lajss.org
Latin American Journal of Solids and Structures
www.lajss.org

Machine Learning Techniques for Low-Carbon Concrete Strength Prediction

Ming Jing[a], Peng Sun[a*]

[a] School of Architecture and Engineering, Jiamusi University, Jiamusi, Heilongjiang, 154000, China. Email: pengsun131@gmail.com

* Corresponding author

Abstract

The strength of Low-Carbon Concrete (LCC) is influenced by multiple factors such as material composition, curing conditions, and Carbon Emission (CE) parameters, making it difficult to capture the variation patterns of strength under the influence of multiple factors. To address this, this study proposes a Low-Carbon Concrete Strength Prediction (LCCSP) method based on Machine Learning (ML) techniques. By modeling the CEs of concrete and identifying three key factors affecting the performance of LCC—age, Water-Cement Ratio (W/C), and the number of Cementitious Materials (CMs)—these factors are used as input samples for an LCCSP model based on Self-Organizing Map-Radial Basis Function Neural Networks (SOM-RBF NNs). First, the SOM network is employed to cluster and extract features from multi-dimensional input data (age, W/C, number of CMs). Then, the cluster centers and radii are used as the Hidden Layer (HL) parameters of the Radial Basis Function Neural Network (RBFNN) to enhance its ability to handle complex nonlinear relationships. The model predicts the strength of LCC based on the input samples. In experiments, the Mean Absolute Error (MAE) and the Coefficient of Determination (R^2) are used as evaluation metrics. The results show that the MAE remains relatively small at various time points, and the R^2 consistently approaches 1, demonstrating good adaptability and robustness for both single-factor and multi-factor inputs, providing more accurate and stable technical support for LCCSP.

Keywords

Machine Learning, Low-Carbon Concrete, Carbon Emission Modeling, Strength Prediction, Influencing Factors, SOM-RBF Neural Network

1 INTRODUCTION

With the global push toward carbon neutrality, the construction industry—a high-energy-consuming and high-carbon-emitting sector—faces an urgent need for green transformation (Kantekin et al., 2024). Concrete, as the most widely used material in construction engineering, consumes large amounts of natural resources such as cement and aggregates during its production process, accompanied by significant CEs (worldwide cement manufacturing is responsible for about 7% of CO_2 emissions caused by human activity). Against this backdrop, LCC technology, which replaces natural aggregates with Recycled Aggregates (RAs) and industrial waste residues (such as Fly Ash (FA) and slag) to partially substitute cement, has emerged as a core pathway for reducing the carbon footprint of buildings (Alam, 2024). However, the fluctuating physical properties of RAs (namely, high-water absorption and low strength) and the varying reactivity of industrial waste materials (Anvari et al., 2023) pose challenges to the strength stability of LCC. How to accurately predict its strength while reducing CEs has become a critical scientific issue in engineering applications and material design (Almeshal et al., 2024).

Khalil et al. previously studied a Backpropagation Neural Network (BPNN)-based prediction method for the Compressive Strength (CS) of CFRP-constrained concrete cylinders. They collected experimental data encompassing concrete properties (e.g., unconfined CS, dimensions), CFRP constraint characteristics (e.g., thickness, hoop tensile strength), and other factors (e.g., specimen chamfering) to construct a dataset. They designed a BPNN structure with an Input Layer (IL), a Hidden Layer (HL), and an Output Layer (OL) to predict CS based on input parameters. Following training and optimization, the model can reveal intricate, nonlinear connections between input and output parameters, surpassing the limitations of traditional models. However, the black-box nature of BPNNs makes it difficult to explain how input parameters synergistically influence prediction results. Additionally, this method requires quantifying the bidirectional constraints between material low-carbonization and strength stability, which traditional single-dimensional models struggle to address. This hinders a deeper understanding of constraint mechanisms and impedes model optimization and parameter sensitivity analysis in engineering applications (AL-Bukhaiti et al., 2023). Gilcyvania et al. studied a concrete Strength Prediction (SP) method based on critical shear cracks, analyzed the shear transfer mechanism of concrete after cracking, clarified the formation, expansion, and shape parameters of critical shear cracks, constructed a mechanical model correlating crack kinematic features with Concrete Strength (ConcS), and verified it through experiments. However, this method does not adequately consider the complex influencing factors of concrete structures (such as different cross-sectional forms and structural continuity) (Abbood, Rahman, & Abu Bakar, 2025; Abbood, Rahman, & Bakar, 2025). When faced with diverse real-world engineering scenarios, its prediction accuracy and stability are easily affected, making it difficult to precisely meet the SP requirements for LCC applications (Costa et al., 2025). Sebastian et al. suggested a concrete SP approach on the basis of random field modeling, using it to describe the spatial variability of ConcS and combining field or experimental data to determine relevant parameters. However, the relevant parameters in random field models are typically assumed to be static constants, while the spatial variability of ConcS may be influenced by dynamic factors such as age, temperature, humidity, and load history. The study may not have fully considered these time-varying characteristics (Geyer et al., 2023). Prayoonwet et al. studied a shear SP method for reinforced concrete beams derived from Neural Networks (NNs) combined with mechanical explanations, using mechanical mechanism features as inputs and shear strength test values as outputs, and utilizing mechanical theory to constrain and optimize the parameters of the NN model. However, the study may have only introduced basic mechanical assumptions, while the shear failure of reinforced concrete beams involves complex nonlinear processes that are difficult to fully describe with existing mechanical theories. If the NN model relies on simplified mechanisms, it may lead to prediction errors (Prayoonwet et al., 2024).

The above studies indicate that traditional methods either rely on empirical assumptions or are limited by data dimensions, making it difficult to simultaneously meet the multi-factor coupling prediction requirements of LCC. ML technology, with its powerful nonlinear fitting capabilities, provides a new paradigm for solving the above problems. Among ML approaches, RBFNNs can approximate complex function relationships with arbitrary precision, while Self-Organizing Mapping Neural Networks (SOMNNs) can achieve unsupervised clustering and reduce the number of dimensions of high-dimensional data. The combination of the two (the SOM-RBFN model) can effectively address the multivariate nonlinear mapping problem of “material composition-strength performance” in LCC. This paper utilizes ML techniques to study SP methods for LCC, primarily employing the SOM-RBFN model to establish a SP model based on multi-factor correlations, thereby overcoming the limitations of traditional methods. The main contributions of this study can be summarized as follows:

(1) A low-carbon concrete strength prediction framework based on a SOM-RBF neural network is developed to effectively capture the nonlinear relationships among age, water-cement ratio, and cementitious material content.

(2) A carbon-emission-oriented analysis is incorporated to identify the key factors that simultaneously influence both the mechanical performance and sustainability characteristics of low-carbon concrete.

(3) By integrating SOM-based feature clustering with RBF neural network prediction, the proposed model improves feature representation and enhances prediction accuracy under both single-factor and multi-factor conditions.

(4) The proposed method provides practical guidance for the design, optimization, and performance assessment of low-carbon concrete materials, supporting sustainable construction and carbon-reduction strategies in engineering applications.

2 LOW-CARBON CONCRETE STRENGTH PREDICTION METHODS

2.1 Concrete Carbon Emission Modeling

Concrete CE modeling is the logical starting point and core foundation for predicting the strength of LCC. Its essence is to transform the core goal of "low carbon" into quantifiable parameters that can be calculated and optimized, forming an organic whole with the factors influencing strength (Devine et al., 2024). The calculation boundary for the CE factor of LCC is depicted in Fig. 1.

