Pol. J. Environ. Stud. Vol. XX, No. X (XXXX), 1-18
DOI: *10.15244/pjoes/219359* ONLINE PUBLICATION DATE: 2026-07-14

Original Research

Carbon Emission Interval Prediction Based on the Difference Creative Search-Convolutional Neural Network-Bidirectional Long Short-Term Memory-Attention-Kernel Density Estimation Model

Haodong Chang, Yipeng Zhao*, Fujun He

College of Management Science, Chengdu University of Technology, Chengdu, 610059, China

Received: 22 August 2025
Accepted: 15 March 2026

Abstract

This study first calculates China's CO_2 emissions from 2000 to 2022 using an emission source classification approach. Second, it constructs an indicator system for CO_2 emission influencing factors, covering five dimensions: economy, society, environment, energy, and technology. Lasso regression is then applied to identify key drivers of carbon emission changes. Based on this, the research represents pioneering work to apply the difference creative search (DCS) algorithm and kernel density estimation (KDE) to the CNN-BiLSTM-Attention model, and thereby construct the DCS-CNN-BiLSTM-Attention-KDE interval prediction model. The prediction results are then compared with those of other point and interval prediction models. Finally, carbon emissions in China from 2023 to 2035 are predicted using the scenario analysis method. The research results indicate that: (1) MAE, RMSE, MAPE, PICP, and PINAW of the DCS-CNN-BiLSTM-Attention-KDE are 10063.0250, 13100.8491, 0.0092, 0.9021, and 0.2817, respectively. In comparison to other models, the proposed model exhibits higher prediction accuracy. (2) The carbon peaking times for the baseline scenario, the green and low-CO_2 scenario, and the strong low-CO_2 scenario are projected to be 2033, 2030, and 2027, respectively. Under the green and low-CO_2 scenarios, as well as the strong low-carbon scenario, China is projected to achieve its carbon peaking target on schedule.

Keywords: carbon emissions, carbon peak, interval prediction, scenario analysis, Lasso regression, kernel density estimation, DCS-CNN-BiLSTM-Attention-KDE model

Introduction

Global climate warming has triggered interconnected crises such as reduced food production and rising sea levels, emerging as humanity's most pressing non-traditional security challenge [1]. To address this, in 2020, China committed to achieving the carbon peak by 2030 and carbon neutrality by 2060 [2], implementing policies like carbon trading markets and carbon tax pilots to advance these "dual carbon" goals [3]. However, China's 30-year transition from peak to neutrality is

*e-mail: 576397663@qq.com

© 2026 This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

uniquely time-constrained and arduous [4], making the scientific identification of CO_2 emission drivers and accurate future forecasting critical for success.

Current research faces three key limitations: outdated CO_2 emission data, inadequate accuracy of traditional point prediction models, and insufficient quantification of prediction uncertainty. To address these gaps, this study first calculates China's 2000-2022 CO_2 emissions via emission source classification, incorporating direct emissions and a life-cycle perspective to enhance data reliability [5], and uses Lasso regression to screen key variables from a five-dimensional indicator system of influencing factors.

Meanwhile, this paper puts forward a Convolutional Neural Network-Bidirectional Long Short-Term Memory-Attention-Kernel Density Estimation interval prediction model that is optimized utilizing the Difference Creativity Search Algorithm (DCS-CNN-BiLSTM-Attention-KDE). This model is not only capable of outputting the prediction interval and quantifying uncertainty, but can also effectively mitigate data noise and complexity through the DCS algorithm and improve its capacity to capture the nonlinear characteristics of CO_2 emissions. In order to evaluate the performance of the model, we conducted a comparison between it and a variety of point prediction models as well as interval prediction models. Finally, by employing the scenario analysis method, we simulate the trajectory of China's CO_2 emissions from 2023 to 2035.

This research addresses three core questions: (1) How to construct a robust CO_2 emission prediction model with limited annual data? (2) Does the DCS algorithm outperform traditional optimizers in this task? (3) How can interval prediction enhance climate policy decision-making?

Despite the small sample size posing deep learning challenges, the hybrid model architecture, attention mechanism, KDE, and DCS optimization mitigate overfitting and boost generalization. At the policy level, the model's interval outputs reflect prediction uncertainty, with interval width indicating policy risks and coverage probability aiding scenario-specific compliance assessments, supporting resilient climate planning.

