scientific reports

Check for updates

OPEN

Experimental analysis and gene expression programming optimization of sustainable concrete containing mineral fillers

Ayesha Rauf[1], Usama Asif[1,2]✉, Kennedy Onyelowe[3,4]✉, Muhammad Faisal Javed[5,6] & Hisham Alabduljabbar[7]

Rapid urbanization has led to a high demand for concrete, causing a significant depletion of vital natural resources, notably river sand, which is crucial in the manufacturing process of concrete. As a result, there is a growing need for environmentally sustainable alternatives to fine aggregate in concrete. Quarry dust (QD) has evolved as a viable and ecologically friendly substitute in response to this demand. In the past, limited experimental investigations and only conventional modeling techniques were used to promote sustainable mineral fillers in concrete. This study proposed a robust soft computing technique using gene-expression programming (GEP) to enhance the usability of sustainable alternatives in concrete. Initially, an experimental study was carried out to examine the feasibility and mechanical characteristics of concrete made from materials including quarry dust and superplasticizer as a partial replacement for fine aggregate. Ten mixed proportions with various proportions (0%, 20%, 40%, and 60%) of quarry dust were used to make M15 and M20 grades of concrete. A series of experimental tests, such as workability, compressive strength (CS), and tensile strength (TS), were conducted to examine the fresh and hardened properties of modified concrete. The established database from the experimental investigations was then used to develop machine learning (ML) models using GEP. The outcomes of the GEP models were validated by comparing them with multi-linear regression (MLR) models and using various statistical metrics such as root mean squared error (RMSE), performance index (PI), correlation coefficient (R), and external validation methods. Finally, sensitivity analysis was performed to investigate the influence of ingredients such as mineral fillers, superplasticizers, and others on the mechanical properties of concrete. To enhance the practical usage of the study, a graphical user interface (GUI) was also created. The study revealed that 40% of the replacement of fine aggregates with mineral filler and superplasticizer shows the optimum properties. GEP models outperformed MLR, achieving R^2 values of 0.96 in CS and 0.92 in TS, compared to MLR's lower values of 0.85 in CS and 0.81 in TS. The proposed GEP equations and user-friendly GUI can be used to develop the pre-mix design of concrete using quarry dust and superplasticizers.

Keywords Quarry dust, Compressive strength, Split Tensile Strength, Superplasticizer, Gene expression programming

Abbreviations

ACI	American Concrete Institute
ASTM	American Society for Testing and Materials
CA	Coarse Aggregate
AIV	Aggregate Impact Value

[1]Department of Civil Engineering, Nazarbayev University, Nazarbayev, Kazakhstan. [2]Department of Civil Engineering, COMSATS University Islamabad, Abbottabad Campus, Abbottabad 22060, Pakistan. [3]Department of Civil Engineering, Michael Okpara University of Agriculture, Umudike 440109, Nigeria. [4]Department of Civil Engineering, Kampala International University, Western Campus, Bushenyi District, Kampala, Uganda. [5]Department of Civil Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Pakistan. [6]Western Caspian University, Baku, Azerbaijan. [7]Department of Civil Engineering, College of Engineering in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia. ✉email: usama.asif@nu.edu.kz; kennedychibuzor@kiu.ac.ug

ML	Machine learning
FA	Fine Aggregate
S	Sand
GEP	Gene expression programming
SSD	Saturated surface dry
R	Coefficient of correlation
PI	Performance Index
RMSE	Root means squared error
OF	Objective function
ANN	Artificial neural network
CS	Compressive strength
TS	Tensile strength
ANFIS	Adaptive neuro-fuzzy inference system
SVM	Support vector machine
QD	Quarry dust

Concrete is a durable building material made by combining aggregates, binders, and water. The water reacts with the binder in a chemical process, creating a solid, cohesive structure. Concrete is experiencing an increase in demand on a global scale due to its advantageous characteristics, such as its simplicity of manufacture, adherence to common design criteria, and relatively affordable cost[1,2]. Consequently, concrete has emerged as the construction material with the highest degree of utilization. The combination of fine and coarse aggregate comprises approximately 65–80% of the total quantity of traditional concrete. Fine aggregates make up about 20–30% of the total concrete volume, with river sand commonly used for this purpose. Globally, the construction industry consumes approximately 5-7.5 billion tons of river sand each year[3]. The extraction of sand occurs through the process of mining from the bottoms of rivers, which can cause significant environmental issues, including the depletion of riverbeds, harm to local flora and fauna, reduced stream water storage capacity, lowered water table levels, destabilized river bridge foundations, and accelerated bedrock erosion. Thus, finding alternative materials for fine aggregate is crucial for meeting construction needs efficiently and economically[4,5].

The primary objective of this study is to promote the use of quarry dust as a sustainable alternative to natural sand by developing predictive mathematical models to estimate the strength properties of concrete modified with mineral fillers and superplasticizers. Previously, various alternative materials, including recycled concrete aggregate, foundry waste, bottom ash, and stone waste, have been employed successfully as a partial substitute for natural aggregates in concrete[6–9]. Recently, there has been a rising interest in utilizing quarry dust, a by-product generated during the quarrying process, as a partial substitute for aggregates in concrete[10,11]. On a global scale, the rock industry generates a yearly productivity of 68 million tons of finished products. This production level presents notable difficulties in waste disposal and imposes a substantial strain on transportation systems and the natural environment, mostly due to the potentially hazardous materials[12]. Moreover, it is worth noting that the stone-crushing industry experiences a significant waste generation ranging from 15 to 25% of the total quantity produced[13], which leads to tons of colloidal waste annually. The management of this fine waste poses an enormous threat to the environment on a global scale, with the predominant method of disposal being landfilling[14]. However, the use of quarry dust as a fine aggregate substitute can preserve natural resources and also provide acceptable-strength concrete with lower environmental concerns.

For instance, N. Khan & Chandrakar et al.[15] investigated the effects of substituting fine aggregates with quarry dust in concrete, with varying proportions ranging from 0 to 100% at increments of 10%. The study employed concrete grades M20 and M25, maintaining a uniform slump of 60 mm throughout the experimentation. CS tests were conducted at both 7 and 28 days under ambient conditions. Additionally, the cubes were subjected to a temperature of 100 °C to evaluate any potential reduction in strength. The results indicated that the highest CS was attained when 50% of the sand was replaced. The concrete exhibited strength properties above the acceptable values even when subjected to elevated temperatures, demonstrating that quarry dust holds potential as a viable alternative to natural fine aggregates at a 50% substitution ratio.

Another study reported that incorporating 30% quarry dust improved both the durability and mechanical properties of concrete[16]. Similarly, Kankam et al.[17] investigated quarry dust as a sand replacement in proportions of 0%, 25%, and 100%, targeting specific CS grades (C25, C30, C35, C40, C45). Their findings revealed that using 25% quarry dust improved strength by 7.10% and overall performance by 16.19%[18]. Other studies using quarry dust at substitution levels of 0-25% with various waste materials also reported enhancements in CS, flexural strength (FS), and impact resistance[13]. However, the introduction of quarry dust up to 30% led to a reduction in slump value, suggesting potential workability challenges[19].

Economically, quarry dust-based concrete was found to reduce costs while maintaining comparable strength to conventional mixes[20]. Specifically, the CS, tensile strength (TS), and modulus of rupture of quarry dust-based concrete were approximately 14% higher than conventional concrete[21]. However, a critical gap remains in the literature regarding the optimal replacement level of quarry dust, as existing studies often report conflicting results or lack comprehensive analysis. Further research is needed to establish the ideal substitution ratio for maximizing both strength and durability in concrete.

Therefore, it is imperative to construct robust models to create a mixed design of concrete containing sustainable additive materials. Recently, machine learning (ML) models have been extensively used to address issues related to structural materials. Various ML techniques, including adaptive neuro-fuzzy inference system (ANFIS), GEP, support vector machine (SVM), artificial neural network (ANN), deep learning (DL), and random forest (RF) have been employed to develop reliable and robust models to anticipate the properties of concrete incorporating additive materials[22–27]. All of the above ML methods involve the process of "training" the solution

by utilizing pre-existing data. Among these techniques, neural network-based models (such as ANN, ANFIS) and SVM are widely used due to their capability to identify complex arrangements and yield a comprehensive pattern as an outcome[28]. However, these models do not provide the empirical formulation for outputs. The complicated architecture of the neural network-based model limits its widespread practical application[29]. For instance, Ebid & Deifalla et al.[30] developed an ANN model for forecasting the punching shear strength of concrete. Overfitting was observed when the anticipated values of ANN were compared to the design codes. Due to their complex composition, they pose challenges in terms of manipulation. Moreover, multicollinearity is another major concern that is often seen in these models.

In recent times, genetic programming (GP) has been distinguished from other ML models by addressing the abovementioned issues due to its independence from pre-existing relationships[31]. In the more comprehensive variant known as GEP, a concise program is created utilizing chromosomes of a particular length. The GEP methodology extends its use by generating mathematical formulations that possess practical applications across many domains. It functions as a viable substitute for conventional ML forecasting methods and demonstrates usefulness across several areas within the field of civil engineering[32,33]. Therefore, considering the widespread use of GEP-based models and their undeniable benefits, this study also used genetic-based models to predict the mechanical properties of concrete modified with mineral filler and superplasticizer. Successful implementation of GEP in the literature can be observed in areas such as the evaluation of properties of concrete incorporating materials like silica fume, fly ash, recycled aggregates, rice husk ash, and other mineral additives[34,35]. For instance, Iftikhar et al.[36] used GEP as a computational technique to predict the CS of paver blocks made with plastic. Likewise, Nazari & Torgal[32] employed GEP to estimate the CS of geopolymer concrete. The efficacy of the models was assessed by using various statistical conditions. The findings of the study indicated that GEP methods provide a robust and reliable predictive model that can be used to enhance the use of mineral additives in concrete. Therefore, considering the widespread use of GEP-based models and their undeniable benefits, this study also used genetic-based models to predict the mechanical properties of concrete modified with mineral filler and superplasticizer.

Considering the above discussion, the provision of reliable equations that establish a connection between the mixed proportion and mechanical qualities of modified concrete with quarry dust has the potential to yield cost and time savings, as well as simplify its sustainable adoption within the construction industry. Although previously various studies have investigated the partial replacement of natural sand with quarry dust, they often lack a holistic approach. For instance, some studies have focused solely on experimental investigations without leveraging advanced modeling techniques to generalize findings beyond specific conditions[10,34]. Others have used machine learning models to predict the mechanical properties of concrete with mineral fillers but have not integrated these predictions with comprehensive experimental data[35]. Furthermore, to date, no study has combined experimental results with machine learning-based predictive models while offering a user-friendly GUI for practical applications, limiting the accessibility and implementation of such research in real-world scenarios.

To address this critical research gap, this study employs GEP to develop reliable and straightforward predictive mathematical formulas and a GUI based on an experimental database for anticipating the strength properties of concrete containing mineral fillers and superplasticizers. The primary objective of this research is to preserve natural sand resources by promoting the use of quarry dust as a sustainable alternative. First, an experimental study was conducted using varying replacement levels of quarry dust (0%, 20%, 40%, and 60%) to assess the workability and mechanical properties of modified concrete. The experimental database was then employed to train the GEP model, and its predictive capability was evaluated using several validation principles. Finally, a sensitivity analysis was performed to examine the effect of input ingredients on the mechanical characteristics of the concrete, providing insights for optimizing mix designs for future applications.

Experimental analysis

Materials

In this study, six main ingredients, such as superplasticizer, water, cement, coarse aggregates, sand, and quarry dust, were used in the production of concrete. A detailed description of the material used is described below.

Binding material

Lucky cement was used as a binding material, which is very popular in Pakistan. The chemical composition of the cement is shown in Table 1. The other basic properties of the cement, such as fineness modulus, standard consistency, and setting time, were determined in the lab using ASTM standards, as shown in Table 2. The standard consistency of Lucy cement was achieved by 32%, which is in the acceptable range. The initial and final setting was found to be 1 h and 5 h, respectively.

Superplasticizer

Superplasticizers are the ingredients used to alter the qualities of concrete, such as workability, flow, and setting time. Sika-ViscoCrete was used in this study, which is a high-range water-reducing and decelerating superplasticizer[36].

SO_3	TiO_2	CaO	SiO_2	MgO	K_2O	Al_2O_3	Fe_2O_3	Na_2O	LOI
2.52	1.35	58.66	21.27	1.94	0	8.79	3.56	0	1.48

Table 1. Chemical composition of cement.

Cement			
Type of Test	Result	Specification	Standard
Fineness modulus	5%	≤10%	ASTM C786
Initial Setting Time	60 min	≥30 min	ASTM C 191-08
Normal Consistency	32%	26-33%	ASTM C 191-08
Final Setting Time	300 min	≤600 min	ASTM C 191-08
Coarse Aggregate			
Type of Test	Result	Specification	Standard
Impact Value	11.01%	10–20% (Strong)	ASTM D-5874
Elongated particles	11.50%	≤15%	ASTM D 4791-99
Flaky particles	14.50%	≤15%	ASTM D 4791-99
Fineness modulus	5.85	5.5-8	ASTM C-316-05
Water Absorption	1%	0.1 to 2%	ASTM C-127-04
Specific Gravity	2.52	2.5-3	ASTM C-127-04
Sand			
Type of Test	Result	Specification	Standard
Fineness modulus	2.5	2.3–3.1	ASTM C-117-05
Water Absorption	1.01%	≤2%	ASTM C-127/128
Specific Gravity	2.67	2 to 3	ASTM C-127/128
Quarry Dust			
Type of Test	Result	Specification	Standard
Fineness modulus	2.83	2.3–3.1	ASTM C-117-05
Water Absorption	2.81	≤2%	ASTM C-127/128
Specific Gravity	1.41%	2 to 3	ASTM C-127/128

Table 2. Summary of test results of materials.

Aggregates

Quarry dust and river sand were used as fine aggregates in this study. When mountains are broken down, all sizes of rocks are crushed into little stones, and the dust created during this process is known as quarry dust. It may be utilized as a cost-effective building material as well as a sand substitute. Quarry dust was gathered from Pakistan's Sargodha Crush Plant. Similarly, locally available gravel sizes ranging from 9.5 to 37.5 mm were used as coarse aggregates. The physical properties of quarry dust, sand, and coarse aggregates, including particle size distribution, impact value, and water absorption, were examined in the laboratory by following ASTM standards as presented in Table 1. The sieve analysis was also performed on these materials, and results are shown in Fig. 1. The Fineness modulus was calculated as 2.5, 2.83, and 5.85 for sand, quarry dust, and gravel, respectively. The proportion of flaky particles in the mixes is 27.5%, which is below the established threshold of 30% for flaky particles. If the proportion of flaky particles exceeds 30%, the aggregate is considered unsuitable for the intended use. In the same way, the proportion of elongated particles is recorded at 24%, while the prescribed threshold for elongated particles in the mixtures stands at 45%. Therefore, if the elongated particles exceed a threshold of 45%, the aggregate is deemed unsuitable for its intended application. The determined aggregate impact value was observed to be 11.01%, which falls within the acceptable range. According to ASTM standards, the impact value for aggregates utilized in concrete, excluding the wearing surface, should not exceed 45%. However, for concrete employed on the wearing surface, the impact value shouldn't exceed 30%. Therefore, the qualities of aggregate materials are within the permissible range.

Mix design and specimen preparation

This study used two different grades of concrete, namely M15 (1:2:4) and M20 (1:1.5:3), with different proportions of quarry dust replacement (0%, 20%, 40%, and 60%) to determine the workability, CS, and TS of concrete. A control mix was designed without replacing sand with quarry dust, and a total of 90 specimens were prepared for each case. A detailed description of the experimental mix design is presented in Table 3. The cube sample sized 150 × 150 × 150 mm was utilized for conducting the CS testing. Likewise, cylindrical moulds with a diameter of 150 mm and a height of 300 mm were utilized to determine the TS testing. The specimens were extracted from the moulds after 24 h and subsequently immersed in water till the time of testing.