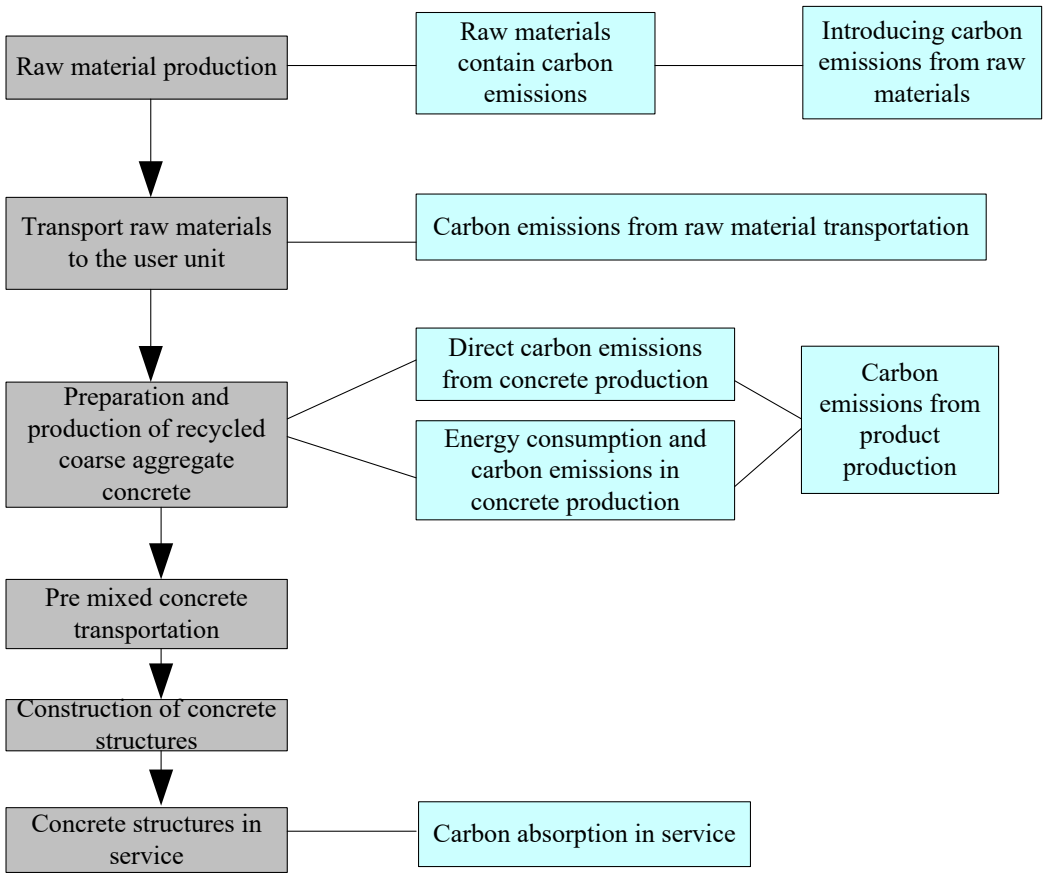

Figure 1 Calculation boundary of the CE factor for LCC

As demonstrated in Fig. 1, LCC, with recycled concrete as a typical representative, requires clear boundaries and components for calculating its CE factor, including CEs during raw material introduction, product manufacturing stages, and carbon absorption during the service life of concrete structures (Tahir et al., 2024). The calculation logic is the net CE per unit volume of concrete, i.e., total CEs minus total carbon absorption. The specific components are as outlined below:

(1) Raw material introduces CEs C_s

CEs C_s introduced by raw materials for low-carbon concrete include CEs C_c embedded in the raw materials, and CEs C_h from the fuel used to transport the raw materials. Different mix ratios of LCC result in different raw material carbon contents C_c (Kanazawa et al., 2024), calculated as:

$$C_c = \sum_{j=1}^{m} c_{sj} F_j \quad (1)$$

In the formula, F_j represents the CE factor of the jth raw material; c_{sj} represents the consumption of the jth raw material; m denotes the total number of raw material types.

The CEs C_h produced by the comprehensive fuel consumption during raw material transportation in LCC generally accounts for a high proportion (Perceka et al., 2023), and its calculation formula is:

$$C_h = \sum_{j=1}^{m} c_{sj} F_j \, c_f \Gamma \quad (2)$$

Here, c_f and Γ are the fuel consumption and transportation distance when transporting the j type of raw material.

Therefore, the calculation equation for the CEs of LCC raw materials, introducing CEs C_s is:

$$C_s = C_c + C_h \quad (3)$$

(2) Product manufacturing CEs C_q

The CEs throughout the production process of LCC mainly originate from energy consumption, especially electricity consumption (Pal et al., 2023). The equation for computing the CEs C_q of concrete produced is:

$$C_q = \sum_{i=1}^{m} c_{ei} F_i \quad (4)$$

Here, c_{ei} denotes the energy i consumed during the 1m^3 LCC production process; i signifies the type of energy consumed; F_i signifies the CE factor of the i type of energy.

(3) Carbon absorption of LCC structures during service life

The carbonation depth reflects the degree of concrete carbonation (Chu et al., 2023). The carbonation depth increases with age, and its relationship can be determined through empirical formulas or experimental data fitting. Based on data such as carbonation depth, the amount of carbon absorption C_b by LCC structures can be calculated using Formula (5):

$$C_b = 0.044 \frac{U_c}{U_0} \quad (5)$$

Where U_c and U_0 are the volume of concrete within the carbonation depth range and the total area of LCC, respectively.

(4) CE factor of LCC

The CE factor for LCC is:

$$C_o = C_s + C_q - C_b \quad (6)$$

2.2 Analysis of Factors Affecting the Strength of Low-Carbon Concrete

Combining the content of section 2.1, it is evident that the preparation of LCC is mainly influenced by raw material composition (such as cement, RA, etc.) and production processes (such as energy consumption, transportation, etc.). This document focuses primarily on analyzing the factors related to raw material composition to predict the strength of LCC. As shown in Eqs. (1) and (4), the main feature of LCC is using industrial waste residues (namely, FA and slag) to replace part of the cement, with the amount of replacement directly influencing the embodied CEs of raw materials and the CEs during product manufacturing. The W/C, as a key parameter of concrete mix design, influences the degree of cement hydration and strength development, and is also related to the quantities of raw materials (such as cement and water), indirectly involving CE calculations. ConcS varies with age, and the carbonation depth is closely related to age, requiring an overall analysis of CE and strength performance. Therefore, this paper designs the age, W/C, and CM amount as factors affecting the strength of LCC, as these three factors have the most direct impact on strength and are strongly correlated with CE parameters (for example, the amount of mix materials simultaneously affects both strength and CEs). Consequently, they are prioritized as the main factors influencing strength.

2.2.1 Age

The strength of LCC depends heavily on its age. ConcS rises with age, but after reaching a certain age, the strength attains a maximum value and then slowly decreases over time. A considerable body of research indicates that changes in the average ConcS over time can be separated into an increasing phase and a decreasing phase, with the rising period lasting approximately 25-30 years or even longer, while the decline phase exhibits a very slow decrease (Qin et al., 2025).

Research indicates that the ConcS in existing structures continues to exhibit a typical distribution pattern. However, the mean value and Standard Deviation (SD) shift based on the structure's age. Consequently, employing a non-stationary normal random process is suitable for illustrating the ConcS in existing structures (Javid et al., 2024). The long-term fluctuation of ConcS — represented by the changes in its mean and SD over time — can be mathematically described using regression models developed from comprehensive long-term exposure tests and field measurement data from real structures. The model for the temporal variation of the mean strength and SD of concrete is outlined below:

$$\mu_f(t) = \zeta(t)\mu_{f_0} \quad (7)$$

$$\ell_f(t) = \vartheta(t)\ell_{f_0} \quad (8)$$

Where μ_{f_0} denotes the average CS of LCC at 28 days; μ_f signifies the mean CS of LCC at the current stage; $\zeta(t)$ indicates a function that changes over time, employed to depict the variation pattern of mean LCC strength, it is precisely determined via this equation:

$$\zeta(t) = 1.4529e^{[-0.0246(\ln t - 1.7154)^2]} \quad (9)$$

Where $\ell_f(t)$ represents the current SD of CS for LCC; ℓ_{f_0} signifies the SD of CS for LCC 28d; $\vartheta(t)$ denotes a function of time, utilized to represent the variation law of the SD of LCC strength, specifically calculated by the following formula:

$$\vartheta(t) = 0.0305t + 1.2368 \quad (10)$$

Age is a key factor in the decline of LCC strength and serves as a vital input when forecasting LCC strength (Jadhav et al., 2024).