This study's contributions are threefold: (1) Expanding interval prediction applications in CO_2 emissions research for small-sample scenarios. (2) Introducing the DCS algorithm to optimize the hybrid model, enhancing feature extraction and accuracy. (3) Leveraging prediction intervals and scenario analysis to provide quantitative decision support for global carbon neutrality efforts.

Literature Review

CO_2 emission prediction models are mainly divided into single-model and combined-model approaches. Single models (e.g., STIRPAT, gray model, BP neural network, LSTM) [6-9] have clear structures but are constrained by model assumptions, limiting their ability to capture data complexity and trends, leading to inaccurate predictions. Scholars have thus proposed combined models (e.g., PSO-ELM, ARIMA-LSTM, TCN-LSTM) [10, 11] that integrate algorithmic strengths to improve point prediction accuracy. However, current optimization algorithms still need improvement in handling data complexity and interference. Research has confirmed that the DCS algorithm effectively reduces data complexity and noise, enabling better capture of non-linear CO_2 emission characteristics.

Both single and combined models are essentially point prediction methods, which only provide fixed future values and fail to quantify prediction uncertainty. This limits their application value in risk assessment and robust decision-making. In contrast, interval prediction outputs potential value ranges, quantifying uncertainty and providing more comprehensive, reliable information for decision-makers [12]. While widely used in photovoltaic power, wind power, and wind speed prediction [12-14], interval prediction has rarely been applied to CO_2 emissions. Most existing studies focus on improving point prediction accuracy, neglecting uncertainty characterization. Integrating interval prediction into CO_2 emission research to forecast future trends and fluctuation ranges is, therefore, more practically significant.

In carbon peak prediction, point prediction models are often combined with scenario analysis [15-17] to examine peak scenarios under different conditions. Selecting key CO_2 emission drivers is critical, as prediction accuracy depends on these factors. Current studies typically select drivers based on five dimensions (economic, social, environmental, energy, and technological) [1, 10, 17] through qualitative methods relying on experience or theory. This subjective selection may omit key variables or include redundant ones, affecting scenario rationality and prediction reliability.

To scientifically select CO_2 emission drivers and address limitations in prediction accuracy and uncertainty characterization, this study builds on existing research by establishing an indicator system and using Lasso regression to screen key factors. It optimizes the CNN-BiLSTM-Attention model's learning rate, hidden layer node count, and regularization parameter via the DCS algorithm, and integrates an adaptive kernel density estimation (KDE) to convert point predictions into interval predictions, developing the DCS-CNN-BiLSTM-Attention-KDE model. This extends the model's applicability and enriches CO_2 emission prediction options, providing upper and lower bound intervals to support stable short-term decision-making.

Materials and Methods

Measurement of Carbon Emissions

Referring to Cong et al. [5], this study classifies CO_2 emission sources into three scopes to calculate China's 2000-2022 emissions:

Scope 1: Direct emissions within urban areas, including energy activities (transportation, industry, buildings), agriculture, industrial processes, land use change, forestry, and waste management.

Scope 2: Indirect energy emissions outside urban boundaries, such as cooling, heating, and electricity purchased to meet urban energy needs.

Scope 3: Other indirect emissions from urban activities outside jurisdiction (not covered by Scope 2), including emissions from production, transportation, use, and disposal of externally procured goods [18].

The CO_2 emission measurement methods for scopes 1-3 are shown in Equations (1)-(3).

$$E_{scope1} = \sum_{i=1}^{n} A_i \times EF_i \tag{1}$$

$$E_{scope2} = \sum_{j=1}^{m} E_j^{import} \times EF_j^{grid} \tag{2}$$

$$E_{scope3} = \sum_{k=1}^{p} \sum_{l=1}^{q} A_{kl} \times EF_{kl} \tag{3}$$

Where, E_{scope1}, E_{scope2}, and E_{scope3} respectively represent the total emissions of Scope 1 to 3. A_i represents the activity level data of class i emission sources. EF_i represents the emission factor of the i-th type of emission source, and n indicates the number of types of emission sources. E_i^{import} represents the electricity or heat purchased from the j-th external power grid or thermal system. E_j^{grid} is the average emission factor of the j-th external power grid or thermal system. Emission factors can use the average emission factor of the regional power grid or the average emission factor of the national power grid. For electricity with traceable sources, the emission factor of a specific power plant should be used. For electricity with untraceable sources, the comprehensive emission factor of the power grid should be used. m represents the quantity of the purchased source. A_{kl} represents the data of the activities of the k-th type of product in the l stage of its life cycle. E_{kl} represents the emission factor corresponding to the stage. p represents the number of product categories. q represents the order of the life cycle.