Testing

Workability

The purpose of this test has been to determine the consistency and workability of newly prepared concrete using ASTM C143[37]. First, the moulds inside the surface were cleaned, greased, and put on a smooth, non-porous horizontal base plate. After that, three equal layers of concrete were poured into the mould. The tamping method was used to compress each layer with the use of a tamping rod, which was used to provide 25 strokes in a regular pattern over the mold's cross-sectional area. Following that, the additional concrete was removed, and

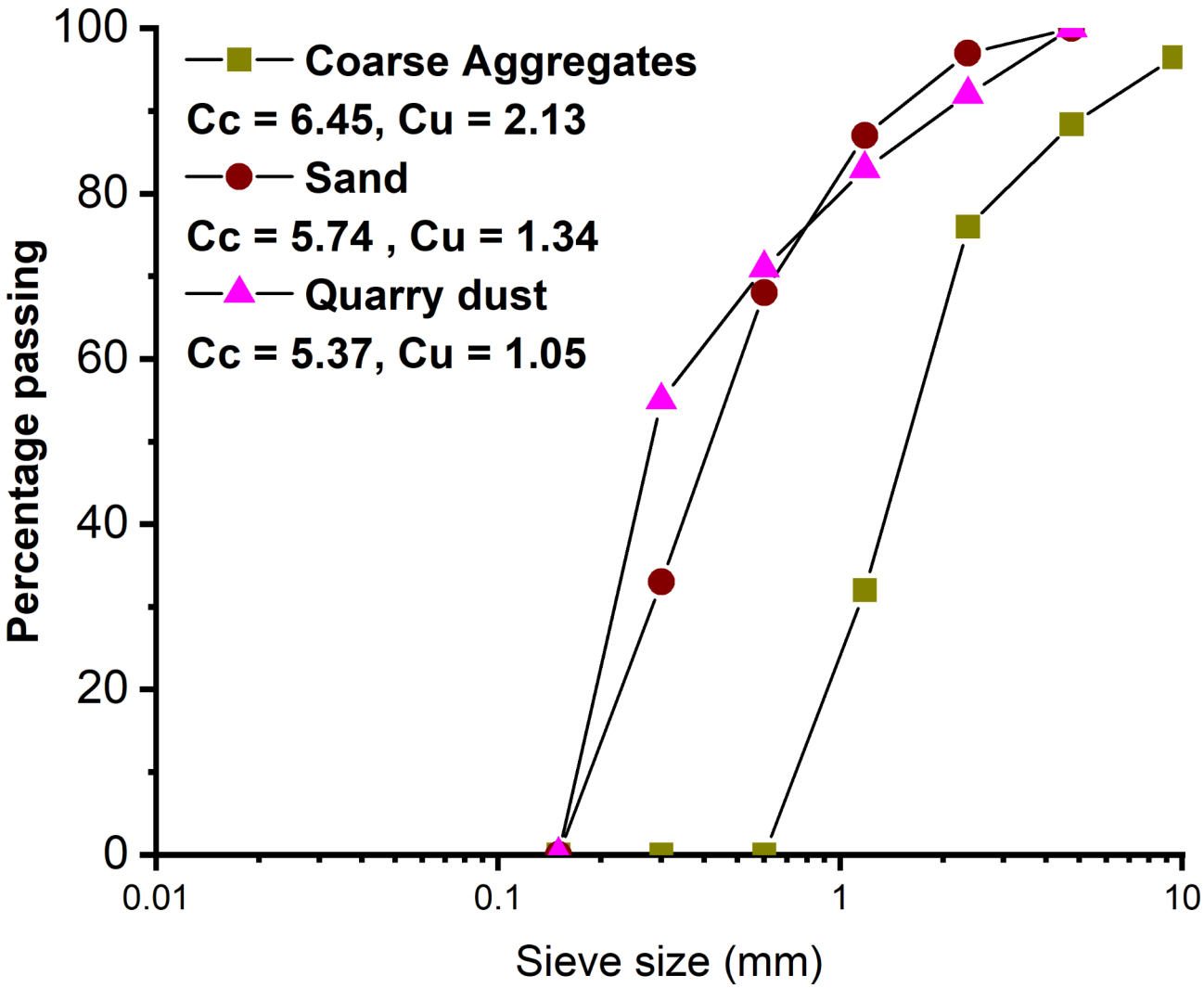

Fig. 1. Sieve analysis for aggregate materials.

Grade 15						
Design	Cement (kg/m^3)	Water (kg/m^3)	Coarse aggregates (kg/m^3)	Quarry dust (kg/m^3)	Sand (kg/m^3)	Superplasticizer (kg/m^3)
M15-CM	293	137	1181	0	713	0
M15-QDO	293	137	1181	0	713	5.86
M15-QD20	293	137	1181	143	570	5.86
M15-QD40	293	137	1181	285	428	5.86
M15-QD60	293	137	1181	428	285	5.86
Grade 20						
Design	Cement (kg/m^3)	Water (kg/m^3)	Coarse aggregates (kg/m^3)	Quarry dust (kg/m^3)	Sand (kg/m^3)	Superplasticizer (kg/m^3)
M20-CM	316	142	1276	0	690	0
M20-QDO	316	142	1276	0	690	5.86
M20-QD20	316	142	1276	138	575	5.86
M20-QD40	316	142	1276	276	437	5.86
M20-QD60	316	142	1276	414	299	5.86

Table 3. Mix design for M15 and M20 concrete.

the surface was made uniform with a trowel. The water which has infiltrated the interface between the mould and the base plate or the mortar was remediated. Finally, immediately raise the concrete and drag it vertically out of the mold. In addition, the slump value was calculated as a decline in concrete height relative to the mold height. M15 had a W/C ratio of 0.47, whereas M20 had a W/C ratio of 0.45.

Compressive strength

To determine the performance of different specimens under compression forces, a compression testing machine (CTM) was employed in accordance with the guidelines outlined in ASTM C192[38]. The three lifts with freshly mixed concrete cylinders were filled, and the cylinders were tamped about 25 times in each lift using a tamping rod. Each lift was also softly tapped about 10–15 times with a hammer. Finally, the tamping rod was used to conclude the preparation of cylindrical specimens. The surplus concrete was knocked off, and the surface was smoothed with a steel trowel. A total of 90 samples were prepared for the purpose of conducting CS tests at three different curing durations: 7 days, 14 days, and 28 days. The concrete specimens were initially subjected to ambient temperature conditions and left undisturbed for 24 h to reach a state of stability. Additionally, measures were taken to cover the upper surface of the specimens in order to avoid any loss of moisture. Afterward, the moulds were extracted from the specimens and then placed in the curing chamber for storage until the time of testing. The specimen was securely covered and positioned within the central testing apparatus, where it was subjected to incremental loading until the point of failure was reached. The most common kind of compression failure found in the specimen was characterized by shearing along a singular plane, as depicted in Fig. 2**(a)**.

Split tensile strength

The split tensile test is commonly known as the "Brazilian Test" in literature. ASTM C39/C39M[39] was used to prepare cylindrical shape specimens. A total of 90 cylinders were prepared for TS testing at 7, 14, and 28 days of curing. The test method involved positioning a cylindrical specimen in a horizontal orientation between the loading plates of a CTM. Subsequently, a load was imposed on the specimen until it experienced failure along its vertical dimension (diameter). The most commonly noted failure was axial splitting, as shown in Fig. 2**(b)**. The TS of the specimens was calculated using the following formula.

$$TS = \frac{2P}{\pi LD} \tag{1}$$

Where P = Load (MPa), L = Specimen Length, and D = Diameter.

Fig. 2. Laboratory testing (**a**) CS testing and failure plan, and (**b**) TS testing and failure plan.

Machine learning analysis

The present study employed GEP to anticipate the CS and TS of concrete modified with mineral filler. The process of development of ML models is depicted in Fig. 3.

GEP

GEP is a form of evolutionary programming technique that utilizes genetic algorithms. It was first proposed by Candida Ferreira[40]. The software is specifically developed for the purpose of symbolic regression, wherein it facilitates the automated evolution of computer programs capable of predicting or modeling correlations within a certain dataset. The concept of GEP draws inspiration from the mechanisms by which genetic information is manifested in live creatures, wherein genes undergo transcription and translation processes to generate functional proteins. The Genetic Programming (GP) theory implies the iterative evolution of a population of computer programs, which are encoded (trained) as genes. This process aims to find the most favorable program that efficiently addresses a specific problem and precisely matches a given database.

The predictive procedure of GEP includes multiple essential phases, which are executed sequentially, as shown in Fig. 4[31]. Initially, a population of expression trees (ETs) is produced by an arbitrary (random) process. Following that, the evaluation of these trees is performed based on their capability to accurately anticipate the target variable within the given database. The selection procedure implies identifying the most proficient gene (parent cell) to form the succeeding generation. Unique programs are developed, and the selection method is conducted using several methodologies, including crossover, mutation, recombination, and transposition. During the crossover stage of the GEP algorithm, the genetic statistics of two parental solutions are linked to produce one or more offspring responses. This process helps in the identification of the most optimal solutions, as represented in Fig. 5**(a)**. The mutation is a biological trend described by the occurrence of small and spontaneous changes in an organism's genetic genome (chromosome), as depicted in Fig. 5**(b)**. This method promotes the advancement of diversity within a population, hence enabling the algorithm to explore unique and inventive solutions (responses) within the designated search space area. The iterative method is continued until the desired fitness level is reached, resulting in the maximum possible accuracy for predicting the target (desired) variable using the provided database. Genetic programming is widely acknowledged for its effectiveness in symbolic regression, data analysis, and function optimization, making it a valuable asset in diverse domains such as ML and data processing.

Database

Before developing the model, a comprehensive database was constructed using experimental results for the CS and TS of concrete modified with quarry dust. Seven key input parameters were identified as significant

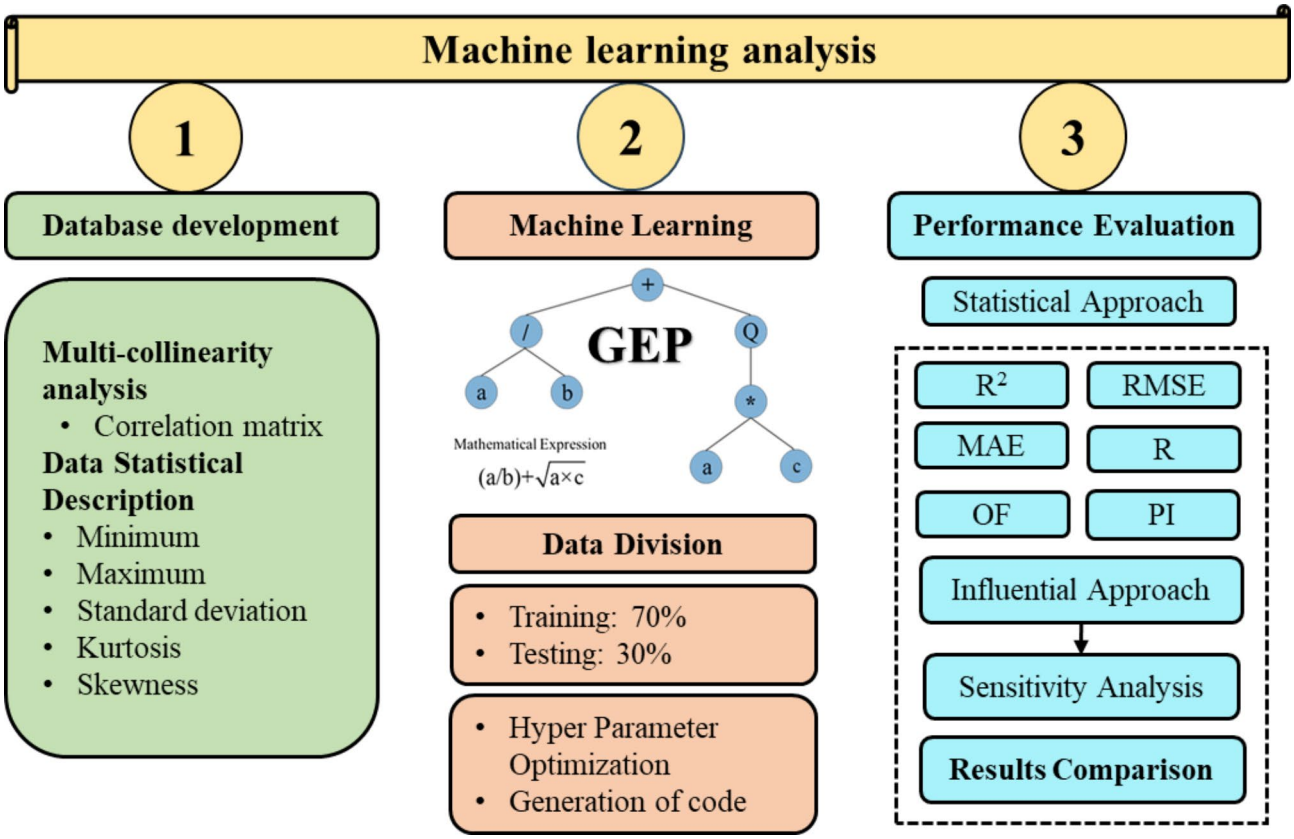

Fig. 3. The sequential process ML analysis used in this study.

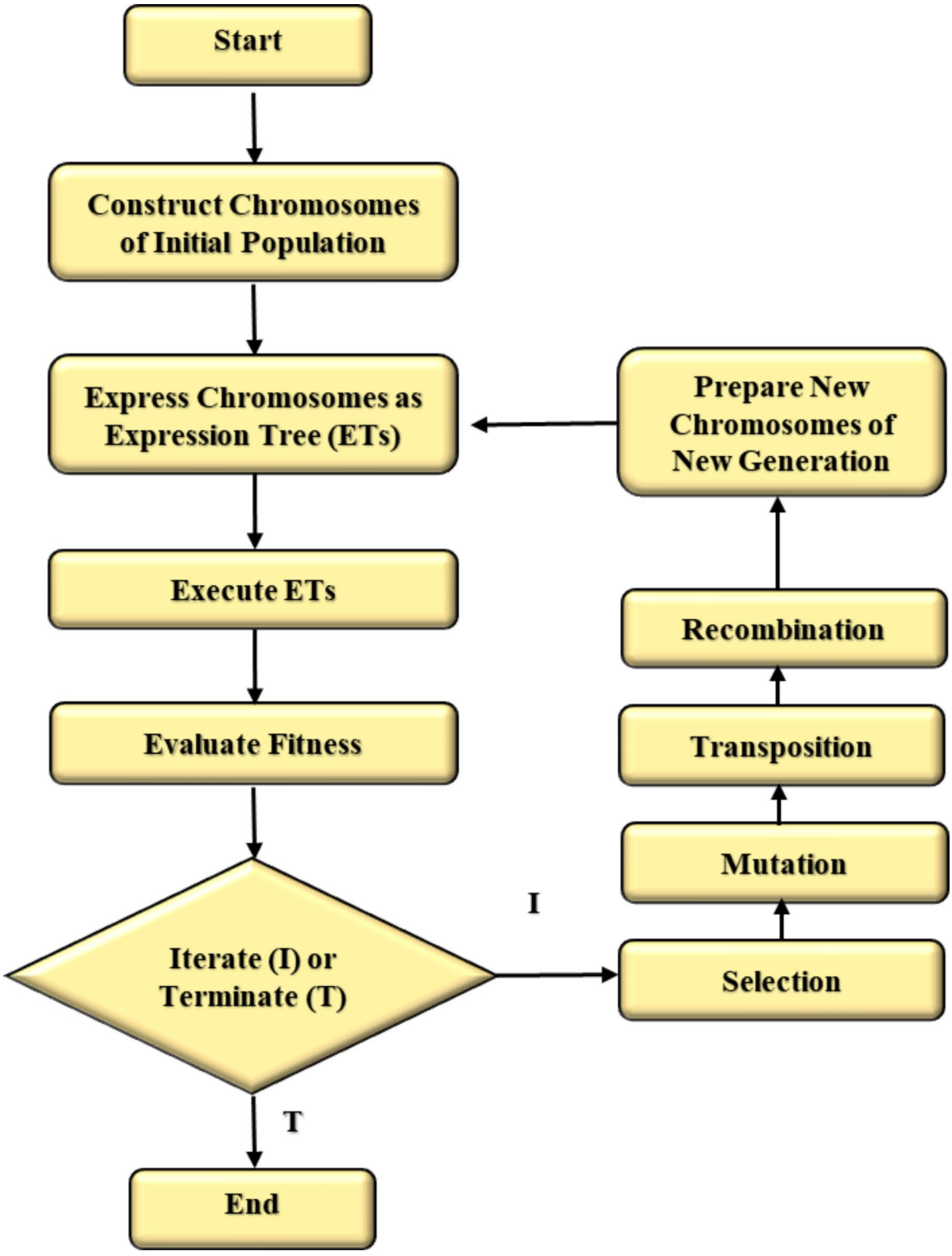

Fig. 4. The process of development of the GEP model used in this study[41].

contributors, including cement (C), coarse aggregates (CA), quarry dust (QD), superplasticizer (SP), sand (S), water (W), and age (A). These parameters were analyzed to model the CS and TS of the concrete using Gene Expression Programming (GEP). The statistical characteristics of the dataset, such as mean, standard deviation (SD), skewness, and range, are summarized in Table 4. The skewness and kurtosis values help describe the shape of the dataset's distribution skewness, indicating asymmetry, while kurtosis measures the data's "tailedness." These factors are critical, as they can influence the model's performance by highlighting potential biases or outliers in the data[42].