2.2.2 Water-Cement Ratio

The W/C symbolized as Ω/E, represents the proportion of water relative to cement within the LCC. It is an important parameter for calculating the mix ratio of LCC and an essential element determining the strength of LCC (Reshi et al., 2024). The cube CS of concrete is positively correlated with the W/C and can be approximately described by the Borromei empirical Formula (Mukhtar et al., 2023). Swiss scholar Borromei was the first to establish an empirical formula relating ConcS to the W/C. After continuous improvement, the following equation is obtained:

$$\mu_c = \phi\mu_e\left(\frac{E}{\Omega} - \varphi\right) \quad (11)$$

Where μ_e is the actual CS of cement at 28-day age; μ_c is the cube CS of concrete at 28-day age; ϕ and φ are regression coefficients pertaining to the cement and aggregate types employed in the concrete.

The W/C is a crucial element influencing the initial strength of LCC. The extent of initial strength greatly affects subsequent strength, since later strength is either improved or diminished depending on the early strength (Yosefi et al., 2025). Given the intricate makeup of contemporary concrete and the common application of CMs, the W/C does not truly represent the engineering importance of the W/C. Consequently, the water-CM ratio (water-binder ratio) is employed instead (Song, 2024). As a result, this study also considers the water-binder ratio as an input variable for predicting the strength of LCC.

2.2.3 Cementitious Material Usage

Cementitious Blended Materials (CBMs) consist of industrial byproducts and natural minerals incorporated during cement manufacturing to modify the cement grade and enhance its properties (Vafaei et al., 2024). CBMs are divided into active and inactive types. Inactive Blended Materials (BMs) refer to those that are not reactive or have very low reactivity; when mixed with water and Portland cement, they rarely or cannot form water-hardening cementitious substances. Common inactive BMs include limestone, ground slag, and quartz sand. Active BMs, in the presence of calcium hydroxide Ca $(OH)_2$ solution, undergo significant hydration reactions to form high-strength water-hardening cementitious substances (Wright et al., 2023). Typical active BMs consist of volcanic ash, granulated blast furnace slag, and FA. The BMs examined in this study are reactive, primarily emphasizing the investigation of slag and FA (Dinesh et al., 2024).

Slag and FA primarily consist of reactive SiO_2 and Al_2O_3. They are not only reactive but also exist as glassy solids or hollow spherical particles. When incorporated into cement, they can improve workability and enhance the density of the cement matrix (Manan et al., 2025). When active BMs are added to $Ca(OH)_2$ solution, significant hydration reactions occur, producing water-hardening cementitious substances. The reaction is as follows:

$$\varsigma Ca(OH)_2 + SiO_2 + \omega H_2O = \varsigma CaO \cdot SiO_2 \cdot \upsilon H_2O \quad (12)$$

$$\varsigma Ca(OH)_2 + Al_2O_3 + \omega H_2O = \varsigma CaO \cdot Al_2O_3 \cdot \upsilon H_2O \quad (13)$$

Where CaO,H_2O,ς represent calcium oxide, water, and the stoichiometric coefficient (reflecting the molar amount of calcium hydroxide involved in the reaction); ω and υ are the molar coefficients of water and bound water; $Ca(OH)_2$ reacts with SiO_2 to produce amorphous calcium silicate hydrate, which progressively develops into a microcrystalline or

weakly crystalline gel. The reaction between $Ca(OH)_2$ and Al_2O_3 generates calcium aluminate hydrate. In the presence of gypsum, this compound further reacts to yield hydrated calcium sulfoaluminate (Haghi et al., 2025). These hydration products can harden in both air and water, exhibiting relatively high strength. The quantity of BM also plays a significant role in determining the LCC's strength and serves as a crucial input factor in concrete SP.

2.3 Low-Carbon Concrete Strength Prediction Model Based on SOM-RBF Neural Network

Using the factors analyzed in Section 2.2 (age, W/C, CM dosage) for the preparation of LCC as input samples, the input is fed into the LCCSP model based on the SOM-RBF NN to perform SP. The RBFNN in this model, as an ML technique, has core advantages such as strong nonlinear fitting ability, excellent generalization performance, and high modeling efficiency, which can effectively solve the problem of traditional methods being unable to capture complex influencing factors. The RBFNN exhibits strong generalization ability, and its network architecture is depicted in Fig. 2.

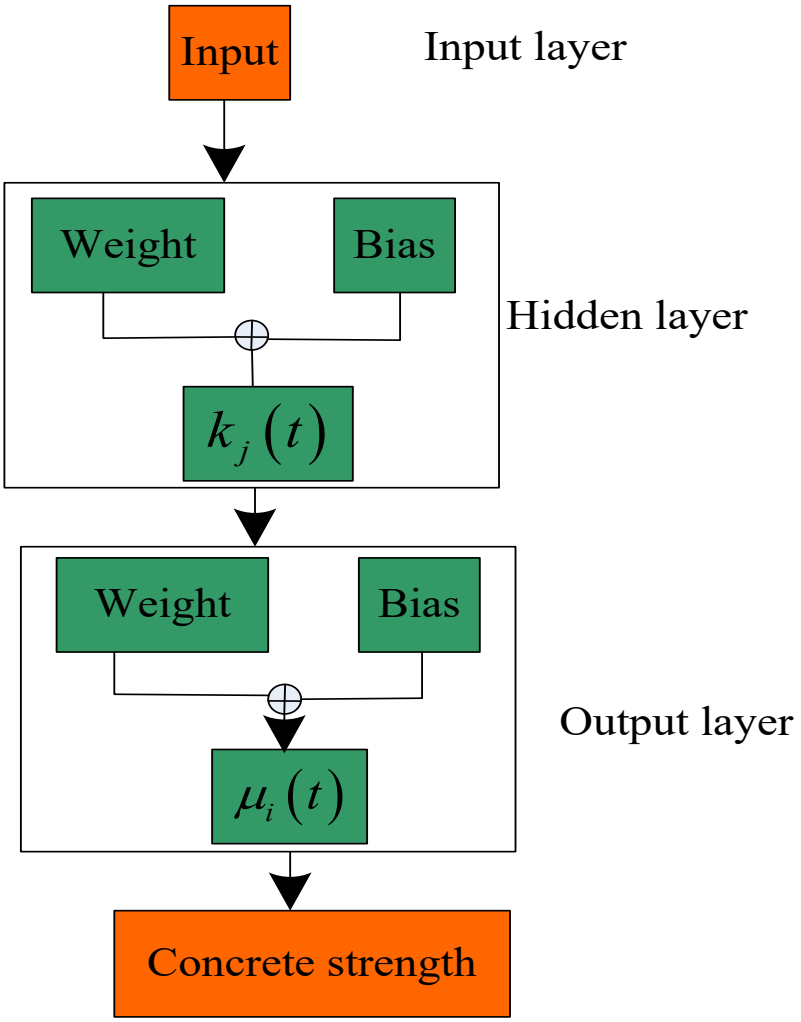

Figure 2 RBFNN structure.

