

Research Article

Research on the Prediction of the Water Demand of Construction Engineering Based on the BP Neural Network

Hao Peng ¹, Han Wu ², and Junwu Wang ²

¹School of Architectural Engineering, Xinyang Vocational and Technical College, Xinyang 464000, China

²School of Civil Engineering and Architecture, Wuhan University of Technology, Wuhan 430070, China

Correspondence should be addressed to Han Wu; wuhan20170620@163.com

Received 8 July 2020; Revised 15 August 2020; Accepted 13 September 2020; Published 1 November 2020

Academic Editor: Tayfun Dede

Copyright © 2020 Hao Peng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The scientific and effective prediction of the water consumption of construction engineering is of great significance to the management of construction costs. To address the large water consumption and high uncertainty of water demand in project construction, a prediction model based on the back propagation (BP) neural network improved by particle swarm optimization (PSO) was proposed in the present work. To reduce the complexity of redundant input variables, this model determined the main influencing factors of water demand by grey relational analysis. The BP neural network optimized by PSO was used to obtain the predicted value of the output interval, which effectively solved the shortcomings of the BP neural network model, including its slow convergence speed and easy to fall into local optimum problems. In addition, the water consumption interval data of the Taiyangchen Project located in Xinyang, Henan Province, China, were simulated. According to the results of the case study, there were four main factors that affected the construction water consumption of the Taiyangchen Project, namely, the intraday amount of pouring concrete, the intraday weather, the number of workers, and the intraday amount of wood used. The predicted data were basically consistent with the actual data, the relative error was less than 5%, and the average error was only 2.66%. However, the errors of the BP neural network model, the BP neural network improved by genetic algorithm, and the pluralistic return were larger. Three conventional error analysis tools in machine learning (the coefficient of determination, the root mean squared error, and the mean absolute error) also highlight the feasibility and advancement of the proposed method.

1. Introduction

The cost management of construction engineering is a complex problem, and different cost management strategies should be carried out according to the diverse characteristics of construction materials [1]. The cost management of construction water has the following characteristics. (1) The unit price of water is low and almost constant. Therefore, the essence of cost analysis related to construction water is to analyze the construction water demand. (2) The total amount of water used for construction is huge, and ensuring the stability of the water supply during construction is, therefore, of great significance to the smooth progress of construction [2]. In addition, although the cost of construction water accounts for a small proportion of the total cost of construction projects, water plays a major role in the

construction process [3]. The scientific and effective prediction of construction water consumption can not only be used to calculate the cost of construction water scientifically and effectively but can also ensure the stability of the water supply during construction as much as possible, which is also of great significance to the smooth progress of construction.

However, the prediction of construction water consumption is complex. Almost all construction operations require water. Operations such as on-site concrete construction and formwork construction are characterized by high water consumption. Moreover, construction engineering has complex stages, and the water consumption in different construction stages is very different.

The traditional methods for the prediction of industrial water consumption include parametric statistics and

deterministic models, which have usually been incorporated into models for engineering calculation or physical analysis after quantifying the factors that affect construction water consumption. He and Tao [4] established a coupled grey system and multivariate regression model to predict water consumption in Wuhan, which was marked by a clear concept and simple structure. Via the power function model, the linear function model, the logarithmic function model, and the parabolic function model, Zhang et al. [5] fitted the curves of water consumption in China in 2015, 2020, and 2030. The results showed that the correlation coefficient of the parabolic curve fitting was the largest, the average absolute percentage error was the smallest, and the fitting effect was the best. Buck et al. [6] used a statistical sampling method to predict residential water consumption in California. Nevertheless, these research methods were time-consuming and prone to predict errors because they assumed that the relationship between independent and dependent variables was a simple linear process. However, regarding water consumption in construction engineering, the interaction of various variables constitutes a large and compound system with continuous and nonlinear changes. A simplified model would affect the accuracy of the analysis results, thereby resulting in the prediction accuracy of water consumption in construction engineering being unable to meet the needs of engineering practice, and posing a substantial threat to the smooth progress of engineering projects. In addition, traditional water demand forecasting methods need a large amount of complete statistical data to obtain consistent research results. However, hefty amounts of complete statistical data are difficult to obtain on construction sites.

In recent years, the emergence of artificial intelligence algorithms, such as artificial neural networks (ANNs), has provided new ideas for conducting the real-time prediction of the complex system of the water demand of construction projects [7]. The ANN method has a good self-adaptive learning ability and nonlinear mapping ability and is able to fully utilize the potential laws of input data, thereby demonstrating significant advantages for the research and analysis of complex systems with multifactor coupling. When an ANN is applied to the water demand prediction of building engineering construction, the model solves the nonlinear problem of water demand prediction by simulating the structural characteristics and action mechanism of biological neurons and uses the limited data measured in the field instead of a large amount of complete statistical data to predict the water demand by using the data-driven method. At present, relevant scholars have carried out research on water demand forecasting and have obtained rich research results.

Donkor et al. [8] summarized the research results related to urban water demand forecasting and pointed out that scientific and effective forecasting variables are the key to successfully forecasting urban water demand. They also pointed out that the soft computing method yielded valuable research results in the short-term forecasting of water demand. Piasecki et al. [9] compared the ANN and the multiple linear regression (MLR) method, and a case study showed

that the ANN method was superior to MLR. Zhang et al. [10] used the main influencing factors of the predicted daily water consumption as the input and the predicted daily water consumption as the output after noise reduction. In this work, the multiscale chaotic genetic algorithm, which is characterized by a strong global searching ability and fast searching speed, was utilized to optimize the parameters of a least-squares support vector machine. By using the ANN method, Santos and Pereira [11] predicted the urban water demand of São Paulo, Brazil. The research indicated that the ANN model could make accurate predictions with a large amount of data, and it was marked by the best performance and a small error. In addition, Santos considered the influences of weather variables on regional urban water consumption. At present, the application of artificial intelligence prediction methods to water consumption prediction mainly includes the following two ideas: a multiparameter prediction model and a time series-based model. Research on the prediction of drinking water demand in Portugal has shown that the univariate time series model based on historical data is useful and can be combined with other prediction methods to reduce errors [12]. The previously mentioned research has demonstrated that soft computing algorithms, such as ANNs, can better deal with the nonlinear problems in water resource demand management. Therefore, the complex and nonlinear mapping between the factors that affect water resource demand and construction water resource demand can be identified by ANNs.

