ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20240721001870
Original scientific paper

IGWO-SVM: An Enhanced Cost Prediction Model for Construction Projects

DaWen ZHANG

Abstracts: Accurate cost prediction is crucial for successful construction project management. This study proposes an improved Grey Wolf Optimizer-Support Vector Machine (IGWO-SVM) model for construction project cost prediction. The model incorporates Tent mapping and quantum well techniques to enhance the global search capability and prediction accuracy. Experimental results show that the IGWO-SVM model achieves a prediction error rate of 0.01%, significantly outperforming traditional methods. The model reduced total construction days by 1.72%, total cost by 1.89%, and improved quality levels by 15.31% on average. The IGWO-SVM model demonstrates high stability and accuracy, providing a reliable tool for construction project cost prediction and optimization.

Keywords: construction project; cost prediction; cost optimization; grey wolf optimizer; support vector machine

1 INTRODUCTION

With the advancement of intelligent engineering technology, prediction models utilizing optimization algorithms have become widely applied in the field of engineering. Cost prediction plays an extremely important role in construction project management, as it is the starting point of project cost management and provides direction and basis for cost control and economic benefits of the entire project. Accurately predicting project costs provides a crucial basis for project decision-making and is essential for the successful implementation of the project [1]. Inaccurate cost forecasting can have a profound impact on construction projects. Amini S. et al. believe that poor on-site management, fluctuating material prices, and improper planning are the key reasons for cost overruns in construction projects. Targeted interventions can reduce the risk of cost overruns [2]. GWO is suitable for handling nonlinear and multi peak optimization problems, which is particularly important for cost prediction in complex engineering projects. SVM can handle various uncertain factors and has strong application capabilities in complex engineering cost prediction. The Grey Wolf Optimizer (GWO) simulates the hierarchical relationships and hunting behavior of grey wolf populations to find the optimal solution. This feature enables GWO to perform excellently in optimizing project schedules, helping to arrange construction sequence and duration reasonably, thereby reducing costs and improving efficiency. Support Vector Machine (SVM) is widely used in cost prediction of construction projects, but there is still room for improvement in its prediction accuracy. The current project cost forecast includes numerous uncertain factors such as human resources, equipment, materials, and time, which are extremely complex. Moreover, traditional methods such as expert prediction and time series analysis are difficult to calculate these uncertain factors, resulting in poor accuracy of the obtained results [3]. To address this, this study combines improved GWO (IGWO) with SVM and innovatively utilizes Quantum Well (QW) and Tent Mapping (TM) techniques to enhance GWO, forming a hybrid optimization prediction model. The study aims to improve the accuracy of cost prediction, optimize engineering costs more effectively, and minimize cost wastage. This study improved the accuracy of engineering cost prediction through the IGWO SVM model, providing more reliable data support for project decision-making.

This study is divided into four parts. The first part reviews the research on cost prediction (CP) in construction projects and the Grey Wolf Optimizer (GWO). It begins with the development of the GWO-SVM model for CP in construction projects, followed by a discussion of the IGWO-SVM model, which incorporates Quantum Well (QW) and Tent Mapping (TM) techniques. The second part presents the construction project CP model based on IGWO-SVM. In the third part, an analysis of the outcomes generated by the IGWO-SVM-based CP model is provided. This section is divided into two subsections: the first focuses on the performance analysis of the IGWO-SVM algorithm, while the second presents the application results of the CP model in construction projects. The fourth part concludes the study, summarizing the findings related to construction project cost prediction using the IGWO-SVM model.

2 LITERATURE REVIEW

As a critical foundation for investment in engineering and construction, cost prediction (CP) for construction projects has garnered significant attention and in-depth research from professionals in the field. To estimate the cost and time of building construction, Ujong et al. utilized artificial neural networks (ANNs) and feed-forward-feedback networks to coordinate and manage projects. They also analyzed the input-output relationships between related data to prevent cost overruns. The model demonstrated a mean absolute error (MAE) of 29.52 and a root mean square error (RMSE) of 56.38 in their results [4]. Meharie's team aimed to automate prediction and reduce prediction error by employing a model that integrated the gradient advancement algorithm, support vector machines (SVM), and linear regression to forecast the cost of road construction projects. Their method showed an 87.8% improvement in accuracy [5]. Ghazal and other scholars addressed the issue of cost overruns in construction projects by applying data mining techniques and knowledge discovery to enhance prediction performance. They used clustering, feature selection, and classification interventions to analyze and quantify factors affecting project costs. The results demonstrated the method's interpretability and accuracy [6]. To calculate cost overruns in construction projects, Annamalaisami and other professionals designed a causal mapping and scenario planning method. This approach was used to measure variances from budgeted costs, identify risk indicators for cost overruns, and ultimately reduce such overruns. The results confirmed the method's effectiveness [7]. In an effort to improve estimation efficiency for project cost decisions in the construction sector, Matel and colleagues proposed an ANN model for estimating the cost of engineering services. Their model increased accuracy by 5%, according to the data [8]. Elhegazy and other researchers developed a multi-story building composite floor system to analyze factors such as structure, materials, and loads for preliminary cost estimation. Their results demonstrated the method's reliability [9].

The GWO-SVM algorithm has demonstrated remarkable application capabilities across various fields, including engineering design, medicine, and geology, achieving impressive results. For instance, Xuan and colleagues designed a GWO-SVM model combined with fuzzy entropy to construct feature vectors for diagnosing faults in wind turbine gearboxes. By identifying and classifying different vibration signals, their method achieved an accuracy of 92.5% [10]. Goyal et al. focused on the classification of brain magnetic resonance images, employing principal component analysis and SVM for feature extraction and classification, with GWO optimization of the parameters. Their method reached an accuracy of 99.61% [11]. Zhang and other researchers applied GWO-SVM fusion information to determine geological orientation and monitor drilling trajectories, demonstrating the feasibility of this approach [12]. In another study, Zhang et al. developed a hybrid kernel GWO-SVM model for predicting defects in metal bent pipes, and the results showed strong performance [13]. To accurately predict the depth of coal seam failure, Dou's team used GWO-SVM to extract data variables and optimize parameters, integrating factor analysis to assess the depth of coal seam floor damage. Their findings indicated that this method offers both accuracy and stability [14]. In the realm of customer churn prediction, Durkaya and colleagues designed a GWO-SVM model combined with principal component analysis to select the optimal features, thereby enhancing prediction accuracy. The results confirmed that the model achieved high accuracy [15].

In summary, both the prediction of construction project costs and the application of GWO-SVM have demonstrated significant value. However, there is relatively limited research on the use of GWO-SVM specifically for construction project cost prediction. Additionally, the presence of fuzzy project definitions and characteristics in cost prediction often leads to low prediction efficiency. To enhance the model's global search capability (GSC) and further improve the efficiency and accuracy of the prediction model, this study integrates GWO with SVM and incorporates Tent Mapping (TM) and Quantum Well (QW) techniques to enhance model performance. By improving the accuracy and stability of the cost prediction (CP) model, the study aims to achieve more precise and controlled project cost predictions, thereby optimizing cost management, improving management efficiency, and increasing the economic benefits of construction projects. A summary table of key related works is shown in Tab. 1.

