Journal of Materials Research and Technology 40 (2026) 4047–4062

Contents lists available at ScienceDirect

Journal of Materials Research and Technology

journal homepage: www.elsevier.com/locate/jmrt

Bio-inspired machine learning for strength prediction of cement mortars incorporating reservoir waste silt: A case study from Bologna, Italy

Hashem Jahangir [a], Danial Rezazadeh Eidgahee [b], Atefeh Soleymani [c], Cesare Sangiorgi [d], Piergiorgio Tataranni [d], Abbas Solouki [e,*]

[a] Department of Civil Engineering, Research Group of Novel Technologies in Civil Engineering, University of Birjand, Birjand, Iran
[b] Department of Civil Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
[c] Department of Civil Engineering, University of Birjand, Birjand, Iran
[d] Department of Civil, Chemical, Environmental and Materials Engineering, University of Bologna, Bologna, Italy
[e] Civil Engineering Division, Department of Engineering, University of Cambridge, Cambridge, United Kingdom

ARTICLE INFO

Keywords:
Cement mortar
Waste silt
Hybrid artificial bee colony-artificial neural networks (ABC-ANN) algorithm
Combinatorial group method of data handling (GMDH-Combi)
Flexural strength
Unconfined compressive strength

ABSTRACT

This study addresses the challenge of accurately predicting the mechanical performance of cement mortars incorporating waste silt, a byproduct from reservoir sedimentation, which has potential for sustainable construction applications. The aim is to develop predictive models for the flexural strength (FS) and unconfined compressive strength (UCS) of such mortars using advanced machine learning techniques. Two hybrid algorithms were applied: artificial neural networks optimized via the artificial bee colony algorithm (ABC-ANN) and the combinatorial group method of data handling (GMDH-Combi). Input parameters included the proportions of cement, water, sand, silt, and additives, while FS and UCS were treated as separate target outputs. The predictive performance of the proposed ABC-ANN and GMDH-Combi models was evaluated and compared against results from a previously studied design of experiments (DOE) methodology. The ABC-ANN model demonstrated superior predictive accuracy compared to the DOE method, achieving coefficients of determination (R^2) values of 0.9948 and 0.9997 and mean absolute percentage error (MAPE) of 1.134 % and 0.319 % for FS and UCS, respectively, while the GMDH model yielded R^2 values of 0.9362 and 0.9629 and MAPE of 6.706 % and 6.150 %. Sensitivity and parametric analyses indicated that water content had the greatest influence on strength predictions, whereas cement content had the least, confirming that both models effectively captured the experimental behavior of the mortars.

1. Introduction

Valorization of secondary by-products and waste is an important challenge, which has been pursued by several sector. Following the guidelines of the sustainable development goals [1], the infrastructure sector has studied various methods that would minimize and reduce the use of raw materials and natural resources. A few strategies for reusing various waste kinds include recycling waste from construction and demolition projects, tire crumb rubber, waste glass, and fibers into different infrastructures. Quarry dust is a major by-product of the aggregate production process, which is mainly landfilled. It had been estimated that by 2019, about 50 billion tons of aggregates will be consumed annually worldwide [2]. Thus, the recycling of quarry dust could substantially decrease the use of the natural resources and provide sustainable solutions for the infrastructure sector. In general, natural aggregates are partially substituted with different types of quarry dust to produce various structures. For instance, the replacement of 60 % of fine aggregates with quarry dust improved the compressive strength of concrete building blocks [3]. The authors indicated that the incorporation of quarry waste was cost effective, efficient and sustainable which reduced the environmental impact by decreasing waste volume. Up to 40 % of marble sludge was replaced with cement binder to produce cement-bound composite. However, the study claimed that the best results were obtained when only 20 % of the waste sludge was used [4]. Different statistical models were used to define the optimum dosage of aggregates, waste and cement to water ratios of concrete paving blocks

* Corresponding author.
E-mail addresses: h.jahangir@birjand.ac.ir (H. Jahangir), danial.rezazadeh@mail.um.ac.ir (D. Rezazadeh Eidgahee), a.soleymani@birjand.ac.ir (A. Soleymani), cesare.sangiorgi4@unibo.it (C. Sangiorgi), piergiorg.tataranni2@unibo.it (P. Tataranni), as3466@cam.ac.uk (A. Solouki).

https://doi.org/10.1016/j.jmrt.2025.12.317
Received 10 August 2025; Received in revised form 21 December 2025; Accepted 30 December 2025
Available online 1 January 2026

[5]. The final mixture design replaced 25 % of the aggregate mass with fine quarry dust waste, which showed acceptable results in terms of mechanical properties.

The process of recycling waste aggregates such as designing stages, sample preparation, and data analysis/validation could become more efficient and accurate if proper statistical methods and models are applied [6]. Moreover, statistical models could maximize the resulting outcomes based on desired parameters. For instance, response surface methodology (RSM) was used to assess the mechanical durability of concrete. The optimized mixtures presented maximum service life, while having the least global warming potential [7]. A RSM was applied to examine the behavior of a pavement recycled aggregate concrete under fatigue and freezing conditions [8].

In recent years, machine learning (ML) techniques have garnered increasing attention among civil engineers as effective alternatives to conventional statistical methods [9]. These data-driven approaches offer significant advantages in terms of processing efficiency, especially when handling large datasets [10]. A comprehensive review by W. Ben Chaabene et al. [11] explored the application of ML in predicting various mechanical properties of concrete mixtures. Techniques such as Artificial Neural Networks (ANN), Decision Trees (DT), Support Vector Machines (SVM), and Linear Regression (LR) have been widely employed to estimate the compressive strength of concrete [12–14]. In these studies, input variables included parameters like water content, Schmidt hammer readings, ultrasonic pulse velocity, and relative humidity of the hardened concrete, while the compressive strength at 7 and 28 days served as the output. Among the models evaluated, the Decision Tree approach demonstrated the highest predictive accuracy, exhibiting the lowest error rate. Similarly, another study implemented eleven different ML algorithms to estimate the flexural and compressive strengths of steel fiber-reinforced concrete, using a dataset compiled from an extensive literature review [15]. The findings showed that predictions for compressive strength were generally more reliable than those for flexural strength, regardless of the algorithm used. In pursuit of sustainable concrete design, Naseri et al. [16] applied six distinct ML algorithms to estimate compressive strength, identifying the water cycle algorithm as the most accurate, with a minimum absolute error of 2.86 MPa.

One of the main concerns for civil engineers is to provide highly accurate prediction and estimation for mechanical and engineering properties of structures and infrastructures. Newly evolving cement-bound mixtures containing various waste and natural aggregates have further influenced and pushed researchers towards producing precise models for predicting the final strength of mixtures [17,18]. Several studies have benefitted from different statistical methods [19,20]. However, due to the complexity of some mixtures, the use of more prediction models could have less accuracy compared to actual results.

This paper investigates the practical characteristics of cement mortars incorporating waste silt by employing advanced machine learning (ML) techniques. The experimental data including 39 cement mortar specimens evaluated for unconfined compressive strength (UCS) and flexural strength (FS) were obtained in authors' previous study [21], and the current research introduces significant novel contributions. The study applies two optimized ML approaches, artificial neural networks enhanced with the artificial bee colony algorithm (ABC-ANN) and the combinatorial group method of data handling (GMDH-Combi), to predict FS and UCS. The GMDH method was selected due to its proven ability to model complex, nonlinear relationships without prior assumptions about the underlying functional forms, making it particularly suitable for datasets with limited sample sizes. The ABC-ANN model was chosen because it combines the adaptive learning capabilities of neural networks with the global optimization efficiency of the artificial bee colony algorithm, enhancing convergence and predictive accuracy. The predictive performance of these models is systematically compared against each other and against the previously reported design of experiments (DOE) method, providing new insights into their relative accuracy and practical applicability. In addition, sensitivity and parametric analyses are conducted to quantify the influence of input variables, including cement, water, sand, silt, and additives, on mortar strength, offering a deeper understanding of material behavior not addressed in the prior work. Fig. 1 presents a schematic overview of the study workflow.

The primary objectives of this study are.

- To develop predictive models for the FS and UCS of cement mortars incorporating waste silt using artificial neural networks optimized with the artificial bee colony algorithm (ABC-ANN) and the combinatorial group method of data handling (GMDH-Combi).
- To systematically compare the predictive performance of ABC-ANN and GMDH-Combi models with each other and with the previously reported DOE method.
- To conduct sensitivity and parametric analyses to quantify the influence of key input variables including, cement, water, sand, silt, and additives, on mortar strength.
- To provide insights into the practical applicability and interpretability of bio-inspired machine learning approaches for sustainable cementitious materials.

2. Research gap and novelty

Despite considerable progress in sustainable construction materials and data-driven prediction of cementitious composites, important knowledge gaps remain. While recent studies contribute to sustainability frameworks in broader infrastructure and resource management contexts [22], most existing work on cement and concrete focuses on conventional mix design or the use of recycled constituents without advanced predictive analytics. Investigations into waste incorporation and mechanical properties, such as the influence of nanomaterials and modified binders on composite performance [23], emphasize material enhancement but do not systematically address strength prediction. Moreover, while waste recycling literature and studies on modified asphalt or cement composites [24–26] and machine learning applications in lightweight or nanomaterial-modified concrete demonstrate the potential of data-driven approaches, there is a lack of targeted research on predicting mechanical strength of cement mortars incorporating reservoir waste silt using bio-inspired optimized ML models. Prior ML studies often employ conventional algorithms without leveraging global optimization capabilities, and closed-form interpretative models are rarely provided.

This study addresses the unresolved question of whether hybrid bio-inspired techniques, such as artificial bee colony–optimized artificial neural networks (ABC-ANN) and the combinatorial group method of data handling (GMDH-Combi), can provide accurate and interpretable predictions for the flexural and compressive strengths of mortars incorporating waste silt.

To fill this gap, the present research applies these advanced machine learning (ML) methods to an experimentally characterized dataset of cement mortars, benchmarks their performance against conventional design of experiments (DOE) results, and conducts sensitivity and parametric analyses to quantify the influence of key mix variables on strength outcomes. The work presents a pioneering case study using waste silt sourced from the reservoir lake of S.A.P.A.B.A. (S.p.A.) in Bologna, Italy, evaluating its potential as a sustainable additive in cement mortar formulations. Waste silt, typically discarded with minimal environmental or structural reuse, represents a novel contribution to construction materials research. While conventional studies rarely focus on such unconventional waste streams, this study demonstrates how ML techniques can effectively predict critical mechanical properties, namely flexural strength (FS) and unconfined compressive strength (UCS).

