frontiers | Frontiers in Materials

TYPE Original Research
PUBLISHED 22 October 2024
DOI 10.3389/fmats.2024.1481871

Check for updates

OPEN ACCESS

EDITED BY
Wenwu Xu,
San Diego State University, United States

REVIEWED BY
Mahmood Ahmad,
University of Engineering and Technology, Peshawar, Pakistan
Masoomeh Mirrashid,
Abu Dhabi University, United Arab Emirates
Mayadah W. Falah,
Al-Mustaqbal University College, Iraq

*CORRESPONDENCE
Marzhan Sadenova,
msadenova@edu.ektu.kz

RECEIVED 16 August 2024
ACCEPTED 30 September 2024
PUBLISHED 22 October 2024

CITATION
Rudenko O, Galkina D, Sadenova M, Beisekenov N, Kulisz M and Begentayev M (2024) Modelling the properties of aerated concrete on the basis of raw materials and ash-and-slag wastes using machine learning paradigm.
Front. Mater. 11:1481871.
doi: 10.3389/fmats.2024.1481871

COPYRIGHT
© 2024 Rudenko, Galkina, Sadenova, Beisekenov, Kulisz and Begentayev. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Modelling the properties of aerated concrete on the basis of raw materials and ash-and-slag wastes using machine learning paradigm

Olga Rudenko[1], Darya Galkina[2], Marzhan Sadenova[2]*, Nail Beisekenov[3], Monika Kulisz[4] and Meiram Begentayev[5]

[1]School of Architecture, Civil Engineering and Energy, D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan, [2]Center of Excellence "VERITAS", D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan, [3]Graduate School of Science and Technology, Niigata University, Niigata, Japan, [4]Faculty of Management, Department of Organisation of Enterprise, Lublin University of Technology, Lublin, Poland, [5]Satbayev University, Almaty, Kazakhstan

The thermal power industry, as a major consumer of hard coal, significantly contributes to harmful emissions, affecting both air quality and soil health during the operation and transportation of ash and slag waste. This study presents the modeling of aerated concrete using local raw materials and ash-and-slag waste in seismic areas through machine learning techniques. A comprehensive literature review and comparative analysis of normative documentation underscore the relevance and feasibility of employing non-autoclaved aerated concrete blocks in such regions. Machine learning methods are particularly effective for disjointed datasets, with neural networks demonstrating superior performance in modeling complex relationships for predicting concrete strength and density. The results reveal that neural networks, especially those with Bayesian Regularisation, consistently outperformed decision trees, achieving higher regression values ($R_{strength}$ = 0.9587 and $R_{density}$ = 0.91997) and lower error metrics (MSE, RMSE, RIE, MAE). This indicates their advanced capability to capture intricate non-linear patterns. The study concludes that artificial neural networks are a robust tool for predicting concrete properties, crucial for producing non-autoclaved curing wall blocks suitable for earthquake-resistant construction. Future research should focus on optimizing the balance between density and strength of blocks by enhancing the properties of aerated concrete and utilizing reliable models.

KEYWORDS

aerated concrete, seismic region, compressive strength, ash and slag waste, machine learning methods

1 Introduction

Advanced Concrete Technology encompasses the science of designing civil engineering projects that are efficient, economical, and safe. Engineers and concrete designers often encounter challenges due to slight changes in project conditions or objectives, complicating the selection of the optimal design. Foamed concrete,

TABLE 1 Cement test results.

Indicator name	Actual value of the indicator
Fineness of grinding by specific surface, cm3/g	2961
True density, kg/m3	3191
Setting time, min	Beginning –160
	Ending - 240
Uniformity of volume change, mm	7.2
Flexural strength, MPa	
2 days	6.1
28 days	9.9
Compressive strength, MPa	
2 days	26.4
28 days	64.5

TABLE 2 Results of physical and mechanical tests of ash and slag wastes.

Name of indicator	Actual value of the indicator
Uniformity of volume change, mm	4.5
Humidity, %	01,140
Bulk density (specific gravity), kg/m3	2112
True density, kg/m3	2530
Specific surface (cm2/g)	1.2
Total residues, % on sieves (mm)	
0.63	2.6
0.315	5.0
0.16	89.5
0.08	98.7
0.06	100
<0.06	

known for its versatility, can be utilized for various construction purposes and, when appropriately formulated, can serve as a structural element. The production of environmentally friendly lightweight foamed concrete (LWFC) necessitates a meticulous selection of the mix composition to achieve the desired physical and mechanical properties.

TABLE 3 Results of physical and mechanical tests of the sand.

Name of indicator	Actual value of the indicator
Bulk density, kg/m3	1557
True density, kg/m3	2874
Humidity, %	0
Volume of intergranular voids, % by volume	35
Total residues, in%, on sieves (mm)	
1.250.63	0.0
0.315	1.0
0.16	10.5
<0.16	91.0
Modulus of sand fineness, Mk	100.0
Sand group	1.025
Total residue on sieve with mesh No. 063, in % by weight	very fine, closer to thin
Content of grains with particle size over 5; 10 mm,% by weight	1.0
Content of grains finer than 0.16 mm, %	0
Content of dust and clay particles, %	9.0
Clay content in lumps, %	0

The most advanced studies include issues of wide variation of the concrete composition by including various wastes from fuel and energy and metallurgical industries, including by reducing the amount of cement or replacing it completely (Wu et al., 2024). Thus, the study (Akor et al., 2023) investigates the suitability of using crushed sandcrete block (CSB) as a partial replacement for fine aggregate in concrete, examining the physical and mechanical properties of the material and its impact on concrete's density and compressive strength at varying mix proportions. The authors (Ahmad and Shokouhian, 2024) study the use of recycled steel fibers as a reinforcing material in concrete.

The paper (Fang et al., 2024) investigated how industrial by-products, mineral admixtures, can improve the crack resistance of concrete mixed with ternary minerals (steel slag powder, mineral powder and fly ash). In this study, the mechanical properties and crack resistance characteristics of each group were preliminarily investigated using the plate method. Then, the G52 group and the control group were selected for comparative analysis based on the temperature stress testing machine (TSTM) and the analytical hierarchy process (AHP) method. The results show that the ternary mineral admixtures reduced the early strength and increased the

10.3389/fmats.2024.1481871

TABLE 4 Sample set composition of the experimental concrete mixtures.

Set no.	Cement (g)	ZShM (g)	Sand (g)	Soda ash (g)	Aluminum powder (g)	Water (g)	Water-solid ratio
1	380	210	60	2.5	0.25	330	0.58
2	370	200	55	2.3	0.23	320	0.56
3	390	220	65	2.6	0.26	340	0.59
4	385	215	62	2.4	0.24	335	0.57
5	375	205	58	2.2	0.22	325	0.55
6	380	210	60	2.5	0.25	330	0.58
7	370	200	55	2.3	0.23	320	0.56
8	390	220	65	2.6	0.26	340	0.59
9	385	215	62	2.4	0.24	335	0.57
10	375	205	58	2.2	0.22	325	0.55
11	380	210	60	2.5	0.25	330	0.58
12	370	200	55	2.3	0.23	320	0.56
13	390	220	65	2.6	0.26	340	0.59
14	385	215	62	2.4	0.24	335	0.57
15	375	205	58	2.2	0.22	325	0.55
16	380	210	60	2.5	0.25	330	0.58
17	370	200	55	2.3	0.23	320	0.56
18	390	220	65	2.6	0.26	340	0.59
19	385	215	62	2.4	0.24	335	0.57
20	375	205	58	2.2	0.22	325	0.55
21	380	210	60	2.5	0.25	330	0.58
22	370	200	55	2.3	0.23	320	0.56
23	390	220	65	2.6	0.26	340	0.59
24	385	215	62	2.4	0.24	335	0.57
25	375	205	58	2.2	0.22	325	0.55
26	380	210	60	2.5	0.25	330	0.58
27	370	200	55	2.3	0.23	320	0.56
28	390	220	65	2.6	0.26	340	0.59
29	385	215	62	2.4	0.24	335	0.57
30	375	205	58	2.2	0.22	325	0.55
31	380	210	60	2.5	0.25	330	0.58
32	370	200	55	2.3	0.23	320	0.56

(Continued on the following page)

TABLE 4 (*Continued*) Sample set composition of the experimental concrete mixtures.

Set no.	Cement (g)	ZShM (g)	Sand (g)	Soda ash (g)	Aluminum powder (g)	Water (g)	Water-solid ratio
33	390	220	65	2.6	0.26	340	0.59
34	385	215	62	2.4	0.24	335	0.57
35	375	205	58	2.2	0.22	325	0.55
36	380	210	60	2.5	0.25	330	0.58
37	370	200	55	2.3	0.23	320	0.56
38	390	220	65	2.6	0.26	340	0.59
39	385	215	62	2.4	0.24	335	0.57
40	375	205	58	2.2	0.22	325	0.55
41	380	210	60	2.5	0.25	330	0.58
42	370	200	55	2.3	0.23	320	0.56
43	390	220	65	2.6	0.26	340	0.59
44	385	215	62	2.4	0.24	335	0.57
45	375	205	58	2.2	0.22	325	0.55
46	380	210	60	2.5	0.25	330	0.58
47	370	200	55	2.3	0.23	320	0.56
48	390	220	65	2.6	0.26	340	0.59
49	385	215	62	2.4	0.24	335	0.57
50	375	205	58	2.2	0.22	325	0.55

28^{th} day strength of concrete. Meanwhile, they reduced the early adiabatic temperature rise, maximum compressive stress and creep effect on the crack resistance characteristics of concrete, thereby enhancing the tensile cracking stress of concrete. Overall, this study provided new information on the mixing proportions of ternary mineral concrete, but the interactions between these components were not fully understood.

Nafees et al. (Nafees et al., 2021) explored the use of silica fume (SF) in concrete, emphasizing its benefits in reducing CO_2 emissions and enhancing concrete durability and mechanical properties. They developed predictive machine learning (ML) models, specifically decision tree (DT) and support vector machine (SVM), to forecast the compressive strength of SF concrete, demonstrating high prediction accuracy and highlighting the effectiveness of these techniques in promoting sustainable concrete solutions.

Compared with traditional concrete materials, high-plasticity concrete, which is also called engineering cementitious composite, has superior properties such as exceptional plasticity, tensile strength with ultimate tensile strain exceeding 3% and excellent workability (Chen et al., 2024). To study the influence of axial compression ratio, torsion-bending ratio and eccentricity on the mechanical performance of reinforced concrete columns under combined loads, a numerical analysis based on ABAQUS was implemented, which is in good agreement with the experimental results (Huang et al., 2021).

In (Song et al., 2024), a data-driven modeling methodology is proposed that combines proven finite element (FE) simulations with machine learning (ML) algorithms to investigate the penetration resistance of RC slabs against projectiles. The methodology uses an extreme learning machine (ELM) network to identify extreme conditions that characterize the penetration resistance, reducing data requirements and improving the prediction performance. The predicted outputs from the ELM network are then used as input to a recurrent neural network (RNN) to predict the residual velocity curves over time, which is an essential metric for evaluating the penetration resistance. To cover different loading scenarios, 64 sets of FE models with different drop velocities, caliber-radius-head (CRH), and projectile masses are developed based on available experimental studies. The Johnson-Holmquist II model and Johnson-Cook model are used for concrete and steel, respectively, to describe the material behavior at high strain rates. The FE prediction results form the training set database, and ML models are fitted using CRH, impact velocity and projectile mass as input layers and residual projectile velocity as output layer. The

10.3389/fmats.2024.1481871

TABLE 5 Results of Kruskal–Wallis statistic for selected input variables.

