frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 27 August 2024
DOI 10.3389/fbuil.2024.1446597

OPEN ACCESS

EDITED BY
Abdullah Mohsen Zeyad,
Jazan University, Saudi Arabia

REVIEWED BY
Panagiotis G. Asteris,
School of Pedagogical and Technological Education, Greece
Amin Al-Fakih,
King Fahd University of Petroleum and Minerals, Saudi Arabia

*CORRESPONDENCE
Kennedy C. Onyelowe,
kennedychibuzor@kiu.ac.ug,
konyelowe@mouau.edu.ng,
k.onyelowe@go.uop.gr,
konyelowe@gmail.com
Ahmed M. Ebid,
ahmed.abdelkhaleq@fue.edu.eg

RECEIVED 10 June 2024
ACCEPTED 30 July 2024
PUBLISHED 27 August 2024

CITATION
Onyelowe KC, Ebid AM, Fernandez Vinueza DF, Estrada Brito NA, Velasco N, Buñay J, Muhodir SH, Imran H and Hanandeh S (2024) Estimating the compressive strength of lightweight foamed concrete using different machine learning-based symbolic regression techniques.
Front. Built Environ. 10:1446597.
doi: 10.3389/fbuil.2024.1446597

Estimating the compressive strength of lightweight foamed concrete using different machine learning-based symbolic regression techniques

Kennedy C. Onyelowe [1,2]*, Ahmed M. Ebid [3]*, Danilo Fernando Fernandez Vinueza [4], Néstor Augusto Estrada Brito [5], Nancy Velasco [5], Jorge Buñay [6], Sabih Hashim Muhodir [7], Hamza Imran [8] and Shadi Hanandeh [9]

[1]Department of Civil Engineering, Michael Okpara University of Agriculture, Umudike, Nigeria, [2]Department of Civil Engineering, Kampala International University, Kampala, Uganda, [3]Department of Civil Engineering, Future University in Egypt, New Cairo, Egypt, [4]Facultad de Administracion de Empresas, Escuela Superior Politecnica de Chimborazo (ESPOCH), Riobamba, Ecuador, [5]Facultad de Informatica y Electronica, Escuela Superior Politecnica de Chimborazo (ESPOCH), Riobamba, Ecuador, [6]Facultad de Mecanica, Escuela Superior Politecnica de Chimborazo (ESPOCH), Riobamba, Ecuador, [7]Department of Architectural Engineering, Cihan University Erbil, Kurdistan Region, Iraq, [8]Department of Construction and Project, Al-karkh University of Science, Baghdad, Iraq, [9]Department of Civil Engineering, Al-Balqa Applied University, As-Salt, Libya

The development of concrete with excellent water and frost resistance providing high level of sound and thermal insulation has triggered the formulation of foamed concrete. However, multiple laboratory studies are required to produce reasonable data to design the relevant codes and mathematics with which design of mixes is made easier at low cost. In this research paper, the artificial intelligence (AI)-based symbolic regression technique estimation of the compressive strength of foamed concrete has been reported. Foamed concrete has been a subject of serious research in sustainable built-environment due to its lightweight and structural functionality. In this research work, data gathering method was applied to gather a globally representative data base comprising concrete density to water density (concrete density g/cm3) (γ/γw), water-cement ratio (W/C), and sand-cement ratio (S/C) as input variable and the compressive strength (Fc) as the study output. The dimensionless factors have been derived to eliminate data handling complexities and improve model performances. The 230 data entries from foamed concrete mixes were partitioned into 75% and 25% for training and validation data, respectively. At the end of the model execution, it was found that the response surface methodology (RSM) produced a symbolic closed-form equation like the genetic programming (GP), evolutionary polynomial regression (EPR), and the group method of data-handling-neural

Abbreviations: SSE, sum of squared error; MAE, mean absolute error; MSE, mean squared error; RMSE, root mean squared error; R2, coefficient of determination; RSM, response surface methodology; GP, genetic programming; EPR, evolutionary polynomial regression; GMDH-NN, group method of data handling-neural networks; Adeq., adequacy; Prec., precision; 3D, three dimensional; Fc, compressive strength of concrete; C.V., coefficient of variance; Std. Dev., Standard deviation; γ/γw, concrete density to water density (concrete density g/cm3); W/C, water-cement ratio; S/C, sand-cement ratio.

network (GMDH-NN). Even though the RSM closed with a minimum error, the GP, EPR and GMDH-NN were faster in runtime. The overall outcomes show that the GP outclassed the EPR, RSM and the GMDH-NN, though with minor margin. Meanwhile the EPR produced the highest outliers from the ±25% test of accuracy envelope. Overall, the present models outperformed those reported in the literature due the parameter reduction through dimensionless factors derivation and provided a decisive model to predict the Fc of foamed concrete.

KEYWORDS

foamed concrete, artificial intelligence (AI), symbolic regression methods, sustainable concrete structures, lightweight concrete (LWC)

1 Introduction

Foamed concrete, also known as cellular concrete or aerated concrete is a lightweight building material that consists of cement, water, and stable foam (Othman et al., 2021). The foam is created by introducing air or gas into cement slurry, resulting in a cellular structure within the concrete (Yang et al., 2022). The foam provides a high volume of stable air voids, which gives foamed concrete its low density and lightweight properties (Kozłowski and Kadela, 2018). The density of foamed concrete typically ranges from 400 kg/m³ to 1,600 kg/m³, making it significantly lighter than traditional concrete (Pan et al., 2007; Yang et al., 2022). Foamed concrete low density makes it useful in applications where weight reduction is important, such as for insulation, filling voids, and reducing structural loads (Wu et al., 2013). Foamed concrete has good thermal insulation properties, making it suitable for insulating walls, roofs, and floors. Due to its low density and chemical composition, foamed concrete exhibits excellent fire resistance and can be used as fire barriers or in fire-rated construction. The cellular structure of foamed concrete helps to absorb sound and reduce noise transmission, making it useful for soundproofing applications (Ayyanar et al., 2023). Foamed concrete can be easily pumped or poured into various shapes and sizes, making it suitable for filling cavities, voids, and irregular spaces. The presence of foam in the concrete mixture improves its workability, making it easier to handle and place (Ayyanar et al., 2023). Foamed concrete requires less raw material compared to traditional concrete, resulting in lower energy consumption and reduced carbon footprint (Bian et al.). Foamed concrete is commonly used in construction for a variety of applications, including lightweight blocks, precast panels, thermal insulation, filling underground voids, trench reinstatement, and road sub-bases (Shawnim and Mohammad, 2019). It is also used in geotechnical applications, such as lightweight backfill and slope stabilization (Kozłowski and Kadela, 2018). Overall, foamed concrete offers a lightweight, durable, and energy-efficient alternative to traditional concrete, with a wide range of applications in the construction industry (Wu et al., 2013). The mix proportion and production process of foamed concrete can vary depending on the specific requirements and application. However, it can be provided with a general guideline for producing foamed concrete (Pan et al., 2007; Yang et al., 2022). The mix proportion will depend on the desired density and strength of the foamed concrete (Ebid and Deifalla, 2022). Cement content could range from 200 to 600 kg/m³, depending on the application and desired strength (Shawnim and Mohammad, 2019). Water-to-cement ratio is usually between 0.3 and 0.6, depending on the consistency required (Ebid and Deifalla, 2022). Fine aggregate content generally, ranges from 400 to 1,200 kg/m³. Foaming agent dosage depends on the specific foaming agent used and the desired foam stability (Bian et al.; Onyelowe et al., 2022a). The production process of foamed concrete involves mixing the materials and generating foam. The foamed concrete should be workable and easily pourable or pumpable. Pour or place the foamed concrete: Transfer the foamed concrete to the desired location using a pump, mixer, or by manual pouring. Once placed, allow the foamed concrete to cure and set according to standard concrete curing practices (Yang et al., 2022). It's important to note that the mix proportion and production process may vary based on the specific requirements, local materials, equipment availability, and climate conditions (Shawnim and Mohammad, 2019). It's recommended to consult with experts or refer to technical guidelines for more precise mix designs and production methods (Ayyanar et al., 2023). Foamed concrete offers several structural benefits that make it a suitable choice for various construction applications (Onyelowe et al., 2022a). Foamed concrete has a significantly lower density compared to traditional concrete (Onyelowe et al., 2022b). This lightweight nature reduces structural loads and allows for the construction of lighter structures (Onyelowe Kennedy C. et al., 2022). It is particularly beneficial in applications where weight reduction is important, such as in high-rise buildings or structures with weak foundations (Pan et al., 2007; Kozłowski and Kadela, 2018; Yang et al., 2022). Foamed concrete has excellent thermal insulation properties (Kozłowski and Kadela, 2018). Its cellular structure traps air within the material, creating an effective barrier against heat transfer (Wu et al., 2013). This makes it useful in constructing energy-efficient buildings and structures that require insulation for temperature control, such as walls, roofs, and floors (Bian et al.). Fire resistance: Foamed concrete exhibits excellent fire resistance due to its low density and composition (Onyelowe et al., 2022b). It has a high fire rating and can act as a fire barrier or provide fire protection in structural elements. Foamed concrete's resistance to high temperatures and ability to insulate against heat make it valuable in fire-resistant construction (Onyelowe et al., 2022b). Sound insulation: The cellular structure of foamed concrete helps to absorb sound and reduce noise transmission (Onyelowe Kennedy C. et al., 2022). This makes it an effective material for soundproofing applications, such as constructing walls between rooms or buildings where noise reduction is desired (Yang et al., 2022). Foamed concrete's lightweight nature can improve the seismic

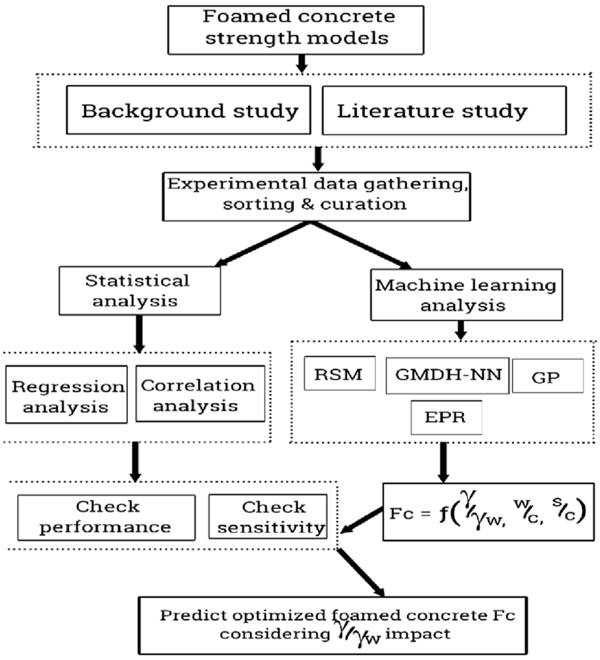

FIGURE 1
Research flowchart.

