

Research Article

Hua Si, Daoming Shen*, Muhammad Nasir Amin*, Siyab Ul Arifeen, Muhammad Tahir Qadir, and Kaffayatullah Khan

Producing sustainable binding materials using marble waste blended with fly ash and rice husk ash for building materials

<https://doi.org/10.1515/rams-2024-0049>

received April 08, 2024; accepted July 11, 2024

Abstract: This study explores the possibilities of a new binding material, *i.e.*, marble cement (MC) made from recycled marble. It will assess how well it performs when mixed with ash from rice husks and fly ash. This research analyzes flexural strength of marble cement mortar (FR-MCM), a mortar that incorporates MC, fly ash, and rice husk ash. A set of machine learning models capable of predicting CS and FS (flexural and compressive strengths) were developed. Gene expression programming (GEP) and multi-expression programming (MEP) are crucial in creating these types of models. Statistics, Taylor's diagrams, R^2 values, and comparisons of experimental and theoretical results were used to evaluate the models. Stress testing also showed how different input features affected the model's outputs. The accuracy of all GEP models was shown to fall within the acceptable range ($R^2 = 0.952$ for CS and $R^2 = 0.920$ for FS), and all MEP prediction models were determined to be exceptionally accurate ($R^2 = 0.970$ for CS and $R^2 = 0.935$ for FS). The statistical testing for error validation also verified that MEP models were more accurate than GEP models. According to sensitivity analysis, curing age and rice husk ash exerted the most significant influence on the prediction of CS and FS, followed by fly ash and MC.

Keywords: marble cement, compressive strength, sustainable binder

1 Introduction

One to eight percent of the world's greenhouse gas emissions arise from making regular Portland cement, widely known as OPC, and it is predicted that this figure will rise by 8% by 2050 [1,2], which casts questions on the practicality of achieving the zero-emissions objective. If the OPC sector is serious about reducing its CO₂ emissions, it must find less harmful alternatives to OPC [3,4]. On the flip side, marble is a material that is gaining popularity all over the globe. Throughout the globe, people are enthused about the marble that is produced. There are about 300 billion tons of marble in the sphere. The quarrying process involves blasting, which results in the waste of over 50% of the marble that is mined [5,6]. The deficiency of a central system for waste disposal means that the quarries continue to collect trash from all over. Marble comes in huge chunks from quarries and is processed into tiles and other valuable stones. Rough blocks should not be refined and trimmed. The exact proportion varies with each processing method, but 20% of these blocks are finely powdered [7]. Companies usually dump waste in open locations. Drying into a fine powder can cause allergic reactions to the skin, cancer of the lungs, and discomfort in the eye. Sludge increases water contamination [8]. Tiny marble particles on plant shrubberies and stems cause dehydration in older hedges and trees, highlighting a drawback of plant diversity [9,10].

The composition of OPC is 60–65% CaO, 20–25% SiO₂, and 4–8% Al₂O₃, rendering to Neville and Brooks [11]. According to the study, limestone is the primary component of cement due to its calcium oxide content. Limestone and water-soluble Portland cement (WMP) have several molecular components, one of which is the excessive content of CaO [12,13]. Marble cement (MC) was made by

* **Corresponding author: Daoming Shen**, School of Civil Engineering and Architecture, Xinxiang University, Xinxiang, Henan, 453000, China, e-mail: shendaoming@163.com

* **Corresponding author: Muhammad Nasir Amin**, Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa, 31982, Saudi Arabia, e-mail: mgadir@kfu.edu.sa

Hua Si: School of Civil Engineering and Architecture, Xinxiang University, Xinxiang, Henan, 453000, China

Siyab Ul Arifeen: Department of Civil Engineering, COMSATS University Islamabad, Abbottabad, 22060, Pakistan

Muhammad Tahir Qadir, Kaffayatullah Khan: Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa, 31982, Saudi Arabia

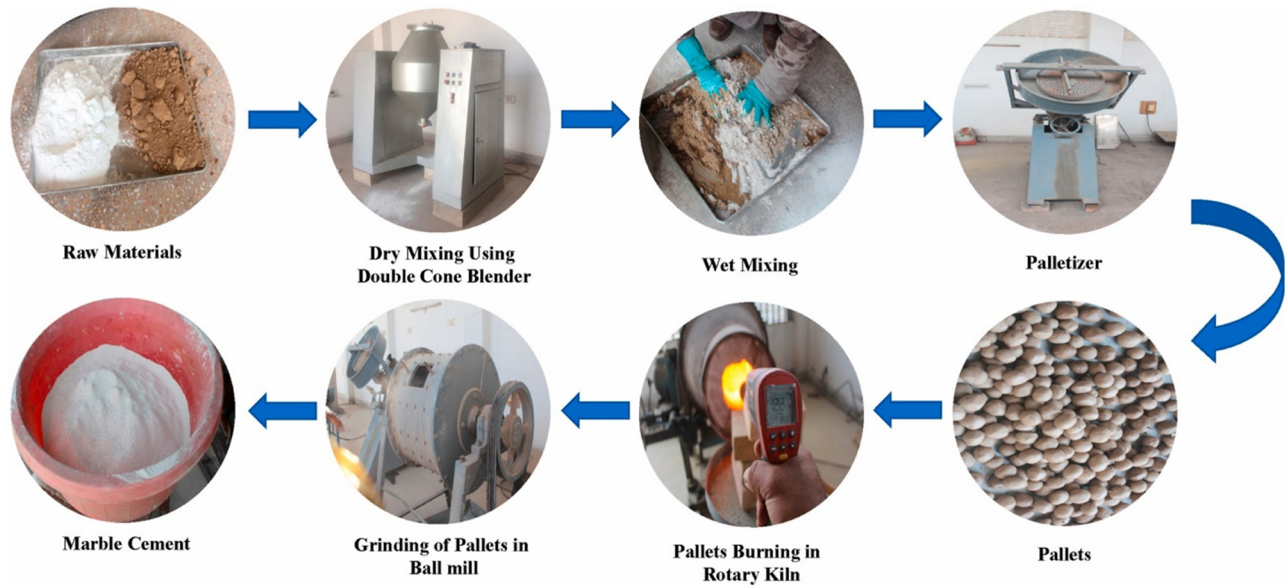


Figure 1: Process for producing MC [15].

blending WMP with a silica-rich substance, like clay, as shown in Figure 1 [14,15]. According to the findings of an X-ray diffraction study that was carried out by Khan *et al.*, MC encompasses 3.7% C_3S , 52.51% C_2S , and 23.11% free lime in its phase chemistries. Low amounts of C_3S and high concentrations of C_2S and CaO resulted from slow cooling. The MC pellets were gently brushed by the chilly wind. C_3S reverted to C_2S and CaO at around $1,100^\circ C$ after that. C_3S that is produced at high temperatures is retained by quickly cooling the cement clinkers during the OPC manufacturing process [16]. The decreased water demand and lack of slaking, which results in a greater concentration of free lime in MC, were the reasons for the shorter setting time of the cement when compared to OPC. Increases in the concentration of free lime are accompanied by an increase in the OPC. The immense internal pressure produced by cement grids accelerates the degradation of mortar and concrete mixed with them. As the permitted lime concentration increases, the compressive strength (CS) of cement generally decreases due to the high amount of ground-up calcium hydroxide ($Ca(OH)_2$) [17].

In order to increase the ingestion of free lime content, researchers suggested using pozzolanic materials as a partial substitute for MC. For a wide range of waste by-products, researchers are currently investigating pozzolanic applications [18–20]. Thermal coal power plants utilize coal powder as their fuel. As a result of the combustion process, fly ash (F) and bottom ash are produced. Minuscule particles of fly ash are carried in automobile emissions. By incorporating precipitators prior to the chimney,

it may be extracted from the exhaust gases. Prior to the exhaust gases entering the chimney, the precipitators filter out any fly ash [21,22]. Globally, rice is also grown. “Rice husk” refers to the brittle outer covering that shields each rice grain [23]. This trash turns to ash after being burned in a controlled environment. The pozzolanic properties of rice husk ash (R) are particularly noteworthy due to the finely powdered nature of the ash [24].

Research on cement-based materials (CBMs) has been extensive due to the significance of their mechanical characteristics [25]. The flexural strength (FS) and CS of CBM offer valuable insights into its properties. These characteristics are inherently linked to a wide array of mechanical and durability properties of the mortar [26,27]. To reduce unnecessary testing, analysts are developing analytical models to determine the material strength. It is common practice to use best-fit curves and other conventional models to reproduce material properties found by regression analysis. However, conventional regression algorithms could incorrectly presume the material’s intrinsic behavior because CBMs are nonlinear [28]. The application of artificial intelligence (AI), i.e., machine learning (ML), has led to the advancement of more complex models in this domain [29–33]. These models must be experimentally validated for accuracy before they can produce trustworthy predictions based on input features. More and more ML algorithms are being used to forecast CBM characteristics [34–36].

This research utilized ML techniques, particularly multi-expression programming (MEP) and gene expression

programming (GEP) to forecast the CS and FS of marble cement mortar (FR-MCM) incorporating rice husk ash and fly ash. This prediction is based on data from studies that have already been published. Statistical tests, the distribution of anticipated outcomes, and the R^2 coefficient were among the metrics used to gauge the enactment of ML algorithms. The main intent of this research was to assess the predictive capabilities of ML methods for material properties. Exploratory experiments or database analysis could provide the data needed for ML algorithms. Machine learning algorithms may be able to learn more about material properties by examining these data. Using experimental data and four input parameters, this study assessed the capability of ML approaches to forecast the FS and CS of FR-MCM. Raw material relevance was further investigated using sensitivity analysis. One possible use for the

newly acquired characteristics and ML models is to aid in the creation of CBM blends or to add to the existing database for environmentally friendly materials.

2 Methods of study

2.1 Data gathering and analysis

Using a dataset from prior research [15] and applying MEP and GEP methodologies, this study intended to approximate the FS and CS of FS-MCM. The estimation of the FS and CS of FR-MCM was conducted using four input parameters: MC, fly ash (F), curing age (A), and rice husk ash (R). The images of the lab work conducted are shown in Figure 2.



Figure 2: Images of specimens for CS and FS within molds, dried, and during testing [15].

Initially comprising 84 data points, the dataset was expanded to 500 points. The executed Python code adhered to a specific process for adding new data points to the dataset. The script begins by launching a Tkinter-powered file dialog box for users to select a database file. Once imported into a Pandas DataFrame, the script verifies the present number of data points. A new file is then created to store the merged dataset, which includes both the synthetic data and the original DataFrame. The script makes statements that clarify the issue as the data are being enhanced. In these declarations, one can find details like how many data points were added, how many data points were synthesized, and where the stored file was located. In addition, the script accounts for cases where resampling is required or when no file is selected. Data gathering and categorization were aided by data preparation. The widely recognized task of discovering fresh insights from available data frequently faces a notable challenge. A common strategy to surmount this hurdle is the preparation of data for data mining. Data preparation is cleaning data of any extraneous or noisy information. The model analysis made use of statistical approaches such as error dispersion and regression. The models' efficacy and reliability were also evaluated. Histograms in Figure 3(a)–(f) depict the frequency distribution of values, offering a visualization of the overall dataset frequency distribution by integrating all components' distributions. The distribution of the collection's value can be better understood with the use of a relative frequency distribution.