Table 1 Summary table of key related work

Author	Research method	Advantage	Limitation
Ujong et al. [4]	ANNs and feed-forward-feedback networks	Prevents cost overruns	Extensive data for training
Meharie et al. [5]	Gradient advancement algorithm, SVM, and linear regression	87.8% improvement in accuracy for road construction costs	Specific to road construction projects
Ghazal et al. [6]	Data mining techniques, clustering, feature selection, and classification interventions	Enhanced prediction performance and interpretability	Low generalization ability
Annamalaisami et al. [7]	Causal mapping and scenario planning method	Effective in identifying and reducing cost overruns	Complexity in implementing
Xuan et al. [10]	GWO-SVM model combined with fuzzy entropy for fault diagnosis in wind turbines	Achieved 92.5% accuracy in classification	Specific to wind turbine applications
Goyal et al. [11]	Principal component analysis and SVM for brain MRI classification with GWO optimization	High accuracy of 99.61%	Limited to medical imaging
Dou et al. [14]	GWO-SVM, integrating factor analysis	Offers accuracy and stability in predictions	Complexity of coal seam data analysis
Durkaya et al. [15]	GWO-SVM model combined with principal component analysis	High prediction accuracy	Focused on customer churn only

3 RESEARCH METHODS

The study initially develops a construction project cost prediction (CP) model based on the Grey Wolf Optimizer (GWO) integrated with Support Vector Machine (SVM). Subsequently, Quantum Well (QW) and Tent Mapping (TM) techniques are incorporated to enhance the global search capability (GSC) of the GWO-SVM model, aiming to optimize engineering costs more effectively. The technical roadmap flowchart of this study is shown in Fig. 1.

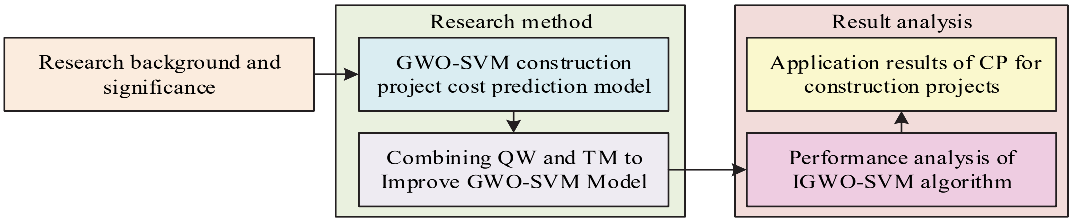


Figure 1 Technical roadmap flowchart of this study

3.1 GWO-SVM Construction Project Cost Prediction Model

Engineering cost prediction (CP) is conducted multiple times at various stages of a project, as engineering projects typically finalize planning decisions in accordance with established protocols, thereby ensuring the reasonableness and accuracy of the project costs [16]. The content of CP for construction projects is illustrated in Fig. 2.

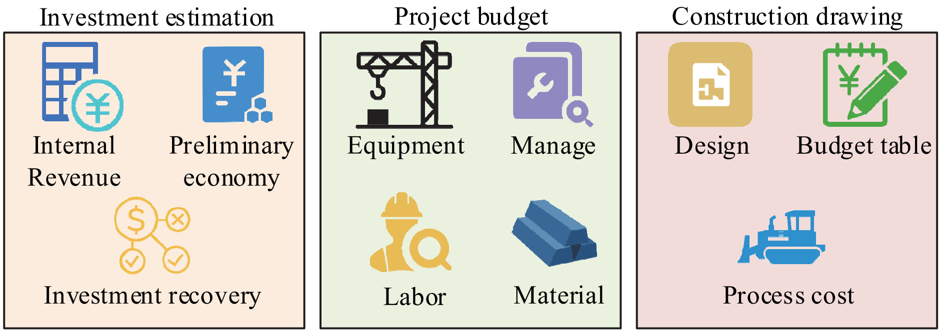


Figure 2 Engineering cost prediction content

The three stages of the cost prediction (CP) project, as shown in Fig. 2, are the project proposal and feasibility study stage, the preliminary design stage, and the construction drawing design stage. These phases correspond to investment estimation, project estimate, and construction drawing budget, respectively. Cost control is one of the most critical aspects in construction works (CW) projects, and the Grey Wolf Optimizer (GWO) can be effectively applied to optimize cost control by adjusting the weights and allocations of various cost factors to achieve optimal cost distribution. GWO is particularly effective in terms of accuracy and convergence speed, excelling in problem-solving by balancing local optimization with global search. Its computational phases include tracking, surrounding, attacking, and establishing a social hierarchy [17]. The hierarchy of grey wolves and their position update method are illustrated in Fig. 3.

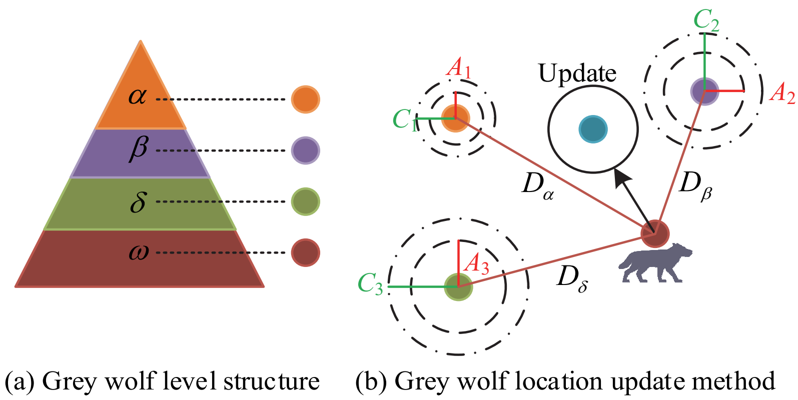


Figure 3 GW level structure and its position update method

To assess the adaptability of each member within the Grey Wolf (GW) population, a GW social hierarchy model was first established, as depicted in Fig. 3a. In this hierarchy, the three GWs with the highest adaptability are labeled as α, β and δ in descending order, with the remaining GWs categorized as ω. In Fig. 3b, during the process of searching for prey, the GW population gradually converges towards the prey, forming encirclement. The mathematical model representing this behavior is provided in Eq. (1).

$$\begin{cases} D = C \bullet X_P(t) - X(t) \\ X(t+1) = X_P - A \bullet D \end{cases} \tag{1}$$

In Eq. (1), D represents the distance between the Grey Wolf (GW) and the prey. The current iteration is denoted by t, while the subsequent iteration is indicated by $t + 1$. X and X_P are the position vectors (PVs) of the GW and prey, respectively. Eq. (2) defines A and C, which are the coefficient vectors.

$$\begin{cases} A = 2a \cdot r_1 - a \\ C = 2r_2 \end{cases} \tag{2}$$

In Eq. (2), both r_1 and r_2 are random vectors in the interval [0, 1], and the parameter a decreases linearly from 2 to 0 over the course of iterations. The recognition of prey by the Grey Wolves (GWs) α, β and δ guides the entire pack. However, when the problem solution lacks distinct spatial features and complexity, the obtained solution may not be sufficiently precise. To address this, the GWs α, β and δ are set to be more proficient in prey localization. After numerous iterations, the three GWs with the best adaptability are retained, and the optimal solution (OS) position is determined based on their actual positions. The distances between the current wolf and the optimal GWs α, β and δ are D_α, D_β, and D_δ, respectively, as shown in Eq. (3).

$$\begin{cases} D_\alpha = C_1 X_\alpha - X \\ D_\beta = C_2 X_\beta - X \\ D_\delta = C_3 X_\delta - X \end{cases} \tag{3}$$

In Eq. (3), X_α, X_β, and X_δ is the PV of the optimal GWs α, β and δ, respectively. The positions of the candidate wolves are determined using the following expression, as shown in Eq. (4).

$$\begin{cases} X_1 = X_\alpha - A_1 D_\alpha \\ X_2 = X_\beta - A_2 D_\beta \\ X_3 = X_\delta - A_3 D_\delta \end{cases} \tag{4}$$

The PV $X(t + 1)$ of the candidate wolf is shown in Eq. (5).