The research is distinguished by the integration of two advanced, optimized ML approaches, ABC-ANN and GMDH-Combi, with ABC-ANN

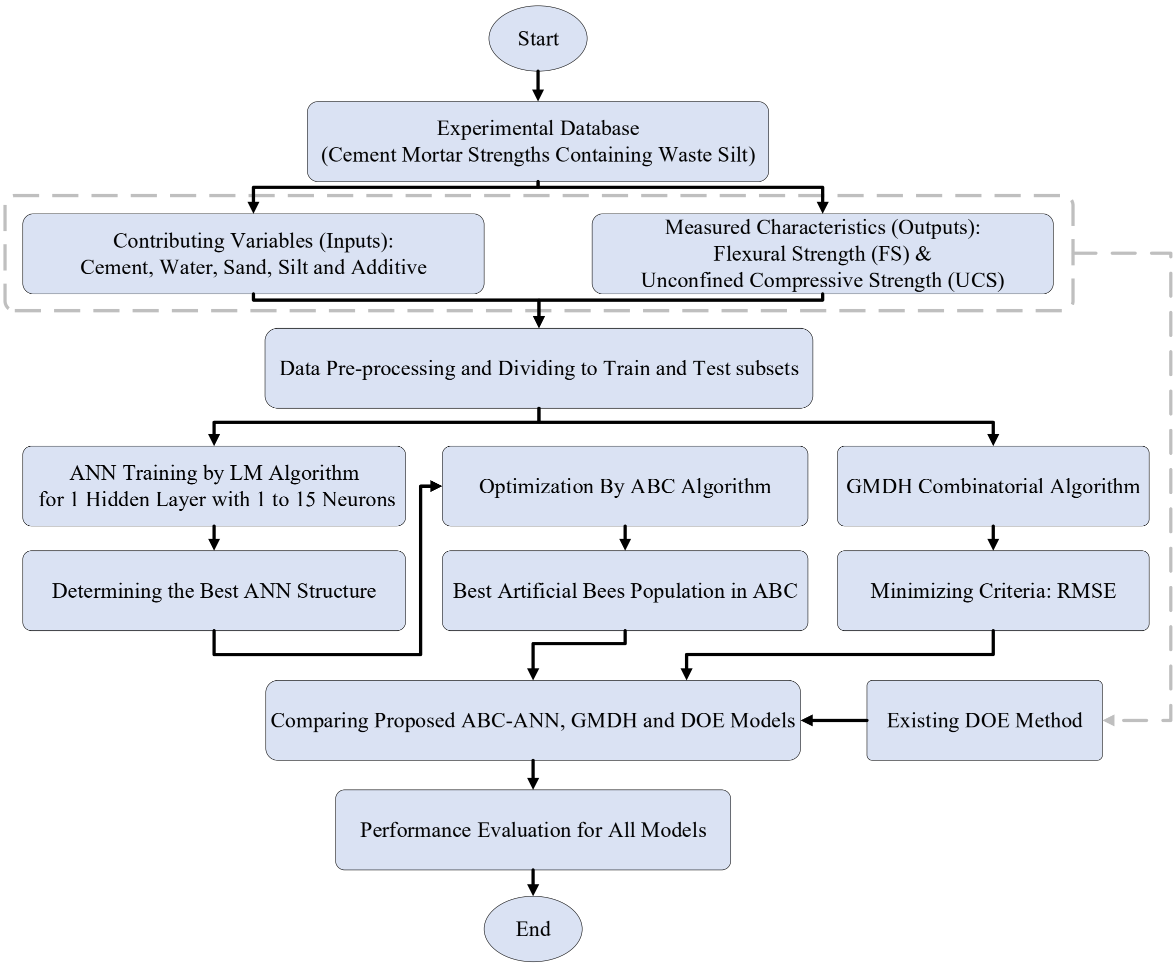

Fig. 1. The schematic overview of the study.

applied for the first time in this context. The ABC-ANN model achieves near-perfect predictive accuracy, outperforming both GMDH and DOE methods, while sensitivity and parametric analyses provide deeper insights into material behavior and model interpretability. Overall, the study offers a novel combination of sustainable material utilization, site-specific waste valorization, and bio-inspired predictive modeling, providing a foundation for eco-efficient construction practices in real-world applications.

3. Materials and sample preparation

The materials and methods described in this section summarize the work conducted previously by the authors, with a full description and characterization available in Ref. [21]. Sand and silt were provided by a local company in Bologna, Italy, while Type 42.5 R cement and a special additive were obtained locally. The additive was used to reduce the swelling behavior of the waste silt, enabling its incorporation into cement mortar mixtures. The silt was sourced from sedimentation lakes designated for landfill use by S.A.P.A.B.A. (Fig. 2). These lakes receive silt-laden water from limestone aggregate washing operations, which is then deposited for landfill. To prepare the raw silt for testing, the material was first extracted from the lakes and oven-dried. It was then sieved and ground into a fine powder using a Los Angeles abrasion machine. This preparation procedure was adapted from standard protocols commonly applied to fine industrial byproducts in cementitious materials, with minor modifications to account for the high moisture content and swelling behavior of the silt. The processed silt was then characterized using X-ray powder diffraction (XRD), as illustrated in Fig. 3.

Fig. 2. Silt collection and settling areas at the S.A.P.A.B.A. landfill site, Italy.

The diffraction patterns shown in Fig. 3 indicate that quartz constitutes approximately 32 % of the material, followed by 28 % calcite, 21 % illite/micas, and 5 % chlorite, with minor amounts of dolomite and feldspars. The sharp peaks visible in Fig. 3 confirm that the silt possesses a predominantly crystalline structure. Quartz, being mechanically stable and chemically inert, contributes to the rigidity of the mortar matrix,

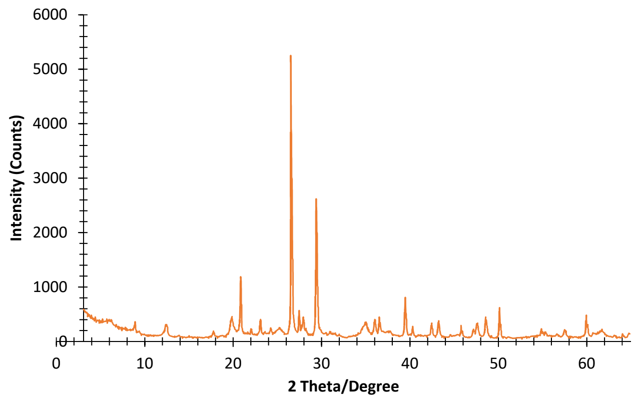

Fig. 3. XRD analysis of waste silt.

while calcite may participate in partial hydration reactions and influence the setting behavior of the cement mortar. Clay minerals such as illite/micas and chlorite can affect water retention and workability of the mixtures. Fig. 3 thus provides a clear visualization of the phases present in the silt and supports the discussion of its potential influence on the mechanical properties of the cement mortars.

The mineralogical composition observed in Fig. 3 has direct implications for the behavior of the waste silt when incorporated into cement mortar. Quartz, which constitutes approximately 32 % of the material, is largely inert and does not participate in cement hydration; instead, it contributes physically by improving particle packing and potentially reducing pore connectivity. Calcite, comprising about 28 %, may interact with the aluminate phases of cement to form carboaluminate phases, which can refine the microstructure and support early-age strength development. Illite and other clay minerals (approximately 21 %) can influence water retention and internal curing due to their surface activity. Given the moderate clay content, these minerals are not expected to negatively affect workability or induce detrimental swelling–shrinkage cycles. Overall, the mineralogical phases indicate that the silt primarily acts as a microfiller, enhancing densification of the mortar matrix without introducing components that could compromise long-term durability or lead to deleterious reactions.

As presented in Fig. 4, Cement mortar beams were produced by blending cement, water, aggregates, additive and waste silt according to a mixture design. The mixes were put into rectangular steel molds, which were left to cure for 28 days before being demolded after 24 h.

4. Machine learning techniques

4.1. Experimental database

The experimental database used in this work was taken from the author's earlier investigation, which evaluated cement mortars' flexural strength (FS) and unconfined compressive strength (UCS) using the design of experiments (DOE) approach [21]. To elaborate, the cement, water, sand, silt and additive contents were considered contributing variables (input) for the DOE, whereas the outputs were FS and UCS values. Statistical properties of variables are presented in Table 1.

4.2. ABC-ANN: Hybrid artificial bee colony - artificial neural networks

Artificial Neural Networks (ANN) are among the most commonly used machine learning algorithms for solving a wide range of engineering problems [27,28]. By utilizing a system of weighted connections and bias values, ANNs map input data to corresponding outputs. The primary objective of ANN training is to iteratively adjust these connection weights in order to minimize the discrepancy between predicted and actual outcomes. A typical ANN structure consists of an input layer, one or more hidden layers, and an output layer, which together form the basis of its computational architecture [29]. As soon as the network errors have been minimized during training, the weight effectiveness obtained from this model is examined in the test set.

Drawing inspiration from the collective foraging behavior of bees, the Artificial Bee Colony (ABC) algorithm serves as a well-performing optimization tool [30,31]. To solve diverse hybrid optimization issues, a colony of artificial bees is created by adopting their behavior during honey gathering. In order to find viable solutions, artificial bees scan the search space [32,33]. After a predetermined number of iterations, the algorithm searches for a new solution. Therefore, it is possible to make use of the ABC to compensate for defects in the feed-forward back propagation method and improve both weights of the ANN and the bias set. In this study, the ability of ABC algorithms is employed to enhance the ANN outcomes and identify the optimal values for FS and UCS parameters.

4.3. GMDH-Combi: Combinatorial group method of data handling algorithm

Drawing inspiration from the Kolmogorov-Gabor polynomial, Ivakhnenko [34] introduced the Group Method of Data Handling

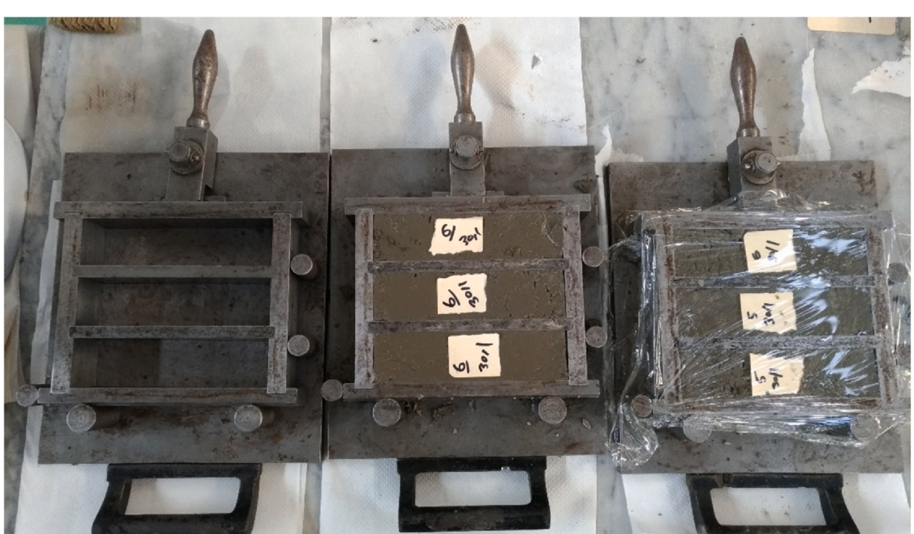

Fig. 4. Cement mortar sample preparation.

Table 1
Statistical features of experimental database.

Feature	Input Parameters					Output Parameters	
	Cement (%)	Water (%)	Sand (%)	Silt (%)	Additive (%)	FS (MPa)	UCS (MPa)
Minimum	22.00	12.00	43.00	2.00	0.17	0.00	0.00
Maximum	28.00	20.00	61.00	20.00	0.35	10.31	55.91
Range	6.00	8.00	18.00	18.00	0.18	10.31	55.91
Average	24.40	15.26	51.00	9.09	0.26	5.68	25.79
SD	2.28	2.94	6.29	6.20	0.07	2.27	11.27
CoV (%)	9.33	19.27	12.33	68.24	28.27	39.90	43.70
Kurtosis	−1.276	−1.144	−1.308	−1.291	−1.761	1.242	1.299
Skewness	0.508	0.557	0.262	0.404	−0.043	−0.849	−0.282

(GMDH), a novel algorithm recognized as one of the heuristic self-organizing techniques. As a result of this technique, Perceptron structures are more precisely defined and data may be categorized as beneficial or detrimental, reducing computation time and requiring less observations [35]. A parametric method known as a combinatorial algorithm is a self-organizing single-layer algorithm which examines all possible models [34]. In the combinatorial algorithm, incremental terms are utilized to specify models, and an external criterion was used to determine which solution was the best for a certain complexity model. Combinatorial GMDH algorithms were known for their simplicity and the ability to sort the whole model structure [34].