Using Pearson's correlation coefficient (r) is a typical approach to discovering parameter dependencies [37]. The results of the distinctive association map plot are shown in Figure 4. When looking to prove multicollinearity or parameter dependency, the R^2 test is a good tool to utilize [38]. The r -value can take values between -1 and $+1$; 0 indicates no connection, while $+1$ indicates a significant positive relationship [39]. The bottom row of Pearson's array displays the association between the independent variables and the dependent variables, in this case, CS and FS. Multicollinearity is an important point to make about ML algorithms [40]. In order to prevent multicollinearity issues, ML models must ensure that no two variables have a correlation coefficient (r -value) greater than 0.8 [41]. As can be observed in Figure 4, the r -value falls inside the acceptable range. Therefore, multicollinearity in correlational models is highly unlikely.

2.2 ML modeling

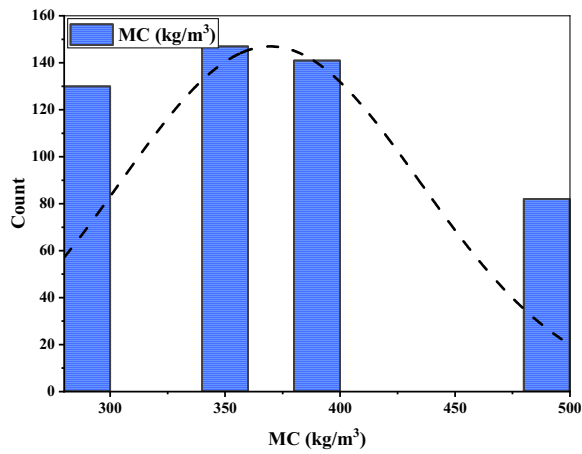
The CS and FS of FR-MCM were tested in a supervised environment. Four inputs were used to generate the

results (CS and FS). More advanced ML methods, like MEP and GEP, were used to predict the CS and FS of FR-MCM. When evaluating ML algorithms, it is a common practice to compare the results with the input data. The ML models were trained using 70% of the dataset, while the remaining 30% was held aside for testing. A high R^2 score for the expected outcome indicates that the model is correct. R^2 is low for a large fluctuation, indicating that the anticipated and actual values differ by a small amount [42]. The correctness of the model can be confirmed through various methods, including statistical analysis and error assessments. An illustration of a scenario model is depicted in Figure 5. All of the GEP and MEP model hyperparameter parameters are included in Table 1.

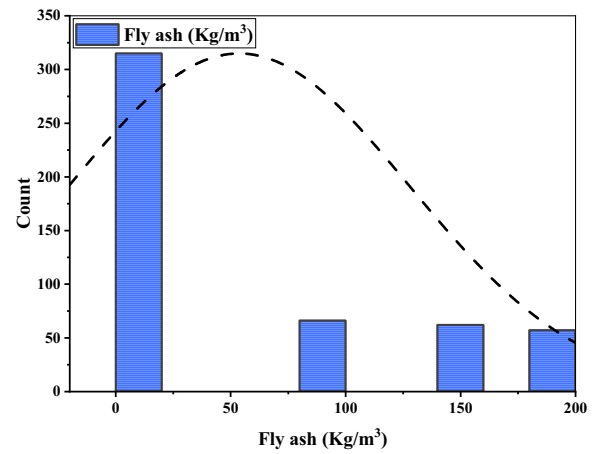
2.2.1 GEP model

J. H. Holland's genetic algorithm (GA) is primarily founded on Darwin's theory of evolution [44]. The chromosomal endpoint is defined by an ordered succession of GAs and contains chromosomes of a particular length. Koza came up with the phrase "gene programming" to describe one novel GA approach [45]. An evolutionary model is generated via generalized problem-solving (GP) using GAs [46]. The adaptation ability of GP is derived from its capacity to replace binary strings of fixed length with nonlinear structures like parse trees. According to Darwin's theory, present AI programs deal with reproduction-related problems by using genetic components that occur naturally, such as procreation, crossover phenomenon, and modification [47]. With GP, wasteful programs are systematically removed from successive iterations. Just like in the previous example, removing the undesired trees is an integral part of replanting the region using the selected method. However, evolution safeguards early convergence [47,48]. Five key considerations must be settled upon before the GP approach can be put into action. Priorities by area, fitness evaluation, key functional operators (including crossovers and populace capacity), and approach-surgical incision results [47]. The majority of the parse trees were created by an integrated genomic processor, even though GP develops models frequently [48]. As a result of the fact that they must function as both genotype and phenotype, nonlinear GP forms provide expressions that are difficult to understand for desired features [48].

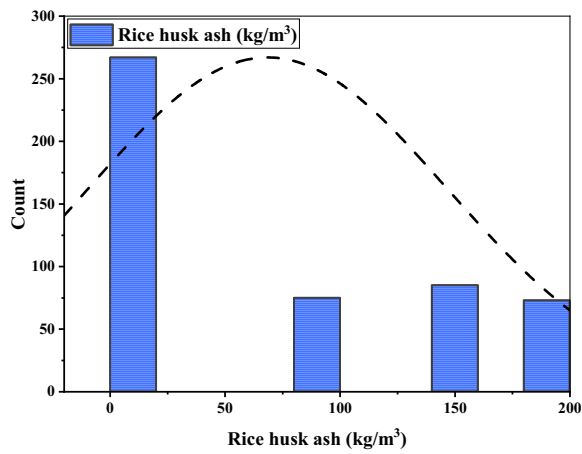
First proposed by Ferreira, GEP is an adaptation of GP that he developed [48]. The GEP model follows the notion of population formation and uses static-length linear chromosomes to build parse trees. Although GEP is an enhanced variant of GP, GP still employs medium-sized



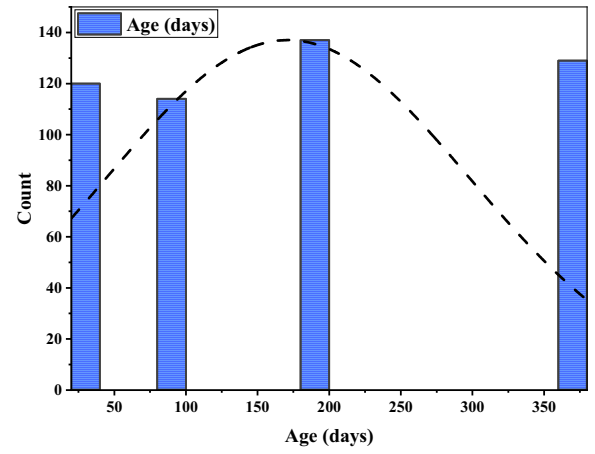
(a)



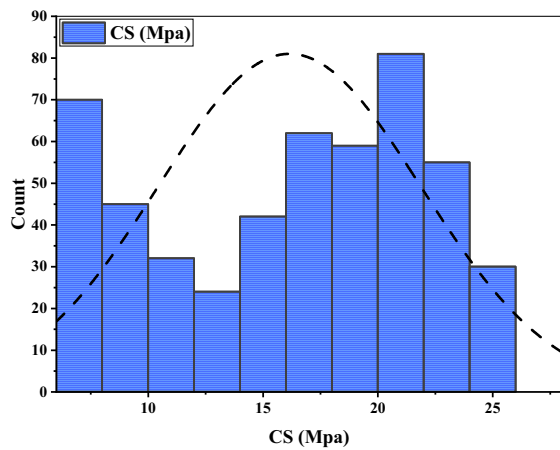
(b)



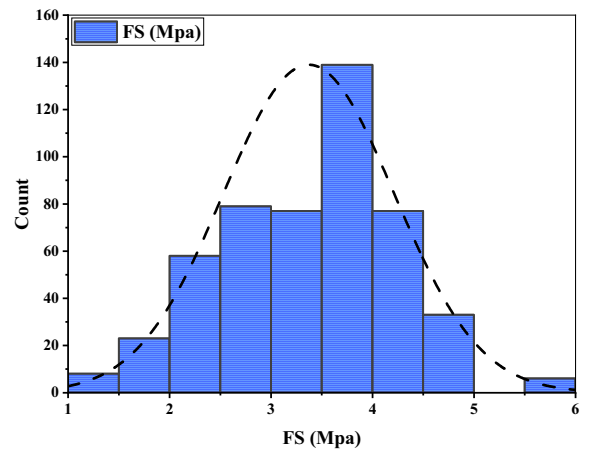
(c)



(d)



(e)



(f)

Figure 3: Frequency distribution plots illustrating the database's input and output characteristics: (a) MC, (b) fly ash, (c) rice husk ash, (d) age, (e) CS, and (f) FS.

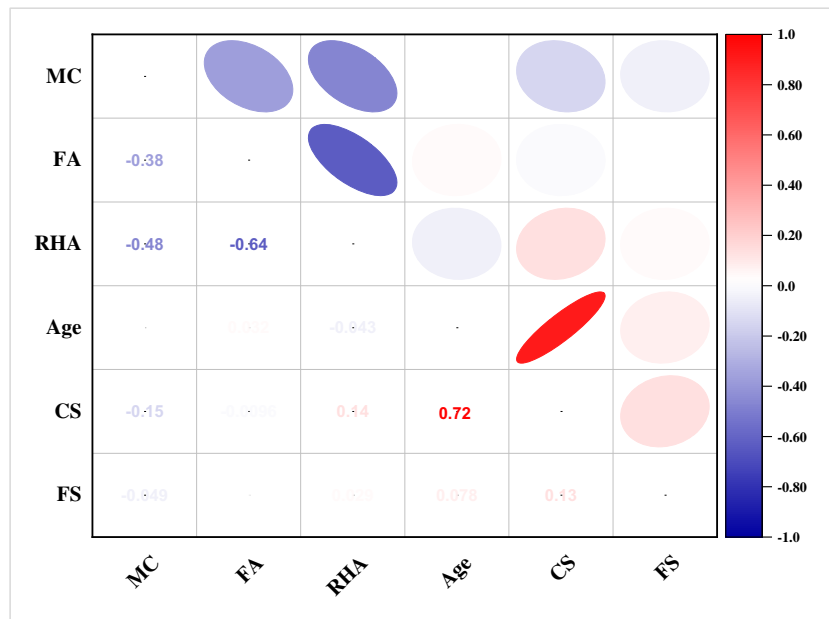


Figure 4: Parameter correlation heat map.

