A Building Carbon Emission Prediction Model Based on Optimized Machine Learning Model

Guo Zhanwei[1*]

[1]Zhongke Software Testing (Guangzhou) Co., Ltd., 510000, China

Corresponding Email: wmeiding@outlook.com

https://doi.org/10.70695/41ts7860

Abstract

To improve the performance of traditional machine learning model, this study employs the chaotic particle swarm optimization (CPSO) and the fuzzy cuckoo search (FCS) algorithm to enhance the BP neural network and support vector machine (SVM), leading to CPSO-BP and FCS-SVM model, respectively. Simulation results indicate that both optimization algorithms significantly enhance the predictive accuracy of their respective models. The relative error of the FCS-SVM decreases by 2.68%, while the CPSO-BP model reduces by 9.75% compared to their pre-benchmark model. Although the CPSO-BP model demonstrates a more substantial improvement in the simulation, the performance of FCS-SVM remains superior to that of CPSO-BP model. Subsequently, the FCS-SVM model was selected to forecast carbon dioxide emissions in China's construction industry from 2022 to 2025, and the results were thoroughly analyzed. The findings indicate that total carbon emissions are expected to rise over the next four years; however, the growth rate is projected to slow significantly. Additionally, carbon emission intensity is anticipated to continue declining until 2023, although the rate of decline will be minimal.

Keywords Building carbon emission; CPSO-BPNN; FCS-SVM; Optimization algorithm; Machine learning; Emission reduction

1 Introduction

During the period of rapid economic development, extensive reliance on fossil fuels has resulted in excessive carbon dioxide emissions, exacerbating issues such as global warming and haze pollution [1]. The latest climate change report released by the Intergovernmental Panel on Climate Change (IPCC) in 2018 highlights that limiting global warming to 1.5 °C offers significant benefits for both humanity and natural ecosystems, fostering a more equitable and sustainable society [2]. To achieve the goal of a 45% reduction in global carbon emissions by 2030 relative to 2010 levels, transformative changes are required across various sectors, including industry and construction [3]. The Chinese government has committed to striving for peak carbon emissions around 2030, with an emphasis on achieving this milestone earlier if possible. Furthermore, it aims for a reduction in carbon emissions per unit of GDP by 60% to 65% compared to levels seen in 2005, underscoring China's resolve in establishing stringent total carbon emission control policies [4]. As a pillar industry of the national economy, the construction industry is a primary contributor to carbon emissions in China. Thus, comprehensive management of its energy consumption and carbon emissions is crucial for promoting low-carbon development within the sector [5]. Effectively controlling carbon dioxide emissions in the construction industry necessitates thorough calculation and forecasting of these emissions [6]. Enhancing the accuracy of carbon dioxide emissions prediction models for the construction sector not only supports the formulation of policies and guidelines relevant to energy conservation and emissions reduction but also aids in environmental pollution control measures [7].

Research on carbon dioxide emission prediction remains limited, with studies employing diverse methodologies. Chung et al. developed a RES model for predicting carbon dioxide emissions in the shipbuilding industry, utilizing scenario analysis to forecast emissions from 2014 to 2020 [8]. Wakivama et al. applied time series analysis to investigate carbon emissions resulting from electricity consumption in Japan's residential sector from 2016 to 2030, concluding that emissions are expected to reach 55.4 million tons by 2030 [9]. Similarly, Mirzaei et al. employed both scenario analysis and system dynamics to predict Iran's carbon dioxide emissions from 2015 to 2025, with projections indicating total emissions of 985 million tons in 2025 [10]. Tudor utilized seven models, including the ETS and STS models, to forecast Bahrain's carbon dioxide emissions for the period from 2016 to 2021, revealing that per capita emissions would reach 20.96 metric tons per year by 2020 [11].

Research in carbon emission prediction predominantly relies on quantitative modeling methods, with the STIRPAT model frequently applied in such analyses. The STIRPAT model, an enhancement of the IPAT environmental pollution model, addresses certain limitations inherent in the original IPAT framework. When examining diverse carbon emission prediction challenges, it is essential to compare and analyze the applicability of these models based on the data and specific requirements of the research question at hand.

Moubarak et al. utilized a long-term balance equation to forecast carbon dioxide emissions from the textile industry under various scenarios — normal, medium, and optimal. Their findings indicated that emissions under the medium scenario were the lowest, with expectations of a reduction of 44.8 million tons by 2025 [12]. Hamzacebi et al. predicted Türkiye's energy consumption for 2015, 2020, and 2025 through an optimized grey prediction model, subsequently calculating future carbon dioxide emissions using an energy consumption-emission coefficient approach [13]. Boran et al. introduced a rolling mechanism to enhance the grey prediction model, applying this improved model to forecast natural gas energy consumption in Türkiye from 2014 to 2020. Dalton et al. employed the STIRPAT model in conjunction with scenario analysis to predict carbon emissions across three scenarios (high aging, medium aging, and low aging) for the U.S. population, concluding that population aging could lead to a significant reduction in long-term carbon emissions, with a nearly 40% decrease projected under the high aging scenario [14].

Existing research on the measurement and prediction of carbon dioxide emissions in the construction industry encompasses various definitions of emission boundaries as well as diverse measurement models, such as the STIRPAT model, system dynamics model, logistic regression, and GM (1,1) [15]. Additionally, while there are some algorithms introduced in the context of carbon emissions prediction, their number is limited. Previous studies have primarily highlighted two critical areas of deficiency in this research field [16]. First, the predominant focus on classical regression statistical models for predicting carbon dioxide emissions from the construction industry, while straightforward and expeditious, often suffers from poor fault tolerance and low predictive accuracy [17]. Second, the design of existing relevant algorithms tends to be relatively simplistic, with many relying on single-algorithm prediction models. While each algorithm has its distinct advantages and disadvantages, combining multiple algorithms has the potential to mitigate the limitations associated with any single algorithm [18-20].

Consequently, our research focuses on the optimization and application of carbon dioxide emission prediction models within the construction industry. Specifically, we address the limitations associated with two prediction models by employing chaos particle swarm optimization and the fuzzy cuckoo algorithm to enhance their performance. Following simulation, we compare various error indicators to identify the optimal model for forecasting carbon dioxide emissions in the construction sector over the next five years.

The organization of our study is as follows: Section 2 details the data collection processes and methodologies employed, while Section 3 presents the performance of the improved methods, alongside a comparison with traditional approaches. Finally, we conclude our findings in Section 4.