At present, the back propagation (BP) neural network is the most commonly used neural network [13]. When applied to complex system analysis, the traditional BP neural network might contain the following shortcomings. (1) The traditional BP neural network is an optimization method of local search, and it can easily fall into the local extremum. The weights can easily converge to local minima, which leads to the failure of network training [14]. Furthermore, the BP neural network is highly sensitive to initial network weights. When the network is initialized with different weights, it tends to converge to different local minima [15]. (2) The structure of the BP neural network can only be selected by experience, and there is no unified and complete theoretical guidance for the selection of a BP neural network structure. If the selected network structure is too large, the training efficiency will not be high and a fitting phenomenon may occur [16], resulting in low network performance and reduced fault tolerance. If the selection is too small, the network may not converge [17]. Based on the preceding analysis, the traditional BP neural network should be strengthened to develop a high-precision model.

Similar to other metaheuristic algorithms such as the genetic algorithm (GA), particle swarm optimization (PSO) is a population-based optimization tool that searches for the optimal solution by updating generations. PSO has no evolutionary operators such as crossover or mutation. Therefore, the advantages of PSO are its very simple concept, low calculation cost, and few parameters that require adjustment [18]. In the fields of bottom hole pressure prediction, ground vibration prediction [19], and asphaltene

precipitation prediction [20], PSO has been applied to optimize the initial weights and thresholds of the BP neural network, which results in higher accuracy.

In addition, most of the existing research results related to water demand forecasting have been aimed at large-scale areas, such as cities [11–14]. To the best of the authors' knowledge, research on the water demand forecasting for construction projects has not yet been reported.

Hence, in this paper, a prediction model of the water consumption of construction projects was established based on grey relational analysis and the BP neural network improved by PSO. The main contributions of this paper are as follows. (1) Previous relevant research has primarily focused on the water demand prediction of large-scale areas, such as cities. However, the present study focused on the water demand prediction of construction projects in the construction stage and presented a detailed case analysis of the construction water of the Taiyangchen Project in Xinyang City, Henan Province, China. This provided new insight for water management in construction engineering. (2) In this paper, the grey relational analysis method was adopted to identify the key factors that affect the construction water consumption of building engineering and to reduce the input variables of the BP model. By setting the threshold of the grey relational degree, it was determined that the key factors that affected the water consumption in the Taiyangchen Project were the intraday amount of pouring concrete, the intraday weather, the number of workers, and the intraday amount of wood used. (3) In view of the shortcomings of the BP neural network model, such as its slow convergence speed and easy to fall into local optimum, PSO, which is characterized by a fast convergence speed and easy realization, was adopted to optimize the initial weights and thresholds of the BP neural network, which effectively solved these problems. Additionally, the error analysis in the case study demonstrated that the calculation results of the BP neural network improved by PSO achieved higher accuracy than the classical BP neural network model, the BP neural network improved by GA, and the pluralistic return.

The remaining chapters of this paper are arranged as follows. Section 2 presents the materials and methods, and the fundamental factor identification method based on grey relational analysis and the BP neural network optimized by PSO are constructed in detail to build a water demand prediction model for construction projects. Section 3 presents the results and discussion and makes a detailed case analysis of the construction water of the Taiyangchen Project in Xinyang City, Henan Province, China. Three common error analysis tools in machine learning are employed in this section to compare the computational accuracy of different models, and the influence of the topological structure of the BP network model on the calculation results is discussed. Section 4 presents the conclusion, which summarizes the research results of this paper and indicates future research directions.

2. Materials and Methods

2.1. Identification of Key Factors Based on Grey Relational Analysis. Grey system theory is a systematic scientific theory

put forward by Professor Deng Julong, a famous Chinese scholar. Grey relational analysis is a quantitative description and comparison method of system development and changing situations. Its basic concept is tantamount to judge whether factors are closely related by the geometric similarity of reference data columns and several comparison data columns, which reflect the correlation degree between curves. In the fields of risk assessment and prediction, relational analysis can determine the weight of each influencing factor by comparing the compactness of each index series with the benchmark series [21].

The grey correlation method is in a position to analyze the development trend of a system [22]. This method can extract the factors that have great influences on the system index in a system with poor information and small samples. Grey relational analysis can overcome the problems of the calculation amount being too large, the samples not obeying a certain probability distribution, and the calculations having different quantitative and directional results.

The steps for the use of grey relational analysis to find the main factors that affect water consumption in construction projects are reproduced below.

Step 1. Raw data processing.

In this paper, the interval-valued processing method is used to process the original data of construction water and its influencing factors [23].

Step 2. Calculate the grey correlation coefficient.

The degree of correlation can reflect the shape of the sequence, and the coefficient of the grey correlation of water used for building engineering construction is

$$\theta_{mn} = \frac{\rho\Delta_{n\max} + \Delta_{n\min}}{\rho\Delta_{n\max} + \Delta_{mn}(i)}, \quad (1)$$

where $\Delta_{n\max}$ and $\Delta_{n\min}$ are the maximum and minimum values in the water consumption data series of construction projects, respectively, and ρ is a resolution function, the function of which is to improve the significance of the difference between correlation coefficients. Generally, a satisfactory resolution result can be obtained when the value is 0.5 [24]. Moreover, m is the reference sequence, n is the comparison sequence, s is the sequence length, and $\Delta_{mn}(i)$ is the absolute difference between the reference sequence m and point i of the comparison sequence n .

Step 3. Calculate the correlation degree.

The correlation degree calculation formula is as follows [25]:

$$\lambda(x_m, y_n) = \left(\frac{1}{s}\right) \sum_{i=1}^s \theta_{mn}(i), \quad (2)$$

where s is the length of the reference sequence, $\theta_{mn}(i)$ is the correlation coefficient between the reference sequence m and the i value of the comparison sequence n , and $\lambda(x_m, y_n)$ is the correlation degree between the reference sequence m on the x curve and the comparison sequence n on the y curve.

Step 4. Rank analysis of correlation degree.

The correlation degree of each factor is sorted depending on the numerical value and describes the relative changes of the reference sequence and comparison sequence. Generally, if the correlation degree between two factors is large, the changes of construction water and the influencing factors are essentially the same [25, 26].