Its structure is simple and can approximate any nonlinear function at arbitrary precision, making it highly regarded for avoiding lengthy and unnecessary computations. The RBFNN consists of three layers: IL, HL, and OL, with the vector dimensions of the IL and OL required to be consistent. The activation functions of the neurons in the HL are composed of RBFs, and the array of computational units in the HL is called HL nodes. Each HL node contains a center vector r, which shares the same dimensionality as the input vector y. The Euclidean Distance (ED) between them is given by $\|r_j(t) - y(t)\|$. The output of the HL is composed of nonlinear activation functions $k_j(t)$:

$$k_j(t) = exp\left(-\frac{\|r_j(t) - y(t)\|^2}{2z_j^2}\right), j = 1,2,\ldots,v \tag{14}$$

Where $y(t)$ represents the factors involved in the preparation of LCC (age, W/C, CM dosage); $r_j(t)$ signifies the center vector of the radial basis network; z_j denotes the width of the Gaussian basis function, generally a positive scalar; v is the number of neurons in the HL. The predicted strength of LCC output by the network is:

$$\mu_i(t) = \sum_{i=1}^{\lambda} \theta_{ji} k_j(t), i = 1,2,\ldots,\lambda \quad (15)$$

Where θ denotes the weights of the OL; λ signifies the number of output nodes.

The RBF (such as the Gaussian function) computes the distance between the input sample and the cluster center, with the center serving as the origin. If the RBF cluster centers and radii are initialized randomly, extensive computation and iteration are required to find suitable values. However, SOM performs unsupervised clustering on samples, quickly obtaining more reasonable initial cluster centers and radii, reducing the parameter search space during RBF training, and accelerating convergence. For example, when dealing with complex multi-factor data of LCC, SOM pre-clustering allows RBF to avoid blind searches in the parameter space and completes training more rapidly.

By combining SOM's clustering ability with RBF's prediction accuracy, a model is constructed using a "two-step" approach: first, SOM is used to group input samples and extract cluster centers and radii; then, these are used as the centers and widths of the RBF-HL, improving the model's capacity to manage intricate information.

Cluster center: Represents the core of a sample group characterized by "similar material composition, CE features, and strength development trends." For example, a center may correspond to a type of concrete with "30% low-carbon cement content + 20% FA addition + slow early strength gain but stable later," with each dimension in the Weight Vector (WV) representing the "typical value" of these features.

Cluster radius: Reflects the degree of variation of this type of concrete sample around the "typical value." For instance, a large radius indicates a wide range of variation in "cement content, FA addition"; a small radius indicates more concentrated features, allowing the RBFs to focus more during prediction.

SOM is also a type of NN, comprising an IL and an OL (also called a competitive layer), with a structure shown in Fig. 3.

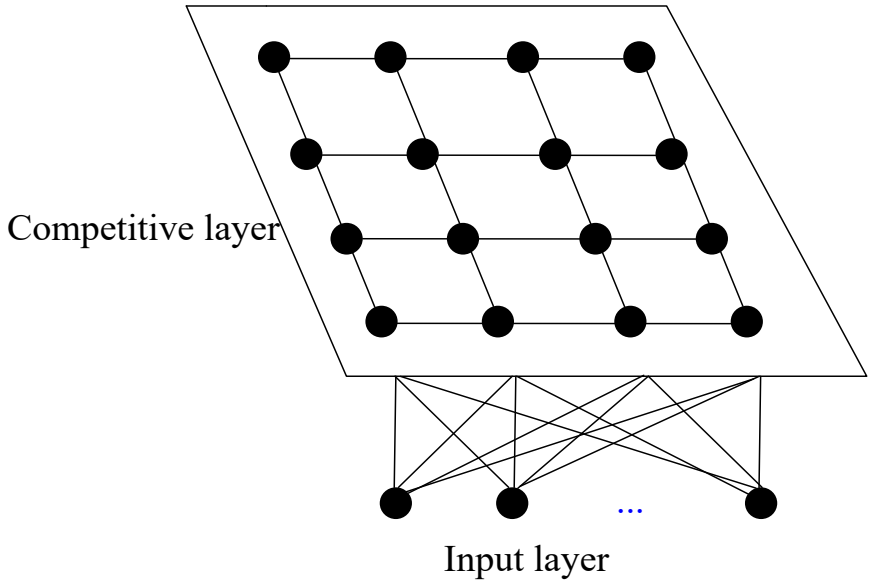

Figure 3 Structure of SOMNN

The number of neurons in the IL is M, and the competitive layer is composed of N neurons arranged in a one-dimensional or two-dimensional array, which can also be higher-dimensional, though this is less common.

The basic principle of SOM is that when a sample of factors affecting LCC preparation (age, W/C, CM amount) is input, the neuron in the OL that receives the strongest stimulus wins, and the neighboring neurons around the Winning Neuron (WN) are also stimulated due to lateral effects. At this point, the network performs a learning operation, adjusting the connection WVs of the WN and its neighbors toward the input pattern. The WN in the competitive layer shifts from one neuron to another when the input pattern category alters. Thus, the network self-organizes by using substantial quantities of data to modify its connection weights, ultimately enabling the output feature map to reflect the distribution of the sample data. The basic SOM learning algorithm's fundamental stages are:

(1) Set variables

The input samples are factors affecting LCC preparation (age, W/C, and CM amount). Using a competitive learning mechanism, samples are mapped to a two-dimensional OL (competitive layer), achieving data dimensionality reduction and clustering. The input sample is $Y = [y_1, y_2, \ldots, y_l]$, the total number of samples is l, and the WV between the j -th input node and output neuron is θ_j. The cluster center corresponds to the typical features of a certain type of LCC preparation factors, and the cluster radius reflects the fluctuation range of these factors.

(2) Initialization

Weights are initialized with small random values, and both input vectors and weights are normalized:

$$Y' = \frac{Y}{\|Y\|} \tag{16}$$

$$\theta_j' = \frac{\theta_j}{\|\theta_j\|} \tag{17}$$

Where $\|Y\|$ and $\|\theta_j\|$ are respectively the Euclidean norms of the input LCC preparation factor sample and the WV.

(3) Input the randomly sampled LCC preparation factor samples into the network. The samples are dot-multiplied with the WVs, and the neuron with the maximum Inner Product (IP) wins the competition. Given that both the sample vectors and the weights have been normalized, identifying the maximum IP is the same as identifying the minimum ED σ:

$$\sigma = \|Y' - \theta_j'\| \tag{18}$$

Identify the neuron with the smallest distance, denoted as the WN.

(4) Update weights

Update the neurons within the topological neighborhood of the WN using the Kohonen rule:

$$\theta'(t+1) = \theta'(t) + \pi(t)\big(Y' - \theta'(t)\big) \tag{19}$$

Where $\pi(t)$ represents the Learning Rate (LR).

(5) Update the LR $\pi(t)$ and the topological neighborhood $\rho_j(t)$ (used to describe the set of neurons adjacent to the WN in the competitive layer), and normalize the weights after learning. The update formula is:

$$\pi(t) = \pi_0 * \left(1 - \frac{\kappa}{T}\right) \tag{20}$$

$$\rho_j(t) = INT\left[\rho_j(0) * \left(1 - \frac{\kappa}{T}\right)\right] \tag{21}$$

Where INT denotes the rounding operation; π_0 is the initial LR; $\rho_j(0)$ is the initial neighborhood radius; κ and T are the current iteration step and total number of iterations, respectively.

(6) Extract the WVs; each neuron corresponds to a set of weights θ (consistent with the IL dimension, for example, input features such as material composition, CEs, curing conditions of LCC, which are multi-dimensional, and the weights are multi-dimensional vectors). After training, these weights will converge to the core distribution area of the corresponding cluster samples. Directly take each neuron's final WV as the center of the cluster, representing the core of the LCC samples in the multi-dimensional feature space. Calculate the ED between each cluster center and all samples belonging to that cluster:

$$\sigma = \|\theta' - Y'\|_2 \tag{22}$$

Use the SD of these distances to represent the radius:

$$z = std(\sigma_1, \sigma_2, \ldots, \sigma_l) \tag{23}$$

Where std indicates the SD of the EDs between samples of a certain class and the corresponding cluster center, which is used to quantify the fluctuation range of the samples in the feature space. The obtained cluster centers and radii are then used as the centers and widths of the RBF-HL, and they are substituted into Equation (14). Combining the output of Equations (15) and (14), the strength of LCC is predicted.