2 Data Source

Carbon dioxide emissions in the construction industry are categorized into direct and indirect emissions [21-22]. Direct carbon dioxide emissions arise from the combustion of energy sources such as raw coal, coke, fuel oil, diesel, natural gas, gasoline, and kerosene during construction and demolition activities [23]. In contrast, indirect carbon dioxide emissions are associated with other industries closely linked to the construction sector, primarily originating from the production of building materials—including cement, steel, wood, glass, and aluminum—that are essential to construction [24-26].

For the purpose of data collection feasibility, the direct energy consumption of the construction industry is quantified based on raw coal, coke, gasoline, kerosene, diesel, fuel oil, and natural gas [27]. Indirect carbon dioxide emissions are attributed to the production processes of five key building materials—cement, steel, glass, wood, and aluminum—derived from other industries [28-30]. The carbon emission coefficients for various energy sources and building materials are presented in Table 1.

Table 1. Carbon emission coefficient of various energy and building materials.

Names	raw coal	coke	fuel oil	diesel oil	natural gas	gasoline	kerosene	cement	steel products	Glass	wood	Aluminum
Carbon emission coefficient	0.5394 (kgC/kg)	0.8303 (kgC/kg)	0.8823 (kgC/kg)	0.8616 (kgC/kg)	0.5956 (kgC/kg)	0.8140 (kgC/kg)	0.8393 (kgC/kg)	0.822 (Kg/kg)	1.789 (kg/kg)	0.966 (Kg/kg)	-842 Kg/m^3	2.6 Kg/kg

The production processes of certain recyclable building materials do generate carbon emissions; however, these emissions can be mitigated through recycling during the building demolition phase. Consequently, the carbon emissions incurred should be incorporated into the life cycle assessment of new buildings, as this portion of emissions does not result in actual environmental harm. Given that aluminum and steel are recyclable materials, it is essential to calculate the carbon emissions of these materials based on actual consumption while excluding the recyclable content. The recovery coefficients for steel and aluminum are reported as 0.8 and 0.85, respectively. Furthermore, wood, a widely used building material, absorbs significant amounts of carbon dioxide through photosynthesis during its growth, thereby contributing positively to environmental protection. As a result, the carbon emission coefficient for wood is considered to be zero.

The data regarding the usage of building materials and various types of energy consumption are sourced from the China Energy Statistical Yearbook and the China Construction Industry Statistical Yearbook, covering the period from 1995 to 2018. The default carbon dioxide emission factors for various energy sources are referenced from the 2006 IPCC National Greenhouse Gas Emission Inventory. Detailed data on carbon dioxide emissions generated by the construction industry in China from 1995 to 2018 can be found in Table 2.

Table 2. CARBON emissions from China's construction industry.

year	Carbon emission (10000 tons)
1995	12651.7
1996	15295.4
1997	16208.8
1998	20590.4
1999	17675.5
2000	176789.5
2001	27569.5
2002	28956.5
2003	38595.6
2004	46049.4
2005	54366.5
2006	57711.4
2007	68494.5
2008	72367.9
2009	91720.5
2010	102982.2
2011	132564.5
2012	245663.9
2013	301256.9
2014	201299.5
2015	175566.5
2016	162566.5
2017	156255.4
2018	145656.5
2019	134669.5
2020	123545.1

The data pertaining to the influencing factors analyzed in this study were sourced from the China Statistical Yearbook, China Construction Statistical Yearbook, and China Energy Statistical Yearbook, covering the period from 1995 to 2020. In total, twelve influencing factors were identified. Using the original data on relevant factors affecting carbon dioxide emissions within the construction industry, we established a random forest feature selection model to compute the Gini index for each influencing factor, subsequently deriving their importance rankings as illustrated in Fig. 1. The hierarchy of importance is as follows: completed area of buildings in the construction industry, GDP, gross output value of the construction industry, number of employees in construction enterprises, labor productivity in the construction industry, primary energy consumption in the construction industry, energy emission intensity, industrial structure, technical equipment rate of construction enterprises, urbanization rate, resident consumption index, and population.

Based on these findings, the completed area of buildings in the construction industry emerged as the most significant factor influencing emissions. Other key determinants include GDP, total output value of the construction industry, number of employees in construction enterprises, labor productivity in the construction industry, and primary energy consumption in the construction sector. To enhance the accuracy of the carbon dioxide emission prediction model for the construction industry, this study retains the top 50% of the influencing factors as input characteristic variables, thereby mitigating potential redundancy that could diminish model precision.

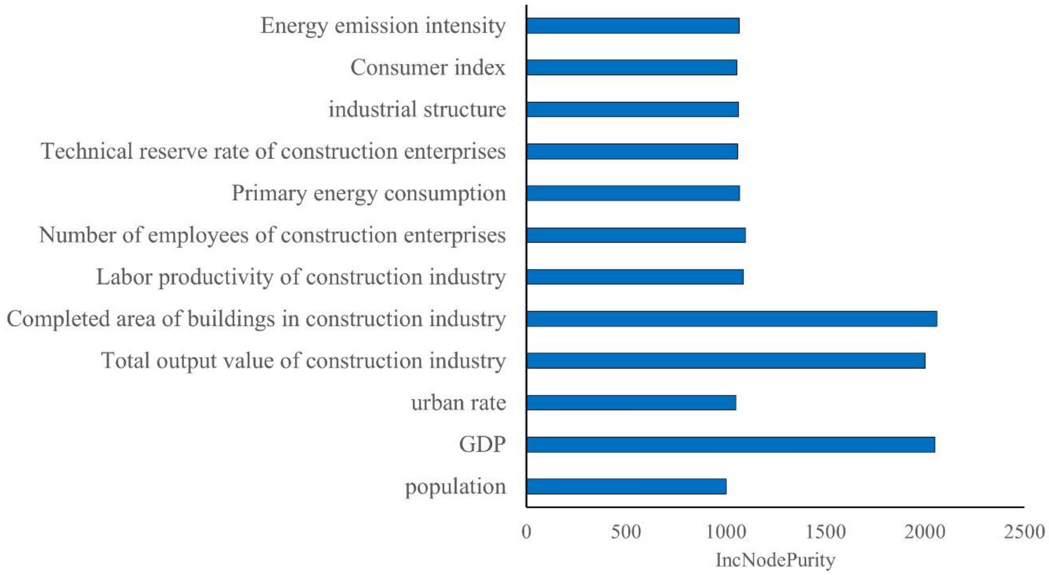

Fig. 1. importance of influential factors.