2.2. Prediction Model of Water Demand Based on the BP Neural Network Optimized by PSO

2.2.1. Introduction of the BP Neural Network. The BP neural network is an artificial neural network model with self-learning and self-adapting abilities and composes of two parts, namely, the forward propagation of input data and the backward propagation of the error value. The standard neural network topology consists of input layer nodes, hidden layer nodes, and output layer nodes that are connected, and the nodes in the same layer do not interact with each other. In this algorithm, n samples $\mathbf{X} = (x_1, x_2, \dots, x_n)$ are taken as the input nodes of the neural network, and the expected result $\mathbf{Y} = (y_1, y_2, \dots, y_n)$ is taken as the corresponding output node. The error value can be obtained by comparing the predicted result with the actual result, and the fitness function is employed to measure whether the error value is consistent. For the calculation results that do not meet the requirements, the network will use the gradient descent method to carry out error back propagation in the weight vector space, for which the correction amount of each weight of the hidden layer and the output layer [27] is shown in equation (3). The error reaches the expected value through repeated iteration, thus completing the establishment of the BP neural network calculation model.

$$F(\psi, \omega, \theta, r) = \left(\sum_{t=1}^{N_1} \sum_{s=1}^M [y_t(s) - \hat{y}_t(s)] \right)^{-1/2}, \quad (3)$$

where ψ is the error value in the BP neural network, ω , θ , and r are the input layer, hidden layer, and output layer of the neural network, respectively, N_1 and M are the weight and threshold number of nodes, respectively, $y_t(s)$ is the expected output of the neural network, $\hat{y}_t(s)$ is the actual output, t is a node that needs to optimize the connection weight, and s is a node that needs to optimize the threshold.

2.2.2. Optimization of the BP Neural Network by PSO. PSO was first proposed by Eberhart and Kennedy in 1995 [28]. Its basic concept originated from the study of birds' foraging behavior, and PSO was inspired by this biological population behavior to solve the optimization problem. In PSO, each particle represents a solution of the problem and corresponds to a fitness value. Particle velocity determines the distance and direction of particle motion and is dynamically adjusted by the motion of itself and other particles, thus realizing the optimization process of individuals in a solvable space.

In the process of adopting PSO, the error between the capacity output and the expected capacity output obtained

by the forward learning of the BP neural network is first initialized by the PSO to determine the individual extremum and group extremum, i.e., to find the weights and thresholds in the BP neural network. The speed and position are then updated, as are the original individual extremum and group extremum after calculating the fitness. Finally, the obtained optimal neural network weights and thresholds are sent to the BP neural network for verification [29].

Supposing that the particle swarm $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is composed of n particles, and the dimension of the particles is usually Q . There are n particles in the swarm, each particle is Q -dimensional, and the swarm composed of n particles searches Q dimensions. Every particle is expressed as $X_i = (X_{i1}, X_{i2}, \dots, X_{iQ})$, which represents the position of the particle i in the Q -dimensional search space and also a potential solution to the problem. According to the objective function, the fitness value corresponding to each particle position X_i can be calculated [30]. The velocity corresponding to each particle can be expressed as $\mathbf{V} = (V_{i1}, V_{i2}, \dots, V_{iQ})$, and each particle should consider two factors when searching:

- (1) The historical optimal value P_i , $\mathbf{P}_i = (P_{i1}, P_{i2}, \dots, P_{iQ})$, $i = 1, 2, \dots, n$.
- (2) The optimal value P_g , $\mathbf{P}_g = (P_{g1}, P_{g2}, \dots, P_{gQ})$, found by all particles. It is worth noting that there is only one P_g here.

During each iteration, particles update their own velocity and position via the individual extremum and global extremum, and the update formula of position velocity optimized by the PSO is as follows [31]:

$$\begin{aligned} v_{id}^{k+1} &= \omega v_{id}^k + c_1 r (p_{id}^k - x_{id}^k) + c_2 r (p_{gd}^k - x_{id}^k), \\ x_{id}^{k+1} &= x_{id}^k + v_{id}^{k+1}, \end{aligned} \quad (4)$$

where ω is the inertia weight, $d = 1, 2, \dots, D$, $i = 1, 2, \dots, n$, k is the current iteration number, v_{id} is the velocity of particles, c_1 is the particle weight coefficient that tracks its own historical optimum value, which represents the particle's own cognition and is called the acceleration factor, c_2 is the weight coefficient of the optimal value of the particle tracking group, which represents the cognition of particles to the whole group knowledge and is called the acceleration factor, and r is a random number that is uniformly distributed in the interval $[0, 1]$. In addition, error analysis should be carried out on the results.

Based on the preceding analysis, the calculation flow chart of the proposed model is presented in Figure 1.

It can be observed in Figure 1 that, via the analysis of the data, the historical data and several factors that have the greatest influences on the water demand of building engineering construction are input into the neural network. After the neurons in each layer act on the influencing factors, they generate the output. The weights and thresholds of the neural network are optimized by PSO, and the fitness value is obtained to determine the individual with the best fitness. After taking the output error as the objective function and

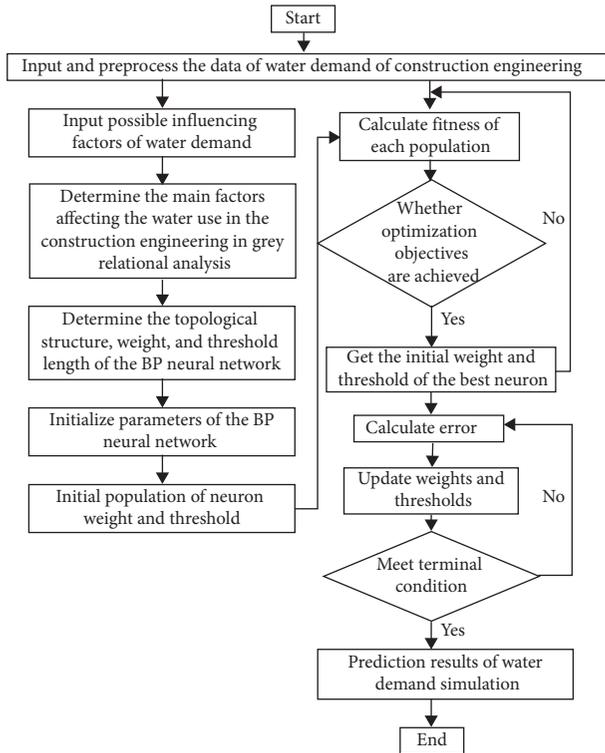


FIGURE 1: Flow chart of the prediction model of water demand.

correcting the error to meet the requirements, the trained neural network can make predictions.

3. Results and Discussion

3.1. Selection of Influencing Factors. There are many factors that affect water consumption in construction areas [32], such as the number of workers (R_1) [33], the intraday amount of pouring concrete (R_2) [34], the highest temperature (R_3) [35], the intraday weather (R_4) [35], the intraday amount of wood used (R_5) [34], and the intraday amount of steel used (R_6) [34].