3 EXPERIMENTAL ANALYSIS

3.1 Experimental Content and Setup

To test the effectiveness of the Proposed Method (PM) in predicting the strength of LCC, in the experiment, recycled concrete was prepared as LCC specimens. For the creation of LCC, P.O42.5 grade standard Portland cement was utilized, and its chemical makeup is presented in Table 1. The cement's technical properties were assessed, with the findings detailed in Table 2.

Table 1 Chemical Composition of Cement

Component	Content /wt%
CaO	62.91
SiO2	21.6
Al2O3	5.41
Fe2O3	2.99
MgO	2.16
SO3	2.57
K2O	0.86
Na2O	0.21
LOI	1.23

Table 2 Technical Indicators of Cement

Indicator Type		Numerical Value
Setting time /min	Initial setting	222
	Final set	286
Compressive strength (CS) /MPa	7d	39.2
	28d	48.6
Flexural Strength /MPa	7d	5.9
	28d	8.2

For the creation of LCC, the chosen natural coarse aggregate consists of crushed stone found in nature, featuring a particle size range of 4.75–26.5 mm and a consistent gradation. The Recycled Coarse Aggregate (RCA) was sourced from a wide range of materials, with a total of 46 batches. Each batch of RCA was screened, and the particle size adjusted. The performance indicators of the RCA are displayed in Table 3, and the RCA utilized meets the particle size requirements specified in the standards. The water utilized is standard tap water. The water-reducing admixture is a high-performance polycarboxylate superplasticizer, possessing a solids content of 20% and a water-reducing capability of about 15%. The liquid has a light-yellow hue. Table 4 shows the basic performance information of FA.

Table 3 Performance Indicators of RCA

Material	Recycled Fine Aggregate	Recycled Coarse Aggregate
24-hour absorption rate /%	36.5	14.2
Bulk density /(kg•m-3)	1194	1181
Apparent density /(kg•m-3)	2413	2518
Crushing value /%	33.5	19.8

Table 4 Basic Performance Information of FA

Type	Numerical Value
Fineness /%	17
Loss on ignition /%	2.9
Bulk density /(kg•m-3)	1.13
Density /(kg•m-3)	2.56
Al2O3 content /%	24.3
SiO2 content /%	45.2
CaO content /%	5.7

To tackle the problem of substantial water uptake in RCA, this study incorporated extra water during the concrete mixing process, beyond the normal mixing water, taking into account the RCA's water absorption capacity. The mix proportions for the C50, C40, C30, and C20 ordinary concrete are depicted in Table 5.

Table 5 Concrete Mix Proportion

Target Strength	C20	C30	C40	C50
Water-cement ratio	0.56	0.48	0.38	0.34
Water /kg•m3	167	179	169	161

Cement /kg•m3	301	376	451	486
Coarse aggregate /kg•m3	1137	1137	1137	1137
Sand /kg•m3	668.6	668.6	668.6	668.6
Water reducing agent /kg•m3	3.7	4	5.2	6.2

After determining the experimental mix proportions, the specific preparation steps for the LCC test specimens are as follows:

(1) First, measure the required quantities of all materials for the experiment. Add the coarse aggregate and additional water to the mixer and stir for 120 seconds. Then, add the fine aggregate and cement in sequence to the mixer and stir for 1 minute each. During this time, mix the measured water-reducing agent and water thoroughly. Subsequently, pour the entire mixture into the mixer and continue stirring for 2–3 minutes.

(2) Pour all the concrete mixture into the test molds. The test blocks prepared in this experiment are cubic CS test blocks with dimensions of 10 cm × 10 cm × 10 cm. To facilitate demolding, the molds are coated with an appropriate amount of release oil before pouring the concrete.

(3) After the molds are filled, place them on a vibrating table and vibrate until the concrete is compacted. Once the surface of the concrete in the molds begins to bleed, stop the vibrating table. Use a trowel to smooth the surface of the concrete, then place the molds on a flat surface for 24 hours before demolding.

(4) After demolding, place the test specimens in a standard curing room for 28 days.

The preparation, curing, and compressive strength testing procedures of concrete specimens were conducted in accordance with the relevant provisions of the latest applicable concrete testing standards to ensure the reliability and repeatability of the experimental results.

After curing to the corresponding age, remove the LCC cube test specimens from the curing room. The CS test is performed with a SYE-2000A universal pressure machine. The load is implemented in a continuous and even manner, keeping the loading rate in the range of 0.5 to 0.8 MPa/s.

Before the test, the initial distance between the universal pressure testing machine's ram and the specimen is precisely adjusted to ensure uniform loading and lay the foundation for obtaining high-precision CS data. The experimental system consists of a universal pressure testing machine, a computer data acquisition terminal, and standard concrete test specimens. The test specimens are prepared in strict accordance with the specifications and precisely placed in the loading area of the testing machine. The computer continuously collects and records the mechanical response parameters during the loading process, thoroughly illustrating how the concrete acts mechanically when compressed. The concrete test specimens and their compression process in the concrete CS test are shown in Fig. 4.

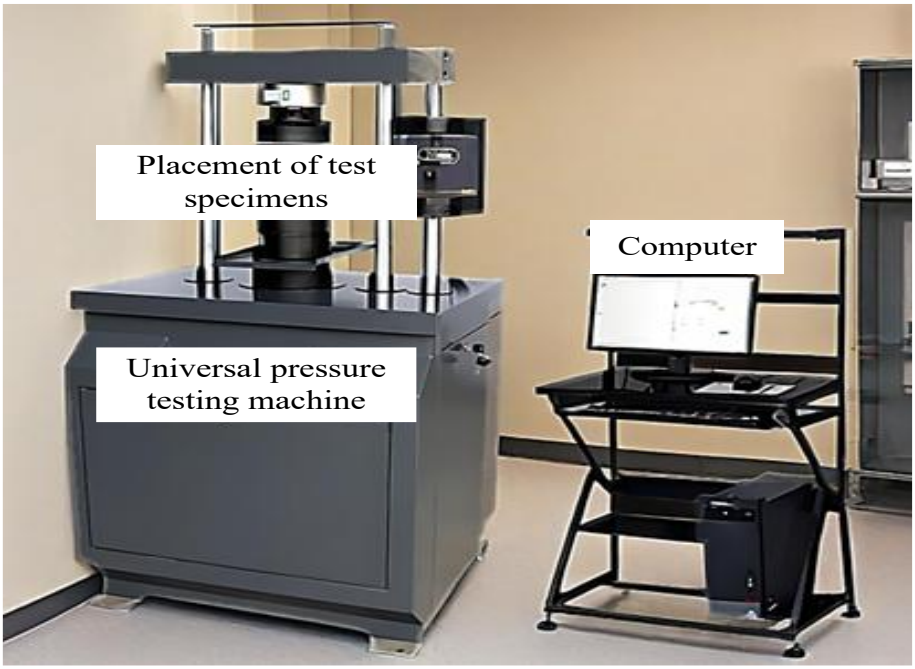

Figure 4 Experimental Scene Diagram

In the experiment, the PM in this document was primarily employed as the core approach for predicting the strength of the prepared LCC in the computer. Before testing the performance of the PM in predicting the strength of LCC, the

training samples shown in Table 6 were set, covering sample codes, strength grades, W/Cs, replacement rates of mixed materials, age, and measured strength values, to establish a data foundation. This lays the experimental groundwork for subsequent investigations into the accuracy and applicability of this method for predicting the strength of LCC, aiding in the analysis of the relationships between various factors (W/C, replacement rate, age, etc.) and ConcS, and advancing the development of LCC performance prediction technology.