3 Methodology

3.1 Chaotic Particle Swarm Optimization Algorithm

The traditional particle swarm optimization (PSO) algorithm has been widely applied in model optimization; however, it exhibits limitations when addressing complex models [31]. As the number of iterations increases, the algorithm is prone to premature convergence and local optima, which significantly impairs the optimization results. To address these issues, researchers have incorporated chaotic mapping into the iterative process of particle swarm optimization, leading to the development of chaotic particle swarm optimization (CPSO). This algorithm leverages various chaotic optimization techniques, wherein the fundamental principle involves projecting the parameters to be optimized into chaotic space through a chaotic mapping relationship. This transformation allows the parameters to be converted into chaotic variables, enabling the search for optimal parameter values within the chaotic space, which are subsequently mapped back to the original space via the chaotic mapping relationship [32].

Typically, logic mapping is employed to facilitate the mapping of parameters into chaotic space, thereby converting them into chaotic variables. The logic mapping can be expressed mathematically in equation (1).

$$Z = a_{n+1} = ka_n(1 - a_n) \tag{1}$$

Z is the corresponding variable of the chaotic space, $a0$ is the initial value of the iterative process, an is the value in its iterative process, and k is the parameter of the control topological ability. Given the initial value of $a0$, the chaotic sequence Z can be iterated according to the above formula: $a0$, $a1$, $a2$... am. According to the above analysis, the chaotic variable will vary due to the different initial value $a0$. Even if the initial value is slightly different, it will also affect the motion track of the chaotic sequence.

The particle swarm optimization algorithm is easy to fall into local optimization, which can be optimized by chaos. If the iterative process falls into the local optimum, the chaos algorithm will set the optimal value at this time as its initial value, generate chaotic sequence through the logic map, and conduct chaos traversal, so that the particle optimization process jumps out of the local optimum. In addition, chaos can also increase the diversity of the initial population. The chaos optimization particle swarm optimization algorithm using Logic mapping can be expressed as:

$$Z \rightarrow Y: X = gbest + R \cdot (Z - 0.5) \tag{2}$$

Where gbest is globally optimal value, R is chaos traversal radius. The chaotic variable Z is generated through the logic mapping, and the chaotic variable Z is a random sequence between [0,1]. Then, according to the mapping relationship, the chaotic traversal search is carried out near the global optimal value, so that the particle swarm optimization algorithm returns to the global optimal value. Set the initial values of $a0$, $a0 \in [0,1]$, get the chaotic variables Z through the logic mapping: $a0$, $a1$, $a2$... am, whose values are between [0,1], and then use the linear mapping relationship to form a new chaotic sequence in the particle definition domain, which can enrich the diversity of initial particles.

$$Z \to X: X = a + (b - a)Z \tag{3}$$

Chaotic particle swarm optimization algorithm uses chaos process to change the defect that traditional particle swarm optimization algorithm is easy to fall into local extremum, and improves the optimization performance of the model. See Fig.2 for parameter optimization process of chaos particle swarm optimization algorithm.

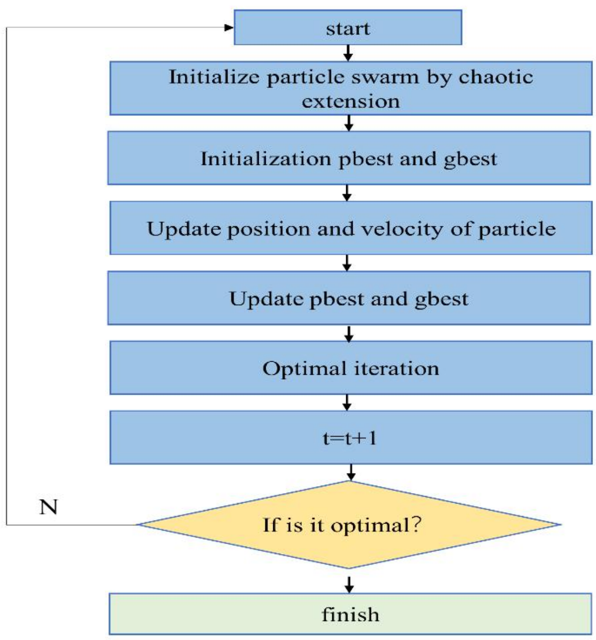

Fig. 2. Chaotic particle swarm optimization algorithm flow.

To address the limitation of the BP neural network that causes it to easily become trapped in local optima, a chaotic particle swarm optimization (CPSO) algorithm is employed for optimization and enhancement. This CPSO algorithm is utilized to optimize the weights within the network structure and the thresholds of the neurons. By mapping data into chaotic space for exploration, the CPSO algorithm helps avoid local optima and enhances the model's predictive performance through chaotic traversal.

The specific optimization approach involves representing the parameters of the optimization problem as vectors in a defined order. These vectors are treated as particles within the swarm, with the dimensionality of each particle corresponding to the number of optimization parameters. The algorithm seeks both individual and global optima, enhancing the chaos of the optimal particle before updating its state. The termination of the iterative process is determined based on a predefined objective function, after which the optimal weights and thresholds are returned. These optimal parameters are then input into the BP neural network to create the optimized BP neural network model. The training process of the BP neural network, optimized using the chaotic particle swarm optimization, is illustrated in Fig. 3.

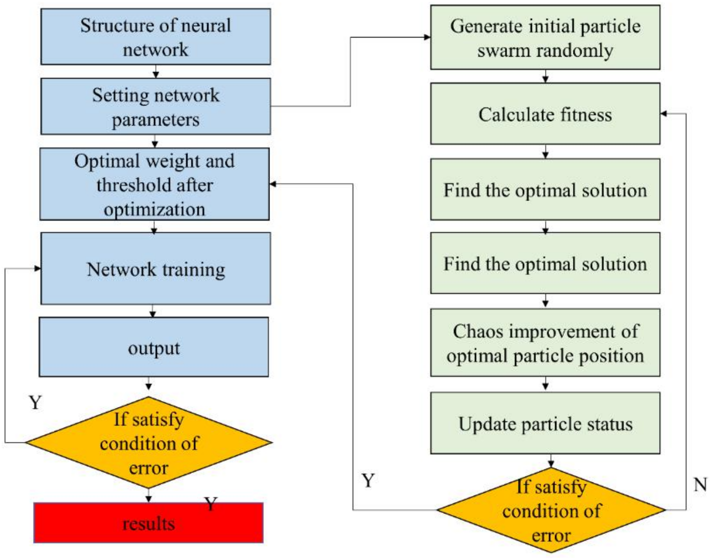

Fig. 3. Chaotic particle swarm optimization of BP neural network process.