performance of structures (Ebid and Deifalla, 2022; Yang et al., 2022; Ayyanar et al., 2023). It reduces the mass of the building, which can help absorb seismic energy and decrease the forces acting on the structure during an earthquake. This can enhance the overall structural integrity and safety of the building (Pan et al., 2007). The cellular structure of foamed concrete provides it with improved impact resistance compared to traditional concrete (Bian et al.). It can absorb and distribute impact forces, making it suitable for applications where impact resistance is important, such as in blast-resistant structures or protective barriers (Shawnim and Mohammad, 2019). Foamed concrete has a lower self-weight, which minimizes settlement and soil consolidation (Ebid and Deifalla, 2022). This can be advantageous in applications where settlement control is critical, such as in lightweight fill materials or in areas with weak or compressible soils (Wu et al., 2013). It's important to note that the specific structural benefits of foamed concrete will depend on factors such as the mix design, density, and application (Onyelowe Kennedy C. et al., 2022). Engineering analysis and design considerations should be undertaken to ensure that foamed concrete is used appropriately and meets the structural requirements of the specific project (Wu et al., 2013). Foamed concrete exhibits certain behaviors under different loading conditions (Ayyanar et al., 2023). Foamed concrete typically performs well under axial loading, which refers to the applied load along the longitudinal axis of a structural element (Wu et al., 2013). It can withstand compressive forces and distribute them evenly throughout the material (Kozłowski and Kadela, 2018; Onyelowe Kennedy C. et al., 2022). The lightweight nature of foamed concrete helps to reduce the overall load on the structure, making it suitable for applications where weight reduction is important (Wu et al., 2013). Foamed concrete may have limited resistance to lateral loads, such as wind or earthquake forces. Its lightweight nature and low modulus of elasticity can result in lower stiffness and reduced lateral load resistance compared to denser concrete (Pan et al., 2007). Additional reinforcement or structural systems like shear walls or bracing might be necessary to enhance its resistance to lateral loads (Wu et al., 2013). Foamed concrete's response to seismic loading depends on various factors, including the density, reinforcement, and structural configuration (Shawnim and Mohammad, 2019). It generally has lower strength and stiffness compared to traditional concrete, but its lightweight nature can provide some advantages (Bian et al.). Foamed concrete can absorb and dissipate seismic energy to some extent due to its ability to deform and absorb vibrations (Ebid and Deifalla, 2022). However, it's crucial to conduct proper engineering analysis and design to ensure that the foamed concrete structures meet the seismic design requirements (Othman et al., 2021; Yang et al., 2022; Kozłowski and Kadela, 2018; Pan et al., 2007; Wu et al., 2013; Ayyanar et al., 2023; Bian et al.). Dynamic Loading: Foamed concrete exhibits different behavior under dynamic loading, such as impact or vibration, compared to static loading (Ebid and Deifalla, 2022). Its ability to absorb and distribute impact forces makes it suitable for applications where impact resistance is important, such as blast-resistant structures or protective barriers (Pan et al., 2007). However, the low density and stiffness of foamed concrete may result in higher deflections and vibrations under dynamic loading, requiring special considerations in design (Wu et al., 2013; Onyelowe Kennedy C. et al., 2022; Deifalla et al., 2020). It's important to note that the specific behavior of foamed concrete under different loading conditions can vary depending on factors such as the mix design, density, reinforcement, and structural configuration (Bian et al.). Proper engineering analysis, design, and testing should be carried out to assess the suitability of foamed concrete in specific structural applications and ensure that it meets the required performance standards (Pan et al., 2007). Foamed concrete offers several sustainability advantages compared to traditional concrete (Bian et al.). Reduced Embodied Carbon: Foamed concrete typically requires less cement compared to traditional concrete. Since cement production is a significant source of carbon dioxide (CO_2) emissions, reducing cement content results in lower embodied carbon in foamed concrete (Wu et al., 2013). The lower density of foamed concrete also means that less raw material is required overall, further reducing the environmental impact (Onyelowe Kennedy C. et al., 2022). The production process of foamed concrete requires less energy compared to traditional concrete (Onyelowe et al., 2022b). Foam generation requires relatively low energy input, and the lightweight nature of foamed concrete reduces transportation and handling energy (Onyelowe et al., 2022a). This results in lower energy consumption during both the manufacturing and construction phases (Onyelowe Kennedy C. et al., 2022). Foamed concrete can be produced on-site using locally available materials, reducing transportation and waste associated with the supply chain (Wu et al., 2013). Additionally, foamed concrete can be easily poured or pumped into desired shapes and sizes, minimizing material waste during construction (Wu et al., 2013). The lightweight nature of foamed concrete also reduces the need for heavy machinery during construction, decreasing fuel consumption and emissions (Kozłowski and Kadela, 2018; Bian et al.). Foamed concrete's cellular structure provides excellent thermal insulation

TABLE 1 Statistical analysis of collected database.

	γ/γ_w	W/C	S/C	Fc
	—	—	—	MPa
	Training set			
Max.	2.07	0.70	3.61	51.20
Min	0.43	0.26	0.00	1.50
Avg	1.53	0.41	0.99	24.15
SD	0.43	0.12	0.68	13.98
Var	0.28	0.30	0.69	0.58
	Validation set			
Max.	2.00	0.83	2.93	48.48
Min	0.62	0.30	0.00	1.80
Avg	1.56	0.43	1.07	23.93
SD	0.39	0.13	0.68	13.78
Var	0.25	0.30	0.64	0.58

TABLE 2 Pearson correlation matrix.

	γ/γ_w	W/C	S/C	Fc
γ/γ_w	1.00			
W/C	−0.52	1.00		
S/C	0.34	−0.10	1.00	
Fc	0.88	−0.58	0.05	1.00

properties (Onyelowe Kennedy C. et al., 2022; Humberg et al., 2019; Marković, 2006; Alyamac et al., 2017; Sambruno et al., 2019; V Bayramov et al., 2022; Bezerra et al., 2008; Box et al., 2005; Kutner et al., 2005; Hoffman et al., 1983). It reduces heat transfer through walls, roofs, and floors, resulting in improved energy efficiency and reduced reliance on heating and cooling systems (Yang et al., 2022). This can lead to energy savings and lower greenhouse gas emissions associated with building operations (Onyelowe et al., 2022a). Foamed concrete has good durability and can withstand environmental exposure (Kozłowski and Kadela, 2018; Asteris et al., 2019; Apostolopoulou et al., 2020; Asteris P. G. et al., 2021; Armaghani and Asteris, 2021; Salami et al., 2022; Shang et al., 2022). Its low permeability and resistance to moisture can reduce the risk of corrosion in reinforced structures, enhancing their longevity (Onyelowe Kennedy C. et al., 2022). Durable structures require less maintenance and repair, resulting in reduced material consumption and waste over the life cycle (Onyelowe et al., 2022a; Onyelowe et al., 2022b; Onyelowe Kennedy C. et al., 2022). Recycling and Reuse: Foamed concrete can be crushed and used as a recycled aggregate in future construction projects (Bian et al.; Bezerra et al., 2008; Daniel et al., 2024; Bardhan et al., 2024; Zhang et al., 2024; Kumar et al., 2023; Alshaeer et al., 2023). This promotes circular economy principles by reducing the demand for virgin materials and minimizing waste generation (Ayyanar et al., 2023).

2 Research significance

The work "estimating the compressive strength of lightweight foamed concrete using different machine learning-based symbolic regression techniques," holds several significant implications and contributions to both the fields of construction material science and machine learning. In the area of material science advancements, lightweight foamed concrete is a material of growing interest due to its potential applications in construction, insulation, and other engineering fields. Accurately estimating its compressive strength was crucial for designing and utilizing it effectively in various structural concrete applications. By employing machine learning-based symbolic regression techniques, the project aims to provide more accurate and efficient methods for predicting the compressive strength of lightweight foamed concrete. This leads to advancements in concrete material science by providing better understanding and control over the properties of this material.

In the area of efficiency in material testing, traditional methods for determining the compressive strength of concrete involve time-consuming and labor-intensive experimental procedures. By leveraging machine learning techniques, particularly symbolic regression, the project seeks to streamline this process, potentially reducing the time and resources required for testing. This efficiency can benefit researchers, engineers, and industries involved in the development and application of lightweight foamed concrete.

10.3389/fbuil.2024.1446597

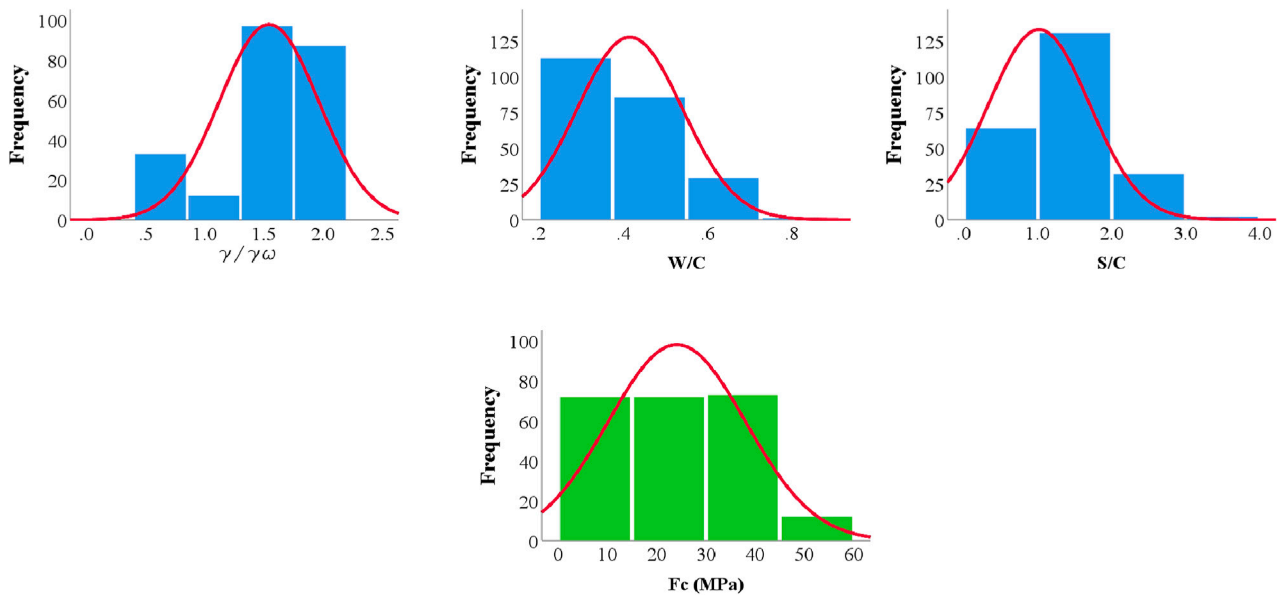

FIGURE 2
Distribution histograms for inputs (in blue) and outputs (in green).

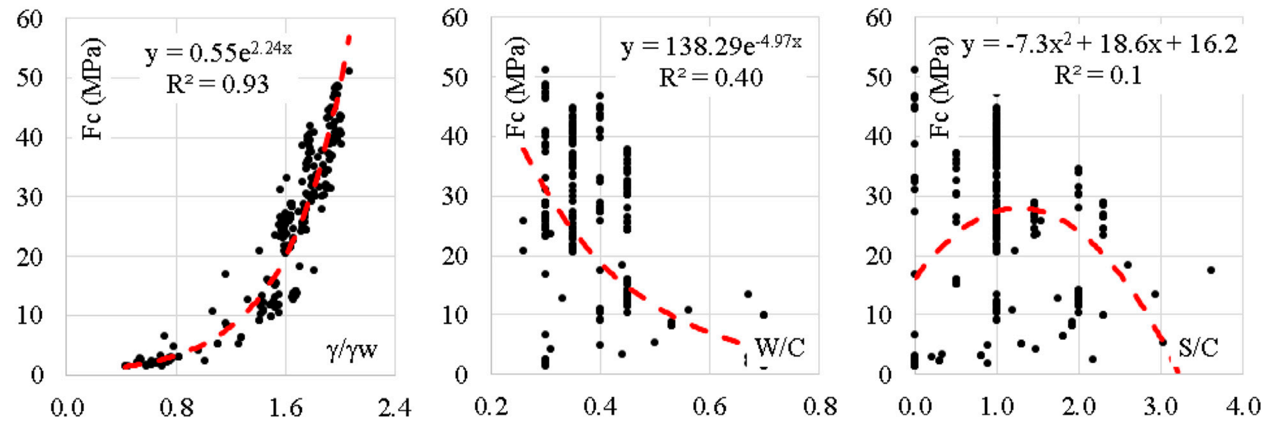

FIGURE 3
Relations between inputs and output (Fc).