Input variable	Density		Strength	
	Kruskal–wallis statistic	*p*-value	Kruskal–wallis statistic	*p*-value
Cement	1,76	0,624	7,14	0,068
ZSHM	15,45	0,0039	27,61	0,000,015
Sand	16,01	0,003	28,85	0,000,008
Soda ash	16,83	0,000,765	8,3	0,04
Aluminum Powder	11,51	0,021	12,03	0,017
Water, g	22,05	0,024	16,24	0,132
Water solid ratio	10,41	0,238	13,85	0,086

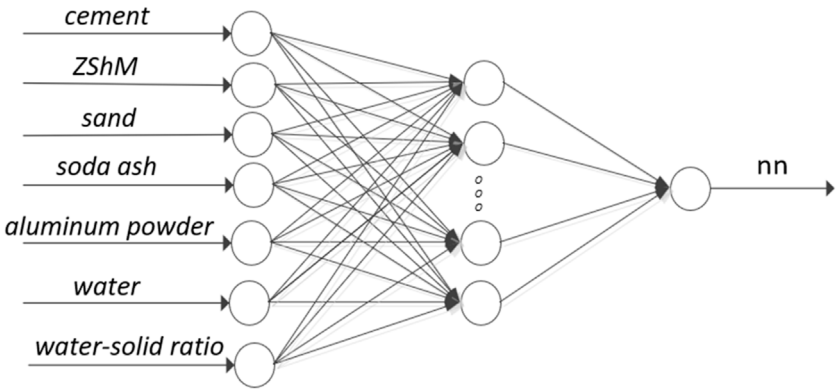


FIGURE 1
Schematic representation of the analyzed models, where nn–strength or density of concrete.

TABLE 6 Quality indicators for the evaluation of the received models.

Quality indicator	Formula	Meaning of symbols
Regression value (R)	$R(y', y^*) = \frac{cov(y', y^*)}{\sigma_{y'}\sigma_{y^*}}, R\epsilon < 0, 1 >$	$\sigma y'$—standard deviation of reference values of the analyzed indicator σy^*—standard deviation of predicted values the analyzed indicator, y_i is the actual value of the analyzed indicator, $\hat{y}_i$ denotes the value of the analyzed indicator for the i-th observation obtained from the model
Mean Squared Error (MSE)	$MSE = \frac{1}{n}\sum_{n=1}^{n}(\hat{y}_i - y_i)^2$	
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$	
Relative Importance of Errors (RIE)	$RIE = \frac{\|y' - y\|}{\|y\|}$	
Mean Absolute Error (MAE)	$MAE = = \frac{1}{N}\sum_{i=1}^{N}(\|y_i - \hat{y}_i\|)$	

results show high Pearson correlation coefficients of 0.984 and 0.991 for the ELM and RNN prediction results, respectively, with the numerical simulation results. The proposed networks provide accurate and efficient alternatives to time-consuming FE modeling for predicting projectile penetration performance. This strategy offers a new ML-based method for fast evaluation, especially for *in situ* applications.

Conducting experimental investigations on all possible compositions is both time-consuming and costly. Attempts have been made to develop unified strength criteria for evaluating

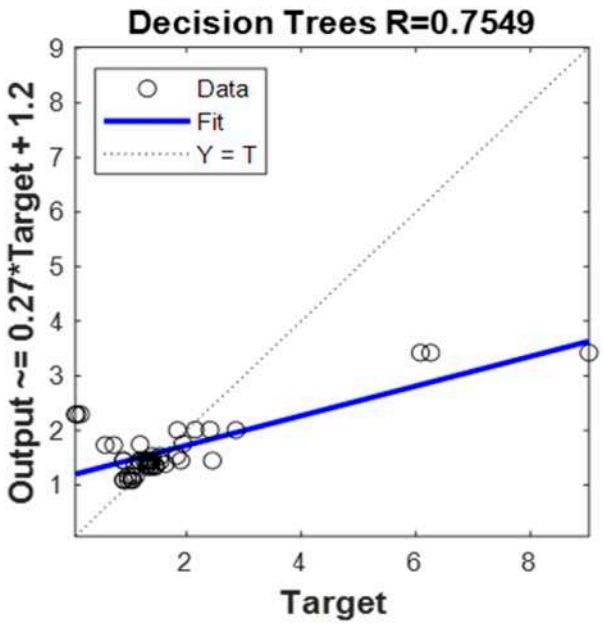


FIGURE 2
Regression statistics for the total set for Decision Trees model for concrete strength.

the performance of concrete mixtures. They are usually based on a mechanistic approach with adjustable failure planes and complex expressions. The authors of (Lu et al., 2017) presented an alternative mechanistic approach to develop a simple unified strength theory with adjustable characteristic stress based on the Drucker-Prager strength theory. The strength curves of this theory are a series of continuous and smooth conical loci, which are located between the Drucker-Prager and Matsuoka-Nakai strength curves in the deviatoric plane in the principal stress space. Another major advantage of the developed strength theory is that its three parameters ($\sigma 0$, φc, and φe) have clear physical meanings and can be determined based on simple laboratory tests. Verifications between the developed theory and experimental triaxial test data available in the literature show that this theory is capable of reasonably representing the three-dimensional (3D) strength properties of various geomaterials.

Thus, the use of Artificial Intelligence (AI) techniques is an effective approach to predict and optimize real-world problems in specific technologies. Currently, AI technologies are emerging as a strategic development vector for companies in the construction sector. The implementation of "smart solutions" across all stages of the life cycle of building materials, products, and structures is becoming increasingly prevalent. AI technology offers ultra-high-dimensional non-linear computational capabilities, intelligent functions for complex analysis and judgement, and the capacity for self-learning and knowledge accumulation.

The application of machine learning (ML) in concrete technology has seen substantial growth, with numerous studies highlighting its potential to enhance concrete mixture design and optimization. Researchers have demonstrated that ML algorithms can effectively predict key properties of concrete, such as compressive strength, workability, and durability, thereby simplifying the formulation process. In this paper (Zhou et al., 2019), a model capable of capturing the complex mechanical behavior of concrete is developed which is evaluated using monotonic and cyclic test data from the literature. Additionally, ML models have been utilized to optimize mix component proportions, improving both efficiency and sustainability in concrete production (Yi et al., 2023). The integration of ML in concrete research has also led to the development of intelligent systems capable of autonomous decision-making and real-time adjustments, showcasing the transformative impact of ML in this field (Li et al., 2022; Moradi et al., 2022; Yang et al., 2022; Haq et al., 2024). As the construction industry increasingly adopts digital technologies, the role of ML in advancing concrete technology continues to expand (Raju et al., 2023; Ullah et al., 2022; Akeke et al., 2023). This review aims to synthesize existing research on ML applications in concrete science, highlighting significant findings, methodological approaches, and future directions. Previous studies have laid a solid foundation for understanding and leveraging ML in concrete research (Qian et al., 2023; Chen et al., 2023; Anjum et al., 2022; Hu et al., 2021).

AI can unlock the potential of multidimensional nonlinear relationship between material components and performance indicators compared to the empirical formula derived from classical statistical approaches. The authors (Razveeva et al., 2023) show that artificial intelligence technology has obvious advantages in measurement accuracy in predicting specific performance indicators compared to traditional statistical methods. It is recommended to use multiple algorithms to cross-validate the model prediction results. Traditionally, support vector machines are used for small datasets. Decision tree evolution methods should be used in model algorithms that require feature optimisation or dispersed index prediction. Artificial neural networks should be used to solve such problems as improving the forecasting model, increasing the forecasting accuracy, increasing the sample data set, etc.

Among the variety of applications of AI methods, a special place is occupied by the development of theory and technology for creating artificial systems that process information from images obtained during construction monitoring of the structural condition of objects. The paper (Elhishi et al., 2023) deals with the process of developing an innovative method for analysing the presence of cracks that appeared after the application of load and delamination as a result of the technological process, with subsequent estimation of the length of cracks and delamination using convolutional neural networks (CNN) in the assessment of the condition of aerated concrete products. Four convolutional neural network models are applied in solving a problem in the field of construction. Defect detection using computer vision is demonstrated.

Elhishi et al. (Farooq et al., 2021) investigated the application of machine learning to predict the strength of high-performance concrete, essential for optimizing performance and safety in construction. Their study evaluated eight machine learning models, including regression methods and tree-based models, finding that ensemble techniques like XGBoost achieved the highest accuracy with an R-Square of 0.91. They also utilized the SHAP technique to provide valuable insights for concrete mix design and construction practices.

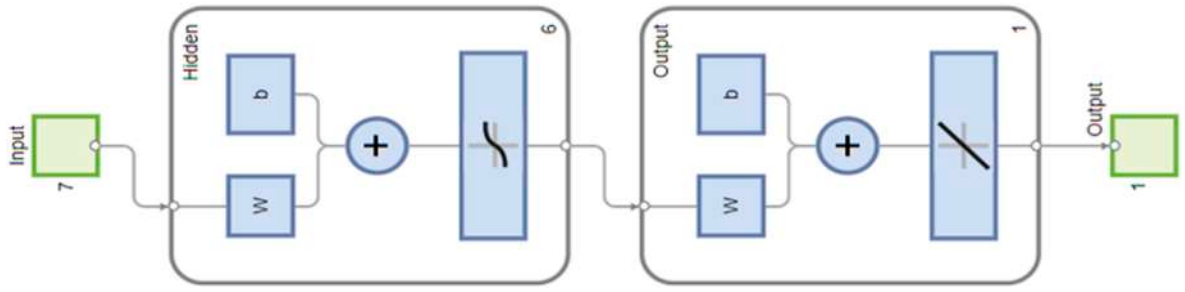


FIGURE 3
The structure of a neural network.

TABLE 7 Results of neural network training for concrete strength.

Network number	1	2	3
Training algorithm	Lavenberg-Marquardt	Scaled Conjugate Gradient	Bayesian Regularization
Epoch	29	23	587
Performance	0.0545	0.1837	0.0838
Best validation performance	0.07799 at epoch 23	0.94392 at epoch 17	0.083782 at apoch 22
Gradient	0.007	0.2744	0.0239

Collectively, these studies underscore the significant advancements and diverse applications of machine learning in concrete technology. The integration of ML not only enhances predictive accuracy for key concrete properties but also optimizes material usage, contributing to more sustainable and cost-effective construction practices (Zheng et al., 2023; Wang et al., 2024; Hameed et al., 2022; Alaneme et al., 2023). As the field progresses, ongoing research and development of innovative ML techniques will be crucial in further advancing concrete technology (Wang et al., 2022; Moein et al., 2023; Gamil, 2023). This review highlights the importance of adopting sophisticated AI models, such as neural networks and ensemble methods, to tackle complex challenges in concrete formulation and performance prediction. Future research should continue to explore and refine these methods, ensuring their effective implementation in real-world construction scenarios, particularly in enhancing the durability and sustainability of concrete in various environmental conditions.

The primary objective of this study is to develop and validate machine learning models for optimizing the properties of aerated concrete using local raw materials and ash-and-slag wastes, specifically for applications in seismic regions. This research aims to address the environmental concerns associated with thermal power plants by promoting the use of sustainable materials in concrete production. By leveraging advanced AI techniques, this study seeks to achieve the following goals:

1. Develop robust machine learning models, particularly neural networks, to accurately predict the strength and density of aerated concrete.
2. Identify the optimal mix proportions of local raw materials and ash-and-slag wastes to enhance the physical and mechanical properties of non-autoclaved aerated concrete.
3. Assess the suitability of the optimized aerated concrete for use in seismic regions, ensuring it meets the required standards for seismic resistance.
4. Demonstrate the environmental benefits of using ash-and-slag wastes in concrete production, contributing to the reduction of harmful emissions from thermal power plants.
5. Highlight the advantages of using AI for modeling complex relationships between material components and performance indicators, surpassing traditional empirical methods.

These objectives are designed to advance the understanding and application of machine learning in the field of concrete technology, ultimately contributing to more sustainable and resilient construction practices.

