



## OPEN Construction cost prediction model for agricultural water conservancy engineering based on BIM and neural network

Kun Han<sup>1</sup>, Tieliang Wang<sup>1</sup>, Wenhe Liu<sup>1</sup>, Chunsheng Li<sup>1✉</sup>, Xiaochen Xian<sup>2</sup> & Yingying Yang<sup>2</sup>

Due to the complex construction conditions, long work cycles, and high uncertainty inherent in agricultural water conservancy projects, accurate construction cost prediction is crucial for investment decisions. This study presents an innovative cost prediction model for these projects, integrating BIM with neural networks. Firstly, BIM technology is utilized to digitize and visualize engineering-related information. Subsequently, a prediction model based on SSA optimized PGNN is constructed. The digital data obtained from BIM is subsequently integrated with the prediction model to estimate the construction costs of agricultural water conservancy projects. In this study, actual engineering projects are selected as case studies, utilizing material price data from January 2016 to February 2021 in Liaoning Province, along with real project data for modeling purposes. The results indicate that the maximum relative error between the predicted and actual values of the combined model is only 2.99%. Furthermore, the RMSE and  $R^2$  of the simulated prediction results are 0.1358 and 0.9819, respectively. The proposed model demonstrates higher prediction accuracy and efficiency. Compared with the PGNN model, the RMSE is reduced by 33%, and  $R^2$  is increased by 6%. These findings suggest that the BIM-SSA-PGNN prediction model provides more accurate and efficient construction cost predictions for agricultural water conservancy projects, promoting technological integration and innovation while optimizing construction project costs. This study provides a scientific basis for management to promote the transformation of the industry towards digital and intelligent sustainable development.

**Keywords** Agricultural water conservancy engineering, Construction cost, BIM technology, Sparrow search algorithm (SSA), Grey BP neural network (PGNN)

As agricultural water conservancy engineering projects expand and become more complex, accurate construction cost prediction is essential for evaluating project feasibility and selecting optimal design schemes, which directly influences the economic viability and overall quality of the project. Due to the inherent complexity and uncertainty of these projects, personnel and machinery costs fluctuate, making construction costs difficult to control and often leading to significant cost overruns<sup>1</sup>. Consequently, developing an accurate method for predicting construction costs is essential.

There are two fundamental approaches to predicting construction costs: (1) traditional mathematical and statistical methods and (2) machine learning methods<sup>2,3</sup>. Traditional mathematical and statistical methods encompass linear regression analysis<sup>4,5</sup>, grey prediction<sup>6</sup>, and time series analysis<sup>7,8</sup>, which have notable limitations. Their prediction accuracy is often constrained, and they struggle to model non-linear relationships in data, making it difficult to meet the industry's increasing demands for better accuracy and efficiency. Furthermore, these methods fail to quickly adapt to changes in the market or project progress<sup>9–11</sup>. In contrast, the rapid advancement of artificial intelligence has led to the growing adoption of machine learning (ML) for construction cost prediction. ML is particularly effective at extracting data features from high-dimensional, non-linear datasets<sup>12–14</sup>, enabling it to capture complex relationships between input and output data. By leveraging past project data, ML can generate reliable and swift cost estimates<sup>15–18</sup>. Kim et al. demonstrated that artificial neural networks (ANNs) outperform traditional statistical methods based on limited historical data for long-term forecasts<sup>19</sup>. Cheng et al.<sup>20</sup> highlighted the challenges of predicting construction costs due to fluctuations in building prices. They proposed a hybrid model combining least squares support vector machines

<sup>1</sup>College of Water Conservancy, Shenyang Agricultural University, Shenyang 110866, China. <sup>2</sup>Liaoning Agricultural Development Service Center, Shenyang 110030, China. ✉email: 496954285@qq.com

(LS-SVMs) and differential evolution (DE) to predict the construction cost index, achieving a root mean square error of just 1.354% on the test dataset. This model holds promise as a valuable tool to aid decision-makers in construction management. Chou et al.<sup>21</sup> developed an ANN model to predict project bidding prices, finding that a model with three neurons in the hidden layer outperformed traditional or case-based reasoning methods, with an average absolute percentage error of 13.09%. El Kholy et al.<sup>22</sup> and Tijanić et al.<sup>23</sup> summarized that ANNs represent a promising approach for estimating construction costs, providing more accurate results and reducing estimation errors. Furthermore, ML models have demonstrated significant applicability across diverse domains. For instance, Kennedy et al. employed ML models to predict the compressive strength of concrete and granular sand, achieving robust predictive performance<sup>24,25</sup>.

In the practical application of neural network models, prediction performance is often influenced by the values of hyperparameters. To address this issue, some scholars have adopted intelligent optimization algorithms to enhance the performance of cost prediction models. For instance, AI et al.<sup>26</sup> developed a model for predicting environmental governance costs by optimizing the parameters of a support vector machine using the particle swarm optimization algorithm. Their findings revealed that this method achieved higher prediction accuracy compared to the BP neural network and LSSVM models, making it more suitable for predicting costs prior to environmental governance. More recently, Zheng et al.<sup>27</sup> constructed a random forest (RF) prediction model optimized with the bird swarm algorithm (BSA) to predict construction costs. Compared to backpropagation neural networks and support vector machines, their model demonstrated superior prediction accuracy and efficiency, providing a solid basis for optimizing construction project cost management.

Despite advancements in using artificial neural networks (ANNs) for cost prediction, these models still face challenges in practical engineering applications. There is an urgent need to integrate advanced information technology and artificial intelligence methods to improve the accuracy and real-time performance of engineering cost prediction<sup>28</sup>. In recent years, the emergence of Building Information Modeling (BIM) technology has opened up new possibilities for improving construction cost forecasting. BIM integrates geometric, material, time, and cost data, providing a powerful data source for accurate cost predictions<sup>29,30</sup>. Abanda et al.<sup>31</sup> emphasized that BIM technology can automate the cost estimation process using standard measurement methods, thereby addressing the inaccuracy of cost prediction. Wang et al.<sup>32</sup> explored the feasibility of using BIM for cost estimation in the construction industry, while Yang et al.<sup>33</sup> highlighted the importance of cost analysis in the early stages of construction project planning and proposed a method combining cost estimation models with BIM technology for more accurate cost predictions, supported by case studies. Li et al.<sup>34</sup> combined BIM data with intelligent modeling to improve the coordination of prefabricated shear wall structures, enhancing design and construction integration. Additionally, some scholars have explored the integration of BIM with machine learning or neural networks. For instance, Hong et al.<sup>35</sup> used BIM technology combined with neural network methods to predict net costs. In addition, Abbasnejad et al.<sup>36</sup> further evaluated the effectiveness of integrating BIM with neural networks through mathematical modeling, demonstrating enhanced decision-making efficiency and improved cost prediction accuracy.

Research by various scholars has consistently highlighted the significant potential of neural networks in predicting construction costs. When combined with data from BIM technology, neural networks can further enhance the accuracy and efficiency of these predictions. This integration offers strong theoretical and practical foundations for future applications in architecture and related engineering fields. However, a notable limitation of these studies is their narrow focus on construction projects within a single domain, which limits their universality and applicability. Therefore, there is an urgent need for comprehensive research to develop scientifically sound construction cost prediction models specifically tailored to the field of agricultural water conservancy engineering.

The Grey BP Neural Network (PGNN), a prominent machine learning model, is renowned for its exceptional capabilities to process time series data and identify patterns. It has proven highly effective in addressing complex, nonlinear problems, making it a valuable tool for predicting the temporal fluctuations in construction costs<sup>37</sup>. Furthermore, PGNN has shown remarkable proficiency in predicting building factor prices<sup>38,39</sup>. Consequently, the objective of this study is to develop and validate a construction cost prediction model for agricultural water conservancy projects that integrates BIM and PGNN neural networks, referred to as the BIM-PGNN model. Additionally, the introduction of the Sparrow Optimization Algorithm (SSA) facilitates the optimization of weights and threshold hyperparameters within the PGNN neural network, overcoming issues like local optima and overfitting, which are common in traditional neural networks. This study aims to provide a theoretical basis and technical support for predicting the construction costs of agricultural water conservancy projects.

## Research method

### Digital analysis of intelligent buildings based on BIM technology

Currently, BIM technology involves the creation and use of digital information models for design management, construction management, and operational maintenance of construction projects. This technology integrates building data and information models, enabling efficient information sharing and transmission across the entire lifecycle of a project, from design to operation and maintenance. The comprehensive application of BIM technology throughout the lifecycle of engineering projects represents the current development trajectory of the construction industry and the dominant trend for future advancements. Fully utilizing BIM technology can significantly enhance project management capabilities and production efficiency within construction enterprises, driving sustainable development across various sectors, including water conservancy engineering. BIM technology is characterized by key features such as visualization, coordination, simulation, and optimization. The BIM methodology facilitates the digitization and visualization of building information through a structured workflow encompassing planning and design, conceptual design, preliminary design, analysis, drawing

production, prefabricated components, 4D/5D construction simulation, construction logistics, operation and maintenance, and renovation and demolition.