In this paper, the data of the Taiyangchen Project in May and June 2019 were collected in units of days. The data of the water use of the Taiyangchen Project was processed by grey relational analysis, and the correlation coefficient and degree were obtained. By comparing the sizes, the main factors that affected the water use of the Taiyangchen Project were determined and then used as input layers and input into the neural network to predict the water use of the project.

The daily water consumption of the Taiyangchen Project in Xinyang City, Henan Province, China, is the research object in this work. The project consists of nine residential buildings with frame-shear wall structures, all of which have 18 floors above ground and 1 floor underground. The height of each building is 52.90 m, and the total construction area of each building is 13449.68 m².

The quantification of the evaluation factors is an important step in the selection and treatment of influencing factors. The number of workers (R_1), the intraday amount of pouring concrete (R_2), the highest temperature (R_3), the

intraday amount of wood used (R_5), and the intraday amount of steel used (R_6) were all determined by actual field investigation and statistics. The score of the intraday weather (R_4) is divided into four situations, namely, sunny (0.9), cloudy (0.6), light rain (0.3), and heavy rain (0). The values of each factor are presented in Table 1. Due to the limitations of the layout, only some samples are reported in Table 1.

The correlation coefficients of the influencing factors were calculated by equations (1) and (2), and the calculation results are shown in Table 2.

It can be seen from Table 2 that the correlation degree of the influencing factors from the greatest to the least is as follows: the intraday amount of pouring concrete (R_2) > the intraday weather (R_4) > the number of workers (R_1) > the intraday amount of wood used (R_5) > the highest temperature (R_3) > the intraday amount of steel used (R_6). This order can be explained by the content and characteristics of the construction work. Concrete pouring is a typical wet operation that requires a lot of water. The weather is another significant factor that affects construction. When it rains, most of the work on the construction site will stop, and the construction water consumption will decrease significantly. The more workers there are, the more water will be used for construction and living. Timber for construction needs to be watered and wetted to ensure that its moisture content is near the optimum, which also requires a large amount of water [36].

When the correlation degree is less than 0.6, the two sequences are considered to be independent, and if the correlation degree is greater than 0.8, the two sequences have a good correlation. A correlation value between 0.6 and 0.8 is beneficial [35]. In Table 2, the factors with a correlation degree greater than 0.8 include the intraday amount of pouring concrete (R_2), the intraday weather (R_4), the number of workers (R_1), and the intraday amount of wood used (R_5), and these are, therefore, the key factors that affect the water consumption of building engineering construction.

3.2. Result of Water Demand Forecasting. In this work, the construction water consumption of the Taiyangchen Project in Xinyang City, Henan Province, China, was taken as the research object, and the data used were sourced from the monitoring data of the municipal pipe network water consumption of the Taiyangchen Project and the on-site construction log.

After the evaluation factor is quantified, the difference of its numerical dimension slows the convergence speed of the algorithm and affects the accuracy of the model. Because all indicators are benefit-based, the data are normalized as follows [37]:

$$x_{ij}^* = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (5)$$

where x_{ij}^* indicates the value of the evaluation index after standardization, $\max(x_j)$ represents the maximum value of indicator j , and $\min(x_j)$ represents the minimum value of indicator j .

TABLE 1: Numerical value of each factor and actual water consumption.

No.	R_1	R_2	R_3	R_4	R_5	R_6	Actual water consumption
Unit	—	m^3	$^{\circ}C$	—	m^3	t	t
1 May	228	310.00	28	0.3	13.00	73.46	836.171
2 May	198	1050.00	24	0.6	57.10	72.67	1200.363
3 May	219	1200.00	26	0.6	9.10	83.07	1428.350
4 May	255	730.00	28	0.6	3.10	165.73	567.857
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
27 June	229	50.00	36	0.3	36.10	84.41	752.873
28 June	200	150.00	35	0.3	3.50	44.68	758.394
29 June	215	780.00	32	0.6	18.30	36.38	1601.371
30 June	219	250.00	32	0.3	4.90	107.60	702.361

TABLE 2: Correlation coefficients of influencing factors.

Factor	R_1	R_2	R_3	R_4	R_5	R_6
Correlation coefficient	0.8454	0.8927	0.5150	0.8589	0.8117	0.4625
Mark	Reserved	Reserved	Deleted	Reserved	Reserved	Deleted

According to grey relational analysis, four main influencing factors of water demand in the construction interval of building engineering were obtained. The number of input nodes is $m = 4$, and the number of hidden layer nodes is $n = 2m + 1 = 9$, so the structure of the BP neural network is 4-9-1 [38], as illustrated in Figure 2. It is worth noting that a variety of BP neural network topologies were constructed in this work, and the calculation results are exhibited in Table 3.

In the process of model formation using the BP neural network, the available data should be divided into two groups, which, respectively, represent training and testing sets. The data of the training set are used for training, while the data of the testing set are used for checking the network. Many researchers choose data in the respective proportions of 90% and 10%, 80% and 20%, or 70% and 30% [39]. In this study, the training set was data of the Taiyangchen Project from May 1 to June 20, 2018, including a total of 51 days of data. The testing set was the data of the Taiyangchen Project from June 21 to June 30, 2018, including a total of 10 days of data. The ratio of training set data to testing set data was therefore 83.61%–16.39%.

To achieve a better prediction effect, the best parameters of the BP neural network and the PSO were set, including the following: the number of training iterations was 1000, the learning rate was 0.1, and the training target was 0.001. The calculation parameters [40] of PSO included 1000 iterations, a population size of 50, the local learning factor $c_1 = 1.49445$, and the global learning factor $c_2 = 1.49445$. The maximum error of iteration termination was 0.00001.

The convergence curve obtained after calculation is presented in Figure 3. In a case analysis, when the number of iterations reaches about 500, the requirements are met. In this study, the population number and the maximum number of iterations were set to relatively large values to ensure that the model could calculate more complex problems. Figure 3 shows the convergence curve after 1000 iterations.