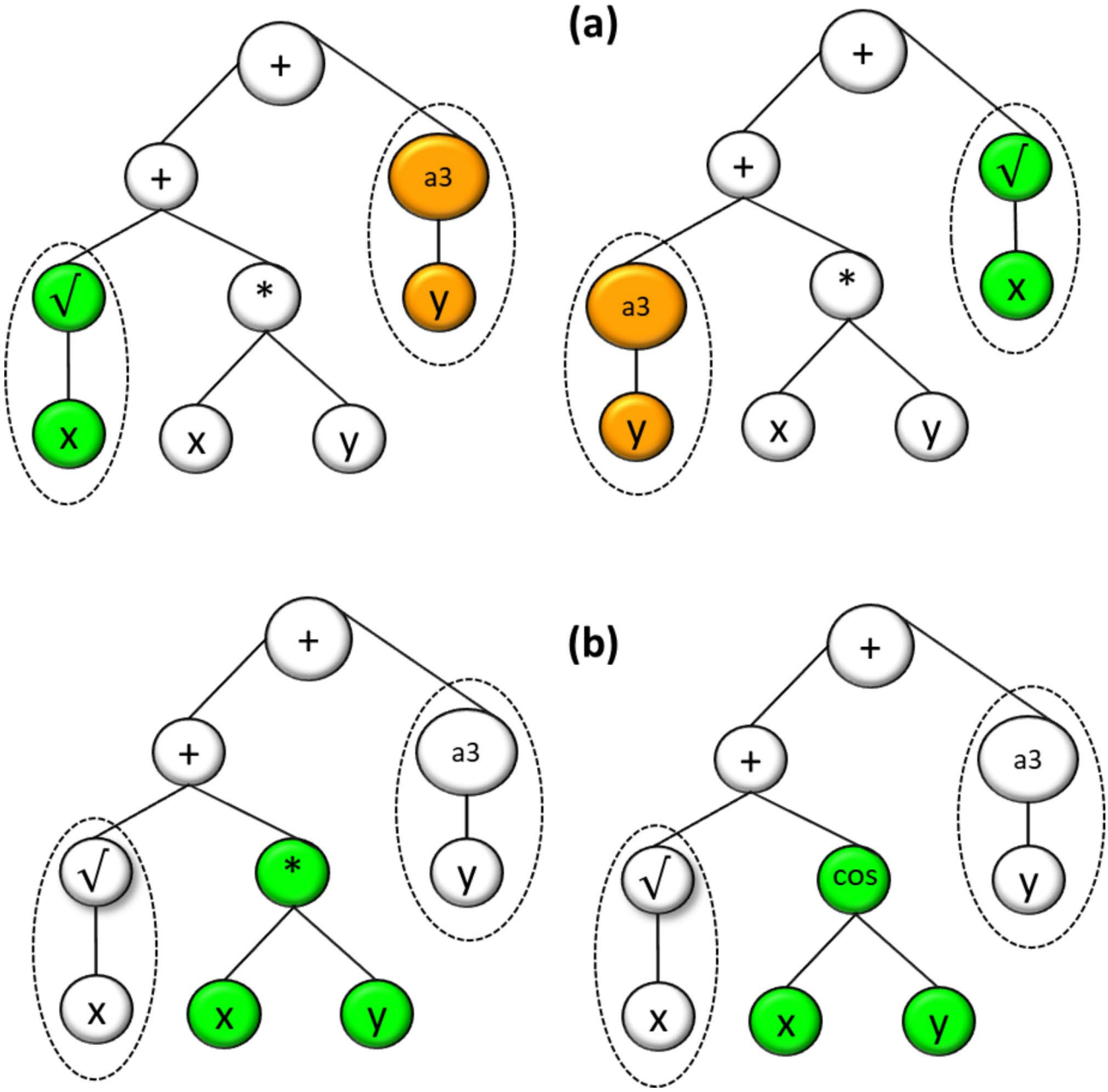

Fig. 5. The demonstration of ETs in LISP representation for (**a**) crossover and (**b**) mutation in GEP.

The distribution of the data plays a critical role in model performance. Contour plots were generated using Python libraries to explore the relationships between the variables, illustrating the interaction between the input parameters and the output (CS and TS), as shown in Figs. 6 and 7. These plots show a well-distributed data pattern across the specified range, further validating the dataset's reliability for model development.

For data preparation, the dataset was randomly divided into training and testing subsets, with 70% of the data allocated for training and 30% reserved for testing. The data was randomized using the built-in function in the GEP to ensure an unbiased and representative split. The pre-processing of the data was also performed using the default settings of the GEP software to bring all variables to a consistent scale (standardization), thereby enhancing model accuracy[43].

GEP model development

In order to create the model, the initial phase involves the identification and development of fundamental equations considering the most important input and output parameters. In this study, the CS and TS of quarry dust concrete are considered as the function of the variable described in Eq. 2.

$$CS\ or\ TS\ (MPa) = f\left(C, S, QD, CA,\ SP, W, A\right) \tag{2}$$

Description	Cement (kg/m^3)	Water (kg/m^3)	Coarse aggregates (kg/m^3)	Quarry dust (kg/m^3)	Sand (kg/m^3)	Plasticizer (kg/m^3)	Age (Days)	CS (Mpa)	TS (Mpa)
Mean	304.63	140.02	1229.03	170.29	538.06	4.74	16.44	14.04	2.60
Mode	316.00	142.00	1276.00	0.00	690.00	5.86	14.00	9.69	2.18
SD	11.56	2.01	47.77	164.54	159.86	2.32	8.77	3.81	0.58
Skewness	-0.02	-0.02	-0.02	0.35	-0.38	-1.60	0.37	0.77	0.55
Minimum	293.00	138.00	1181.00	0.00	285.00	0.00	7.00	9.09	1.50
Kurtosis	-2.05	-2.05	-2.05	-1.39	-1.35	0.57	-1.53	-0.38	-0.47
Median	316.00	142.00	1276.00	143.00	570.00	5.86	14.00	13.44	2.43
Maximum	316.00	142.00	1276.00	428.00	713.00	5.86	28.00	22.67	3.90

Table 4. Statistical summary of the database.

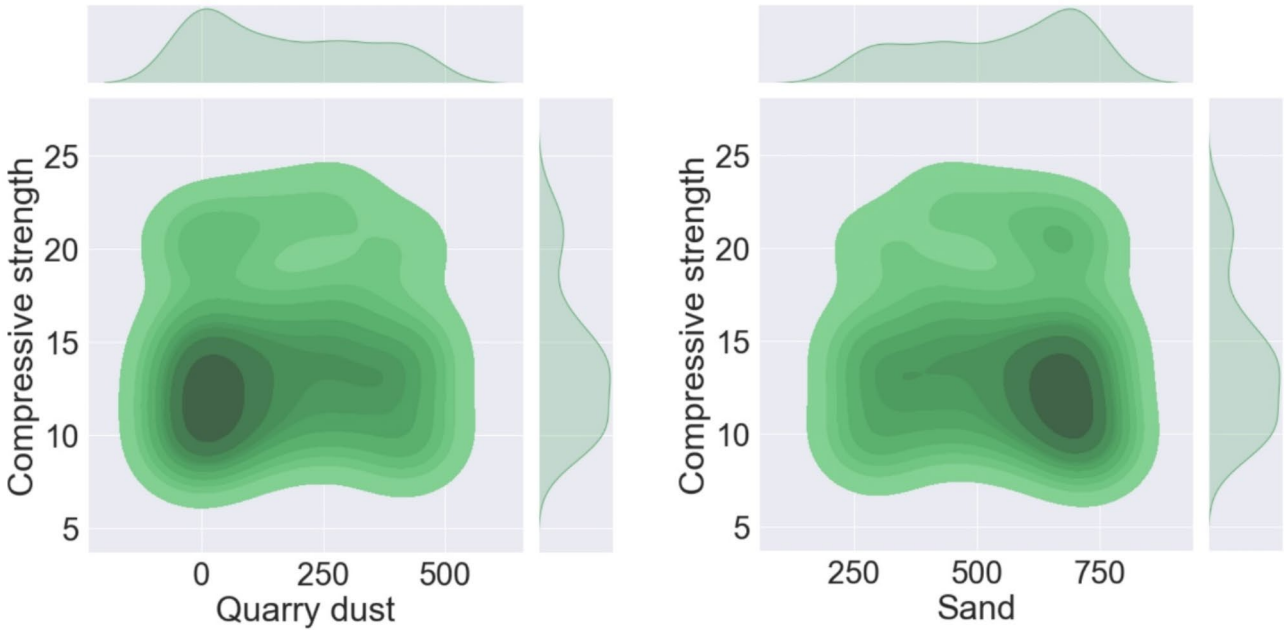

Fig. 6. Contour plots representing the relative distribution of (**a**) quarry dust (kg/m^3) and (**b**) sand (kg/m^3) in the CS (MPa) database.

For the GEP model development, optimizing hyperparameters is critical to achieving efficient and accurate results. The model's behavior and its ability to find optimal solutions are influenced by key hyperparameters such as mutation rate, number of chromosomes, genes, head size, linking function, and termination criteria. Previous studies often use a trial-and-error approach to fine-tune these parameters, where different values are systematically tested, and performance is assessed to find the best configuration.

In this study, hyperparameter optimization followed a trial-and-error method, as recommended in the literature[44], to identify the best combination of settings. Parameters such as mutation rates (ranging from 0.01 to 0.1), and head sizes (ranging from 8 to 16) were tested, among others. Lower mutation rates provided stability but slowed convergence, while higher rates improved exploration but increased the risk of premature convergence. Similarly, larger head sizes enhanced the model's complexity and accuracy but at the expense of increased computational demand.

Considering the balance between accuracy and model complexity, the optimal settings were identified and are presented in Fig. 8. Tables 5 and 6 further summarize the various combinations of hyperparameter settings tested. A total of 20 GEP models were executed to meet the termination criteria, which was based on minimizing the root mean square error (RMSE). This process ensured that the final configuration produced the best balance between accuracy and computational efficiency.

GEP model assessment

The correlation coefficient (R) is mainly used to assess the efficiency of a model. However, it is important to note that this measure cannot be primarily relied upon as an indicator of model capacity for prediction, as it does not consider the effects of dividing or multiplying results by a constant. Hence, this study also incorporates other measures such as mean absolute error (MAE), performance index (PI), and RMSE. The mathematical equations for these metrics are presented below[45,46].

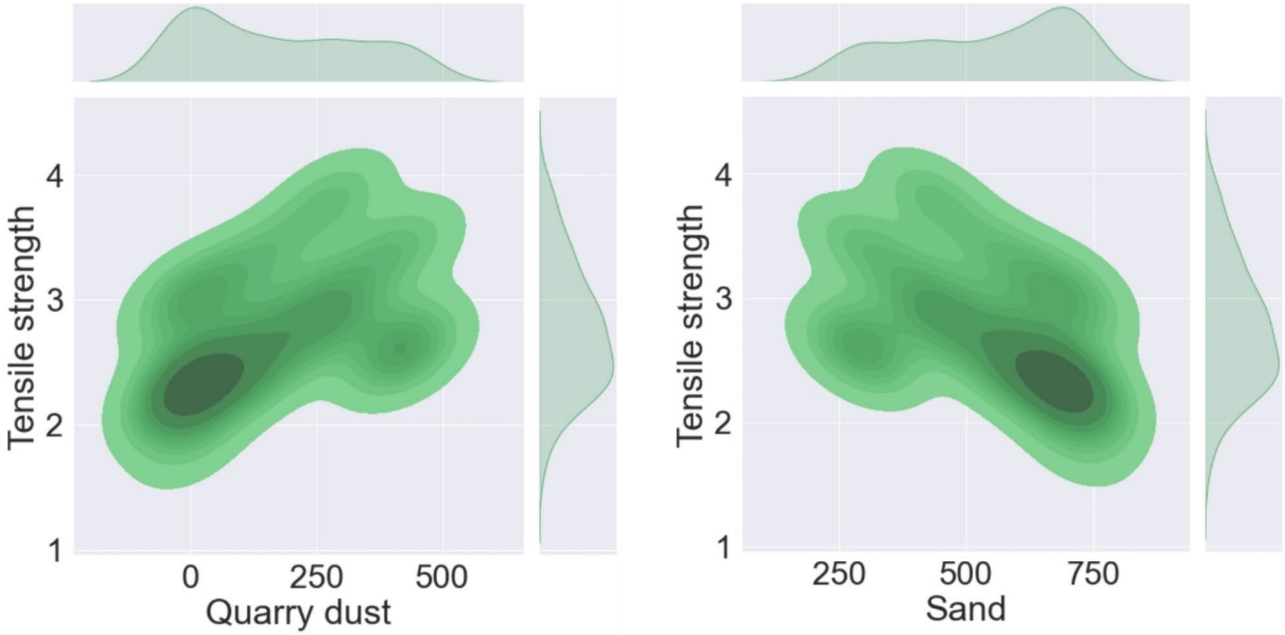

Fig. 7. Contour plots representing the relative distribution of (**a**) quarry dust (kg/m^3) and (**b**) sand (kg/m^3) in the TS (MPa) database.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(exp_i - mo_i)^2}{n}} \tag{3}$$

$$R = \frac{\sum_{i=1}^{n}\left(exp_i - \bar{exp}_i\right)^2\left(mo_i - \bar{mo}_i\right)^2}{\sqrt{\sum_{i=1}^{n}\left(exp_i - \bar{exp}_i\right)^2\sum_{i=1}^{n}\left(mo_i - \bar{mo}_i\right)^2}} \tag{4}$$

$$RRMSE = \frac{1}{\left|\bar{exp}\right|}\sqrt{\frac{\sum_{i=1}^{n}(exp_i - mo_i)^2}{n}} \tag{5}$$

$$PI = \frac{RRMSE}{1+R} \tag{6}$$

$$MAE = \frac{\sum_{i=1}^{n}|exp_i - mo_i|}{n} \tag{7}$$

In the equations mentioned above, the variable "exp" represents the target value, the variable "mo" denotes the model anticipated value, and the variable "n" represents the total number of assembled data points. Likewise, $\bar{\mathrm{exp}}_\mathrm{i}$ and $\bar{\mathrm{mo}}_\mathrm{i}$ Indicates the mean values of the data record.