Table 6 Training Sample Setting Details

Sample Code	Strength	Water-Cement Ratio	Substitution Rate of the Mixture	Age /d	Strength Measured Value /MPa
1	C20	0.51	0	120	55.4
2	C20	0.36	0.53	28	32.2
3	C30	0.45	0.15	28	40.6
4	C30	0.51	0.23	28	51.8
5	C40	0.51	0.45	28	47.8
6	C40	0.41	0.53	28	29.4
7	C50	0.51	0.23	28	58.7
8	C50	0.37	0.18	90	56.1

3.2 Performance Testing of the Proposed Method

The proposed SOM-RBF neural network framework was implemented using Python on a personal workstation equipped with an Intel Core i7 processor, 16 GB RAM, and the Windows operating system. The model development and data processing procedures were conducted using commonly adopted scientific computing and machine learning libraries to ensure computational efficiency and reproducibility. Using the training samples in Table 6, an LCCSP model on the basis of the SOM-RBF NN is trained. After model training, the data shown in Table 7 are used as test samples. First, the prediction effect of the LCC strength model based on the SOM-RBF NN is tested before applying the SOM network, with details shown in Fig. 5.

Table 7 Details of Test Samples for SP of LCC

Sample Code	Strength	Water-Cement Ratio	Substitution Rate of the Mixture	Age /d	Strength Measured Value /MPa
1	C20	0.47	0.25	730	57.5
2	C30	0.47	0.3	730	57
3	C40	0.36	0.35	360	53.5
4	C40	0.36	0.4	180	52.5
5	C50	0.41	0.45	180	56.5
6	C50	0.45	0.5	180	55
7	C50	0.49	0.45	180	55.5
8	C50	0.39	0.5	90	55
9	C50	0.43	0.5	60	60
10	C50	0.41	0.55	28	55.5

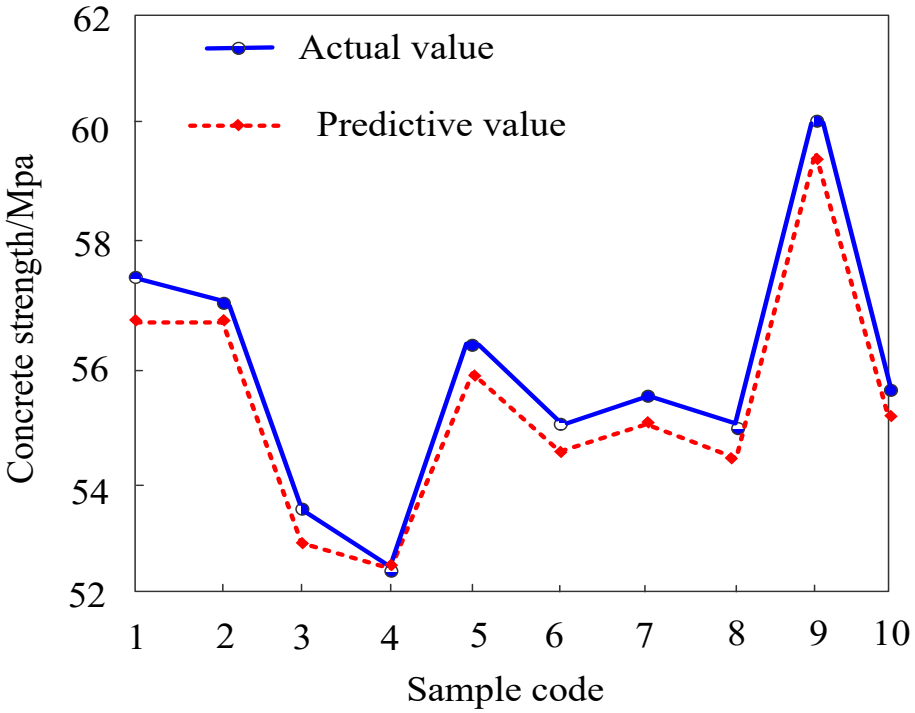

Figure 5 Prediction effect of LCC strength before using the SOM network

From Fig. 5, it is evident that the predicted values of LCC strength generally correlate with the actual values, but some samples (such as codes 4 and 8) show significant deviations. These deviations reflect that, without using the SOM network, the SP model for LCC under complex factor coupling effects struggles to fit accurately. This may be due to the complex laws governing the combined influence of elements, namely, W/C, replacement rate, and curing age on strength. Conventional RBFNNs have limitations in data mapping and pattern recognition, making it difficult to effectively handle nonlinear relationships among multiple factors, leading to prediction inaccuracies for some samples and highlighting issues with insufficient prediction accuracy.

After employing the SOM network, the prediction effect of the LCCSP model on the basis of the SOM-RBF NN is depicted in Fig. 6.

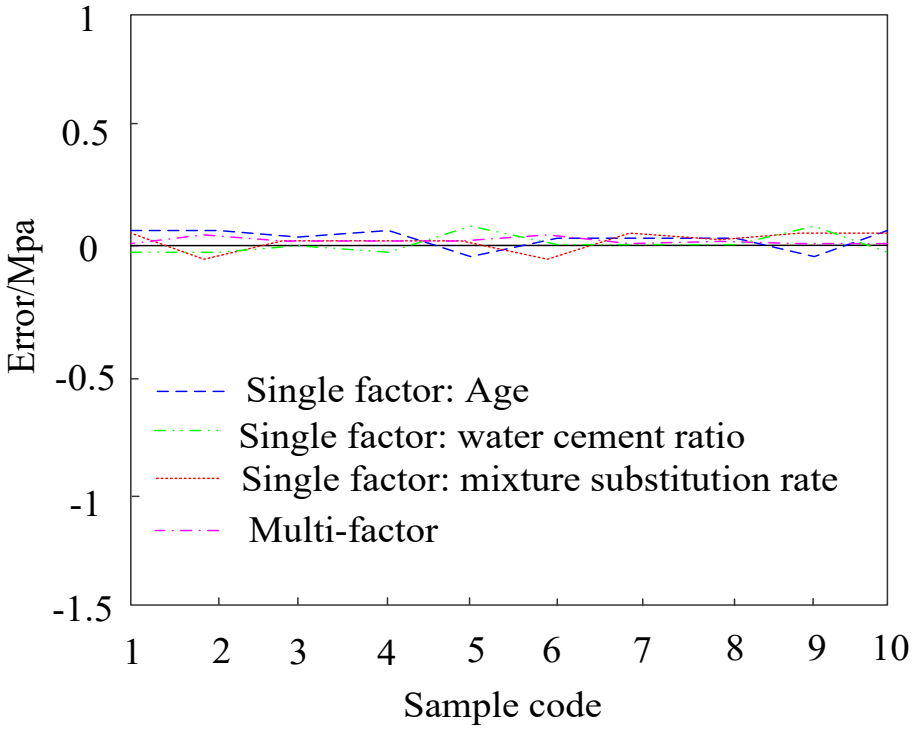

Figure 6 Prediction effect of LCC strength

As shown in Fig. 6, under single-factor (age, W/C, replacement rate of aggregate) and multi-factor inputs, the prediction errors of LCC strength fluctuate within a small range around zero. Compared to the significant deviations observed in some samples when the SOM network was not introduced, the error distribution after introduction is more concentrated and stable, indicating an improved adaptability of the model to different influencing factors (single-factor or multi-factor coupling). When dealing with the complex nonlinear relationships in LCCSP, the clustering and feature optimization provided by the SOM network enable the RBFNN to map the input data more accurately. With multi-factor inputs, the error curves are consistent with those of single factors (age, W/C, replacement rate of aggregate) in overall trend, and the fluctuation amplitudes are similar. This shows that after optimization by the SOM network, the model does not significantly increase the prediction error despite the increased complexity of factor coupling when utilizing multi-factor collaborative information, demonstrating that the SOM network effectively clarifies the correlations among multiple factors and helps the RBFNN to more efficiently extract multi-dimensional features. Meanwhile, the prediction errors for single factors are not noticeably worse than those for multiple factors, reflecting that the SOM-RBF model responds more accurately to a key single factor, and the model's robustness to different input patterns is enhanced.