3.2 Support Vector Machine Model Optimized by Fuzzy Cuckoo Search Algorithm

The standard cuckoo search algorithm uses Levi flight for global search. Levi flight has infinite mean and variance, which ensures that the cuckoo search algorithm can search the search space more effectively and find the global optimum [33]. However, CS algorithm has problems such as slow convergence speed, low solution accuracy, and poor population diversity in the later stage of evolution. In order to overcome this defect, the fuzzy cuckoo search algorithm (FCS) adjusts the discovery probability P_a by calling the fuzzy controller to make it have faster convergence speed and calculation accuracy [34]. That is, in each generation of CS update, first use two indicators to evaluate the diversity of the population as the input of the fuzzy controller, then use the fuzzy reasoning in the fuzzy controller to analyze the information in the knowledge base and adjust the discovery probability P_a in combination with the fuzzy control rules in the rule base, and finally carry out the fuzzy solution to obtain the dynamic discovery probability P_a, and finally get the global optimal solution.

There are two core parameters that have a great impact on the prediction performance of the construction industry carbon dioxide emission prediction model based on support vector machines [35], one is the core parameter σ, and the other is the penalty parameter C. The optimization direction of the support vector machine model is to use the fuzzy cuckoo search algorithm to search the best parameters of the model's decision function, get the best values of the penalty parameters and the kernel parameters, avoid the impact of subjective settings, and use this idea to build the optimized support vector machine construction carbon dioxide prediction model [36].

Establish the FCS-SVM carbon dioxide emission prediction model for the construction industry, carry out data pretreatment first, normalize the data to eliminate dimensional differences, and divide the data into two parts, one is the training sample, the other is the test sample. Then set the other parameters besides the two core parameters of the support vector machine algorithm. The kernel function is RBF function, the SVM setting type is epsilon-SVR (3), and the loss function value of e-SVR is 0.01. Finally, the fuzzy cuckoo search algorithm and support vector machine are combined to find the optimal values of the two core parameters, and the optimized support vector machine prediction model is established. The specific modeling process is divided into the following steps (details in Fig.4):

Set the relevant parameters of FCS, in which the number of bird nests is 25, the setting accuracy is 0.00001, the discovery probability is 0.25, and the search space dimension is 2.

Normalize the carbon dioxide emission data of the construction industry and relevant influencing factors.

Retain the optimal position of the previous generation, and update the remaining solution with levy flight, thus obtaining a new solution. Compare and analyze the two sets of solutions, and replace the poor solution with the better solution in the new solution, so as to obtain a new set of solutions (n=1,2,3, n).

Call the fuzzy logic controller to adjust the discovery probability P_a. Comparing the random number r with the adjusted discovery probability P, r follows a uniform distribution [0,1]. If $P_a>r$, a solution is randomly generated; If $P_a<r$, the solution is not changed, and a set of solutions is obtained. Compare the two sets of solutions, and replace the difference solution in the previous generation with the good solution in this set of solutions, so as to obtain a new set of solutions.

If the good solution meets the accuracy requirements at this time, the result will be output; otherwise, go back to step 2.

The optimal parameters obtained by FCS will be returned to the support vector machine model for predictive learning training.

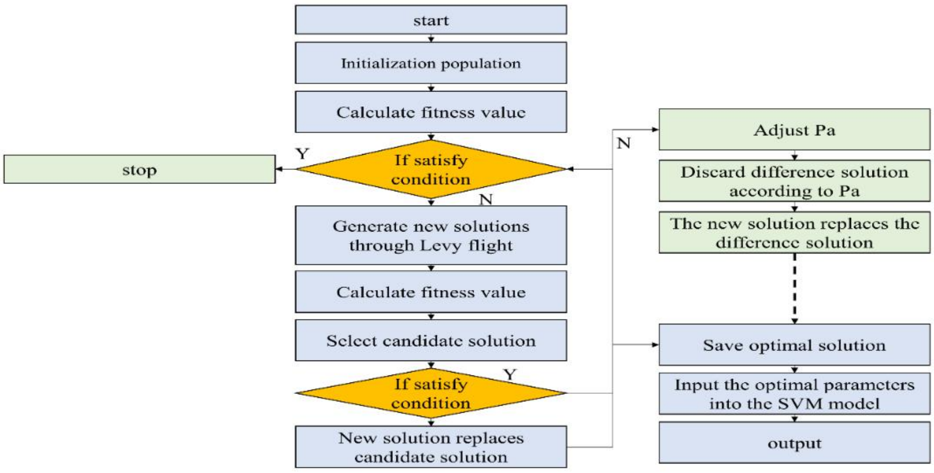

Fig. 4. Flow chart of support vector machine optimized by fuzzy cuckoo search algorithm.

4 Results and Discussion

4.1 Chaotic Particle Swarm Optimization Algorithm

The optimized BP neural network prediction model achieves the established accuracy requirements, with a mean square error (MSE) for the prediction results of 0.000919278. The fuzzy cuckoo search algorithm was employed to optimize the two core parameters of the support vector machine (SVM) prediction model, yielding a kernel parameter of 0.01 and a penalty parameter of 52.6481. The prediction outcomes for both the chaotic particle swarm optimization-based BP neural network (CPSO-BP) and the fuzzy cuckoo search-optimized support vector machine (FCS-SVM) are illustrated in Fig. 5. To evaluate the performance of the models, we employed several metrics: average relative error (ARE), average absolute error (AAE), root mean square error (RMSE), and the coefficient of determination (R^2). The detailed results are presented in Table 3.

According to the data in Table 3, both the ARE and AAE metrics for the optimized BP neural network model and the support vector machine model show improvements compared to their pre-optimization values. Although the average relative error for the BP neural network model remains higher than that of the SVM model both before and after optimization, the reduction achieved in the BP neural network model is more significant than that observed in the SVM model, suggesting that there is still potential for further optimization of the BP neural network. The root mean square error of the optimized BP neural network is markedly reduced, indicating that the predicted values generated by this model are closer to the actual values, thereby enhancing the model's prediction accuracy. Similarly, the root mean square error for the support vector machine also shows a significant reduction, reflecting an improvement in the prediction accuracy of the optimized model. Overall, the optimized prediction models demonstrate clear advantages over traditional prediction models.