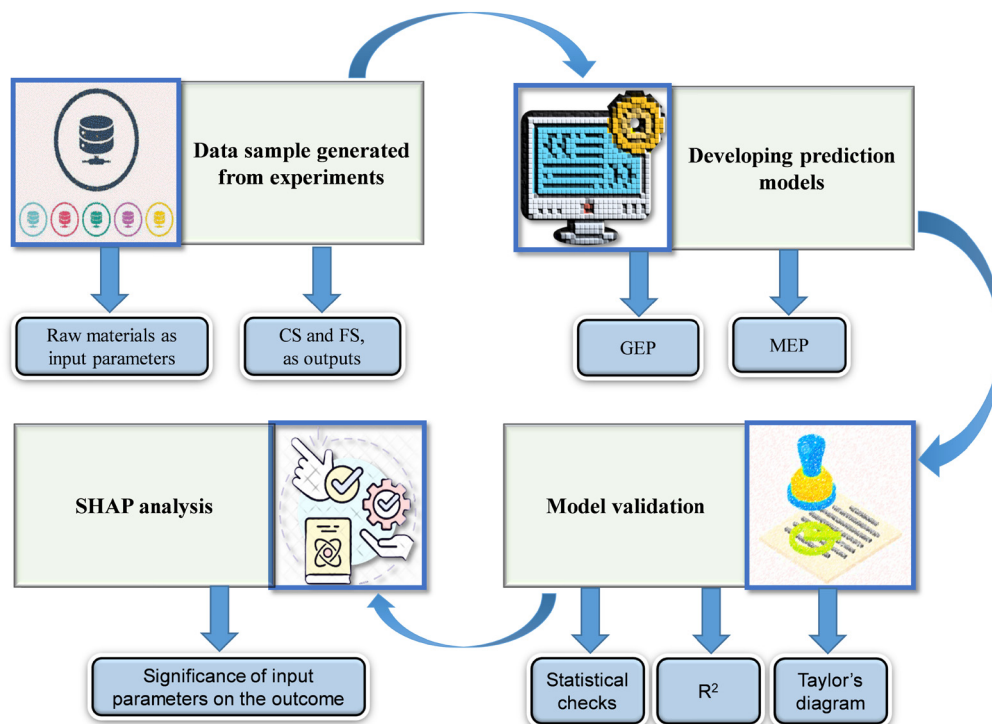


Figure 5: A comprehensive overview of the study's method.

program development using simple chromosomes of a specific length. Among GEP's several advantages is its high-reliability prediction capabilities for complex and nonlinear problems [49,50]. As with GP, this one has a fitness function, an end set, and specified termination

criteria. The “Karva” dialect really recognizes chromosomes before they are formed, even if the GEP process produces them with seemingly random numbers. GEP is fundamentally based on a constant-length linear structure. Conversely, in GP's data processing algorithm, parse

Table 1: Hyperparameters for GEP and MEP models (parameters comparable to the study of Amin et al. [43])

GEP		MEP	
Parameters	Settings	Parameters	Settings
Stumbling mutation	0.00141	Code length	40
Chromosomes	200	Number of generations	250
Leaf mutation	0.00546	Number of runs	15
Random chromosomes	0.0026	Replication number	15
Data type	Floating number	Cross over probability	0.9
General	CS, FS	Mutation probability	0.01
Constant per gene	10	Number of generations	500
Linking function	Addition	Number of treads	2
Lower bound	-10	Error	MSE, MAE
Upper bound	10	Operators/variables	0.5
Mutation rate	0.00138	Number of sub-populations	50
Genes	4	Sub-population size	100
Inversion rate	0.00546	Terminal set	Problem input
One-point recombination rate	0.00277	Function set	+, −, ×, ÷, square root
IS transposition rate	0.00546	Problem type	Regression
Gene recombination rate	0.00277		
Function set	+, −, ×, ÷, square root		
RIS transposition rate	0.00546		
Gene transposition rate	0.00277		
Two-point recombination rate	0.00277		

trees of varying lengths are observable. Chromosomes with a constant length are the original name for these unique cords. Thus, dynamic manifestation/parse trees with prong morphologies of varied sizes are used to represent chromosomes [47]. Several phenol strains and genotypes have unique genetic codes [48]. GEP eliminates costly structural alterations and duplications by preserving the genome.

Normally, a chromosome will have two halves that complement each other, called the “head” and the “tail” sections. Interestingly, this can occur when a single chromosome produces many genes [47]. These genes include the instructions for performing Boolean operations, as well as mathematical reasoning and numeracy. Programmers provide genetic coding cells with specialized functions. Inferring these chromosomes’ contents using the new language Karva can establish the framework for empirical formulations. The journey starts in Karva on the ET, which is followed by a leading revolution. Rendering to Eq. (1), ET places the connected nodes in the sublayer [49]. The amount of ETs dictates two parameters: the duration and intensity of GEP gene K expression:

$$ET_{GEP} = \log\left(i - \frac{3}{j}\right). \quad (1)$$

The independence of GEP’s results from pre-existing relationships highlights its advanced nature as an ML

technique. The development of a GEP mathematical equation involves several stages, as illustrated in Figure 6. An individual’s chromosomal count remains constant at birth. After verifying these chromosomes as expression trees (ETs), a comprehensive evaluation of each individual’s fitness is performed. The most accomplished individuals undergo an iterative method to refine and achieve the optimal solution, granting reproductive rights to the fittest and strongest individuals. Ultimately, breeding, mutation, and crossover yield the final numerical expression.

2.2.2 MEP model

There is a cutting-edge, linear-based genetic programming technology known as MEP. This technique relies on linear chromosomes. As far as their core software is concerned, GEP and MEP are completely interchangeable. A number of software components (substitutes) may be encoded into an individual’s genome using MEP, in contrast to its forerunner, the GP method. The desired result is achieved by selecting the most advantageous chromosomes through fitness evaluation [40,51]. Rendering to Grosan and Oltean [52], this hypothetical scenario occurs after a bipolar arrangement re-establishes itself, resulting in two successive generations adopting one parent as their own. Until the condition for termination is satisfied or the best

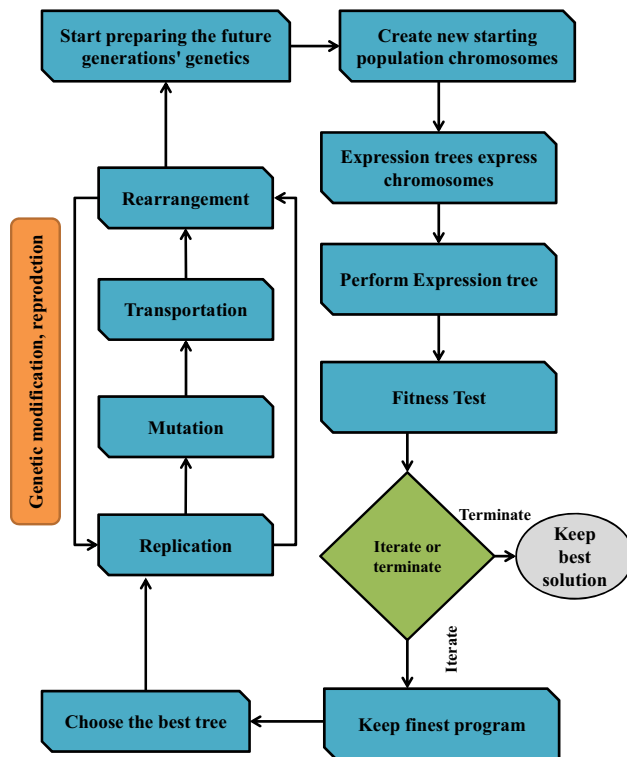


Figure 6: Methodology of the GEP flowchart [43].

program is identified, the process keeps running, as shown in Figure 7. Modifications in infants take place during this process. Similar to the GEP model, the MEP paradigm enables the combination of numerous components. Key considerations in MEP include algorithm or code length, number of subpopulations, number of functions, and the potential for crossover [53]. When there are as many bundles as there are people in a population, evaluation and management get quite complicated and time-consuming. Code length has a significant impact on the number of mathematical expressions. MEP parameters required to construct a solid model of mechanical properties are shown in Table 1.

In both the assessment and simulation phases, both methods extensively utilize literature datasets [54,55]. There is evidence suggesting that established linear GP methods like GEP and MEP excel in estimating the attributes of ecological mortar and concrete. Grosan and Abraham asserted that a fusion of linear genomic programming and maximal entropy programming (METP) represents the utmost active neural network-based method [56]. The mode of operation utilized by the GEP is slightly more complicated than that utilized by the MEP [53]. Despite MEP's lower density in comparison to GEP, several notable distinctions exist: (i) MEP explicitly encodes references to function arguments;

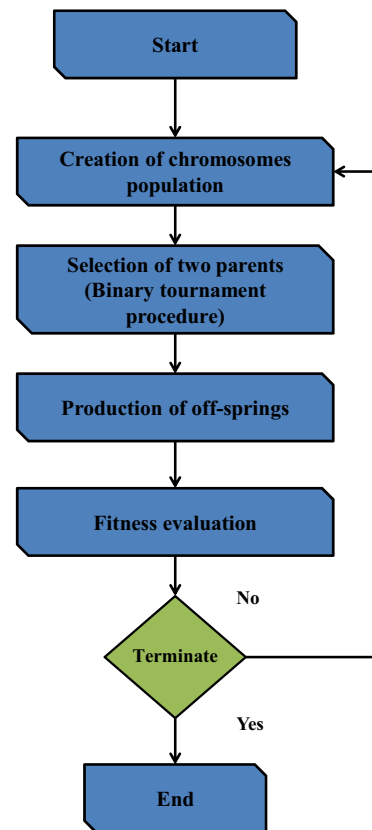


Figure 7: Process flow diagram for MEP operation [43].

(ii) there is no necessity for non-coding sections to be positioned at a fixed point within the genetic factor; and (iii) MEP permits code reprocessing [57]. Further thorough evaluations of these genetic approaches to engineering challenges are, hence, absolutely needed. Signals at the “head” and “tail” of the GEP chromosome facilitate the creation of syntactically correct software programs, leading many to believe it is more capable [52]. Therefore, comprehensive evaluations of these genomic engineering methods are essential.

2.3 Model's validation

A test set was utilized to conduct a statistical analysis of the MEP and GEP models. From the literature [55,58–61], seven numerical measures were figured for every one of the models: Pearson's correlation coefficient (R), relative root mean square error (RRMSE), mean absolute error (MAE), Nash–Sutcliffe efficiency (NSE), relative squared error (RSE), root mean square error (RMSE), and mean absolute percentage error (MAPE). Numerous statistical

measures can be calculated using the following formulas (Eqs (2)–(8)):

$$R = \frac{\sum_{i=1}^n (a_i - \bar{a}_i)(p_i - \bar{p}_i)}{\sqrt{\sum_{i=1}^n (a_i - \bar{a}_i)^2 \sum_{i=1}^n (p_i - \bar{p}_i)^2}}, \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - a_i|, \quad (3)$$

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(p_i - a_i)^2}{n}}, \quad (4)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|p_i - a_i|}{a_i}, \quad (5)$$

$$RSE = \frac{\sum_{i=1}^n (a_i - p_i)^2}{\sum_{i=1}^n (\bar{a} - a_i)^2}, \quad (6)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (a_i - p_i)^2}{\sum_{i=1}^n (a_i - \bar{p}_i)^2}, \quad (7)$$

$$RRMSE = \frac{1}{|\bar{a}|} \sqrt{\frac{\sum_{i=1}^n (a_i - p_i)^2}{n}}. \quad (8)$$

In a data set with n points, the i th actual value is a_i and the average predicted value is p_i ; for each data point i , n represents the total number of points. One reliable method of gauging a model's predictive capability (a_i and p_i) is to look at its correlation coefficient, R . A large value for R designates a highly correlated connection between the anticipated and real levels of production [62]. Nevertheless, R is unaltered by division and multiplication operations. At the intersection of the actual and anticipated outcomes, the R^2 statistic was calculated in order to ascertain the actual value. Greater efficiency in model construction is indicated by R^2 values that are closer to 1 compared to other methods [63,64]. The proposed model demonstrates significantly improved performance with smaller errors, where both RMSE and MAE approach zero. Similarly, they exhibit substantial improvement as error sizes increase [65,66]. Although, after additional investigation, it turns out to be abundantly evident that MAE operates exceptionally well in databases that are both uninterrupted and uniform [67]. Reducing the levels of the mistakes mentioned before often leads to an improvement in the model's performance.