To further evaluate the performance of the PM, three comparison methods were chosen: the BPNN-based compressive SP method (Comparison Method 1), the random field modeling-based concrete SP method (Comparison Method 2), and the NN prediction method combined with mechanical interpretation (Comparison Method 3). The MAE and R^2 were adopted as evaluation metrics to assess the prediction performance of the four methods on test sample 4 under universal pressure testing machine at different loading times. Figs. 7 and 8 present the MAE and R^2 test results for the SP of LCC using different methods. All comparison methods were trained using the same input feature set (age, W/C, and admixture content) to ensure comparability.

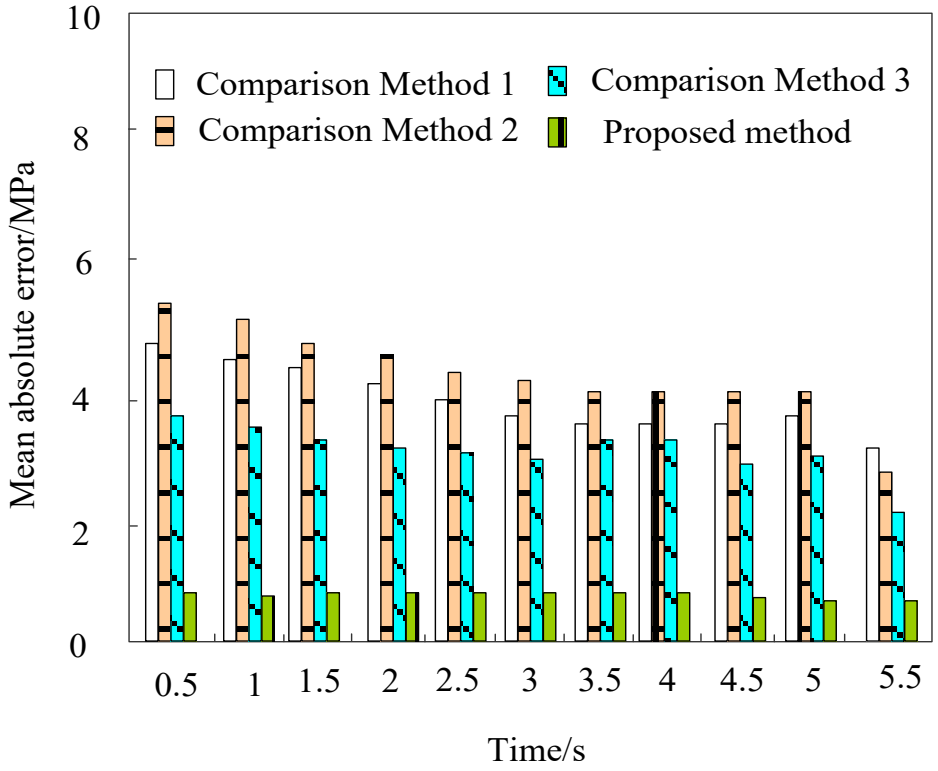

Figure 7. Average absolute error of SP results for LCC using different methods

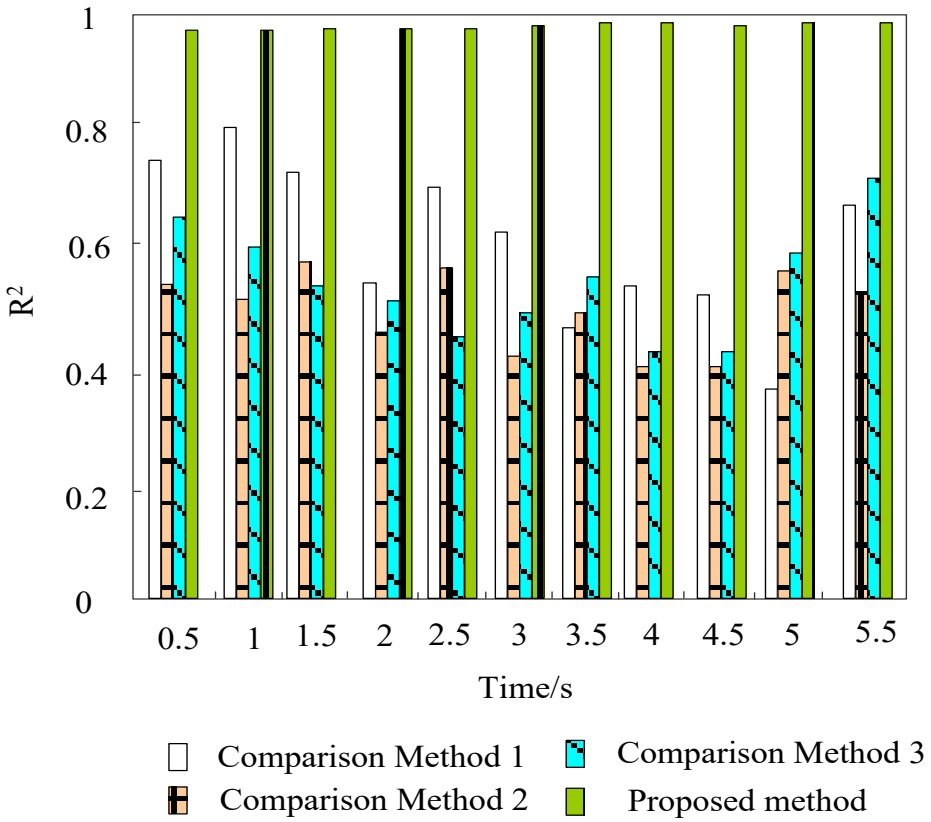

Figure 8 R2 values of different methods for predicting the strength of LCC

Figs. 7 and 8 demonstrated that the MAE of the three comparison methods for predicting the strength of LCC is generally at a high level and fluctuates significantly with changes in the loading time. However, the MAE of the PM is considerably less than that of the comparison approaches and maintains a small value at all time points, indicating that the absolute error between the predicted values and the measured values is smaller, and the PM can more accurately reflect the actual strength of LCC, effectively reducing prediction errors. Additionally, the R^2 test results of the PM consistently approach 1, far exceeding those of the comparison methods. The R^2 values of the three comparison methods are not only low but also exhibit significant fluctuations over time, indicating poor fitting to the sample data and weak ability to explain data variability. Through dual-indicator validation, the PM demonstrates higher accuracy and stability in the prediction of LCC strength.

4 CONCLUSIONS

Low-Carbon Concrete (LCC) utilizes recycled aggregates and industrial waste residues, such as fly ash and slag, to reduce carbon emissions associated with conventional concrete production. However, the variability of recycled materials and supplementary cementitious materials can significantly affect strength development, making accurate strength prediction essential for engineering applications. To address this issue, this study proposed a Low-Carbon Concrete Strength Prediction (LCCSP) method based on a Self-Organizing Map–Radial Basis Function Neural Network (SOM-RBFNN). The proposed model incorporates three key factors affecting LCC strength, namely age, water-cement ratio, and cementitious material content. The SOM network was employed to perform feature clustering and extract representative cluster centers and radii, which were subsequently used to optimize the hidden-layer parameters of the RBF neural network. Experimental results demonstrated that the introduction of the SOM network significantly improved prediction performance by enhancing the model's ability to capture complex nonlinear relationships among multiple influencing factors. Compared with the conventional RBF neural network and the selected benchmark methods, the proposed approach produced lower prediction errors and more stable prediction results under both single-factor and multi-factor input conditions. The Mean Absolute Error (MAE) remained at a relatively low level throughout the testing process, while the coefficient of determination (R^2) consistently approached 1, indicating excellent fitting capability, robustness, and prediction accuracy. Overall, the proposed SOM-RBFNN framework provides an effective and reliable

tool for LCC strength prediction. The method offers practical support for material design, performance evaluation, and quality control of low-carbon concrete, contributing to the broader adoption of sustainable construction materials in engineering practice. Although the proposed method demonstrated satisfactory prediction accuracy, several limitations remain. The experimental dataset used in this study was relatively limited in size and primarily focused on specific low-carbon concrete mixtures. Therefore, the model's generalization capability should be further validated using larger and more diverse datasets. Future studies will investigate additional influencing factors, including curing temperature, aggregate characteristics, and environmental exposure conditions, while exploring advanced machine learning and hybrid artificial intelligence techniques to further improve prediction accuracy and robustness.

