
MAE	1.777991
RSE	5.96E-02

(b) Regression analysis models

To appreciate the use of GEP algorithm for prediction, a linear equation was formulated using the classical statistical software package SPSS. The result is give in equation (9).

$$y = -757.4 + 0.646x_1 + 0.610x_2 + 0.398x_3 + 0.367x_4 - 0.359x_5 \quad (9)$$

Table 5
Comparison of Model Performance Metrics.

Performance Metric	Models	
	GEP	SPSS
R Square	0.94	0.91

4.2. Discussion of results

It can be seen from Figures 3 and 4 that there exist a close fit between the target and the model curves. The function from the GEP algorithm is able to closely follow the pattern of the actual data with an accurate correlation.

The R^2 and MSE values are 0.94 and 5.15 respectively, showing a reasonably good fit of the model. The R^2 value is 0.14 more than the suggested good fit. (that is R^2 more than or equal to 0.8) as proposed by Chopra,Sharma [23]. Table 4 gives the full performance metrics for the GEP model.

The R^2 value for Equations (9) is 0.91, as compared to 0.94 obtained from the GEP Model. With these values, the GEP algorithm can be appreciated as a more approximate tool for modelling concrete compressive strength

As it can be seen from the models, the GEP model is highly nonlinear, therefore it is quite difficult to evaluate by conventional techniques. The model based on the regression analysis is a linear function and relatively easier to solve. From the statistical details it is obvious that the model from the GEP algorithm is more accurate for the prediction of concrete compressive strength.

It can be seen from Table 5 that the GEP model result appear very close to the target value as given by the performance metric as compared to the regression analysis based model. This is to clearly show that the predictive ability of GEP is more approximate and accurate than the classical statistical regression analysis.

5. Conclusion

Gene expression programming has been used in this study to model the compressive strength of concrete. The following conclusions are drawn from the study:

Mathematical equations have been derived for the prediction of compressive strength of concrete, this is done using the GEP algorithm which is a major setback and disadvantage in artificial neural network (ANN). Although, ANN is a powerful predictive tool, it lacks the ability to express the relationships between the independent variables and the response using a mathematical equation as seen in Gene Expression Programming.

With R^2 value of 0.94 from the GEP model, the GEP algorithm has shown to be a good prediction program for modelling the compressive strength of concrete.

On the comparative study of the prediction models, GEP algorithm gave a more accurate model as compared with the regression analysis from the SPSS. The equation from the GEP algorithm appears to be complex than the simple linear function of the stepwise regression analysis.

This is as a result of the fact that the relationship that exist between concrete compressive strength and the constituents is nonlinear rather than linear, so it is best to represent the numerical modelling nonlinearly.

Also, the GEP algorithm gave the resulting model in various programming languages. This has made it easier for easy usage and analysis on other programming tools, especially for optimization on any optimization tool.

References

[1] Armstrong T. The cement industry in figures. An overview of global cement sector trends. Int Cem Rev 2012.

[2] Greer, W.L., J.G H, T.B C. Air emission and control measures Chapter 6.1 in Innovation in portland cement manufaturing. Portland cement assosiation. Portland cement assosiation 2004.

[3] Çakır Ö, Aköz F. Effect of curing conditions on the mortars with and without GGBFS. Constr Build Mater 2008;22:308–14. doi:10.1016/j.conbuildmat.2006.08.013.

[4] Bágel L. Strength and pore structure of ternary blended cement mortars containing blast furnace slag and silica fume. Cem Concr Res 1998;28:1011–22. doi:10.1016/S0008-8846(98)00078-7.

[5] ACI-CommitteeE-701, Cementitious Materials for Concrete. America Concrete Institute., 2003.

[6] Maruyama I, Kanematsu M, Noguchi T, Tomosawa F. Optimization of mix proportion of concrete under various severe conditions by applying the genetic algorithm. 3rd (CONSEC'01) Vancouver 2004.

[7] Naderpour H, Nagai K, Fakharian P, Haji M. Innovative models for prediction of compressive strength of FRP-confined circular reinforced concrete columns using soft computing methods. Compos Struct 2019;215:69–84. doi:10.1016/j.compstruct.2019.02.048.

[8] Hosseini G. Capacity Prediction of RC Beams Strengthened with FRP by Artificial Neural

Networks Based on Genetic Algorithm. J Soft Comput Civ Eng 2017;1:93–8.

[9] Ebtehaj I, Bonakdari H. No-deposition sediment transport in sewers using gene expression programming. J Soft Comput Civ Eng 2017;1:29–53.

[10] Kosmatka SH, Kerkhoff B, Panarese WC. Design and control of concrete mixtures. vol. 5420. Portland Cement Association Skokie, IL; 2002.

[11] Bilim C, Atiş CD, Tanyildizi H, Karahan O. Predicting the compressive strength of ground granulated blast furnace slag concrete using artificial neural network. Adv Eng Softw 2009;40:334–40. doi:10.1016/j.advengsoft.2008.05.005.

[12] Chidiac SE, Panesar DK. Evolution of mechanical properties of concrete containing ground granulated blast furnace slag and effects on the scaling resistance test at 28days. Cem Concr Compos 2008;30:63–71. doi:10.1016/j.cemconcomp.2007.09.003.

[13] Ferreira C. Gene expression programming: a new adaptive algorithm for solving problems. ArXiv Prepr Cs/0102027 2001.

[14] Koza JR, Koza JR. Genetic programming: on the programming of computers by means of natural selection. vol. 1. MIT press; 1992.

[15] Kusy M, Obrzut B, Kluska J. Application of gene expression programming and neural networks to predict adverse events of radical hysterectomy in cervical cancer patients. Med Biol Eng Comput 2013;51:1357–65. doi:10.1007/s11517-013-1108-8.

[16] Deng S, Yuan C, Yang J, Zhou A. Distributed Mining for Content Filtering Function Based on Simulated Annealing and Gene Expression Programming in Active Distribution Network. IEEE Access 2017;5:2319–28. doi:10.1109/ACCESS.2017.2669106.

[17] Gholampour A, Gandomi AH, Ozbakkaloglu T. New formulations for mechanical properties of recycled aggregate concrete using gene expression programming. Constr Build Mater 2017;130:122–45. doi:10.1016/j.conbuildmat.2016.10.114.

[18] Baykasoğlu A, Öztaş A, Özbay E. Prediction and multi-objective optimization of high-strength concrete parameters via soft computing approaches. Expert Syst Appl 2009;36:6145–55. doi:10.1016/j.eswa.2008.07.017.

[19] Gandomi AH, Yang X-S, Talatahari S, Alavi AH. Metaheuristic algorithms in modeling and optimization. Metaheuristic Appl Struct Infrastructures 2013:1–24.

[20] Siddique R, Kaur D. Properties of concrete containing ground granulated blast furnace slag (GGBFS) at elevated temperatures. J Adv Res 2012;3:45–51. doi:10.1016/j.jare.2011.03.004.

[21] Oner A, Akyuz S. An experimental study on optimum usage of GGBS for the compressive strength of concrete. Cem Concr Compos 2007;29:505–14. doi:10.1016/j.cemconcomp.2007.01.001.

[22] Sarıdemir M. Genetic programming approach for prediction of compressive strength of concretes containing rice husk ash. Constr Build Mater 2010;24:1911–9. doi:10.1016/j.conbuildmat.2010.04.011.

[23] Chopra P, Sharma RK, Kumar M. Prediction of compressive strength of concrete using artificial neural network and genetic programming. Adv Mater Sci Eng 2016;2016.