Two of the most useful ways to evaluate a model's predictive power are the Taylor diagram and statistical validation. This statistic helps show which models are more accurate and dependable by showing how far they are from the truth, which serves as a benchmark [68,69]. Determining the best model placement requires the integration of three essential metrics: the standard deviation, represented on the model's horizontal and vertical axes;

the correlation coefficient, shown by radial lines; and the RMSE, depicted as concentric circles centered around the point of the true value. In prediction tasks, the most reliable model is identified by its highest accuracy rate [68].

3 Results and interpretation

3.1 GEP models

3.1.1 CS-GEP model

Using mathematical conclusions based on genomic number and head dimensions, the GEP technique (as shown in Figure 8(a)–(d)) created models that calculated the CS using ETs. Most sub-ETs in predicting FR-MCM's CS utilized fundamental arithmetic operations such as dot product, multiplication, subtraction, addition, and square root. Deciphering these sub-ETs through the GEP approach yields equations. By inputting data into Eq. (9), future CS values of FR-MCM can be predicted. The resulting model outperforms an ideal model under perfect conditions when sufficient data are available. As demonstrated in Figure 9(a), the continuous line signifies a seamless fit to the data, whereas broken lines indicate variations of up to 20% from this line, visually demonstrating the arrangement between trial and estimated CS findings. Predicted CS values from the GEP model closely matched measured values, indicating its effectiveness. An R^2 of 0.952 and a CS prediction inside the 20% threshold 96% of the time indicate that the model achieved significantly improved accuracy. Figure 9(b) plots the absolute error compared to trial data, showing how the GEP model aligns with experimental results. The predictions showed minimal deviation, with an absolute error ranging from 0.00 to 3.477 MPa and an average error of 1.01 MPa. Additionally, Figure 10 illustrates that the distribution of error values closely resembles that of a bell-shaped curve. The pressure measurements ranged from 54 (<0.5 MPa) to 75 (>1.0 MPa). The range from 37 to 54 was between 0.5 and 1.0 MPa. It is important to note that maximal error frequencies occur very seldom:

$$\begin{aligned} CS(\text{MPa}) = & (\sqrt{3.635}) + (MC + (\sqrt{\sqrt{\sqrt{(MC \times R) + 2F}}})) \\ & + (\sqrt{A} - MC) + \left(\frac{(3R + (MC + F))}{(12.125 - A) \times (A - 8.554)} \right), \end{aligned} \quad (9)$$

where MC represents the marble cement, A represents the curing age, R represents the rice husk ash, F represents the fly ash, and CS represents the compressive strength.

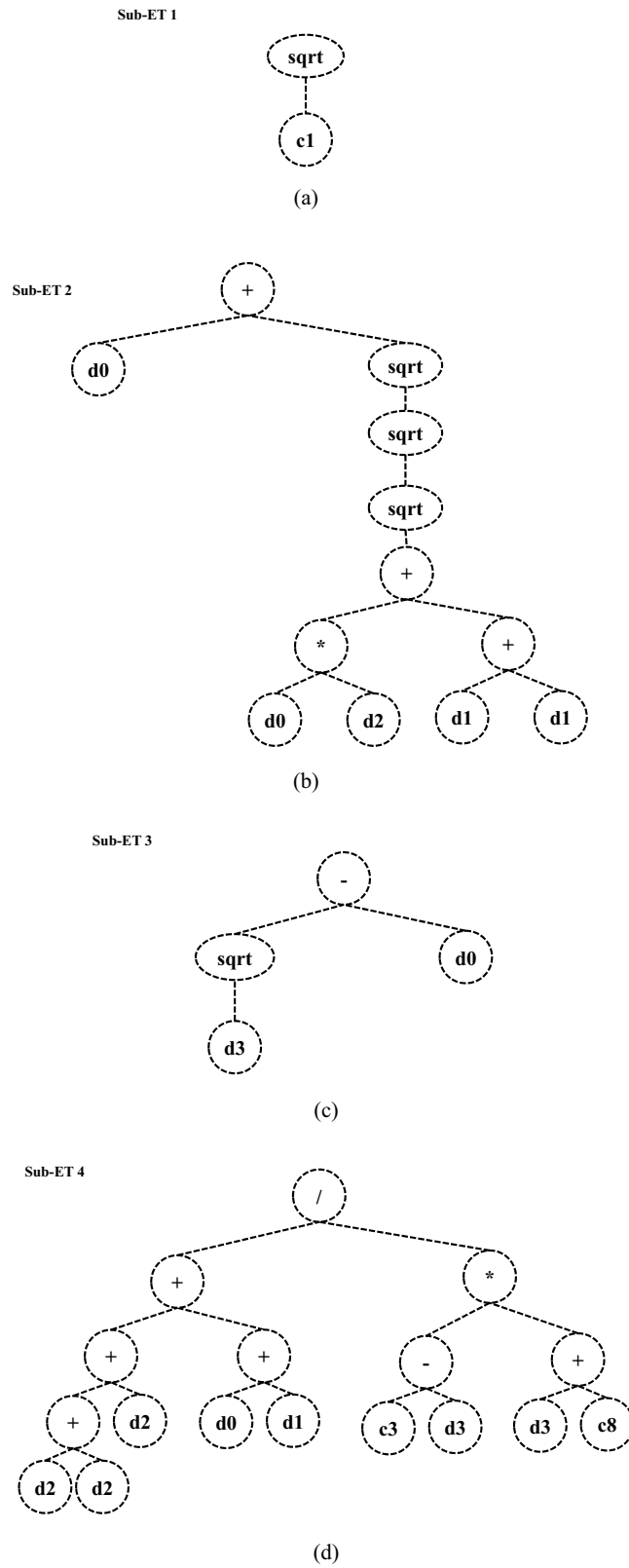
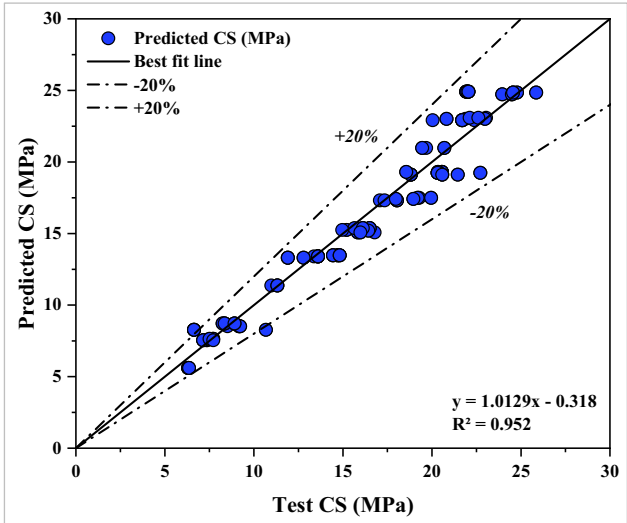
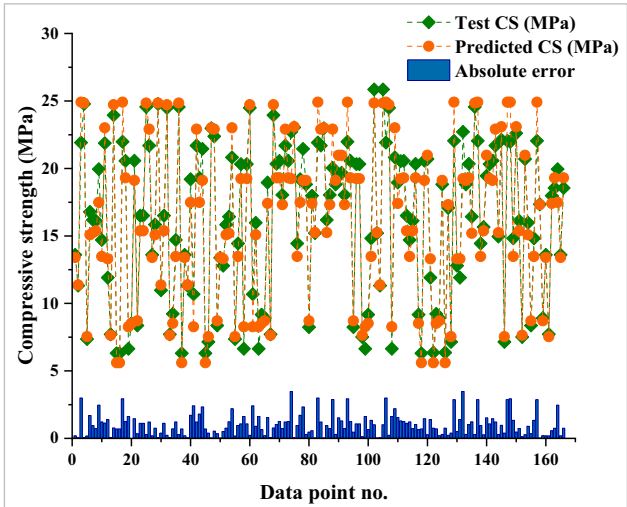


Figure 8: Illustration of the CS-GEP model's ET: (a) sub-ET 1, (b) sub-ET 2, (c) sub-ET 3, and (d) sub-ET 4.



(a)



(b)

Figure 9: GEP model for CS-FR-MCM: (a) trial-projected CS correlation and (b) trial-projected CS error spread.

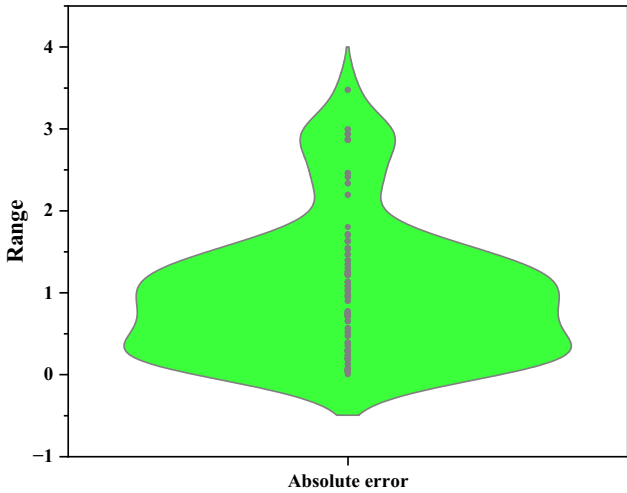


Figure 10: Violin diagram for CS-GEP error dispersion.

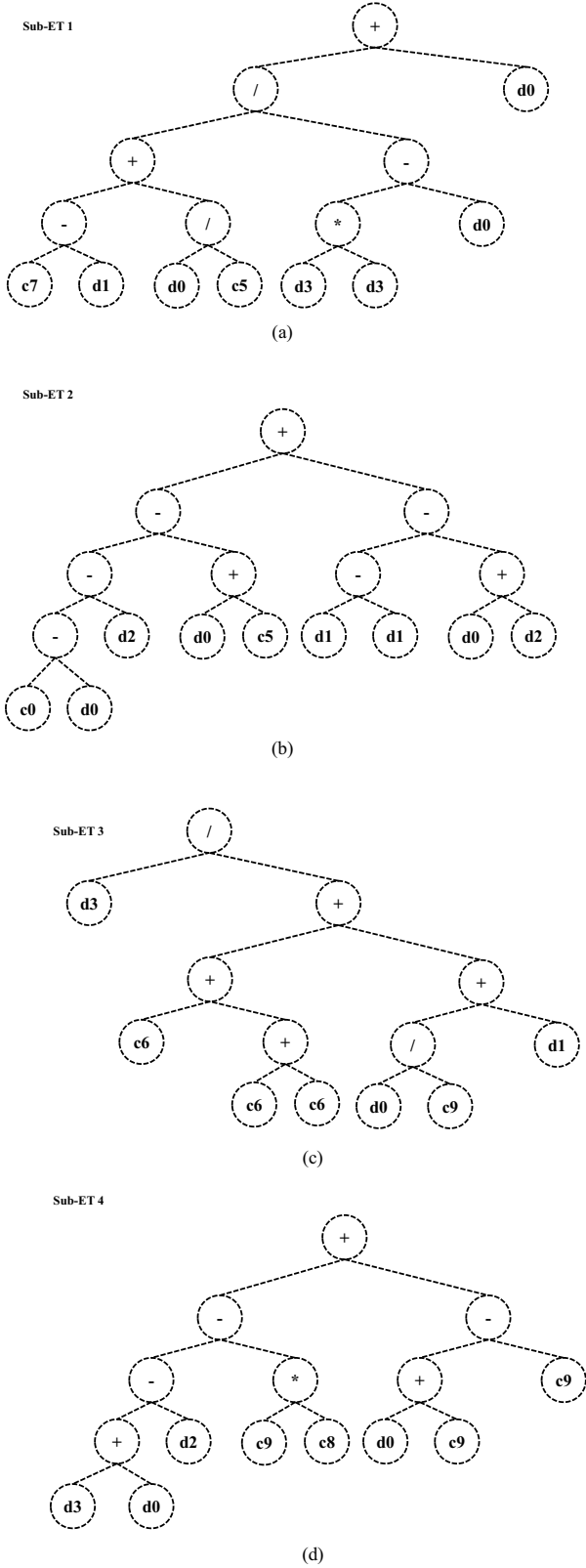


Figure 11: Illustration of the FS-GEP model's ET: (a) sub-ET 1, (b) sub-ET 2, (c) sub-ET 3, and (d) sub-ET 4.