Acknowledgments

This research was supported by the Basic Scientific Research Funds for Provincial Higher Education Institutions of Heilongjiang Province (Project No. 2023-KYYWF-0573), titled “Research on the Frost Resistance of Recycled Coarse Aggregate Concrete in Severe Cold Regions”.

Author’s contributions: Peng SUN: Supervision, Conceptualization, Project administration, Writing-Original draft preparation. Ming JING: Software, Methodology

Data Availability: Available upon request

Editor: Marco L. Bittencourt

References

Abbood, I. S., Rahman, N. A., & Abu Bakar, B. H. (2025). Enhanced data-driven shear strength predictive modeling framework for RCDBs using explainable boosting-based ensemble learning algorithms coupled with Bayesian optimization. *Results in Engineering*, *27*, 106556. doi: https://doi.org/10.1016/j.rineng.2025.106556

Abbood, I. S., Rahman, N. A., & Bakar, B. H. A. (2025). Shear Strength Prediction for RCDBs Utilizing Data-Driven Machine Learning Approach: Enhanced CatBoost with SHAP and PDPs Analyses. *Applied System Innovation*, *8*(4). doi: 10.3390/asi8040096

Alam, M. S. (2024). Proposed prediction models for shear strength of fiber reinforced polymer reinforced concrete deep members without stirrups. *Advances in Structural Engineering*, *27*(2), 314–332.

AL-Bukhaiti, K., Liu, Y., Zhao, S., & Abas, H. (2023). An application of BP neural network to the prediction of compressive strength in circular concrete columns confined with CFRP. *KSCE Journal of Civil Engineering*, *27*(7), 3006–3018.

Almeshal, I., Özkılıç, Y. O., Aksoylu, C., Karalar, M., & Alharthai, M. (2024). Ductility and strength of reinforced concrete beams strengthened with aluminum CNC waste. *Structural Concrete*, *25*(5), 3232–3245.

Anvari, A. T., Babanajad, S., & Gandomi, A. H. (2023). Data-driven prediction models for total shear strength of reinforced concrete beams with fiber reinforced polymers using an evolutionary machine learning approach. *Engineering Structures*, *276*, 115292.

Chu, S. H., Kurumisawa, K., & Kong, Y. K. (2023). Physically explicable mathematical model for strength prediction of UHPFRC. *Engineering Structures*, *275*, 115191.

Costa, G., Campos, C. M. O., Martha, L. F., & Cardoso, D. C. T. (2025). Novel Multiaction Model for Shear Strength Prediction of Reinforced Concrete Beams: Development and Application to Elements with FRP Bars and FRC. *Journal of Composites for Construction*, *29*(1), 04024092.

Devine, R. D., Barbachyn, S. M., Kurama, Y. C., & Thrall, A. P. (2024). Numerical Modeling of Squat Reinforced Concrete Shear Walls with High-Strength Materials. *ACI Structural Journal*, *121*(5), 175–188.

Dinesh, A., & Prasad, B. R. (2024). Predictive models in machine learning for strength and life cycle assessment of concrete structures. *Automation in Construction*, *162*, 105412.

Geyer, S., Papaioannou, I., & Straub, D. (2023). Spatial modeling of concrete strength based on data. *Structural Safety*, *103*, 102345.

Haghi, N., Efe, S., & Epackachi, S. (2025). Seismic response prediction of composite plate shear walls-concrete filled (C-PSW/CF) using machine learning methods. *Engineering Structures*, *322*, 119228.

Jadhav, S. P., Selvam, M., Zunain, S. M., & Bugalia, N. (2024). Evaluating external generalizability of machine learning models for recycled aggregate concrete property prediction. *Journal of Cleaner Production*, *469*, 143166.

Javid, A., & Toufigh, V. (2024). Utilizing ensemble machine learning and gray wolf optimization to predict the compressive strength of silica fume mixtures. *Structural Concrete*, *25*(5), 4048–4074.

Kanazawa, T., Nagai, K., & Bolander, J. E. (2024). Optimization-Based Analysis of Diagonal Tension Failure of Reinforced Concrete Dapped-End Beams. *Journal of Structural Engineering*, *150*(10), 04024137.

Kantekin, Y., & Bakir, B. B. (2024). Joint shear strength prediction for fiber reinforced concrete beam-to-column connections. *Engineering Structures*, *308*, 117959.

Manan, A., Pu, Z., Ahmad, J., & Umar, M. (2025). Multi-targeted strength properties of recycled aggregate concrete through a machine learning approach. *Engineering Computations*, *42*(1), 388–430.

Mukhtar, F., & Deifalla, A. (2023). Shear strength of FRP reinforced deep concrete beams without stirrups: Test database and a critical shear crack-based model. *Composite Structures*, *307*, 116636.

Pal, B., & Ramaswamy, A. (2023). A multi-physics-based approach to predict mechanical behavior of concrete element in a multi-scale framework. *Mechanics of Materials*, *176*, 104509.

Perceka, W., & Liao, W.-C. (2023). Shear Strength Model for Steel Fiber-Reinforced Concrete Columns. *ACI Structural Journal*, *120*(6).

Prayoonwet, W., Koshimizu, R., Ozaki, M., Sato, Y., Jirawattanasomkul, T., & Yodsudjai, W. (2024). Shear strength prediction for RC beams without shear reinforcement by neural network incorporated with mechanical interpretations. *Engineering Structures*, *298*, 117065.

Qin, X., & Kaewunruen, S. (2025). Surrogated modelling for structural reliability and safety assessments of high-strength concrete beams with industrial and recycled steel fibers. *Engineering Structures*, *322*, 119214.

Reshi, I. A., Shah, A. H., Jan, A., Tariq, Z., Sholla, S., Rashid, S., & Wani, M. U. (2024). Machine learning enhanced modeling of steel-concrete bond strength under elevated temperature exposure. *Structural Concrete*, *25*(6), 4609–4622.

Song, X. (2024). Development of Compressive Fracture Energy Model for Confined Normal-Strength Concrete. *ACI Structural Journal*, *121*(2), 5–18.

Tahir, M., Karam, M. S., Hassam, M., Umer, M., & Hameed, R. (2024). Numerical evaluation of axial compressive behavior of hollow concrete columns reinforced with GFRP bars and spirals. *Structural Concrete*, *25*(1), 566–582.

Vafaei, A., Davoudi-Kia, A., Kutanaei, S. S., & Taslimi, M. (2024). Data-driven investigation of compressive strength of FRP-confined concrete columns using a unified model based on RSM considering interactions between parameters. *Structural Concrete*, *25*(3), 2183–2205.

Wright, S., & Redmond, L. (2023). Methodology for Predicting Concrete Strength after Short-Term Heat Exposure. *ACI Materials Journal*, *120*(6), 33–46.

Yosefi, A., Mojtahedi, F. F., & Bahrami, M. (2025). Identification of Damages in Concrete and Steel Structures: A Comprehensive Review. *Damage Detection and Structural Health Monitoring of Concrete and Masonry Structures: Novel Techniques and Applications*, 207–247.