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{5}$$

When $|A| < 1$, the GWs begin to concentrate and search for prey within a specific search range. Conversely, $|A| > 1$ indicates that the GWs are dispersed, necessitating an expansion of the search range to locate prey. To enhance the efficiency of cost prediction (CP) for construction projects and achieve global search optimization, this study combines the Grey Wolf Optimizer (GWO) with Support Vector Machine (SVM). The SVM regression function attempts to find a hyperplane such that most of the data points are located near this hyperplane. The regression function $f(x_i)$ for SVM is defined in Eq. (6).

$$f(x_i) = w_i \cdot \phi(x_i) + b \tag{6}$$

In Eq. (6), w_i represents a definable weight vector, $\phi(x_i)$ is a vector mapping, and b is a bias. To evaluate the error between the regression function $f(x_i)$ and the predictor variable y_i, the study employs the insensitive loss function L_ε. The insensitive loss function L_ε aims to reduce sensitivity to small errors, as shown in Eq. (7).

$$L_\varepsilon = (f(x_i), y_i) = \begin{cases} |-y_i + f(x_i)| - \varepsilon, |-y_i + f(x_i)| > \varepsilon \\ 0, |-y_i + f(x_i)| \le \varepsilon \end{cases} \tag{7}$$

In Eq. (7), ε influences the number of SVMs involved, thereby affecting the fitting capability of the approximation function. In SVM, hyperplanes are used to distinguish data points. When combined with hyperplane analysis, the relationship between w_i and b is expressed in Eq. (8).

$$\begin{cases} \min G(w, b, \xi_i, \xi_i^*) = \frac{1}{2}\|w_u\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ s.t \begin{cases} \xi_i, \xi_i^* \ge 0, i = 1, 2, ..., n \\ -y_i + b + w \cdot \phi(x_i) \le \varepsilon + \xi_i^* \\ y_i - b - w \cdot \phi(x_i) \le \varepsilon + \xi_i \end{cases} \end{cases} \tag{8}$$

To quantify the sample bias outside of the insensitive region, slack variables ξ_i and ξ_i^* are introduced into Eq. (8). In the case of linear differentiability, the OSs for w_i and b are determined when the objective function value is minimized. The objective function is used to minimize the complexity and error of the model, as shown in Eq. (9).

$$\min \phi(x_i) = \frac{1}{2}\|w_i\|^2 \tag{9}$$

The Karush Kuhn Tucker (KKT) condition is an important part of optimization problems, used to solve constrained optimization problems. The dual form of the Lagrange multiplier is introduced. This transformation converts the original objective and constraints into a quadratic equation with two variables and a linear function, as represented in Eq. (10).

$$\begin{cases} \max f(a_i, a_i^*) = \sum_{i=1}^{n}(a_i - a_i^*) - \varepsilon\sum_{i=1}^{n}(a_i + a_i^*) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_i^*)(a_j - a_j^*)K(x_i, x_j) \\ \text{s.t.} \begin{cases} \sum_{i=1}^{n}(a_i - a_i^*) = 0 \\ a_i, a_i^* \in [0, C] \end{cases} \end{cases} \tag{10}$$

In Eq, (10), a_i^* and a_i are the values of the Lagrange multipliers. Consequently, the regression function of the SVM is formulated as shown in Eq. (11).

$$\begin{cases} w^* = \sum_{i=1}^{n}\left(a_i - a_i^*\right)K\left(x, x_i\right) \\ f\left(x\right) = \sum_{i=1}^{n}\left(a_i - a_i^*\right)K\left(x, x_i\right) + b \end{cases} \tag{11}$$

In Eq. (11), $K\left(x, x_i\right)$ represents the kernel function. The SVM algorithm selects the Radial Basis Function (RBF) as the kernel function due to its strong nonlinear modeling capability, its independence from input dimensions, and its ability to perform high-level feature mapping. The corresponding formula for $K\left(x, x_i\right)$, is provided in Eq. (12).

$$K\left(x, y\right) = \exp\left(\frac{-\left\|x - y\right\|^2}{2\sigma^2}\right) \tag{12}$$

In Eq. (12), σ is the width of the RBF kernel. In the GWO-SVM construction cost prediction model, data preprocessing and feature selection are two key steps. Perform data preprocessing such as data cleaning and standardization to ensure data quality and consistency. The evaluation of feature importance involves methods such as analysis of variance and mean comparison to assess the relationship between each feature and the target variable. Feature selection helps the model focus on the most predictive features, thereby improving the accuracy and efficiency of the model.

3.2 Combining QW and TM to Improve GWO-SVM Model

In cost prediction (CP) for construction projects, the GWO-SVM algorithm faces challenges due to its single global search method and more random initialization, which increases the risk of converging to local optima in the later stages of iteration. In GWO, the positions of wolves are updated sequentially based on the established social hierarchy, which leads to the aggregation of wolves. If the alpha wolf α falls into the local extreme point, it will affect the whole wolf pack to fall into the local extreme, so in the GWO, it is necessary to focus on the position of wolf α, β and δ. To enhance the probability of obtaining the global optimum, quantum search is utilized to update the positions of these wolves. In this study, the quantum space is defined as the feasible domain for the GWO solution, where the particles represent individual GWs, and the potential well points correspond to the positions of wolves α, β and δ. The state of each individual is represented by the wave function $\Psi\left(X_i, t\right)$, and X_i is the PV of the individual GW. In addition, the probability density of wolf α, β and δ appearing at a specific point in space is determined using Schrödinger's equation [18]. The QW update mechanism process is illustrated in Fig. 4.