3.1.2 FS-GEP model

Figure 11(a)–(d) illustrates how the GEP method utilizes arithmetical associations based on genomic number and head dimension to construct ET-based models for FS. In FR-MCM's FS forecasts, the majority of sub-ETs are constructed using basic arithmetic operations ($\div$, $\times$, $-$, $+$, and square root). Deciphering these sub-ETs through the GEP

3.2 MEP models

3.2.1 CS-MEP model

To ascertain the CS of FR-MCM, an empirical formula was derived based on MEP findings that incorporate the effects of the four primary constituents. The final set of formulated equations is presented in Eq. (11):

$$\begin{aligned} \text{CS(MPa)} = & (2(\sqrt{\sqrt{\sqrt{2A}} * \sqrt{\sqrt{2A}}}) - \left(\frac{(\sqrt{\sqrt{\sqrt{2A}} * \sqrt{\sqrt{2A}}}) + (FA + \sqrt{\sqrt{2A}} + \text{RHA}) + \left(\frac{FA + \sqrt{\sqrt{2A}} + \text{RHA}}{\sqrt{\sqrt{\sqrt{2A}} * \sqrt{\sqrt{2A}}}} \right)}{(FA + \sqrt{\sqrt{2A}}) - \text{RHA}} \right) \Bigg) \\ & + \frac{FA + \sqrt{\sqrt{2A}}}{\text{MC}}, \end{aligned} \quad (11)$$

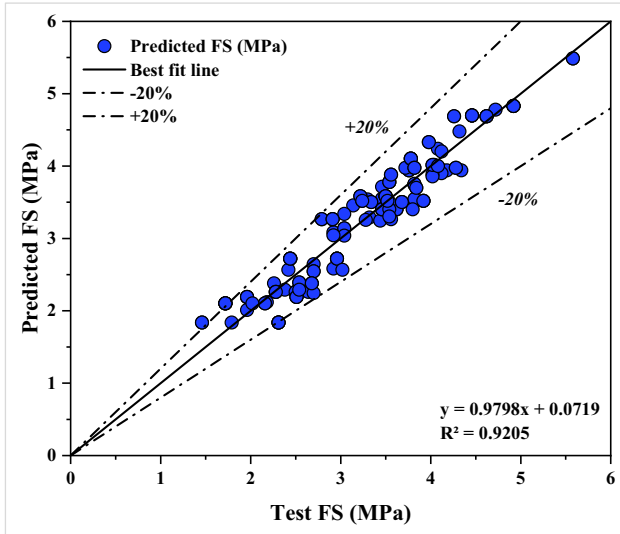
method yields arithmetic formulas. By inputting data into Eq. (10), the future FS of FR-MCM can be estimated. Predicted FS values from the GEP model closely matched the measured values, highlighting its effectiveness. The model achieved an R^2 of 0.920 and predicted FS inside the 20% limits 97% of the time, representing significantly improved exactness. Figure 12(b), which plots the absolute error versus the experimental data, shows the GEP model. The absolute error range for the test findings and the predictions provided by the GEP equation was 0.001–0.476 MPa, with an average of 0.207 MPa. A bell-curve distribution was observed for the error value (Figure 13). Overall, 45 values were more than 0.3 MPa, 70 values were in the range of 0.1–0.3 MPa, and 51 values were lower than 0.1 MPa. Note that maximal error frequencies are extremely rare occurrences:

$$\begin{aligned} \text{FS(MPa)} = & \left(\frac{((-15.657 - F) + \left(-\frac{\text{MC}}{1.325} \right))}{A^2 - \text{MC}} \right) + \text{MC} \\ & + (((5.109 - \text{MC}) - R) - (\text{MC} + 2.174)) \\ & + (2F - (\text{MC} + R)) + \left(\frac{A}{(32.406 + (\frac{\text{MC}}{1.518} + F))} \right) \\ & + (((\text{RHA} \times \text{MC}) + R) + 0.150) \\ & + ((\text{MC} + 0.108) - 0.108), \end{aligned} \quad (10)$$

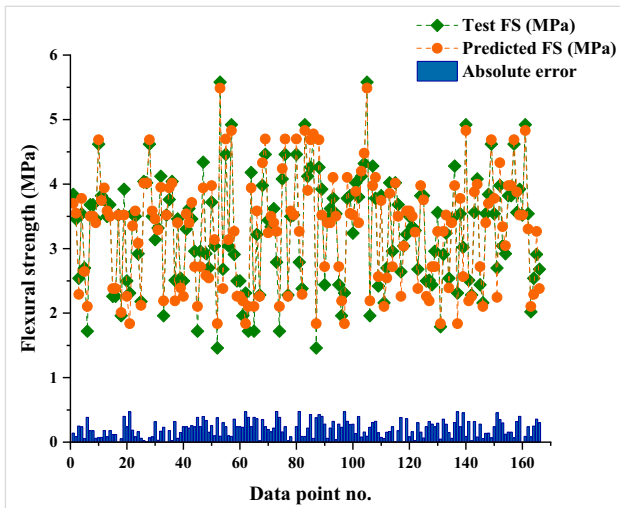
where MC represents marble cement, A represents curing age, R represents rice husk ash, F represents fly ash, and CS represents compressive strength.

where MC represents the marble cement, A represents the curing age, R represents the rice husk ash, F represents the fly ash, and CS represents the compressive strength.

Figure 14(a) illustrates that the MEP model exhibits robustness and is well-trained, as indicated by its high R^2 value of 0.971. To top it all off, it works fine with new, untested data. The R^2 value of the CS-MEP model is higher than that of the CS-GEP model, indicating that the former is more accurate. Figure 14(a) displays a line that is perfectly in line with the data, which is the solid black line, and dotted lines that are off by as much as 20%. With very similar predicted and observed values of CS, the MEP model was able to accurately predict the CS of FR-MCM. The MEP method successfully detected CS, with 100% of its predictions falling inside the 20% cutoff, showing remarkable precision. The results of comparing the target and actual values in the MEP simulations are shown in Figure 14(b). The data indicate that MEP estimate error margins ranged from 0.016 to 2.806 MPa, with an average of 0.841 MPa. To be more specific, 59 errors were less than 0.5 MPa, 47 were in the range of 0.5–1 MPa, and 73 were higher than 1 MPa. The MEP model is superior to the GEP model in predicting values that are deemed outliers. The MEP model decreases error correlation and standard deviation, as shown in Figure 15's violin figure. Due to its generalizability and ease of use, the MEP equation is a popular tool. With fewer mistakes and a greater correlation coefficient, the MEP model beats the GEP model.



(a)



(b)

Figure 12: GEP model for FS-FR-MCM: (a) correlation between trial and projected FS and (b) error spread between trial and projected FS.

3.2.2 FS-MEP model

An empirical formula was developed based on MEP findings to predict the FS of FR-MCM and incorporate the effects of the four primary constituents. The final set of formulated equations is presented in Eq. (12):

$$FS(MPa) = \left(\sqrt{\sqrt{(2RHA + \sqrt{A})(1 + FA)} + A} - \frac{\sqrt{A}}{2RHA + \sqrt{A}} \right) + \left(\frac{\sqrt{\sqrt{(2RHA + \sqrt{A})(1 + FA)} + A}}{MC + FA - 2RHA - \sqrt{A} - FA(2RHA + \sqrt{A})} \right), \quad (12)$$

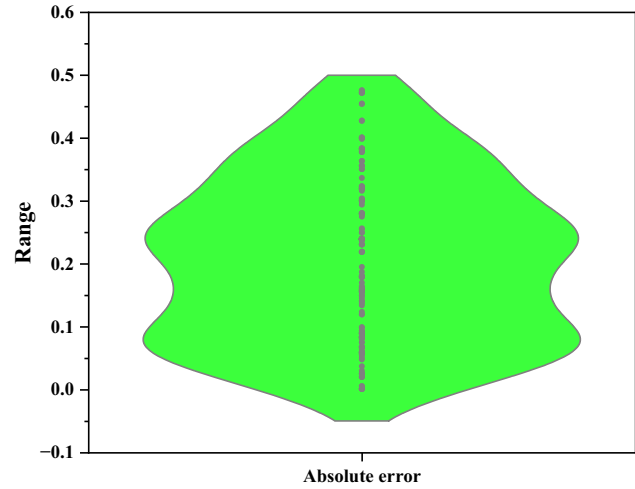
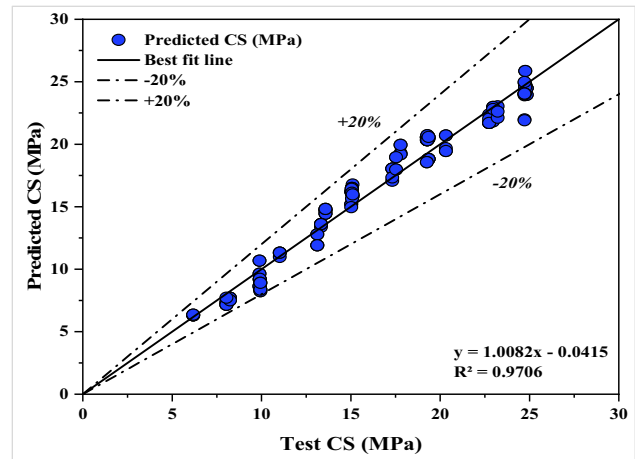
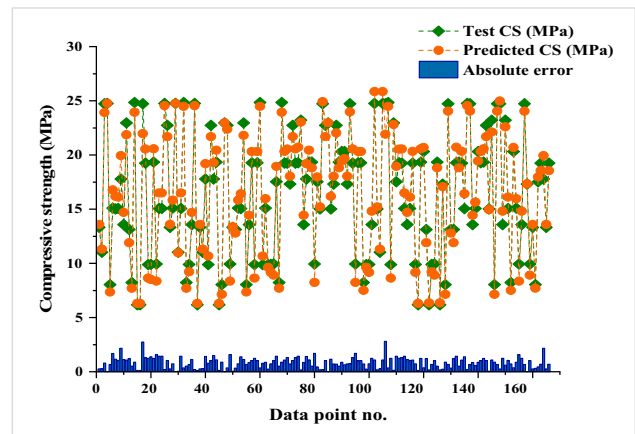


Figure 13: Violin plot for FS-GEP error dispersion.