2 Materials and methods

2.1 Raw materials

An attempt was made to produce aerated concrete blocks with the required density and strength for use in seismically hazardous areas at minimal cost by selecting locally available raw materials. Portland cement of grade CEM I 52.5H from Bukhtarma Cement Company (East Kazakhstan), which complies with the

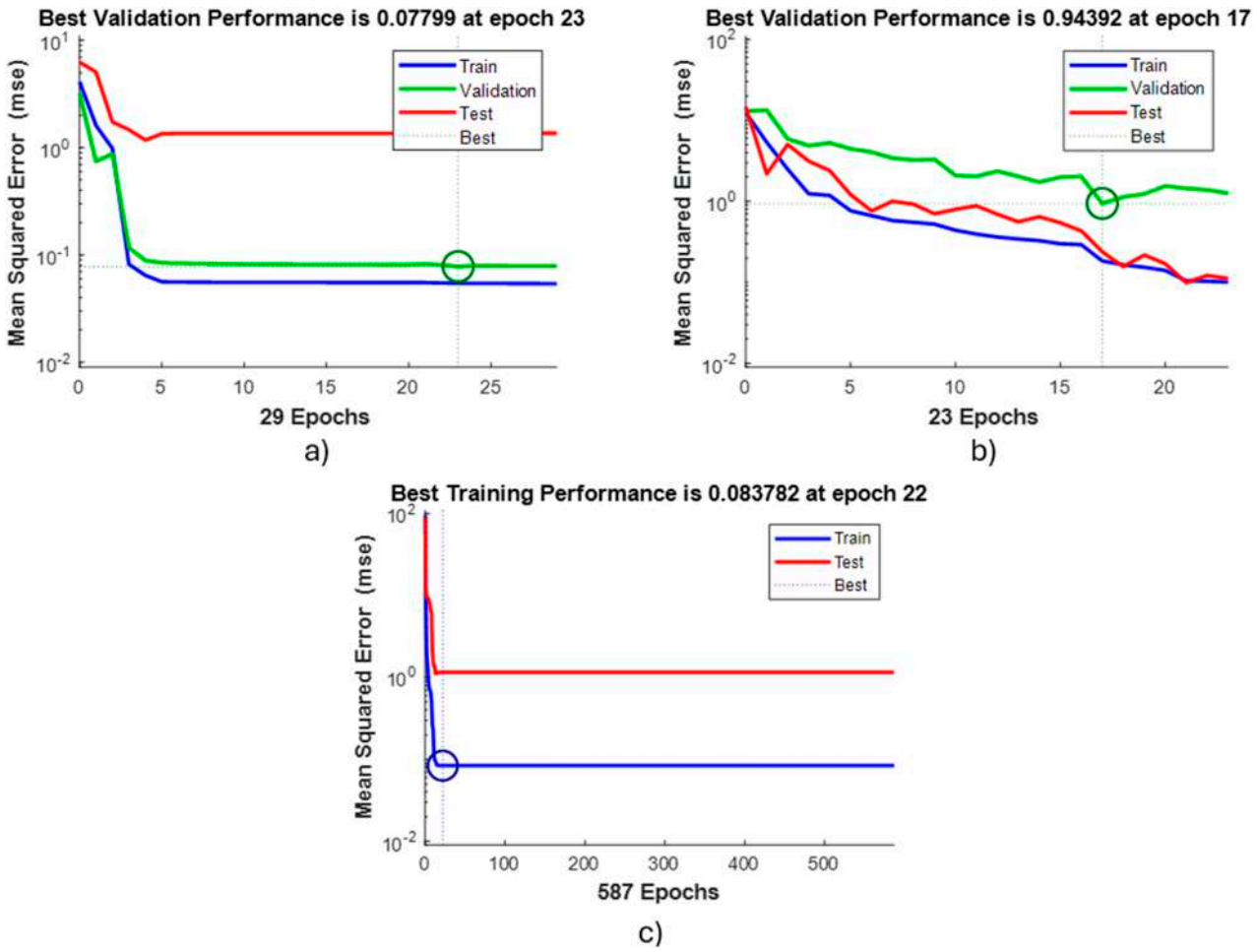


FIGURE 4
Network training process for concrete strength - best validation performance of predictive model: **(A)** L-M, **(B)** SCG and **(C)** BR.

TABLE 8 Results of neural network quality indicators.

Training algorithm	Lavenberg-marquardt	Scaled conjugate gradient	Bayesian regularization
R (all data)	0.9587	0.938	0.9587
MSE	0.2624	0.3105	0.2499
RMSE	0.5122	0.5572	0.4999
RIE	0.2208	0.2403	0.2155
MAE	0.2527	0.4034	0.2458

relevant standard (EN 197-1, 2014), was used as the binder. The chemical composition of the cement includes 19.95% silicon dioxide (SiO_2), 5.58% aluminum oxide (Al_2O_3), 4.98% ferric oxide (Fe_2O_3), 63.07% calcium oxide (CaO), 4.5% magnesium oxide (MgO), and 0.36% sulfur trioxide (SO_3). The results of the cement tests are presented in Table 1.

For studies involving animals or humans, or other research requiring ethical approval, the authority that granted approval and the corresponding ethical approval code must be listed.

Ash-and-slag waste from the coal-fired thermal power plant of Ust-Kamenogorsk (East Kazakhstan) was used as silica aggregate. Preliminary ash and slag waste was sieved on a sieve with a mesh size of 1.25 mm. The chemical composition of the ash and slag waste includes 35.3%–45.6% silicon dioxide (SiO_2) in quartz form, 53.9%–64.0% mullite ($3Al_2O_3{\cdot}2SiO_2$), 0.3%–1.8% hematite (Fe_2O_3), and 0.2%–1.1% magnetite (Fe_2O_3). The results of physical and mechanical tests of ash and slag waste are given in Table 2.

The question of partial replacement of ash-and-slag aggregate with natural local raw material - very fine sand of II class according to GOST 8736 (GOST 8736-2014, 2015), with SiO2 content more than 90% was investigated. The results of physical and mechanical tests of sand are given in Table 3.

As a pore-forming agent, in the work we used gas-forming aluminium powder of PAP-1 grades, corresponding to the normative requirements (GOST 5494-95, 2006). Calcined soda ash was also used as an additive.

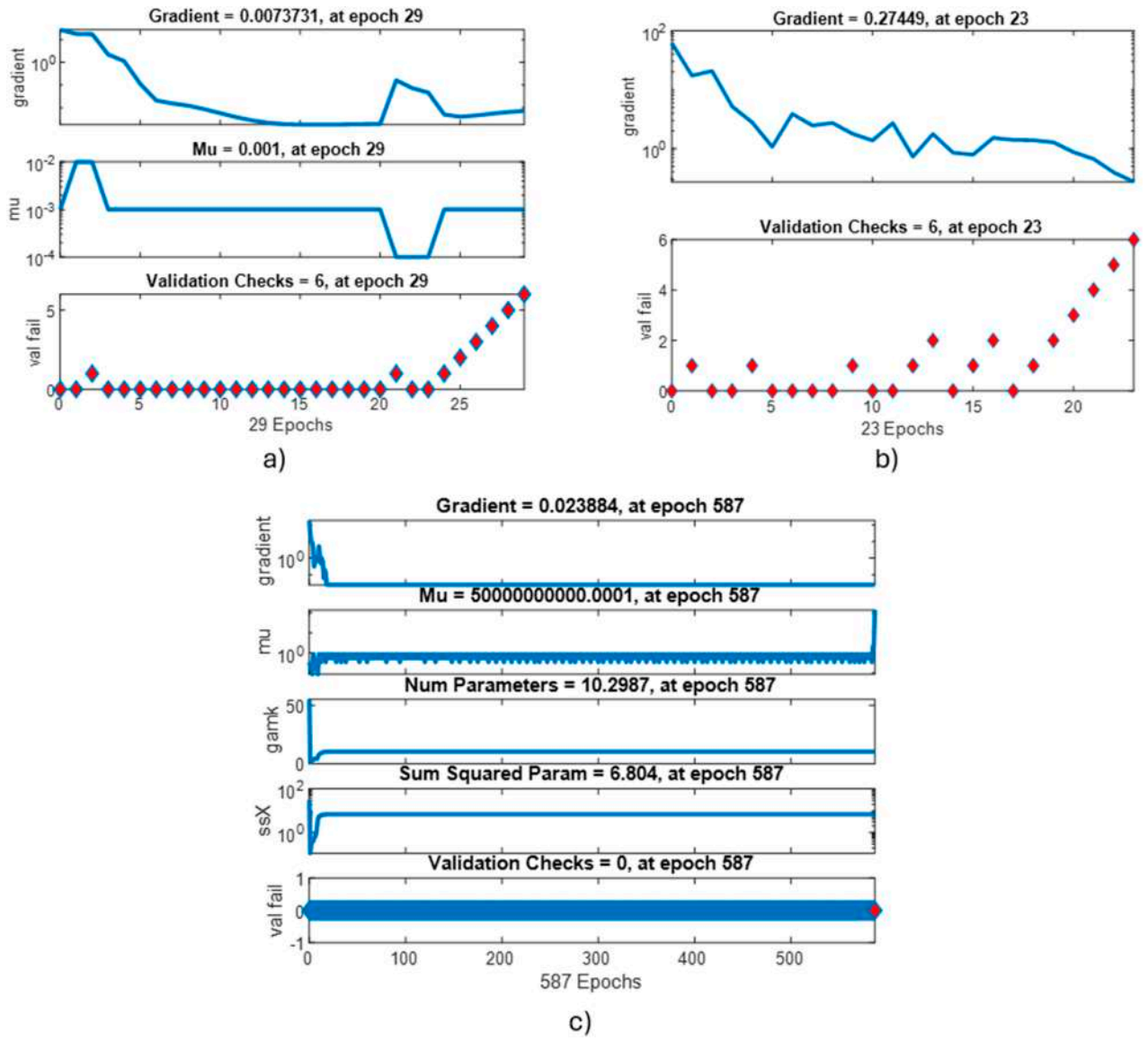


FIGURE 5
Network training process for concrete strength - validation checks: **(A)** L-M, **(B)** SCG and **(C)** BR.

To study the influence of the components on the strength and density of aerated concrete, a series of samples were made with different variations in the content of components. In a number of samples the cement content was changed (from −10 to +20%). Some samples were made with the same amount of cement, but different content of aluminium powder from 0.15 to 0.65 g). The ash and slag component was partially replaced by traditional sand (from 10% to 100%).

2.2 Study of the characteristics of the initial materials of experimental samples

Chemical and phase compositions of initial materials were determined using an Agilent ISP-MS inductively coupled plasma mass spectrometer (Agilent Technologies, USA). The specific surface area was determined on a PSX-10a device using the Kozeny and Karman gas permeability method according to the instructions for the device. Physical and mechanical tests were carried out according to standard techniques.

2.3 Manufacture of experimental samples

The experimental samples (Table 4) were produced in the laboratory of the Center of Competence and Technology Transfer in Construction of EKTU named after D.Serikbayev. The amount of dry components was selected so that the volume of finished aerated concrete was at least 0.01 m^3. The solid components were used in the dry state. Dosing of dry constituents was carried out by weight. First, cement and aggregate were mixed. Mixing was carried out using a construction drill-mixer. Soda ash was separately dissolved in a