The GMDH-Combi algorithm inherently generates closed-form polynomial expressions, enhancing the interpretability of model predictions. This feature allows engineers to understand the relative contributions of each input variable to the predicted flexural strength (FS) and unconfined compressive strength (UCS). In this study, the GMDH-Combi model was used alongside the ABC-ANN to balance interpretability and predictive performance. While ABC-ANN delivered the highest accuracy in strength predictions, GMDH-Combi provided explicit polynomial equations that clearly relate input proportions to output properties. This dual-model strategy enables both precise prediction and practical interpretability, supporting informed engineering decisions regarding mixture design and the use of reservoir waste silt in cement mortars.

4.4. Multicollinearity considerations

Given the mixture-based nature of the input parameters, some degree of correlation among cement, water, sand, silt, and additive contents was expected. To assess this, pairwise correlation coefficients were computed for all input variables. No correlations exceeded commonly accepted multicollinearity thresholds, suggesting that the interdependencies among mixture components were moderate and did not pose a significant risk to model stability. Moreover, the modeling approaches adopted in this study, namely ABC-ANN and GMDH, are inherently resilient to moderate multicollinearity because they utilize non-linear learning processes rather than direct coefficient estimation. Furthermore, the use of train–test data partitioning confirmed that both models maintained strong predictive performance in the presence of these correlations, indicating that multicollinearity did not adversely affect model training or validation.

5. Results and discussion

5.1. Data scaling

As part of the construction of ABC-ANN and GMDH-Combi models, acquired data were standardized and scaled between 0.1 and 0.9 based on the minimum and maximum values as follow:

$$V_{scaled} = \left[0.80 \times \frac{V - V_{\min}}{V_{\max} - V_{\min}}\right] + 0.1 \tag{1}$$

In Equation (1), V indicates the included variable, and the $V_{\min}$ and $V_{\max}$ represent the variable's minimum and maximum values, respectively.

5.2. The hybrid ABC-ANN models

The artificial neural network (ANN) model was developed using input variables that included the contents of cement, water, sand, silt, and additives. The ANN was designed to predict two target properties of cement mortar incorporating waste silt: flexural strength (FS) and unconfined compressive strength (UCS). Following the approach recommended by Shahin et al. [36], the complete dataset was randomly shuffled, with 80 % allocated for training and the remaining 20 % reserved for testing to ensure well-performing model evaluation and generalization performance. The ANN models were constructed using the feed-forward backpropagation (FFBP) algorithm. The network architecture consisted of an input layer, a hidden layer utilizing a tangent sigmoid activation function ($f_n^{tan\text{-}sigmoid}$) as defined in Equation (2), and an output layer employing a linear activation function (f_n^{linear}), as shown in Equation (3).

$$f_n^{tan-sigmoid} = \frac{2}{1 + e^{-2x}} - 1 \tag{2}$$

$$f_n^{linear} = ax + b \tag{3}$$

A single hidden layer was selected to balance model complexity with the limited size of the experimental dataset (39 specimens). The number of neurons in the hidden layer was optimized through systematic trials, with performance metrics confirming strong predictive accuracy for both flexural and compressive strength. While deeper or alternative network structures could potentially improve generalization with larger datasets, the current architecture provides reliable and interpretable predictions without risk of overfitting.

The Levenberg-Marquardt method [37] were employed to train the network neurons' linking weights and biases in MATLAB software [38]. Trial-and-error approach was employed to determine the optimal artificial neural network configuration, including the optimized number of hidden neurons. The range explored for the number of neurons in the single layer were considered between 1 and 15. The correlation coefficient (R), Mean Square Error (MSE), and the Mean Absolute Percentage Error (MAPE) are standard statistical error and performance metrics that are used to select the optimal configuration. For every configuration, an assessment index was computed, and the responses' quality was taken into account while ranking the outcomes. A final evaluation of each pattern's summations of the given rankings was performed and the optimal network design was selected. Tables 2 and 3 present the performance and ranking of artificial neural network (ANN) models with different numbers of neurons in the hidden layer for predicting the flexural strength (FS) and unconfined compressive strength (UCS) of cement mortars incorporating waste silt. The rankings are based on multiple performance metrics, including the correlation coefficient (R), MSE, and MAPE for both training and testing datasets.

From the analysis of Tables 2 and 3, it is evident that networks with

Table 2
Performances and ranking of ANNs for FS estimation optimized model.

Values							Ranking						
Neuron No.	R		MSE		MAPE (%)		R		MSE		MAPE (%)		Sum
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	
1	0.598	0.425	0.021	0.013	41.186	19.460	13	12	13	10	14	13	75
2	0.927	0.895	0.005	0.003	15.394	11.186	2	1	2	1	4	3	13
3	0.655	0.599	0.018	0.009	40.101	15.587	11	10	9	9	13	8	60
4	0.920	0.704	0.005	0.007	13.952	16.253	3	9	3	5	3	10	33
5	0.660	0.294	0.019	0.024	35.613	27.609	10	13	10	14	10	15	72
6	**0.944**	**0.859**	**0.003**	**0.004**	**8.956**	**8.931**	**1**	**3**	**1**	**2**	**1**	**2**	**10**
7	0.816	0.830	0.010	0.004	25.873	11.595	6	4	5	3	7	4	29
8	0.705	0.721	0.016	0.008	37.087	15.522	9	8	7	7	11	7	49
9	0.507	0.765	0.029	0.008	52.656	15.259	14	6	15	6	15	6	62
10	0.758	0.774	0.016	0.013	29.248	18.685	7	5	8	11	8	12	51
11	0.615	0.194	0.021	0.014	38.732	17.690	12	14	12	12	12	11	73
12	0.828	0.739	0.011	0.009	16.096	13.420	5	7	6	8	6	5	37
13	0.729	0.883	0.019	0.005	13.361	5.138	8	2	11	4	2	1	28
14	0.858	0.493	0.010	0.014	15.809	16.200	4	11	4	13	5	9	46
15	0.435	−0.171	0.029	0.025	33.631	25.866	15	15	14	15	9	14	82

* The optimal quantity of neurons in the hidden layer is identified in the row highlighted.

Table 3
Performances and ranking of ANNs for UCS estimation optimized model.

Values							Ranking						
Neuron No.	R		MSE		MAPE (%)		R		MSE		MAPE (%)		Sum
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	
1	0.607	0.381	0.018	0.017	38.792	25.747	13	14	13	13	14	13	80
2	0.885	0.603	0.006	0.010	17.634	22.030	5	10	5	10	6	11	47
3	0.640	0.600	0.017	0.010	32.417	21.779	12	11	11	9	13	10	66
4	0.900	0.832	0.005	0.005	15.699	15.030	4	6	4	5	2	6	27
5	0.940	0.768	0.004	0.006	16.632	18.257	2	8	2	6	3	9	30
6	**0.957**	**0.952**	**0.002**	**0.001**	**12.809**	**5.501**	**1**	**1**	**1**	**1**	**1**	**1**	**6**
7	0.874	0.533	0.007	0.012	17.180	22.161	6	13	6	12	5	12	54
8	0.913	0.774	0.005	0.007	16.884	17.218	3	7	3	7	4	8	32
9	0.835	0.622	0.013	0.011	19.230	13.140	8	9	9	11	8	4	49
10	0.758	0.906	0.013	0.005	24.570	14.277	10	3	8	4	11	5	41
11	0.772	0.373	0.015	0.018	28.064	29.565	9	15	10	14	12	14	74
12	0.452	0.544	0.032	0.026	52.921	29.720	15	12	15	15	15	15	87
13	0.841	0.833	0.010	0.008	17.780	15.520	7	5	7	8	7	7	41
14	0.523	0.921	0.025	0.004	19.480	6.772	14	2	14	3	9	2	44
15	0.743	0.895	0.018	0.004	22.512	12.142	11	4	12	2	10	3	42

* The optimal quantity of neurons in the hidden layer is identified in the row highlighted.

six hidden neurons achieved the highest overall ranking for both FS and UCS predictions, indicating the optimal balance between model complexity and predictive performance. The optimized ANN for FS achieved correlation coefficients of 0.944 for training and 0.859 for testing, with normalized MSE values of 0.003 and 0.004, and MAPE values of 8.956 % and 8.931 %, respectively. For UCS, the ANN achieved correlation coefficients of 0.957 for training and 0.952 for testing, with normalized MSE values of 0.002 and 0.001, and MAPE values of 12.809 % and 5.501 %, respectively. These results demonstrate that the selected six-neuron ANN structure provides strong predictive accuracy for both mechanical properties, effectively capturing the relationship between input variables (cement, water, sand, silt, and additives) and mortar strength. Fig. 5 illustrates the optimized ANN architecture for FS and UCS estimation. Overall, the analysis confirms that careful tuning of hidden neurons is critical for maximizing model performance, and the selected ANN configuration offers reliable predictions for practical application in cement mortar design with waste silt incorporation.

In this study, the Artificial Bee Colony (ABC) optimization algorithm was employed to determine the optimal weight values for the ANN model by enhancing the search process within the solution space. To identify the best configuration, a trial-and-error approach was applied systematically. Initially, multiple ABC-ANN models were developed, each incorporating a different number of artificial bees. The internal

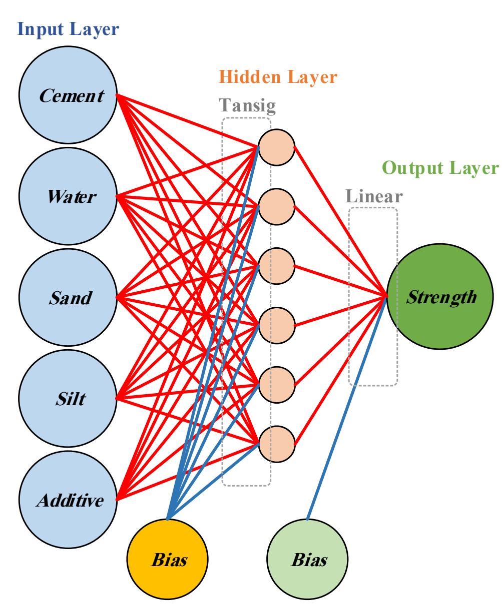

Fig. 5. The optimal structure of FS and UCS ANN models.

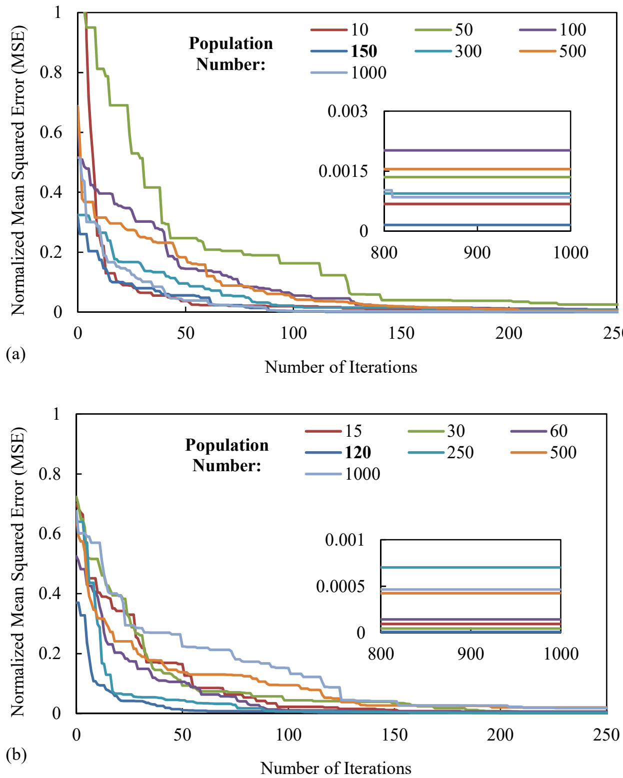

Fig. 6. Model performance evolution with increasing iterations and population: a) FS and b) UCS.

parameters of the ABC algorithm, including the number of iterations, limit values, and employed/observer/onlooker bee ratios, were incrementally adjusted to evaluate their effect on the model's predictive accuracy. For each configuration, the model's normalized mean squared error (MSE) was monitored over 1000 iterations. The bee population was gradually increased from 10 to 1000, and the configuration that yielded the lowest MSE was selected as the optimal model. This trial-and-error procedure ensured that the ABC-ANN model converged effectively to the global minimum, providing the most accurate predictions for flexural strength (FS) and unconfined compressive strength (UCS). Fig. 6 illustrates the evolution of normalized MSE during 1000 iterations for different bee populations, highlighting the impact of population size on convergence and performance.