The emission scope classification method, first proposed by the World Resources Institute in corporate GHG inventory guidelines [19-21], aligns with direct emission measurement [22], urban carbon footprint assessment [5, 23], and life-cycle perspectives in Scope 3 [24]. This ensures more accurate CO_2 quantification and reliable trend analysis.

DCS-CNN-BiLSTM-Attention-KDE Interval Prediction Model

Differentiated Creative Search Algorithm

The DCS algorithm is a novel intelligent optimization method proposed by Duankhan in 2024 [25]. It uses a team-performance-based iterative model, categorizing team members into high-achieving, medium-performing, and low-performing groups. Its implementation steps are as follows [25]:

(1) Population initialization. At the initial stage, the DCS algorithm selects a set of initial solutions, denoted as $X_{i,m}$, and iteratively optimizes them. The methodology for selecting these initial solutions, $X_{i,m}$, is outlined in Equation (4).

$$X_{i,m} = L_m + U(0,1) \times (U_m - L_m) \tag{4}$$

In this context, $U(0,1)$ denotes the uniform distribution defined over the interval (0,1). The symbols L and U indicate the lower and upper limits, respectively. According to Equation (1), the initial solution is randomly generated between U_m and L_m of the data.

(2) Differentiated knowledge acquisition (DKA). DKA primarily investigates the rate at which individuals acquire new knowledge [25], exerting various influences. It assigns a specific role to each initial solution, denoted as $X_{i,m}$. The impact of DKA on each initial solution $X_{i,m}$ is quantified using Equations (5) and (6).

$$j_{rand} = randint(1, D) \tag{5}$$

$$v_{i,m} = \begin{cases} v_{i,m}, if\ U(0,1) \le \eta_{i,t}\ or\ m = j_{rand} \\ X_{i,m}, \quad otherwise \end{cases} \tag{6}$$

Where, $v_{i,m}$ is the $v_{i,s}$ in the m-th dimension of the initial solution. $v_{i,s}$ is the initial solution v_i of the s-th iteration. j_{rand} is a randomly chosen integer between 1 and D, which is generated once for each sample X. $\eta_{i,s}$ is the quantitative knowledge acquisition rate of the sample at the s-th iteration, and its calculation formula is Equation (7).

$$\eta_{i,s} = \frac{[U(0,1) \times \lambda_{i,s}] + (U(0,1) \le \lambda_{i,s})}{2} \tag{7}$$

Where, the symbol [•] indicates the meaning of rounding. $\lambda_{i,s}$ represents the coefficient value, which is the value of the variable λ at the s-th iteration of the sample. The formula for the calculation of $\lambda_{i,s}$ is given by Equation (8).

$$\lambda_{i,s} = 0.55 \times \left(\frac{R_{i,s}}{N}\right)^{0.5} + 0.25 \quad (8)$$

Where, N is the total number of samples. $R_{i,s}$ is the dimension of the i-th sample at the beginning of the s-th iteration.

(3) Realistic creation. Convergent thinking facilitates the continuous alteration of element $v_{i,m}$ on the m-th dimension of the vector via Equation (9), thereby ensuring both diversity and convergence within the resultant solution.

$$v_{i,m} = w \times X_{best,m} + \rho_s \times (X_{r2,m} - X_{i,m}) + \omega_{i,s} \times (X_{r1,m} - X_{i,m}) \quad (9)$$

Where, $X_{best,m}$ represents the m-th position of the best-performing sample in the iterative process. w stands for cognitive weight, and the default value is 1. $\omega_{i,s}$ is the coefficient value of sample i-th at the s-th iteration. $X_{r1,d} \neq X_{r2,d} \neq X_{i,d} \neq X_{best,d}$. The nsg is calculated by p. When the population is large, p is 0.2. When the total sample size is small, the value of nsg is not less than 6. ρ_s represents the coefficient value at the s-th iteration. The divergent mind is represented by Equation (10).

$$v_{i,d} = X_{r1,d} + Linnik(\alpha, \sigma) \quad (10)$$

Where, the divergent mind is updated by adding a *Linnik* distributed random number generator with control parameters α and σ. $X_{i,m}$ may move to a new position within the divergence region under the influence of *Linnik*(α, σ).