Overfitting poses a possible concern in numerous ML-based models. Therefore, in order to address this problem, this study incorporated an objective function as presented in Eq. (8). The incorporation of an OF makes it possible to establish an appropriate balance between the complexity of the model and its capacity to effectively extrapolate unfamiliar data.

$$OF = \left(\frac{n_T - n_V}{n}\right)PI_{i\,T} + 2\left(\frac{n_V}{n}\right)PI_{i\,V} \tag{8}$$

Here,

OF = objective function.

$\mathrm{n_T}$ and $\mathrm{n_V}$ = number of training and validation dataset

Results and discussion

Experimental results

Workability

The workability of the concrete was evaluated using the slump test, as illustrated in Fig. 9. The concrete grades in this study are categorized as M15 (15 MPa CS) and M20 (20 MPa CS), with QD representing the quarry

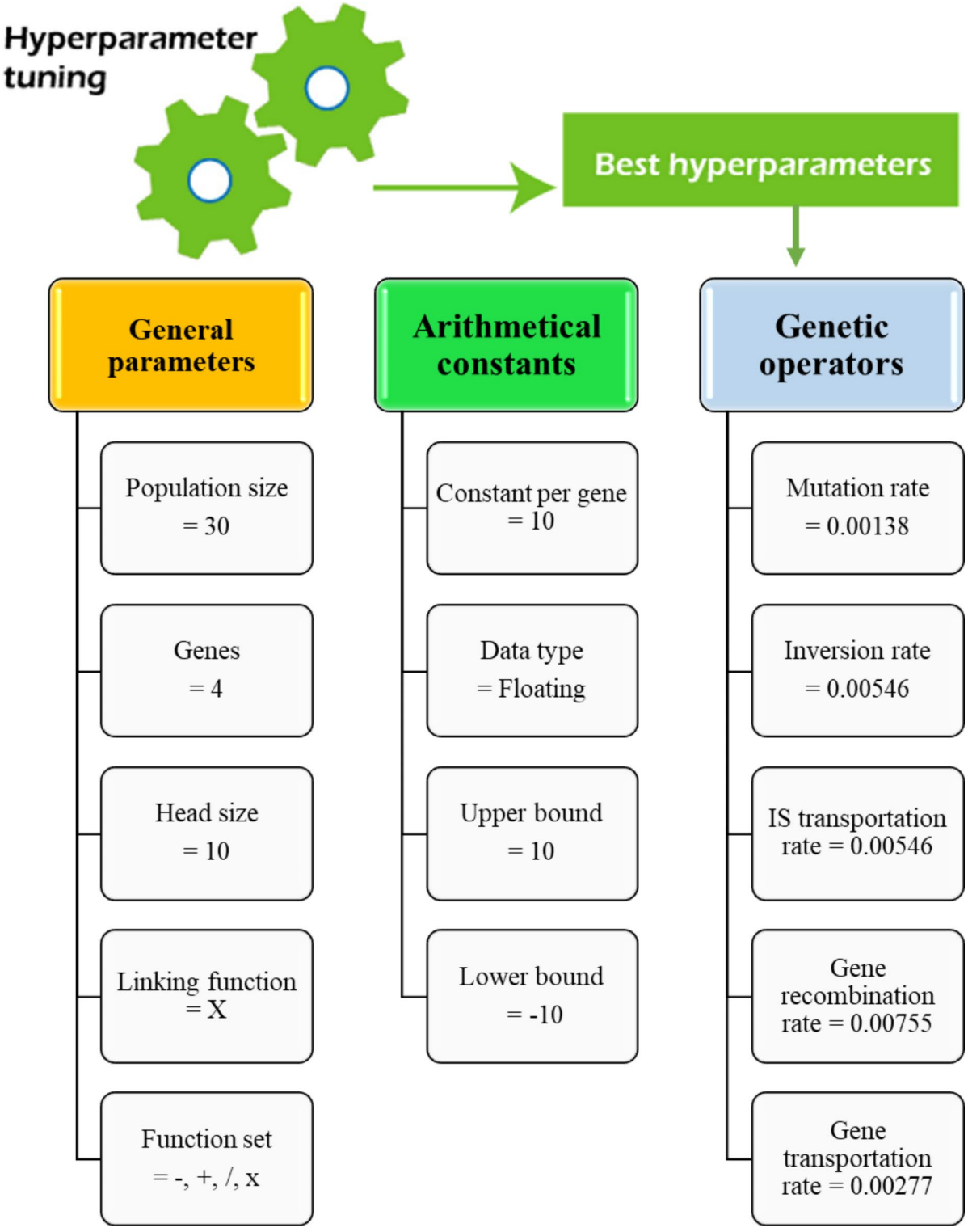

Fig. 8. Optimum hyperparameter employed in GEP models.

dust content in the mix. The results demonstrate that as the proportion of quarry dust increases, the slump values for both M15 and M20 concretes decrease significantly. For example, the slump value for M15-QD20 decreased by approximately 10%, while M20-QD60 showed a slump reduction of nearly 35%. This reduction in workability can be attributed to the increased surface area of the quarry dust particles, which requires more water to maintain the same consistency[47]. Quarry dust's finer particles reduce the amount of free water available, thus lowering the concrete's overall fluidity.

Models	No. of chromosomes	Variable used	Head size	Number of genes	Linking function	Function set	Duration (minutes)	RMSE MPa
CS1	30	7	8	3	Addition	(÷, ×, +, -)	20	1.047
CS2	50	7	10	4			23	1.079
CS3	80	6	12	5			25	1.060
CS4	100	7	14	6			30	0.934
CS5	150	7	16	7			40	0.959
CS6	30	7	8	3	Division	(÷, ×, +, -)	23	1.069
CS7	50	6	10	4			26	1.087
CS8	80	5	12	5			29	0.871
CS9	100	6	14	6			35	1.046
CS10	150	7	16	7			45	1.002
CS11	30	6	8	3	Multiplication	(÷, ×, +, -)	18	0.998
CS12	**30**	**7**	**10**	**4**			**22**	**0.84**
CS13	80	7	12	5			24	0.92
CS14	100	7	14	6			27	1.094
CS15	150	5	16	7			35	0.859
CS16	30	6	8	3	Multiplication	(÷, ×, +, -, Pow, √)	30	0.912
CS17	50	7	10	4			34	0.875
CS18	80	7	12	5			38	0.881
CS19	100	6	14	6			42	0.991
CS20	150	6	16	7			50	0.951

Table 5. Descriptive summary of parametric settings of GEP models used for CS.

Models	No. of chromosomes	Variable used	Head size	Number of genes	Linking function	Function set	Duration (minutes)	RMSE MPa
TS1	30	5	8	3	Addition	(÷, ×, +, -)	22	0.165
TS2	50	6	10	4			24	0.179
TS3	80	7	12	5			26	0.196
TS4	100	7	14	6			32	0.153
TS5	150	7	16	7			40	0.176
TS6	30	6	8	3	Division	(÷, ×, +, -)	24	0.169
TS7	50	6	10	4			26	0.127
TS8	80	7	12	5			28	0.176
TS9	100	6	14	6			33	0.188
TS10	150	7	16	7			45	0.167
TS11	30	9	8	3	Multiplication	(÷, ×, +, -)	20	0.145
TS12	**30**	**7**	**10**	**4**			**22**	**0.142**
TS13	80	7	12	5			26	0.152
TS14	100	6	14	6			27	0.191
TS15	150	6	16	7			37	0.187
TS16	30	7	8	3	Multiplication	(÷, ×, +, -, Pow, √)	31	0.191
TS17	50	7	10	4			37	0.181
TS18	80	6	12	5			38	0.191
TS19	100	5	14	6			44	0.156
TS20	150	7	16	7			54	0.213

Table 6. Descriptive summary of parametric settings of GEP models used for TS.

In the control mixes (M15-CM and M20-CM) without superplasticizers, slump values of 85 mm and 81 mm were recorded, respectively. The addition of a 2% superplasticizer significantly improved the slump, resulting in values of 92 mm for M15 and 88 mm for M20 concretes without quarry dust. Superplasticizers enhance workability by reducing the water demand for the same fluidity, which is particularly beneficial in compensating for the loss of workability caused by quarry dust. Specifically, superplasticizers improve the dispersion of cement particles, reducing agglomeration and increasing fluidity, which is crucial in mixes with higher quarry dust content[48,49]. However, the minimum slump value recorded was 54 mm for the M20-QD60 mix, indicating that excessive quarry dust content, even with superplasticizers, can significantly impact workability.

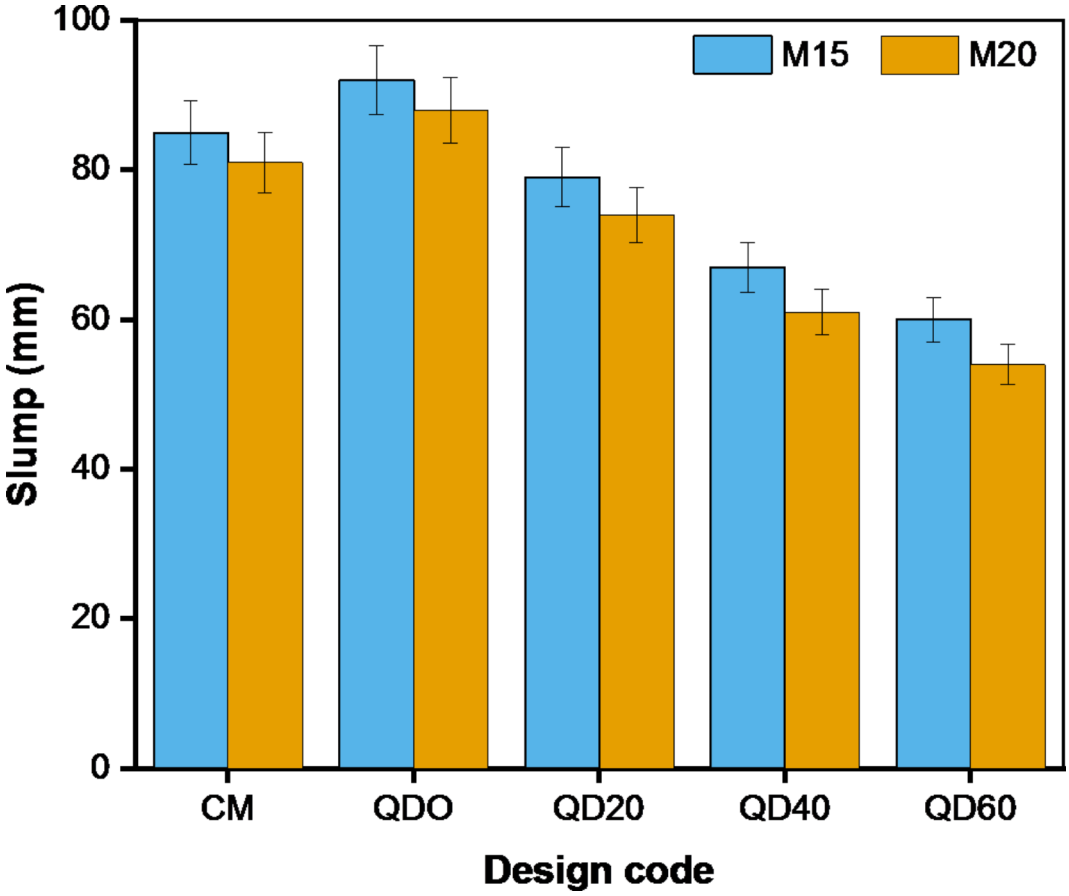

Fig. 9. Slump test laboratory results for M15 and M20 grade concrete.

Compressive strength

The CS results for all mix designs of M15 and M20 at 7, 14, and 28 days of curing are illustrated in Fig. 10(**a**, **b**). The CS shows a linear increase as the curing time extends from 7 to 28 days, confirming the beneficial effects of prolonged curing on concrete strength. Additionally, the CS increases with the replacement level of sand with quarry dust up to 40%, after which it begins to decrease. The observed decline in CS for mixtures with a 60% substitution ratio of quarry dust can be attributed to insufficient workability in these concrete mixes, which hinders proper hydration and strength development.

The maximum CS achieved was 17.2 MPa for M15 and 22.23 MPa for M20 at 28 days, with 40% incorporation of quarry dust and 2% superplasticizer. This represents enhancements of 15% and 13% in CS for M15-QD40 and M20-QD40 compared to their respective control mixes. Such increments in CS can be attributed to the effective filling of voids between sand particles with quarry dust, resulting in a denser and more cohesive concrete matrix, as supported by previous studies[47]. For instance, Bahoria et al.[50] observed similar improvements in CS with quarry dust incorporation, noting that up to 30–35%, replacement led to optimal strength gains due to enhanced particle packing and reduced void content.

Moreover, the addition of superplasticizer contributes significantly to early strength development, enhancing the CS of M15 by 5% and M20 by 6% compared to conventional concrete mixes. This aligns with findings by Pereira et al.[46], who highlighted that superplasticizer not only improve workability but also promote better cement particle dispersion, leading to increased hydration and strength gains in concrete. Overall, these results emphasize the importance of optimal quarry dust replacement levels and the use of superplasticizers in achieving desirable concrete performance.

Split tensile strength

Figure 11(**a**, **b**) presents the laboratory results for the TS of M15 and M20 concrete grades at 7, 14, and 28 days. The data indicate a clear linear trend in TS corresponding to the increasing age of the specimens, reflecting the typical hydration process of concrete. As expected, TS generally improves with curing time as the cement continues to hydrate and gain strength[15].

The results also show that TS increases with the substitution of sand with quarry dust up to a replacement level of 40%. Beyond this threshold, a decline in TS is observed, indicating an optimal proportion for quarry dust use. The peak TS values of 3.74 MPa for M15 and 3.83 MPa for M20 at 28 days demonstrate the effective role of quarry dust, particularly in conjunction with a 2% superplasticizer. This enhancement in TS can be attributed to

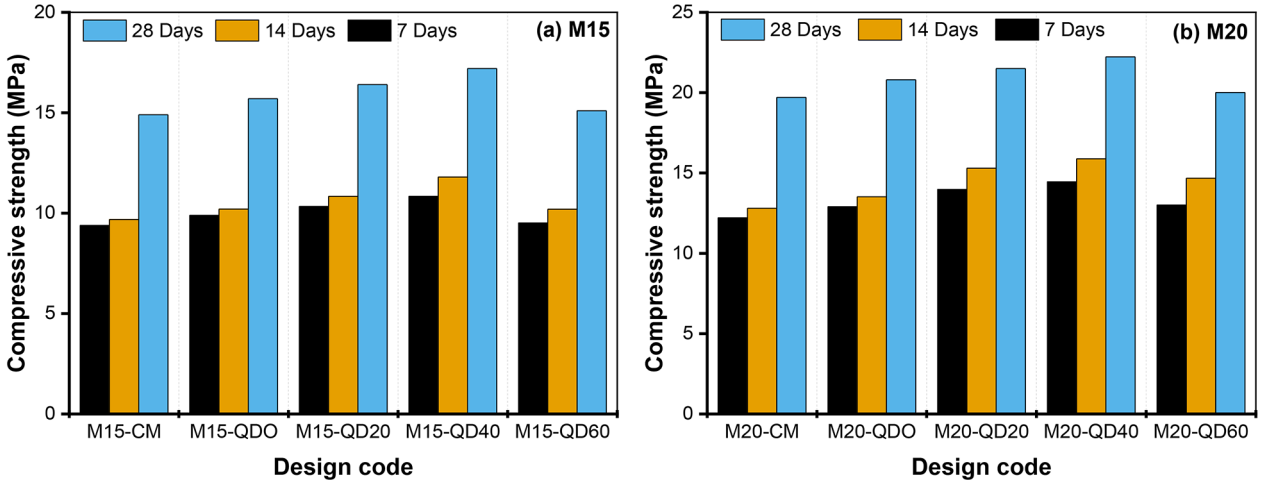

Fig. 10. CS test laboratory results for (**a**) M15 and (**b**) M20 grade concrete.