Following the comparison of prediction accuracies, the FCS-SVM prediction model will be employed to forecast and analyze carbon dioxide emissions in China's construction industry for the period from 2021 to 2025.

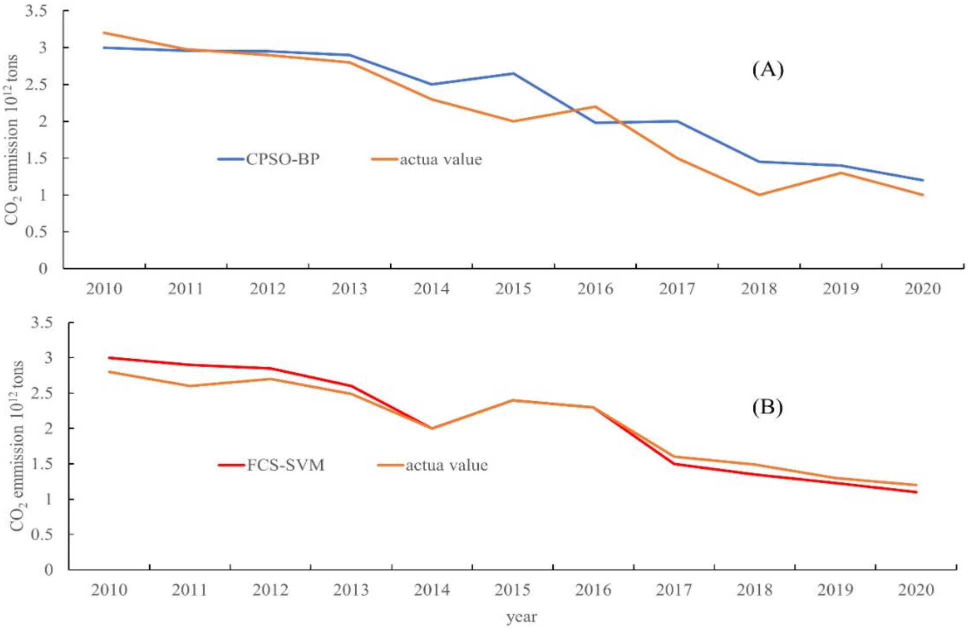

Fig. 5. Prediction results of BP neural network based on chaotic particle swarm optimization (CPSO-BP, (A)); fuzzy cuckoo search algorithm optimizes support vector machine prediction results (FCS-SVM, (B)).

Table 3. performance of BPNN, SVM, CPSO-BPNN, and FCS-SVM model.

error	BPNN	CPSO-BPNN	SVM	FCS-SVM
ARE	14.1%	2.93%	5.1%	1.45%
AAE	28806.1	8221.3	8790.1	5563.1
RMSE	3654.5	8780.1	9440.1	6541.2
R^2	0.72	0.981	0.968	0.994

4.2 Chaotic Particle Swarm Optimization Algorithm

Given the superior performance of the FCS-SVM model, we utilize this approach to forecast carbon dioxide emissions from China's construction industry for the period of 2021 to 2025. Initially, it is essential to predict the values of various influencing factors for the years 2019 to 2021, which will be accomplished using the GM (1,1) grey prediction method for fitting and forecasting each indicator.

Figure 6 illustrates the fitting curve for the predicted values of the influencing factors as generated by the GM (1,1) model, demonstrating satisfactory fitting performance. The results indicate that the average relative error for primary energy consumption in the construction industry is 0.034%, while the average relative errors for the number of employees in construction enterprises, GDP, and the total output value of the construction industry are 0.041%, 0.075%, and 0.077%, respectively—all of which are below the first-level critical threshold of 0.01. A posterior error ratio of less than 0.35 is required to achieve a first-level prediction effect. The posterior error ratios for the four influencing factors are as follows: 0.2229, 0.2147, 0.1694, and 0.1736, all of which satisfy the established criteria. Further testing of the remaining two influencing factors also yielded results that complied with the requirements. Consequently, we conclude that the prediction models for all influencing factors meet the accuracy standards, thereby demonstrating high reliability.

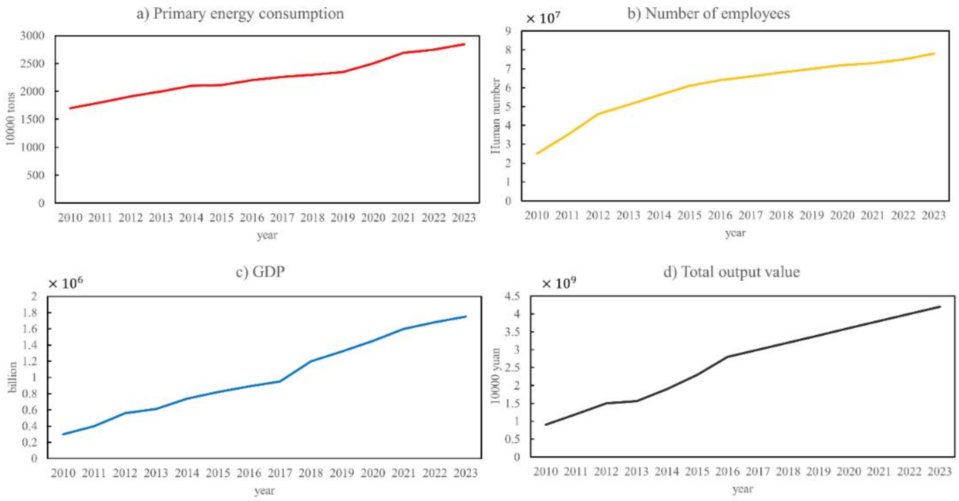

Fig. 6. predictions of GM (1,1) model for four influential factors a) primary energy consumption, b) number of employees, c) GDP and d) total output value.

Through GM (1,1) model, the predicted values of various indicators from 2021-2025 can be obtained as shown in Table 4.