A normalized MSE drop was seen in the FS and UCS models up to about 150 iterations, as shown in Fig. 6. As found via trial and error, the ABC-ANN network only requires 250 cycles to perform effectively, indicating that the 1000 cycles used for the optimization phase is excessive. Although the iteration limit was set to 1000, Fig. 6 shows that the normalized MSE for FS and UCS models dropped sharply within the first 150–250 iterations, indicating that convergence was effectively achieved much earlier. The extended iteration count served as a conservative safety margin to ensure global convergence. The process for choosing the best configuration is comparable to the detailed process for figuring out how many neurons are in the hidden layer given in the preceding section. Thus, the optimal number of bees was also determined by ranking the values of R, MSE, and MAPE. According to Tables 4 and 5, the number of optimum bees is 150 for the FS parameter and 120 for the UCS parameter, based on the overall rank value.

As shown in Tables 4 and 5, selecting a bee population of 150 for the training and testing phases in the flexural strength (FS) estimation yielded high performance, with correlation coefficients (R) of 0.9995 and 0.9738, normalized mean squared error (MSE) values of 0.0061 and 0.0240, and mean absolute percentage errors (MAPE) of 0.9872 % and 2.4460 %, respectively. Similarly, for unconfined compressive strength (UCS) estimation, using 120 bees in both training and testing phases resulted in R values of 0.9999 and 0.9995, normalized MSEs of 0.0023 and 0.0037, and MAPE values of 0.2496 % and 0.4196 %, respectively. Compared to other configurations, these settings achieved the lowest overall ranking values, indicating superior model performance.

By applying the weights and biases derived from the proposed ABC-ANN models, the mathematical relationships between the input parameters (cement, water, sand, silt, and additive content) and the output variables, flexural strength (FS) and unconfined compressive strength (UCS), can be expressed as follows:

$$CementMortarStrength = f_n^{linear}\left\{ b_0 + \sum_{k=1}^{h}\left[W_k f_n^{tan-sigmoid}\left(b_{hk} + \sum_{i=1}^{m} W_{ik} X_i \right)\right]\right\} \tag{4}$$

In Equation (4), the $f_n^{tan\text{-}sigmoid}$ and f_n^{linear} represent the activation functions defined in Equations (2) and (3), respectively. The variable h denotes the number of neurons in the hidden layer, which is set to 6 in this paper. X_i refers to the input variables of the network (cement, water, sand, silt, and additive), while m, representing the number of input features, is equal to 5. W_{ik} indicates the weight connecting the ith input to the kth neuron in the hidden layer, and W_k corresponds to the weight linking the kth hidden neuron to the output layer. b_{hk} is the bias term associated with the kth hidden neuron, and b_0 denotes the bias applied at the output layer. For the purpose of promoting transparency, reproducibility, and broader application, the weights and biases used in the ABC-ANN models developed in this study are provided to support future research efforts. Based on Equation (4) and the acquired weights and biases, the estimated FS and UCS parameters using ABC-ANN models could be stated as follows:

$$FS_{ABC-ANN} = 87.06A_1 - 87.14A_2 - 4.30A_3 - 4.04A_4 + 59.35A_5 + 4.10A_6 - 63.00 \tag{5}$$

Table 4
Selected artificial bees population for ABC-ANN model of FS estimation.

Values							Ranking						
Population	R		MSE		MAPE (%)		R		MSE		MAPE (%)		Sum
	TR	TS	TR	TS	TR	TS	TR	TS	TR	TS	TR	TS	
10	0.9894	0.5820	0.0278	0.1359	7.1954	11.3235	2	7	2	7	7	7	32
50	0.9790	0.9669	0.0390	0.0284	5.6465	4.9349	5	5	5	5	4	5	29
100	0.9689	0.9482	0.0474	0.0356	6.6554	6.5660	7	6	7	6	6	6	38
150	**0.9995**	**0.9738**	**0.0061**	**0.0240**	**0.9872**	**2.4460**	**1**	**3**	**1**	**3**	**1**	**1**	**10**
300	0.9851	0.9800	0.0329	0.0220	4.2342	2.8111	4	1	4	1	3	2	15
500	0.9741	0.9785	0.0432	0.0224	6.6064	3.4079	6	2	6	2	5	3	24
1000	0.9881	0.9717	0.0296	0.0276	3.6315	4.1699	3	4	3	4	2	4	20

* The highlighted row shows the selected population in the ABC-ANN model.
Note: TR denotes Training Data Division, and TS denotes Testing Data Division.

$$UCS_{ABC-ANN} = 18.54B_1 - 18.59B_2 - 167.17B_3 + 0.14B_4 + 0.15B_5 + 0.35B_6 - 166.62 \tag{6}$$

In this case, the response coefficients of hidden neurons, A_1 to A_6 and B_1 to B_6, are determined by applying the following equations:

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \end{bmatrix} = Tansig\left(\left(\begin{bmatrix} -8.72 & -4.77 & -24.47 & -11.42 & 6.59 \\ 17.77 & 30.30 & 55.04 & 67.95 & 7.22 \\ -52.99 & -62.56 & -159.97 & -174.39 & 2.54 \\ -28.25 & -51.63 & -83.45 & -85.81 & -28.84 \\ 600.65 & 803.53 & 1806.36 & 1798.49 & 17.05 \\ 20.26 & -19.25 & 29.56 & -8.02 & -14.65 \end{bmatrix}\begin{bmatrix} Cement \\ Water \\ Sand \\ Silt \\ Additive \end{bmatrix}\right) + \begin{bmatrix} 22.31 \\ -73.29 \\ 188.17 \\ 130.24 \\ -2165.54 \\ 21.35 \end{bmatrix}\right) \tag{7}$$

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \end{bmatrix} = Tansig\left(\left(\begin{bmatrix} 133.13 & 189.24 & 419.76 & 431.36 & 26.70 \\ 123.99 & 177.23 & 392.50 & 403.95 & 19.43 \\ -254.17 & -336.04 & -763.51 & -765.79 & -7.10 \\ -150.26 & 6.90 & -51.95 & 149.40 & -33.99 \\ 336.49 & 443.37 & 1001.87 & 1000.99 & 12.66 \\ -16.25 & -21.73 & -48.88 & -50.45 & -0.76 \end{bmatrix}\begin{bmatrix} Cement \\ Water \\ Sand \\ Silt \\ Additive \end{bmatrix}\right) + \begin{bmatrix} -514.31 \\ -479.35 \\ 912.98 \\ 46.85 \\ -1208.89 \\ 59.93 \end{bmatrix}\right) \tag{8}$$

5.3. The GMDH-Combi models

To make closed-form equations more trustworthy, the GMDH combinatorial technique, like ANN models, randomly selects 20 % of the database to be utilized as a test set in the algorithm and not for model training and optimization. It was found that Equations (9) and (10) can be used to estimate FS and UCS parameters using the GMDH combinatorial method after a great deal of testing.

$$FS_{GMDH} = \frac{1}{Wa \cdot Sa \cdot Ad}\Big(Wa \cdot Sa(-0.013Ce - 0.953Wa + 0.300Sa) + Sa \cdot Ad^2(1273.44 + 4.73Wa \cdot Sa - 16.43Wa^2) + Ad(-90043.90 + Sa(-98913.50 + 1027.41Ce + 963.50Wa + 1023.94Si) + 979.89Sa^2 + 309.72Wa^2)\Big) \tag{9}$$

$$UCS_{GMDH} = \frac{1}{Wa^2 \cdot Sa^2 \cdot Si^2 \cdot Ad^2}\Big(-9.72Ce^2 \cdot Sa \cdot Si^2 + 5896.29Ce^2 \cdot Sa \cdot Si^2 \cdot Ad^4 - 40468.00Ce^2 \cdot Sa \cdot Si^2 \cdot Ad^5 + Ce^2 \cdot Sa \cdot Ad^3(334818.60 + 4135.73Si^2) + Ce^2 \cdot Ad(114.29Si^3 + Sa(17618.60 - 94.79Si^2) + 0.00113Ce \cdot Wa \cdot Sa^3 \cdot Si) + Ad^2(Ce^2(-171600.00Sa + Si^2(-23024.80 + 532.64Si)) + 19.75Wa^2 \cdot Sa^2 \cdot Si^2 - 0.00094Ce \cdot Wa^5 \cdot Sa^2)\Big) \tag{10}$$

Where Ce, Wa, Sa, Si and Ad are the representative of cement (%), water (%), sand (%), silt (%) and additive (%) contents, respectively.

5.4. Performance of proposed models

Table 6represents frequently employed parameters in assessing error and performance [39] based on Equations (11)–(17) for R (Correlation Coefficient), R^2 (Coefficient of Determination), MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error) and NMAE (Normalized Mean Absolute Error). These error performance and model criteria were employed to assess the outcomes of the improved ABC-ANN and GMDH models for FS and UCS estimation.

$$R = \frac{\sum_{i=1}^{n}(OS_i - OS_{ave.})(ES_i - ES_{ave.})}{\sqrt{\sum_{i=1}^{n}(OS_i - OS_{ave.})^2 \sum_{i=1}^{n}(ES_i - ES_{ave.})^2}} \tag{11}$$

Table 5
Selected artificial bees population for ABC-ANN model of UCS estimation.

Values							Ranking						
Population	R		MSE		MAPE (%)		R		MSE		MAPE (%)		Sum
	TR	TS	TR	TS	TR	TS	TR	TS	TR	TS	TR	TS	
15	0.9981	0.9989	0.0108	0.0049	1.0170	0.7043	3	2	3	2	3	2	15
30	0.9995	0.9958	0.0058	0.0092	0.8635	1.3994	2	3	2	3	2	4	16
60	0.9976	0.9934	0.0120	0.0119	1.7286	1.3589	4	5	4	4	4	3	24
120	**0.9999**	**0.9995**	**0.0023**	**0.0037**	**0.2496**	**0.4196**	**1**	**1**	**1**	**1**	**1**	**1**	**6**
250	0.9860	0.9882	0.0291	0.0151	3.5255	2.0103	7	6	7	6	7	5	38
500	0.9923	0.9854	0.0216	0.0168	2.5798	2.3425	5	7	5	7	5	7	36
1000	0.9908	0.9955	0.0237	0.0120	3.0805	2.1775	6	4	6	5	6	6	33

* The highlighted row shows the selected population in the ABC-ANN model.
Note: TR denotes Training Data Division, and TS denotes Testing Data Division.

$$R^2 = \left(\frac{\sum_{i=1}^{n} (OS_i - OS_{ave.})(ES_i - ES_{ave.})}{\sqrt{\sum_{i=1}^{n} (OS_i - OS_{ave.})^2 \sum_{i=1}^{n} (ES_i - ES_{ave.})^2}} \right)^2 \quad (12)$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (OS_i - ES_i)^2 \quad (13)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (OS_i - ES_i)^2} \quad (14)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |OS_i - ES_i| \quad (15)$$

$$MAPE = \frac{1}{n} \left[\frac{\sum_{i=1}^{n} |OS_i - ES_i|}{\sum_{i=1}^{n} |OS_i|} \right] \times 100 \quad (16)$$

$$NMAE = \left[\frac{\frac{1}{n} \sum_{i=1}^{n} |OS_i - ES_i|}{\max(OS_i) - \min(OS_i)} \right] \times 100 \quad (17)$$

In Equation (11) through (17), OS_i denotes the observed values, while ES_i represents the estimated values for the FS and UCS parameters. The variable n refers to the total number of data points considered. $OS_{ave.}$ and ES_{ave} indicate the mean of the observed and estimated values, respectively.