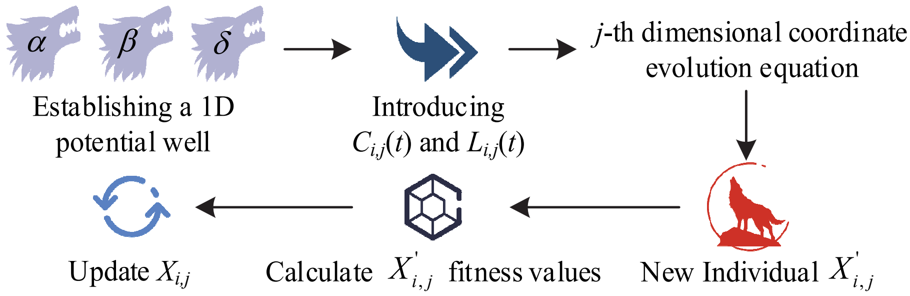


Figure 4 QW update mechanism

In Fig. 4, the one-dimensional (1D) potential well is established with GWs α, β, and δ at the center. The average optimal position $C_{i,j}(t)$ and the characteristic length BB of the 1D potential well are introduced, as represented in Eq. (13).

$$\begin{cases} C_{i,j}\left(t\right) = \left(C_1\left(t\right), C_2\left(t\right), \ldots, C_n\left(t\right)\right) = \frac{1}{M}\sum_{i=1}^{M} P_i\left(t\right) \\ L_{i,j}\left(t\right) = 2d \cdot \left|C_j\left(t\right) - X_{i,j}\left(t\right)\right| \end{cases} \tag{13}$$

In Eq. (13), $i = \alpha, \beta, \delta, j = 1, 2, \ldots, M$. P_i is the center of the potential trap. Using this, the evolution equation for the j-th dimension coordinate is derived. The new individual $X'_{i,j}$ is obtained according to the equation and the $X'_{i,j}$ fitness value *Fitness* is calculated as shown in Eq. (14).

$$Fitness = \sum_{i=1}^{N} \omega_T \cdot \frac{\left(X_{i,1} - T_s\right)}{\left(T_n - T_s\right)} + \omega_C \cdot \frac{\left(X_{i,2} - C_s\right)}{\left(C_n - C_s\right)} - \omega_Q \cdot \frac{\left(X_{i,3} - Q_s\right)}{\left(1 - Q_s\right)} \tag{14}$$

In Eq. (14), N represents the engineering sample data. $X_{i,1}$, $X_{i,2}$, and $X_{i,3}$ correspond to the duration, cost, and quality of project i, respectively. T is the duration of the construction project, measured in days. C is the cost of the construction project, measured in tens of thousands of yuan. Q is the quality of the construction project, with a quality score range of 0 to 1. Finally, the update for $X_{i,j}$ is performed. In dimension M, the i only wolf the j dimension location equation $X_{i,j}$, is given in Eq. (15).

$$X_{i,j}(t+1) = X_{i,j} \pm d \cdot \left|-X_{i,j}(t)+C_{i,j}(t)\right| \cdot \ln\left[1/u_{i,j}(t)\right] \quad (15)$$

In Eq. (15), d is the expansion-contraction coefficient and $C_{i,j}(t)$. In interval (0, 1), $u_{i,j}(t)$ is a uniformly distributed random number. Using specific mapping rules, the problem space of construction project cost prediction (CP) is transformed into a chaotic solution set space. The Tent Mapping (TM) method is employed for this conversion due to its simplicity, more uniform distribution, and convenience, allowing the problem to acquire a distinct nature and structure within the chaotic space. In this chaotic solution set space, TM equations are iteratively applied to compute the mapping objects and locate the optimal solution (OS). These TM equations, as a form of dynamical system equations with stochastic and nonlinear properties, can generate complex and seemingly random sequences through iterative operations. The stochastic and nonlinear characteristics of chaotic iteration diversify and broaden the search process, helping to overcome the limitations of local optimal solutions and increasing the likelihood of finding global optimal solutions [19, 20]. Tent chaotic mapping is incorporated in the study to update the wolf positions during the GWO solution process, as illustrated by Eq. (16), with the aim of enhancing the algorithm's efficiency in identifying the OS.

$$x_{n+1} = \begin{cases} 2\left(x_n + 0.1 \times \text{rand}(0,1)\right) & 0 \le x_n \le 0.5 \\ 2\left(1-\left(x_n + 0.1 \times \text{rand}(0,1)\right)\right) & 0.5 < x_n \le 1 \end{cases} \quad (16)$$

The GW ω individual X_i is mapped to the interval (0, 1). $X_{\min,j}$ and $X_{\max,j}$ are the maximum and minimum values, respectively, to obtain $Z_{i,j}$, as shown in Eq. (17).

$$Z_{i,j} = \left(-X_{\min,j} + X_{i,j}\right) / \left(-X_{\min,j} + X_{\max,j}\right) \quad (17)$$

Generate a Tent chaotic mapping sequence from Eq. (17). Map the chaotic sequence back to the feasible domain to generate a new gray wolf ω individual X_i. The adaptation value of the new Grey Wolf X_i is calculated, and the Grey Wolf ω individual is updated accordingly. By integrating Quantum Well (QW) and Tent Mapping (TM), a new model called IGWO-SVM is obtained. The flow of this process is illustrated in Fig. 5.

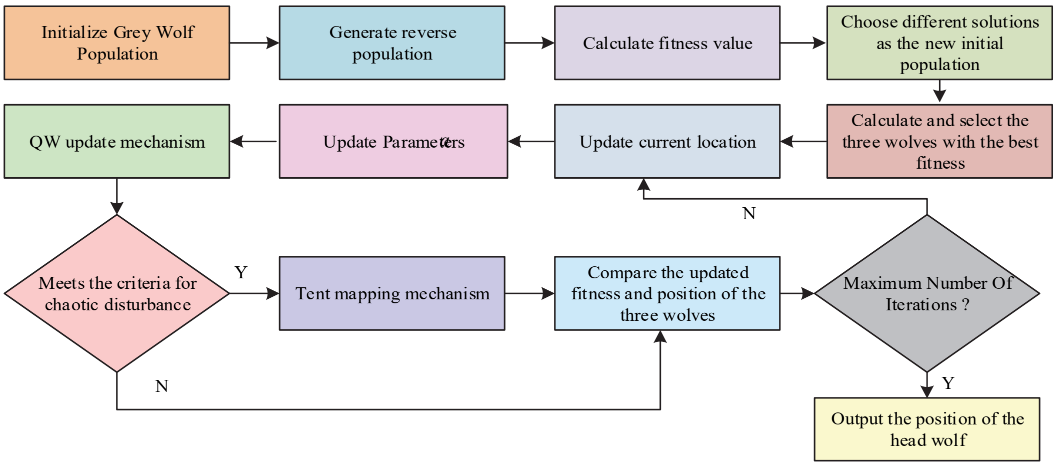


Figure 5 IGWO-SVM model solving process

In Fig. 5, the process begins by initializing the grey wolf population and generating a reverse population. The fitness values of both the initial and reverse populations are calculated, and different solutions are compared and selected to form a new initial population. Next, the three wolves with the best fitness are identified, and their current positions and parameters are updated. The Quantum Well (QW) update mechanism is then introduced to assess whether the chaotic disturbance criteria are met. If these criteria are satisfied, the Tent Mapping mechanism is applied. After the Tent Mapping process, the fitness and positions of the three updated wolves are compared. If the maximum number of iterations is reached, the position of the head wolf is output as the final solution. Combining QW and TM techniques in the GWO-SVM model can effectively improve the prediction accuracy of the model. The global search capability of QW and the chaotic characteristics of TM enable the model to better adapt to data features, avoid cost waste, and optimize engineering costs when processing complex construction cost data.