Table 4. Predicted value of influencing factors of carbon dioxide emission in construction industry.

year	Labor productivity (yuan/person)	GDP(billion yuan × $\mathbf{10^6}$)	Total value (10000 yuan× $\mathbf{10^9}$)	Employees (human number × $\mathbf{10^7}$)	House area (10000 m^3× $\mathbf{10^5}$)	Primary energy consumption/10000 tons of standard coal
2021	68452	1.11	2.7	6.2	4.43	2456
2022	69255	1.24	2.9	6.5	4.56	2512
2023	701265	1.29	3.2	6.8	4.61	2569
2024	712355	1.38	3.4	7.2	4.72	2645
2025	731256	1.51	3.7	7.5	4.78	2715

By substituting the predicted values of the indicators, presented in Table 4, as input variables into the FCS-SVM-based carbon dioxide emission prediction model for the construction industry, the estimated carbon dioxide emissions for China's construction sector from 2019 to 2023 are projected to be 181.205 million tons, 1834.56 million tons, 181.024 million tons, 1901.625 million tons, and 2132.163 million tons, respectively.

As illustrated in Fig. 7, carbon dioxide emissions from the construction industry between 2017 and 2020 exhibit relative stability with a gradual upward trend. In 2021, a slight decline in emissions is observed, followed by a period of slow growth; however, this growth rate is lower than that observed prior to 2015, indicating a flatter trend overall. The emission reduction performance of China's construction industry is analyzed from two perspectives: total carbon dioxide emissions and carbon dioxide emission intensity. The carbon dioxide emission intensity for China's construction sector is determined based on the predicted and actual carbon dioxide emissions, as well as the total output value of the construction industry. The resulting data are presented in Fig. 8.

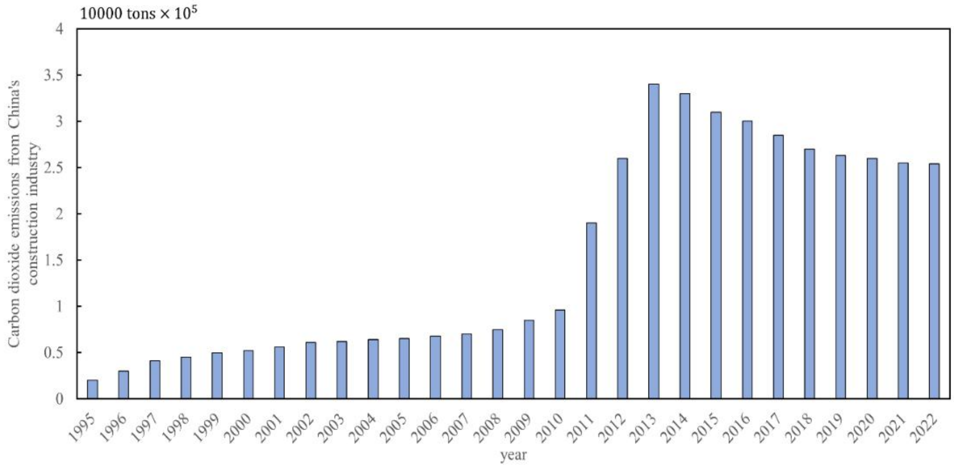

Fig. 7. Carbon dioxide emissions from China's construction industry during 1995-2022.

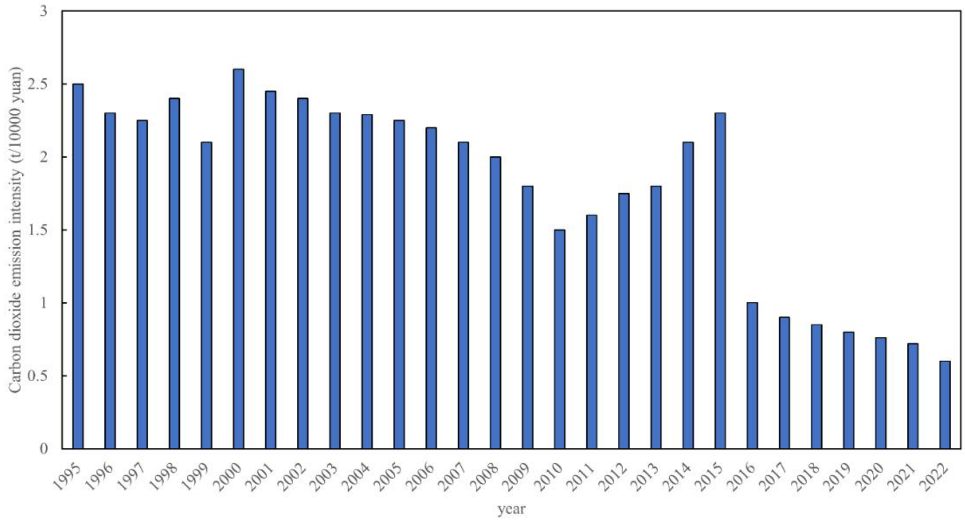

Fig. 8. Carbon dioxide emission intensity of China's construction industry during 1995-2022.

The carbon dioxide emission intensity of China's construction industry has exhibited a gradual decline since 2013, achieving relative stability since 2018. By 2020, the sector successfully reached the objective of reducing carbon emission intensity by 40% to 50% compared to the levels recorded in 2005. However, since 2016, despite a continued decrease in carbon dioxide emission intensity within the construction industry, the rate of decline has been minimal, and there was a rebound in 2023. This trend underscores the persistent challenges associated with energy conservation and emission reduction that lie ahead in the coming years. Given that the construction industry is integral to China's green development initiatives, there is a pressing need to actively revise development strategies, facilitate low-carbon transitions across aspects such as energy structure and building materials, and thereby contribute significantly to national efforts in energy conservation and emission reduction.

4.3 Chaotic Particle Swarm Optimization Algorithm

In this study, the CPSO optimization algorithm is employed to enhance the performance of the BP neural network. This method concurrently optimizes the weight and threshold parameters, thereby mitigating the tendency of the model to converge on local extrema, which in turn improves the prediction accuracy. Furthermore, the incorporation of the fuzzy cuckoo search algorithm is introduced to optimize the support vector machine model. Simulation results demonstrate that both optimization algorithms contribute to an enhancement in the prediction accuracy of the models. Notably, compared to previous research findings, the algorithm proposed herein exhibits a greater degree of robustness. Please refer to Table 5 for further details.

Table 5. Comparison of model performance in prediction of carbon emission with others.