As presented in Table 6, the ABC-ANN models outperformed both the proposed GMDH and previously studied DOE models in estimating both FS and UCS parameters. Specifically, the ABC-ANN models achieved coefficients of determination (R^2) of 0.9948 for FS and 0.9997 for UCS, which are higher than those obtained from the GMDH models (0.9362 and 0.9629, respectively) and the DOE models (0.9115 and 0.9552, respectively). Additionally, the ABC-ANN models exhibited lower error values across all evaluated metrics, including MSE, RMSE, MAE, MAPE, and NMAE, compared to the GMDH and DOE models. Nonetheless, the GMDH models still performed better than the DOE models in terms of accuracy.

To further evaluate the robustness and generalization capability of the proposed models, a k-fold cross-validation (k = 5) was conducted, and the resulting MSE values were compared with those obtained from the final ABC-ANN and GMDH models. The results of the 5-fold cross-validation are presented in Fig. 7. For FS, the average MSE obtained from k-fold cross-validation is 0.0256 for the ABC-ANN model and 0.0338 for the GMDH model. These values are in close agreement with the MSE values reported for the final ML models in Table 6 (0.027 for ABC-ANN and 0.333 for GMDH), indicating stable predictive performance across different data partitions. Similarly, for UCS, the average MSE values from 5-fold cross-validation are 0.3173 for the ABC-ANN model and 4.8032 for the GMDH model, which closely match the corresponding MSE values of 0.035 and 4.749 reported for the final models. The strong agreement between the cross-validated MSE values shown in Fig. 7 and the final model errors demonstrates that the predictive performance of both models is not sensitive to data ordering or the use of a single training–testing split. Overall, the k-fold cross-validation analysis confirms the robustness of the proposed modeling framework and further highlights the superior generalization performance of the ABC-ANN model compared with the GMDH approach for predicting both FS and UCS.

Fig. 8 provides a graphical comparison of the predicted versus experimentally measured FS and UCS values for the ABC-ANN, GMDH, and DOE models. In the figure, the yellow line represents the ideal 1:1 correspondence line, which reflects perfect agreement between predicted and observed values and is theoretically achieved when the correlation coefficient (R) equals 1.0 [29]. It can be seen from Fig. 8 that the predicted values for both FS and UCS parameters obtained from GMDH models are closer to the line of equality than DOE estimates. Although the results of all DOE, GMDH and ABC-ANN models are reliable, the hybrid ABC-ANN models outperformed both GMDH and DOE methods, resulting in the most reliable estimates.

Fig. 9 shows the true error values for each technique providing a thorough comparison between the residual error of each observation. For both FS and UCS estimation, the outcomes of ABC-ANN models are positioned extremely close to the red line representing the zero-error value.

Fig. 10 show the FS and UCS estimation error distribution diagrams

Table 6
Summary of FS and UCS model results using conventional performance indicators.

Parameter	Method	R	R^2	MSE	RMSE	MAE	MAPE (%)	NAME (%)
FS	DOE	0.9547	0.9115	0.455	0.674	0.576	8.852	5.591
	GMDH	0.9676	0.9362	0.333	0.577	0.440	6.706	4.142
	ABC- ANN	0.9974	0.9948	0.027	0.164	0.075	1.134	0.727
UCS	DOE	0.9773	0.9552	5.692	2.386	1.987	7.307	3.554
	GMDH	0.9813	0.9629	4.749	2.179	1.697	6.150	3.067
	ABC- ANN	0.9999	0.9997	0.035	0.187	0.081	0.319	0.145

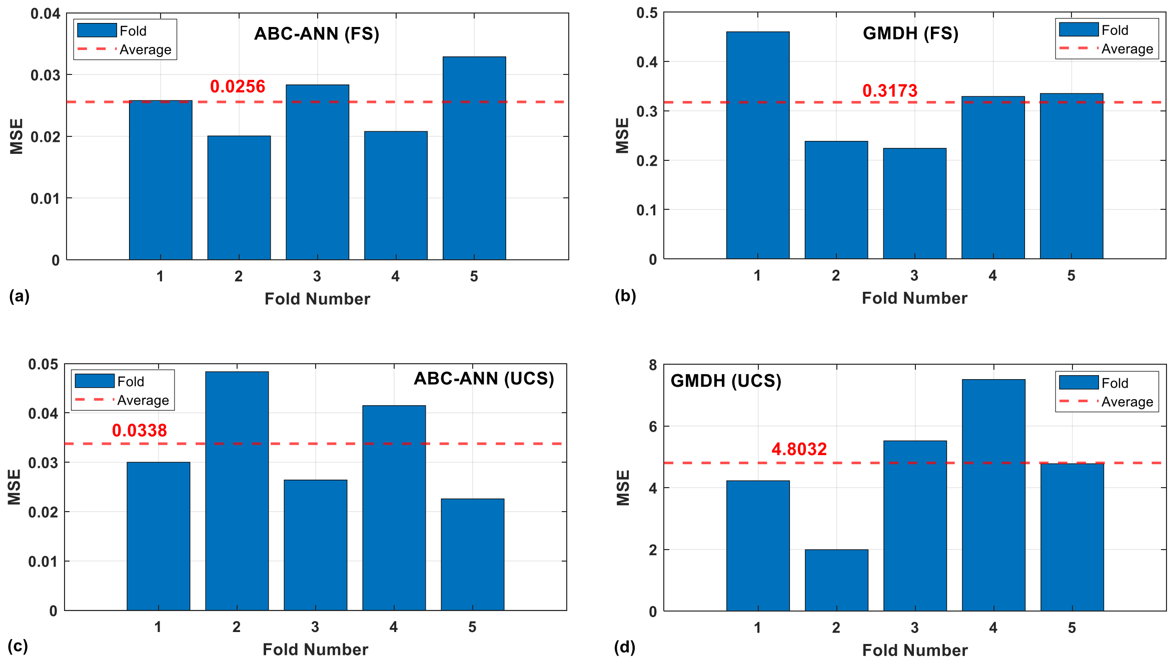

Fig. 7. MSE results of 5-fold cross-validation of the proposed ABC-ANN and GMDH models respectively for: (a and b) FS; and (c and d) UCS.

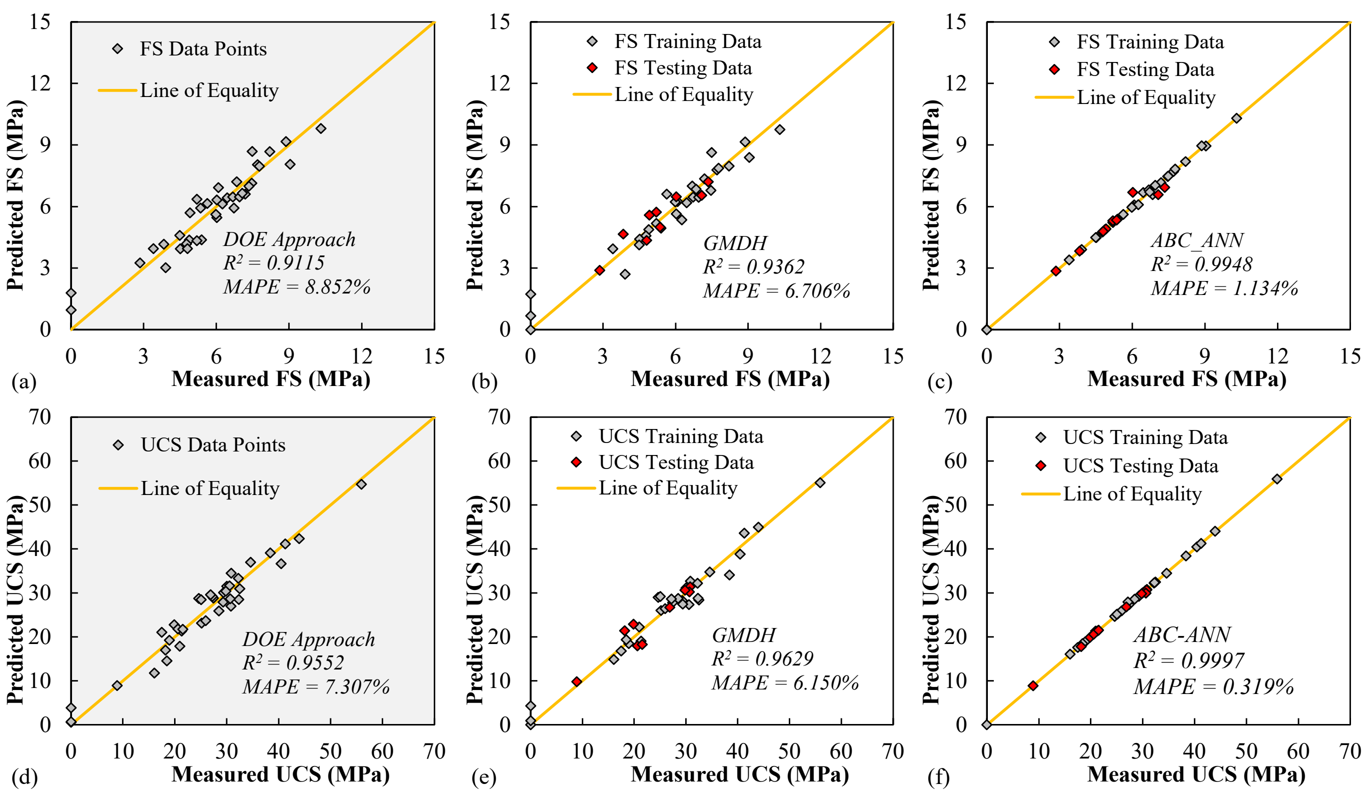

Fig. 8. Measured vs. predicted FS and UCS using different models: (a and d) DOE; (b and e) GMDH and (c and f) ABC-ANN.

in ABC-ANN, GMDH, and DOE models, respectively. As it could be inferred from Fig. 10, the proposed ABC-ANN models outperform both GMDH and DOE models by an error range between −0.6 and 0.6 MPa for FS estimation and between −0.6 and 1.2 MPa for UCS estimation.

As a result, it can be stated that the ABC-ANN models outperform other models and have the best accuracy and consistency. In spite of the fact that ABC-ANN models are in decent condition than GMDH models, the GMDH models have the advantage of giving a closed-form representation that is easy to understand.

5.5. Sensitivity and parametric analyses

Although models calibrated with high-performing machine learning techniques can yield accurate results, this does not necessarily guarantee

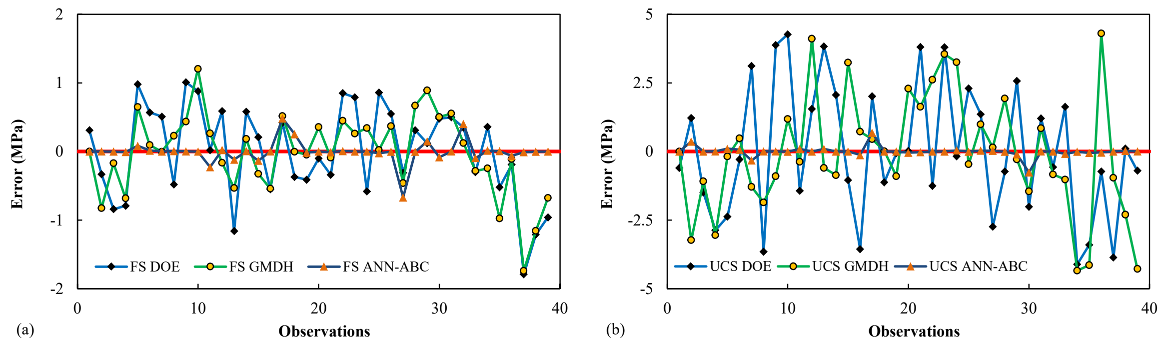

Fig. 9. Error-values of proposed DOE, GMDH and ABC-ANN models for estimating: a) FS and b) UCS.