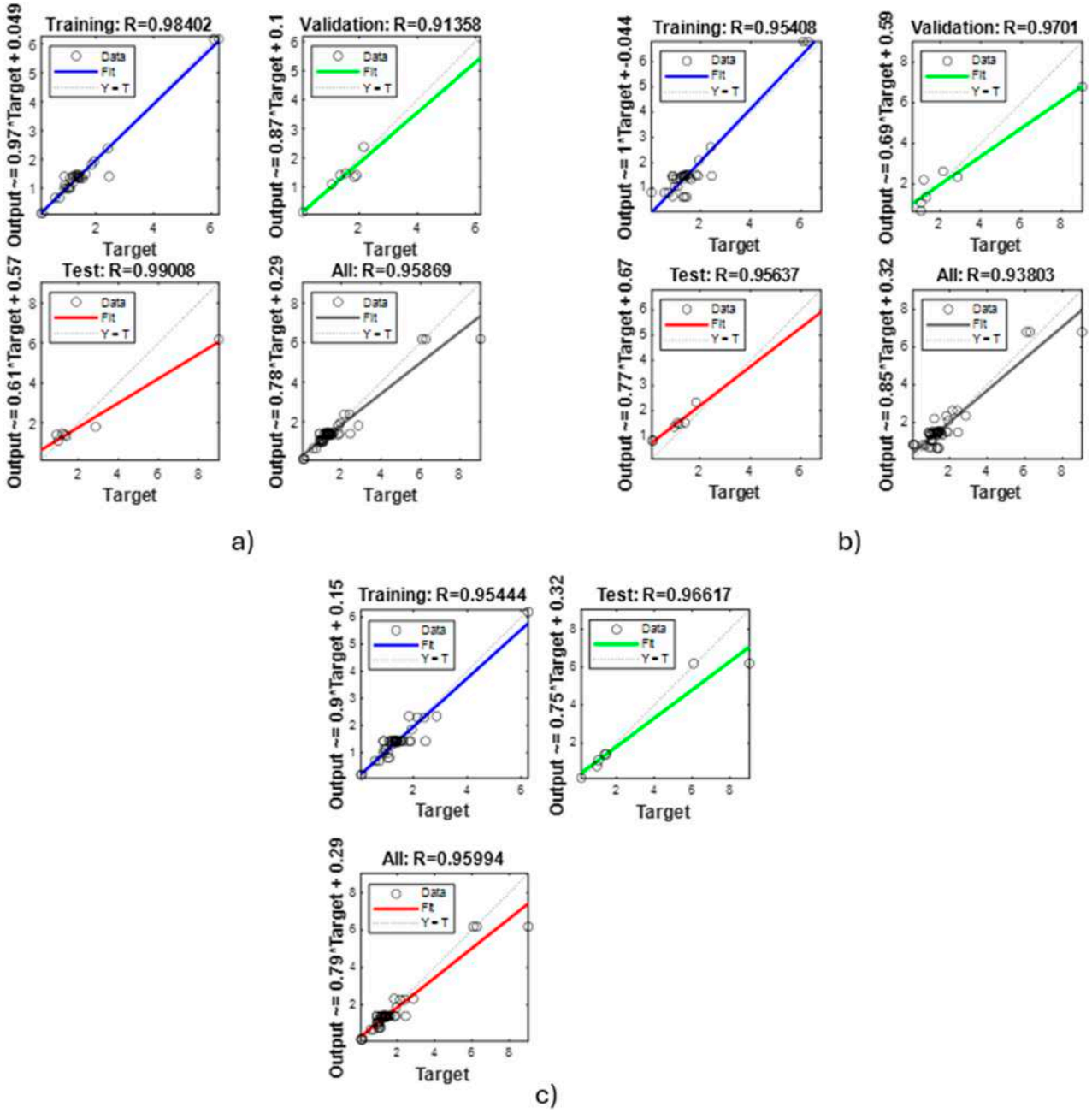


FIGURE 6
Regression statistics for individual sets and the total set for neural network model for concrete strength: **(A)** L-M, **(B)** SCG and **(C)** BR.

small amount of water (50 mL). Water for making aerated concrete was heated to 50°C. Heated water and soda ash solution were added to the dry mixture and mixed for 3 min. After that, aluminum powder was added to the obtained solution and stirring was carried out for 2 min. The ready mixture was poured into the molds to 2/3 of their height. The pulling of aerated concrete took place under normal conditions, then the samples were demolded and placed for storage into a moist chamber for 28 days.

Table 4 Sample set composition of the experimental concrete mixtures. To investigate the influence of the components on the strength and density of aerated concrete, a set of samples were made with different variations in the content of components. In a number of samples, the cement content was changed (from −10 to +20%). Some of the samples were made with the same amount of cement but different content of aluminum powder (from 0.15 to 0.65 g). The ash-and-slag component was partially replaced by traditional sand (from 10% to 100%).

The Kruskal–Wallis test was applied to determine the statistical significance of differences between the input variables and the output variables, namely, Strength and Density. This non-parametric test was selected due to the non-normal distribution of the data and the presence of small group sizes in many categories. The

TABLE 9 Comparison of ElasticNet and neural network models for concrete strength.

Quality indicators	Decision trees model	NN model with bayesian regularization
R (all data)	0.7549	0.9587
MSE	1.2989	0.2499
RMSE	1.1397	0.4999
RIE	0.4914	0.2155
MAE	0.5936	0.2458

Kruskal–Wallis test does not require the assumption of normality and is used to compare the median ranks of three or more independent groups. The assumptions of the Kruskal–Wallis test include the independence of samples and the assumption that the dependent variable is at least ordinal. The test is particularly well-suited for this dataset, where normality and equal variances are difficult to ensure due to small sample sizes. However, the Kruskal–Wallis test only determines whether there is a statistically significant difference among the groups, and does not specify which groups differ from one another.

The Table 5 presents the results of the Kruskal–Wallis test for each input variable, detailing the test statistics and *p*-values for both Strength and Density. The analysis indicates significant differences in the influence of some input variables, such as ZSHM, Sand, Soda ash, and Aluminum powder on both Strength and Density. Other variables, such as Cement and Water-solid ratio, did not show significant differences.

The Kruskal–Wallis test identified significant differences for several input variables with respect to both Strength and Density. Specifically, ZSHM, Sand, Soda ash, and Aluminum powder were found to significantly influence both Strength and Density. Conversely, Cement and Water-solid ratio did not exhibit statistically significant effects on either output variable. Water significantly affected Density, but not Strength. These results underscore the importance of material composition in influencing the mechanical properties of the samples, particularly in terms of density, which appears to be more sensitive to variations in input variables than strength.

2.4 Mechanical testing of aerated concrete

To measure compressive strength and average density, cubic shaped samples with 150 mm rib were used in the amount of 3 pieces per each composition at the age of 28 days. A correction factor was applied to account for the moisture content of the samples at the time of testing. The volume of the samples was calculated from their geometric dimensions, the mass of the samples was determined by weighing. Strength measurements were made on a 2 PG-10 automatic press machine with two ranges of 500/1000 kN.

2.5 Modelling properties of aerated concrete blocks using machine learning methods

The modeling process was executed employing decision trees and artificial neural networks through the use of Matlab 2023b software. The hardware setup included an Intel Core i7 CPU with 16 GB of RAM, which was sufficient for the computational needs of this study. The input parameters comprised cement, ZShM, sand, soda ash, aluminum powder, water, and the water-solid ratio, whereas the output parameters were focused on assessing concrete strength and concrete density. Given the investigation of two distinct output parameters, the study analyzed two variations of models. The schematic representation of these models is illustrated in Figure 1.

Two machine learning techniques, decision trees (DT) and neural networks (NN), were adopted for the prediction of concrete strength and density. The choice of these methods was motivated by the need to determine the most effective approach for forecasting concrete characteristics using a limited dataset of 50 samples. Decision trees were favored for their straightforward structure and ease of interpretation, which is crucial for enabling stakeholders to understand the basis for predictions. Neural networks were chosen for their ability to discern intricate, non-linear relationships in data, making them well-suited for capturing the complexities involved in concrete behavior. The employment of both DT and NN allows for a comprehensive assessment of predictive capabilities, ensuring the models are robust and precise despite the constraints of small data quantity. This side-by-side evaluation is instrumental in establishing which method yields the most reliable and accurate predictions of concrete strength and density.

In the domain of statistics, data mining, and machine learning, decision trees emerge as a predictive modeling algorithm characterized by its versatility and efficacy in addressing both classification and regression problems. These algorithms construct a structured hierarchy of logical “if, then” statements, leading to accurate classification or value estimation. Decision trees are distinguished by their predictive capacity and straightforward interpretability, segmenting datasets into a branched tree structure. This structure facilitates the differentiation between classification trees, which assign categorical labels, and regression trees, which forecast continuous outcomes. The iterative construction of decision trees entails the analysis of each variable and its potential divisions to optimize selection, focusing on enhancing predictive accuracy based on the mean-squared error (MSE) optimization criterion.

The development of Decision Tree (DT) models was rigorously orchestrated, with a focal emphasis on minimizing the mean-squared error (MSE) as the pivotal optimization criterion. This criterion was chosen to improve the accuracy of node splits and reduce prediction errors relative to the training dataset. Key parameters, such as the tree’s maximum depth (ranging from 5 to 50), the minimum number of samples required to initiate a node split (tested between 2 and 20), and the minimum number of samples required for a leaf node (ranging from 1 to 10), were carefully adjusted. These adjustments were essential to strike a balance between model complexity and mitigating the risk of overfitting. Additionally, the number of predictors sampled at each

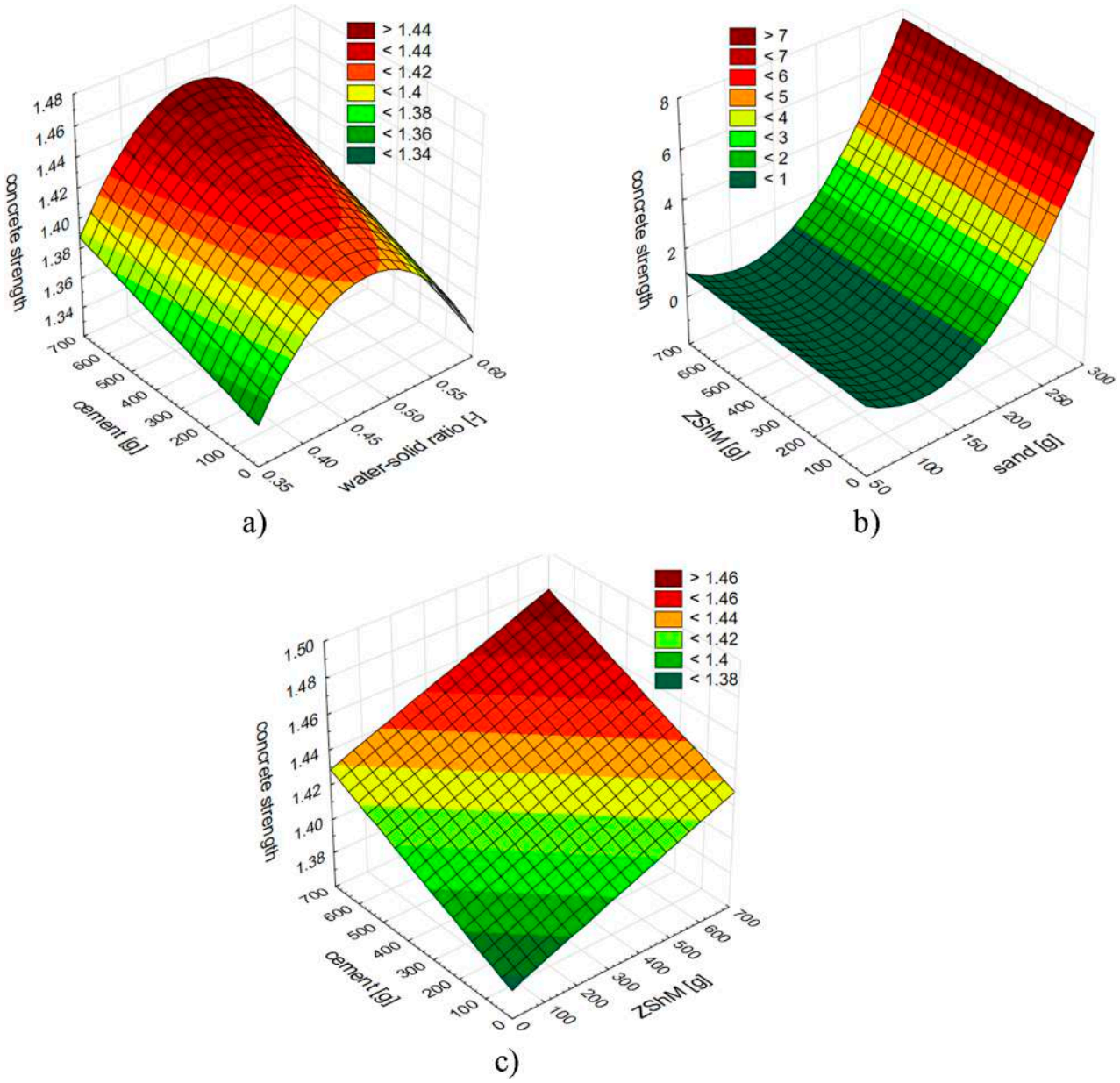


FIGURE 7
The network performance results for trained NN model with Bayesian Regularization depending on water-solid ratio and cement **(A)** depending on ZShM and sand **(B)** depending on ZShM and cement **(C)**.

split was set to 7, in alignment with the number of features in the dataset. This parameter, known as NumPredictorsToSample, was chosen to optimize the model's performance without introducing unnecessary complexity. The structural configuration of the model was systematically evaluated by varying the number of trees from 50 to 500, in increments of five, to identify the optimal model configuration. Each combination of these parameters was evaluated using MSE, while cross-validation was employed to ensure that the chosen parameters generalized well to unseen data. The dataset used for model training was split into training and testing sets in an 80/20 ratio, with a fixed random seed applied to ensure repeatability. This random split ensured a consistent division of the data between the training and testing sets.