Considering insights into concrete behavior, the use of machine learning techniques allows for the analysis of complex relationships and patterns within the data that may not be apparent through traditional statistical methods. By applying symbolic regression, which can uncover mathematical expressions representing these relationships, the project may reveal novel insights into the factors influencing the compressive strength of lightweight foamed concrete. These insights could contribute to a deeper understanding of concrete behavior and aid in the development of more robust predictive models. However, on the generalizability and transferability of the research, machine learning-based symbolic regression techniques have the potential to be applied beyond the specific context of lightweight foamed concrete. The methodologies and insights gained from this project may be transferable to other materials with similar characteristics or even to different domains altogether. This generalizability enhances the broader impact of the research and underscores the importance of exploring innovative machine learning approaches in various scientific and engineering disciplines. Also, the project represents a fusion of expertise from both material science and machine learning domains. By integrating knowledge and methodologies from these disparate fields, the research not only advances our understanding of lightweight foamed concrete but also demonstrates the value of interdisciplinary collaboration in tackling complex scientific and engineering challenges. This interdisciplinary approach may serve as a model for future research endeavors seeking to address multifaceted problems. Overall, the project holds significance in advancing material science, improving efficiency in material testing, providing insights into concrete behavior, fostering generalizable methodologies, and promoting interdisciplinary collaboration.

10.3389/fbuil.2024.1446597

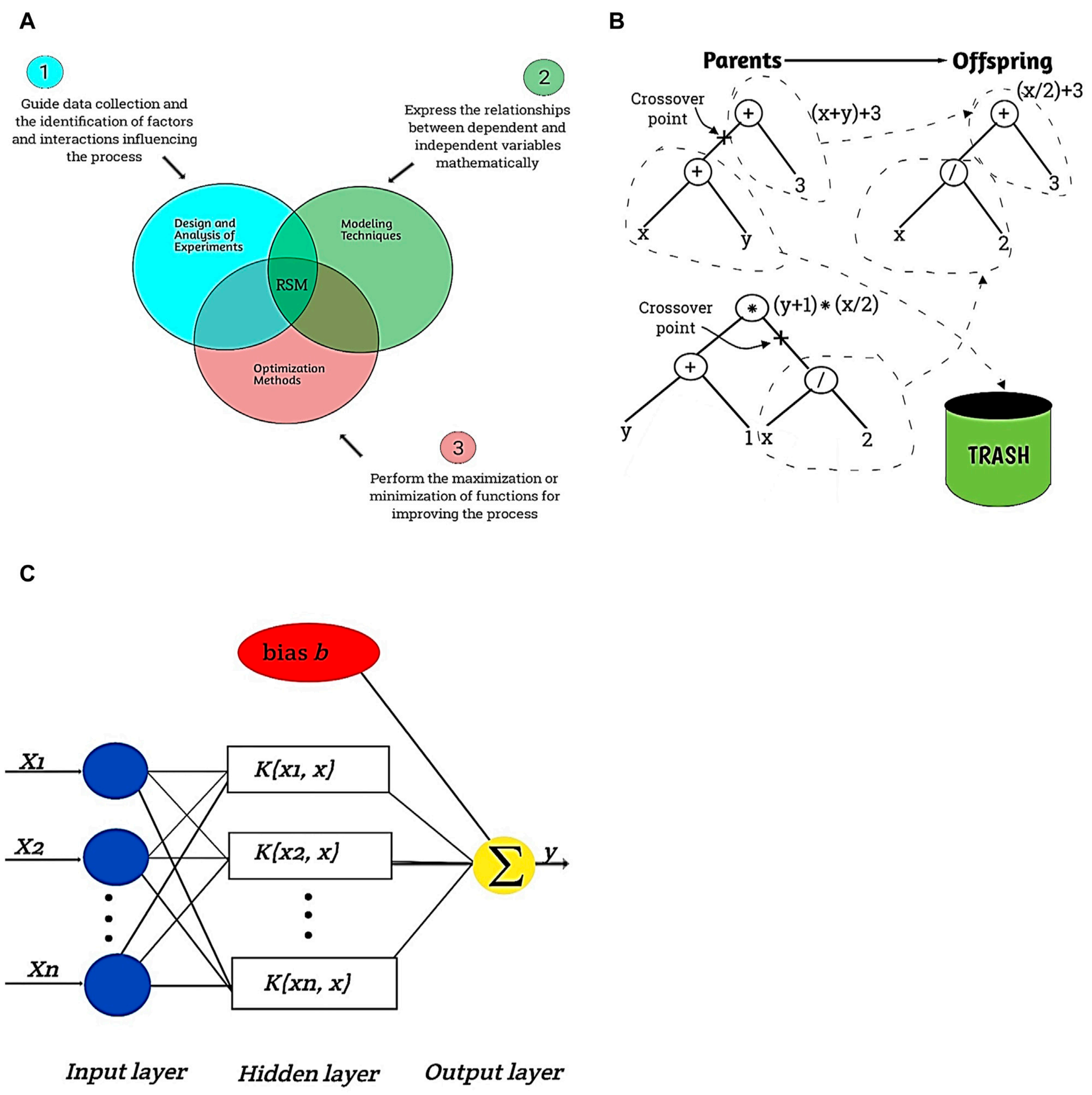

FIGURE 4
(A) Response surface methodology overview, **(B)** Genetic programming framework, and **(C)** Group Method of Data Handling (GMDH) framework.

3 Literature reviews

While foamed concrete offers sustainability advantages, it's important to consider the specific project requirements, local conditions, and life cycle assessment for a comprehensive evaluation (Bian et al.). Some aspects, such as the energy required for foam generation or the environmental impact of foaming agents, should also be considered and optimized for a more sustainable application of foamed concrete (Ayyanar et al., 2023). Symbolic regression techniques have been applied in the study of the strength behavior of foamed concrete. Some of such are the response surface methodology (RSM), ANN, GP and EPR (Shawnim and Mohammad, 2019; Ebid and Deifalla, 2022; Ayyanar et al., 2023; Bian et al.). Response Surface Analysis (RSA) allows researchers to examine intricate psychological phenomena, such as determining whether the alignment between two psychological categories is linked to elevated values in an outcome variable (Humberg et al., 2019). The utilization of RSA in the field of personality and social psychology has been on the rise (Marković, 2006). However, certain oversimplifications and misconceptions have raised concerns over the validity of the findings reported in published literature (Humberg et al., 2019). In this paper, we elucidate the foundational mathematical principles necessary for comprehending RSA outcomes, and we furnish a

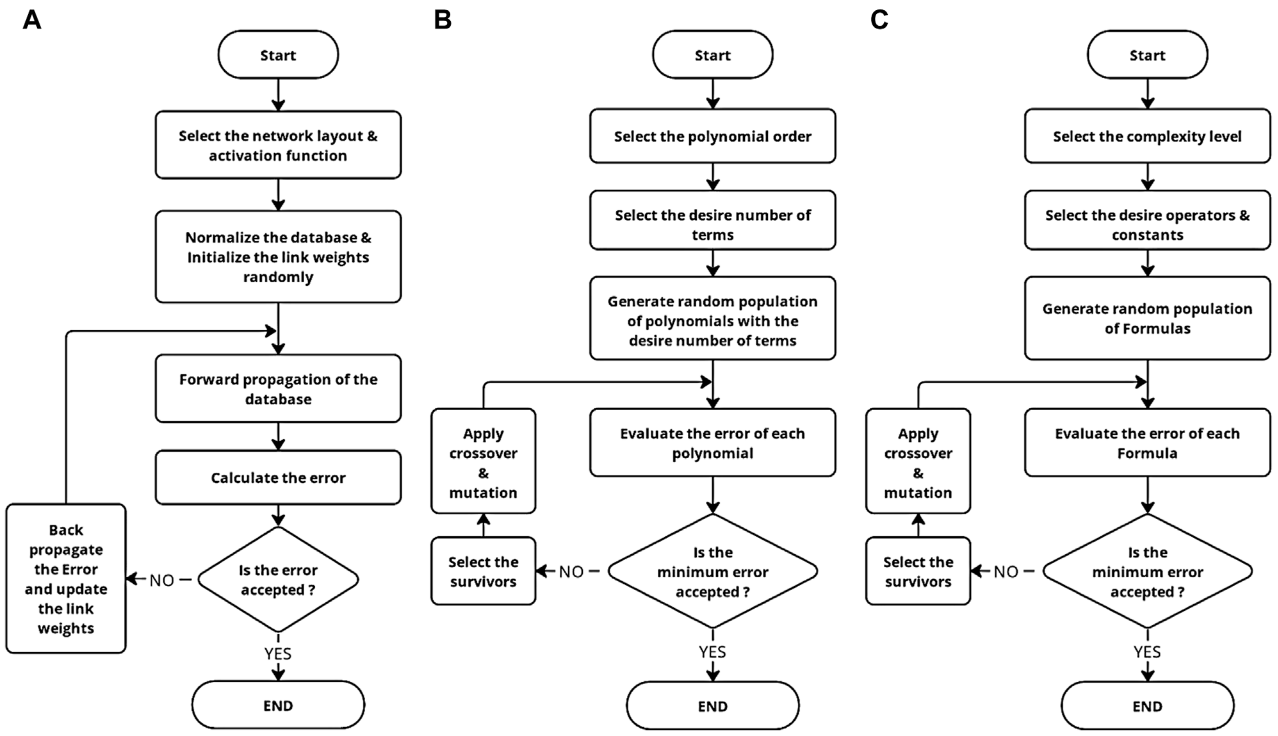

FIGURE 5
(A) GMDH-NN. **(B)** EPR. **(C)** GP. Flowcharts for different (AI)-based symbolic regression techniques.

comprehensive guide for accurately discerning congruence effects. Humberg et al. (Humberg et al., 2019) elucidated two prevailing mistakes by demonstrating that the evaluation of a solitary RSA parameter is insufficient in determining the presence of a congruence effect. Furthermore, we establish that in cases where a congruence effect is observed, RSA is incapable of discerning the relative superiority or inferiority of an interpreter mismatch in one direction compared to a mismatch in the (underestimation) opposite direction (Marković, 2006). It is anticipated that this involvement will augment the strength and robustness of experimental research that employ this potent methodology. Response Surface Methodology (RSM) is an influential experimental design procedure utilized for the analysis and modeling of issues where multiple variables have an impact on a response of interest (Marković, 2006; Alyamac et al., 2017; Sambruno et al., 2019; V Bayramov et al., 2022; Bezerra et al., 2008). While the utilization of this approach has been extensively employed for the purpose of optimizing experimental processes, its application within the concrete industry has been relatively restricted. In their study, Khayat et al. (Alyamac et al., 2017) employed a composite central response surface methodology to evaluate the impact of various parameters of self-consolidating concrete (SCC) mixtures on multiple responses, including V-funnel flow time, filling capacity, and slump flow (Hoffman et al., 1983; Box et al., 2005; Kutner et al., 2005; Salami et al., 2022; Shang et al., 2022). In their study, Simon et al. (Sambruno et al., 2019) employed the Response Surface Methodology to optimize the composition of high performance concrete mixtures. The objective was to achieve the highest possible compressive strength while concurrently minimizing chloride permeability and cost. Bayramov and colleagues (V Bayramov et al., 2022) proposed an analytical model utilizing response surface methodology to enhance fracture parameters in reinforced steel fiber concretes, aiming to enhance their ductility.