4 RESULTS AND DISCUSSION

To evaluate the performance of the enhanced algorithm, the study compares the adaptation values of several algorithms and selects various evaluation indices for the IGWO-SVM algorithm. To validate the applicability of the revised model, real-world construction work (CW) projects are also used for testing.

4.1 Performance Analysis of IGWO-SVM Algorithm

The hardware configuration for this study is Intel Core i7-9750H processor, 16 GB of memory, and Windows 10 operating system. Choose PyCharm 2020 development platform for software configuration, version is Python 3.8. The IGWO-SVM algorithm was set with a maximum of 200 iterations, a population size of 30, and a Tent Mapping

(TM) factor of 2. Setting the maximum number of iterations to 200 is to ensure that the algorithm has enough time to converge to an ideal solution. 200 times is a balance point that can achieve good results without causing unnecessary computation time. The population size determines the exploration ability of the algorithm in the search space. A larger population can increase the diversity of solutions and avoid getting stuck in local optima. 30 individuals can provide sufficient diversity in many situations without making the calculations too complicated. To evaluate the performance of IGWO, its convergence and stability were compared with other algorithms, including PSO, GWO, GWO+Tent, and GWO+QW. The results, showing a comparison of the fitness values for these different algorithms, are presented in Fig. 6.

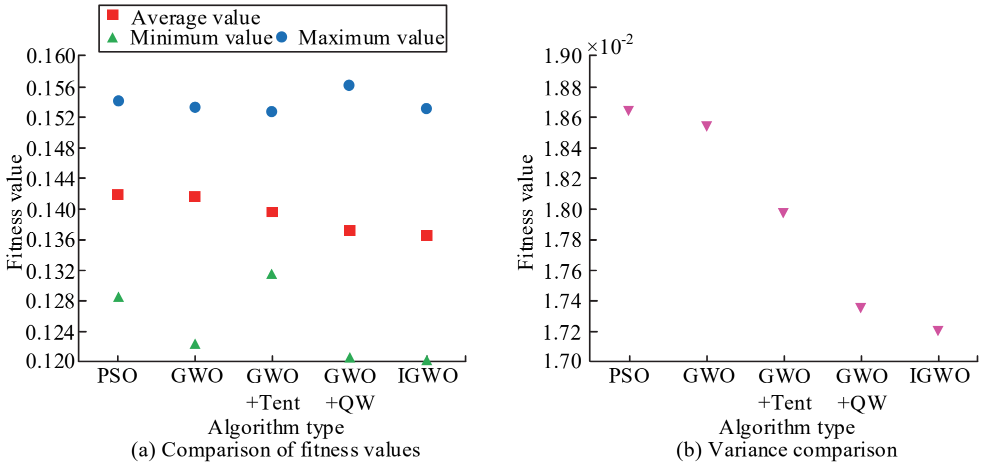


Figure 6 Comparison results of fitness values for different algorithms

In Fig. 6a, the average, minimum, and maximum values of IGWO are 0.136, 0.120, and 0.153, respectively, all of which are smaller than those of GWO. The maximum value of GWO+QW is 0.156, which is higher than the maximum values of the remaining four algorithms. This improvement is attributed to the introduction of QW, which enhances the algorithm's diversity and expands its search range. This helps avoid local optima and allows for a more thorough exploration of the global solution space. GWO+QW discovers higher quality solutions in the solution space, thereby improving fitness values. But this enhanced exploration ability also means that the algorithm lacks the ability to finely adjust the solutions found, resulting in a higher maximum value but not always stable within a better range. In Fig. 6b, the variances of PSO and GWO are 1.866×10^{-2} and 1.856×10^{-2}, respectively. The variance of IGWO is the smallest, 1.721×10^{-2}, which is 7.27% lower than that of GWO. The variance of GWO with the addition of QW alone is reduced by 6.47% compared to GWO, and the variance of GWO with the addition of TM alone is reduced by 3.06% compared to GWO. This indicates that the performance optimization and stability of the combined improved IGWO are further enhanced compared to the single IGWO. The study also examines the fitness values of several algorithms as the number of iterations increases to analyze the convergence behavior of IGWO. The associated curves are then generated, as illustrated in Fig. 7.

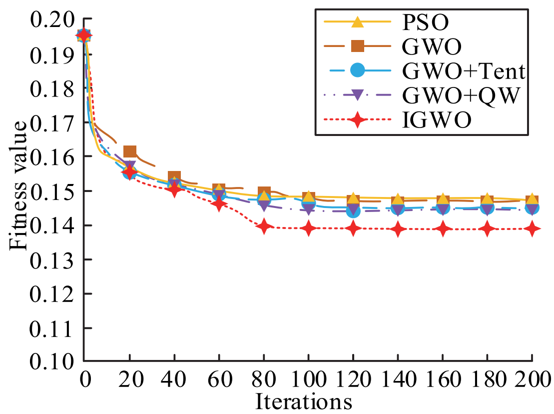


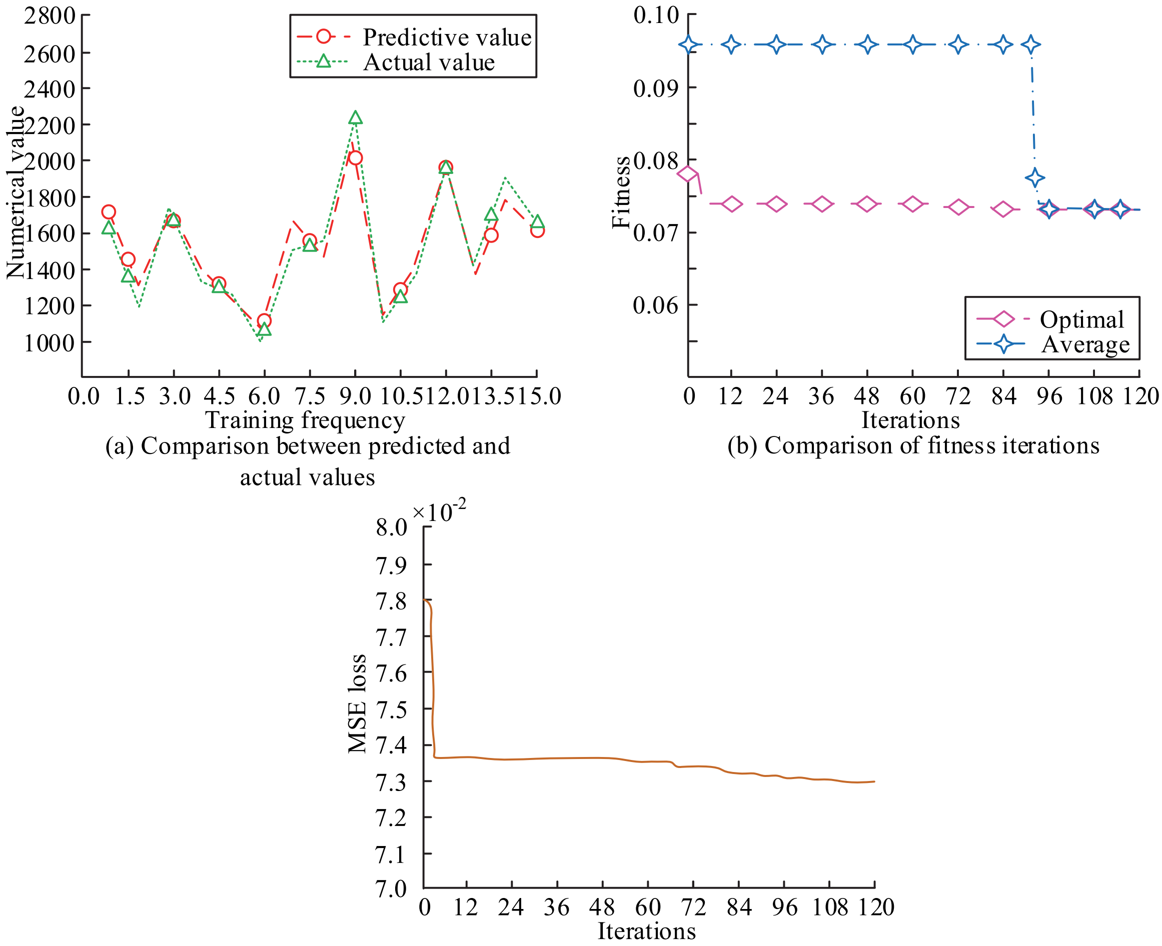
Figure 7 Comparison of changes in fitness values of different algorithms