Neural networks represent a fundamental category within the domain of machine learning, extensively utilized for both classification and regression challenges. The architecture of artificial neural networks (ANNs) is intricate, featuring multiple layers that transition from input to hidden stages, composed of densely interconnected neurons. These neurons process inputs by applying specific weights and biases, utilizing activation functions to introduce non-linearity - a key aspect for discerning complex patterns in data. Within the structure of a neural network, each

TABLE 10 Results of Decision Trees model quality indicators for concrete density.

Quality indicators	Decision trees model
R (all data)	0.86314
MSE	7512.87
RMSE	86.6768
RIE	0.1185
MAE	66.6475

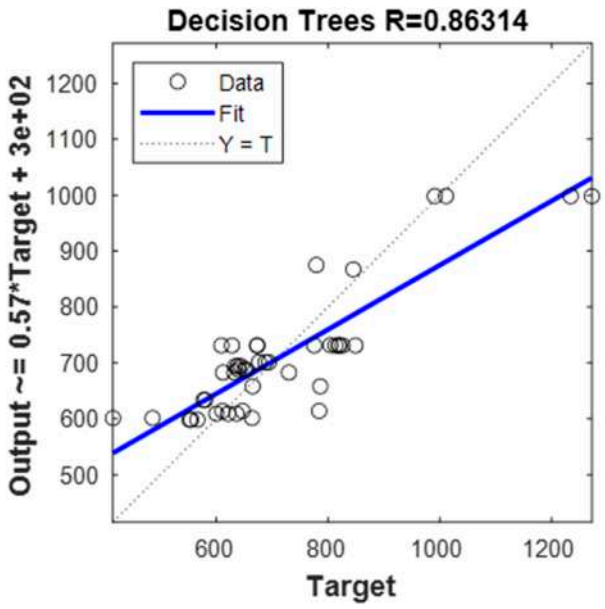


FIGURE 8
Regression statistics for the total set for Decision Trees model for concrete density.

neuron in a given layer computes a summation of inputs. This calculation involves the aggregation of each input multiplied by its respective weight, plus a bias unique to the neuron. Activation functions play a critical role by enabling neurons to encode and represent complex data patterns through non-linear dynamics. This capability significantly enhances the network's performance in tasks that surpass linear discrimination, pivotal for addressing complex challenges in machine learning. Neural networks excel in processing intricate and non-linear patterns, essential for complex applications such as visual and auditory recognition, and predictive analysis. They evolve through iterative optimization, fine-tuning connection weights to minimize discrepancies between actual and predicted outputs, thereby reducing error in the training dataset. This adaptability and advanced learning capacity are crucial for advancements in machine learning and artificial intelligence.

An experiment involving a neural network with a single hidden layer, varying neuron counts from 2 to 15, was conducted to determine the optimal balance between model complexity and predictive accuracy. This selection was based on iterative testing of various neuron configurations. The study assessed three learning algorithms: Levenberg-Marquardt (L-M), Bayesian Regularization (BR), and Scaled Conjugate Gradient (SCG), selected for their respective efficiencies. The L-M algorithm is fast but memory-demanding, halting when validation errors no longer decrease to prevent overfitting. In contrast, BR, though slower, offers better generalization, particularly for complex datasets, by adjusting to prevent overfitting. SCG, efficient in memory usage, also ceases training upon no further reduction in validation error. The dataset comprised 50 observations, with a 75% allocation for training and the remainder equally split between validation and testing.

For both DT and ANN models, the dataset was split into training, validation, and testing sets. The DT models used an 80/20 (training/testing) split, while the ANN models used a 70/15/15 (training/validation/testing) split. A fixed random seed was applied to ensure repeatability of the results, and although stratified sampling was not used, the random split was sufficient to maintain statistical consistency (e.g., mean and standard deviation) across the sets .

The evaluation of the quality of the models was carried out using specific metrics, as shown in Table 6. The adoption of these quality indicators is justified by their effectiveness in assessing the robustness of predictive models. The Regression Value (R) quantifies both the magnitude and the direction of the linear correlation between observed and predicted values. MSE and RMSE are essential for assessing the average severity of prediction errors, with RMSE emphasizing larger discrepancies due to its squaring operation. RIE provides a normalized evaluation of the errors, allowing analysis of the relative magnitude of the prediction errors. MAE, on the other hand, calculates the mean of the absolute discrepancies, providing valuable insight into the practical implications of prediction errors, particularly in the area of energy recovery. Together, these indicators provide a holistic view of model effectiveness, balancing the magnitude of error with the consistency of the model's predictive accuracy.

To ensure the reproducibility of the results, all code and scripts used in this study are available in a public repository at https://doi.org/10.5281/zenodo.13755460. The repository includes detailed documentation on how to run the code, as well as a list of dependencies and instructions for replicating the experiments.

2.6 Manufacturing of experimental samples

Experimental samples were produced in the laboratory of the Centre of Competence and Technology Transfer in the field of construction of D.Serikbayev EKTU. The amount of dry components was selected so that the volume of finished aerated concrete was not less than 0.01 m3. Solid components were used in a dry state. Dosing of dry components was carried out by weight. At the first stage cement and aggregate were mixed. Mixing was carried out using a construction drill-mixer. Separately, soda ash was dissolved in a small amount of water (50 mL). The water for making aerated concrete was heated to 500C. The heated water and soda ash solution was added to the dry mixture and mixed for 3 min. After that, aluminium powder was added to the obtained solution and stirred for 2 min. The ready mixture was poured into the moulds to 2/3 of the height. The aerated concrete was lifted under normal conditions,

TABLE 11 Results of neural network training for concrete density.

Network number	1	2	3
Training algorithm	Lavenberg-Marquardt	Scaled Conjugate Gradient	Bayesian Regularization
Epoch	17	40	178
Performance	2927.86	4740.52	4238.18
Best validation performance	9158.92 at epoch 11	5753.69 at epoch 34	4665.42 at apoch 15
Gradient	731	3471.71	1345.02

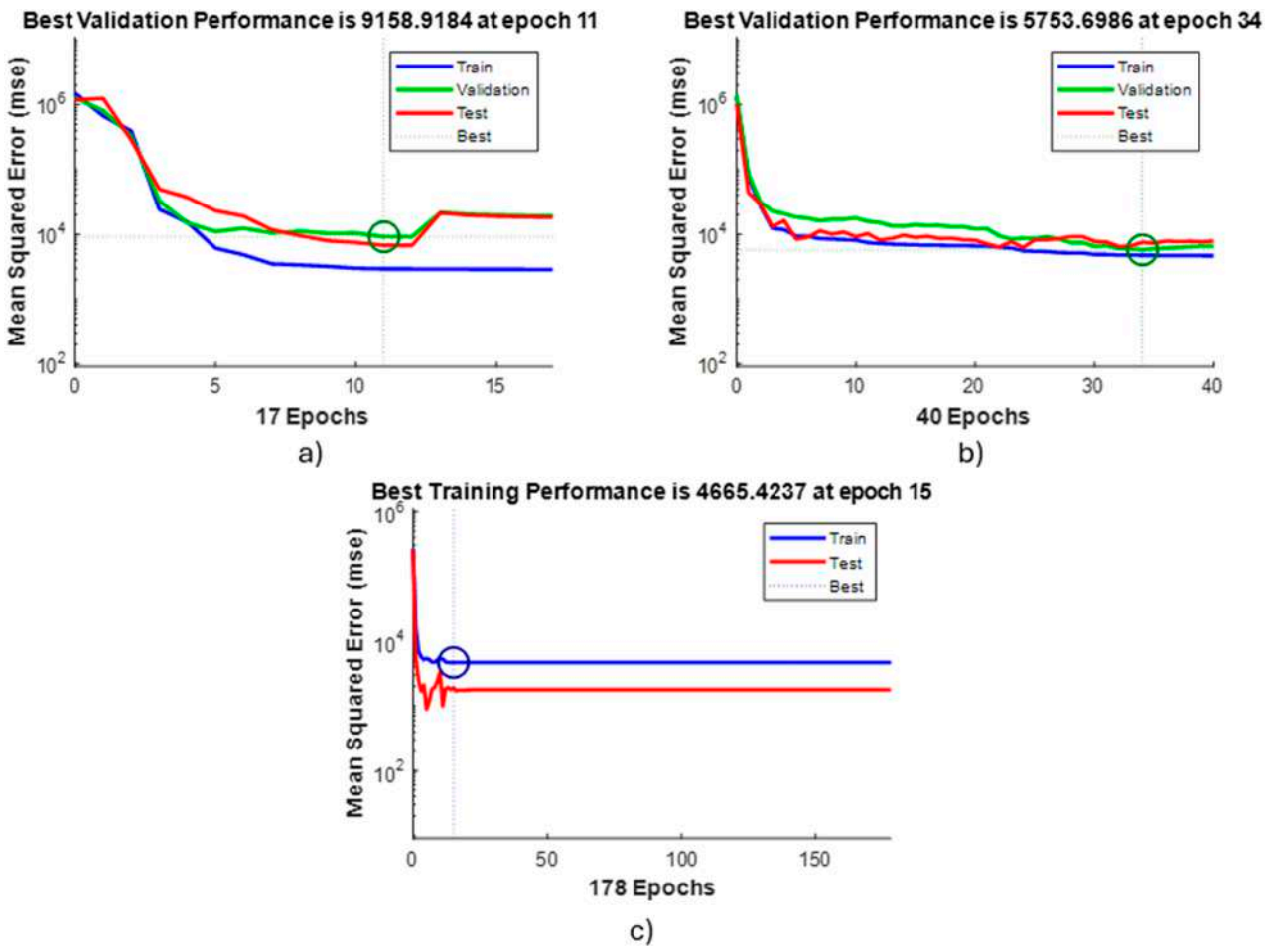


FIGURE 9
Network training process for concrete density - best validation performance of predictive model: **(A)** L-M, **(B)** SCG and **(C)** BR.

then the specimens were demoulded and placed for storage in a normal curing chamber for 28 days.

3 Results

3.1 Results for concrete strength

The first analyzed method for modeling the concrete strength is decision trees (DT). The process involved varying the number of trees within a range from 50 to 500 trees, with a step of 5 trees. The best DT modeling results were obtained for 55 trees. The model's consistent predictive ability was underscored by a Mean Absolute Error (MAE) of 0.5936, along with a Mean Squared Error (MSE) of 1.2989, indicating its ability to accurately fit the dataset. The Root Mean Squared Error (RMSE) was 1.1397, capturing the spread of prediction errors, while the Relative Importance of Errors (RIE) was 0.4914, further highlighting the model's prediction accuracy. The regression analysis of the entire dataset with the DT model is shown in Figure 2.

TABLE 12 Results of neural network quality indicators for concrete density.

Training algorithm	Lavenberg-marquardt	Scaled conjugate gradient	Bayesian regularization
R (all data)	0.91879	0.8972	0.91997
MSE	4487.43	5329.91	4238.18
RMSE	66.98	73.006	65.101
RIE	0.0916	0.0998	0.089
MAE	46.453	57.692	51.597

The second method of analysis was neural networks (NN). The ANN, when utilizing the Levenberg-Marquardt (L-M) learning algorithm, demonstrated optimal performance with a configuration of 6 neurons in the hidden layer, a milestone achieved following 29 iterations. Subsequent to establishing the efficacy of a 6-neuron hidden layer configuration, the training regimen was expanded to incorporate two additional learning algorithms: the Scaled Conjugate Gradient (SCG) and Bayesian Regularization (BR). The training structure of the neural network is shown in Figure 3. The Table 7 displays the results of neural network training, including network numbers, training algorithms used, number of epochs, performance outcomes, best validation performances along with the epochs they were achieved in, and gradients.