(4) Retrospective assessment (RA). This study streamlines the RA framework, emphasizing the selection and tracking of optimal performance [25]. The procedural details of RA are outlined in Equation (11).

$$X_{i,t+1} = \begin{cases} v_{i,t}, f(v_{i,t}) \le f(X_{i,t}) \\ X_{i,t}, \qquad otherwise \end{cases} \quad (11)$$

DCS-CNN-BiLSTM-Attention-KDE

The CNN-BiLSTM model is widely used in prediction research [26-29]. CNN extracts local features via convolutional layers, while BiLSTM handles sequential data and captures long-term dependencies, enabling the model to understand both local features and temporal evolution [29]. This study integrates a global average pooling layer and an SE channel attention mechanism into CNN-BiLSTM to form the CNN-BiLSTM-Attention model. This mechanism recalibrates channel feature responses by learning feature importance, enhancing sensitivity to key information, and improving representational capacity [28].

To address the limitation of point prediction (inability to quantify uncertainty) [12], this study optimizes the CNN-BiLSTM-Attention model's learning rate, hidden layer node count, and regularization parameter using the DCS algorithm. It also integrates an adaptive kernel density estimation (AKDE) to convert point predictions into interval predictions. AKDE uses local weighted regression to adjust bandwidth parameters automatically, adapting to local data characteristics and improving density estimation accuracy compared to fixed bandwidth methods. The model implementation steps are shown in Fig. 1:

Input independent variables (influencing factors) and the dependent variable (CO_2 emissions).

Split data into a training set (70%) and a test set (30%).

Normalize the features and labels to enhance training efficiency.

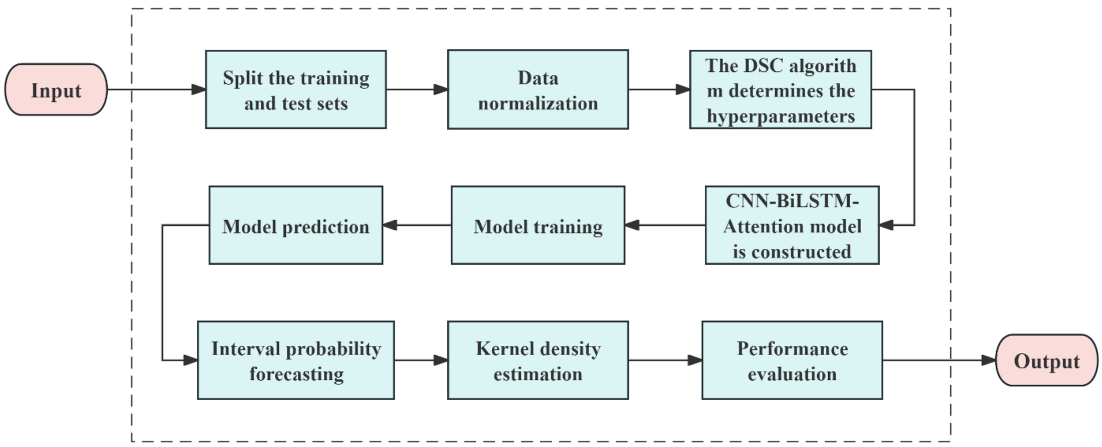


Fig. 1. DCS-CNN-BiLSTM-Attention-KDE model implementation process.

Use the DCS algorithm to optimize hyperparameters.

Construct the CNN-BiLSTM-Attention model (CNN for feature extraction, BiLSTM for sequential dependencies, attention for key information).

Train the model with optimized parameters using the training set.

Predict on the training and test sets, then reverse-normalize the results.

Calculate the training set errors, establish confidence intervals, and perform AKDE.

Identify the sampling points, conduct kernel density estimation, and generate the density plots.

Evaluate performance via visualization (comparison and density plots) and metrics (MAE, RMSE, MAPE, PICP, PINAW).