Following the optimization calculation process of PSO, the error between the 498th iteration and 499th iteration was

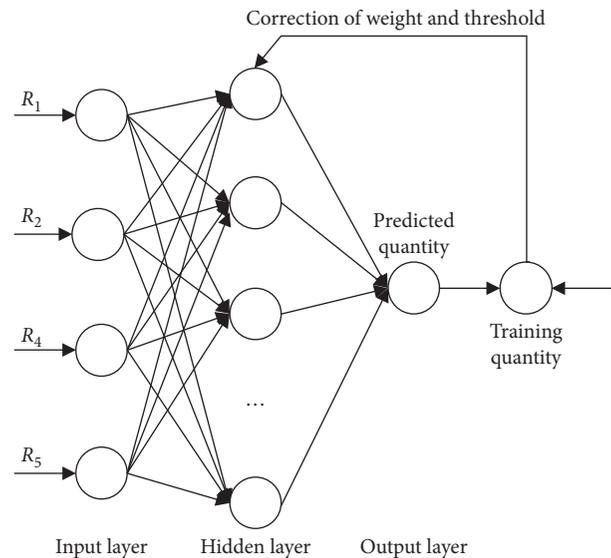


FIGURE 2: Topological structure diagram of the BP neural network.

TABLE 3: Comparison of three error representations.

Error representations	R^2	RMSE	MAE
BP	0.7921	731.2692	45.4554
PSO-BP	0.9959	96.0900	15.7467
GA-BP	0.9853	130.1673	20.3815
Pluralistic return	0.3767	1938.1279	73.0130

greater than the minimum acceptable precision (0.00001), while the error between the 499th iteration and 500th iteration was less than 0.00001. After that, the errors of the calculation results were all less than 0.00001. The calculation was arrested at the 1000th iteration with a very small error. These findings indicate that the PSO found the optimal neural network weights and thresholds at the 500th iteration, which is illustrated by both Figure 3 and Table 4. In addition, it can be qualitatively judged from Figure 3 that the GA

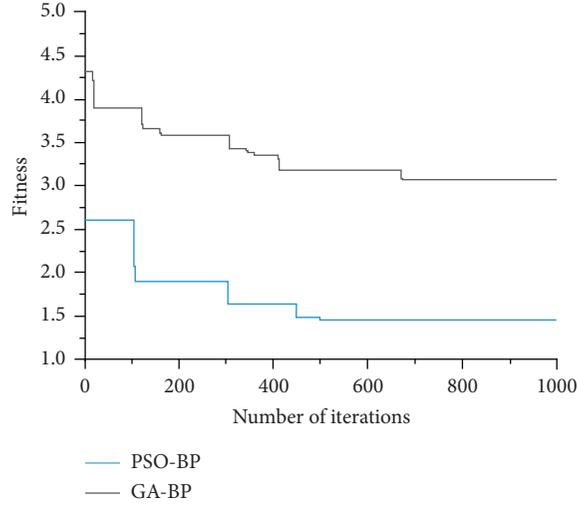


FIGURE 3: The convergence curve of PSO and GA.

TABLE 4: Precision level and calculation termination of PSO at the 1000th iteration.

Iteration (n)	Fitness ($n-1$)	Fitness (n)	Fitness (n) - Fitness ($n-1$)	Result
498	1.488344	1.488344	$0 < 0.00001$	Continue
499	1.488344	1.455848356	$0.032495395 > 0.0001$	Continue
500	1.455848356	1.455848356	$0 < 0.0001$	Continue
1000	1.455848	1.455848	$0 < 0.0001$	Stop

converged between 700 and 800 generations. The PSO converged faster than GA. This is an advantage of PSO compared with GA.

The prediction results obtained by using the proposed calculation model are reported in Table 5. In addition, the BP model, the pluralistic return, and the BP model improved by GA were also used to predict the results. The parameters of GA were set as follows: population size was 100, ending evolution algebra was 1000, crossover probability was 0.5, and mutation probability was 0.001.

In addition, this paper used the standard calculation method which was commonly used in engineering practice, and multivariate linear return method to calculate the quantity used in the construction site.

According to Chinese national standard (the Code for Fire Protection Design of Buildings, GB 50016-2014), the calculation process of construction site water consumption is as follows:

- (1) Water consumption for site operation:

$$q_1 = K_1 \sum \left(\frac{Q_1 \cdot N_1}{T_1 \cdot t} \right) \cdot \left(\frac{K_2}{8 \times 3600} \right), \quad (6)$$

where q_1 (L/s) is the water consumption of construction site, K_1 is the unexpected construction water consumption coefficient, Q_1 (L/s) is the annual engineering quantity, N_1 is the construction water quota (L/m³), T_1 (days) is the annual effective working day, t (hours) is the number of working

shifts per day, and K_2 is the imbalance coefficient of water consumption.

According to the field investigation, K_1 was 1.10, T_1 was 300, and t was 8 in this project. The calculated Q_1 and N_1 were brought into equation (6), and the q_1 was 0.52 L/s.

- (2) Water for construction machinery:

$$q_2 = K_1 \sum Q_2 \cdot N_2 \cdot \left(\frac{K_3}{8 \times 3600} \right), \quad (7)$$

where q_2 is the water consumption of machinery, K_1 is the unexpected construction water consumption coefficient, Q_2 is the same number of machinery, N_2 is the water quota of construction machinery machine-team, and K_3 is the water imbalance coefficient of construction machinery.

According to field investigation, K_3 was 1.5. According to the construction site statistics, Q_2 and N_2 were taken into equation (8), and $q_2 = (0.04 \text{ L/s})$ could be calculated.

- (3) Domestic water consumption on the construction site:

$$q_3 = \frac{P_1 \cdot N_3 \cdot K_4}{t \times 8 \times 3600} \quad (8)$$

where q_3 is the domestic water consumption of the construction site, P_1 is the domestic water

TABLE 5: Prediction results of different models.

Time	Actual water consumption	BP	PSO-BP	GA-BP	Pluralistic return
21 June	1430.37	1397.56	1446.50	1449.50	967.29
22 June	860.32	605.91	857.48	883.10	554.43
23 June	610.36	702.29	581.46	633.88	571.73
24 June	604.13	876.12	615.22	557.52	516.61
25 June	1560.54	1367.86	1528.54	1619.99	463.06
26 June	1910.00	2315.52	1888.37	1876.97	468.56
27 June	752.87	618.75	786.22	779.86	485.53
28 June	758.39	722.45	763.90	746.23	420.71
29 June	1601.37	1409.33	1668.54	1719.87	551.14
30 June	702.36	1029.35	673.01	655.37	450.15

consumption of the construction site, N_3 is the water quota of the construction site, K_4 is the imbalance coefficient of the construction site, and t is the number of working shifts per day.

According to the field investigation, K_4 was 1.3 and t was 1. With equation (8), q_3 was 2.08 L/s.