However, in similar research papers, density and compressive strength relationship have been studied (Othman et al., 2021), strength, durability, and microstructure relationship in foamed concretes have also been investigated (Yang et al., 2022), the mechanical characteristics of lightweight foamed concrete was equally studied (Kozłowski and Kadela, 2018), the material components impact such as cement, sand, mineral admixtures, etc. on the behavior of foamed concrete has also been investigated (Pan et al., 2007), and the impact of polystyrene on the strength performance of foamed concrete was studied and report as well (Wu et al., 2013). Moreso, the strength properties of a normal foamed concrete produced from the primary concrete components was studied (Ayyanar et al., 2023), as well as the pore size distribution impact on the foamed concrete compressive strength has been modeled and the porosity analysis using SEM images in relation to the compressive strength was also studied (Shawnim and Mohammad, 2019). Various other research works have studied different forms of concrete materials including the application of biocementation in foamed concrete and also the application of different machine learning techniques to design sustainable techniques in concrete and foamed concrete production (Asteris et al., 2019; Apostolopoulou et al., 2020; Asteris P. G. et al., 2021; Armaghani and Asteris, 2021; Nguyen et al., 2021; Emad et al., 2022; Alshaeer et al., 2023; Kumar et al., 2023; Alkayem et al., 2024; Bardhan et al., 2024; Daniel et al., 2024; Zhang et al., 2024). Conversely, many more investigations on the modeling of

TABLE 3 Response surface model tab.

File version	13.0.5.0		
Study type	Response surface	Subtype	Randomized
Design type	Blank spreadsheet	Runs	229.00
Design model	Quadratic	Blocks	No Blocks
Build time (ms)	1.0000		

TABLE 4 RSM factors and numeric coding statistics.

Factor	Name	Units	Type	Sub-type	Minimum	Maximum	Coded low	Coded high	Mean	Std. Dev
A	g/gw	—	Numeric	Continuous	0.4300	2.07	−1 ↔ 0.43	+1 ↔ 2.07	1.54	0.4204
B	W/C	—	Numeric	Continuous	0.2600	0.8300	−1 ↔ 0.26	+1 ↔ 0.83	0.4155	0.1249
C	S/C	—	Numeric	Continuous	0.0000	3.61	−1 ↔ 0.00	+1 ↔ 3.61	1.01	0.6855

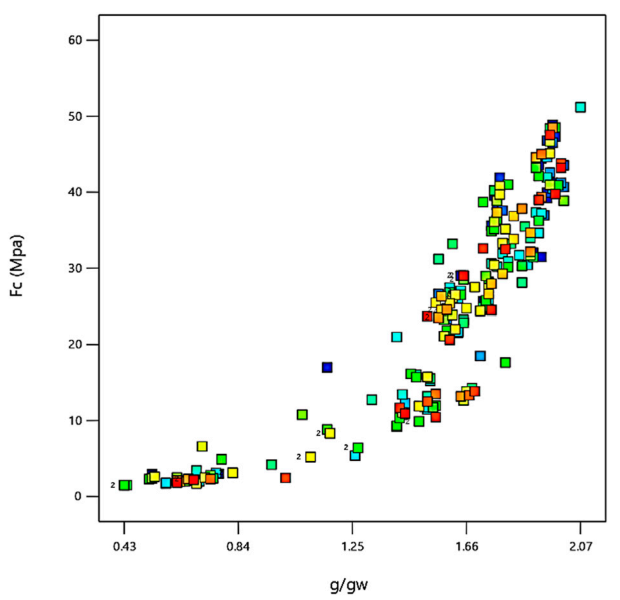

FIGURE 6
Fc versus g/gw relationship scatter plot.

different concrete behaviors have been reported in the literature, which had applied different machine learning techniques such as the artificial neural network (ANN) and these methods tried to train the techniques with the improved and the adaptive particle swarm optimization (IPSO and APSO) (Cavaleri et al., 2022; Armaghani et al., 2021; Kardani et al., 2022). Also, the Linear and Non-Linear Multivariate Adaptive Regression Splines (MARS-L and MARS-C), Gaussian Process Regression (GPR), and Minimax Probability Machine Regression (MPMR) were applied in the concrete properties model protocol as reported in the literature (Asteris Panagiotis G. et al., 2021). Results showed that the application of the metaheuristic algorithms in the training of the basic machine learning techniques increase speed and performance. However, these studies did not consider the impact of the dimensionless quantities such as concrete density to water density ratio, water-cement ratio, and sand-cement ratio on the foamed concrete strength behavior. Hence, the focus of this research project is to predict the compressive strength (Fc) of foamed concrete considering only three (3) input parameters derived from a dimensionless operation, which gave rise to density ratio, water-cement ratio and sand-cement ratio. This operation reduced the data handling complexities in this project. The research flowchart of this project is illustrated in Figure 1.

4 Methods

4.1 Collection of database and statistical analysis

A globally 230 records were collected from experimentally tested samples of foam concrete with mixes deposited in the literature (Salami et al., 2022). Each record contains the following data: concrete density to water density (concrete density g/cm^3) (γ/γw), water-cement ratio (W/C), sand-cement ratio (S/C), and compressive strength (MPa) (Fc). The collected records were divided into 75% training set (170 records) and 25% validation set (60 records). The appendix includes the complete dataset, while Tables 1, 2 summarize their statistical characteristics and the Pearson correlation matrix, respectively. Overfitting was overcome due to the data size, cross-validation, and data shuffling (Asteris Panagiotis G. et al., 2021). Finally, Figure 2 shows the histograms for both inputs and outputs and Figure 3 shows the relations between the inputs and the outputs. It can be shown that γ/γw produced the most reliable consistency with the Fc of the foamed concrete making that ratio the most important in the production of foamed concrete for a sustainable concrete structure construction. The skewed distribution in Figure 2 except for the S/C ration histogram, shows the inconsistency in the entries of the concrete component ratios, which requires a machine learning handling with superior workstation. In Figure 3, the parametric relations between the output and the studied ratios have been

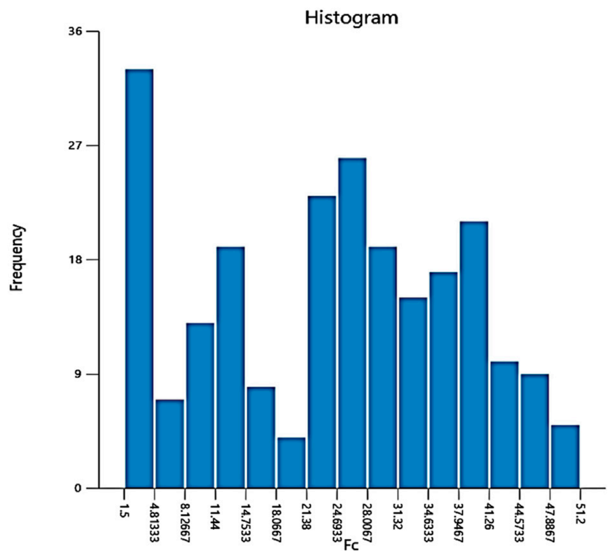

FIGURE 7
Histogram chart of the Fc entries.

presented depicting the point distributions, respectively. It can be shown that the unit weight ratio shows a more realistic parametric relationship with the compressive strength of the studied foamed concrete with a correlation of above 90%. This shows the importance of the ratio in the sustainable prediction of the strength of a lightweight foamed concrete, as agreeable with the previous study (Salami et al., 2022).

4.2 Research program

Response surface methodology (RSM) is a comprehensive set of statistical and mathematical approaches that include fitting a polynomial equation to trial data (Box et al., 2005; Kutner et al., 2005; Bezerra et al., 2008). The primary purpose of RSM is to accurately explain the performance of a given data set, with the ultimate goal of producing statistical predictions (Box et al., 2005). This approach is particularly applicable in situations when the outcome or outcomes of attention are influenced by multiple variables (Kutner et al., 2005). The aim is to concurrently improve the levels of these variables to achieve optimal system performance (Bezerra et al., 2008). Prior to implementing the RSM approach, it is imperative to carefully select an appropriate investigational design that will effectively delineate the specific tests to be conducted inside the designated investigational region under investigation (Bezerra et al., 2008). Several experimental matrices have been developed for this specific purpose. First-order models, such as factorial designs, are suitable experimental designs to employ in cases where the dataset lacks curvature. In order to model a response function for experimental data that cannot be well represented by linear functions, it is recommended to employ investigational designs that account for quadratic response surfaces (Box et al., 2005). Examples of such designs contain three-level factorial, central composite, Doehlert designs, and Box-Behnken (Box et al., 2005; Kutner et al., 2005; Bezerra et al., 2008). Statistically, RSM solves Equation 1:

$$\max f(x) \equiv E(Y(x)) \tag{1}$$

let Y be a random variable with an unknown mean function that depends on the d-dimensional factor vector x. Additionally, the variance of Y, which is caused by experimental error, is an unknown constant value. The development of response surface approach can be attributed to Box and his colleagues throughout the 1950s (Box et al., 2005; Kutner et al., 2005; Bezerra et al., 2008). The phrase in question has its origins in the graphical representation that arises from evaluating the fitness of a scientific model (Box et al., 2005). Its usage has been prevalent in the literature on chemometrics (Bezerra et al., 2008). The RSM methodology encompasses a collection of statistical and mathematical methodologies that rely on the fitting of experimental models to investigational data acquired through experimental design (Box et al., 2005). In pursuit of this goal, the utilization of linear or square polynomial functions is considered. Figure 4A illustrates the integration of these components in the context of RSM (Bezerra et al., 2008). The aforementioned confluence of techniques necessitates that researchers exercise caution and attentiveness throughout all three stages of Response Surface Methodology (RSM) (Box et al., 2005; Kutner et al., 2005). Without exercising due caution this practice is likely to encounter failure and may not yield the anticipated or intended outcomes.

Furthermore, three different symbolic regression techniques were used to predict the compressive strengths (Fc) of foam concrete using the collected database (Onyelowe et al., 2022a; Onyelowe et al., 2022b; Onyelowe Kennedy C. et al., 2022; Ebid and Deifalla, 2022). These techniques are “Genetic programming” (GP), three models of “Group method of data handling Neural Network” (GMDH-NN)and “Evolutionary Polynomial Regression” (EPR).

Genetic programming (GP) is a type of evolutionary algorithm and a machine learning technique inspired by biological evolution as shown the framework in Figure 4B. It belongs to the broader category of evolutionary algorithms, which are computational models inspired by the processes of natural selection and genetics. Inspiration from Evolution: Genetic programming draws inspiration from the process of natural evolution. It is based on the idea that a population of candidate computer programs, represented as structures such as syntax trees, can evolve over time to solve a

TABLE 5 Fc fit summary response for RSM linear and quadratic expressions.

Source	Sequential p-value	Lack of Fit p-value	Adjusted R^2	Predicted R^2	
Linear	<0.0001	<0.0001	0.8638	0.8620	
2FI	<0.0001	<0.0001	0.9171	0.9146	
Quadratic	<0.0001	<0.0001	0.9529	0.9494	Suggested

TABLE 6 Fc sequential model sum of squares [Type I] response.