Model Evaluation Index

To evaluate the accuracy of the CO_2 emission prediction model, we selected MAE, RMSE, and MAPE as evaluation metrics for the point prediction. The formulae for these evaluation indices are presented in Equations (12)-(14).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}| y_i - \hat{y}_i | \tag{12}$$

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2\right)^{\frac{1}{2}} \tag{13}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{| y_i - \hat{y}_i |}{y_i} \tag{14}$$

Where, y_i is the actual value. $\hat{y}$ is the predicted value.

In selecting evaluation metrics for the interval prediction model, we opted for the PI Coverage Probability (PICP) and the PI Normalized Average Width (PINAW). The corresponding calculation formulae are provided in Equations (15) and (16).

$$PICP = \frac{1}{N}\sum_{i=1}^{N}e_i,\ \ e_i = \begin{cases} 1, y_i \in [L_i, U_i] \\ 0, y_i \notin [L_i, U_i] \end{cases} \tag{15}$$

$$PINAW = \frac{1}{N \times R}\sum_{i=1}^{N}(U_i - L_i) \tag{16}$$

Where, L_i and U_i denote the lower and upper limits of the *i*-th performance indicator, respectively, while R indicates the range value of the target.

Data sources

Data on CO_2 emissions, influencing factors, and energy consumption are sourced from the National Bureau of Statistics (NBS, https://www.stats.gov.cn/). Industrial process data comes from the China Industrial Statistical Yearbook, agricultural/forestry/land use data from the China Agricultural Statistical Yearbook, and outsourced electricity/heating/cooling data from the NBS website. The study covers 32 Chinese regions (excluding Hong Kong, Macao, and Taiwan) from 2000 to 2022.

Results

Analysis of The Influencing Factors of Carbon Emission

Construction of the Carbon Emission Index System

Scholars differ in their selection of factors influencing CO_2 emissions. Nevertheless, the primary factors influencing CO_2 emissions generally encompass five dimensions: economic, societal, environmental, energetic, and technological. This study provides an overview of existing research on CO_2 emission predictions. Regarding the economic dimension, four indices were selected: per capita GDP, the share of the primary sector in GDP, the share of the secondary sector in GDP, and the share of the tertiary sector in GDP [4]. Concerning the societal dimension, four indices were selected: total population, urbanization rate, private car ownership, and investment in fixed assets [1]. The forest coverage rate serves as an indicator of environmental factors. With respect to the energetic dimension, four indices were chosen: energy intensity, total energy consumption, energy structure, and the proportion of non-fossil energy [10]. Regarding the technological dimension, three indices were selected: research and development intensity, technical market turnover, and the harmless disposal rate of domestic waste [1]. The index system for influencing factors of CO_2 emissions is presented in Table 1.

Screening of Influencing Factors

In this study, Lasso regression was employed to identify the key factors influencing CO_2 emissions. Before employing Lasso regression, we carried out a multicollinearity diagnosis on the original 16 influencing factors. Through calculating the variance inflation factor (VIF) and tolerance, we observed that multiple variables exhibited strong collinearity (as shown in Table 2). Lasso regression is capable of performing precise variable selection through L1 regularization, efficiently addressing multicollinearity problems, preventing overfitting, and improving the generalization ability of the model. The optimal parameters were determined

Table 1. Index system of CO_2 emission influencing factors.

First level index	Second level index	Unit
Economic factors	Per capita GDP	yuan
	Share of primary sector in GDP	%
	Share of secondary sector in GDP	%
	Share of tertiary sector in GDP	%
Social factors	Total population	Ten thousand people
	Urbanization rate	%
	Private car ownership	Ten thousand vehicles
	Expenditure on durable assets across the entire community	Hundred million yuan
Environmental factors	Forest coverage rate	%
Energy factor	Energy intensity	Ten thousand tons of benchmark coal per 100 million yuan
	Total energy consumption	Ten thousand tons of benchmark coal
	Energy structure	%
	Proportion of non-fossil energy	%
Technical factors	Research and development intensity	%
	Technology market turnover	100 million yuan
	Harmless disposal rate of domestic waste	%

through k-fold cross-validation. The cross-validation residual plot is presented in Fig. 2.