In Fig. 7, the convergence of the five algorithms accelerates at varying degrees before 50 iterations. Both GWO and PSO experience local optima after 50 iterations and begin to converge around the 100th iteration. In contrast, IGWO starts to converge earlier, at around 80 iterations, with an adaptability value of 0.136, indicating that IGWO has better convergence performance. The study focuses on duration as the primary consideration, and the comparison of the cost optimization (CO) results of different algorithms is presented in Tab. 2.

Table 2 Cost optimization results of different algorithms (unit: $\times 10^4$ yuan)

Engineering sample	Before optimization	IGWO	GWO	GWO+Tent	GWO+QW	PSO
1	2745.2	2671.5	2681.8	2672.4	2671.7	2677.6
2	5386.2	5321.7	5315.3	5330.1	5338.2	5318.7
3	8538.6	8493.8	8497.5	8495.2	8498.2	8493.6
4	6843.8	6806.8	6805.2	6807.1	6806.2	6807.9
5	3400.7	3351.0	3360.7	3353.7	3357.3	3354.2
P value	/	/	0.041	0.053	0.027	0.013

In Tab. 2, a comparison between the pre-optimization and post-optimization costs reveals that the costs are reduced across different algorithms. The IGWO method, in particular, reduces costs by an average of 539,400 yuan from the pre-optimization cost, representing a 1.02% reduction. Additionally, the GWO method with only Tent Mapping (TM) and the GWO method with only Quantum Well (QW) reduced costs by 1.00% and 0.98%, respectively. Among the five methods compared, the IGWO demonstrated the most effective cost optimization (CO). The *P* values of the IGWO algorithm and different algorithms, except for the GWO+Trent algorithm, are all less than 0.05. The IGWO algorithm does have significant differences in cost optimization performance compared to other algorithms. To further validate the efficacy of the revised approach, the study summarized the outcomes of 20 model training cycles. The initial number of wolves was fixed at 90, and the number of iterations was increased to 120. Fig. 8 presents the comparison between the actual values and the predicted values (PreV), alongside the findings of the optimal and average fitness, and the model's mean square error (MSE).



(a) Comparison between predicted and actual values

(b) Comparison of fitness iterations

(c) MSE loss

Figure 8 Comparison of engineering cost forecast results

In Fig. 8a, the comparison between the actual value (AV) and the predicted value (PreV) shows that their trends are essentially aligned. At the 9th iteration of training, the PreV is 2025, while the AV is 2230. In Fig. 8b, the average fitness remains around 0.096 during the first 92 iterations and then decreases to 0.074 between the 96th and 120th iterations. The optimal fitness shows minimal variation within the 120 iterations, decreasing slightly from approximately 0.078 to 0.074. In Fig. 8c, the mean square error (MSE) loss value of the improved model experiences a sharp decline from 7.80×10^{-2} to 7.37×10^{-2} at 0 to 4 iterations. Its change after 4 iterations is much smoother and does not change much. By 120 iterations, the MSE loss value is about 7.30×10^{-2}. This indicates that the average squared difference between the predicted and actual values of the IGWS-SVM model is 0.073. MSE is a measure of model performance, and the smaller the value, the higher the degree of agreement between the model's predicted results and actual values. IGWO SVM combines improved GWO and SVM, which results in higher computational complexity when training on large-scale datasets. To further confirm the applicability of the SVM model in the enhanced IGWO-SVM model, the study inserted the engineering data into alternative models, including the decision tree, BP neural network, and random forest (RF), while keeping other conditions constant. This study used standard error (SE) to estimate confidence intervals. The resulting cost prediction (CP) comparisons are summarized in Tab. 3.

Table 3 Comparison of applicability results of SVM models

Model	SVM	Decision tree	BPNN	RF
Actual value	1859.73	1859.73	1859.73	1859.73
Predictive value	1859.62	1643.38	1862.61	1800.54
Absolute value of error	0.11	216.35	2.88	59.19
Error rate	0.01%	11.63%	0.15%	3.18%
SE of Error Rate	0.001%	0.45%	0.005%	0.07%
95% Confidence Interval (CI)	0.01% ± 0.002%	11.63% ± 0.90%	0.15% ± 0.01%	3.18% ± 0.14%

In Tab. 3, the error rate of the SVM prediction value in the IGWO-SVM model is 0.01%, which is nearly identical to the actual value (AV) of the project. This indicates that the SVM model has high prediction accuracy and precision. Compared with the other three prediction models, the integration of Quantum Well (QW) and Tent Mapping (TM) ensures that the SVM model searches for the optimal solution (OS) more efficiently, thereby enhancing the model's functionality and stability.

4.2 Application Results of CP for Construction Projects

A total of 126 construction work (CW) sample data were selected for the study, all sourced from the Guanglianda Index Network. Guanglianda Index Network is a professional construction engineering data platform that provides rich engineering cost and construction management data. Select projects closely related to cost forecasting to ensure sufficient representativeness. This includes considering factors such as the type, scale, and region of the project. The study selected projects closely related to cost prediction to ensure that the sample has certain representativeness. The study employed methods such as data cleaning, standardization, and cross validation to minimize the impact of biases in the dataset on the model. Of these, 102 samples were used in the training set, and 24 samples were allocated to the test set. The study considered 22 pertinent cost prediction (CP) indicators, with 6 indicators specifically related to the engineering cost category. Fig. 9 presents a comparison between the actual values and the predicted values (PreVs) in both the training and test sets.