The Figure 4 shows three graphs corresponding to the training process of different neural networks, illustrating the mean square error in different numbers of epochs for the training, validation and test sets, highlighting the best validation results for each network for different training algorithms: L-M (Figure 4A), SCG (Figure 4B) and BR (Figure 4C). Indicators for assessing the quality of the network are delineated in Table 8.

To mitigate the risk of overfitting, a condition wherein the model exhibits high accuracy on training data but fails to generalize effectively to new data, the training regimen was closely monitored. The cessation of training was triggered upon the identification of six consecutive increases in error rates during the validation phase, or in the absence of any amelioration in error metrics. This strategy, commonly known as "early stopping," aims to forestall overfitting by discontinuing the training at the juncture where a decrement in the model's validation performance is observed. Figure 5 delineates the evolution of the network's training progression.

The graph in Figure 6 shows the correlation between observed and predicted values through regression analysis for various data sets using three different model training algorithms, each shown in sub-figures a, b, and c. Each graph plots actual target values on the horizontal axis against model predictions on the vertical axis, and includes a best-fit line demonstrating model performance. The ideal prediction scenario, where predictions and targets are equal, is represented by a dashed diagonal line labeled Y = T. The degree to which the scatter points and best fit line approach the Y = T line indicates the accuracy of the model's predictions. The graph indicates that each model has a commendable level of predictive ability.

For models trained with the Levenberg-Marquardt algorithm, the calculated regression coefficients (R values) for various data partitions were: 0.984 for training, 0.914 for validation, 0.99 for testing, and 0.9587 for all data combined. Models utilizing the Scaled Conjugate Gradient algorithm displayed R values of 0.954 for training, 0.97 for validation, 0.956 for testing, and 0.938 for the combined data sets. Lastly, models employing Bayesian Regularization algorithm showed R values of 0.954 for training, 0.966 for validation, and 0.9587 for the combined data sets. All models yielded R values exceeding the 0.9 benchmark, indicating a high fidelity in the models' ability to echo the actual data trends.

When considering the regression coefficient (R) and the network performance measures—MSE, RMSE, RIE, and MAE - it is evident that the network utilizing the Bayesian Regularization training algorithm achieved the most favorable outcomes (Table 9). In comparison to the Decision Tree (DT) model, the neural network with Bayesian Regularization surpassed the DT across all mentioned performance metrics. The neural network demonstrated enhanced accuracy with a greater R-value (0.9587 compared to 0.7549) and notably reduced error indices, including MSE, RMSE, RIE, and MAE. This indicates that, within the given scenario, the neural network employing Bayesian Regularization has a higher proficiency in modeling the intricacies of the dataset and delivering predictions with greater precision than the DT model.

As a result of the modeling, it was possible to predict the concrete strength using the trained NN model with Bayesian regularization by entering the input data into Statistica. The results of the networks are shown in the following figures, for the NN model with Bayesian regularization as a function of water-solid ratio and cement (Figure 7A), as a function of ZShM and sand (Figure 7B), and as a function of ZShM and cement (Figure 7C).

3.2 Results for concrete density

Modeling was performed using the Decision Trees (DT) algorithm, with the number of trees adjusted in increments of 5 in the range from 50 to 500. The most effective DT model was identified with a configuration of 85 trees, which had a Mean Absolute Error (MAE) of 66.6475, indicating reliable prediction quality. The model also showed an impressive ability to fit the data set with a Mean Squared Error (MSE) of 7512.87. The Root Mean Squared Error (RMSE) of 86.6768 reflected the distribution of prediction errors, while the model's accuracy in prediction was evident from a Relative Importance Error (RIE) of 0.1185. These metrics are detailed in Table 10, and the regression analysis applying the DT model to the entire dataset can be observed in Figure 8.

For the NN and the Levenberg-Marquardt (L-M) learning algorithm, the best results were achieved using a 8-neuron hidden layer, culminating in optimal performance within 17 iterations. The investigation proceeded to evaluate the network under two additional algorithms: Bayesian Regularization (BR) and Scaled Conjugate Gradient (SCG). Details of the neural network's training outcomes, are systematically documented in Table 11.

Figure 9 presents a series of plots that chart the evolution of the mean square error across various epochs in the training,

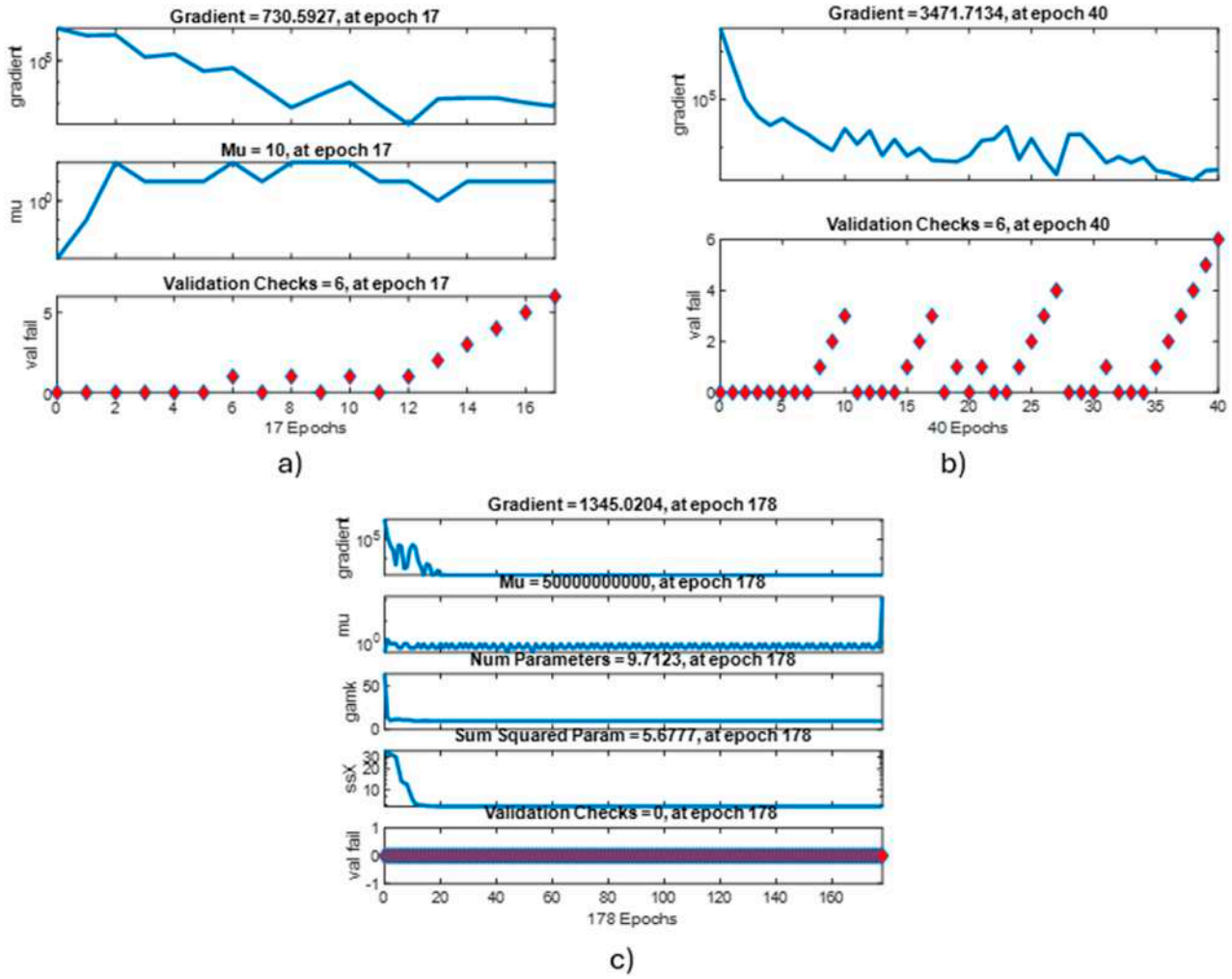


FIGURE 10
Network training process for concrete density - validation checks: **(A)** L-M, **(B)** SCG and **(C)** BR.

testing, and validation datasets of neural networks. These plots serve to accentuate the optimal validation performances attained using different training algorithms: BR (Figure 9C), SCG (Figure 9B), and L-M (Figure 9A). The parameters used for evaluating the quality of the network's training are specified in Table 12. Figure 10 illustrates the development of the network's training over time.

Figure 11 contains a graphs that illustrates the relationship between actual and forecasted data points, determined by regression analysis across multiple data sets. This relationship is visualized using three distinct model training algorithms. For the L-M algorithm (Figure 11A), the training set displays an R-value of 0.91495, indicating a strong positive linear relationship, which is slightly lower in the validation set with an R-value of 0.8659, and shows a notable decrease in the test set with an R-value of 0.95937. The overall R-value across all data sets is 0.91879, signifying high predictive accuracy of the model. The SCG algorithm (Figure 11B) exhibits a similar trend with the training set having an R-value of 0.91325, the validation set at 0.91183, and the test set at 0.92481. The composite R-value for all data is slightly less compared to the L-M at 0.89724. Figure 11C, depicting the BR algorithm, shows a training R-value of 0.91824, which is comparable to the other algorithms. The test set performs better than both L-M and SCG with an R-value of 0.97588. The overall R-value for the BR algorithm is the highest among the three at 0.91997, indicating the most consistent and closest predictions to the actual values across all sets. The analysis of these R-values indicates that while all three models perform well, the BR algorithm seems to provide a marginally better fit for the data across all subsets, as reflected in the composite R-value, which is the highest among the algorithms evaluated.

Examining the regression coefficient (R) along with various indicators of network performance such as MSE, RMSE, RIE, and MAE, it can be seen that the networks trained with Bayesian regularization produced the most advantageous results (see Table 13). Compared to the Decision Tree (DT) model, the performance of the Bayesian Regularized neural network is superior in all evaluated metrics. This neural network exhibits significantly higher accuracy, as evidenced by an R-value of 0.91997 compared to 0.86314 for the DT, and significant reductions in error measures, including MSE, RMSE, RIE, and MAE. These results suggest that, for the specific case at