In this study, BIM5D technology, an extension of BIM, was utilized to extract construction quantities for agricultural water conservancy projects. BIM5D technology primarily relies on the BIM platform to establish a three-dimensional model of water conservancy projects, forming a five-dimensional building information model that integrates 3D (three-dimensional model), 1D (schedule), and 1D (cost budget). The specific construction workflow of the BIM5D model is shown in Fig. 1. In the context of cost prediction for agricultural water conservancy engineering construction, BIM5D technology provides essential data, including project progress, costs, funds, resources, and construction organization, enabling coordinated and shared change information with high visualization accuracy. By dynamically extracting precise data on construction quantities for fixed periods, different materials, and various flow segments, BIM5D technology lays the foundation for accurate cost prediction.

### PGNN neural network model and its optimization

The BP neural network is the most widely used multi-layer feed-forward neural network model in machine learning. Trained on sample data using the error back-propagation algorithm offers advantages such as good fault tolerance and strong associative memory, particularly for solving nonlinear problems and handling limited sample data<sup>40</sup>. This network typically consists of input, hidden, and output layers, with the propagation process primarily divided into forward propagation and back-propagation.

The PGNN model, also known as the Grey BP Neural Network model, combines the Grey GM (1,1) model and the BP Neural Network (BPNN) model. It incorporates the grayscale features of Grey System Theory and the self-learning adaptability of the BP Neural Network, enabling effective handling of nonlinear data sequences<sup>41,42</sup>. The formulation of the time series material unit price prediction problem based on PGNN is as follows: Let the original data sequence of material unit price eigenvalues for  $n$  months be denoted as  $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ . After applying the GM (1,1) model, the predicted unit price sequence is obtained as follows:

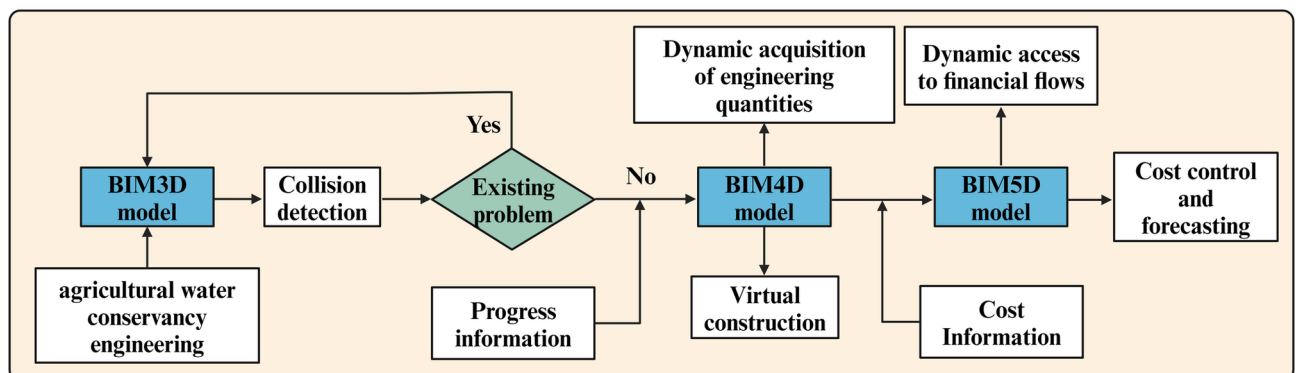
$$x^{(0)}(t+1) = (1 - e^{-\alpha}) \left( x^{(0)}(1) - \frac{\mu}{\alpha} \right) e^{-\alpha t} \quad (1)$$

where  $\alpha$  is the development coefficient; and  $\mu$  is the grey action quantity.

The specific steps for predicting material unit prices using the Grey BP Neural Network model are as follows: ① Preprocess the material unit price data and establish a Grey GM (1,1) model to obtain the unit price prediction sequence; ② Subtract the predicted unit price sequence from the original sequence to obtain the residual sequence of material unit prices; ③ Use the residual sequence as the output sample to train a Grey BP Neural Network model, which generates the unit price correction sequence; ④ Add the predicted unit price sequence to the correction sequence to obtain the final predicted material unit prices using the Grey BP Neural Network model.

The PGNN neural network model offers distinct advantages, particularly in modeling dynamic changes in sequence data, a key feature for intelligent construction cost prediction. Construction costs are influenced by various factors, such as market fluctuations and material price changes, which often exhibit temporal variation. The grey model effectively captures these time series characteristics, providing more accurate and real-time predictions. Furthermore, BP neural networks are highly effective in processing nonlinear and dynamic data, making them particularly useful for addressing challenges in the construction field, where projects often involve complex nonlinear relationships and dynamic fluctuations<sup>43</sup>.

While PGNN neural networks can simulate dynamic data changes and predict the prices of agricultural water conservancy construction materials in real-time, they face challenges in practical applications, such as slow convergence, low computational efficiency, and difficulty in achieving global optimality. To enhance the global optimization ability of the network, the Sparrow Search Algorithm (SSA) is introduced to optimize the PGNN neural network. The SSA algorithm optimizes the weights and thresholds of the PGNN network by simulating the hunting and anti-hunting behavior of sparrow populations, thereby improving the training efficiency and



**Fig. 1.** BIM5D modeling process diagram.

predictive performance of the network<sup>44</sup>. Additionally, the SSA algorithm avoids the random assignment of network weights and thresholds, guiding the network away from local minima toward a global or near-global optimal solution<sup>45,46</sup>.

This algorithm adopts two behavioral strategies: searching and following. Searchers actively seek food sources, while followers obtain food from the searchers. When sparrows detect a predator, they emit an alarm signal and move to a safer location. Sparrows that find better food sources may transition to searchers, while the overall ratio of searchers to followers remains constant.

#### 1) Search behavior.

Searchers with higher fitness values are responsible for locating food and guiding the entire population's actions. They find food more quickly and cover a larger search area than followers during the hunting process. The location of the searcher is updated using formula (2):

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot iter_{max}}\right) & \text{if } R_2 < ST \\ X_{i,j}^t + Q \cdot L \cdot fR_2 & \text{if } R_2 \geq ST \end{cases} \quad (2)$$

where  $X_{i,j}$  represents the position of the  $j$ -th dimension of the  $i$ -th search sparrow;  $t$  is the current iteration number;  $\alpha$  is a random number between 0 and 1;  $iter_{max}$  is the maximum number of iterations of the algorithm;  $Q$  is a random number that following a normal distribution;  $L$  is a  $1 \times d$ -dimensional matrix;  $R_2$  is the warning value within the range of 0 to 1; and  $ST$  is the safe value ranging from 0.5 to 1.

#### 2) Following behavior.

Followers constantly monitor the searchers during the hunting process and compete for high-quality food discovered by the searchers. The position of the follower is generated by formula (3):

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{worst}^t - X_{i,j}^t}{i^2}\right) & \text{if } i > \frac{n}{2} \\ X_P^{t+1} + |X_{i,j}^t - X_P^{t+1}| \cdot A^+ & \text{Otherwise} \end{cases} \quad (3)$$

where  $X_{worst}^t$  represents the worst position of the searcher;  $X_P^{t+1}$  represents the best position of the searcher; and  $A$  is a  $1 \times d$ -dimensional matrix randomly assigned  $-1$  or  $1$ .

Some sparrows in the population serve as "warning signs," guiding others to safer hunting areas. These sparrows position themselves between the edge and the safe zones. The location of a warning sparrow is randomly generated according to formula (4):

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta \cdot |X_{i,j}^t - X_{best}^t| & \text{if } f_i > f_g \\ X_{i,j}^t + K \cdot \left(\frac{|X_{i,j}^t - X_{worst}^t|}{(f_i - f_w) + \varepsilon}\right) & \text{if } f_i = f_g \end{cases} \quad (4)$$

where  $X_{best}^t$  is the optimal position of the sparrow population;  $\beta$  is a random number that follows a normal distribution (0,1);  $K$  is a random number with a value range of  $-1-1$ ;  $\varepsilon$  is a constant.  $f_i$  is the current individual fitness value of the sparrow in the algorithm;  $f_g$  is the optimal fitness value of the sparrow in the algorithm; and  $f_w$  is the worst fitness value.

The process of optimizing the PGNN neural network using the SSA algorithm is illustrated in Fig. 2, and the specific steps involved are as follows:

1) Neural Network Structure Determination and Initialization Parameter Setting: Establish the population size, maximum iteration count, the proportion of searchers and followers within the population, and the warning values. The topology of the PGNN neural network is determined based on the input and output data from the training unit price sample.

2) Fitness Function Calculation: The average mean square error between the output values of the training and testing samples and the expected values is used as the fitness value for the sparrow algorithm to determine the current optimal initial position:

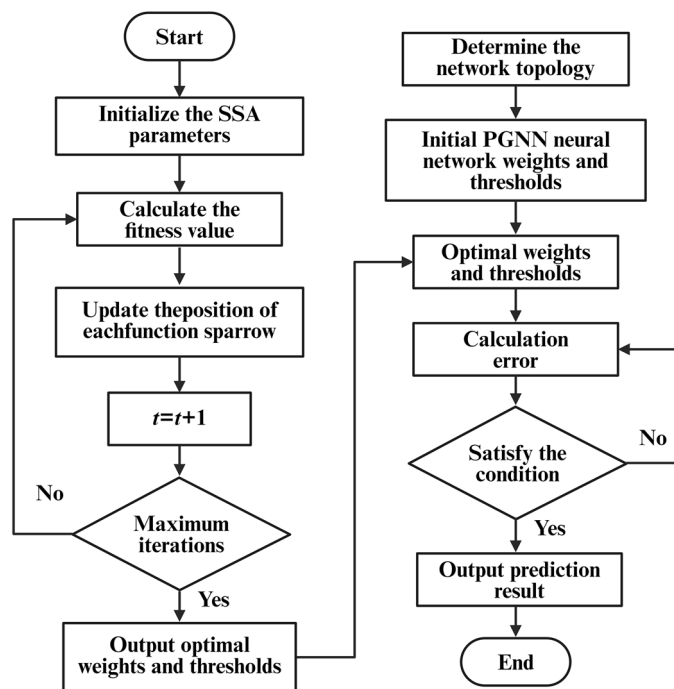
$$fitness = \frac{\sum_{i=1}^n [(Y_i - Z_i)^2 + (Y_i' - Z_i')^2]}{2n} \quad (5)$$

where  $n$  is the number of samples;  $Y_i$  and  $Z_i$  are the output and expected values of the training samples for unit price; and  $Y_i'$  and  $Z_i'$  are the output and expected values of the testing samples for unit price.