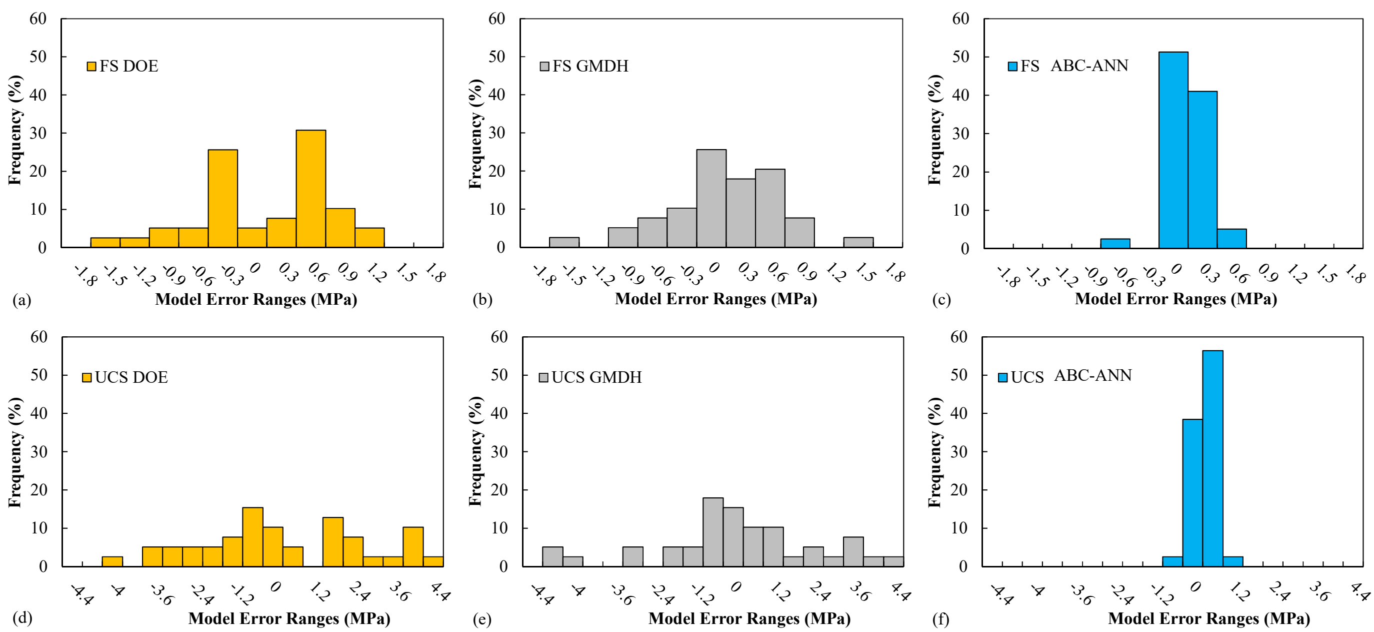

Fig. 10. Error distribution diagrams of FS and UCS in numerical models: (a and d) DOE; (b and e) GMDH and (c and f) ABC-ANN.

reliable performance across a wide range of input variables. The robustness of these models should therefore be evaluated by assessing how well their predictions align with established physical principles under varying conditions. In this paper, Figs. 11 and 12 illustrate the sensitivity analysis conducted to examine the generalization capability of the proposed ABC-ANN and GMDH models. This analysis involved systematically varying one input variable at a time (cement, water, sand, silt, or additive) while holding all other inputs constant at their mean training values. Artificial datasets were generated by incrementally increasing the selected variable across its observed minimum and maximum range. These synthetic inputs were then processed through the ABC-ANN and GMDH models to obtain the corresponding predicted outputs for flexural strength (FS) and unconfined compressive strength (UCS).

It appears that the predicted FS and UCS responses in both models (ABC-ANN and GMDH) are identical to the experimental data, as shown in Figs. 11 and 12. The efficiency of ABC-ANN and GMDH models can generally be validated as the outcomes of experiments and models for both FS and UCS along with different inputs are similar. It also can be deduced that models spanned the range of input variable changes, corresponding to experimental and physical expected behavior. Therefore, the generality and robustness of models can be derived from sensitivity analyses that have been conducted.

Silt and e) Additive.

After ensuring the generality and robustness of suggested models, as illustrated in Fig. 13, a parametric analysis was conducted to examine the influence of each input parameter (Cement, Water, Sand, Silt and Additive) on ABC-ANN and GMDH models of FS and UCS outputs. The approach results are derived for each input variable by considering the minimum and maximum interval values. As a result, the influence of each variable on R and RMSE is assessed. As presented in Equation (18), if all the parameters are replaced with their mean values, the impact on R and RMSE is zero and if all them are replaced, the impact would be 100 % [40].

$$\text{Input Impact} = \left(\frac{U_x - U_{none}}{U_{all} - U_{none}}\right) \times 100\% \tag{18}$$

In Equation (18), U_x denotes the correlation coefficient (R) and root mean squared error (RMSE) values associated with a specific input parameter, while U_{none} represents the impact of excluding that parameter (i.e., setting it to zero) on the R and RMSE metrics. Meanwhile, U_{all} corresponds to the R and RMSE values obtained when all input variables are fixed at their mean values. Fig. 13 shows that water exerts the most significant influence on both R and RMSE across the ABC-ANN and GMDH models for FS and UCS predictions, followed by Sand, Additive, and Silt, while Cement has the least impact. The dominant effect of water can be attributed to its role in controlling the water-to-cement (w/

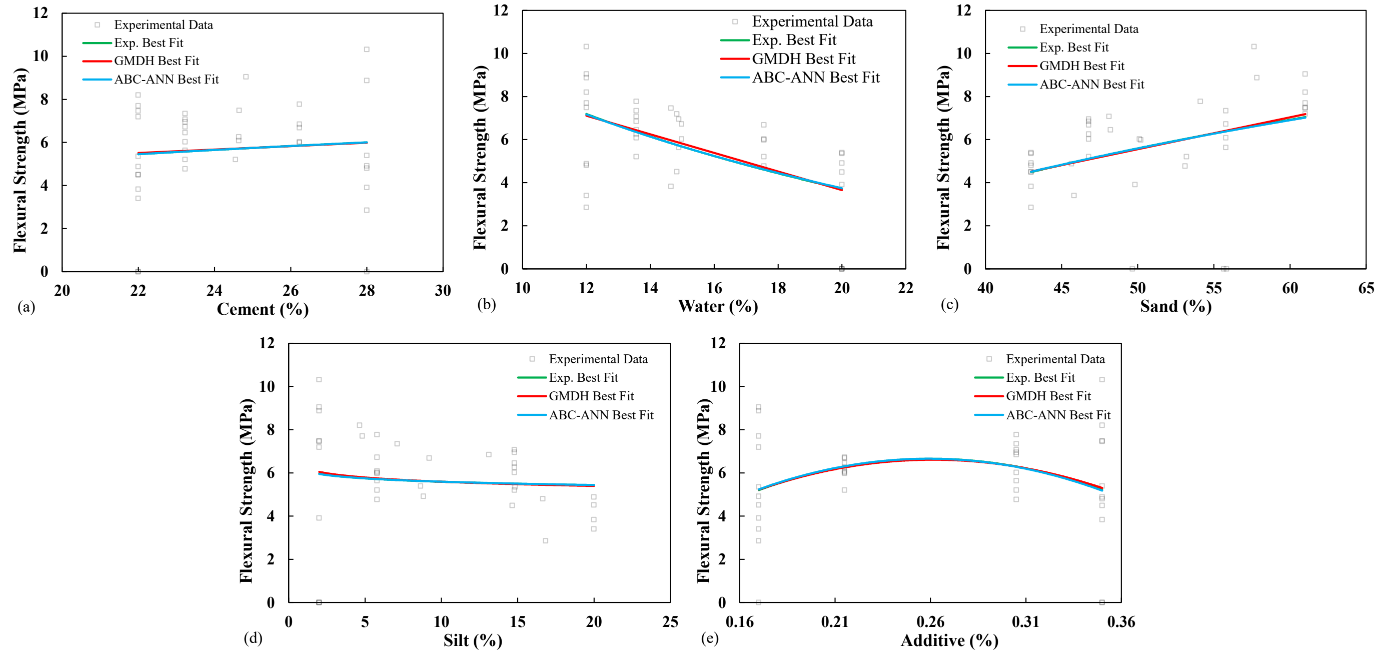

Fig. 11. The FS GMDH and ABC-ANN models sensitivity analysis considering a) Cement; b) Water; c) Sand; d).

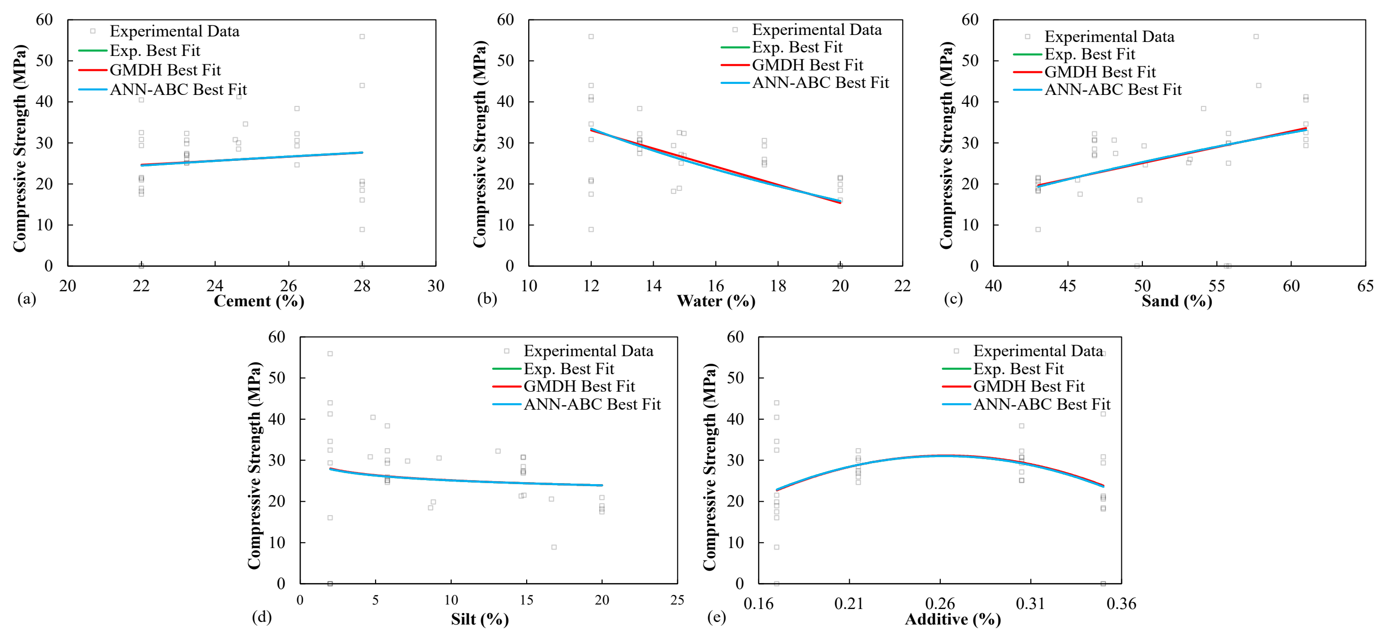

Fig. 12. The UCS GMDH and ABC-ANN models sensitivity analysis considering a) Cement; b) Water; c) Sand; d) Silt and e) Additive.

c) ratio, which governs hydration, microstructural development, and ultimately the mechanical strength of cement mortars. In contrast, variations in cement content within the studied range produce relatively minor changes in w/c ratio, explaining the comparatively smaller influence. Silt influences particle packing and microstructure, sand provides granular skeleton support, and additives modify workability or mitigate swelling effects. All input parameters were found to meaningfully affect model predictions, but the theoretical significance of water content is consistent with well-established cement chemistry principles.