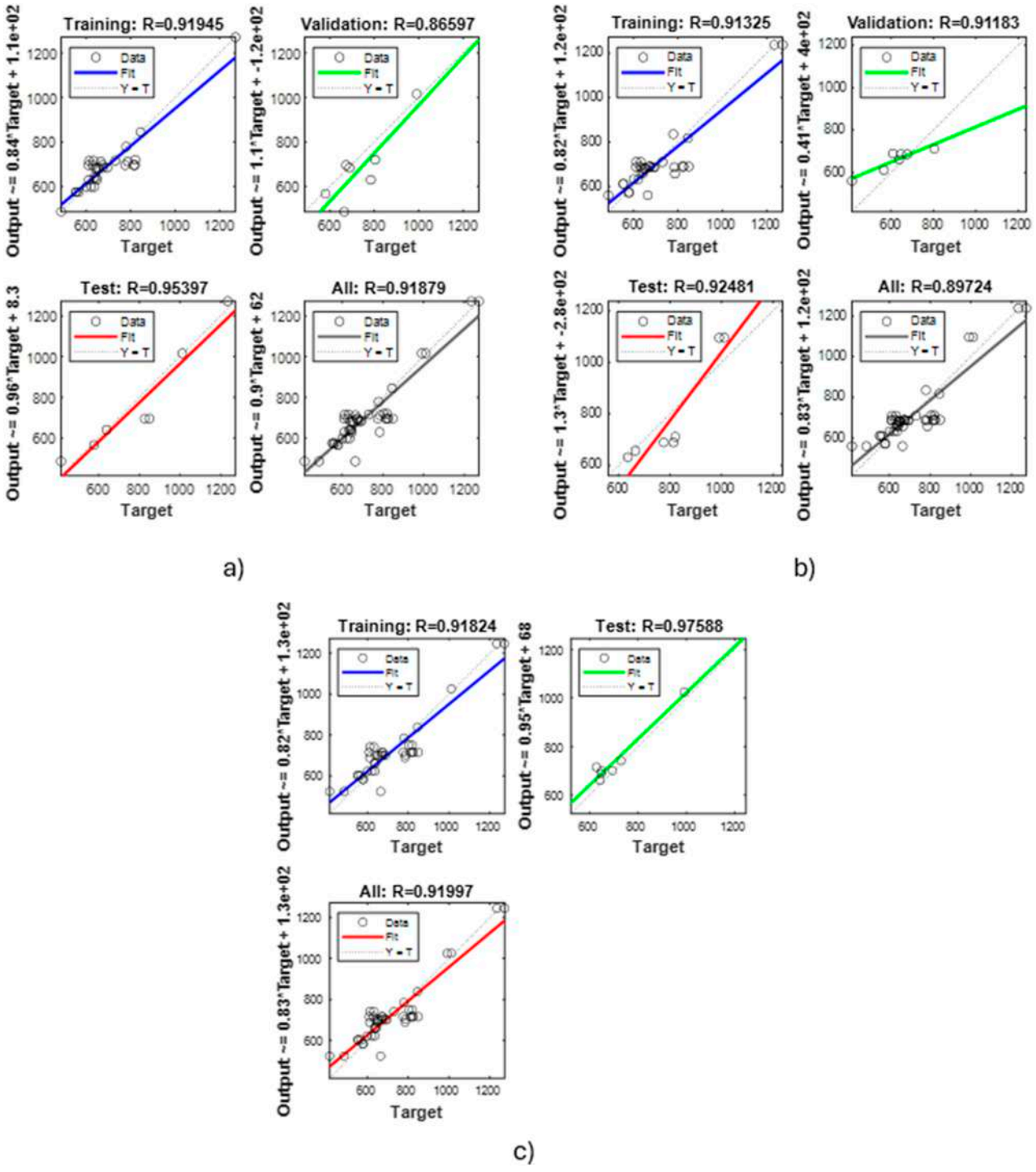


FIGURE 11
Regression statistics for individual sets and the total set for neural network model for concrete density: **(A)** L-M, **(B)** SCG and **(C)** BR.

hand, neural networks applying Bayesian regularization have an enhanced ability to capture the complexity inherent in the data, thereby providing predictions with increased accuracy compared to those of the DT model.

As a result of the modeling, it was possible to predict the concrete density using the trained NN model with Bayesian regularization by entering the input data into Statistica. The results of the networks are shown in the following figures, for the NN model with Bayesian regularization as a function of water-solid ratio and aluminum powder (Figure 12A), as a function of ZShM and aluminum powder (Figure 12B), and as a function of ZShM and sand (Figure 12C).

TABLE 13 Comparison of ElasticNet and neural network models for concrete density.

Quality indicators	Decision trees model	NN model with bayesian regularization
R (all data)	0.86314	0.91997
MSE	7512.87	4238.18
RMSE	86.6768	65.101
RIE	0.1185	0.089
MAE	66.6475	51.597

4 Discussion

This study focused on modeling the properties of aerated concrete using local raw materials and ash-and-slag wastes, employing advanced machine learning techniques with the aim of achieving an optimal balance between compressive strength and density. This balance is especially important for applications in seismic regions, where building materials must meet high standards of performance and safety. The results obtained from this investigation provide substantial insights with implications for both the construction industry and environmental sustainability. The comparative analysis conducted between Decision Trees (DT) and Artificial Neural Networks (ANN) highlighted a significant advantage in using neural networks for predicting concrete properties. While Decision Trees offered reasonable accuracy, ANN models, particularly those incorporating Bayesian Regularization, consistently outperformed DT across all evaluated metrics. Specifically, ANN models yielded a regression value (R) of 0.9587 for compressive strength and 0.91997 for density, indicating far more reliable and accurate predictions when compared to Decision Tree models.

The superior performance of ANN models can largely be attributed to their ability to model complex non-linear relationships between input variables and the target outcomes. In concrete technology, interactions between materials such as cement, ZSHM, sand, soda ash, aluminum powder, and water are highly intricate and non-linear. Traditional models like Decision Trees often struggle to capture these complexities, while ANN models are adept at generalizing and learning from such non-linear relationships, making them more effective in predicting concrete properties. Bayesian Regularization played a pivotal role in enhancing the generalization capacity of ANN models. This technique mitigates overfitting by introducing regularization terms in the optimization process, which penalize excessive weights and biases within the network, ensuring that the model captures general patterns in the data rather than merely memorizing the training data. The improved performance metrics, such as lower Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Relative Information Entropy (RIE), and Mean Absolute Error (MAE), highlight the robustness of the Bayesian Regularized ANN models in predicting the properties of aerated concrete.

Future work in this area could greatly benefit from the inclusion of interpretability techniques such as SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-agnostic Explanations) to better understand the contributions of individual input variables to the model's predictions. These techniques would enhance the transparency of the model, allowing for practical applications in construction scenarios where understanding the decision-making process is crucial. One of the most significant contributions of this study is the successful integration of ash-and-slag waste from thermal power plants into aerated concrete mixtures. This innovation not only addresses waste management issues but also significantly reduces the environmental footprint of concrete production. By replacing traditional raw materials like cement with industrial by-products, the study contributes to lowering the carbon footprint associated with cement production, which is one of the largest contributors to greenhouse gas emissions in the construction industry. This use of waste materials, without compromising the mechanical properties of aerated concrete, highlights the feasibility of adopting sustainable construction practices on a larger scale.

The reduction in greenhouse gas emissions, combined with the diversion of waste materials from landfills, aligns with global sustainability objectives and demonstrates the potential for eco-friendly innovations in construction. By optimizing the mix proportions of local raw materials and ash-and-slag waste, the study presents a viable path for producing non-autoclaved aerated concrete blocks that are suitable for earthquake-resistant structures. The lightweight nature and increased compressive strength of these blocks make them particularly beneficial for use in seismic regions, where structural resilience is critical. The findings indicate that these aerated concrete blocks could serve as a sustainable and cost-effective alternative to traditional building materials, offering both safety and environmental benefits in earthquake-prone areas. This research opens several promising directions for future work. One essential area is the continuous enhancement of machine learning models by incorporating larger, more diverse datasets. Expanding the dataset size would improve the models' accuracy and reliability, while also helping to mitigate the risk of overfitting, especially in complex models such as neural networks. Incorporating data from different geographical regions and varying material compositions could further broaden the applicability of these models, ensuring they generalize well across different conditions.

Exploring more advanced machine learning techniques, such as ensemble methods like Random Forests, Gradient Boosting, or deep learning architectures such as Convolutional Neural Networks, could potentially provide even better predictive capabilities. Additionally, future research should focus on the long-term durability and performance of aerated concrete blocks under real-world conditions, particularly in seismic regions. Testing these materials under various environmental factors such as moisture, temperature fluctuations, and seismic loads would provide more comprehensive insights into their real-world applicability. Another significant direction for future studies is investigating the use of additional sustainable materials that could be incorporated into concrete mixtures. Incorporating other types of waste materials could further reduce the environmental impact of concrete production, while maintaining or even improving material performance. Such innovations would contribute to the development of more environmentally friendly building materials that not only meet structural requirements but also align with sustainability objectives.

Despite the promising results of this study, several limitations must be acknowledged. The relatively small sample size (50 observations) limits the generalizability of the findings. While cross-validation was employed to mitigate the risk of overfitting, a larger dataset is needed to fully validate the predictive power of the models

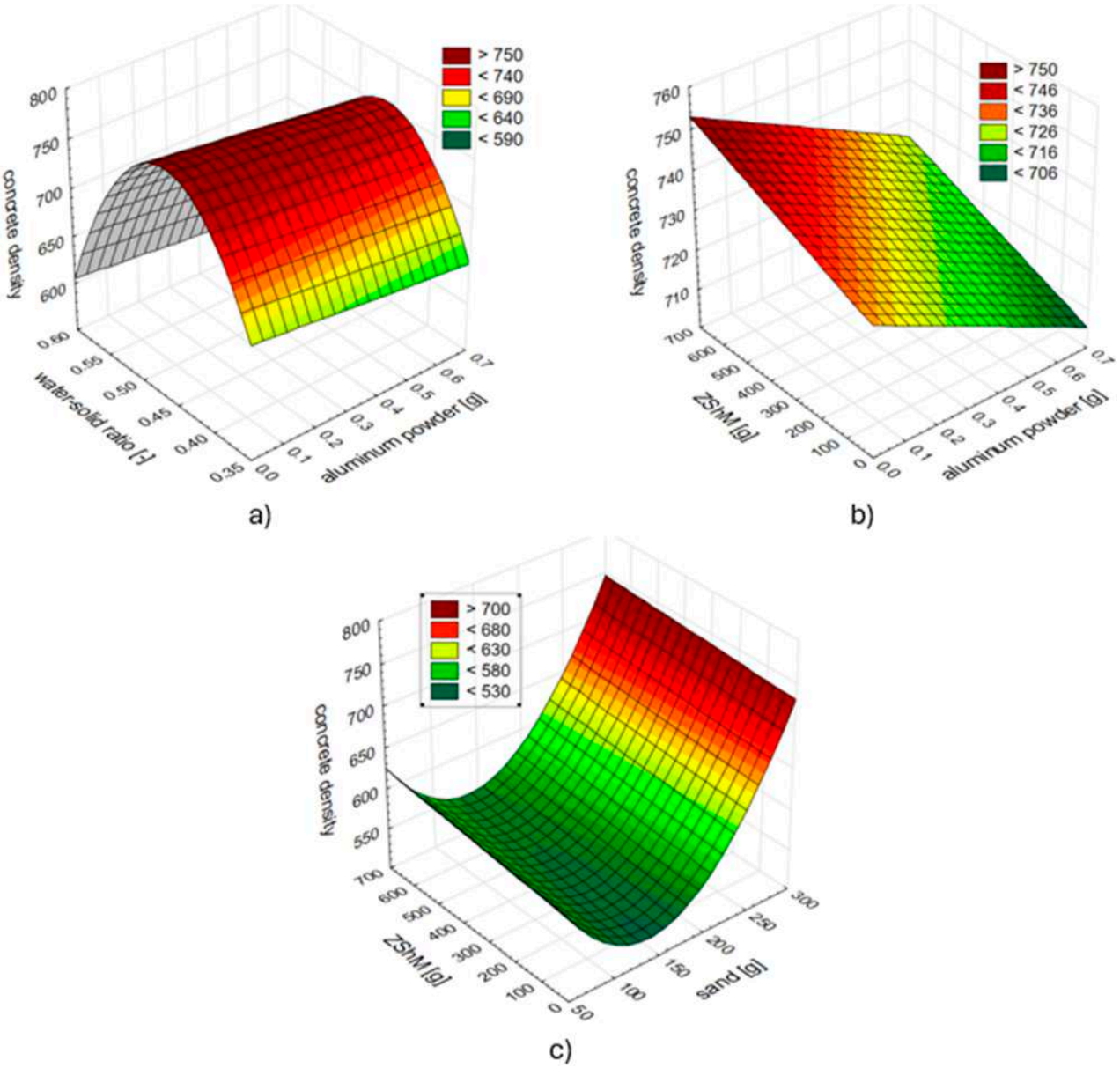


FIGURE 12
The network performance results for trained NN model with Bayesian Regularization depending on water-solid ratio and aluminum powder **(A)** depending on ZShM and aluminum powder **(B)** depending on ZShM and sand **(C)**.

across different material compositions and contexts. Expanding the dataset in future studies would provide a more robust basis for model generalization and accuracy. Another limitation is the inherent risk of overfitting in the ANN models. Although techniques like Bayesian Regularization and early stopping were implemented to mitigate this issue, the complexity of neural networks can still lead to overfitting, particularly when working with small datasets. Future research should focus on expanding the dataset and exploring additional interpretability techniques to further minimize the risk of overfitting and improve model transparency.