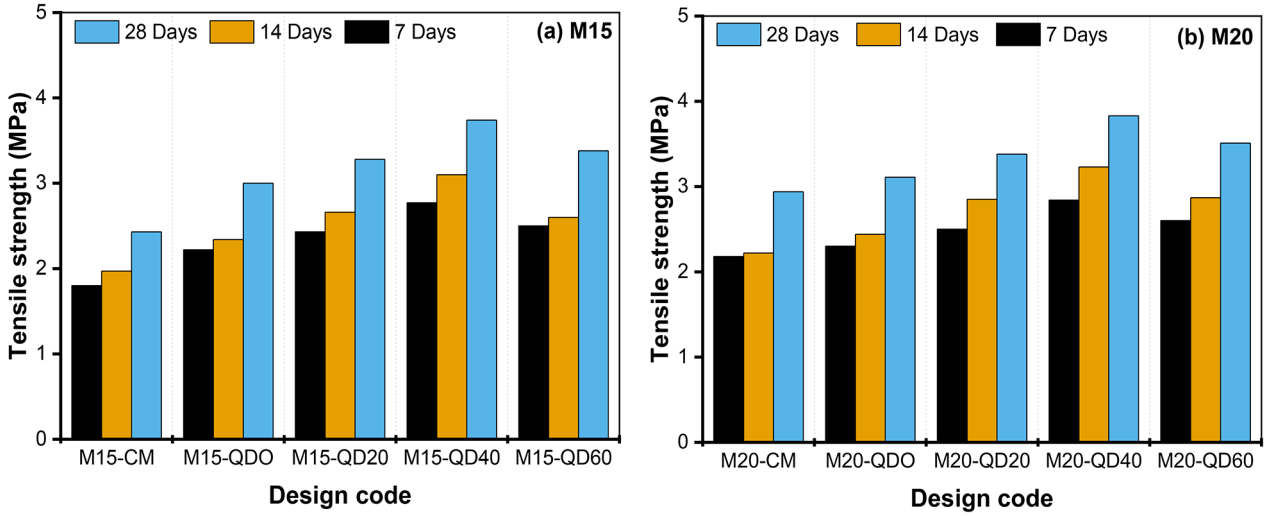

Fig. 11. TS test laboratory results for (**a**) M15 and (**b**) M20 grade concrete.

the improved particle packing and increased surface area provided by quarry dust, which leads to better bonding and strength development[10].

The observed improvements in TS with the use of superplasticizers further highlight their role in optimizing concrete properties. Specifically, the superplasticizer contributed to an 11% and 6% increase in TS for M15 and M20 grades, respectively, when compared to control mixes.

Machine learning results

GEP models predictive performance

The statistical analysis outcomes of the best GEP models for CS and TS are depicted in Fig. 12(**a**, **b**). When the R^2 values exceed 0.7, it indicates that the anticipated outcomes are in closer alignment with the experimental findings. Furthermore, as the R^2 value approaches 1, the degree of agreement between actual and model outcomes increases. The R^2 provides an indicator of the fitness efficiency of a specific function on the database being used[51]. Specifically, a higher R^2 value during the training stage is indicative of a highly favorable fitness of the resultant function. Nevertheless, this phenomenon can result in the model deviating from its original purpose, making it incapable of accurately predicting unseen data. Consequently, a substantial difference between errors observed during the training and evaluation (testing) stages appears, ultimately leading to higher regression values in both the training and evaluation datasets being examined. In the proposed models, it was observed that the R^2 value exhibited greater values during the training stage compared to the evaluation stage in most cases, and also, minimal difference was observed between the two stages. The R^2 values associated with estimating CS exhibit a value equal to 0.96 and 0.94, whereas for TS models, they are equal to 0.92 and 0.91 in the training

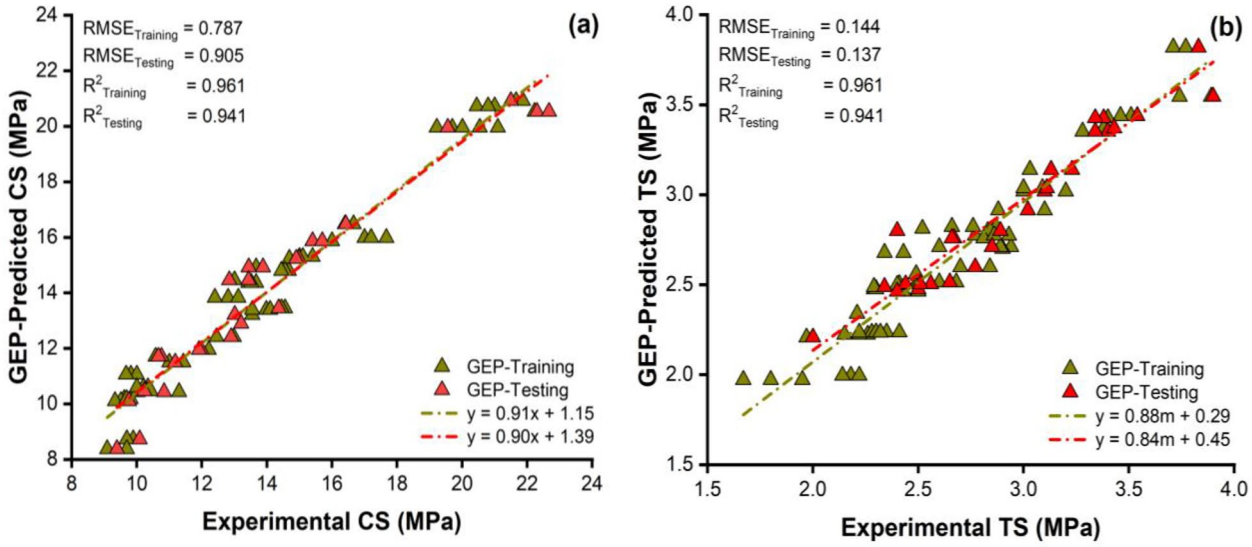

Fig. 12. Experimental Vs GEP-anticipated outcomes (**a**) for CS-GEP and (**b**) for TS-GEP models.

and testing stages, respectively. These high R^2 values strongly support the acceptance of the proposed models, as they align with benchmarks found in similar studies. For instance, Ullah et al.[26] reported R^2 values in the range of 0.85 to 0.95 for concrete strength predictions using ML techniques, underscoring the robustness of proposed models. While high R^2 values are indicative of model performance, it is important to also consider the RMSE and MAE, which provide additional insights into model efficiency. A lower RMSE and MAE signify better predictive performance of the proposed GEP models in estimating CS and TS in concrete mixes. It can be noted that the least RMSE for CS-GEP are (0.787 Mpa and 0.905 MPa) and MAE (0.65 Mpa and 0.703 MPa). Likewise, these errors for TS-GEP are (0.144 Mpa and 0.137 MPa) and (0.115 MPa and 0.102 MPa) in the training and testing stages, respectively. In addition, comprehensive information regarding further statistical errors (including RSE, RRMSE, R) can be found in Table 6 for CS-GEP and TS-GEP models. On a statistical basis, it can be stated that there were minimal discrepancies between the actual and projected results. Hence, it is evident that the proposed models possess the potential to be utilized in the prediction of the CS and TS of concrete incorporating mineral filler.

Furthermore, Figs. 13 and 14 illustrate the range of absolute errors for CS and TS GEP models with a mean error of 0.65 MPa and 0.115, respectively. The smallest error values in the case of CS-GEP and TS-GEP are recorded as 0.03 MPa and 0.003 MPa, while the greatest error values can be identified as 1.71 MPa and 0.34 MPa, respectively. Additionally, the analysis of the projected error results indicates that most of the data points, approximately 95%, fall under 2 MPa and 0.3 MPa for both cases. This observation can be seen as evidence of the model's excellent performance in forecasting the CS and TS of concrete incorporating mineral fillers. While the GEP demonstrated excellent results, it has limitations such as sensitivity to input parameters, potential overfitting, and complex evolutionary trees (ETs). In this study, regression models were also utilized for cross-validation to compare the GEP results more effectively.

GEP-based formulations

Following the statistical analysis of various GEP trials, the trial that yielded the highest value of R^2 and the smallest value of RMSE was selected for further analysis. This selected optimum trial was then utilized to extract ETs, as depicted in Figs. 15 and 16 for CS and TS models, respectively. The (ETs) consist of 'ds' and 'Cs', which indicate ingredients of concrete (inputs) and constants. These variables are connected by a linkage function, as "multiplication" is used in this study. The other settings used to develop these ETs are already discussed in the methodology section. The exact formula to estimate the CS of modified concrete was derived from decoding given ETs, which are expressed in Eqs. (9)-(13).

$$CS\ (MPa) = A \times\ B \times\ C \times\ D \tag{9}$$

$$A = \left[\frac{1}{-18.99 + A\,(C - SP) - A.SP}\right] \tag{10}$$

$$B = \left[C - \frac{3CA}{-267.16 + C} - 0.081(-CA + QD)\right] \tag{11}$$

$$C = \left[-3.16 + \frac{(-A - 1.67C)W}{CA}\right] \tag{12}$$

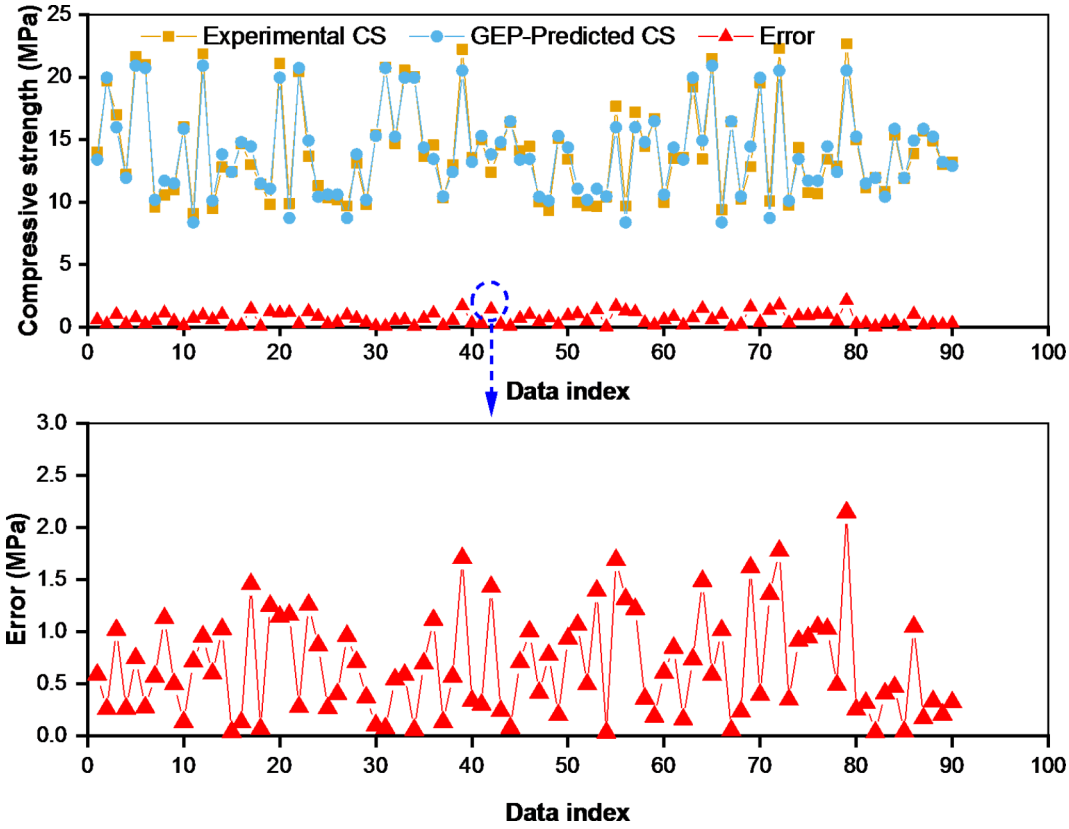

Fig. 13. Error distribution in GEP-anticipated CS outcomes.

$$D = \left[-8.15 - A - \frac{S}{C - CA - QD + S}\right] \tag{13}$$

Similar to CS, the GEP-based mathematical expression for TS is decoded from ETs given in Fig. 2 and is presented in the form of Eqs. (14)-(18).

$$TS\ (MPa) = E \times\ F \times\ G \times\ H \tag{14}$$

$$E = \left[1.23 - \frac{7.793(22.022726006712 + \mathrm{C})}{\mathrm{CA}(\mathrm{C} - \mathrm{QD} - 0.12\mathrm{S})}\right] \tag{15}$$

$$F = \left[11.78 + 0.11\left(2\mathrm{A} - \frac{\mathrm{QD}}{\mathrm{W}}\right)\right] \tag{16}$$

$$G = \left[5.37 - \frac{5.47\mathrm{C}}{(9.11 + \mathrm{QD} + \mathrm{SP})(\mathrm{C} - \mathrm{W})}\right] \tag{17}$$

$$H = \left[\frac{-8.77 + C - \frac{QD}{S} + SP}{A + 8.44(1.16 + CA)}\right] \tag{18}$$

The proposed formulations can be utilized for the predesign of concrete containing quarry dust by simply providing the relevant input parameters and solving the equations. It is important to note that these equations are applicable only within the range of values used in the database. For instance, to determine the CS and TS of concrete with quarry dust, practitioners can input the necessary data into Eqs. 9 and 14, to get the values of CS and TS, respectively.

External validation of GEP-based formulations

Various statistical analysis (described by the equations in Table 7) were performed as independent external validation checks on these proposed models for CS and TS of modified concrete with mineral filler. In addition, the acceptable limits of each check are outlined in Table 7. The equations used to determine whether or not a model is correct depend on the conditions of each validation check. The values of k or k' denote the slopes of the regression plots that pass within the point of origin (0, 0) and are required to be nearly equal to one[52]. The

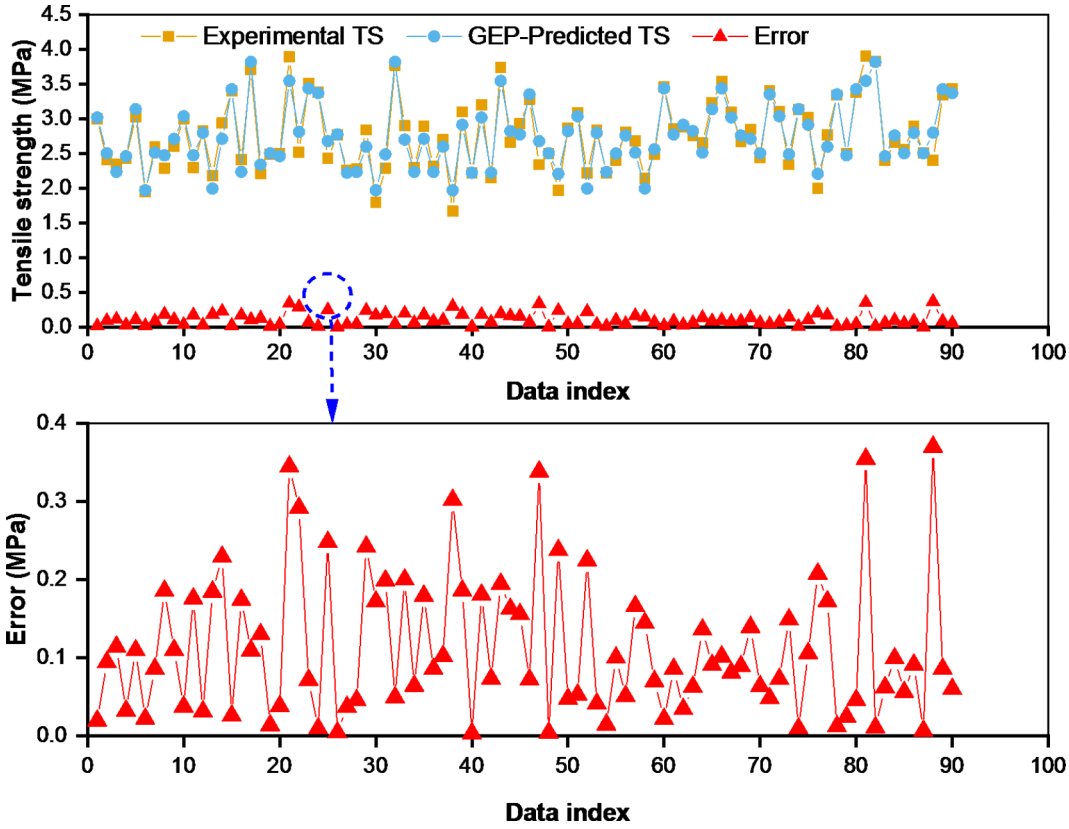

Fig. 14. Error distribution in GEP-anticipated TS outcomes.

verification index R_o^2 and $R'_o{}^2$was proposed by[53] means to assess the external dependability and predicted efficacy of the model. If these values are closer to 1, then the specified condition is satisfied. It can be observed that all equations fall within the acceptable range; hence, it can be concluded that both models are correct and have a good level of prediction.