5.6. Practical implications

The experimental results and predictive models presented in this study demonstrate the feasibility of using reservoir waste silt in cement mortars. At a practical scale, silt can be incorporated provided that preprocessing steps, including drying, sieving, and additive treatment to control swelling, are applied. The ABC-ANN and GMDH models can guide practitioners in optimizing mix designs, ensuring that mechanical performance targets for flexural and compressive strength are met.

From an environmental and durability perspective, the processed silt showed low levels of potential contaminants, and the cement mortar matrix is expected to immobilize these substances effectively. Nevertheless, for large-scale implementation, it is recommended to conduct site-specific chemical analysis of silt, long-term durability testing, and leachate assessments to verify safety and regulatory compliance. Overall, these findings support the sustainable use of waste silt in

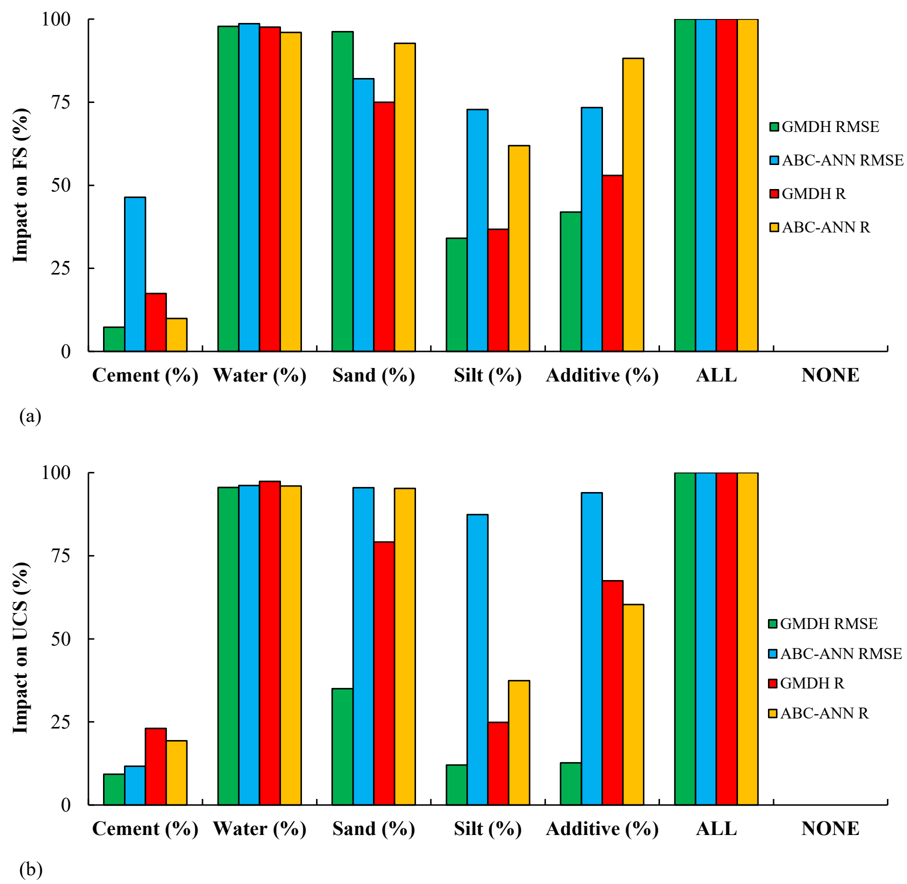

Fig. 13. The impact of the input parameters on ABC-ANN and GMDH models for different cement mortar strengths: a) FS and b) UCS.

construction, promoting circular economy practices while maintaining structural integrity and environmental protection.

The predictive models presented in this study were developed using laboratory mortars under strictly controlled conditions. While this ensures reliable and consistent predictions, industrial-scale applications involve greater variability in materials, mixing, and curing environments, which may influence model performance. To enhance the applicability of the models in practical settings, future research should incorporate field-scale data that capture the inherent variability of raw materials and environmental conditions. Including additional input parameters such as ambient temperature, humidity, and curing time can improve model generalization. Furthermore, quantifying prediction uncertainty through confidence intervals can help practitioners understand the potential range of outcomes. Implementing basic quality control protocols during mixing and curing will also reduce the impact of uncontrolled variability, allowing the models to provide more reliable guidance for the design and evaluation of mortars containing reservoir waste silt in real-world construction projects.

5.7. Limitations and future research directions

Despite the promising results obtained in this study, several limitations should be noted, and avenues for future research are suggested to further enhance the predictive modeling of cement mortars incorporating reservoir waste silt as listed below.

- The experimental database used for model development consisted of 39 mortar specimens, representing the complete set of samples produced in the authors' prior experimental study. Although the sample size is modest, several measures were employed to reduce the risk of overfitting, including the use of independent training–testing subsets, objective model selection criteria, and validation through sensitivity and parametric analyses. The comparable performance observed between training and testing phases confirms that the selected ABC-ANN and GMDH models were not adversely affected by sample size limitations. Nonetheless, a larger dataset would further enhance the generalizability of future models. Expanding the database was not feasible in the present work because additional material production and laboratory testing were beyond the scope of the completed experimental program, and no external datasets are available for the specific waste silt source. Future research should therefore focus on generating broader experimental datasets to support model extension and cross-validation under a wider range of material and mixture conditions.
- The experimental database employed in this study was derived from a prior laboratory program in which each mortar composition was produced as a single batch. While this approach ensured consistency in materials and procedures, it does not capture within-mix variability that could arise from batching, mixing, and testing uncertainties. As a result, performance indicators such as R^2 may be optimistic. Future experimental studies should incorporate multiple replicate batches for each mix design to quantify inherent variability, improve statistical representativeness, and provide more rigorous datasets for training and validating machine learning models.
- The present study is limited to 28-day strength data, as the experimental database was derived from a previously conducted testing program. While 28-day strength is a standard reference for mortar performance evaluation, incorporating additional curing ages (e.g., 7, 14, and 90 days) in future experimental campaigns would allow the development of time-dependent predictive models capable of capturing strength-gain kinetics. Including curing age as an explicit input variable would further enhance the applicability of the

proposed machine learning framework for construction planning and long-term performance assessment.

- Parametric analysis showed that, variations in cement had the least impact on the predicted flexural and compressive strengths. This outcome can be attributed to the limited range of cement percentages, which caused only minor changes in the water-to-cement ratio and, consequently, minimal influence on the hydration process and strength development. It is acknowledged that extending the range of cement content in future studies could alter this trend, potentially increasing its relative contribution to strength outcomes. This consideration is noted as a limitation of the current study and a recommendation for subsequent research efforts.
- The current predictive models were developed based on laboratory-cured mortar specimens under controlled temperature and humidity. While this ensures consistent and reliable data, it does not capture the variability of field curing conditions. Future work could extend the models by including environmental parameters, such as curing temperature, relative humidity, and exposure conditions, as additional input variables. Incorporating these factors would enhance the models' capability to predict flexural and compressive strengths under practical construction scenarios and support more robust design and quality control of mortars containing reservoir waste silt.
- The predictive models presented in this study were developed using laboratory-scale mortar specimens, and their current applicability is limited to similar small-scale mixtures. Scaling to full-size structural concrete would require consideration of additional factors, such as coarse aggregates, heterogeneous microstructures, variable curing conditions, and environmental influences. Incorporating these parameters into the training datasets and retraining the ABC-ANN and GMDH models would enable reliable predictions for full-scale concrete. The present work therefore establishes a proof-of-concept framework for machine learning-based prediction of cementitious materials containing reservoir waste silt, which can be extended to structural applications through further experimentation and dataset expansion.
- The present study focuses on mortars incorporating waste silt as a partial replacement for cement. Introducing supplementary cementitious materials (SCMs) such as fly ash or slag would modify the hydration process, microstructure, and mechanical performance, potentially affecting both predictive accuracy and sustainability outcomes. The current ABC-ANN and GMDH models would require retraining with a dataset that includes SCM proportions to capture these new interactions accurately. Nevertheless, combining silt with SCMs could further enhance the environmental benefits of cementitious mixtures, supporting circular economy principles and improved eco-efficiency in construction applications.
- While this study emphasizes the sustainable reuse of reservoir waste silt in cement mortars, no formal life cycle assessment (LCA) or embodied carbon analysis was performed. The primary focus was on predicting mechanical properties using machine learning models. Nevertheless, incorporating silt as a partial cement replacement reduces cement demand, which directly lowers CO_2 emissions associated with cement production. Future research could include a preliminary LCA or embodied carbon evaluation to quantify the environmental benefits of using waste silt, supporting broader sustainability claims and facilitating adoption in eco-efficient construction practices.
- The present models were developed using reservoir waste silt from a single location in Bologna, Italy, with specific mineralogical and particle size characteristics. Because silt properties can vary with reservoir location and seasonal factors, the current predictive models may not be directly applicable to silts from other sources. Extending the models to new silts would require incorporating additional experimental data representing the variability in mineralogy, particle size distribution, and chemical composition. Retraining the ABC-ANN and GMDH models with such data would enhance their generalizability and ensure reliable predictions for a broader range of site-specific materials.

6. Conclusion

Flexural strength (FS) and unconfined compressive strength (UCS) of cement mortars containing waste silt estimated by the artificial neural network optimized by artificial bee colony algorithm (ABC-ANN) and also combinatorial GMDH methods. Consequently, data including 39 cement mortar specimens tested in the previous study by authors, were gathered. Moreover, the results of the previously examined design of experiments method (DOE) models for estimating FS and UCS parameters are investigated and compared to proposed ABC-ANN and combinatorial GMDH models using regular criteria for error and performance evaluation, including the R, R^2, MSE, RMSE, MAE, MAPE and NMAE. In the proposed models, the cement, water, sand, silt and additive contents were considered as the input parameters and corresponding flexural and unconfined compressive strength of cement mortars (FS and UCS) were considered as outputs. In model training, 80 % of the available experimental data was utilized, with the remaining 20 % reserved for the testing phase, yielding more precise results compared to the training phase. The outcomes showed that the suggested hybrid ABC-ANN models effectively predicted both flexural strength (FS) and unconfined compressive strength (UCS) parameters for cement mortars, surpassing the performance of GMDH and DOE models. However, all models were able to reliably predict the FS and UCS values. The following are the quantitative outcomes of the analysis.