Source	Sum of squares	df	Mean square	F-value	p-value	
Mean vs. Total	1.329E + 05	1	1.329E + 05			
Linear vs. Mean	38451.40	3	12817.13	483.19	<0.0001	
2FI vs. Linear	2383.76	3	794.59	49.21	<0.0001	
Quadratic vs. 2FI	1574.98	3	524.99	57.21	<0.0001	Suggested
Residual	2009.61	219	9.18			
Total	1.774E + 05	229	774.52			

TABLE 7 Model summary statistics.

Source	Std. Dev	R^2	Adjusted R^2	Predicted R^2	Press	
Linear	5.15	0.8656	0.8638	0.8620	6131.04	
2FI	4.02	0.9193	0.9171	0.9146	3793.32	
Quadratic	3.03	0.9548	0.9529	0.9494	2248.73	Suggested

TABLE 8 Fc Response ANOVA for Quadratic model.

Source	Sum of squares	df	Mean square	F-value	p-value	
Model	42410.15	9	4712.24	513.52	<0.0001	significant
A-γ/γ_w	50.89	1	50.89	5.55	0.0194	
B-W/C	131.34	1	131.34	14.31	0.0002	
C-S/C	22.32	1	22.32	2.43	0.1203	
AB	77.70	1	77.70	8.47	0.0040	
AC	100.51	1	100.51	10.95	0.0011	
BC	2.13	1	2.13	0.2326	0.6301	
A^2	1406.07	1	1406.07	153.23	<0.0001	
B^2	259.52	1	259.52	28.28	<0.0001	
C^2	0.0078	1	0.0078	0.0009	0.9767	
Residual	2009.61	219	9.18			
Lack of fit	1877.72	150	12.52	6.55	<0.0001	significant
Pure error	131.89	69	1.91			
Cor total	44419.76	228				

specific problem or perform a particular task. Representation of Solutions: In genetic programming, candidate solutions are typically represented as hierarchical structures, often in the form of syntax trees. These trees encode the structure and behavior of computer programs. Initialization: A population of random candidate solutions (trees) is generated to kickstart the evolutionary process. Evaluation: The fitness of each candidate solution is assessed based on its performance in solving the given problem. A fitness function quantifies how well a solution meets the specified criteria. Selection: Solutions are selected from the population based on their fitness, with a higher probability of selection for those with better fitness. This mimics the natural selection process. Crossover (Recombination): Pairs of selected solutions are combined to create new offspring through a process similar to genetic recombination. This involves swapping subtrees between parent solutions. Mutation: Random changes are introduced into the offspring solutions to mimic genetic mutations. These changes can include subtree replacement or modification of specific elements. Replacement: The new offspring solutions replace some of the existing solutions in the population, creating the next-generation. Termination: The evolution continues for a predefined number of generations or until a stopping criterion is met, such as finding a

TABLE 9 Fit statistics.

Std. dev	3.03	R^2	0.9548
Mean	24.09	Adjusted R^2	0.9529
C.V. %	12.57	Predicted R^2	0.9494
		Adeq Precision	92.7003

TABLE 10 Coefficients in terms of actual factors.

Factor	Coefficient estimate	df	Standard error	95% CI low	95% CI high	VIF
Intercept	26.57	1	7.03	12.71	40.42	
A-γ/γ_w	−14.45	1	6.14	−26.54	−2.36	165.34
B-W/C	−95.19	1	25.16	−144.78	−45.60	245.52
C-S/C	5.06	1	3.24	−1.33	11.45	122.93
AB	−24.29	1	8.35	−40.74	−7.84	48.38
AC	−4.45	1	1.34	−7.10	−1.80	56.25
BC	−1.74	1	3.61	−8.85	5.37	36.79
A^2	22.23	1	1.80	18.69	25.77	95.16
B^2	108.78	1	20.45	68.47	149.09	164.82
C^2	0.0098	1	0.3369	−0.6541	0.6737	9.01

solution that satisfies the desired criteria. Result: The final evolved solution, often represented by the best individual in the last generation, is considered the output of the genetic programming process. Genetic programming is used in various fields, including optimization problems, symbolic regression, automatic programming, and evolving control strategies. It is particularly well-suited for problems where the structure of the solution is not known in advance and can be effectively evolved. Overall, genetic programming provides a flexible and powerful approach for automatically discovering solutions to complex problems through the principles of evolution and natural selection.

Similarly, the Group Method of Data Handling (GMDH) is a type of neural network that falls under the category of machine learning models, with a typical architectural framework illustrated in Figure 4C. It is used for modeling and predicting complex relationships within data sets. Developed by the Soviet mathematician Alexey G. Ivakhnenko, GMDH is specifically designed for automatic feature selection and model optimization. Model Architecture: GMDH is a self-organizing, polynomial neural network. It automatically selects relevant input features and constructs a polynomial model based on the data. Self-Organization: GMDH organizes itself by iteratively selecting and combining input features. During the training process, the model identifies the most relevant features and their interactions. Layered Structure: GMDH typically has multiple layers, each representing a level of polynomial expansion. Each layer adds a new polynomial term to the model, and the network evolves to capture complex relationships in the data. Training Process: The training process involves finding the most relevant features and their combinations to create a polynomial model. GMDH uses a series of algorithms to select and optimize features, and it evaluates the quality of the model at each stage. Multiple Models: GMDH often generates multiple candidate models during the training process. These models compete, and the best-performing model is selected based on certain criteria, such as accuracy or generalization ability. Automatic Feature Selection: One of the strengths of GMDH is its ability to perform automatic feature selection. The model decides which input features are most relevant for accurate predictions. Polynomial Model: GMDH constructs a polynomial model based on the selected features. The polynomial terms represent the mathematical relationships between the input variables and the output. Prediction and Generalization: Once trained, the GMDH model can be used for making predictions on new, unseen data. The model aims to generalize well to capture the underlying patterns in the data. GMDH has been applied in various fields, including engineering, economics, and data analysis. Its ability to automatically select relevant features and create polynomial models makes it useful for tasks where the relationship between variables is not known beforehand. It's important to note that GMDH is just one approach among various neural network architectures, each with its own strengths and applications. And lastly, Evolutionary Polynomial Regression (EPR) is a computational intelligence technique used for modeling and predicting complex relationships within data. It combines the principles of evolutionary algorithms and polynomial regression to automatically evolve polynomial equations that best fit the given data set. The goal is to discover mathematical expressions that describe the underlying patterns in the data.

Evolutionary Algorithms: EPR employs evolutionary algorithms, such as genetic algorithms or genetic programming, to search through a space of potential polynomial equations. These algorithms mimic the process of natural selection, including selection, crossover (recombination), and mutation, to evolve a population of candidate

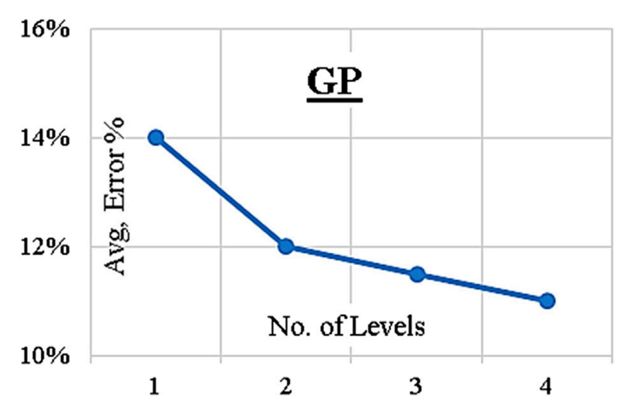

FIGURE 8
GP Model accuracy against complexity level.

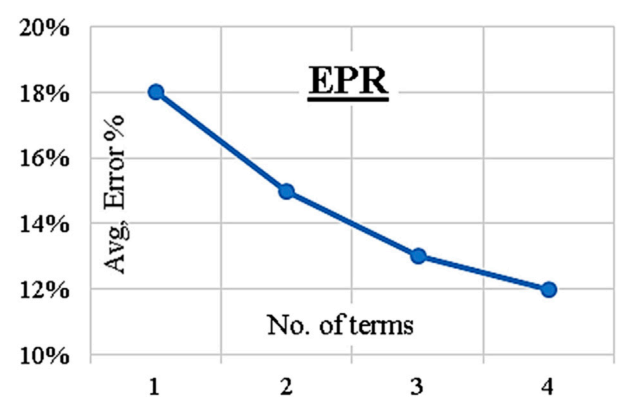

FIGURE 9
EPR model accuracy versus number of terms.

solutions. Polynomial Regression: Polynomial regression involves fitting a polynomial equation to the data, allowing for the modeling of non-linear relationships. The polynomial equation takes the form; $y = a_0 + a_1x + a_2x^2 + \ldots\ldots + a_nx^n$, where y is the output variable, x is input variable and a_0, a_1, a_2, and a_n are the coefficients.

Encoding Equations: Candidate polynomial equations are encoded in a way that allows the evolutionary algorithm to manipulate and evolve them. Various representations can be used, such as binary trees or strings, depending on the specific evolutionary algorithm employed. Fitness Evaluation: The fitness of each candidate solution (polynomial equation) is evaluated based on its ability to accurately predict the output variable for the given input data. A fitness function quantifies how well the equation fits the data, and it serves as a guide for the evolutionary algorithm in selecting better solutions. Evolutionary Process: The evolutionary algorithm iteratively evolves the population of polynomial equations over multiple generations. Selection mechanisms, crossover, and mutation operations are applied to create new generations of candidate solutions. Automatic Model Generation: EPR is capable of automatically generating mathematical models without prior knowledge of the underlying relationships in the data. It adapts the structure of polynomial equations based on the data patterns. Generalization: The evolved polynomial equations aim to generalize well to new, unseen data, capturing the essential features of the relationships within the dataset. Applications: EPR has been applied in various fields, including engineering, finance, biology, and data analysis, where complex relationships need to be uncovered from empirical data. EPR is particularly useful when dealing with problems where the relationship between input and output variables is not known or is highly non-linear. It offers a data-driven approach to model building and has the advantage of being capable of handling complex and non-linear relationships.

Flowcharts for the used techniques are presented in Figure 5. All the three developed models were used to predict (Fc) in (MPa) using(γ/γw, W/C, S/C). The following section discusses the results of each model. The accuracies of developed models were evaluated by comparing the sum of squared errors (SSE), mean absolute error (MAE), root mean squared error (RMSE) and R-squared (R2) between predicted and calculated compressive strength parameters values. These indices agree with the applied new methods in previous research projects (Apostolopoulou et al., 2020; Asteris P. G. et al., 2021). Also, this project employed the combined impact of data augmentation and hyperparameter tuning to overcome overfitting (Asteris et al., 2019; Armaghani and Asteris, 2021; Bardhan et al., 2024; Daniel et al., 2024).

4.3 Sensitivity analysis

A preliminary sensitivity analysis was carried out on the collected database to estimate the impact of each input on the (Fc) values. "Single variable per time" technique is used to determine the "Sensitivity Index" (SI) for each input using Hoffman and Gardener (Hoffman et al., 1983) formula as shown in Equation 2:

$$SI\,(X_n) = \frac{Y\,(X_{max}) - Y\,(X_{min})}{Y\,(X_{max})} \tag{2}$$

Accordingly, the (SI) values are (0.97, 0.91, 0.53) for (γ/γw, W/C, S/C) respectively. A sensitivity index of 1.0 indicates complete sensitivity, a sensitivity index less than 0.01 indicates that the model is insensitive to changes in the parameter.