(a)



(b)

Figure 14: MEP model for CS-FR-MCM: (a) correlation between trial and projected CS and (b) error spread between trial and projected CS.

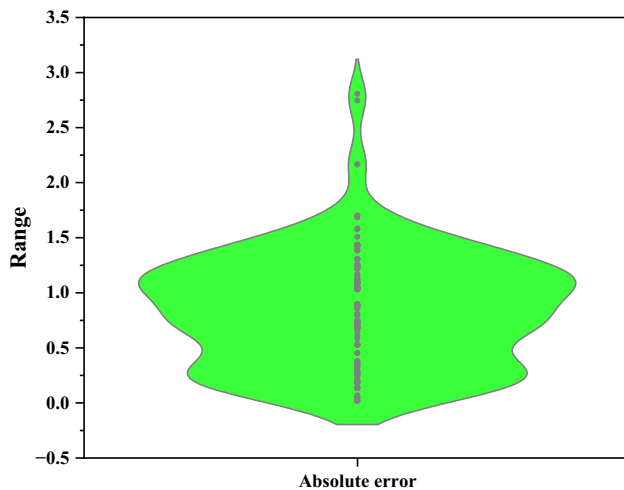
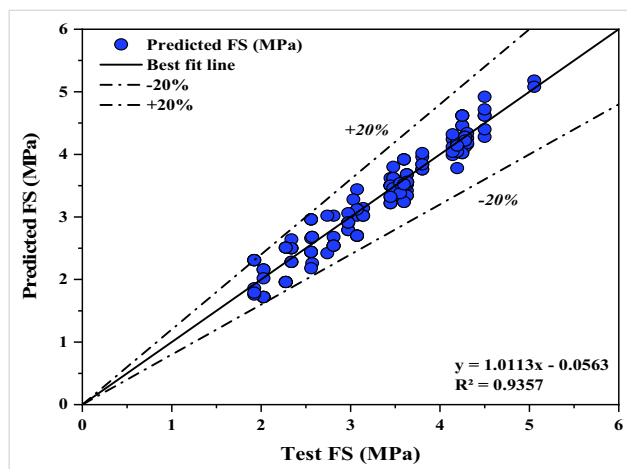
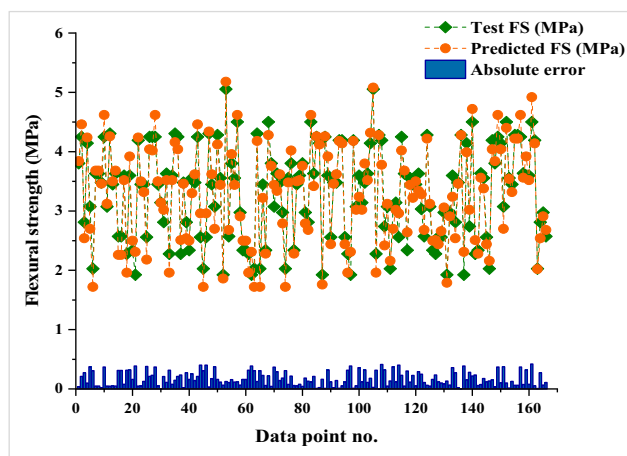


Figure 15: An error distribution violin plot for CS-MEP models.



(a)



(b)

Figure 16: MEP model for FS-FR-MCM: (a) correlation between trial and projected FS and (b) error spread between trial and projected FS.

where MC represents the marble cement, A represents the curing age, R represents the rice husk ash, F represents the fly ash, and CS represents the compressive strength. Figure 16(a) demonstrates a well-trained MEP model with an R^2 value of 0.935, indicating its capability to handle complexity effectively. With a higher R^2 value, the FS-MEP model also proves to be more accurate than the FS-GEP model. The predicted values of FS are closely aligned with the measured values, highlighting the MEP model's efficiency in predicting the FS of FR-MCM. The MEP method achieved remarkable accuracy in predicting FS, which was consistently inside the 20% benchmark. Figure 16(b) compares the target and actual values calculated in MEP simulations. The results showed that 0.001–0.419 MPa were the values that made up the MEP forecasts, with a standard deviation of 0.172 MPa. Specifically, 56 errors occurred at pressures below 0.1 MPa, 71 at pressures between 0.1 and 0.3 MPa, and 39 at pressures over 0.3 MPa. The MEP model outperformed the other in terms of predicting outlier values. The MEP model reduced the standard deviation of errors, as shown in the violin plot in Figure 17. The MEP equation has received a lot of praise for being both simple and widely used. It is evident that the MEP model works better than the GEP model due to its greater R^2 and reduced error levels.

3.3 Model's validation

Table 2 presents the results of efficacy and inaccuracy metrics (R , RSE, RMSE, NSE, MAE, and RRMSE) derived from calculations using Eqs (2)–(8). A lower error score indicates that the produced models are more accurate in their predictions. The performance metrics of the GEP and MEP models for predicting the properties of FR-MCM highlight notable differences. The FS-GEP model achieved an MAE of 0.207 MPa, while the CS-GEP model recorded a higher MAE of 1.011 MPa, with corresponding RMSE values of 7.10 and 6.50%, respectively. In contrast, the MEP models demonstrated improved accuracy, with the FS-MEP model achieving a lower MAE of 0.172 MPa and the CS-MEP model showing an MAE of 0.841 MPa. These results indicate that the MEP approach consistently outperformed the GEP method in both flexural and CS predictions for FR-MCM. However, the RMSE values of the CS-MEP and FS-MEP models were significantly reduced to 5.70 and 5.90%, respectively. There was a surprising similarity between the CS and FS models, as demonstrated by additional error-based statistical metrics. These metrics encompassed

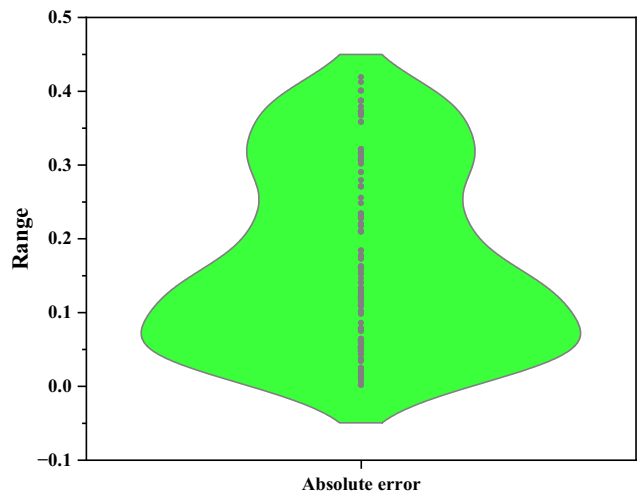
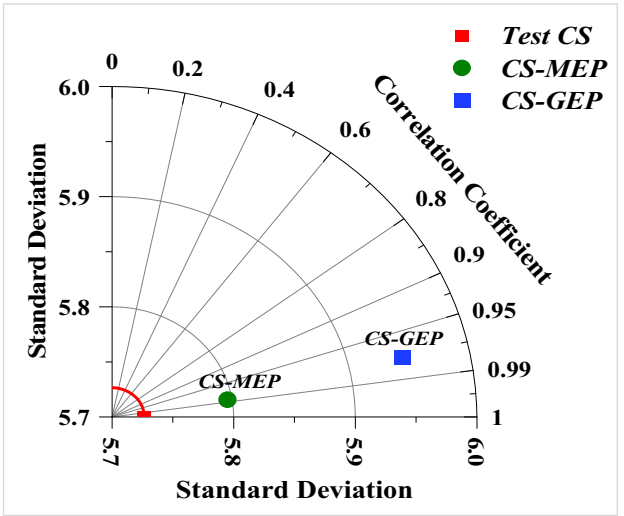


Figure 17: FS-MEP error dispersion violin plot.

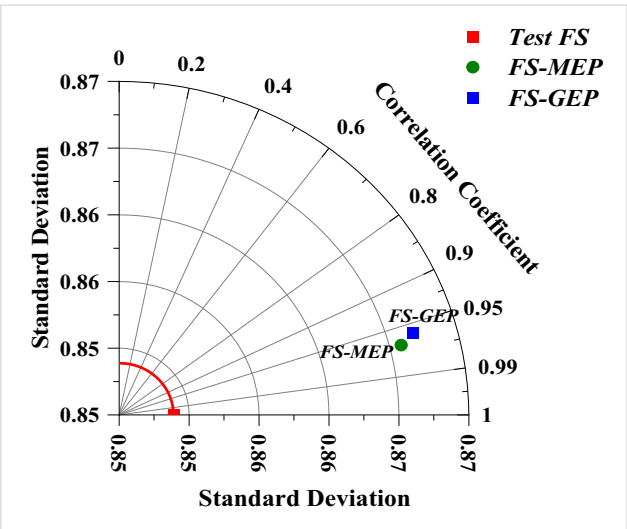
Table 2: Results of statistical analysis

Property	CS model		FS model	
	GEP	MEP	GEP	MEP
MAE (MPa)	1.011	0.841	0.207	0.172
MAPE (%)	6.50	5.70	7.10	5.90
RMSE (MPa)	1.305	0.987	0.244	0.210
R	0.976	0.985	0.960	0.967
RSE (MPa)	0.284	0.244	0.266	0.248
NSE	0.948	0.969	0.917	0.929
RRMSE (MPa)	0.602	0.322	0.548	0.372

RSE, RMSE, and relative root-squared error (RRMSE). Alongside error-based validation, two measures, namely NSE and Pearson's coefficient (R), were employed to assess the effectiveness of the constructed models. If a model can become more efficient, it indicates that it can make more accurate predictions. With respect to the CS-GEP model, the NSE value was 0.948, whereas that of the FS-GEP model was 0.917. Nevertheless, these values considerably increased to 0.969 and 0.929, respectively, in the CS-MEP and FS-MEP models. When examining the generated CS and FS models using Pearson's coefficient (R), the outcomes were comparable. Figure 18(a) and (b) illustrates the Taylor diagram comparing all the different prediction models. MEP models exhibit notably higher accuracy compared to GEP models in calculating the CS and FS of FR-MCM, which holds across the various types of models. Previous research has established that MEP models are the most accurate ML-based analysis technique for forecasting the CS and FS of FR-MCM. Their R^2 values are the highest, their error rates are the lowest, their efficiency is the highest, and their standard deviation is the lowest.



(a)



(b)

Figure 18: Taylor diagrams: (a) CS-models and (b) FS-models.

3.4 Sensitivity analysis

Examining how various input parameters impact CS and FS prediction for FR-MCM is the primary objective of this research. The predicted outcomes are highly associated with the input variables [70]. Figure 19 illustrates the impression of various variables on the CS and FS of FR-MCM, offering insights into the long-term development of mortar and construction materials. The extrapolation of CS for FR-MCM was most influenced by a curing age of 86%, followed by 2.0% fly ash, 4% MC, and 8% rice husk ash. Similarly, for FS prediction, the curing age (A) had the greatest impact at 80%, followed by fly ash at 5.0%, MC at 3.0%, and rice husk ash at 12.0%. The outcomes were proportional to the number of model parameters and data points utilized in the sensitivity analyses. It was discovered that varied input parameters, including the amounts of mortar mix used while applying the ML approach, had distinct effects on the analytical outcomes. By comparing the values of the data input features in Eqs (13) and (14) their relative significance was determined.

$$N_i = f_{\max}(x_i) - f_{\min}(x_i), \quad (13)$$

$$S_i = \frac{N_i}{\sum_{j=1}^n N_j}, \quad (14)$$

where $f_{\min}(x_i)$ is the minimum expected value and $f_{\max}(x_i)$ is the maximum predicted value for all i th outputs.