(4) Domestic water consumption in living quarters:

$$q_4 = \frac{P_2 \cdot N_4 \cdot K_5}{24 \times 3600}, \quad (9)$$

where q_4 is the domestic water consumption in the living area, P_2 is the number of residents in the living area, N_4 is the daily domestic water quota in the living area, and K_5 is the unbalanced coefficient of water consumption in the living area.

After field calculation, P_2 was 150 and K_5 was 2. Bringing complex N_4 into equation (9), q_4 was 0.8 L/s.

(5) Total water consumption Q :

$$Q = q_1 + q_2 + q_3 + q_4. \quad (10)$$

Bring q_1 , q_2 , q_3 , and q_4 into equation (10), and the Q is 3.44 L/s. It is worth mentioning that the fire water was not considered in our calculation here because the fire water was only used in the fire.

Considering that the daily construction time is 8 hours, the water consumption calculated according to Chinese construction codes was 990.72 m³. Comparing the data in Tables 1 and 5, the water consumption calculated by Chinese construction codes was often less than the actual water demand. This situation would easily lead to water shortage and shutdown in the construction site, which was one of the important backgrounds of the research work in this paper. In addition, under the constraints of limited time and timeliness of data collection, it took a long time to investigate the values of more than a dozen parameters by adopting Chinese national calculation standard, which had the disadvantage of low calculation efficiency.

Using the return analysis function of Excel 2016 Software, the expression of multivariate return was calculated as follows:

$$Q = 211.1739 + 0.8247r_1 + 0.1152r_2 + 72.2757r_4 + 1.6083r_5, \quad (11)$$

Bring the data of the test set into equation (11), and the prediction result is as shown in Table 5.

3.3. Analysis and Discussion of Calculation Results. Error analysis was conducted to verify the accuracy of this algorithm, and the relative error value was obtained according to the predicted and actual values. The formula is as follows:

$$E = \frac{|\gamma_p - \gamma_a|}{\gamma_a}, \quad (12)$$

where E is the relative error, γ_p is the predicted value, and γ_a is the true value.

According to equation (12), the relative error value can be calculated by the values present in Table 6. Compared with the actual water consumption, the error of the calculations of the proposed method was less than 5%, and the average error was only 2.47%. The average error of the GA-BP model was 4.06%, and the maximum error was 8.36%. In contrast, the error of the calculations of the BP model was larger, the maximum error was 46.56%, and the average error was 22.39%. The maximum error of the results calculated by multivariate return analysis was 75.47%, and the average error was 41.61%. This proves that the proposed method is effective and advanced in predicting the water demand of construction projects.

The four calculation results of PSO-BP with the biggest error appeared in the last four times. That was to say, the first six predictions were very accurate, and the last four predictions had large errors.

To better compare the prediction results of the two methods and highlight the advantages of the proposed PSO-BP neural network model, three common error analysis tools in machine learning were used to compare several algorithms.

TABLE 6: Prediction results of different models.

Time	BP (%)	PSO-BP (%)	GA-BP (%)	Pluralistic return (%)
21 June	2.29	1.13	1.32	32.38
22 June	26.10	0.33	2.58	35.55
23 June	15.06	0.73	3.71	6.33
24 June	45.02	1.84	8.36	14.49
25 June	12.35	2.05	3.67	70.33
26 June	21.23	1.13	1.76	75.47
27 June	17.82	4.43	3.46	35.51
28 June	25.53	4.74	1.63	44.53
29 June	11.99	4.19	6.89	65.58
30 June	46.56	4.18	7.17	35.91

The determination coefficient (R^2) indicates the degree of correlation between the measured value and the predicted value. The closer R^2 is to 1, the higher the correlation. On the contrary, the closer R^2 is to 0, the lower the correlation. As listed in Table 3, the R^2 values of the BP model, the PSO-BP model, the GA-PSO model, and the pluralistic return were found to be 0.7921, 0.9959, 0.9853, and 0.3767, respectively, thereby demonstrating that the PSO-BP model is better than the BP model, the GA-BP model, and the pluralistic return. The root mean square error (RMSE) is an important standard by which to measure the prediction results of machine learning models. The RMSE of the PSO-BP model was 96.0900, which is notably less than the BP model, the GA-BP model, and the pluralistic return. The mean absolute error (MAE) is the average of absolute errors, which can better reflect the actual situation of predicted value errors. The RMSE of the PSO-BP model was 15.7467, which is notably less than the BP model, GA-PSO, and pluralistic return. The comparative analysis of calculation errors indicates that the PSO-BP model achieved better prediction accuracy and optimization performance. This excellent calculation result is also consistent with previous benchmark test results [41].

The number of hidden layers is another important factor that affects the accuracy of the BP neural network [42]. Therefore, the influences of the topological structures of different network models on the prediction results were also analyzed [38].

Referring to the classical research results in related fields [43], the topological structures of six different network models were designed, and the related calculation results are shown in Table 7.

It can be seen from Table 3 that the calculation accuracy of the PSO-BP neural network model was significantly higher than that of the BP model or GA-BP model when the same network model topology was adopted. Regardless of the topology of the network model, the average error of the calculation results of the PSO-BP neural network model was very small. The calculation results demonstrate the effectiveness and advancement of the PSO-BP neural network model in forecasting the water demand of construction projects. In addition, overfitting or underfitting is a qualitative phenomenon that occurs in artificial neural network algorithms, and there is no tool with which to quantitatively describe them. Referring to the details of previous research studies [41, 44], it is believed that the prediction accuracy of the PSO-

TABLE 7: Calculation results of topological structures of different network models.

Model	Number of hidden layers	Number of hidden layer nodes	Average error
BP	1	9	11.66
	2	9-5	17.37
	2	9-7	9.53
GA-BP	1	9	4.06
	2	9-5	3.65
PSO-BP	2	9-7	2.37
	1	9	2.47
	2	9-5	1.38
	2	9-7	2.31

TABLE 8: Calculation results with different numbers of input variables.

Input variables	Average error (%)	Maximum error (%)
$r_1, r_2, r_4,$ and r_5	4.06	4.74
$r_1, r_2,$ and r_4	7.43	10.15
$r_1, r_2,$ and r_5	9.55	14.28
$r_1, r_4,$ and r_5	8.90	15.69
$r_2, r_4,$ and r_5	7.05	12.73
$r_1, r_2, r_3, r_4,$ and r_5	3.67	4.21
$r_1, r_2, r_3, r_4, r_5,$ and r_6	3.31	3.96

BP neural network model is higher than the calculation accuracy of the BP neural network model, which also demonstrates that the proposed model reduces the phenomenon of overfitting or underfitting. After overfitting or underfitting, the prediction accuracy is often inadequate [45].