Sensitivity analysis (SA)

It is important to perform several experiments on ML-based models to ascertain the dependability and efficacy of the proposed models across different datasets. In this study, sensitivity analysis (SA) was carried out to quantitatively assess the impact of input parameters on the CS and TS of concrete containing mineral fillers. The SA of the prediction model on the complete database provides an indication of the degree to which a created model responds to changes in a specific input ingredient of concrete. Equations (19) and (20) were utilized to determine the impact of each input ingredient on the intended objective.

$$Z_i = f_{max}(z_i) - f_{min}(z_i) \tag{19}$$

$$SA = \frac{Z_i}{\sum_n^{j=1} Z_j} \tag{20}$$

The highest and lowest model projected values are denoted as $f_{max}(z_i)$ and $f_{min}(z_i)$, respectively. The results of SA are depicted in Fig. 17**(a, b)** for CS and TS GEP-based models. It can be observed that among the seven inputs considered in this study, the content of cement (C) has the most significant impact on the CS and TS of the concrete, accounting for 49.5% and 25% of the predicted values, respectively. This high influence can be attributed to cement's critical role in the hydration process, which directly affects the formation of the concrete matrix and its overall strength. Previous studies have shown that an increase in cement content enhances the bonding properties and density of the concrete, thereby improving its mechanical performance[54]. In contrast, the quantity of superplasticizer (SP) is the least influential input, contributing only 20.61% and 4% to the predicted CS and TS. This limited impact can be explained by the fact that while superplasticizers enhance workability, their effect on strength is more indirect, often depending on the overall mix design and the hydration characteristics of the cement used[55].

Moreover, the analysis reveals that age (A) also contributes significantly to the strength of concrete. Age plays a crucial role as the hydration process continues over time, allowing for increased strength development. These findings are well-supported by laboratory testing, which confirms the efficacy of the GEP models in predicting the mechanical properties of concrete[10,56]. For a clearer understanding of these trends, Fig. 17(a, b) presents

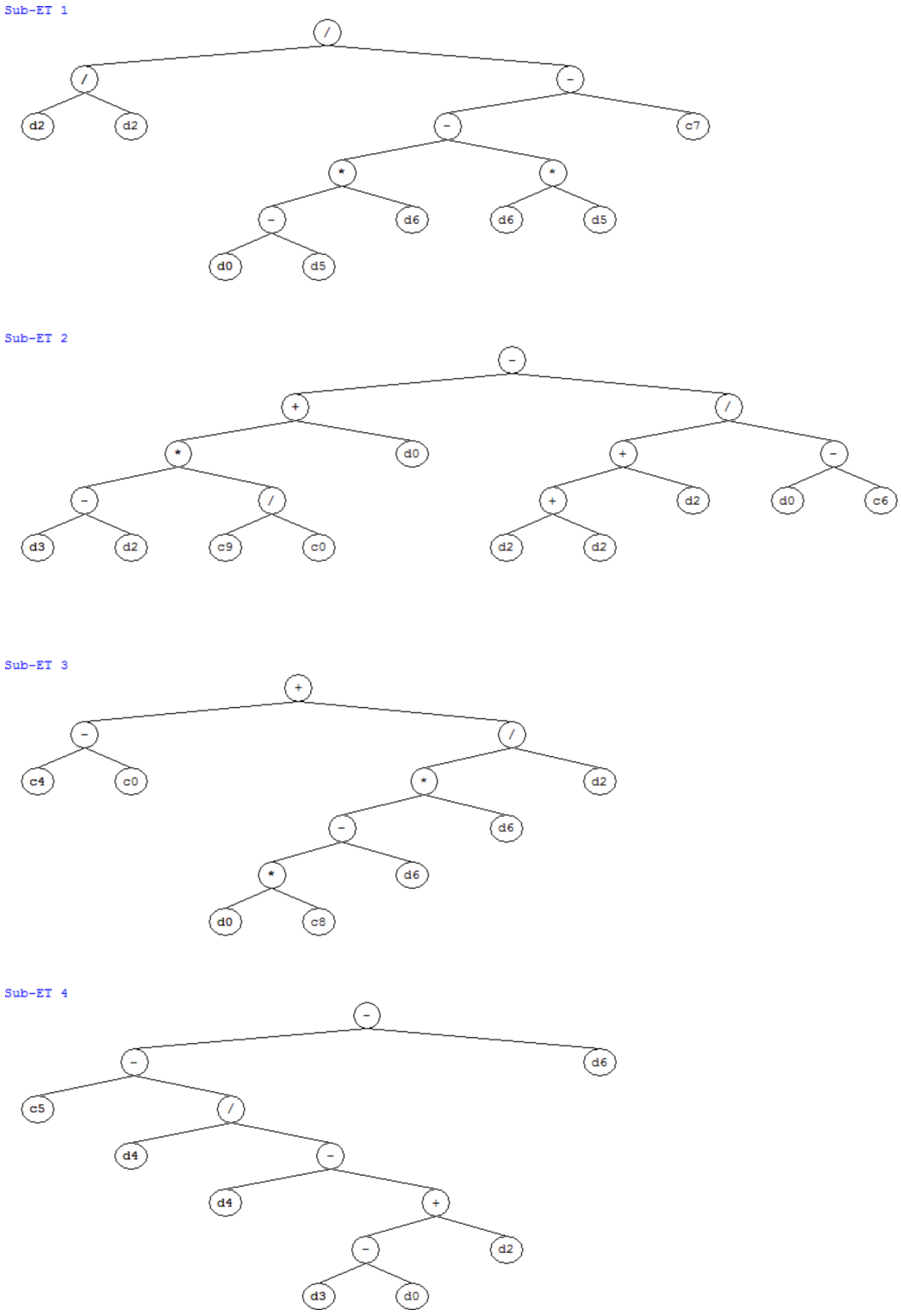

Fig. 15. Expression tree (ETs) of the CS-GEP model.

the SA results. The figure effectively illustrates the varying degrees of influence each input has on CS and TS, providing a visual representation of the data.

GEP model comparison with multi-linear regression

In order to further validate the predictive capacity of the GEP models, a comparative analysis was performed between the effectiveness of the suggested GEP model and that of the conventional Multi-Linear Regression

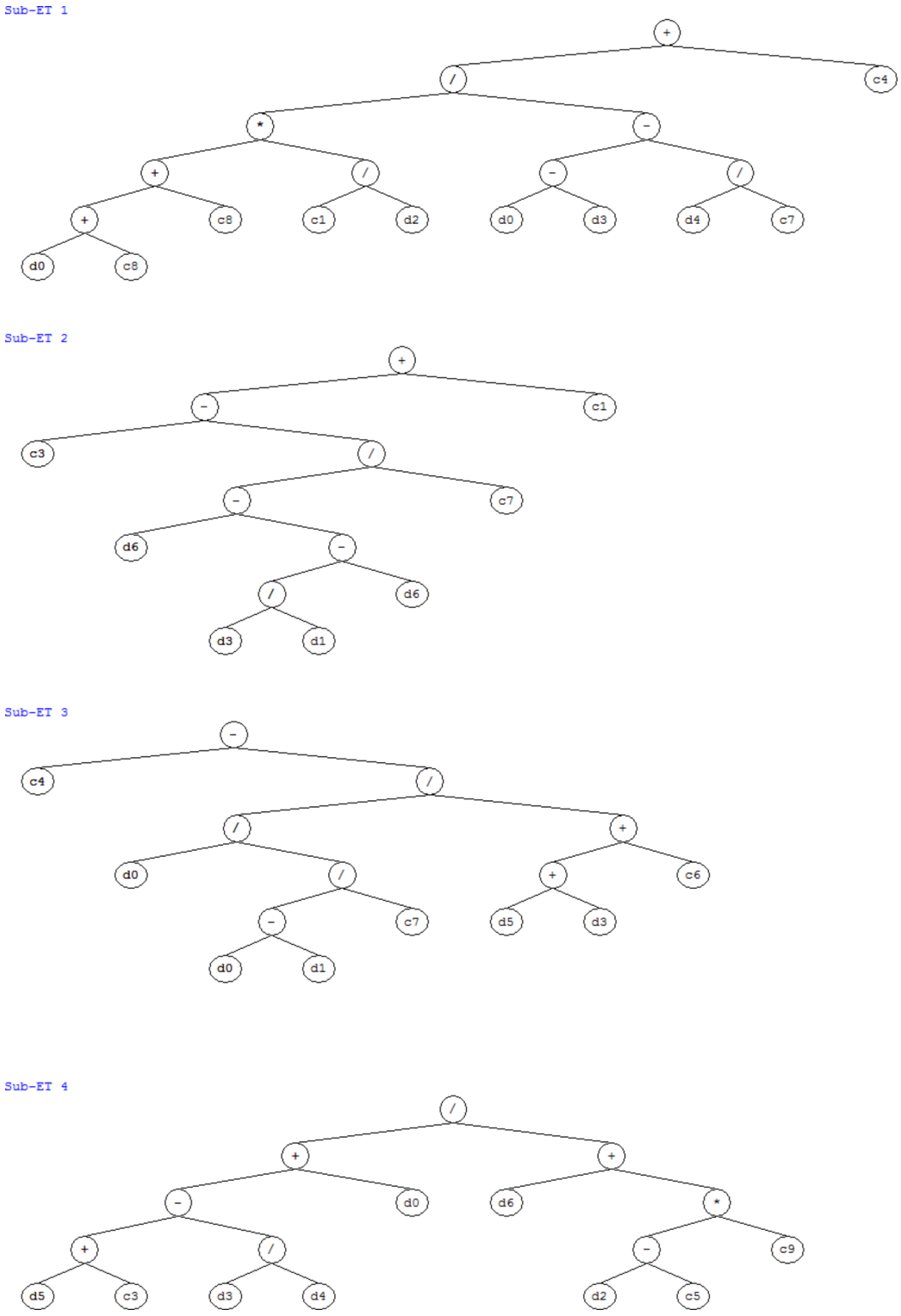

Fig. 16. Expression tree (ETs) of the TS-GEP model.

(MLR). Equations (21) and (22) present the MLR empirical formulations for CS and TS of concrete, respectively, that have been developed using an identical dataset.

$$CS\ (MPa) = 0.019C + 0.045.CA + 0.061QD + 0.062S + 0.0199SP + 0.294A - 92 \quad (21)$$

$$\mathrm{TS\ (MPa)} = 0.013\mathrm{C} + 0.0012.\mathrm{CA} - 0.0016.\mathrm{QD} - 0.002\mathrm{S} + 0.07\mathrm{SP} + 0.041\mathrm{A} + 1.92 \quad (22)$$

S. No.	Equation	Acceptable range	Model	
			CS	TS
1	R	$R > 0.8$	0.971	0.983
2	$R_o^2 = 1 - \frac{\sum_{i=1}^{n}(mo_i - exp_i^o)^2}{\sum_{i=1}^{n}(mo_i - mo_i^o)^2}$, $exp_i^o = k \times mo_i$	$R_o^2 \cong 1$	0.989	0.980
3	$R\prime_o^2 = 1 - \frac{\sum_{i=1}^{n}(exp_i - mo_i^o)^2}{\sum_{i=1}^{n}(exp_i - exp_i^o)^2}$, $mo_i^o = k\prime \times exp_i$	$R\prime_o^2 \cong 1$	0.981	0.991
4	$k = \sum_{i=1}^{n} \frac{(exp_i \times mo_i)}{exp_i^2}$	$0.85 < k < 1.15$	0.981	0.973
5	$k' = \sum_{i=1}^{n} \frac{(exp_i \times mo_i)}{mo_i^2}$	$0.85 < k\prime < 1.15$	1.012	1.021
6	$m = \frac{(R^2 - R_o^2)}{R^2}$	$m < 1$	0.0541	0.054
7	$n = \frac{(R^2 - R\prime_o^2)}{R^2}$	$n < 1$	0.0035	0.004

Table 7. External validation of the GEP models.

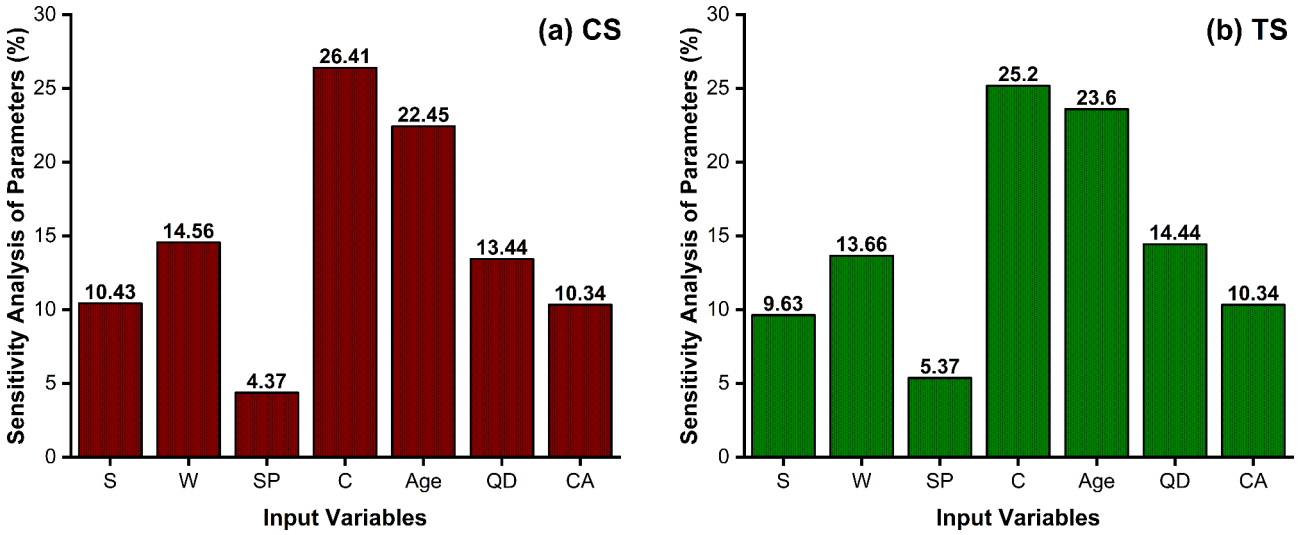

Fig. 17. Sensitivity analysis outcomes (**a**) CS-GEP and (**b**) TS-GEP model.