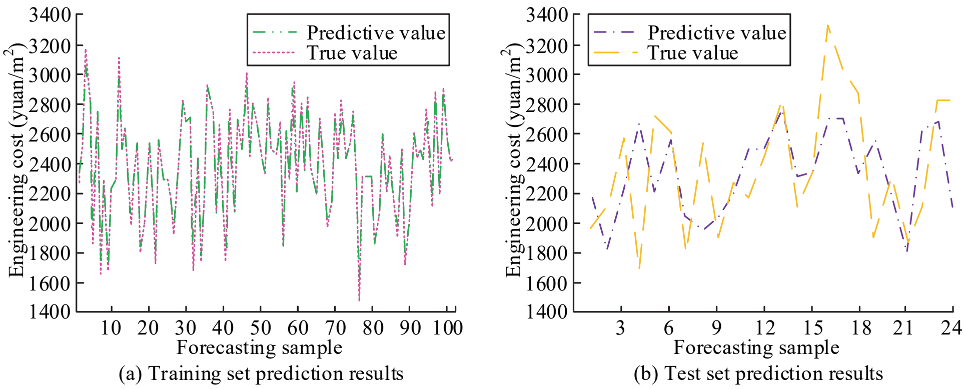


Figure 9 Comparison of prediction results between training and testing sets

In Fig. 9a, the IGWO-SVM model shows a high degree of overlap between the true values and the predicted values (PreVs) in the training set, with a Root Mean Square Error (RMSE) of 53.51 and an R^2 value of 0.978. The RMSE of the model designed in reference [4] is 56.38. In comparison, the IGWS-SVM model performs better than the reference model in similar tasks. In Fig. 9b, the test set consists of 24 samples, where the RMSE is higher at 409.43, indicating a greater prediction bias, although the R^2 value remains 0.978. The IGWO-SVM model performs well on the training set, but overfitting occurs on the test set. The model learns too many details on the training data, resulting in a decrease in its predictive ability on the test set. To further refine the interval for cost prediction (CP) in construction projects, the study compares the predictions made by the IGWO-SVM and GWO-SVM models. Absolute error (AE) is the absolute difference between the predicted value and the true value, and the size and volatility of AE can intuitively demonstrate the predictive performance of the model. The comparison between model prediction performance and absolute error is shown in Fig. 10.

The variance in the prediction results (PRs) of the GWO-SVM model is larger and shows a slight deviation from the true value, while the PRs of the IGWO-SVM model in Fig. 10(a) are closer to the true value. For instance, the GWO-SVM model's PR for the 4th sample data is 1600 yuan/m^2, which deviates by 1200 yuan/m^2 from the true value of 2800 yuan/m^2. In Fig. 10b, all 24 engineering samples are considered, and the absolute error variation of the IGWO-SVM model shows minimal fluctuation, with a maximum absolute error of 240.41. This indicates that the IGWO-SVM model performs more stably in different samples. Meanwhile, smaller AEs can help predict project costs more accurately, thereby enabling better resource allocation and budget control. On the other hand, the GWO-SVM model's absolute error fluctuates more significantly, with the 4th sample showing a maximum deviation of −1,245.34. To achieve cost optimization (CO) management and ensure that quality and duration are maintained within a reasonable range by appropriately adjusting the schedule while minimizing construction costs, the study randomly selects five project sample data for analysis. With cost as the primary consideration, the CO results of the IGWO-SVM algorithm are compared, as shown in Tab. 4.

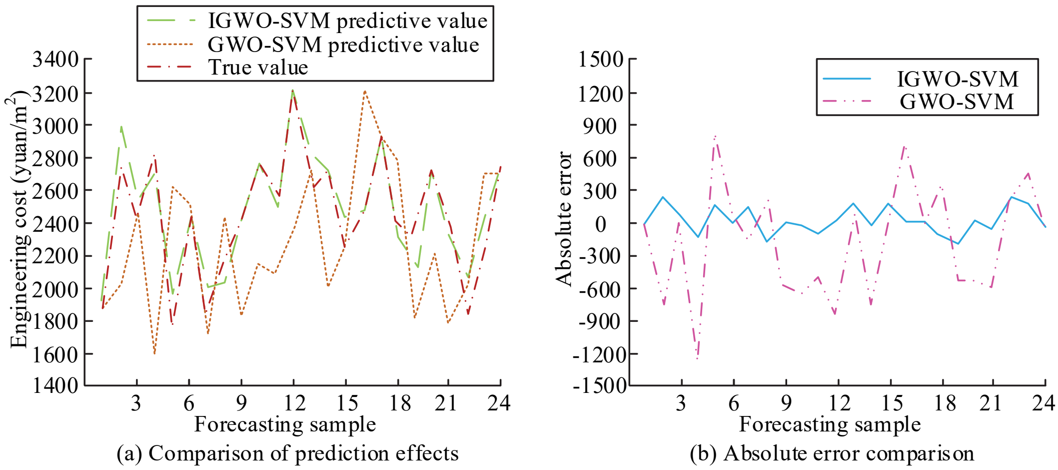


Figure 10 Comparison of model prediction performance and absolute error before and after improvement

Table 4 Optimization results for cost minimization

Engineering sample		1	2	3	4	5
Before optimization	Total construction period (days)	121.00	45.00	91.00	45.00	14.00
	Total cost (× 10^4 yuan)	2745.20	5386.20	8538.60	6843.80	3400.70
	Total quality	0.81	0.80	0.82	0.83	0.79
After optimization	Total construction period (days)	120.11	44.48	89.79	43.65	13.29
	Total cost (× 10^4 yuan)	2661.50	5285.90	8501.00	6775.70	3336.90
	Total quality	0.96	0.94	0.95	0.96	0.93

In Tab. 4, the optimized duration, cost, and quality are better than the pre-optimization results, and the ratings of total quality are above 0.93, which is a more significant improvement. While the total duration days are reduced by 1.72% on average, while the total cost is reduced by 1.89% and the quality level is improved by 15.31%.

5 CONCLUSION

This study presents an enhanced cost prediction model for construction projects using an improved Grey Wolf Optimizer-Support Vector Machine (IGWO-SVM) approach. The integration of Tent mapping and quantum well techniques significantly improves the model's global search capability and prediction accuracy. Experimental results demonstrate that the IGWO-SVM model achieves a prediction error rate of 0.01%, outperforming traditional methods. The model's application to real-world construction projects resulted in average reductions of 1.72% in total construction days and 1.89% in total cost, while improving quality levels by 15.31%. These improvements highlight the potential of the IGWO-SVM model to enhance construction project management practices by providing more accurate cost predictions and optimization strategies. Future work should focus on testing the model's performance across a wider range of project types and sizes, and exploring its integration with other project management tools and techniques. This comprehensive improvement enhances the global search capability of the model, making it perform well in optimizing construction CP. The IGWO-SVM model can provide project managers with more reliable cost prediction results, which helps optimize budget formulation and resource allocation, thereby reducing the risk of cost overruns. However, the addition of QW and TM to the IGWS-SVM model may increase the complexity and computational burden of the model, which may result in higher time costs for model training and optimization processes. Future research can improve the computational efficiency of algorithms, such as parallel computing and distributed computing, to enhance the speed of model training and optimization.