3) The iterative process is initiated to adjust the weights and thresholds. Continuously update the position coordinates of each functional sparrow using Eqs. (2)-(4), and optimize the optimal individual position of the population at various iterations. Use the best individual and the global optimal solution from the sparrow population in the current iteration to determine the weights and thresholds of the PGNN model.

4) Network Training: Repeat steps (2) to (3) until the maximum number of iterations is reached, outputting the minimum fitness value to obtain the optimal weights and thresholds for the network. Alternatively, continue the process until either the iteration limit is reached or the training results meet the specified error accuracy requirements.

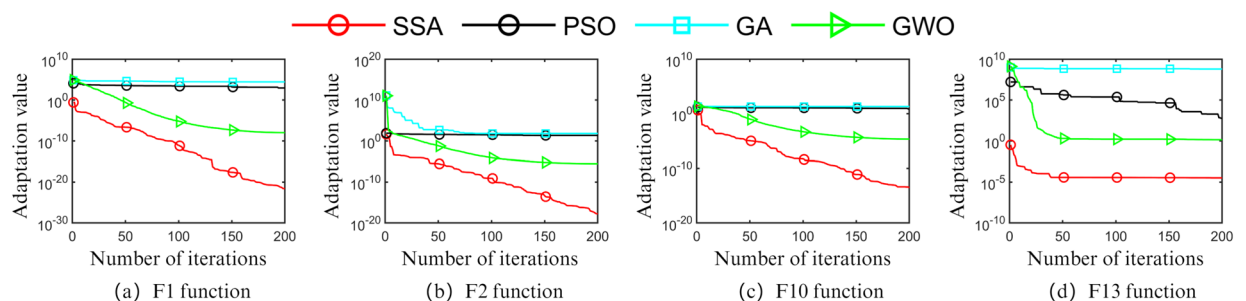
To verify the superiority of the Sparrow Search Algorithm (SSA) in hyperparameter optimization, performance tests were conducted using benchmark functions, including unimodal test functions (F1, F2) and high-dimensional multimodal functions (F10, F13). The unimodal functions were selected to evaluate the exploitation ability of the algorithm, i.e., rapidly locating the optimal solution, while the high-dimensional



**Fig. 2.** Flowchart of PGNN model optimized by SSA algorithm.

Basis function	Range of $x$ values	Dimension $n$	Theoretical optimum
$F_1 = \sum_{i=1}^n x_i^2$	$[-100, 100]$	30	0
$F_2 = \sum_{i=1}^n  x_i  + \prod_{i=1}^n  x_i $	$[-10, 10]$	30	0
$F_{10} = -20 \sum_{i=1}^n \exp \left( -0.2 \sqrt{1/n} \sum_{i=1}^n x_i^2 \right) - \exp \left( \sum_{i=1}^n \cos(2\pi x_i) / n \right) + 20 + e$	$[-32, 32]$	30	0
$F_{13} = 0.1 \left\{ \sin^2(3\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_i + 1)] + (x_n - 1) [1 + \sin^2(2\pi x_n)] \right\} + \sum_{i=1}^n u(x_i, 5, 100, 4)$	$[-50, 50]$	30	0

**Table 1.** Selected benchmark function.



**Fig. 3.** Convergence curves for different algorithms.

multimodal functions aimed to assess its exploration capability for avoiding local optima. The mathematical definitions of these benchmark functions are provided in Table 1. This study compares SSA with three other metaheuristic algorithms: Genetic Algorithm (GA)<sup>47</sup>, Particle Swarm Optimization (PSO)<sup>48</sup>, and Whale Optimization Algorithm (WOA). To ensure statistical robustness, all algorithms were configured with identical parameters: a population size of 30, a maximum iteration count of 200, and 20 independent runs. The average results were adopted as final performance metrics, with convergence curves illustrated in Fig. 3.



As illustrated in Fig. 3, SSA demonstrated superior performance in two critical dimensions: (1) For unimodal test functions (F1, F2), SSA exhibited faster convergence speed and final solutions closer to the theoretical optimum, confirming its enhanced optimization efficiency; (2) In high-dimensional multimodal functions (F13, F15) characterized by complex search spaces, SSA achieved global optima with minimal iterations, whereas comparative algorithms frequently stagnated in local optima or exhibited slower search dynamics. These empirical findings substantiate that SSA effectively balances exploration and exploitation mechanisms compared to GA, PSO, and WOA, thereby improving both search precision and convergence acceleration, which contribute to elevated predictive accuracy in subsequent model.

### Construction cost prediction model for agricultural water conservancy engineering based on BIM and SSA-PGNN

To effectively predict the construction cost of agricultural water conservancy projects, this study established a prediction model that integrates BIM technology and the SSA-PGNN neural network. After normalizing the data, the model consists of two key components: extracting actual engineering quantities using BIM and constructing an SSA-PGNN unit price prediction model. Firstly, a 3D model of the agricultural water conservancy engineering is established. Based on this model, the scheduling plan and cost budget are imported to construct the BIM5D database, and the required engineering quantities are extracted based on specific criteria. Subsequently, the PGNN neural network model, optimized using the SSA algorithm, is utilized to predict the unit price for labor and machinery, facilitating real-time, dynamic cost predictions for agricultural water conservancy projects. The specific process framework is shown in Fig. 4.

### Case study and result analysis

#### Data acquisition and pre-processing

To assess the performance of the proposed construction cost prediction model for agricultural water conservancy projects based on BIM and the SSA-PGNN neural network, an agricultural water conservancy project in Yanghe Town, Anshan City, Liaoning Province, was used as a case study. The primary focus of our study was on the construction period of specific structures, including "mortar masonry rectangular groove", "concrete road (3 m wide)", "agricultural bridge (6 m × 4 m × 1.5 m)", "agricultural bridge (6 m × 6 m × 2 m)", "agricultural culvert (φ 1000 × 4 m)", "agricultural culvert (φ 600 × 6 m)", and "high mortar masonry retaining walls (1.2 m and 1.5 m)". Due to the substantial contribution of steel reinforcement and concrete to the direct construction cost, the study concentrated on evaluating the total construction cost of grade III seismic steel bars (HRB400 φ 18–25 mm) and C25 and C30 concrete materials. The data was sourced from open data platforms of the government and industry, encompassing a total of 62 material information prices for C25, C30 concrete, and HRB400 steel bars in Liaoning Province from January 2016 to February 2021, as shown in Fig. 5.

The sample input data comprised the prices of concrete and steel bars over the first 12 months, with the corresponding output data representing the price in the 13th month. A PGNN neural network model, optimized using the SSA algorithm, was then developed. The 62 material information price data points were divided into 50 groups, with the *randperm* function used to randomly assign 40 groups as the training set and 10 groups as the testing set in the ratio of 8:2. Additionally, K-fold cross-validation with K=5 was implemented on the training set, where different subsets of the training data were alternately utilized for model training and validation, thereby enabling a more reliable assessment of the model's generalization performance.

The original unit price data were normalized to the interval [0, 1] using Min–Max normalization as specified in Eq. (6), a preprocessing step implemented to mitigate significant errors caused by substantial discrepancies in input data magnitudes. This transformation accelerated both the convergence rate and prediction accuracy of the