Regarding the PSO-BP neural network model, with the increase of the number of hidden layers, the average error was considered to be further reduced. There are two hidden layers, and the model with the 4-9-5-1 network structure achieved the highest calculation accuracy. The accuracy of calculation with a hidden layer model was 2.47%, which was the lowest in the PSO-BP neural network model and can also meet the needs of engineering practice [46].

This paper discussed the influence of the number of input variables on the calculation results. The analysis here adopted the calculation model of PSO-BP, and the number of hidden layer was 1. The average error and maximum error are shown in Table 8.

When there were three input variables, the calculation error of the PSO-BP model was obviously larger than that of four input variables. When the input variables were increased to 5 or 6, the calculation accuracy was not significantly improved. Considering the availability of construction site data, the field investigation time would increase significantly if there were 5 or 6 input variables. Therefore, it could be considered that the four input variables obtained by grey relational analysis in this paper were reasonable.

4. Conclusions

The purpose of this research was to use the BP neural network to accurately predict the water consumption of

construction projects. First, via the use of the data, it was found that there are four factors that affect the water consumption in construction projects, namely, the intraday amount of pouring concrete, the intraday weather, the number of workers, and the intraday amount of wood used. Then, after taking these four key factors as the input layer and using the optimal neural network weights and thresholds obtained by PSO, a predictive model of construction water consumption based on the neural network model was constructed. Finally, a case study of the construction water consumption of the Taiyangchen Project in Xinyang City, Henan Province, China, revealed that, compared with the actual water consumption, the error calculated by the proposed method was less than 5%, and the average error was only 2.47%. In addition, the three common error analysis tools used in machine learning (the coefficient of determination, the root mean squared error, and the mean absolute error) all highlighted that the calculation accuracy of the proposed method was significantly higher than the BP algorithm, the GA-BP, and the pluralistic return. There were two hidden layers in the PSO-BP neural network model, and the model with the 4-9-5-1 network structure was found to have the highest calculation accuracy. The calculation accuracy of the model with a hidden layer and a network structure of 4-9-1 was 2.47%, which can also meet the needs of engineering practice. The model proposed in this paper can effectively predict the water consumption of building engineering construction and determine abnormal water in a timely manner to rationally dispatch the water supply and ultimately achieve the purpose of saving water. In future research, the sample set can be expanded, the learning effect of the model can be improved, and a more perfect prediction model of construction water consumption can be trained.

Data Availability

The case analysis data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the Science and Technology Project of Wuhan Urban and Rural Construction Bureau, China (201943).

References

- [1] J.-S. Chou, "Cost simulation in an item-based project involving construction engineering and management," *International Journal of Project Management*, vol. 29, no. 6, pp. 706–717, 2011.
- [2] I. Ghalekhondabi, E. Ardjmand, W. A. Young et al., "Water demand forecasting: review of soft computing methods," *Environmental Monitoring and Assessment*, vol. 189, no. 7, Article ID 313, 2017.
- [3] P. W. Jayawickrama and S. Rajagopalan, "Alternative water sources in Earthwork construction," *Transportation Research Record*, vol. 2004, pp. 88–96, 2007.
- [4] F. He and T. Tao, "An improved coupling model of grey-system and multivariate linear regression for water consumption forecasting," *Polish Journal of Environmental Studies*, vol. 23, no. 4, pp. 1165–1174, 2014.
- [5] X. Zhang, M. Yue, Y. Yao, and H. Li, "Regional annual water consumption forecast model," *Desalination and Water Treatment*, vol. 114, pp. 51–60, 2018.
- [6] S. Buck, M. Auffhammer, H. Soldati et al., "Forecasting residential water consumption in California: rethinking model selection," *Water Resources Research*, vol. 56, no. 1, 2020.
- [7] X. Wu, V. Kumar, J. Ghosh et al., "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.
- [8] E. A. Donkor, T. A. Mazzuchi, R. Soyer, and J. Alan Roberson, "Urban water demand forecasting: review of methods and models," *Journal of Water Resources Planning and Management*, vol. 140, no. 2, pp. 146–159, 2014.
- [9] A. Piasecki, J. Jurasz, and B. Kazmierczak, "Forecasting daily water consumption: a case study in town, Poland," *Periodica Polytechnica-Civil Engineering*, vol. 62, no. 3, pp. 818–824, 2018.
- [10] W. Zhang, Q. Yang, M. Kumar, and Y. Mao, "Application of improved least squares support vector machine in the forecast of daily water consumption," *Wireless Personal Communications*, vol. 102, no. 4, pp. 3589–3602, 2018.
- [11] C. C. does Santos and A. J. Pereira, "Water demand forecasting model for the metropolitan area of São Paulo, Brazil," *Water Resources Management*, vol. 28, no. 13, pp. 4401–4414, 2014.
- [12] A. Sardinha-Lourenço, A. Andrade-Campos, A. Antunes, and M. S. Oliveira, "Increased performance in the short-term water demand forecasting through the use of a parallel adaptive weighting strategy," *Journal of Hydrology*, vol. 558, pp. 392–404, 2018.
- [13] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of Microbiological Methods*, vol. 43, no. 1, pp. 3–31, 2000.
- [14] Z. Zhao, Q. Xu, and M. Jia, "Improved shuffled frog leaping algorithm-based BP neural network and its application in bearing early fault diagnosis," *Neural Computing and Applications*, vol. 27, no. 2, pp. 375–385, 2016.
- [15] Z. Ma, X. Song, R. Wan, L. Gao, and D. Jiang, "Artificial neural network modeling of the water quality in intensive Litopenaeus vannamei shrimp tanks," *Aquaculture*, vol. 433, pp. 307–312, 2014.
- [16] P. Bansal, S. Gupta, S. Kumar et al., "MLP-LOA: a metaheuristic approach to design an optimal multilayer perceptron," *Soft Computing*, vol. 23, no. 23, pp. 12331–12345, 2019.
- [17] S. Yang, X. Zhu, L. Zhang, L. Wang, and X. Wang, "Classification and prediction of Tibetan medical syndrome based on the improved BP neural network," *IEEE Access*, vol. 8, pp. 31114–31125, 2020.
- [18] D. T. Bui, N. D. Hoang, T. D. Pham et al., "A new intelligence approach based on GIS-based multivariate adaptive regression splines and metaheuristic optimization for predicting flash flood susceptible areas at high-frequency tropical typhoon area," *Journal of Hydrology*, vol. 575, pp. 314–326, 2019.