5 Results and discussion

5.1 Response surface methodology (RSM) foamed concrete model

The maximum model order was set to quadratic for process factors. The selected model on the Model tab may be the design model or lower in order. The fit summary calculation was ended prematurely based on options set on the Transform tab. Select the highest order polynomial where the additional terms are significant and the model is not aliased. Focus on the model maximizing the Adjusted R^2 and the Predicted R^2. These are clearly presented in Tables 3, 4; Figures 6, 7, and Tables 5–7. The response surface interface has been randomized and run within 1.0 min under a quadratic design model as presented in Table 3. Table 3 only shows the settings and parameter indices with which the RSM was executed stating also the version of the software, while Table 4 presents the

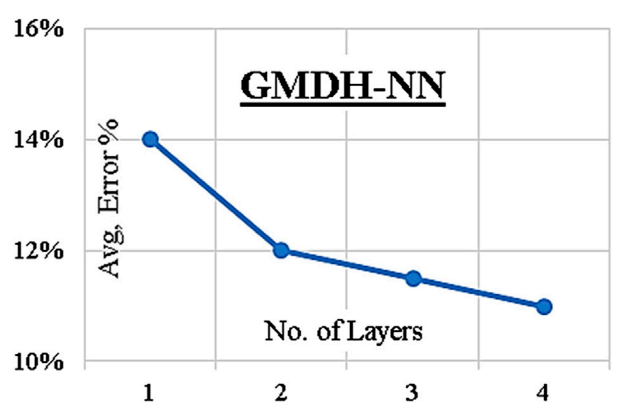

FIGURE 10
GMDH-NN Model accuracy Vs number of layer.

software coding statistics, which analyzed the three parameters. Table 5 further shows the RSM model results for the linear and quadratic modes displaying the R2, adjusted R2 and predicted R2. A continuous numeric model has also been applied on the parameters between a low and high coding interface.

It has been presented in Tables 8–10 that the RSM factor coding is actual, sum of squares is Type III – Partial and the model F-value of 513.52 implies that the model is significant. There is only a 0.01% chance that an F-value this large could occur due to noise. *P*-values less than 0.0500 indicate model terms are significant. In this case A, B, AB, AC, A^2, B^2 are significant model terms. Values greater than 0.1000 indicate the model terms are not significant. If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model. The lack of fit F-value of 6.55 implies the lack of fit is significant. There is only a 0.01% chance that a Lack of Fit F-value this large could occur due to noise. Significant lack of fit is bad -- we want the model to fit. The predicted R^2 of 0.9494 is in reasonable agreement with the Adjusted R^2 of 0.9529; i.e., the difference is less than 0.2. Adeq precision measures the signal to noise ratio. A ratio greater than 4 is desirable. Your ratio of 92.700 indicates an adequate signal. This model can be used to navigate the design space. The coefficient estimate represents the expected change in response per unit change in factor value when all remaining factors are held constant. The intercept in an orthogonal design is the overall average response of all the runs. The coefficients are adjustments around that average based on the factor settings. When the factors are orthogonal the VIFs are 1; VIFs greater than 1 indicate multi-colinearity, the higher the VIF the more severe the correlation of factors. As a rough rule, VIFs less than 10 are tolerable. The final equation in terms of actual factors is presented in Equation 3. The equation in terms of actual factors can be used to make predictions about the response for given levels of each factor. Here, the levels should be specified in the original units for each factor. This equation should not be used to determine the relative impact of each factor because the coefficients are scaled to accommodate the units of each factor and the intercept is not at the center of the design space. Figs in supplementary material present the RSM models' color points by value of Fc for normal percentage probability plot, color points by predicted value of Fc for extremely studentized residuals plot, color points by value of Fc for extremely studentized residuals versus run plot, Box-Cox plot of Fc for power transforms residuals, color points by actual and predicted value of Fc plot, color points by residual value of Fc versus g/gw plot, color points by value of Fc for leverage versus run plot and color points by value of Fc for degree of fitness (DFFITS) versus run plot (Humberg et al., 2019). These show the color shade behavior of the loading on the foamed concrete Fc, which agrees with previous RSM research, works (Humberg et al., 2019; Marković, 2006; Alyamac et al., 2017; Sambruno et al., 2019; V Bayramov et al., 2022; Bezerra et al., 2008). The RSM model constraints and coefficients are presented in tables in supplementary material and the desirability validation of the optimized foamed concrete Fc, the actual factor coding for the foamed concrete Fc and the 3D surface configuration of the foamed concrete Fc with respect to w/c and γ/γw are presented in Figs in supplementary material. It can be shown that the Fc of 53.6108 MPa has been optimized at the γ/γw of 2.04099, W/C of 0.270519, S/C of 1.08914, and a standard error of 0.739442 MPa.

$$\begin{aligned}Fc = &+22.23269_{\gamma/\gamma w^2} + 108.77855W/C^2 + 0.009830S/C^2\\ &- 24.28764_{\gamma/\gamma w} * W/C - 4.44771_{\gamma/\gamma w} * S/C\\ &- 1.73927W/C * S/C - 14.45057_{\gamma/\gamma w} - 95.19159W/C\\ &+ 5.05998S/C + 26.56823\end{aligned} \tag{3}$$

5.2 GP foamed concrete model

Four GP models were developed with complexity levels ranged between two and five. The population size, survivor size and number of generations were 1,000, 300 and 2,000 respectively. Figure 8 shows the improvement in accuracy with increasing the complexity. Equation 4 presented the output formula for (Fc) from the second trial. The average error (%) of total dataset is 12%, while the R^2 value is 0.965.

$$Fc = \left(\frac{\gamma}{\gamma_w}\right)\left(\frac{C}{W}e^{\left(\frac{\gamma}{\gamma_w}\right)} + \frac{\gamma}{\gamma_w}\right) - \left(\frac{S}{C}\right) \tag{4}$$

5.3 EPR foamed concrete model

Four developed EPR models were limited to sixth level polynomial, for three (30 inputs; there are 84 possible terms (28 + 21 + 15 + 10 + 6 + 3 + 1 = 5,005) as follows:

$$\begin{aligned}&\sum_{n=1}^{n=3}\sum_{m=1}^{m=3}\sum_{l=1}^{l=3}\sum_{k=1}^{k=3}\sum_{j=1}^{j=3}\sum_{i=1}^{i=3} X_n.X_m.X_l.X_k.X_j.X_i\\ &+\sum_{m=1}^{m=3}\sum_{l=1}^{l=3}\sum_{k=1}^{k=3}\sum_{j=1}^{j=3}\sum_{i=1}^{i=3} X_m.X_l.X_k.X_j.X_i\\ &+\sum_{l=1}^{l=3}\sum_{k=1}^{k=3}\sum_{j=1}^{j=3}\sum_{i=1}^{i=3} X_l.X_k.X_j.X_i + \sum_{k=1}^{k=3}\sum_{j=1}^{j=3}\sum_{i=1}^{i=3} X_k.X_j.X_i\\ &+\sum_{j=1}^{j=3}\sum_{i=1}^{i=3} X_j.X_i + \sum_{i=1}^{i=3} X_i + C\end{aligned} \tag{5}$$

GA technique was applied on these 84 terms to select the most effective terms to predict the values of (Fc). The process began

10.3389/fbuil.2024.1446597

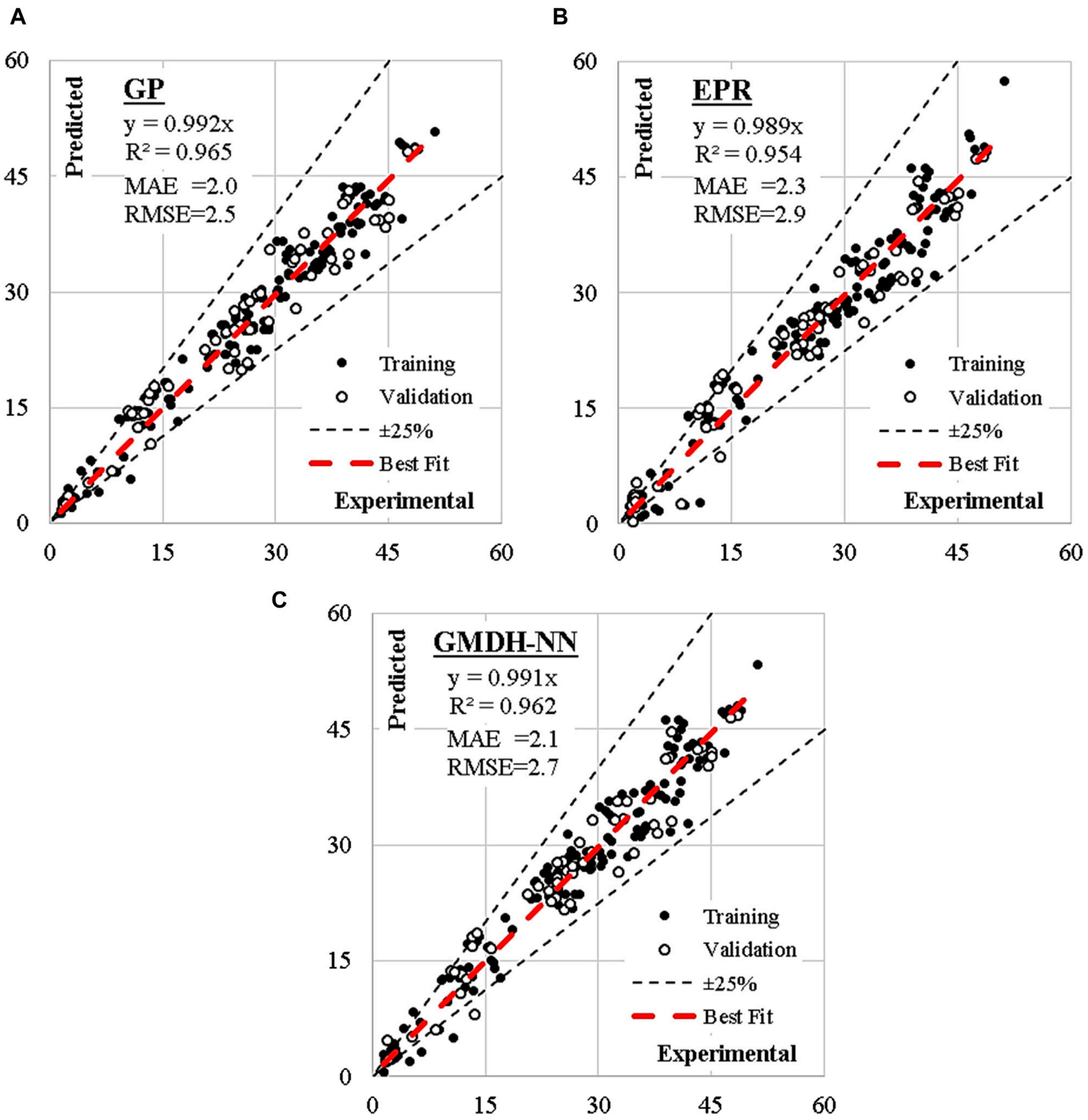

FIGURE 11
Relation between predicted and calculated (Fc) values using the developed models. **(A)** GP model, **(B)** EPR model, **(C)** GMDH-NN model.

with only 1 term and increased gradually up to 4 terms, Figure 9 presents the enhancement of fitness with increasing the number of terms and indicates that 5 is the optimum number of terms. The output of the second model is illustrated in Equation 6. The average error (%) and R^2 values were 15% and 0.954, respectively. The closed-form equation shows the importance of the density ratio ($\frac{\gamma}{\gamma_w}$) over other factor considerations in the optimization of the Fc of the studied foamed concrete.

$$Fc = \left(\frac{\gamma}{\gamma_w}\right)^2\left(\frac{4.34\,C}{W}\right) - \frac{S.C}{2\,W^2} - 1.88 \tag{6}$$

5.4 GMDH-NN foamed concrete model

Four GMDH-NN models based on the number of layers were developed to predict the Fc values using “GMDH Shell-3” software. The process began with only one layer and increased to four layers. The quadratic activation function was considered for the entire model. The error (%) values of the four models are illustrated in Figure 10. The average error (%) of total dataset is 12% and the R^2 value is 0.962. Also, Equations 7, 8 show the symbolic closed-form equations proposed by the GMDH-NN model technique, which further shows the importance and influence of the density ratio in

TABLE 11 Summary of the performance accuracies of the developed models.