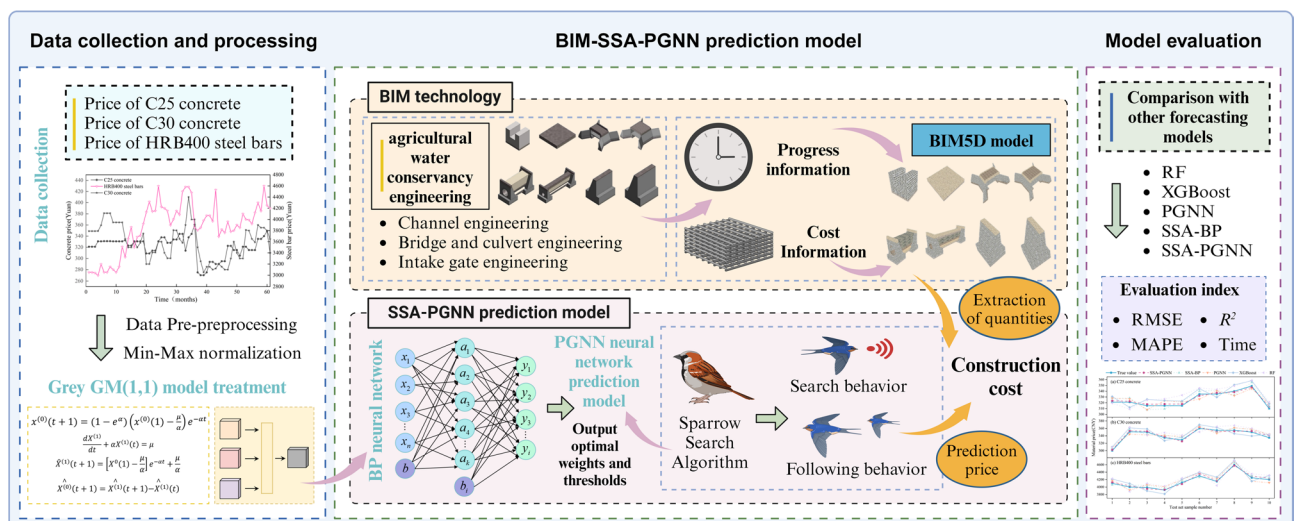
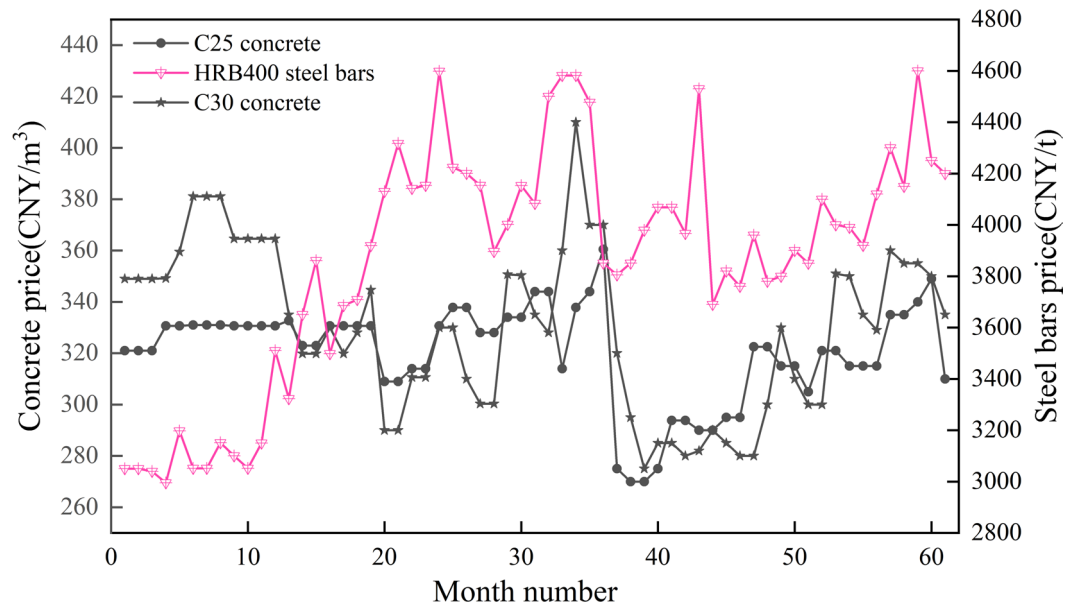


Fig. 4. Analysis and modelling process framework.



**Fig. 5.** Monthly average price chart of concrete and steel bars.

training network by eliminating dimensional heterogeneity and standardizing data scales across measurement units.

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (6)$$

where  $x'_i$  is the normalized data;  $x_i$  is the unit price data before normalization;  $x_{max}$  is the maximum value of each group of unit price data, and  $x_{min}$  is the minimum value.

### Extracting agricultural water conservancy engineering quantities based on BIM technology

Revit design software is currently the most widely used BIM modeling platform, supporting tasks such as 3D modeling, 3D visualization, and interaction between drawings and models. The family library function enables the integration of pre-built components from various projects, thereby reducing the modeling workload through the utilizing of shared family libraries and improving modeling efficiency. Revit-generated 3D models of individual agricultural water conservancy projects are shown in Fig. 6a.

After establishing the Revit3D models for each individual project, the 3D models are exported as E5D files using the BIM5D plugin and imported into the digital project management platform. Individual engineering models of roads, bridges, and culverts are imported, as shown in Fig. 6b. Following the import process, project schedules from the scheduling plan and cost budgets from pricing software are integrated into the platform to establish the BIM5D model.

The construction simulation function in the BIM+technology management system simulates the planned construction period for each project. It dynamically displays the construction status of each project component over time, providing real-time updates to facilitate progress adjustments and ensure timely completion.

After completing the simulation construction, the BIM5D platform allows dynamic querying of quantities for each project component. The BIM5D platform extracts data for all components, classifies them, and efficiently determines the required quantities. Associating construction progress with the model facilitates real-time calculation of work quantities, material usage, and flow period in real-time. Table 2 shows the usage of steel bars and concrete for each project during the construction period from March 20th to September 20th, 2021, as extracted from the BIM5D platform.

### Parameter setting and model evaluation

In this study, the computer configuration used for network training and testing is based on Matlab2023b software running on a Windows 10 64 bit operating system environment. The system is equipped with an NVIDIA GeForce GTX1650 graphics card with 4 GB of graphics memory, 16 GB of RAM, a 1 TB of hard drive, and an Intel (R) Core (TM) i5-9300H CPU operating at 2.40 GHz.

Firstly, the data are normalized to generate a training dataset with random numbers between 0 and 1. This preprocessing step minimized material cost data oscillation during gradient updates, accelerated network convergence, and reduced training time. Based on empirical equations, the number of hidden layer nodes was calculated using a trial-and-error approach, with the selection range determined to be<sup>5,14</sup>. Simulation experiments were performed with varying node counts to determine the optimal number that minimized deviation. Table 3 shows the average relative error values for different node counts after network training, indicating that the grey

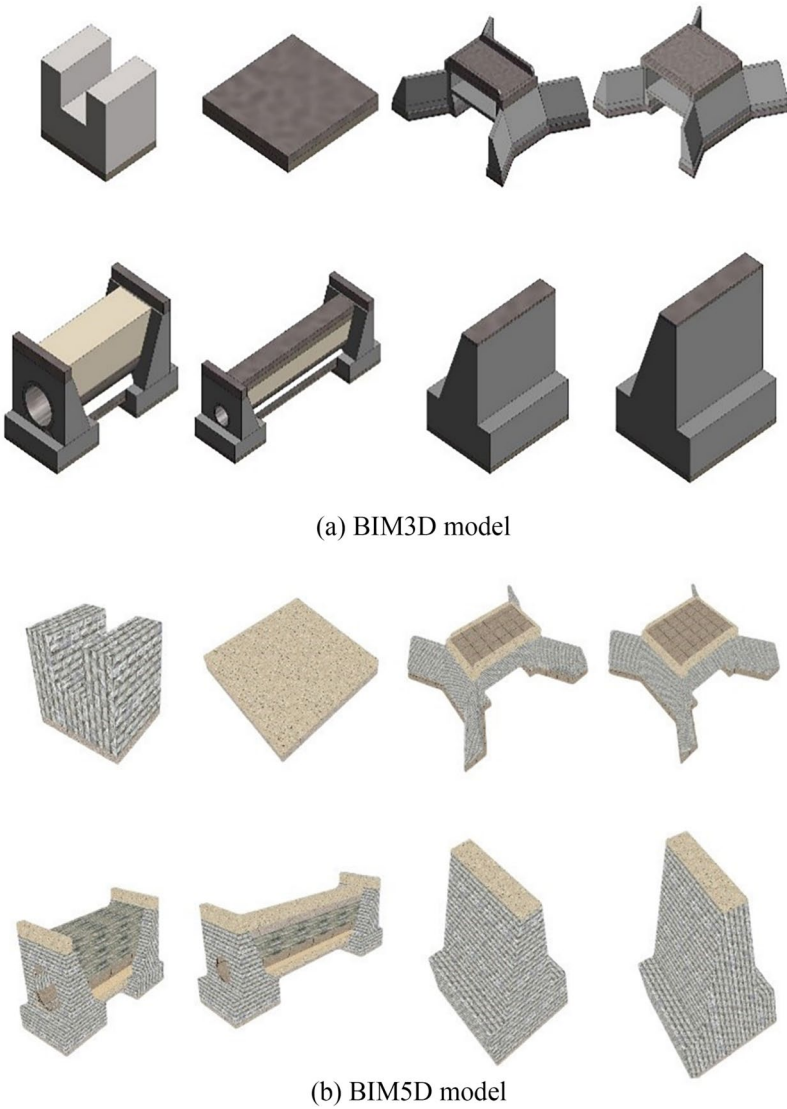


Fig. 6. BIM model of each project.

Start time	End time	Amount of HRB400 rebar (t)	Amount of C25 concrete (m <sup>3</sup> )	Amount of C30 concrete (m <sup>3</sup> )
2021.03.20	2021.03.31	0	819.00	0
2021.04.01	2021.04.30	2.72	85.58	20.81
2021.05.01	2021.05.31	3.00	72.32	21.53
2021.06.01	2021.06.30	3.07	65.54	24.60
2021.07.01	2021.07.31	1.99	664.42	8.25
2021.08.01	2021.08.31	0	732.00	0
2021.09.01	2021.09.20	0	942.00	0

Table 2. Quantity of reinforcement and concrete in each project.