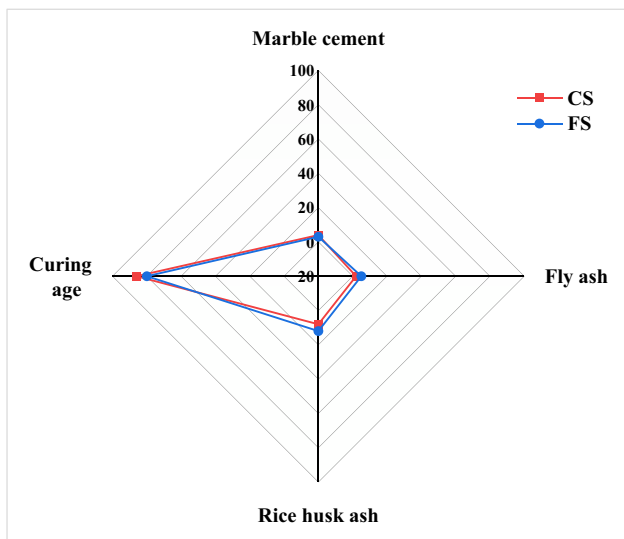


Figure 19: Radar plot for the sensitivity analysis.

4 Discussion

The findings of this research are specific to FR-MCM because the GEP and MEP models employed are tailored to handle values within a defined range of four input parameters. The accuracy of CS and FS projections is ensured by the use of consistent unit measurements and testing procedures across all models. MEP achieves higher accuracy than GEP due to its linear representation of multiple expressions within a single chromosome, which simplifies the evolutionary process. MEP's efficient genetic operations and ability to simultaneously evaluate multiple solutions enhance its optimization capabilities and predictive performance, leading to more accurate outcomes. For the most part, the models can only understand the mix's design and the impact of each input component by consulting mathematical formulas. All four inputs to the composite analysis, when used in any combination, will make the predicted models meaningless. Incorrect pairing with the training data could cause these models to fail to perform as anticipated. Inconsistent or out-of-date units of the input parameters put the models in danger of under or over-predicting the results. Consistently maintaining the unit sizes is crucial for the models to perform successfully. Improvements in energy efficiency, risk assessment, quality control, predictive maintenance, and material strength prediction are just a few of the many construction-related uses for ML models. These models must, however, conquer certain challenges. One potential pitfall is relying on human input, which might lead to erroneous outcomes due to human error. Possible directions for future study include incorporating IoT gadgets, creating hybrid models, making sustainability a top priority, using explainable AI techniques, and tailoring data collection and delivery to certain businesses. These strategies could significantly enhance ML-based solutions and address existing limitations. These efforts aim to enhance ML-based solutions and address current limitations effectively. Emerging technology has the potential to radically alter the construction industry by facilitating more efficient, transparent, and interpretable processes and better-informed decision-making. These enhancements have the potential to promote more sustainable practices, lessen project delays, and increase safety precautions. The results of this study have the potential to promote the use of more sustainable building practices and materials in the future.

5 Conclusions

This work aims to assess and predict the FS and CS of marble cement mortar (FR-MCM) that contains fly ash and rice husk ash using GEP and MEP. Using 500 CS and FS data sets from lab studies, the models were built, tested, and validated. A summary of the main points from the study is as follows:

- 1) The research showed that compared to other GEP models, MEP models had the best data prediction accuracy. The most effective model, MEP, achieved an R^2 value of 0.971 for the mechanical properties of FR-MCM.
- 2) The efficiency and predictability of the created models were confirmed by statistical testing; the MEP models were determined to be the most accurate in estimating the mechanical properties of FR-MCM.
- 3) MEP models had a lower standard deviation compared to GEP models, indicating that the MEP approach was more precise according to Taylor's diagram evaluation. Specifically, for CS prediction, the standard deviation for MEP models was 5.80 compared to 5.94 for GEP models, and for FS prediction, the standard deviation for MEP models was 0.865 compared to 0.868 for GEP models.
- 4) The sensitivity study revealed that curing age was the most important input parameter positively correlated with FR-MCM's mechanical properties.

The unique mathematical methodologies offered by GEP and MEP are essential for predicting features in other databases. Engineers and scientists leverage the mathematical models derived from this study to optimize, improve, and evaluate mortar mixture proportions effectively. These models serve to refine and enhance the formulation process, ensuring optimal performance and reliability in construction applications.

Acknowledgments: This work was supported by the Henan Province Science and Technology Research Project (No. 242102321100) and the National Natural Science Foundation of China (No. 41877251). The authors acknowledge the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Grant No. KFU241384). The authors extend their appreciation for the financial support that made this study possible.

Funding information: This research was funded by Henan Province Science and Technology Research Project (No. 242102321100); the National Natural Science Foundation of China (No. 41877251) and the Deanship of Scientific

Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Grant No. KFU241384).

Author contributions: H.S.: conceptualization, data acquisition, methodology, visualization, and writing – original draft. D.S.: supervision, resources, formal analysis, validation, writing, reviewing, and editing. M.N.A.: funding acquisition, investigation, supervision, project administration, writing, reviewing, and editing. S.U.A.: software, methodology, data acquisition, formal analysis, and writing – original draft. M.T.Q.: visualization, project administration, conceptualization, writing, reviewing, and editing. K.K.: validation, resources, writing, reviewing, and editing. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: The datasets generated and/or analyzed during the current study are available as supplementary materials to this article.

References

- [1] Castro-Alonso, M. J., L. E. Montañez-Hernandez, M. A. Sanchez-Muñoz, M. R. Macías Franco, R. Narayanasamy, and N. Balagurusamy. Microbially induced calcium carbonate precipitation (MICP) and its potential in bioconcrete: microbiological and molecular concepts. *Frontiers in Materials*, Vol. 6, 2019, id. 126.
- [2] Habert, G., S. A. Miller, V. M. John, J. L. Provis, A. Favier, A. Horvath, et al. Environmental impacts and decarbonization strategies in the cement and concrete industries. *Nature Reviews Earth & Environment*, Vol. 1, 2020, pp. 559–573.
- [3] Wang, D., M. N. Amin, K. Khan, S. Nazar, Y. Gamil, and T. Najeh. Comparing the efficacy of GEP and MEP algorithms in predicting concrete strength incorporating waste eggshell and waste glass powder. *Developments in the Built Environment*, Vol. 17, 2024, id. 100361.
- [4] Khan, M. and C. McNally. A holistic review on the contribution of civil engineers for driving sustainable concrete construction in the built environment. *Developments in the Built Environment*, Vol. 16, 2023, id. 100273.
- [5] Khan, Z., M. Umar, K. Shahzada, and A. Ali. Utilization of marble dust in fired clay bricks. *Journal of Environmental Monitoring*, Vol. 17, 2017, pp. 1–10.
- [6] Ahmad, T., M. Hussain, M. Iqbal, A. Ali, W. Manzoor, H. Bibi, et al. Environmental, energy, and water footprints of marble tile production chain in a life cycle perspective. *Sustainability*, Vol. 14, 2022, id. 8325.
- [7] Bilir, T., Ö. Karadağ, and B. F. Aygün. Waste marble powder. In *Sustainable concrete made with ashes and dust from different sources*, Woodhead Publishing Series in Civil and Structural Engineering, Cambridge, USA, 2022, pp. 479–506.