Technique	Model	SSE	MAE	MSE	RMSE	Error	R2
		%	MPa	MPa	MPa	%	—
RSM	Equation 3	131.89	—	1.91	—	3.118	0.949
GP	Equation 4	1,479	2.0	6.4	2.5	12	0.965
EPR	Equation 6	1957	2.3	8.5	2.9	15	0.954
GMDH-NN	Equation 7	1,650	2.1	7.2	2.7	12	0.962

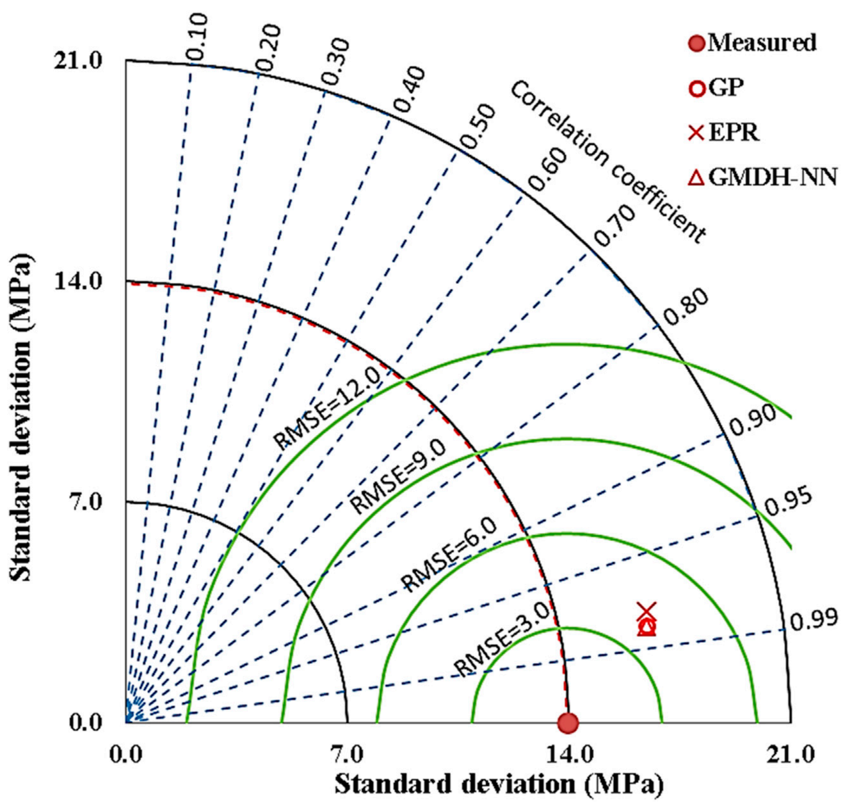

FIGURE 12
Comparison of the accuracies of the developed models using Taylor charts.

the design, production and use of foamed concrete in sustainable concrete structures construction. This exerted importance of the density of the concrete agrees with the RSM model (Humberg et al., 2019; Marković, 2006; Alyamac et al., 2017; Sambruno et al., 2019; V Bayramov et al., 2022; Bezerra et al., 2008).

$$Fc = \left(\frac{9.15\,W}{C}\right)^2 + \left(1.44 - \frac{1.17\,W}{C}\right)X_1 - \left(\frac{86\,W}{C}\right) + 20.1 \quad (7)$$

$$X_1 = 22.2\left(\frac{\gamma}{\gamma_w}\right)^2 - 26.9\left(\frac{\gamma}{\gamma_w}\right) - 1.1\left(\frac{S}{C}\right)^2 + 10.4 \quad (8)$$

Figure 11 shows the relationship between predicted and calculated foamed concrete $_{Fc}$ values using the developed models. It can be shown that the ±25% line of fit envelope has been applied to show the models accuracy and the outliers' consistency with the foamed concrete Fc. The GP model produced a parametric line of fit equation of 0.992x, R^2 of 0.965, MAE of 2.0 MPa, and RMSE of 2.5 MPa. Similarly, the EPR produced a parametric line of fit equation of 0.989x, R^2 of 0.954, MAE of 2.3 MPa, and RMSE of 2.9 MPa. Finally, the GMDH-NN produced a parametric line equation of 0.991x, R^2 of 0.962, MAE of 2.1 MPa, and RMSE of 2.7 MPa. The RSM produced an R2 of 0.949 with a standard error of computation of 0.739,442 MPa. These outcomes show that the GP outclassed the EPR, RSM and the GMDH-NN, though with minor margin (Shang et al., 2022). Meanwhile the EPR produced the highest outliers from the ±25% test of accuracy envelope. Table 11 presents the summary of the performance accuracies of the three AI-based symbolic model techniques. The comparisons of the accuracies of the developed models using Taylor charts and the variance distribution for the developed models have been presented in Figures 12, 13.

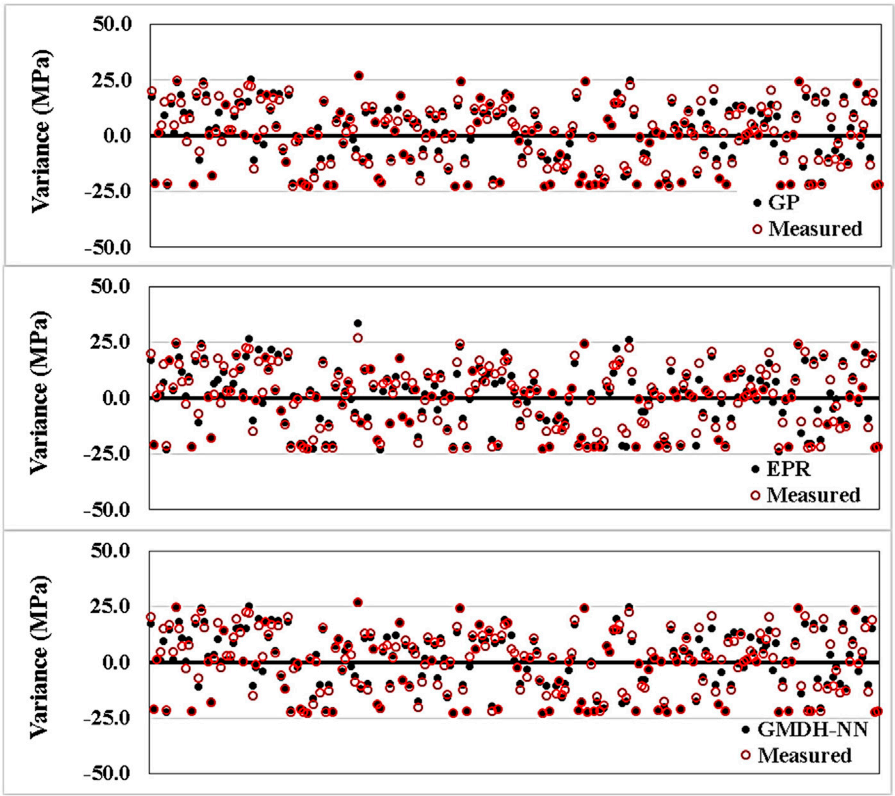

FIGURE 13
Variance distribution for the developed models.

6 Conclusion

This research aims to predict the compressive strength (Fc) of foamed concrete using γ/γw, W/C, and S/C ratios of the concrete component as independent variables. Three AI-based symbolic regression techniques were used, which were the response surface methodology (RSM), genetic programming (GP), evolutionary polynomial regression (EPR) and the group method of data handling-neural network (GMDH-NN). Also, the response surface methodology (RSM) technique was also applied as a symbolic interface producing field applicable expressions in graphs in this prediction work. The results of comparing the accuracies of the developed models could be concluded in the following points:

- The present research paper has reported the application of dimensionless parameters derived from the concrete components against what was used earlier in the previous work. This was to reduce data handling complexities and improve the performance.
- The sensitivity of the parameters showed that the density ratio is the most impactful parameter that influenced the behaviour of the foamed concrete strength and this corroborates with previous results deposited in the literature.
- GP, EPR and GMDH-NN models showed the same accuracy (85%–88%), while the RSM model showed higher accuracy (94.16%) based on applicable standard error of computation.
- Despite of the developed models using the four techniques showed better accuracies up to 89%, but the developed formulas are too complicated to be practical. That is why the selected simpler and less-accurate formulas are considered.
- Both correlation and sensitivity analysis showed that the density (γ/γw) has the main influence on the compressive strength, then the water-cement ratio (W/C) and finally the sand-cement ratio (S/C).
- The developed models are valid within the considered range of parameter values, and when it is beyond this range, the prediction accuracy should be re-verified.
- Generally, even though the RSM has its practical advantage, the GP produced the most decisive model and can be used side by side with the RSM for a more practical application.
- Overall, the data size is expected to be increased by a more extensive search to collect more recent entries on the production of foamed concrete for sustainable concrete structures. Also, more machine learning techniques should be deployed to study this database especially the metaheuristic methods.

7 Limitations and future work

The research work "estimating the compressive strength of lightweight foamed concrete using different machine learning-based symbolic regression techniques" is an interesting endeavor aimed at predicting the compressive strength of lightweight foamed concrete through various machine learning-based symbolic regression approaches. A knowledge of some potential limitations and avenues for future work will be vital to potential researchers.

7.1 Limitations

The effectiveness of machine learning models heavily relies on the quality and quantity of the data available. Limited or biased data could lead to models that are not sufficiently robust or generalizable. The success of symbolic regression models depends on the selection of appropriate features. If important features related to the compressive strength of lightweight foamed concrete are not included or poorly chosen, it can lead to inaccurate predictions. While symbolic regression techniques can provide mathematical expressions to represent relationships between input features and output, the interpretability of these models might be limited. Understanding why certain features contribute more to the prediction than others could be challenging. Models developed in this project may perform well on the specific dataset used for training but might struggle to generalize to unseen data or different contexts. Ensuring the generalizability of the models is crucial for practical applications. Some symbolic regression techniques might be computationally expensive, especially with large datasets or complex models. This could limit their practicality in real-world applications or require significant computational resources.