Table 8 illustrates the values of the various statistical indices (R, RMSE, RRMSE, OF, PI, MAE, and RSE) for the developed MLR models of CS and TS. The GEP-based model demonstrates superior efficacy than that of the MLR model, with an increase in the overall R^2 value of 13% and 11.7% in the case of CS and TS models, respectively. Similar to the GEP model, the PI readings in all cases of the MLR models are likewise below 0.2, which is an indication of the better validity of the model. The average PI values of the CS-GEP and TS-GEP models are seen to be 27% and 25% lower, respectively, compared to the CS-MLR and TS-MLR. Based on the OF values, the CS-GEP and TS-GEP models demonstrate a 26% and 23.5% improvement compared to the MLR models, respectively. Overall, GEP shows lesser mean error by 18.7% and 29.6% as compared to CS-MLR and TS-MLR models, respectively. Additionally, Fig. 18 visually compares the actual results versus the GEP and MLR predicted outcomes for both the training and evaluation sets. The scatter plots clearly show that the GEP-predicted values align much more closely with the actual values than those predicted by MLR models. Therefore, the suggested GEP models show enhanced and stable predictive capability when evaluating CS and TS of concrete using data that has not been previously seen.

The results of this study show that the GEP-predicted model exhibits superior performance compared to standard MLR models while utilizing the same dataset. The aforementioned limitations of traditional regression techniques arise from the requirement for prediction models to be connected to certain pre-specified declarations, which can be potentially linear or non-linear, as well as the assumption of residue normalization[57]. On the other hand, the outcomes of the GEP methods demonstrate that the developed model has been effectively trained and gained knowledge of the non-linear correlation between the inputs and output variables[58]. This has resulted in a significant reduction in error statistical analysis and an improved capacity to generalize, surpassing the performance of MLR models.

Model	Phase	R^2	R	RMSE	RRMSE	RSE	MAE	PI	OF
MLR-CS	Training	0.851	0.922	0.949	0.076	0.15	0.801	0.039	0.037
	Testing	0.877	0.936	1.318	0.074	0.126	1.2	0.038	
GEP-CS	Training	0.96	0.979	0.787	0.056	0.042	0.65	0.028	0.027
	Testing	0.941	0.97	0.905	0.064	0.06	0.703	0.032	
MLR-TS	Training	0.817	0.904	0.205	0.078	0.183	0.164	0.041	0.04
	Testing	0.776	0.881	0.191	0.062	0.227	0.124	0.033	
GEP-TS	Training	0.913	0.955	0.144	0.053	0.087	0.115	0.027	0.026
	Testing	0.923	0.96	0.137	0.046	0.084	0.102	0.023	

Table 8. Statistical summary of MLR and GEP models.

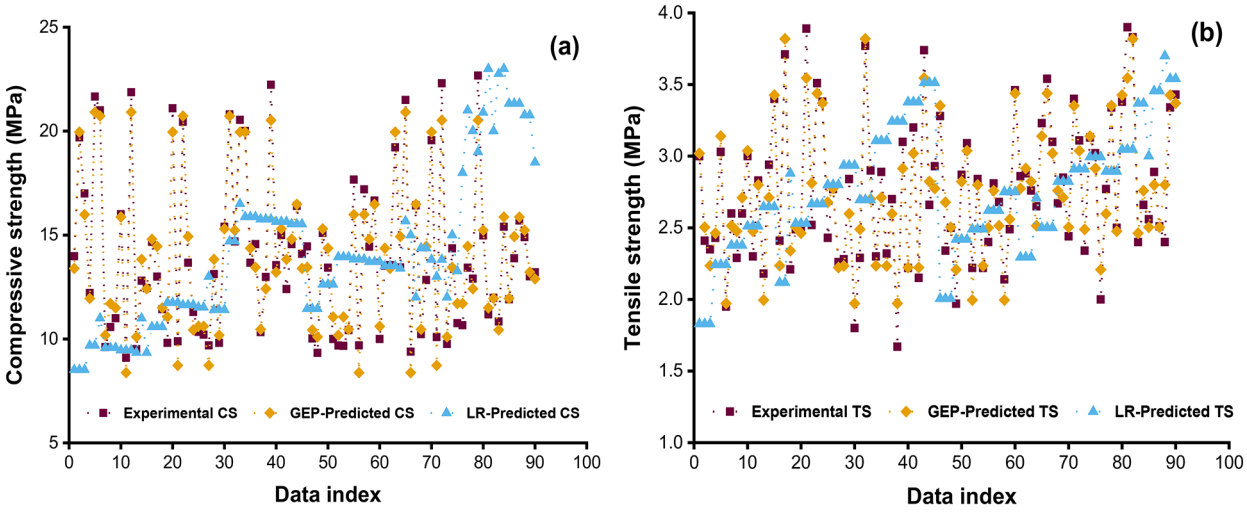

Fig. 18. Comparison between actual, MLR, and GEP-anticipated outcomes for (**a**) CS and (**b**) TS models.

Graphical user interface (GUI)

To enhance the practical application of this study, a GUI was developed in MATLAB, as shown in Fig. 19, specifically designed for handling concrete compositions that include quarry dust. The interface is divided into two sections: an input parameter section and an output result section, ensuring clear and user-friendly navigation. Users can input key parameters such as cement, quarry dust, water, age, plasticizer, and gravel and quickly obtain predictions by pressing the calculate button.

The prediction model integrated into the GUI is based on the XGBoost algorithm, chosen for its higher efficiency and accuracy in regression tasks, especially when dealing with complex, non-linear relationships like those found in concrete properties. XGBoost, a gradient boosting algorithm, builds multiple decision trees in sequence, where each tree corrects the errors of the previous one, leading to improved model performance. This iterative approach makes XGBoost highly effective in handling structured data, managing missing values, and preventing overfitting critical challenges in predicting concrete performance based on a wide range of input parameters.

MATLAB's Python integration was used to implement the XGBoost model within the GUI. This allows the trained model, developed in Python, to process the user-provided inputs and deliver accurate predictions displayed in the output section of the GUI. This integration enables users to interact with the machine learning model without needing advanced programming skills.

While initially developed for the dataset used in this study, the GUI framework can be further enhanced by incorporating additional data points and parameters. This will improve the accuracy and reliability of the model's predictions, making the GUI adaptable to a broader range of concrete mix design applications.

Comparison with literature

As mentioned earlier, this study utilizes a database for model creation based on laboratory testing. Hence, there are no models available on the same database to compare the efficacy of the suggested models. For example, Ray et al.[59] studied the combined effects of stone dust (SD) and nylon fiber (NF) on concrete's compressive and splitting tensile strength, along with varying water-cement (WC) ratios, using ANN-based models. Their study provides accurate models for predicting concrete strength with R^2 values of greater than 0.8. Similarly, K. Khan et al.[60] investigated the use of manufactured sand (MSC) in concrete to address river sand depletion.

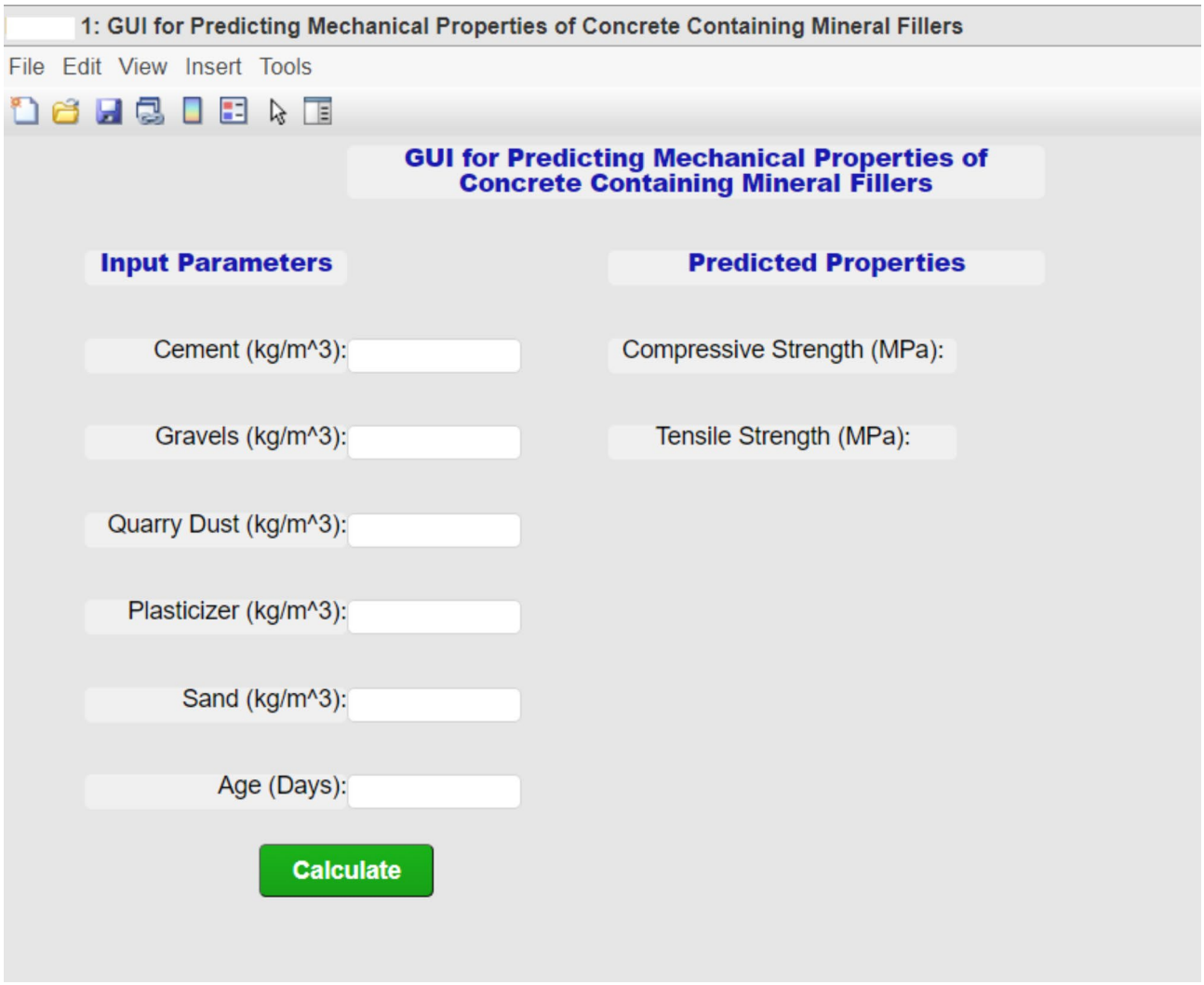

Fig. 19. GUI for estimating CS and TS of concrete.

Proposed models	Technique	Material used	R^2	R	MAE	References
CS	**GEP**	**Quarry dust and SP**	**0.961**	**0.979**	**0.654**	**This study**
TS	**GEP**		**0.912**	**0.951**	**0.115**	
CS	ANN	Stone dust and fiber	0.848	0.921	0.203	[60]
TS	ANN		0.852	0.923	0.116	
CS	GEP	Manufactured sand	0.906	0.951	3.211	[61]

Table 9. Comparison of present models with literature.

Using GEP models based on 275 experimental results, they predicted MSC's compressive strength. The best-performing model (M5) achieved R^2 values of 0.919 and 0.906 in training and testing, respectively. Key factors influencing strength were the water-cement ratio, curing period, and stone powder content, with the water-cement ratio being the most significant. The summary of the results of current and existing studies is shown in Table 9. The results produced by the suggested model exhibit a strong resemblance to those obtained from the literature for other materials. The alignment seen in this study indicates that equations based on the GEP hold affirmed as effective pre-design indicators for sustainable concrete design in the construction sector. The potential application of this innovation can bring about considerable reductions in time, cost, and resource allocation, representing a noteworthy progression for the respective field.

Conclusion

The current study addresses the increasing demand for sustainable alternatives to river sand in concrete production by examining the potential of quarry dust as a feasible and eco-friendly replacement. Through a combination of experimental evaluation and robust predictive modeling using GEP, this research highlights the suitability of quarry dust in concrete applications and presents GEP as a practical tool for future design. Several key findings from this study can be drawn as follows;

1. The optimal replacement of fine aggregates with 40% quarry dust and 2% superplasticizer exhibited superior properties to conventional mixes.
2. Maximum CS of 17.2 MPa (M15) and 22.3 MPa (M20) was achieved at 28 days, with enhancements of 15% and 13% over control mixes.
3. TS peaked at 3.74 MPa (M15) and 3.83 MPa (M20), demonstrating the beneficial effect of quarry dust and superplasticizer.
4. The addition of superplasticizer accelerated early strength gain and boosted the overall strength by 5–6%.
5. GEP models demonstrated excellent predictive capabilities, achieving R^2 values of 0.96 and 0.94 (CS) and 0.92 and 0.91 (TS) during training and testing, respectively.
6. RMSE and MAE for CS and TS predictions were notably low, underscoring the accuracy of the models.
7. Comparative analysis revealed that GEP models outperformed MLR models, with improvements of 5% (CS) and 6% (TS) in predictive accuracy.
8. Sensitivity analysis (SA) showed that cement content had the greatest influence on CS and TS predictions, while superplasticizers had the least impact.

The findings of this study demonstrate that quarry dust can be effectively used as a replacement for river sand in concrete, addressing environmental concerns related to resource depletion. The validated GEP-based predictive equations and GUI serve as valuable tools for future concrete mix designs, facilitating the incorporation of quarry dust and promoting more sustainable construction practices. To improve the usage of quarry dust in concrete, a few recommendations for future research are discussed in the following section.

Recommendations for future work

Future studies should expand the experimental scope by incorporating higher-grade concrete mixes beyond M20 and conducting more comprehensive testing on the long-term durability of quarry dust-modified concrete, such as resistance to chemical attacks and freeze-thaw cycles. The existing database should be enhanced with additional data points, including different types of supplementary materials, to improve the robustness and generalizability of the models. Advanced machine learning techniques, such as ANN, MEP, and SVM, could be explored to compare performance and accuracy with GEP models. Furthermore, future research should investigate the practical implementation of these models in concrete mix design, ensuring optimization for specific environmental and material conditions.

Data availability

The datasets used during the current study are available from the corresponding author upon reasonable request.