6 REFERENCE

[1] Karatas, I. & Budak, A. (2024). Development and comparative of a new meta-ensemble machine learning model in predicting construction labor productivity. *Engineering, Construction and Architectural Management*, *31*(3), 1123-1144. https://doi.org/10.1108/ECAM-08-2021-0692

[2] Amini, S., Rezvani, A., Tabassi, M., & Sadati, S. (2023). Causes of cost overruns in building construction projects in Asian countries; Iran as a case study. *Engineering, Construction and Architectural Management*, *30*(7), 2739-2766.

[3] Ahmed, H. U., Mostafa, R. R., Mohammed, A., Sihag, P., & Qadir, A. (2023). Support vector regression (SVR) and grey wolf optimization (GWO) to predict the compressive strength of GGBFS-based geopolymer concrete. *Neural Computing and Applications*, *35*(3), 2909-2926. https://doi.org/10.1007/s00521-022-07724-1

[4] Ujong, J. A., Mbadike, E. M., & Alaneme, G. U. (2022). Prediction of cost and duration of building construction using artificial neural network. *Asian Journal of Civil Engineering*, *23*(7), 1117-1139.

https://doi.org/10.1007/s42107-022-00474-4

[5] Meharie, M. G., Mengesha, W. J., Gariy, Z. A., & Mutuku, R. N. (2022). Application of stacking ensemble machine learning algorithm in predicting the cost of highway construction projects. *Engineering, Construction and Architectural Management*, *29*(7), 2836-2853. https://doi.org/10.1108/ECAM-02-2020-0128

[6] Ghazal, M. M. & Hammad, A. (2022). Application of knowledge discovery in database (KDD) techniques in cost overrun of construction projects. *International Journal of Construction Management*, *22*(9), 1632-1646. https://doi.org/10.1080/15623599.2020.1738205

[7] Annamalaisami, C. D. & Kuppuswamy, A. (2022). Reckoning construction cost overruns in building projects through methodological consequences. *International Journal of Construction Management*, *22*(6), 1079-1089. https://doi.org/10.1080/15623599.2019.1683689

[8] Matel, E., Vahdatikhaki, F., Hosseinyalamdary, S., Evers, T., & Voordijk, H. (2022). An artificial neural network approach for cost estimation of engineering services. *International journal of construction management*, *22*(7), 1274-1287. https://doi.org/10.1080/15623599.2019.1692400

[9] Elhegazy, H., Chakraborty, D., Elzarka, H., Ebid, A. M., Mahdi, I. M., Aboul Haggag, S. Y., & Abdel Rashid, I. (2022). Artificial intelligence for developing accurate preliminary cost estimates for composite flooring systems of multi-storey buildings. *Journal of Asian Architecture and Building Engineering*, *21*(1), 120-132.

[10] Xuan, H. U., Chun, L. I, & Kehua, Y. E. (2022). Application of gwo-svm in wind turbine gearbox fault diagnosis. *Journal of Mechanical Strength*, *43*(5), 1026-1034.

[11] Goyal, S., Yadav, N., Rani, A., & Singh, V. (2022). Classification of SGS-SRAD Denoised MRI Using GWO Optimized SVM. *IETE Journal of Research*, *68*(6), 4383-4393. https://doi.org/10.1080/03772063.2020.1792360

[12] Mao, M., Yang, H., Xu, F., Ni, P., & Wu, H. (2022). Development of geosteering system based on GWO-SVM model. *Neural Computing and Applications*, *34*(15), 12479-12490. https://doi.org/10.1007/s00521-021-06583-6

[13] Zhang, S., Yuan, Y., Wang, Z., Lin, Y., Jiang, L., & Fu, M. (2022). A hierarchical prediction method based on hybrid-kernel GWO-SVM for metal tube bending wrinkling detection. *The International Journal of Advanced Manufacturing Technology*, *121*(7), 5329-5342. https://doi.org/10.1007/s00170-022-09691-2

[14] Dou, Z., Han, R., & Wang, Y. (2023). A prediction method of failure depth of coal seam floor based on FA-GWO-SVM model. *Journal of Engineering Science and Technology Review*, *16*(1), 161-169. https://doi.org/10.25103/jestr.161.21

[15] Durkaya Kurtcan, B. & Ozcan, T. (2023). Predicting customer churn using grey wolf optimization ‐ based support vector machine with principal component analysis. *Journal of Forecasting*, *42*(6), 1329-1340. https://doi.org/10.1002/for.2960

[16] Zhang, K. & Dong, C. (2023). Using neural network to automatic manufacture product label in enterprise under IoT environments. *Computer Science and Information Systems*, *20*(2), 701-722. https://doi.org/10.2298/CSIS220703019Z

[17] Liu, Y., As'arry, A., Hassan, M.K., Hairuddin, A. A., & Mohamad, H. (2024). Review of the grey wolf optimization algorithm: variants and applications. *Neural Computing and Applications*, *36*(6), 2713-2735.

[18] Wang, X. B. &Tian, S. F. (2022). Exotic localized waves in the shifted nonlocal multicomponent nonlinear Schrödinger equation. *Theoretical and Mathematical Physics*, *212*(3), 1193-1210. https://doi.org/10.1134/S0040577922090033

[19] Chinthamu, N. & Karukuri, M. (2023). Data Science and Applications. *Journal of Data Science and Intelligent Systems*, *1*(1), 83-91. https://doi.org/10.47852/bonviewJDSIS3202837

[20] Wang, X. & Guan, N. (2023). 2D sine-logistic-tent-coupling map for image encryption. *Journal of Ambient Intelligence and Humanized Computing*, *14*(10), 13399-13419. https://doi.org/10.1007/s12652-022-03794-0

[21] Doskočil, R. & Doubravský, K. (2023). Using Trend Modelling to Support Project Scheduling, Management, and Prediction. Economic Computation and Economic *Cybernetics Studies and Research*, *57*(4), 75-90. https://doi.org/10.24818/18423264/57.4.23.05

[22] Garai, S., Paul, R.K., Kumar, M., & Choudhury, A. (2023). Intra-Annual National Statistical Accounts Based on Machine Learning Algorithm. *Journal of Data Science and Intelligent Systems*, *2*(2), 12-15.

[23] Zhu, X. Q. (2023). Collaborative modelling and simulation for manufacturing cost analysis. *International Journal of Simulation Modelling*, *22*(2), 338-349. https://doi.org/10.2507/IJSIMM22-2-CO9

[24] Munoz-Ibanez, C., Chairez, I., Jimenez-Martinez, M., Molina, A., & Alfaro-Ponce, M. (2023). Hybrid forecasting modelling of cost and time entities for planning and optimizing projects in the die-cast aluminium industry, *Advances in Production Engineering & Management*, *18*(2), 163-174. https://doi.org/10.14743/apem2023.2.464

[25] Zhang, Y. D., Liao, L., Yu, Q., Ma, W. G., & Li, K. H. (2021). Using the gradient boosting decision tree (GBDT) algorithm for a train delay prediction model considering the delay propagation feature. *Advances in Production Engineering & Management*, *16*(3), 285-296. https://doi.org/10.14743/apem2021.3.400

Contact information:

DaWen ZHANG
School of Civil Engineering, Southwest Jiaotong University Hope College,
ChengDu, SiChuan, China, 610400
E-mail: 18980852708@163.com