Number of nodes in the hidden layer	5	6	7	8	9	10	11	12	13	14
Average relative error(%)	5.25	3.88	5.38	5.79	8.23	9.47	7.79	8.87	3.63	7.22

Table 3. Average relative error values corresponding to the number of nodes in different hidden layers.



Serial number	Population		Iteration	
	Population size	Algorithm error	Number of iterations	Algorithm error
1	10	0.018436	80	0.003069
2	20	0.016377	90	0.001163
3	30	0.012069	100	0.000514
4	40	0.012239	110	0.000640
5	50	0.020448	120	0.000614
6	60	0.017388	130	0.001142
7	70	0.022617	140	0.001163
8	80	0.035267	150	0.001758

**Table 4.** Comparison of algorithm errors for different population sizes and number of iterations.

Model	Hyper-parameters	Value
BP	Number of iterations	2000
	Learning rate	0.01
	Number of hidden layer nodes	11
	Minimum error of training objective	0.00001
	Learning function	trainlm
RF	Estimator	100
	Max depth	5
XGBoost	Estimator	500
	Max depth	3
	Learning rate	0.01
SSA	Early warning value	0.6
	Population size	30
	Maximum number of iterations	100
	Upper boundary of weight threshold	5
	Weight threshold lower boundary	−5
	Proportion of searchers	0.7
	Proportion of early warning personnel	0.2

**Table 5.** Quantity of reinforcement and concrete in each project.

BP neural network achieved the best fitting performance when the hidden layer nodes were set to 13. Based on these results, the PGNN neural network topology was finalized as 12–13–1.

The runtime performance of the SSA algorithm depends on the population size and the number of iterations. The population size was tested within the range of 10 to 80, and the number of iterations ranged from 80 to 150. The fitness function error, defined as the mean square error (MSE) of the training set, was used to identify the optimal parameters. As shown in Table 4, the minimum errors of 1.2069 and 0.0514 were achieved when the population size was set to 30 and the number of iterations to 100. At these values, the algorithm converged.

To evaluate the performance of the SSA- PGNN model proposed in this study, Random Forest (RF), XGBoost, PGNN and SSA-BP models were established as comparative models using the same dataset. Specifically, Random Forest (RF) employs an ensemble of multiple decision trees, where the final output is derived from averaging all individual tree predictions, effectively mitigating overfitting while ensuring stable predictions. XGBoost, as a representative gradient boosting framework-based ensemble learning algorithm, achieves high-precision predictions through the integration of multiple weak learners, demonstrating particular efficacy in processing large-scale datasets. These two models were selected as baseline representatives of traditional machine learning (ML) approaches for comparative validation. All machine learning models underwent systematic parameter optimization through a trial-and-error approach to identify optimal parameter combinations, and the specific parameter settings for each model are shown in Table 5.

Finally, to assess the model's performance, we use the following metrics: root mean square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination ( $R^2$ ). Lower RMSE and MAPE values indicate higher prediction accuracy of the model.  $R^2$  values range from 0 to 1, and values closer to 1 indicate a better fit between the predicted and actual values, suggesting superior model performance. The formulas for calculating RMSE, MAPE and  $R^2$  are provided below:

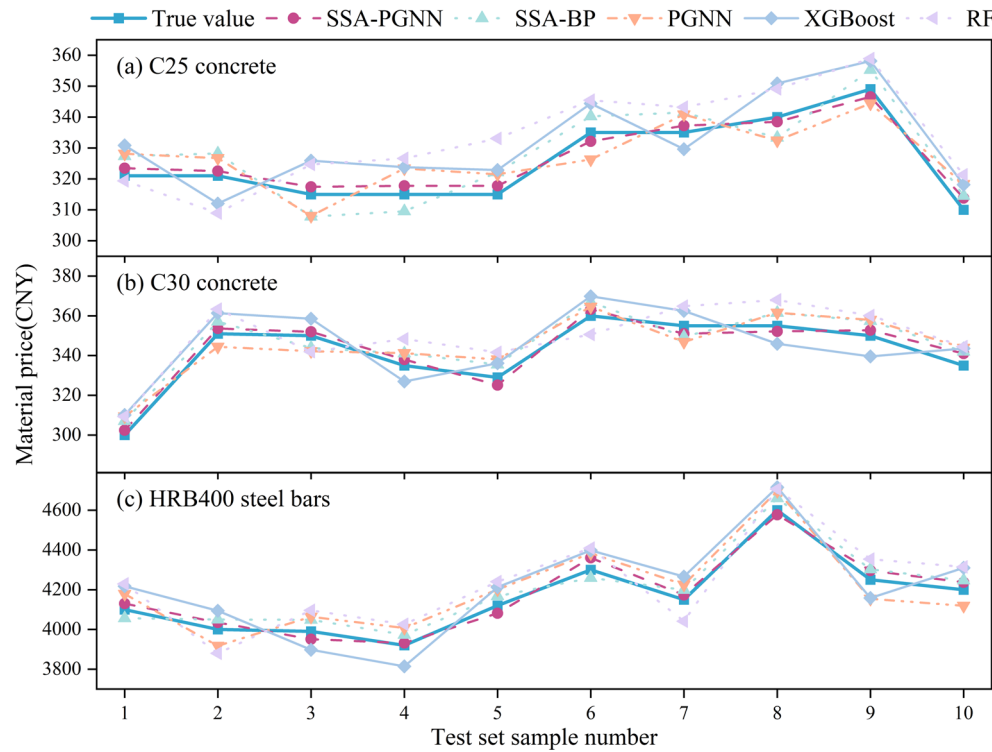


Fig. 7. Comparison between predicted and actual values of different model.

Model	C25 concrete				C30 concrete				HRB400 steel bars			
	MAPE	RMSE	R <sup>2</sup>	Time/s	MAPE	RMSE	R <sup>2</sup>	Time/s	MAPE	RMSE	R <sup>2</sup>	Time/s
RF	6.27%	0.452	0.861	45.1	6.73%	0.167	0.873	43.5	6.31%	0.172	0.887	42.5
XGBoost	5.23%	0.424	0.893	40.7	5.71%	0.142	0.902	41.4	5.44%	0.154	0.909	38.6
PGNN	4.42%	0.395	0.939	69.7	4.90%	0.132	0.947	68.0	5.31%	0.143	0.958	65.3
SSA-BP	3.95%	0.216	0.961	30.4	4.73%	0.097	0.970	32.6	3.60%	0.125	0.986	30.4
SSA-PGNN	3.29%	0.140	0.979	29.8	4.04%	0.079	0.983	30.7	2.43%	0.114	0.992	28.7

Table 6. Comparison of performance evaluation index of different models.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(Y_i - \hat{Y}_i\right)^2} \tag{7}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100\% \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^n \left(Y_i - \overline{Y}_i\right)^2} \tag{9}$$

where  $Y_i$  is the real values;  $\hat{Y}_i$  is the predicted values; and  $\overline{Y}_i$  is the average of the predicted values.

Results and discussion

Prediction results for both the training and testing sets of all five models demonstrated no significant signs of overfitting or underfitting. In the unit price prediction of C25 concrete, C30 concrete, and HRB400 steel bars, the predicted values of the five models were compared with the true values, and the results are depicted in Fig. 7. The comparison of performance evaluation indicators for various prediction models is shown in Table 6.

Figure 7 demonstrates that the proposed SSA-PGNN model provides superior fitting to the nonlinear changes in unit prices for agricultural water conservancy materials, including C25 concrete, C30 concrete, and

HRB400 steel bars, compared to the PGNN and SSA-BPNN models. Its predicted unit price curves align more closely with actual values, achieving the highest degree of fitting and the most accurate prediction performance. Furthermore, the statistical results of the performance evaluation indicators in Table 6 indicate that the SSA-PGNN prediction model proposed in this study has the smallest RMSE and MAPE values, with an  $R^2$  value closest to 1, thus outperforming other prediction models. Under the same conditions, the PGNN model demonstrates superior predictive performance compared to traditional ML models (RF, XGBoost), indicating that PGNN exhibits smaller overall prediction deviations and more stable accuracy, particularly showing significant advantages in handling nonlinear temporal data such as material unit prices. Simultaneously, the SSA-PGNN model exhibits significant improvements in performance evaluation metrics compared to the SSA-BPNN model. For the C25 concrete unit price prediction model, its RMSE and MAPE decreased by 35% and 17%, respectively, while  $R^2$  increased by 2%. In the C30 concrete unit price prediction model, its RMSE and MAPE decreased by 15% and 18%, respectively, while  $R^2$  increased by 1%. In the HRB400 steel reinforcement unit price prediction model, its RMSE and MAPE decreased by 9% and 32%, respectively, while  $R^2$  increased by 1%. These results confirm that incorporating grey GM(1,1) processing of historical material price data into the BP neural network effectively extracts temporal fluctuation trends, mitigates data randomness, and achieves higher prediction accuracy and operational efficiency by synergistically correcting nonlinear errors through neural networks.