Author	Model	Objective	Relative error (%)
Wang et al. (2022) [37]	SSA-LSTM Algorithm	carbon emission of Thermal Power Plant	3.5
Han et al. (2022) [38]	improved residual neural network	Novel economy and carbon emissions	3.75
Feng et al. (2022) [39]	optimized extreme learning machine	A daily carbon emission	3.88
Javier et al. (2022) [40]	Machine learning	Black carbon emission	5.64
Chu et al. (2022) [41]	PSO-SVR method	Building carbon emission	5.86%
This study	FCS-SVM	Building carbon emission	2.89

Wang et al. (2022) employed the SSA-LSTM algorithm to develop a regression prediction model for carbon emissions from coal-fired power plants. Their findings indicate that the incorporation of boiler feed water influencing factors into the SSA-LSTM model leads to a significant improvement in the regression prediction accuracy of carbon emissions from these facilities [37]. Han et al. (2022) introduced an innovative prediction model for economic factors and carbon emissions based on a residual neural network (RESNET), aimed at optimizing and analyzing the energy structures of various countries or regions globally [38]. Feng et al. (2022) proposed a novel daily carbon emission forecasting model, demonstrating that their two-stage feature selection method enhances prediction accuracy. Following this feature selection, the metrics R^2, MAPE, and RMSE improved by 0.55%, 30.23%, and 28.46%, respectively [39]. Javier et al. (2022) presented a methodological approach for predicting black carbon emissions during industrial furnace operations using machine learning predictive models. The evaluation results confirmed that these models can effectively forecast undesirable black carbon emissions in advance [40]. In the study conducted by Chu et al. (2022) [41], eleven influential factors affecting building carbon emissions were identified, and a support vector regression prediction model was developed to enhance prediction accuracy, generalization, and robustness. The results indicated that the particle swarm optimization model outperformed four other predictive models (e.g., linear regression and decision tree) in terms of accuracy (with an average improvement of 1.01% R^2 on the training subset), generalization (an average improvement of 19.89% R^2 on the testing subset), and robustness (an average improvement of 18.93% R^2 under varying levels of noise intensity).

5 Results and Discussion

The construction industry significantly contributes to global carbon emissions and regional environmental degradation. Accurate predictions of carbon dioxide emissions are crucial for informing policy adjustments and fostering technological innovations in energy conservation and emission reduction. This study establishes two machine learning models: the BP neural network and the support vector machine. Each model is subsequently optimized using the chaos particle swarm optimization algorithm and the fuzzy cuckoo search algorithm, respectively. The performance of these models is evaluated through simulation and comparative analysis against traditional prediction models, with carbon dioxide emissions from China's construction industry predicted for the period 2023 to 2025 based on the FCS-SVM model. The main conclusions of this paper are as follows:

The nonlinear relationship between carbon dioxide emissions and their influencing factors poses limitations for traditional prediction methods. Utilizing BP neural networks and support vector machines, this study forecasts carbon dioxide emissions in the construction industry. A comparison of average absolute error, average relative error, mean square error, and R^2 among the BP neural network, support vector machine, and traditional ARIMA prediction model indicates that the overall error in traditional models exceeds that of both the BP neural network and support vector machine. This highlights the inadequacy of traditional models in predicting nonlinear relationships with multiple influencing factors.

The support vector machine's predictions more closely align with actual values, demonstrating superior accuracy compared to the BP neural network in small sample data predictions.

The CPSO algorithm is utilized to optimize the BP neural network, while the FCS algorithm optimizes the support vector machine. A comparison of average relative error, average absolute error, mean square error, and R^2 reveals that CPSO significantly enhances the accuracy of the BP neural network, leading to a 9.8% reduction in average relative error compared to pre-optimization levels. Despite this improvement, the prediction error of the optimized support vector machine remains lower than that of the BP neural network. Consequently, the FCS-SVM model is employed to predict and analyze carbon dioxide emissions from the construction industry for the period 2022 to 2025. The results indicate a rise in total carbon emissions over the next four years, albeit at a noticeably reduced growth rate. Additionally, the carbon emission intensity is projected to decline prior to 2023, although the rate of decline is minimal, indicating ongoing challenges in future emission reduction efforts.

The efficacy of the BP neural network model improves with the quantity of data available for training. This paper analyzes data collected from 1995 to 2020. Increasing the volume of training data is expected to enhance model robustness, thus facilitating more comprehensive model training. Future research could focus on predicting and examining carbon dioxide emissions from the construction industry across different pro.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Jun Li, Michel Colombier, Managing carbon emissions in China through building energy efficiency, Journal of Environmental Management,2009,90(8):2436-2447.
2. Raymond J Cole. Energy and greenhouse gas emissions associated with the construction of alternative structural systems[J].Building and Environment,2013(34):335-348.
3. Leif Gustavsson, Anna Joelsson, Roger Sathre. Life cycle primary energy use and carbon emission of an eight-storey wood-framed apartment building[J]. Energy and Buildings,2016(42):230-242.
4. Hui Yan, Linda C.H. Fan, Qinping Chen. Greenhouse gas emissions in building construction: A case study of One Peking in Hong Kong[J]. Building and Environment,2016(45)949-955.
5. Christodoulakis N M, Kalyvitis S C, Lalas D P, et al. Forecasting energy consumption and energy related Carbon dioxide emissions in Greece: an evaluation of the consequences of the Community Support Framework II and natural gas penetration. Energy Economics, 2013,22(4):395-422.
6. Sun Yuhao,Hao Shengyue,Long Xiaofeng. A study on the measurement and influencing factors of carbon emissions in China's construction sector[J]. Building and Environment,2023,229.
7. Guo Xiaomin,Fang Chuanglin. Spatio-temporal interaction heterogeneity and driving factors of carbon emissions from the construction industry in China.[J]. Environmental science and pollution research international,2022.
8. Chung Y, Paik C, Kim H, et al. Bottom-up Analysis of GHG Emissions from Shipbuilding Processes for Low-carbon Ship Production in Korea[J]. Journal of Ship Production & Design,2017, 33(3): 221-226.
9. Wakiyama T, Kuramochi T. Scenario analysis of energy saving and Carbon dioxide emissions reduction potentials to ratchet up Japanese mitigation target in 2030 in the residential sector[J]. Energy Policy, 2017, 103: 1-15.
10. Mirzaei M, Bekri M. Energy consumption and Carbon dioxide emissions in Iran,2025[J]. Environmental Research, 2017, 154: 345-351.
11. Tudor C. Predicting the Evolution of Carbon dioxide Emissions in Bahrain with Automated Forecasting Methods[J]. Sustainability, 2016, 8(9): 923.
12. Hamzacebi C, Es H A. Forecasting the annual electricity consumption of Turkey using an optimized grey model[J]. Energy, 2014, 70: 165-171.
13. Boran F E. Forecasting Natural Gas Consumption in Turkey Using Grey Prediction[J]. Energy Sources, Part B: Economics, Planning, and Policy, 2015, 10(2): 208-213.
14. Zhao Qinfeng,Gao Weijun,Su Yuan,Wang Tian. Carbon emissions trajectory and driving force from the construction industry with a city-scale: A case study of Hangzhou, China[J]. Sustainable Cities and Society,2023,88.
15. Zhiyuan Guo,Jie Yin. Carbon Emission Forecast of Construction Industry Based on Grey Theory[J]. Academic Journal of Computing & Information Science,2022,5.0(10.0).
16. Zhou Yingbin,Lv Siqi,Wang Jianlin,Tong Junbo,Fang Zhong. The Impact of Green Taxes on the Carbon Emission Efficiency of China's Construction Industry[J]. Sustainability,2022,14(9).