Fig. 2 illustrates that the abscissa represents the Lasso regression parameter λ after logarithmic transformation, while the ordinate depicts the MSE of Lasso regression across various λ values. The top horizontal axis indicates the number of non-zero variables selected for each specific λ value. The λ value associated with the leftmost dashed line denotes the one that minimizes the model's MSE. In this study, the λ value corresponding to the leftmost dotted line was chosen as the optimal parameter for Lasso regression, specifically $\lambda = 0.00043$.

Therefore, seven key factors influencing CO_2 emissions were identified through Lasso regression analysis, namely: per capita GDP, the proportion of the secondary industry in GDP, total social investment in fixed assets, forest coverage rate, energy intensity, energy mix, and research and development (R&D) intensity. Additionally, these seven key factors can be categorized into five major dimensions that impact CO_2 emissions.

Construction and Comparison of Carbon Emission Prediction Model

DCS-CNN-BiLSTM-Attention-KDE Interval Prediction Model

In this study, the model was developed and validated using MATLAB 2023a software. The dataset was partitioned into a training set and a test set in a 7:3 ratio. Network training was configured for a maximum of 1000 iterations, with a learning rate set to 0.01. For the DCS algorithm, the number of groups was set to 5, with a maximum iteration limit of 500. DCS was utilized to optimize three parameters: the learning rate, the number of hidden layer nodes, and the regularization coefficient. Specifically, the learning rate ranged from 0.00001 to 0.01, the number of hidden layer nodes was set between 10 and 100, and the regularization coefficient was constrained to the range of 0.000001 to 0.01. To verify the advantages of the DCS algorithm in the optimization process, this study applies the WOA, PSO, and GWO algorithms, respectively, to optimize the CNN-BiLSTM-Attention model and contrast the evolutionary curves of the optimization results obtained by the DCS algorithm. Fig. 3 presents the evolutionary curves of the respective algorithms.

It is evident from Fig. 3a) that the DCS attains the optimal value after approximately 50 iterations. A meticulous comparison of Fig. 3(b-d) demonstrates that the DCS exhibits the most superior optimization performance in comparison with the other optimization algorithms. The point prediction results for CO_2 emissions, obtained after model training, are presented in Fig. 4.

As depicted in Fig. 4, the point predictions generated by the DCS-CNN-BiLSTM-Attention-KDE model exhibited a high level of agreement with the actual values. The mean prediction error for the test set was

Table 2. The results of the multicollinearity test.

Variable	VIF	Tolerance
GDP per capita	2939.644	0.000
Proportion of primary industry in GDP	561.172	0.002
Proportion of secondary industry in GDP	500.303	0.002
Proportion of tertiary industry in GDP	1504.223	0.001
Total population	7032.463	0.000
Urbanization rate	16598.866	0.000
Private car ownership	2495.430	0.000
Investment in fixed assets of the whole society	1783.135	0.001
Forest coverage rate	29.643	0.034
Energy intensity	1618.146	0.001
Total energy consumption	2484.687	0.000
Energy mix	1043.278	0.001
Proportion of non-fossil energy	1284.696	0.001
Research and development intensity	712.473	0.001
Technology market turnover	366.997	0.003
Harmless disposal rate of domestic waste	978.919	0.001