Under the same conditions, the SSA- PGNN model shows significant improvements in evaluation metrics such as RMSE, MAPE, and  $R^2$  compared to the PGNN model. In the C25 concrete unit price prediction model, its RMSE and MAPE decreased by 64% and 25%, respectively, while  $R^2$  increased by 4%. In the C30 concrete unit price prediction model, its RMSE and MAPE decreased by 39% and 18%, respectively, while  $R^2$  increased by 4%. In the HRB400 steel reinforcement unit price prediction model, its RMSE and MAPE decreased by 20% and 54%, respectively, while  $R^2$  increased by 3%. Regarding computational costs, training durations for non-optimized models (RF, XGBoost, PGNN) ranged between 40–60 s, whereas SSA-optimized models (SSA-BP, SSA-PGNN) achieved reduced training times of approximately 10 s. These improvements highlight the importance of using optimization algorithms to optimize the weights and thresholds of the PGNN model, which significantly enhances predictive performance and efficiency, as randomly generated hyperparameters can negatively impact the model's generalization ability, training times and prediction accuracy, preventing it from reaching optimal performance.

This study employed SPSS software to conduct independent samples  $t$ -tests on the prediction results of the SSA-PGNN and PGNN models, with the confidence interval set at 95% for significance analysis of test set results. The computational results for different materials are presented in Table 7. We observed that the Sig values for all three materials exceeded 0.05, indicating that the prediction results of the two models met the homogeneity of variance test. Subsequently, the  $t$ -test results revealed that the sig (Bilateral) values were all below 0.05, demonstrating a statistically significant difference between the predicted values of the two models.

In summary, the comparative analysis demonstrates that the SSA-PGNN model outperforms other models in key metrics such as prediction accuracy, RMSE, and  $R^2$ . The SSA-PGNN prediction model has better accuracy and stability in unit price prediction, confirming its scientific validity.

The prediction results of HRB400 steel bars, C25 concrete, and C30 concrete unit prices from March 2021 to September 2021, based on the established SSA-PGNN prediction model, are shown in Fig. 8. Combined with the steel bars and concrete consumption extracted from the BIM5D platform, the main material cost information of the actual project was obtained. Subsequently, the BIM-SSA-PGNN model developed in this study was utilized to predict the construction cost of the actual project in Yanghe Town, Anshan City, Liaoning Province. These predictions were then compared with those of the PGNN model and the SSA-BPNN model, with the results presented in Table 8.

Performance analysis of the BIM-SSA-PGNN model revealed that its prediction accuracy remained consistently high across different months, achieving a MAPE of only 2.99%. This accuracy surpasses that of the BIM-ANN model proposed by Zhang et al.<sup>49</sup>, which reported a MAPE of 4.29% when predicting the price of HRB400 steel bars. The data spanned from April 2019 to October 2022, while this study employed a broader temporal range, demonstrating the importance of data sample size for neural network model performance. This also validates the perspective of Cheng et al.<sup>20</sup>, who emphasized that construction prices fluctuate over time, leading to difficulties in predicting construction costs, thus highlighting the criticality of prediction models'

Material	Assumption	Levene test of variance equation		t-test pf mean–variance				
		F	Sig	t	df	sig (Bilateral)	Mean difference	Standard error
C25 concrete	Variance is equal	0.203	0.658	-2.630	18.000	0.017	-16.419	6.243
	Variance inequality	–	–	-2.630	15.229	0.019	-16.419	6.243
C30 concrete	Variance is equal	1.952	0.179	-2.508	18.000	0.022	-16.146	6.437
	Variance inequality	–	–	-2.508	14.106	0.025	-16.146	6.437
HRB400 steel bars	Variance is equal	2.718	0.117	4.201	18.000	0.001	297.914	70.921
	Variance inequality	–	–	4.201	13.633	0.001	297.914	70.921

Table 7. The calculation results of t-test.

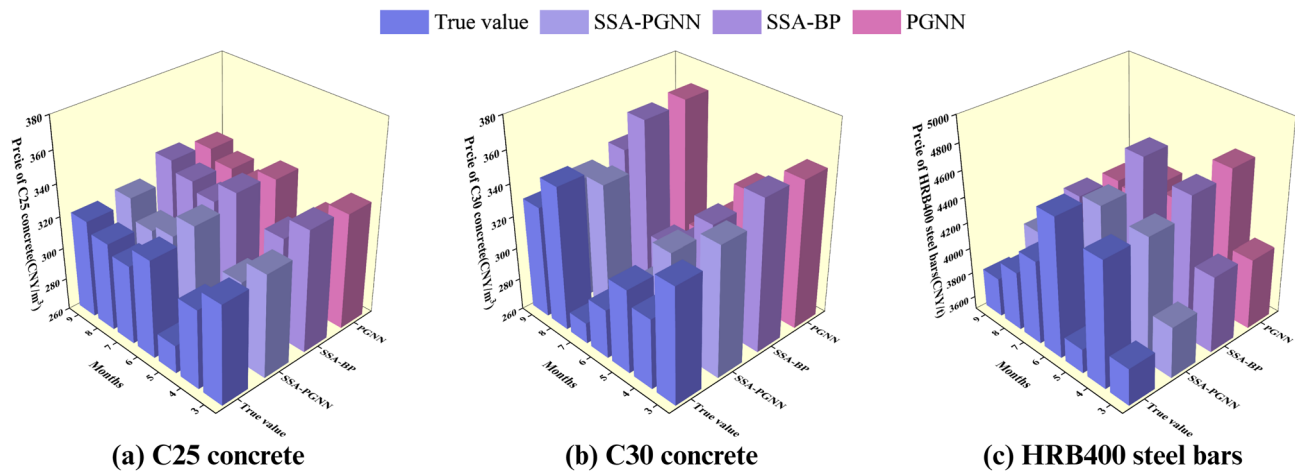


Fig. 8. Comparison of PGNN, SSA-BP and SSA-PGNN prediction results.

Month	True value	SSA-PGNN prediction value	SSA-BP prediction value	PGNN prediction value
3	263,324.88	265,356.00	274,987.44	272,039.04
4	44,951.66	44,975.43	46,698.14	46,406.48
5	37,861.71	38,043.00	39,309.70	39,303.79
6	42,269.57	42,449.07	43,860.44	42,221.06
7	215,982.71	219,069.40	225,061.55	223,178.05
8	230,316.48	226,773.60	240,520.56	237,943.92
9	302,193.60	304,266.00	315,579.42	312,197.64
MAPE		2.99%	3.88%	5.15%
R <sup>2</sup>		0.9819	0.9643	0.9290
RMSE		0.1358	0.1683	0.2041

Table 8. Comparison of construction cost prediction results of different models.

capability to capture temporal data variations. On the other hand, Baduge et al.<sup>50</sup> argued that neural network models could be extensively applied in performance prediction of construction materials including concrete, steel, and timber to optimize cost-effectiveness. These findings substantiate the feasibility and applicability of the BIM-SSA-PGNN prediction model established in this study, particularly in generating satisfactory outcomes for material price forecasting.

Furthermore, the BIM-SSA-PGNN model demonstrates optimal performance among the compared models, with an  $R^2$  value of 0.9819 and an RMSE of 0.1358, underscoring its stability in cost prediction, effectiveness in capturing cost fluctuations, and viability for supporting intelligent construction cost management. Compared with the SSA-BP model, this model reduced RMSE by 19% and improved  $R^2$  by 2%, demonstrating superior performance. This conclusion aligns with the viewpoints of Pham et al.<sup>51</sup> and Alshboul et al.<sup>52</sup>, indicating that the proposed model effectively captures construction cost fluctuations and explains that such variations primarily originate from historical price volatility.

Therefore, the BIM-SSA-PGNN model exhibits significant advantages over traditional methods in construction cost prediction. The digital and visualization capabilities of BIM technology enable accurate extraction of engineering data information from projects. Combined with the rapid processing capacity of the SSA-PGNN model for time-series data, it achieves precise prediction of construction costs for agricultural water conservancy projects, providing the agricultural water conservancy industry with more accurate and reliable cost management tools.

Practical application and future work  
Practical application

Management application aspect. The proposed BIM-SSA-PGNN model provides a scientific decision-making tool for the whole-lifecycle management of agricultural water conservancy projects. By integrating engineering data (e.g., quantities, schedules, and resource demands) through BIM5D technology with the predictive capabilities of the SSA-PGNN model, managers can dynamically generate cost baselines during the planning phase, monitor cost deviations in real time during construction, and adjust resource allocation strategies based on prediction results. Specifically, the model can combine construction schedule simulations to predict the cost impacts of different schedule compression schemes, assisting managers in achieving multi-objective

optimization among quality, cost, and schedule. Furthermore, the high-precision prediction results generated by the model can support budget preparation and contract risk allocation during the bidding phase, reducing claims risks caused by cost estimation deviations and enhancing the standardization and foresight of project management processes.

**Technical application aspect.** The technical value of this model lies in the innovative integration of BIM and machine learning. By automatically extracting structured quantity data (e.g., concrete volume, steel reinforcement specifications) via BIM5D technology and combining it with the SSA-optimized PGNN network, an end-to-end automated workflow from data acquisition to predictive analysis is achieved. The model synchronizes BIM component attributes with real-time market price databases, allowing technical teams to visualize the impact weights of material price fluctuations on costs and inversely optimize designs. This integration overcomes limitations in traditional cost estimation, enabling self-updating prediction models with enhanced real-time adaptability.

**Economic application aspect.** The model effectively controls hidden cost overrun risks in agricultural water conservancy projects through accurate predictions. The SSA-PGNN model can forecast material price fluctuations across different construction seasons, helping financial departments optimize cash flow scheduling and reduce financial costs caused by advance project payments.