17. Erdogan Seyfettin. Dynamic nexus between technological innovation and buildings Sector's carbon emission in BRICS countries.[J]. Journal of environmental management,2021,293.
18. Wang Zhenshuang,Zhou Yanxin,Zhao Ning,Wang Tao,Zhang Zhongsheng. Spatial Correlation Network and Driving Effect of Carbon Emission Intensity in China's Construction Industry[J]. Buildings,2022,12(2).
19. Zhang Hui,Du Liu Jie,Wang Bai Tian,Qi Ping. Carbon Emission Efficiency Measurement of Construction Industry and Its Treatment Measures-A Case Study of Henan Province, China[J]. Nature Environment and Pollution Technology,2021,20(2).
20. Zhang Mengna,Li Lianshui,Cheng Zhonghua. Research on carbon emission efficiency in the Chinese construction industry based on a three-stage DEA-Tobit model.[J]. Environmental science and pollution research international,2021,28(37).
21. Li Dezhi,Huang Guanying,Zhu Shiyao,Chen Long,Wang Jiangbo. How to peak carbon emissions of provincial construction industry? Scenario analysis of Jiangsu Province[J]. Renewable and Sustainable Energy Reviews,2021,144.
22. Dezhi Li,Guanying Huang,Guomin Zhang,Jiangbo Wang. Driving factors of total carbon emissions from the construction industry in Jiangsu Province, China[J]. Journal of Cleaner Production,2020,276.
23. Qing jing Shi,Chun Bai. Carbon Emission Efficiency in the Construction Industry and Its Carbon Emission Control Measures: A Case Study of Henan Province, China[J]. Nature Environment & Polution Technology,2020,19(2).
24. Na Lu,Shuyi Feng,Ziming Liu,Weidong Wang,Hualiang Lu,Miao Wang. The Determinants of Carbon Emissions in the Chinese Construction Industry: A Spatial Analysis[J]. Sustainability,2020,12(4).
25. Wei Lu,Vivian W.Y. Tam,Heng Chen,Lei Du. A holistic review of research on carbon emissions of green building construction industry[J]. Engineering, Construction and Architectural Management,2020,27(5).
26. Liu Hua,Zhu Min. Carbon Emission Efficiency of Construction Industry in Hunan Province and Measures of Carbon Emission Reduction[J]. Nature Environment & Polution Technology,2019,18(3).
27. Peng Wu,Yongze Song,Jianbo Zhu,Ruidong Chang. Analyzing the influence factors of the carbon emissions from China's building and construction industry from 2000 to 2015[J]. Journal of Cleaner Production,2019,221.
28. Qiang Du,Jie Zhou,Ting Pan,Qiang Sun,Min Wu. Relationship of carbon emissions and economic growth in China's construction industry[J]. Journal of Cleaner Production,2019,220.
29. Lai Xiaodong,Lu Chao,Liu Jixian. A synthesized factor analysis on energy consumption, economy growth, and carbon emission of construction industry in China.[J]. Environmental science and pollution research international,2019,26(14).
30. Xianchun Tan,Haiping Lai,Baihe Gu,Yuan Zeng,Hui Li. Carbon emission and abatement potential outlook in China's building sector through 2050[J]. Energy Policy,2018,118.
31. Du Qiang,Lu Xinran,Li Yi,Wu Min,Bai Libiao,Yu Ming. Carbon Emissions in China's Construction Industry: Calculations, Factors and Regions.[J]. International journal of environmental research and public health,2018,15(6).
32. Roh Seungjun,Tae Sungho,Kim Rakhyun. Developing a Green Building Index (GBI) Certification System to Effectively Reduce Carbon Emissions in South Korea's Building Industry[J]. Sustainability,2018,10(6).
33. FENG Y Y, CHEN S Q, ZHANG L X. System dynamics modeling for urban energy consumption and CARBON emissions: A case study of Beijing, China[J]. Ecological Modelling, 2013, 252: 44-52.
34. Abdel Karim Baareh. Evolutionary Design of a Carbon Dioxide Emission Prediction Model using Genetic Programming[J]. International Journal of Advanced Computer Science and Applications (IJACSA),2018,9.
35. Pruethsan Sutthichaimethee,Apinyar Chatchorfa,Surapol Suyaprom. A Forecasting Model for Economic Growth and CARBON Emission Based on Industry 4.0 Political Policy under the Government Power: Adapting a Second-Order Autoregressive-SEM[J]. Journal of Open Innovation: Technology, Market, and Complexity,2019,5(3).
36. Wang Xindong,Yan Chun,Liu Wei,Liu Xinhong. Research on Carbon Emissions Prediction Model of Thermal Power Plant Based on SSA-LSTM Algorithm with Boiler Feed Water Influencing Factors[J]. Sustainability,2022,14(23).
37. Han Yongming,Cao Lian,Geng Zhiqiang,Ping Weiying,Zuo Xiaoyu,Fan Jinzhen,Wan Jing,Lu Gang. Novel economy and carbon emissions prediction model of different countries or regions in the world for energy optimization using improved residual neural network.[J]. The Science of the total environment,2022.
38. Kong Feng,Song Jianbo,Yang Zhongzhi. A daily carbon emission prediction model combining two-stage feature selection and optimized extreme learning machine[J]. Environmental Science and Pollution Research,2022,29(58).
39. RubioLoyola Javier,PaulFils Wolph Ronald Shwagger. Applied Machine Learning in Industry 4.0: Case-Study Research in Predictive Models for Black Carbon Emissions[J]. Sensors,2022,22(10).
40. Chu Xiaolin,Zhao Ruijuan. A building carbon emission prediction model by PSO-SVR method under multi-criteria evaluation[J]. Journal of Intelligent & Fuzzy Systems,2021,41(6).