The interpretability of ANN models remains a challenge, despite their superior predictive capabilities. Unlike simpler models such as Decision Trees, which offer clear and understandable decision paths, ANN models often function as black boxes, making it difficult to directly explain the contributions of individual input variables to the predictions. Incorporating interpretability tools such as SHAP or LIME in future studies could provide valuable transparency, allowing for a deeper understanding of how input variables influence the model's outcomes. Finally, implementing these machine learning models in real-world applications presents additional challenges. The variability in local raw materials and environmental conditions could affect the accuracy and applicability of these models in different regions. Further validation across various geographic regions and material compositions is necessary

to ensure that these models generalize well in practical construction scenarios. Addressing these challenges will be crucial for confidently applying machine learning models in the construction industry.

In conclusion, this study successfully demonstrated the potential of advanced machine learning techniques for modeling the properties of aerated concrete using local raw materials and ash-and-slag waste. The results show that Bayesian Regularized ANN models are particularly effective for optimizing concrete mixtures for seismic regions. The environmental benefits, coupled with the enhanced mechanical properties of the concrete, underscore the importance of sustainable construction practices and highlight the role of AI in advancing concrete technology. Future research should continue exploring new materials and innovative modeling techniques to further enhance the sustainability and resilience of concrete structures.

5 Conclusion

The conducted research demonstrates the appropriateness of employing machine learning methods for disjointed datasets. A comparative analysis between Decision Trees and Neural Networks, particularly those using Bayesian Regularization, clearly indicates the superiority of Neural Networks in modeling complex relationships involved in predicting concrete strength and density. Neural Networks consistently outperformed Decision Trees across all performance metrics, achieving higher regression values ($R_{strength}$ = 0.9587 and $R_{density}$ = 0.91997) and significantly lower error indices (MSE, RMSE, RIE, MAE). This showcases their advanced capability to capture intricate non-linear patterns in the data. The intricate architecture and application of Bayesian Regularization in Neural Networks have proven highly effective in understanding and predicting concrete properties, making them the preferred choice over simpler Decision Trees.

The comprehensive evaluation metrics solidified the understanding that Neural Networks offer a more accurate and reliable modeling approach. Robust validation techniques, including early stopping, effectively mitigated overfitting, ensuring the models' reliability and applicability to new data.

The results of the analysis indicate that producing non-autoclaved aerated concrete blocks for construction in seismic regions is feasible; however, the use of local materials alone is insufficient. Future research should focus on maintaining the correct ratio between density (weight) and strength of the blocks through a combination of methods to improve the properties of aerated concrete and the application of reliable models. The results of the work made it possible to identify key characteristics that could potentially provide unique advantages in modeling the properties of aerated concrete (containing various components, including ash and slag waste) and other building mixtures using the machine learning paradigm for a wide range of applications.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

OR: Writing–original draft, Writing–review and editing, Funding acquisition. DG: Writing–original draft, Writing–review and editing. MS: Conceptualization, Investigation, Methodology, Writing–review and editing. NB: Formal Analysis, Software, Supervision, Visualization, Writing–review and editing. MK: Data curation, Formal Analysis, Resources, Software, Writing–review and editing. MB: Conceptualization, Project administration, Supervision, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research is funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR21882292 – "Integrated development of sustainable construction industries: innovative technologies, optimization of production, effective use of resources and creation of technological park").

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahmad, I., and Shokouhian, M. (2024). Promoting sustainable green infrastructure: experimental and numerical investigation of concrete reinforced with recycled steel fibers. *Archives Adv. Eng. Sci.* doi:10.47852/bonviewaaes42022837

Akeke, G. A., Inem, P. E. U., Alaneme, G. U., and Nyah, E. E. (2023). Experimental investigation and modelling of the mechanical properties of palm oil fuel ash concrete using Scheffe's method. *Sci. Rep.* 13 (1), 18583. doi:10.1038/s41598-023-45987-3

Akor, J., Onjefu, L., Edeh, J., and Olubambi, A. (2023). Suitability of crushed sandcrete block (CSB) as a partial replacement for fine aggregate in concrete structures. *Archives Adv. Eng. Sci.* 26. doi:10.47852/bonviewaaes32021741

Alaneme, G. U., Olonade, K. A., and Esenogho, E. (2023). Critical review on the application of artificial intelligence techniques in the production of geopolymer-concrete. *SN Appl. Sci.* 5 (8), 217. doi:10.1007/s42452-023-05447-z

Anjum, M., Khan, K., Ahmad, W., Ahmad, A., Amin, M. N., and Nafees, A. (2022). Application of ensemble machine learning methods to estimate the compressive strength of Fiber-Reinforced Nano-Silica modified concrete. *Polymers* 14 (18), 3906. doi:10.3390/polym14183906

Chen, R. S., Zhang, H. Y., Hao, X. K., Yu, H. X., Shi, T., Zhou, H. S., et al. (2024). Experimental study on ultimate bearing capacity of short thin-walled steel tubes reinforced with high-ductility concrete. *Structures* 68, 107109. doi:10.1016/j.istruc.2024.107109

Chen, S., Zhou, M., Shi, X., and Huang, J. (2023). A novel MBAS-RF approach to predict mechanical properties of Geopolymer-Based compositions. *Gels* 9 (6), 434. doi:10.3390/gels9060434

Elhishi, S., Elashry, A. M., and El-Metwally, S. (2023). Unboxing machine learning models for concrete strength prediction using XAI. *Sci. Rep.* 13 (1), 19892. doi:10.1038/s41598-023-47169-7

EN 197-1: 2014. Cement – Part 1: composition, specifications and conformity criteria for common cements.

Fang, B., Qian, Z., Song, Y., Diao, X., Shi, T., Cai, X., et al. (2024). Evaluation of early crack resistance performance of concrete mixed with ternary minerals using temperature stress testing machine (TSTM). *J. Clean. Prod.* 465, 142780. doi:10.1016/j.jclepro.2024.142780

Farooq, F., Ahmed, W., Akbar, A., Aslam, F., and Alyousef, R. (2021). Predictive modeling for sustainable high-performance concrete from industrial wastes: a comparison and optimization of models using ensemble learners. *J. Clean. Prod.* 292, 126032. doi:10.1016/j.jclepro.2021.126032

Gamil, Y. (2023). Machine learning in concrete technology: a review of current researches, trends, and applications. *Front. Built Environ.* 9. doi:10.3389/fbuil.2023.1145591

GOST 5494-95 (2006). *Aluminum powder. Specifications*. Moscow: Standartinform, 12.

GOST 8736-2014 (2015). *Sand for construction works*. Moscow: Standartinform, 14.

Hameed, M. M., Abed, M. A., Al-Ansari, N., and Alomar, M. K. (2022). Predicting compressive strength of concrete containing industrial waste materials: novel and hybrid machine learning model. *Adv. Civ. Eng.* 2022, 1–19. doi:10.1155/2022/5586737

Haq, M. Z. U., Sood, H., Kumar, R., and Ricciotti, L. (2024). Sustainable geopolymers from polyethylene terephthalate waste and industrial by-products: a comprehensive characterisation and performance predictions. *J. Mater. Sci.* 26, 3858–3889. doi:10.1007/s10853-024-09447-1

Hu, X., Li, B., Mo, Y., and Alselwi, O. (2021). Progress in artificial intelligence-based prediction of concrete performance. *J. Adv. Concr. Technol.* 19 (8), 924–936. doi:10.3151/jact.19.924

Huang, H., Guo, M., Zhang, W., Zeng, J., Yang, K., and Bai, H. (2021). Numerical investigation on the bearing capacity of RC columns strengthened by HPFL-BSP under combined loadings. *J. Build. Eng.* 39, 102266. doi:10.1016/j.jobe.2021.102266

Li, Z., Yoon, J., Zhang, R., Rajabipour, F., Srubar, W. V., Dabo, I., et al. (2022). Machine learning in concrete science: applications, challenges, and best practices. *Npj Comput. Mater.* 8 (1), 127. doi:10.1038/s41524-022-00810-x

Lu, D., Ma, C., Du, X., Jin, L., and Gong, Q. (2017). Development of a new nonlinear unified strength theory for geomaterials based on the characteristic stress concept. *Int. J. Geomechanics* 17 (2). doi:10.1061/(asce)gm.1943-5622.0000729

Moein, M. M., Saradar, A., Rahmati, K., Mousavinejad, S. H. G., Bristow, J., Aramali, V., et al. (2023). Predictive models for concrete properties using machine learning and deep learning approaches: a review. *J. Build. Eng.* 63, 105444. doi:10.1016/j.jobe.2022.105444

Moradi, N., Tavana, M. H., Habibi, M. R., Amiri, M., Moradi, M. J., and Farhangi, V. (2022). Predicting the compressive strength of concrete containing binary supplementary cementitious material using machine learning approach. *Materials* 15 (15), 5336. doi:10.3390/ma15155336

Nafees, A., Amin, M. N., Khan, K., Nazir, K., Ali, M., Javed, M. F., et al. (2021). Modeling of mechanical properties of silica Fume-Based green concrete using machine learning techniques. *Polymers* 14 (1), 30. doi:10.3390/polym14010030

Qian, Y., Sufian, M., Hakamy, A., Deifalla, A. F., and El-Said, A. (2023). Application of machine learning algorithms to evaluate the influence of various parameters on the flexural strength of ultra-high-performance concrete. *Front. Mater.* 9. doi:10.3389/fmats.2022.1114510

Raju, M. R., Rahman, M., Hasan, M. M., Islam, M. M., and Alam, M. S. (2023). Estimation of concrete materials uniaxial compressive strength using soft computing techniques. *Heliyon* 9 (11), e22502. doi:10.1016/j.heliyon.2023.e22502

Razveeva, I., Kozhakin, A., Beskopylny, A. N., Stel'makh, S. A., Shcherban, E. M., Artamonov, S., et al. (2023). Analysis of geometric characteristics of cracks and delamination in aerated concrete products using convolutional neural networks. *Buildings* 13 (12), 3014. doi:10.3390/buildings13123014

Song, X., Wang, W., Deng, Y., Su, Y., Jia, F., Zaheer, Q., et al. (2024). Data-driven modeling for residual velocity of projectile penetrating reinforced concrete slabs. *Eng. Struct.* 306, 117761. doi:10.1016/j.engstruct.2024.117761

Ullah, H. S., Khushnood, R. A., Ahmad, J., and Farooq, F. (2022). Predictive modelling of sustainable lightweight foamed concrete using machine learning novel approach. *J. Build. Eng.* 56, 104746. doi:10.1016/j.jobe.2022.104746

Wang, M., Kang, J., Liu, W., Su, J., and Li, M. (2022). Research on prediction of compressive strength of fly ash and slag mixed concrete based on machine learning. *PLoS ONE* 17 (12), e0279293. doi:10.1371/journal.pone.0279293

Wang, R., Zhang, J., Lu, Y., Ren, S., and Huang, J. (2024). Towards a reliable design of geopolymer concrete for green landscapes: a comparative study of tree-based and regression-based models. *Buildings* 14 (3), 615. doi:10.3390/buildings14030615

Wu, P., Liu, X., Zhang, Z., Wei, C., Wang, J., and Gu, J. (2024). The harmless and value-added utilization of red mud: recovering iron from red mud by pyrometallurgy and preparing cementitious materials with its tailings. *J. Industrial Eng. Chem.* 132, 50–65. doi:10.1016/j.jiec.2023.11.038

Yang, D., Zhao, J., Suhail, S. A., Ahmad, W., Kamiński, P., Dyczko, A., et al. (2022). Investigating the ultrasonic pulse velocity of concrete containing waste marble dust and its estimation using artificial intelligence. *Materials* 15 (12), 4311. doi:10.3390/ma15124311

Yi, G., Ye, X., and Li, Q. (2023). Empirical study of surface deterioration analysis based on random fields for reinforced concrete structures in marine environment. *Materials* 16 (11), 4150. doi:10.3390/ma16114150

Zheng, W., Shui, Z., Xu, Z., Gao, X., and Zhang, S. (2023). Multi-objective optimization of concrete mix design based on machine learning. *J. Build. Eng.* 76, 107396. doi:10.1016/j.jobe.2023.107396

Zhou, X., Lu, D., Du, X., Wang, G., and Meng, F. (2019). A 3D non-orthogonal plastic damage model for concrete. *Comput. Methods Appl. Mech. Eng.* 360, 112716. doi:10.1016/j.cma.2019.112716