- The architecture of an artificial neural network (ANN) with a single hidden layer was investigated by varying the number of neurons from 1 to 15. Results indicated that a configuration with six neurons in the hidden layer provided the best performance for predicting both flexural strength (FS) and unconfined compressive strength (UCS). For FS prediction, the model achieved correlation coefficients (R) of 0.944 and 0.859 on the training and testing datasets, respectively, alongside normalized mean squared errors (MSE) of 0.003 and 0.004 and mean absolute percentage errors (MAPE) of 8.956 % and 8.931 %. Regarding UCS estimation, the optimized ANN yielded R values of 0.957 and 0.952 for training and testing, with normalized MSE values of 0.002 and 0.001, and corresponding MAPE values of 12.809 % and 5.501 %.
- Tested the ABC algorithm with different numbers of artificial bees, from 10 to 1000; it was found that using 150 bees for FS estimation and 120 bees for UCS estimation could lead to the best results with high correlation and the least errors in ANN training optimization;
- For both FS and UCS estimation, ABC-ANN models outperform proposed combinatorial GMDH and previously investigated DOE models due to higher R^2 values and lower MSE, RMSE, MAE, MAPE and NMSE and NMAE errors. Aside from that, the findings of the combinatorial GMDH models are superior to those of the DOE models;

- The findings of the sensitivity analysis reveal that the FS and UCS values predicted in both ABC-ANN and GMDH models are exactly consistent with those of experimental results. Therefore, the models' accuracy in predicting FS and UCS can provide insights into their practical use, and their generalized responses can be inferred between the maximum and minimum values of the input variables;
- The parametric analysis examining the influence of each input variable on the correlation coefficient (R) and root mean squared error (RMSE) revealed that, for both the ABC-ANN and combinatorial GMDH models predicting flexural strength (FS) and unconfined compressive strength (UCS), water is the most influential parameter while cement is the least. Sand, additive, and silt contents rank as the second, third, and fourth most impactful inputs, respectively.

Disclaimer

The results and findings presented in this study are solely the outcome of an independent academic research project conducted by the authors. The data related to the reservoir waste silt were obtained with permission for research purposes only. S.A.P.A.B.A. (S.p.A.), Bologna, Italy, did not participate in the design, execution, analysis, or interpretation of the research and holds no responsibility for the conclusions drawn herein. The authors bear full responsibility for the content, interpretations, and any opinions expressed in this paper, which do not reflect the views or positions of S.A.P.A.B.A. (S.p.A.).

Funding

This work was carried out within the framework of the SAFERUP! Project, supported by the European Union#8217; s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 765057.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to express their sincere gratitude to S.A.P.A.B.A. (S.p.A.), Bologna, Italy, for providing access to the reservoir site and providing the waste silt material used in this research. The support enabled the experimental component of this study but did not influence the research design or analysis. This research was conducted independently, and all interpretations and conclusions are the sole responsibility of the authors.

References

[1] UN. Sustainable development goals. 2015.
[2] Kurad R, Silvestre JD, de Brito J, Ahmed H. Effect of incorporation of high volume of recycled concrete aggregates and fly ash on the strength and global warming potential of concrete. J Clean Prod 2017;166:485–502. https://doi.org/10.1016/j.jclepro.2017.07.236.
[3] Febin GK, Abhirami A, Vineetha AK, Manisha V, Ramkrishnan R, Sathyan D, et al. Strength and durability properties of quarry dust powder incorporated concrete blocks. Constr Build Mater 2019;228:116793. https://doi.org/10.1016/j.conbuildmat.2019.116793.
[4] Mashaly AO, El-Kaliouby BA, Shalaby BN, El-Gohary AM, Rashwan MA. Effects of marble sludge incorporation on the properties of cement composites and concrete paving blocks. J Clean Prod 2016;112:731–41. https://doi.org/10.1016/j.jclepro.2015.07.023.
[5] Vianna Bahiense A, Alexandre J, de Castro Xavier G, de Azevedo ARG, Monteiro SN. Dosage of interlocking paving with ornamental rock waste: an experimental design approach, particle packing and polluting potential. Case Stud Constr Mater 2021;15:e00596. https://doi.org/10.1016/j.cscm.2021.e00596.
[6] Wu L, Sun Z, Cao Y. Modification of recycled aggregate and conservation and application of recycled aggregate concrete: a review. Constr Build Mater 2024;431:136567. https://doi.org/10.1016/j.conbuildmat.2024.136567.
[7] Habibi A, Ramezanianpour AM, Mahdikhani M. RSM-based optimized mix design of recycled aggregate concrete containing supplementary cementitious materials based on waste generation and global warming potential. Resour Conserv Recycl 2021;167:105420. https://doi.org/10.1016/j.resconrec.2021.105420.
[8] Li W, Cai L, Wu Y, Liu Q, Yu H, Zhang C. Assessing recycled pavement concrete mechanical properties under joint action of freezing and fatigue via RSM. Constr Build Mater 2018;164:1–11. https://doi.org/10.1016/j.conbuildmat.2017.12.219.
[9] Khatir A, Capozucca R, Khatir S, Magagnini E, Le Thanh C, Riahi MK. Advancements and emerging trends in integrating machine learning and deep learning for SHM in mechanical and civil engineering: a comprehensive review. J Braz Soc Mech Sci Eng 2025;47:419. https://doi.org/10.1007/s40430-025-05697-5.
[10] Xu Y, Zhou Y, Sekula P, Ding L. Machine learning in construction: from shallow to deep learning. Dev Built Environ 2021;6:100045. https://doi.org/10.1016/j.dibe.2021.100045.
[11] Ben Chaabene W, Flah M, Nehdi ML. Machine learning prediction of mechanical properties of concrete: critical review. Constr Build Mater 2020;260:119889. https://doi.org/10.1016/j.conbuildmat.2020.119889.
[12] Abdulrahman PI, Jaf DKI, Malla SK, Mohammed AS, Kurda R, Asteris PG, et al. Predictive modeling of compressive strength in silica fume-modified self-compacted concrete: a soft computing approach. Struct Concr 2025;26:2340–65. https://doi.org/10.1002/suco.202400931.
[13] Adebanjo AU, Shafiq N, Razak SNA, Kumar V, Farhan SA, Singh P, et al. Design and modeling the compressive strength of high-performance concrete with silica fume: a soft computing approach. Soft Comput 2024;28:6059–83. https://doi.org/10.1007/s00500-023-09414-z.
[14] Nadgouda PA, Kumar DR, Sharma AK, Wipulanusat W. Optimizing compressive strength prediction of sustainable concrete using ternary-blended agro-waste ash, sugarcane bagasse ash, and rice husk ash with soft computing techniques. Struct Concr 2025. https://doi.org/10.1002/suco.202401562.
[15] Kang M-C, Yoo D-Y, Gupta R. Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. Constr Build Mater 2021;266:121117. https://doi.org/10.1016/j.conbuildmat.2020.121117.
[16] Naseri H, Jahanbakhsh H, Hosseini P, Moghadas Nejad F. Designing sustainable concrete mixture by developing a new machine learning technique. J Clean Prod 2020;258:120578. https://doi.org/10.1016/j.jclepro.2020.120578.
[17] Yang G, Bi Y, Wang Z, Si C, Li T, Pu C, et al. Investigating the shrinkage mechanism, characteristics, and environmental benefits of waste tire rubber-modified cement-stabilized macadam. Environ Res 2025;285:122293. https://doi.org/10.1016/j.envres.2025.122293.
[18] Pietrzak A, Ulewicz M, Kozień E, Pietraszek J. Application of a mixture of fly ash and solid waste from gas treatment from municipal solid waste incineration in cement mortar. Materials 2025;18:481. https://doi.org/10.3390/ma18030481.
[19] Jahangir H, Esfahani MR. Numerical study of bond – slip mechanism in advanced externally bonded strengthening composites. KSCE J Civ Eng 2018;22:4509–18. https://doi.org/10.1007/s12205-018-1662-6.
[20] Jahangir H, Esfahani MR. Experimental analysis on tensile strengthening properties of steel and glass fiber reinforced inorganic matrix composites. Sci Iran 2021;28:1152–66. https://doi.org/10.24200/sci.2020.54787.3921.
[21] Solouki A, Viscomi G, Tataranni P, Sangiorgi C. Preliminary evaluation of cement mortars containing waste silt optimized with the design of experiments method. Materials 2021;14. https://doi.org/10.3390/ma14030528.
[22] Mashaal NM, Abdelzaher MA, Awad MM, Elwakeel KZ, Hassan AA. Towards sustainability and circular economy: conceptual design model for the water-energy-food nexus in Ferdaws Village, Western desert, Egypt. Innov Infrastruct Solut 2025;10:223. https://doi.org/10.1007/s41062-025-02002-0.
[23] Abdelzaher MA, Farghali AA, Hamouda AS. Effective impact of nano-plastic-waste incorporated with nanotitina on the physical, mechanical and microstructural properties of white cement pastes composites for progressing towards sustainability. Sci Rep 2024;14:12581. https://doi.org/10.1038/s41598-024-62661-4.
[24] Owaid KA, Ghazal RY, Abdelzaher MA. Study of the effect of modification of asphalt on the rheological properties employing microwave radiation—An aging study. Recycling 2023;8:65. https://doi.org/10.3390/recycling8050065.
[25] Abdelzaher M, Hamouda A, El-Kattan I, Baher A. Laboratory study for accelerating the CKD mineral carbonation. Egypt J Chem 2021. https://doi.org/10.21608/ejchem.2021.93980.4425. 0–0.
[26] Abdelzaher M, Mohamed E, Shehata N, Salah H, Abbas R. Environmental safe disposal of cement kiln dust for the production of geopolymers. Egypt J Chem 2021;0–0. https://doi.org/10.21608/ejchem.2021.89060.4276.
[27] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943;5:115–33. https://doi.org/10.1007/BF02478259.
[28] Jahangir H, Rezazadeh Eidgahee D. A new and robust hybrid artificial bee colony algorithm – ANN model for FRP-concrete bond strength evaluation. Compos Struct 2021;257:113160. https://doi.org/10.1016/j.compstruct.2020.113160.
[29] Rezazadeh Eidgahee D, Rafiean AH, Haddad A. A novel formulation for the compressive strength of IBP-based geopolymer stabilized clayey soils using ANN and GMDH-NN approaches. Iran J Sci Technol - Trans Civ Eng 2019;44:219–29. https://doi.org/10.1007/s40996-019-00263-1.
[30] Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J Global Optim 2007;39:459–71. https://doi.org/10.1007/s10898-007-9149-x.

[31] Le LT, Nguyen H, Dou J, Zhou J. A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning. Appl Sci 2019;9:2630. https://doi.org/10.3390/app9132630.

[32] Karaboga D, Akay B. A comparative study of Artificial Bee Colony algorithm. Appl Math Comput 2009;214:108–32. https://doi.org/10.1016/j.amc.2009.03.090.

[33] Liang J, Xue Y. An adaptive GP-based memetic algorithm for symbolic regression. Appl Intell 2020;50:3961–75. https://doi.org/10.1007/s10489-020-01745-w.

[34] Ivakhnenko AG. The group method of data of handling; a rival of the method of stochastic approximation. Soviet Automatic Control 1968;13:43–55.

[35] Ivakhnenko AG. Heuristic self-organization in problems of engineering cybernetics. Automatica 1970;6:207–19. https://doi.org/10.1016/0005-1098(70)90092-0.

[36] Shahin MA, Maier HR, Jaksa MB. Data division for developing neural networks applied to geotechnical engineering. J Comput Civ Eng 2004;18:105–14. https://doi.org/10.1061/(ASCE)0887-3801(2004)18:2(105).

[37] Marquardt DW. An Algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 1963;11:431–41. https://doi.org/10.1137/0111030.

[38] Demuth H, Beale M. Matlab neural network toolbox user's guide version 6. 2009.

[39] Rezazadeh Eidgahee D, Soleymani A, Jahangir H, Nikpay M, Arora HC, Kumar A. Estimating the load carrying capacity of reinforced concrete beam-column joints via soft computing techniques. Artificial intelligence applications for sustainable construction. Elsevier; 2024. p. 367–87. https://doi.org/10.1016/B978-0-443-13191-2.00014-6.

[40] Naderpour H, Rezazadeh Eidgahee D, Fakharian P, Rafiean AH, Kalantari SM. A new proposed approach for moment capacity estimation of ferrocement members using Group Method of Data Handling. Eng Sci Technol Int J 2020;23:382–91. https://doi.org/10.1016/j.jestch.2019.05.013.