Received: 8 August 2024; Accepted: 7 November 2024
Published online: 26 November 2024

References

1. Collivignarelli, M. C. et al. The Production of Sustainable Concrete with the Use of Alternative Aggregates: A Review, Sustainability., Vol. 12, Page 7903, vol. 12, no. 19, p. 7903, Sep. 2020, doi: (2020). https://doi.org/10.3390/SU12197903
2. Dawood, A. O., AL-Khazraji, H. & Falih, R. S. Jun., Physical and mechanical properties of concrete containing PET wastes as a partial replacement for fine aggregates, Case Studies in Construction Materials, **14**, doi: (2021). https://doi.org/10.1016/J.CSCM.2020.E00482
3. Singh, S., Nagar, R., Agrawal, V., Rana, A. & Tiwari, A. Sustainable utilization of granite cutting waste in high strength concrete. *J. Clean. Prod.* **116**, 223–235. https://doi.org/10.1016/J.JCLEPRO.2015.12.110 (Mar. 2016).
4. Singh, S. & Chourasia, A. Alternative fine aggregates in production of sustainable concrete-A review, doi: (2020). https://doi.org/10.1016/j.jclepro.2020.122089
5. Kou, S. C. & Poon, C. S. Properties of concrete prepared with crushed fine stone, furnace bottom ash and fine recycled aggregate as fine aggregates, Constr Build Mater, vol. 23, no. 8, pp. 2877–2886, Aug. doi: (2009). https://doi.org/10.1016/J.CONBUILDMAT.2009.02.009
6. Kirthika, S. K., Surya, M. & Singh, S. K. Effect of clay in alternative fine aggregates on performance of concrete, doi: (2019). https://doi.org/10.1016/j.conbuildmat.2019.116811
7. Dong, J. F., Xu, Y., Guan, Z. W. & Wang, Q. Y. Freeze-thaw behaviour of basalt fibre reinforced recycled aggregate concrete filled CFRP tube specimens. *Eng. Struct.* **273**, 115088 (2022).
8. Iftikhar, B. et al. Experimental study on the eco-friendly plastic-sand paver blocks by utilising plastic waste and basalt fibers, doi: (2023). https://doi.org/10.1016/j.heliyon.2023.e17107
9. Rauf, A., Tussupbekova, A. & Moon, S. W. Jong-Kim, and Effect of Drying-Wetting Cycles on the Mechanical Behavior of Cement-Treated Soil, 대한토목학회 학술대회, pp. 259–260, Accessed: Mar. 04, 2024. [Online]. Available: (2023). https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11627589
10. Kang, F., Wu, Y., Ma, J. & Li, J. Structural identification of super high arch dams using Gaussian process regression with improved salp swarm algorithm. *Eng. Struct.* **286**, 116150 (2023).
11. Khan, S. U. et al. Effects of different Mineral admixtures on the properties of fresh concrete, doi: (2014). https://doi.org/10.1155/2014/986567
12. Shyam Prakash, K. & Rao, C. H. Study on Compressive Strength of Quarry Dust as Fine Aggregate in Concrete, Advances in Civil Engineering, vol. 2016, doi: (2016). https://doi.org/10.1155/2016/1742769

13. Vijayalakshmi, M., Sekar, A. S. S. & Ganesh Prabhu, G. Strength and durability properties of concrete made with granite industry waste. *Constr. Build. Mater.* **46**, 1–7. https://doi.org/10.1016/J.CONBUILDMAT.2013.04.018 (Sep. 2013).
14. Palanisamy, C. et al. Experimental investigation on self-compacting concrete with waste marble and granite as fine aggregate. *Mater. Today Proc.* **65**, 1900–1907. https://doi.org/10.1016/J.MATPR.2022.05.159 (Jan. 2022).
15. Khan, N. & Chandrakar, R. An experimental study on uses of Quarry Dust to replace sand in concrete. *Int. Res. J. Eng. Technol.*, (2017). Accessed: Oct. 28, 2023. [Online]. Available: www.irjet.net.
16. Gupta, T., Kothari, S., Siddique, S., Sharma, R. K. & Chaudhary, S. Influence of stone processing dust on mechanical, durability and sustainability of concrete, Constr Build Mater, vol. 223, pp. 918–927, Oct. doi: (2019). https://doi.org/10.1016/J.CONBUILDMAT.2019.07.188
17. Kankam, C. K., Meisuh, B. K., Sossou, G. & Buabin, T. K. Stress-strain characteristics of concrete containing quarry rock dust as partial replacement of sand, Case Studies in Construction Materials, vol. 7, pp. 66–72, Dec. doi: (2017). https://doi.org/10.1016/J.CSCM.2017.06.004
18. Shen, W. et al. Cleaner production of high-quality manufactured sand and ecological utilization of recycled stone powder in concrete. *J. Clean. Prod.* **375**, 134146. https://doi.org/10.1016/J.JCLEPRO.2022.134146 (Nov. 2022).
19. Celik, T., Marar, K. & OF CRUSHED STONE DUST ON SOME PROPERTIES OF CONCRETE,. PI1 SOOOS-8846(96)00078 – 6 EFFECTS. *Cem. Concr Res.* **26** (7), 1121–1130 (1996).
20. Ingalkar, R. S. & Harle, S. M. Replacement of natural sand by Crushed Sand in the concrete. *Landsc. Archit. Reg. Plann.* **2** (1), 13–22. https://doi.org/10.11648/j.larp.20170201.12 (2017).
21. Hameed, M. S., Sekar, A. S. S. & PROPERTIES OF GREEN CONCRETE CONTAINING QUARRY ROCK DUST AND MARBLE SLUDGE POWDER AS FINE AGGREGATE., vol. 4, no. 4, 2009, Accessed: Oct. 28, 2023. [Online]. Available: www.arpnjournals.com.
22. Amin, M. N. et al. Split Tensile Strength Prediction of Recycled Aggregate-Based Sustainable Concrete Using Artificial Intelligence Methods, Materials., Vol. 15, Page 4296, vol. 15, no. 12, p. 4296, Jun. 2022, doi: (2022). https://doi.org/10.3390/MA15124296
23. Zhu, Y. et al. Predicting the Splitting Tensile Strength of recycled aggregate concrete using individual and Ensemble Machine Learning Approaches. *Cryst. (Basel).* **12** (5), 569. https://doi.org/10.3390/CRYST12050569/S1 (May 2022).
24. Liu, K., Zheng, J., Dong, S., Xie, W. & Zhang, X. Mixture optimization of mechanical, economical, and environmental objectives for sustainable recycled aggregate concrete based on machine learning and metaheuristic algorithms. *J. Building Eng.* **63**, 105570. https://doi.org/10.1016/J.JOBE.2022.105570 (Jan. 2023).
25. Bansal, T., Talakokula, V. & Mathiyazhagan, K. Equivalent structural parameters based non-destructive prediction of sustainable concrete strength using machine learning models via piezo sensor. *Measurement*. **187**, 110202. https://doi.org/10.1016/J.MEASUREMENT.2021.110202 (Jan. 2022).
26. Ullah, H. S., Khushnood, R. A., Ahmad, J. & Farooq, F. Predictive modelling of sustainable lightweight foamed concrete using machine learning novel approach. *J. Building Eng.* **56**, 104746. https://doi.org/10.1016/J.JOBE.2022.104746 (Sep. 2022).
27. Shah, S. A. R. et al. Predicting Compressive Strength of Blast Furnace Slag and Fly Ash Based Sustainable Concrete Using Machine Learning Techniques: An Application of Advanced Decision-Making Approaches, Buildings 2022, Vol. 12, Page 914, vol. 12, no. 7, p. 914, Jun. doi: (2022). https://doi.org/10.3390/BUILDINGS12070914
28. Dong, J. F., Wang, Q. Y., Guan, Z. W. & Chai, H. K. High-temperature behaviour of basalt fibre reinforced concrete made with recycled aggregates from earthquake waste. *J. Build. Eng.* **48**, 103895 (2022).
29. Ahmad, A. et al. Compressive Strength Prediction via Gene Expression Programming (GEP) and Artificial Neural Network (ANN) for Concrete Containing RCA, Buildings 2021, Vol. 11, Page 324, vol. 11, no. 8, p. 324, Jul. doi: (2021). https://doi.org/10.3390/BUILDINGS11080324
30. Ebid, A. & Deifalla, A. Using Artificial Intelligence Techniques to Predict Punching Shear Capacity of Lightweight Concrete Slabs, Materials Vol. 15, Page 2732, vol. 15, no. 8, p. 2732, Apr. 2022, doi: (2022). https://doi.org/10.3390/MA15082732
31. Iqbal, M. F. et al. Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming. *J. Hazard. Mater.* **384**, 121322. https://doi.org/10.1016/J.JHAZMAT.2019.121322 (Feb. 2020).
32. Nazari, A. & Torgal, F. P. Modeling the compressive strength of geopolymeric binders by gene expression programming-GEP, Expert Syst Appl, vol. 40, no. 14, pp. 5427–5438, Oct. doi: (2013). https://doi.org/10.1016/J.ESWA.2013.04.014
33. Iftikhar Faraz, M. et al. A comprehensive GEP and MEP analysis of a cement-based concrete containing metakaolin, Structures, vol. 53, pp. 937–948, Jul. doi: (2023). https://doi.org/10.1016/J.ISTRUC.2023.04.050
34. Sundaralingam, K., Peiris, A., Anburuvel, A. & Sathiparan, N. Quarry dust as river sand replacement in cement masonry blocks: Effect on mechanical and durability characteristics. *Mater. (Oxf).* **21**, 101324. https://doi.org/10.1016/J.MTLA.2022.101324 (Mar. 2022).
35. Liu, Y. et al. Variable fatigue loading effects on corrugated steel box girders with recycled concrete. *J. Constr. Steel Res.* **215**, 108526 (2024).
36. Sika® ViscoCrete®-3110 | Ultra High Range Water Reducers. Accessed: Oct. 09, 2024. [Online]. Available: https://pak.sika.com/en/construction/concrete-technology/ready-mixed-concrete-admixtures/ultra-high-rangewaterreducers/sika-viscocrete-3110.html
37. C143/C143M Standard Test Method for Slump of Hydraulic-Cement Concrete. Accessed: Oct. 08, 2024. [Online]. Available: https://www.astm.org/c0143_c0143m-12.html
38. C192/C192M Standard Practice for Making. and Curing Concrete Test Specimens in the Laboratory. Accessed: Oct. 08, 2024. [Online]. Available: https://www.astm.org/c0192_c0192m-14.html
39. C39/C39M Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens. Accessed: Oct. 08, 2024. [Online]. Available: https://www.astm.org/c0039_c0039m-21.html
40. Ferreira, C. Gene Expression Programming: a New Adaptive Algorithm for Solving Problems, Feb. Accessed: Jul. 31, 2023. [Online]. Available: (2001). https://arxiv.org/abs/cs/0102027v3
41. Asif, U. et al. Predictive Modeling and Experimental Validation for Assessing the Mechanical Properties of Cementitious Composites Made with Silica Fume and Ground Granulated Blast Furnace Slag, Buildings 2024, Vol. 14, Page 1091, vol. 14, no. 4, p. 1091, Apr. doi: (2024). https://doi.org/10.3390/BUILDINGS14041091
42. Khan, M. & Javed, M. F. Towards sustainable construction: machine learning based predictive models for strength and durability characteristics of blended cement concrete. *Mater. Today Commun.* **37**, 107428. https://doi.org/10.1016/J.MTCOMM.2023.107428 (Dec. 2023).
43. Javed, M. F., Siddiq, B., Onyelowe, K., Khan, W. A. & Khan, M. Metaheuristic optimization algorithms-based prediction modeling for titanium dioxide-assisted photocatalytic degradation of air contaminants. *Results Eng.* **23**, 102637. https://doi.org/10.1016/J.RINENG.2024.102637 (Sep. 2024).
44. Asif, U. et al. Apr., Predicting the Mechanical properties of plastic concrete: an optimization method by using genetic programming and ensemble learners, Case Studies in Construction Materials, p. e03135, doi: (2024). https://doi.org/10.1016/J.CSCM.2024.E03135
45. Asif, U., Javed, M. F., Alyami, M. & Hammad, A. W. Performance evaluation of concrete made with Plastic Waste using Multi-expression Programming. *Mater. Today Commun.* 108789. https://doi.org/10.1016/J.MTCOMM.2024.108789 (Apr. 2024).
46. Jalal, F. E. et al. ANN-based swarm intelligence for predicting expansive soil swell pressure and compression strength, Scientific Reports 2024 14:1, vol. 14, no. 1, pp. 1–34, Jun. doi: (2024). https://doi.org/10.1038/s41598-024-65547-7
47. Dehwah, H. A. F. Mechanical properties of self-compacting concrete incorporating quarry dust powder, silica fume or fly ash. *Constr. Build. Mater.* **26** (1), 547–551. https://doi.org/10.1016/J.CONBUILDMAT.2011.06.056 (Jan. 2012).

48. Pereira, P., Evangelista, L. & De Brito, J. The effect of superplasticizers on the mechanical performance of concrete made with fine recycled concrete aggregates. *Cem. Concr Compos.* **34** (9), 1044–1052. https://doi.org/10.1016/J.CEMCONCOMP.2012.06.009 (Oct. 2012).
49. Shamsabadi, E. A., Ghalehnovi, M., de Brito, J. & Khodabakhshian, A. Performance of concrete with Waste Granite Powder: the Effect of Superplasticizers. *Appl. Sci. 2018*. **8** (10), 1808. https://doi.org/10.3390/APP8101808 (Oct. 2018). Page 1808.
50. Bahoria, B. V., Parbat, D. K., Nagarnaik, P. B. & Waghe, U. P. Sustainable utilization of quarry dust and waste plastic fibers as a sand replacement in conventional concrete.
51. Quan Tran, V., Quoc Dang, V. & Si Ho, L. Evaluating compressive strength of concrete made with recycled concrete aggregates using machine learning approach, doi: (2022). https://doi.org/10.1016/j.conbuildmat.2022.126578
52. Golbraikh, A. & Tropsha, A. Beware of q2! *J. Mol. Graph Model.* **20** (4), 269–276. https://doi.org/10.1016/S1093-3263(01)00123-1 (Jan. 2002).
53. Alavi, A. H., Ameri, M., Gandomi, A. H. & Mirzahosseini, M. R. Formulation of flow number of asphalt mixes using a hybrid computational method, Constr Build Mater, vol. 25, no. 3, pp. 1338–1355, Mar. doi: (2011). https://doi.org/10.1016/J.CONBUILDMAT.2010.09.010
54. Rahla, K. M., Mateus, R. & Bragança, L. Comparative sustainability assessment of binary blended concretes using supplementary Cementitious materials (SCMs) and ordinary Portland Cement (OPC). *J. Clean. Prod.* **220**, 445–459. https://doi.org/10.1016/j.jclepro.2019.02.010 (2019).
55. Cartuxo, F., De Brito, J., Evangelista, L., Jiménez, J. R. & Ledesma, E. F. Rheological behaviour of concrete made with fine recycled concrete aggregates – Influence of the superplasticizer, Constr Build Mater, vol. 89, pp. 36–47, Aug. doi: (2015). https://doi.org/10.1016/J.CONBUILDMAT.2015.03.119
56. Chitkeshwar, A. K. & Naktode, P. L. Concrete with rock quarry dust with partial replacement of fine aggregate. *Mater. Today Proc.* **62**, 6455–6459. https://doi.org/10.1016/J.MATPR.2022.04.195 (Jan. 2022). no. P12.
57. Jin, R., Chen, Q. & Soboyejo, A. B. O. Non-linear and mixed regression models in predicting sustainable concrete strength. *Constr. Build. Mater.* **170**, 142–152. https://doi.org/10.1016/J.CONBUILDMAT.2018.03.063 (May 2018).
58. Rostami, A., Hemmati-Sarapardeh, A. & Mohammadi, A. H. Jul., Estimating n-tetradecane/bitumen mixture viscosity in solvent-assisted oil recovery process using GEP and GMDH modeling approaches, (2019). https://doi.org/10.1080/10916466.2018.1531885, vol. 37, no. 14, pp. 1640–1647, doi: 10.1080/10916466.2018.1531885.
59. Ray, S. et al. Predicting the strength of concrete made with stone dust and nylon fiber using artificial neural network, Heliyon, vol. 8, no. 3, p. e09129, Mar. doi: (2022). https://doi.org/10.1016/J.HELIYON.2022.E09129
60. Khan, K. et al. Prediction Models for Estimating Compressive Strength of Concrete Made of Manufactured Sand Using Gene Expression Programming Model, Materials., Vol. 15, Page 5823, vol. 15, no. 17, p. 5823, Aug. 2022, doi: (2022). https://doi.org/10.3390/MA15175823

Author contributions

A.R: Conceptualization, Methodology, Investigation, Validation, Writing-Original Draft, Writing – Review and Editing, Visualization.U.A: Conceptualization, Methodology, Software, Machine learning, Data Curation, Investigation, Validation, Writing-Original Draft, Writing – Review and Editing, Visualization.K.O: Supervision, Project administration, Funding acquisition, Resources.F.J: Conceptualization, Methodology, Validation, Investigation, Writing – Review and Editing, Supervision, Project administration, Funding acquisition, Resources.H.A: Project administration, Funding acquisition, Investigation, Validation.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to U.A. or K.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.