7.2 Future work

Increasing the size and diversity of the dataset through techniques like data augmentation could improve the robustness and generalizability of the models. Exploring additional features or transforming existing ones could enhance the performance of the models. Domain knowledge could be leveraged to engineer features that better capture the underlying relationships in the data. Conducting rigorous evaluation of the models on independent datasets and comparing their performance with other machine learning techniques can provide insights into their effectiveness and limitations. Developing techniques to improve the interpretability of symbolic regression models can facilitate better understanding and trust in the predictions, making them more useful in practical applications. Deploying the models in real-world scenarios and validating their performance in different contexts can demonstrate their utility and identify areas for further improvement. Exploring ensemble methods that combine multiple symbolic regression models or different types of machine learning techniques could potentially improve predictive performance and robustness. Investigating and developing more efficient optimization algorithms tailored to symbolic regression could reduce computational complexity and enable the scalability of the models to larger datasets.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

KO: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. AE: Investigation, Methodology, Project administration, Software, Visualization, Writing–original draft. DF: Data curation, Investigation, Project administration, Resources, Software, Writing–review and editing. NE: Data curation, Formal Analysis, Investigation, Project administration, Resources, Writing–original draft. NV: Data curation, Investigation, Methodology, Project administration, Resources, Visualization, Writing–review and editing. JB: Investigation, Methodology, Project administration, Resources, Writing–original draft, Writing–review and editing. SM: Investigation, Methodology, Resources, Writing–review and editing. HI: Investigation, Methodology, Project administration, Resources, Writing–review and editing. SH: Investigation, Methodology, Project administration, Writing–original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2024.1446597/full#supplementary-material

References

Alkayem, N. F., Shen, L., Mayya, A., Asteris, P. G., Fu, R., Luzio, G. D., et al. (2024). Prediction of concrete and FRC properties at high temperature using machine and deep learning: a review of recent advances and future perspectives. *J. Build. Eng.* 83, 108369. doi:10.1016/j.jobe.2023.108369

Alshaeer, H. A. Y., Irwan, J. M., Alshalif, A. F., Noman, E. A., Amran, M., Gamil, Y., et al. (2023). Optimisation of compressive strength of foamed concrete with a novel Aspergillus iizukae EAN605 fungus. *Case Stud. Constr. Mater.* 19, e02400. doi:10.1016/j.cscm.2023.e02400

Alyamac, K. E., Ghafari, E., and Ince, R. (2017). Development of eco-efficient self-compacting concrete with waste marble powder using the response surface method. *J. Clean. Prod.* 144, 192–202. doi:10.1016/j.jclepro.2016.12.156

Apostolopoulou, M., Asteris, P. G., Armaghani, D. J., Douvika, M. G., Lourenço, P. B., Cavaleri, L., et al. (2020). Mapping and holistic design of natural hydraulic lime mortars. *Cem. Concr. Res.* 136, 106167. doi:10.1016/j.cemconres.2020.106167

Armaghani, D. J., and Asteris, P. G. (2021). A comparative study of ANN and ANFIS models for the prediction of cement-based mortar materials compressive strength. *Neural Comput. Appl.* 33 (9), 4501–4532. doi:10.1007/s00521-020-05244-4

Armaghani, D. J., Mamou, A., Maraveas, C., Roussis, P. C., Siorikis, V. G., Skentou, A. D., et al. (2021). Predicting the unconfined compressive strength of granite using only two non-destructive test indexes. *Geomechanics Eng.* 25 (Number 4), 317–330. doi:10.12989/gae.2021.25.4.317

Asteris, P. G., Armaghani, D. J., Hatzigeorgiou, G. D., Karayannis, C. G., and Pilakoutas, K. (2019). Predicting the shear strength of reinforced concrete beams using artificial neural networks. *Comput. Concr.* 24 (5), 469–488. doi:10.12989/cac.2019.24.5.469

Asteris, P. G., Koopialipoor, M., Armaghani, D. J., Kotsonis, E. A., and Lourenço, P. B. (2021a). Prediction of cement-based mortars compressive strength using machine learning techniques. *Neural Comput. and Applic* 33, 13089–13121. doi:10.1007/s00521-021-06004-8

Asteris, P. G., Skentou, A. D., Bardhan, A., Samui, P., and Pilakoutas, K. (2021b). Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. *Cem. Concr. Res.* 145, 106449. doi:10.1016/j.cemconres.2021.106449

Ayyanar, D., Puvaneshwaran, J. N., Vipurajan, E., Shahul Hameed, M., and Jeyaseela, J. (2023). Experimental investigation on strength properties of foam concrete. *East Asian J. Multidiscip. Res.* 2 (4), 1517–1528. doi:10.55927/eajmr.v2i4.3861

Bardhan, A., Singh, R. K., Alatiyyah, M., and Alateyah, S. A. (2024). An effective hybrid model of ELM and enhanced GWO for estimating compressive strength of metakaolin-contained cemented materials. *Comput. Model. Eng. Sci.* 139 (2), 1521–1555. doi:10.32604/cmes.2023.044467

Bezerra, M. A., Santelli, R. E., Oliveira, E. P., Villar, L. S., and Escaleira, L. A. (2008). Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta* 76, 965–977. doi:10.1016/j.talanta.2008.05.019

Bian, P., Zhang, M., Yu, Q., Zhan, B., Gao, P., Guo, B., et al. Prediction model of compressive strength of foamed concrete considering pore size distribution. Available at: https://ssrn.com/abstract=4525291.

Box, G. E. P., Hunter, J. S., and Hunter, W. G. (2005) "Statistics for experimenters: design, discovery and innovation,". Wiley-Interscience.

Cavaleri, L., Barkhordari, M. S., Repapis, C. C., Armaghani, D. J., Ulrikh, D. V., and Asteris, P. G. (2022). Convolution-based ensemble learning algorithms to estimate the bond strength of the corroded reinforced concrete. *Constr. Build. Mater.* 359, 129504. doi:10.1016/j.conbuildmat.2022.129504

Daniel, C., Khatti, J., and Grover, K. S. (2024). Assessment of compressive strength of high-performance concrete using soft computing approaches. *Comput. Concr.* 33 (Number 1), 55–75. doi:10.12989/cac.2024.33.1.055

Deifalla, A., Awad, A., Seleem, H., and Abdelrahman, A. (2020). Investigating the behavior of lightweight foamed concrete T-beams under torsion, shear, and flexure. *Engineering Structures* 219, 110741.

Ebid, A., and Deifalla, A. (2022). Using artificial intelligence techniques to predict punching shear capacity of lightweight concrete slabs. *Materials.* 15 (8), 2732. doi:10.3390/ma15082732

Emad, W., Mohammed, A. S., Bras, A., Asteris, P. G., Kurda, R., Muhammed, Z., et al. (2022). Metamodel techniques to estimate the compressive strength of UHPFRC using various mix proportions and a high range of curing temperatures. *Constr. Build. Mater.* 349, 128737. doi:10.1016/j.conbuildmat.2022.128737

Hoffman, F. O., and Gardner, R. H. (1983). "Evaluation of uncertainties in environmental radiological assessment models," in *Radiological assessments: a textbook on environmental dose assessment*. Editors J. E. Till and H. R. Meyer (Washington, DC: U.S. Nuclear Regulatory Commission). Report No. NUREG/CR-3332.

Humberg, S., Nestler, S., and Back, M. D. (2019). Response surface analysis in personality and social psychology: checklist and clarifications for the case of congruence hypotheses. . *Sci.* 10, 409–419. doi:10.1177/1948550618757600

Kardani, N., Bardhan, A., Samui, P., Nazem, M., Asteris, P. G., and Zhou, A. (2022). Predicting the thermal conductivity of soils using integrated approach of ANN and PSO with adaptive and time-varying acceleration coefficients. *Int. J. Therm. Sci.* 173, 107427. doi:10.1016/j.ijthermalsci.2021.107427

Kozłowski, M., and Kadela, M. (2018). Mechanical characterization of lightweight foamed concrete. *Adv. Mater. Sci. Eng.* 2018, 8. doi:10.1155/2018/6801258

Kumar, A., Arora, H. C., Kapoor, N. R., Kontoni, D. P. N., Kumar, K., Jahangir, H., et al. (2023). Practical applicable model for estimating the carbonation depth in fly-ash based concrete structures by utilizing adaptive neuro-fuzzy inference system. *Comput. Concr.* 32 (2), 119–138. doi:10.12989/cac.2023.32.2.119

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005) "Applied linear statistical models,", 5. New York: McGraw-Hill Irwin.

Marković, I. (2006). *High-performance hybrid-fibre concrete: development and utilisation*. Amsterdam, Netherlands: IOS Press.

Nguyen, N.-H., Vo, T. P., Lee, S., and Asteris, P. G. (2021). Heuristic algorithm-based semi-empirical formulas for estimating the compressive strength of the normal and high-performance concrete. *Constr. Build. Mater.* 304, 124467. doi:10.1016/j.conbuildmat.2021.124467

Onyelowe, K. C., Ebid, A. M., Ariel, R., Atefeh, S., Haci, B., Kontoni Denise-Penelope, N., et al. (2022c). Global warming potential-based life cycle assessment and optimization of the compressive strength of fly ash-silica fume concrete; environmental impact consideration. *Front. Built Environ.* 8. doi:10.3389/fbuil.2022.992552

Onyelowe, K. C., Ebid, A. M., Riofrio, A., Baykara, H., Soleymani, A., Mahdi, H. A., et al. (2022b). Multi-objective prediction of the mechanical properties and environmental impact appraisals of self-healing concrete for sustainable structures. *Sustainability.* 14 (15), 9573. doi:10.3390/su14159573

Onyelowe, K. C., Kontoni, D.-P. N., Ebid, A. M., Dabbaghi, F., Soleymani, A., Jahangir, H., et al. (2022a). Multi-objective optimization of sustainable concrete containing fly ash based on environmental and mechanical considerations. *Buildings.* 12 (7), 948. doi:10.3390/buildings12070948

Othman, R., Jaya, R. P., Muthusamy, K., Sulaiman, M., Duraisamy, Y., Abdullah, MMAB, et al. (2021). Relation between density and compressive strength of foamed concrete. *Materials* 14 (11), 2967. doi:10.3390/ma14112967

Pan, Z., Hiromi, F., and Wee, T. (2007) Preparation of high performance foamed concrete from cement, sand and mineral admixtures, *J. Wuhan. Univ. Technol.* 22, 295–298. doi:10.1007/s11595-005-2295-4

Salami, B. A., Iqbal, M., Abdulraheem, A., Jalal, F. E., Alimi, W., Jamal, A., et al. (2022). Estimating compressive strength of lightweight foamed concrete using neural, genetic and ensemble machine learning approaches. *Cem. Concr. Compos.* 133, 104721. doi:10.1016/j.cemconcomp.2022.104721

Sambruno, A., Bañon, F., Salguero, J., Simonet, B., and Batista, M. (2019). Kerf Taper defect minimization based on abrasive waterjet machining of low thickness thermoplastic carbon fiber composites C/TPU. *Mater. (Basel)* 12, 4192. doi:10.3390/ma12244192

Shang, X., Qu, Na, and Li, J. (2022). Development and functional characteristics of novel foam concrete. *Constr. Build. Mater.* 324, 126666. doi:10.1016/j.conbuildmat.2022.126666

Shawnim, P. A., and Mohammad, F. (2019). Compressive strength of foamed concrete in relation to porosity using SEM images. *J. Civ. Eng. Sci. Technol.* 10 (1), 34–44. ISSN 2462-1382. doi:10.33736/jcest.1005.2019

V Bayramov, A., V Ermakova, G., V Kucheryavyy, A., V Meintser, I., and Zaraisky, A. G. (2022). Lamprey as laboratory model for study of molecular bases of ontogenesis and evolutionary history of vertebrata. *J. Ichthyol.* 62, 1213–1229. doi:10.1134/s0032945222060029

Wu, Z., Chen, B., and Liu, N. (2013). Fabrication and compressive properties of expanded polystyrene foamed concrete: experimental research and modeling. *J. Shanghai Jiaot. Univ. (Sci.)* 18, 61–69. doi:10.1007/s12204-013-1369-2

Yang, X., Xu, S., Zhao, Z., and Lv, Y. (2022). Strength, durability, and microstructure of foamed concrete prepared using special soil and slag. *Sustainability* 14, 14952. doi:10.3390/su142214952

Zhang, C., Zhu, Z., Shi, L., Kang, X., Wan, Y., Huo, W., et al. (2024). Compressive strength and sensitivity analysis of fly ash composite foam concrete: efficient machine learning approach. *Adv. Eng. Softw.* 192, 103634. doi:10.1016/j.advengsoft.2024.103634