- [19] D. J. Armahani, M. Hajihassani, E. T. Mohamad et al., "Blasting-induced flyrock and ground vibration prediction through an expert artificial neural network based on particle swarm optimization," *Arabian Journal of Geosciences*, vol. 7, no. 12, pp. 5383–5396, 2014.
- [20] M. A. Ahmadi and S. R. Shadzadeh, "New approach for prediction of asphaltene precipitation due to natural depletion by using evolutionary algorithm concept," *Fuel*, vol. 102, pp. 716–723, 2012.
- [21] Z. Zang, Y. Wang, and Z. X. Wang, "A grey TOPSIS method based on weighted relational coefficient," *Journal of Grey System*, vol. 26, no. 2, pp. 112–123, 2014.
- [22] Z. X. Wang, "Correlation analysis of sequences with interval grey numbers based on the kernel and greyness degree," *Kybernetes*, vol. 42, no. 1-2, pp. 309–317, 2013.
- [23] G. Deng, J. Qiu, G. Liu, and K. Lv, "Environmental stress level evaluation approach based on physical model and interval grey association degree," *Chinese Journal of Aeronautics*, vol. 26, no. 2, pp. 456–462, 2013.
- [24] N. M. Xie and S. F. Liu, "Research on evaluations of several grey relational models adapt to grey relational axioms," *Journal of Systems Engineering and Electronics*, vol. 20, no. 2, pp. 304–309, 2009.
- [25] B. Zhu, L. Yuan, and S. Ye, "Examining the multi-timescales of European carbon market with grey relational analysis and empirical mode decomposition," *Physica A: Statistical Mechanics and its Applications*, vol. 517, pp. 392–399, 2019.
- [26] K. Zhang, Y. Chen, and L. F. Wu, "Grey spectrum analysis of air quality index and housing price in Handan," *Complexity*, vol. 2019, Article ID 8710138, , 2019.
- [27] S. M. Huang, "Production capacity prediction based on neural network technology in an efficient economic and management environment of oil field," *Fresenius Environmental Bulletin*, vol. 29, no. 4, pp. 2442–2449, 2020.
- [28] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 256–279, 2004.
- [29] C. Hou, X. Yu, Y. Cao, C. Lai, and Y. Cao, "Prediction of synchronous closing time of permanent magnetic actuator for vacuum circuit breaker based on PSO-BP," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 6, pp. 3321–3326, 2017.
- [30] Z. H. Zhan, J. Zhang, Y. Li et al., "Adaptive particle swarm optimization," *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, vol. 39, no. 6, pp. 1362–1381, 2009.
- [31] F. vandenBergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 225–239, 2004.
- [32] A. Piasecki, J. Jurasz, and W. Marszelewski, "Application of multilayer perceptron artificial neural networks to mid-term water consumption forecasting - a case study," *Ochrona Srodowiska*, vol. 38, no. 2, pp. 17–22, 2016.
- [33] M. E. Banihabib and P. Mousavi-Mirkalaei, "Extended linear and non-linear auto-regressive models for forecasting the urban water consumption of a fast-growing city in an arid region," *Sustainable Cities and Society*, vol. 48, Article ID 101585, 2019.
- [34] M. Y. Han, G. Q. Chen, J. Meng, X. D. Wu, A. Alsaedi, and B. Ahmad, "Virtual water accounting for a building construction engineering project with nine sub-projects: a case in E-town, Beijing," *Journal of Cleaner Production*, vol. 112, pp. 4691–4700, 2016.
- [35] L. N. Yang, W. Z. Li, and X. Liu, "On campus interval water demand prediction based on grey genetic BP neural network," *Journal of Water Resources and Water Engineering*, vol. 30, no. 3, pp. 133–138, 2019.
- [36] R. Walker, S. Pavia, and M. Dalton, "Measurement of moisture content in solid brick walls using timber dowel," *Materials and Structures*, vol. 49, no. 7, pp. 2549–2561, 2016.
- [37] D. Liu, J. P. Feng, H. Li et al., "Spatiotemporal variation analysis of regional flood disaster resilience capability using an improved projection pursuit model based on the wind-driven optimization algorithm," *Journal of Cleaner Production*, vol. 241, Article ID 118406, 2019.
- [38] W. Cheng, M. Zhou, J. J. Ye et al., "PSO-BP modeling research on fee rate measurement of construction project safe construction cost," *China Safety Science Journal*, vol. 26, no. 5, pp. 146–151, 2016.
- [39] G. Zhang, B. Eddy Patuwu, and M. Y. Hu, "Forecasting with artificial neural networks," *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.
- [40] I. C. Trelea, "The particle swarm optimization algorithm: convergence analysis and parameter selection," *Information Processing Letters*, vol. 85, no. 6, pp. 317–325, 2003.
- [41] H. H. Li, Y. D. Lu, C. Zheng et al., "Groundwater level prediction for the arid oasis of Northwest China based on the artificial Bee Colony algorithm and a back-propagation neural network with double hidden layers," *Water*, vol. 11, no. 4, Article ID 860, 2019.
- [42] A. Kaveh and H. Servati, "Design of double layer grids using backpropagation neural networks," *Computers & Structures*, vol. 79, no. 17, pp. 1561–1568, 2001.
- [43] K. M. Neaupane and S. H. Achet, "Use of back propagation neural network for landslide monitoring: a case study in the higher Himalaya," *Engineering Geology*, vol. 74, no. 3-4, pp. 213–226, 2004.
- [44] Q. Shen, W.-m. Shi, X.-p. Yang, and B.-x. Ye, "Particle swarm algorithm trained neural network for QSAR studies of inhibitors of platelet-derived growth factor receptor phosphorylation," *European Journal of Pharmaceutical Sciences*, vol. 28, no. 5, pp. 369–376, 2006.
- [45] X. F. Yan, "Hybrid artificial neural network based on BP-PLSR and its application in development of soft sensors," *Chemosometrics and Intelligent Laboratory Systems*, vol. 103, no. 2, pp. 152–159, 2010.
- [46] F. Y. Jiang, Y. L. Zhao, S. Dong et al., "Prediction of submarine pipeline damage based on genetic algorithm," *Transactions of Oceanology and Limnology*, vol. 3, pp. 52–59, 2019.