**Scalability aspect.** By adjusting BIM models and parameters, the model can adapt to diverse engineering scenarios such as bridges, tunnels, and residential buildings. Combined with real-time market prices of materials and machinery in different regions, it enables cross-engineering domain and cross-regional application capabilities. For instance, by constructing a BIM model for residential building projects, concrete quantity data can be automatically extracted, and the SSA-PGNN model can predict concrete prices in the target region, ultimately deriving concrete costs for regional residential projects.

### Future work

Dataset quality and quantity critically influence machine learning model performance. Limited by the current dataset size, this study has not yet employed deep learning models such as LSTM and Transformer, which excel in handling complex, large-scale datasets. Future studies should prioritize systematic collection and organization of engineering-related data to improve cost prediction accuracy using deep learning. Integrating SHAP value analysis could further visualize the contribution of engineering features to cost predictions, enhancing project planning and management.

Additionally, future research will focus on integrating IoT with BIM models, with a particular emphasis on exploring real-time data interaction mechanisms among on-site sensor networks, inspection systems, and BIM5D platforms. To achieve this, researchers should develop a predictive system architecture capable of synchronizing multi-source heterogeneous data—such as construction progress, resource consumption, and market price fluctuations—thereby supporting dynamic cost control throughout project lifecycles. By extending this framework to diverse project types, intelligent prediction systems can be established across engineering domains, ultimately advancing the construction industry toward real-time collaboration and autonomous decision-making.

### Conclusion

This study addresses the complexity of predicting construction costs in agricultural water conservancy projects by developing a construction cost prediction model based on BIM technology and a PGNN neural network. The model's accuracy and generalization ability are further enhanced through the integration of the SSA algorithm. The following conclusions are drawn from case studies:

- 1) By utilizing BIM technology to create a 3D model of an actual project, introducing the schedule plan and cost budget to establish the BIM5D model, and dynamically extracting key engineering quantity information within the required time period, a reasonable adaptation between BIM5D technology functions and cost prediction task requirements is achieved.
- 2) To address the nonlinear and periodic characteristics of unit price fluctuations in agricultural water conservancy projects, the grey GM (1,1) model combined with the BP neural network was selected as the prediction model. The SSA algorithm is employed to swiftly search for model weights and thresholds, adjusting the population positions to minimize errors, resulting in an optimized process. Experimental results demonstrate that the SSA-PGNN model outperforms the PGNN model in prediction accuracy.
- 3) Using the agricultural water conservancy project in Yanghe Town, Anshan City, Liaoning Province, as a case study, the BIM-SSA-PGNN model was employed to quickly and accurately process relevant engineering quantity data. The model's performance yielded an average prediction accuracy of 97.01% across different months, with an RMSE of 0.1358 and an  $R^2$  of 0.9819. The model evaluation indicators surpass those of the PGNN model and SSA-BPNN model, verifying its high accuracy and reliability. This model can provide reference and technical support for predicting the construction cost of agricultural water conservancy projects, while enhancing the efficiency and effectiveness of construction cost management and project decision-making.
- 4) The proposed construction cost prediction model has significant practical implications for the agricultural water conservancy industry. The model effectively captures the dynamic trends in cost changes, accounting for complex factors such as design modifications and market price fluctuations, while enhancing the real-time accuracy of predictions. This model allows project managers to optimize project planning, adjust labor, modify work hours, and reallocate tasks as needed, thereby providing more comprehensive decision support and optimization solutions.



## Data availability

The Supplementary Materials include all data. Some or all of the data, models, or codes that support the findings of this study are available from the first author Kun Han (email: 2,373,903,326@qq.com) upon reasonable request.

Received: 12 December 2024; Accepted: 2 July 2025

Published online: 07 July 2025

## References

- Annamalaisami, C. D. & Kuppuswamy, A. Reckoning construction cost overruns in building projects through methodological consequences. *Int. J. Constr. Manag.* **22**, 1079–1089. <https://doi.org/10.1080/15623599.2019.1683689> (2022).
- Chakraborty, D., Elhegazy, H., Elzarka, H. & Gutierrez, L. A novel construction cost prediction model using hybrid natural and light gradient boosting. *Adv. Eng. Inform.* **46**, 101201. <https://doi.org/10.1016/j.aei.2020.101201> (2020).
- Han, K. et al. Predicting the construction cost of high standard farmland irrigation projects using NGO-CNN-SVM. *Trans. Chin. Soc. Agric. Eng. (Trans. CSAE)*. **40**, 62–72. <https://doi.org/10.11975/j.issn.1002-6819.202312025> (2024).
- Elmoussalami, H. H. Artificial intelligence and parametric construction cost estimate modeling: State-of-the-art review. *J. Constr. Eng. Manag.* **146**, 03119008. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001678](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001678) (2020).
- Lin, L., Jiang, W., Chen, B., Yu, J. & Zheng, C. Construction and application of cost prediction model based on multiple linear regression analysis. *Procedia Comput. Sci.* **247**, 617–623. <https://doi.org/10.1016/j.procs.2024.10.074> (2024).
- Desse, E. M. & Mengesha, W. J. Predicting construction cost under uncertainty using grey-fuzzy earned value analysis. *Heliyon*. **10**, e27662. <https://doi.org/10.1016/j.heliyon.2024.e27662> (2024).
- Alshamrani, O. S. Initial cost forecasting model of mid-rise green office buildings. *J. Asian Arch. Build. Eng.* **19**, 613–625. <https://doi.org/10.1080/13467581.2020.1778005> (2020).
- Elfahham, Y. Estimation and prediction of construction cost index using neural networks, time series, and regression. *Alex. Eng. J.* **58**, 499–506. <https://doi.org/10.1016/j.aej.2019.05.002> (2019).
- Kumar, P., Nakkeeran, G., Onyelowo, K. C. & Krishnaraj, L. Comparative study on net-zero masonry walls made of clay and fly ash bricks and grouts/mortars/stuccos with the effect of super fine fly ash blended cement—low carbon cement. *Int. J. Low-Carbon Tech.* **18**, 1008–1014. <https://doi.org/10.1093/ijlct/ctad087> (2023).
- Onyelowo, K. C. & Ebid, A. M. The influence of fly ash and blast furnace slag on the compressive strength of high-performance concrete (HPC) for sustainable structures. *Asian J. Civ. Eng.* **25**, 861–882. <https://doi.org/10.1007/s42107-023-00817-9> (2024).
- Onyelowo, K. C., Ebid, A. M. & Ghadikolaei, M. R. GRG-optimized response surface powered prediction of concrete mix design chart for the optimization of concrete compressive strength based on industrial waste precursor effect. *Asian J. Civ. Eng.* **25**, 997–1006. <https://doi.org/10.1007/s42107-023-00827-7> (2024).
- Indraratna, B., Armaghani, D. J., Gomes Correia, A., Hunt, H. & Ngo, T. Prediction of resilient modulus of ballast under cyclic loading using machine learning techniques. *Transp. Geotech.* **38**, 100895. <https://doi.org/10.1016/j.trgeo.2022.100895> (2023).
- Onyelowo, K. C. et al. Runtime-based metaheuristic prediction of the compressive strength of net-zero traditional concrete mixed with BFS, FA, SP considering multiple curing regimes. *Asian J. Civ. Eng.* **25**, 1241–1253. <https://doi.org/10.1007/s42107-023-00839-3> (2024).
- Onyelowo, K. C., Ebid, A. M. & Hanandeh, S. The influence of nano-silica precursor on the compressive strength of mortar using Advanced Machine Learning for sustainable buildings. *Asian J. Civ. Eng.* **25**, 1135–1148. <https://doi.org/10.1007/s42107-023-00832-w> (2024).
- Alshboul, O., Shehadeh, A., Almasabha, G. & Almuflih, A. S. Extreme Gradient Boosting-Based Machine Learning Approach for Green Building Cost Prediction. *Sustainability* <https://doi.org/10.3390/su14116651> (2022).
- Cheng, M.-Y., Tsai, H.-C. & Sudjono, E. Conceptual cost estimates using evolutionary fuzzy hybrid neural network for projects in construction industry. *Exp. Syst. Appl.* **37**, 4224–4231. <https://doi.org/10.1016/j.eswa.2009.11.080> (2010).
- Juszczyk, M. The Challenges of Nonparametric Cost Estimation of Construction Works with the use of Artificial Intelligence Tools. *Procedia Eng.* **196**, 415–422. <https://doi.org/10.1016/j.proeng.2017.07.218> (2017).
- Kumar, P. S. & Gururaj, S. Conceptual cost modelling for sustainable construction project planning—a levenberg-marquardt neural network approach. *Appl. Math. Inf. Sci.* **13**, 201–208. <https://doi.org/10.18576/amis/130207> (2019).
- Kim, G.-H., An, S.-H. & Kang, K.-I. Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Build. Environ.* **39**, 1235–1242. <https://doi.org/10.1016/j.buildenv.2004.02.013> (2004).
- Cheng, M.-Y., Hoang, N.-D. & Wu, Y.-W. Hybrid intelligence approach based on LS-SVM and Differential Evolution for construction cost index estimation: A Taiwan case study. *Auto. Constr.* **35**, 306–313. <https://doi.org/10.1016/j.autcon.2013.05.018> (2013).
- Chou, J.-S., Lin, C.-W., Pham, A.-D. & Shao, J.-Y. Optimized artificial intelligence models for predicting project award price. *Auto. Constr.* **54**, 106–115. <https://doi.org/10.1016/j.autcon.2015.02.006> (2015).
- El-Kholy, A. M., Tahwia, A. M. & Elsayed, M. M. Prediction of simulated cost contingency for steel reinforcement in building projects versus regression-based models. *Int. J. Constr. Manag.* **22**, 1675–1689. <https://doi.org/10.1080/15623599.2020.1741492> (2022).
- Tijanić, K., Car-Pušić, D. & Šperac, M. Cost estimation in road construction using artificial neural network. *Neural Comput. Appl.* **32**, 9343–9355. <https://doi.org/10.1007/s00521-019-04443-y> (2020).
- Onyelowo, K. C. et al. Utilization of GEP and ANN for predicting the net-zero compressive strength of fly ash concrete toward carbon neutrality infrastructure regime. *Int. J. Low-Carbon Tech.* **18**, 902–914. <https://doi.org/10.1093/ijlct/ctad081> (2023).
- Onyelowo, K. C., Ebid, A. M. & Hanandeh, S. Advanced machine learning prediction of the unconfined compressive strength of geopolymer cement reconstituted granular sand for road and liner construction applications. *Asian J. Civ. Eng.* **25**, 1027–1041. <https://doi.org/10.1007/s42107-023-00829-5> (2024).
- Ai, D. & Yang, J. A machine learning approach for cost prediction analysis in environmental governance engineering. *Neural Comput. Appl.* **31**, 8195–8203. <https://doi.org/10.1007/s00521-018-3860-z> (2019).
- Zheng, Z., Zhou, L., Wu, H. & Zhou, L. Construction cost prediction system based on Random Forest optimized by the Bird Swarm Algorithm. *Math. Biosci. Eng.* **20**, 15044–15074. <https://doi.org/10.3934/mbe-2023674> (2023).
- Chen, Y., Lu, G., Wang, K., Chen, S. & Duan, C. Knowledge graph for safety management standards of water conservancy construction engineering. *Auto. Constr.* **168**, 105873. <https://doi.org/10.1016/j.autcon.2024.105873> (2024).
- Durdyyev, S., Ashour, M., Connelly, S. & Mahdiyar, A. Barriers to the implementation of Building Information Modelling (BIM) for facility management. *J. Build. Eng.* **46**, 103736. <https://doi.org/10.1016/j.jobe.2021.103736> (2022).
- Huang, Y., Wu, L., Chen, J., Lu, H. & Xiang, J. Impacts of building information modelling (BIM) on communication network of the construction project: A social capital perspective. *PLoS ONE* **17**, e0275833. <https://doi.org/10.1371/journal.pone.0275833> (2022).
- Abanda, F. H., Kamsu-Foguem, B. & Tah, J. H. M. BIM – New rules of measurement ontology for construction cost estimation. *Eng. Sci. Tech., Int. J.* **20**, 443–459. <https://doi.org/10.1016/j.jestech.2017.01.007> (2017).
- Wang, K.-C. & Tung, S.-H. Cost estimation model based on building information modeling and virtual reality for customizing presold homes. *KSCE J. Civ. Eng.* **27**, 1397–1411. <https://doi.org/10.1007/s12205-023-0825-2> (2023).