- [8] Fawad, M., F. Ullah, W. Shah, M. Irshad, Q. Mehmood, A. A. Tahir, et al. Impacts of the marble waste slurry on ground water quality and its reuse potential. *Fresenius Environmental Bulletin*, Vol. 30, 2021, pp. 2077–2086.
- [9] Kushwah, E. R. P. S. Scientific disposal system of marble slurry for clean and green environment. *International Journal Of Engineering Sciences & Research Technology*, Vol. 3, No. 10, 2014, pp. 500–503.
- [10] Zornoza, R., A. Faz, D. M. Carmona, S. Martínez-Martínez, and J. A. Acosta. Plant cover and soil biochemical properties in a mine tailing pond five years after application of marble wastes and organic amendments. *Pedosphere*, Vol. 22, 2012, pp. 22–32.
- [11] Neville, A. M. and J. J. Brooks. *Concrete technology*, Vol. 438, Longman Scientific & Technical, England, 1987.
- [12] Bouazza, N., A. El Mrihi, and A. Maâte. Geochemical assessment of limestone for cement manufacturing. *Procedia Technology*, Vol. 22, 2016, pp. 211–218.
- [13] Kore, S. D. and A. K. Vyas. Impact of marble waste as coarse aggregate on properties of lean cement concrete. *Case Studies in Construction Materials*, Vol. 4, 2016, pp. 85–92.
- [14] Khan, M. A., B. Khan, K. Shahzada, S. W. Khan, N. Wahab, and M. I. Ahmad. Conversion of waste marble powder into a binding material. *Civil Engineering Journal*, Vol. 6, 2020, pp. 431–445.
- [15] Khan, M. A., S. A. Khan, B. Khan, K. Shahzada, F. Althoey, and A. F. Deifalla. Investigating the feasibility of producing sustainable and compatible binder using marble waste, fly ash, and rice husk ash: A comprehensive research for material characteristics and production. *Results in Engineering*, Vol. 20, 2023, id. 101435.
- [16] Wesselsky, A. and O. M. Jensen. Synthesis of pure Portland cement phases. *Cement and Concrete Research*, Vol. 39, 2009, pp. 973–980.
- [17] Mtarfi, N. H., Z. Rais, and M. Taleb. Effect of clinker free lime and cement fineness on the cement physicochemical properties. *Journal of Materials and Environmental Science*, Vol. 8, 2017, pp. 2541–2548.
- [18] Sepehri, A. and M.-H. Sarrafzadeh. Effect of nitrifiers community on fouling mitigation and nitrification efficiency in a membrane bioreactor. *Chemical Engineering and Processing-Process Intensification*, Vol. 128, 2018, pp. 10–18.
- [19] Lao, J.-C., B.-T. Huang, L.-Y. Xu, M. Khan, Y. Fang, and J.-G. Dai. Seawater sea-sand Engineered Geopolymer Composites (EGC) with high strength and high ductility. *Cement and Concrete Composites*, Vol. 138, 2023, id. 104998.
- [20] Mohammed, A., R. Kurda, D. J. Armaghani, and M. Hasanipناه. Prediction of compressive strength of concrete modified with fly ash: Applications of neuro-swarm and neuro-imperialism models. *Computers and Concrete*, Vol. 27, 2021, pp. 489–512.
- [21] Sabri, M. Z. H. A., R. A. Malek, A. A. Omar, and K. N. Ismail. *Study of fly ash concrete exposed to elevated temperature*, 2022, pp. 645–650.
- [22] American Coal Ash, A. *Fly ash facts for highway engineers*, US Department of Transportation, Federal Highway Administration, 2003.
- [23] Fuad, M. Y. A., Z. Ismail, Z. A. M. Ishak, and A. K. M. Omar. Rice husk ash. *Plastics Additives: An AZ reference*, Polymer Science and Technology Series, Springer, Dordrecht, Vol. 1, 1998, pp. 561–566.
- [24] Kone, B., J. N. Mwere, and E. K. Ronoh. Experimental effect of cassava starch and rice husk ash on physical and mechanical properties of concrete. *International Journal of Engineering Trends and Technology*, Vol. 70, 2022, pp. 343–350.
- [25] Yuan, X., Y. Tian, W. Ahmad, A. Ahmad, K. I. Usanova, A. M. Mohamed, et al. Machine learning prediction models to evaluate the strength of recycled aggregate concrete. *Materials*, Vol. 15, 2022, id. 2823.
- [26] Singh, N., P. Kumar, and P. Goyal. Reviewing the behaviour of high volume fly ash based self compacting concrete. *Journal of Building Engineering*, Vol. 26, 2019, id. 100882.
- [27] Althoey, F. Compressive strength reduction of cement pastes exposed to sodium chloride solutions: Secondary ettringite formation. *Construction and Building Materials*, Vol. 299, 2021, id. 123965.
- [28] Awoyera, P. O. Nonlinear finite element analysis of steel fibre-reinforced concrete beam under static loading. *Journal of Engineering Science and Technology*, Vol. 11, 2016, pp. 1669–1677.
- [29] Amin, M. N., W. Ahmad, K. Khan, M. N. Al-Hashem, A. F. Deifalla, and A. Ahmad. Testing and modeling methods to experiment the flexural performance of cement mortar modified with eggshell powder. *Case Studies in Construction Materials*, Vol. 18, 2023, id. e01759.
- [30] Chen, Z., M. N. Amin, B. Iftikhar, W. Ahmad, F. Althoey, and F. Alsharari. Predictive modelling for the acid resistance of cement-based composites modified with eggshell and glass waste for sustainable and resilient building materials. *Journal of Building Engineering*, Vol. 76, 2023, id. 107325.
- [31] Nazar, S., J. Yang, M. N. Amin, K. Khan, M. F. Javed, and F. Althoey. Formulation of estimation models for the compressive strength of concrete mixed with nanosilica and carbon nanotubes. *Developments in the Built Environment*, Vol. 13, 2023, id. 100113.
- [32] Jaf, D. K. I., P. I. Abdulrahman, A. S. Mohammed, R. Kurda, S. M. A. Qaidi, and P. G. Asteris. Machine learning techniques and multi-scale models to evaluate the impact of silicon dioxide (SiO₂) and calcium oxide (CaO) in fly ash on the compressive strength of green concrete. *Construction and Building Materials*, Vol. 400, 2023, id. 132604.
- [33] Salih, A., S. Rafiq, P. Sihag, K. Ghafor, W. Mahmood, and W. Sarwar. Systematic multiscale models to predict the effect of high-volume fly ash on the maximum compression stress of cement-based mortar at various water/cement ratios and curing times. *Measurement*, Vol. 171, 2021, id. 108819.
- [34] Javed, M. F., M. N. Amin, M. I. Shah, K. Khan, B. Iftikhar, F. Farooq, et al. Applications of gene expression programming and regression techniques for estimating compressive strength of bagasse ash based concrete. *Crystals*, Vol. 10, 2020, id. 737.
- [35] Abdalla, A. and A. S. Mohammed. Hybrid MARS-, MEP-, and ANN-based prediction for modeling the compressive strength of cement mortar with various sand size and clay mineral metakaolin content. *Archives of Civil and Mechanical Engineering*, Vol. 22, 2022, id. 194.
- [36] Abdalla, A. and A. Salih. Implementation of multi-expression programming (MEP), artificial neural network (ANN), and M5P-tree to forecast the compression strength cement-based mortar modified by calcium hydroxide at different mix proportions and curing ages. *Innovative Infrastructure Solutions*, Vol. 7, 2022, id. 153.
- [37] Puth, M.-T., M. Neuhäuser, and G. D. Ruxton. Effective use of Pearson's product-moment correlation coefficient. *Animal behaviour*, Vol. 93, 2014, pp. 183–189.
- [38] Pearson, K. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Vol. 50, 1900, pp. 157–175.
- [39] Gravier, J., V. Vignal, S. Bissey-Breton, and J. Farre. The use of linear regression methods and Pearson's correlation matrix to identify mechanical-physical-chemical parameters controlling the micro-

- electrochemical behaviour of machined copper. *Corrosion Science*, Vol. 50, 2008, pp. 2885–2894.
- [40] Iqbal, M. F., M. F. Javed, M. Rauf, I. Azim, M. Ashraf, J. Yang, et al. Sustainable utilization of foundry waste: Forecasting mechanical properties of foundry sand based concrete using multi-expression programming. *Science of The Total Environment*, Vol. 780, 2021, id. 146524.
- [41] Jalal, F. E., Y. Xu, M. Iqbal, B. Jamhiri, and M. F. Javed. Predicting the compaction characteristics of expansive soils using two genetic programming-based algorithms. *Transportation Geotechnics*, Vol. 30, 2021, id. 100608.
- [42] Lee, B. C. and D. M. Brooks. Accurate and efficient regression modeling for microarchitectural performance and power prediction. *ACM SIGOPS Operating Systems Review*, Vol. 40, 2006, pp. 185–194.
- [43] Amin, M. N., W. Ahmad, K. Khan, and A. F. Deifalla. Optimizing compressive strength prediction models for rice husk ash concrete with evolutionary machine intelligence techniques. *Case Studies in Construction Materials*, Vol. 18, 2023, id. e02102.
- [44] Holland, J. H. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, MIT Press, Cambridge, Massachusetts, USA, 1992.
- [45] Koza, J. On the programming of computers by means of natural selection. *Genetic programming*, MIT Press, Cambridge, Massachusetts, United States, 1992.
- [46] Gholampour, A., T. Ozbakkaloglu, and R. Hassanli. Behavior of rubberized concrete under active confinement. *Construction and Building Materials*, Vol. 138, 2017, pp. 372–382.
- [47] Topcu, I. B. and M. Saridemir. Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. *Computational Materials Science*, Vol. 41, 2008, pp. 305–311.
- [48] Ferreira, C. *Gene expression programming: Mathematical modeling by an artificial intelligence*, Berlin, Germany, Springer, 2006. p. 21.
- [49] Gandomi, A. H., G. J. Yun, and A. H. Alavi. An evolutionary approach for modeling of shear strength of RC deep beams. *Materials and Structures*, Vol. 46, 2013, pp. 2109–2119.
- [50] Gandomi, A. H., S. K. Babanajad, A. H. Alavi, and Y. Farnam. Novel approach to strength modeling of concrete under triaxial compression. *Journal of Materials in Civil Engineering*, Vol. 24, 2012, pp. 1132–1143.
- [51] Wang, H.-L. and Z.-Y. Yin. High performance prediction of soil compaction parameters using multi expression programming. *Engineering Geology*, Vol. 276, 2020, id. 105758.
- [52] Oltean, M. and C. Grosan. A comparison of several linear genetic programming techniques. *Complex Systems*, Vol. 14, 2003, pp. 285–314.
- [53] Fallahpour, A., E. U. Olugu, and S. N. Musa. A hybrid model for supplier selection: Integration of AHP and multi expression programming (MEP). *Neural Computing and Applications*, Vol. 28, 2017, pp. 499–504.
- [54] Alavi, A. H., A. H. Gandomi, M. G. Sahab, and M. Gandomi. Multi expression programming: a new approach to formulation of soil classification. *Engineering with Computers*, Vol. 26, 2010, pp. 111–118.
- [55] Mohammadzadeh S, D., S.-F. Kazemi, A. Mosavi, E. Nasseralshariati, and J. H. M. Tah. Prediction of compression index of fine-grained soils using a gene expression programming model. *Infrastructures*, Vol. 4, 2019, id. 26.
- [56] Grosan, C. and A. Abraham. Stock market modeling using genetic programming ensembles. In *Genetic systems programming: Theory and experiences*, Berlin, Germany, Springer, 2006, pp. 131–146.
- [57] Oltean, M. and D. Dumitrescu. Multi expression programming. *Journal of Genetic Programming and Evolvable Machines*, 2002.
- [58] Iqbal, M. F., Q.-f. Liu, I. Azim, X. Zhu, J. Yang, M. F. Javed, et al. Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming. *Journal of hazardous materials*, Vol. 384, 2020, id. 121322.
- [59] Shahin, M. A. *Genetic Programming for Modelling of Geotechnical Engineering Systems*, Springer International Publishing, Switzerland, 2015.
- [60] Çanakçı, H., A. Baykasoğlu, and H. Güllü. Prediction of compressive and tensile strength of Gaziantep basalts via neural networks and gene expression programming. *Neural Computing and Applications*, Vol. 18, 2009, pp. 1031–1041.
- [61] Alade, I. O., M. A. Abd Rahman, and T. A. Saleh. Predicting the specific heat capacity of alumina/ethylene glycol nanofluids using support vector regression model optimized with Bayesian algorithm. *Solar Energy*, Vol. 183, 2019, pp. 74–82.
- [62] Alade, I. O., A. Bagudu, T. A. Oyeohan, M. A. Abd Rahman, T. A. Saleh, and S. O. Olatunji. Estimating the refractive index of oxygenated and deoxygenated hemoglobin using genetic algorithm–support vector regression model. *Computer Methods and Programs in Biomedicine*, Vol. 163, 2018, pp. 135–142.
- [63] Zhang, W., R. Zhang, C. Wu, A. T. C. Goh, S. Lacasse, Z. Liu, et al. State-of-the-art review of soft computing applications in underground excavations. *Geoscience Frontiers*, Vol. 11, 2020, pp. 1095–1106.
- [64] Alavi, A. H., A. H. Gandomi, H. C. Nejad, A. Mollahasani, and A. Rashed. Design equations for prediction of pressuremeter soil deformation moduli utilizing expression programming systems. *Neural Computing and Applications*, Vol. 23, 2013, pp. 1771–1786.
- [65] Kisi, O., J. Shiri, and M. Tombul. Modeling rainfall-runoff process using soft computing techniques. *Computers & Geosciences*, Vol. 51, 2013, pp. 108–117.
- [66] Alade, I. O., M. A. Abd Rahman, and T. A. Saleh. Modeling and prediction of the specific heat capacity of Al₂O₃/water nanofluids using hybrid genetic algorithm/support vector regression model. *Nano-Structures & Nano-Objects*, Vol. 17, 2019, pp. 103–111.
- [67] Shahin, M. A. Use of evolutionary computing for modelling some complex problems in geotechnical engineering. *Geomechanics and Geoengineering*, Vol. 10, 2015, pp. 109–125.
- [68] Band, S. S., E. Heggy, S. M. Bateni, H. Karami, M. Rabiee, S. Samadianfard, et al. Groundwater level prediction in arid areas using wavelet analysis and Gaussian process regression. *Engineering Applications of Computational Fluid Mechanics*, Vol. 15, 2021, pp. 1147–1158.
- [69] Taylor, K. E. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, Vol. 106, 2001, pp. 7183–7192.
- [70] Ahmad, A., K. A. Ostrowski, M. Maślak, F. Farooq, I. Mehmood, and A. Nafees. Comparative study of supervised machine learning algorithms for predicting the compressive strength of concrete at high temperature. *Materials*, Vol. 14, 2021, id. 4222.