33. Yang, S.-W., Moon, S.-W., Jang, H., Choo, S. & Kim, S.-A. Parametric method and building information modeling-based cost estimation model for construction cost prediction in architectural planning. *Appl. Sci.* **12**, 9553 (2022).
34. Li, S. et al. Intelligent modeling of edge components of prefabricated shear wall structures based on BIM. *Buildings* **13**, 1252 (2023).
35. Hong, Y., Hammad, A. W. A., Akbarnezhad, A. & Arashpour, M. A neural network approach to predicting the net costs associated with BIM adoption. *Auto. Constr.* **119**, 103306. <https://doi.org/10.1016/j.autcon.2020.103306> (2020).
36. Abbasnejad, B. et al. Measuring BIM implementation: A mathematical modeling and artificial neural network approach. *J. Constr. Eng. Manag.* **150**, 04024032. <https://doi.org/10.1061/JCEMD4.COENG-14262> (2024).
37. Li, Y., Tong, Z., Tong, S. & Westerdahl, D. A data-driven interval forecasting model for building energy prediction using attention-based LSTM and fuzzy information granulation. *Sustain. Cities Soc.* **76**, 103481. <https://doi.org/10.1016/j.scs.2021.103481> (2022).
38. Huang, L. Combination of model-predictive control with an Elman neural for optimization of energy in office buildings. *Multiscale Multidisciplinary Model., Experiments Des.* **5**, 183–197. <https://doi.org/10.1007/s41939-021-00111-8> (2022).
39. Xu, L., Wang, L., Xue, W., Zhao, S. & Zhou, L. Elman neural network for predicting aero optical imaging deviation based on improved slime mould algorithm. *Optoelectron. Lett.* **19**, 290–295. <https://doi.org/10.1007/s11801-023-2137-7> (2023).
40. Yan, J., Xu, Z., Yu, Y., Xu, H. & Gao, K. Application of a hybrid optimized bp network model to estimate water quality parameters of beihai lake in Beijing. *Appl. Sci.* **9**, 1863 (2019A).
41. Chen, C.-I. & Huang, S.-J. The necessary and sufficient condition for GM(1,1) grey prediction model. *Appl. Math. Comput.* **219**, 6152–6162. <https://doi.org/10.1016/j.amc.2012.12.015> (2013).
42. Chen, L. & Pai, T.-Y. Comparisons of GM (1,1), and BPNN for predicting hourly particulate matter in Dali area of Taichung City, Taiwan. *Atmos. Pollut. Res.* **6**, 572–580. <https://doi.org/10.5094/APR.2015.064> (2015).
43. Li, X. et al. Prediction model of farmland water conservancy project cost index based on PCA–DBO–SVR. *Sustainability* **17**, 2702. <https://doi.org/10.3390/su17062702> (2025).
44. Meng, X.-B., Gao, X. Z., Lu, L., Liu, Y. & Zhang, H. A new bio-inspired optimisation algorithm: Bird Swarm Algorithm. *J. Expt. Theor. Artificial Intel.* **28**, 673–687. <https://doi.org/10.1080/0952813X.2015.1042530> (2016).
45. Wang, J. et al. Short-term power forecasting of fishing-solar complementary photovoltaic power station based on a data-driven model. *Energy Rep.* **10**, 1851–1863. <https://doi.org/10.1016/j.egy.2023.08.039> (2023).
46. Zhang, S. et al. Thermal deformation behavior investigation of Ti–10V–5Al–2.5Fe–0.1B titanium alloy based on phenomenological constitutive models and a machine learning method. *J. Mater. Res. Technol.* **29**, 589–608. <https://doi.org/10.1016/j.jmrt.2024.01.120> (2024).
47. Xu, W. et al. Prediction model of drilling costs for ultra-deep wells based on GA-BP neural network. *Energy Eng.* **120**, 1701–1715. <https://doi.org/10.32604/ee.2023.027703> (2023).
48. Wang, W., Liang, R., Qi, Y., Cui, X. & Liu, J. Prediction model of spontaneous combustion risk of extraction borehole based on PSO-BPNN and its application. *Sci. Rep.* **14**, 5. <https://doi.org/10.1038/s41598-023-45806-9> (2024).
49. Zhang, Y. & Mo, H. Intelligent building construction cost optimization and prediction by integrating BIM and elman neural network. *Heliyon* **10**, e37525. <https://doi.org/10.1016/j.heliyon.2024.e37525> (2024).
50. Baduge, S. K. et al. Artificial intelligence and smart vision for building and construction Machine and deep learning methods and applications. *Auto. Constr.* **141**, 104440. <https://doi.org/10.1016/j.autcon.2022.104440> (2022).
51. Pham, T. Q. D., Le-Hong, T. & Tran, X. V. Efficient estimation and optimization of building costs using machine learning. *Int. J. Constr. Manag.* **23**, 909–921. <https://doi.org/10.1080/15623599.2021.1943630> (2023).
52. Alshboul, O. et al. Deep and machine learning approaches for forecasting the residual value of heavy construction equipment: a management decision support model. *Eng., Constr. Arch. Manag.* **29**, 4153–4176. <https://doi.org/10.1108/ECAM-08-2020-0614> (2022).

## Acknowledgements

This work has received support from Shenyang Agricultural University and the Liaoning Provincial Local Standard Project. We sincerely appreciate your support. The author expresses gratitude to the editor and reviewers for their suggestions for improvement.

## Author contributions

Kun Han: Conceptualization, Methodology, Software, Formal analysis, Data Curation, Writing-Original Draft. Chunsheng Li: Supervision, Project Administration, Funding acquisition. Tieliang Wang: Resources, Data Curation, Investigation, Validation. Wenhe Liu: Software, Validation, Writing-Review & Editing. Xiaochen Xian: Writing-Review & Editing, Validation. Yingying Yang: Writing-Review & Editing, Validation. All authors have read and agreed to the published version of the manuscript.

## Funding

This work was supported by the Local Standard Approval Plan Projects in Liaoning Province (20242223, 20242224, 20242225).

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-10153-4>.

**Correspondence** and requests for materials should be addressed to C.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025, corrected publication 2025