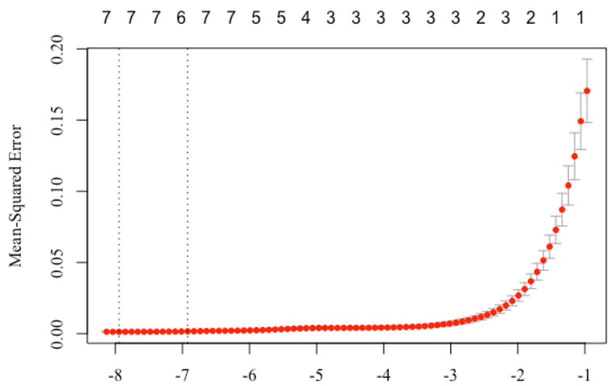


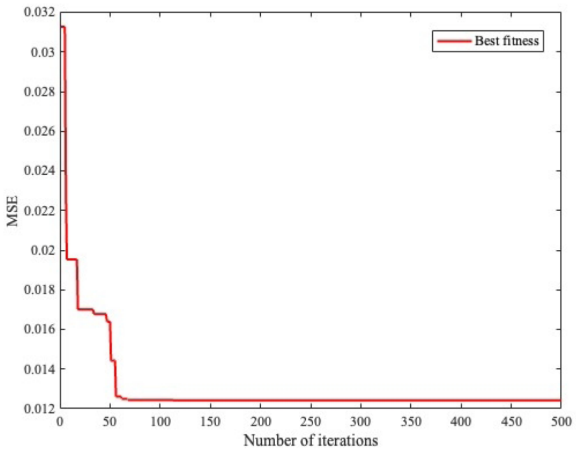
Fig. 2. Residual plot of CV.

0.92%, indicating a satisfactory predictive performance. Additionally, the KDE interval prediction method was integrated to generate the interval prediction results for the CO_2 emission test set. This method rests on the assumption of an independent and homogeneous distribution of training errors and applies to scenarios where the error distribution remains relatively stable. Consequently, in this study, the Lagrange multiplier test was initially employed for residual analysis, yielding an F-value of 2.701 and a p-value of 0.077. The test results indicate that the model residuals exhibit no significant heteroscedasticity or temporal autocorrelation, thereby substantiating the rationality of employing the global error distribution for the construction of KDE intervals.

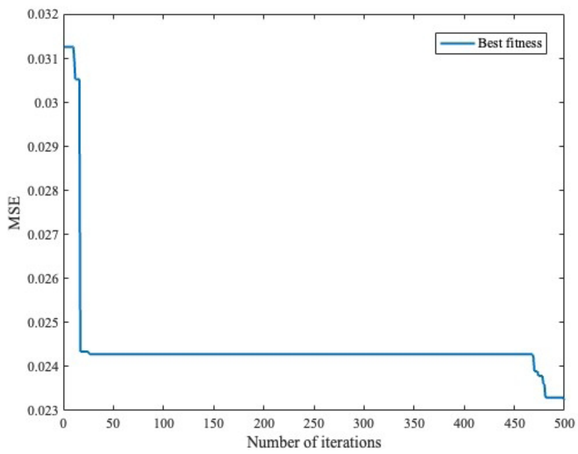
Fig. 5 illustrates the distribution of CO_2 emission prediction intervals for the test set, across confidence levels of 95%, 90%, 75%, 50%, 25%, 10%, and 5%. As illustrated in Fig. 5, an increase in the confidence level resulted in a corresponding widening of the prediction interval, thereby encompassing a greater number of actual CO_2 emission values. This suggests that the model exhibits strong performance in capturing the variability of CO_2 emission data. The evaluation results for forecasts across various confidence intervals are presented in Table 3.

Table 3 indicates that the values of the PICP and the PINAW exhibit an ascending trend as the confidence level rises. Furthermore, the PICP values closely approximate the respective confidence levels, suggesting a satisfactory interval prediction performance of the model. The model thus provides comprehensive prediction information and fulfills practical requirements.

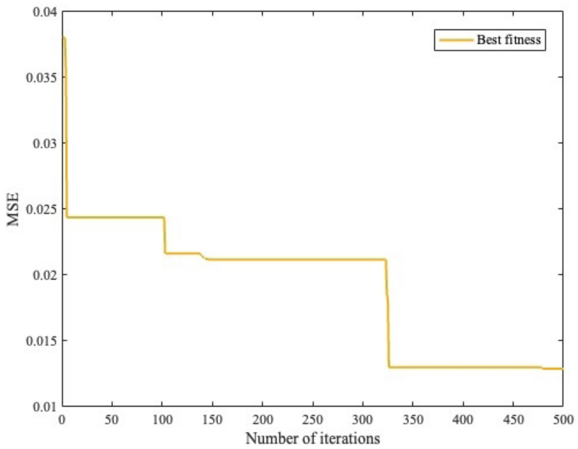
Fig. 6 illustrates the probability density functions corresponding to the first, third, fourth, and sixth prediction points generated by the DCS-CNN-BiLSTM-Attention-KDE model. The vertical line depicted in Fig. 6 signifies the actual observed value at the respective prediction time point. The peak of the probability density function on the vertical axis indicates the modal value, which corresponds to the most likely outcome of the deterministic forecast.



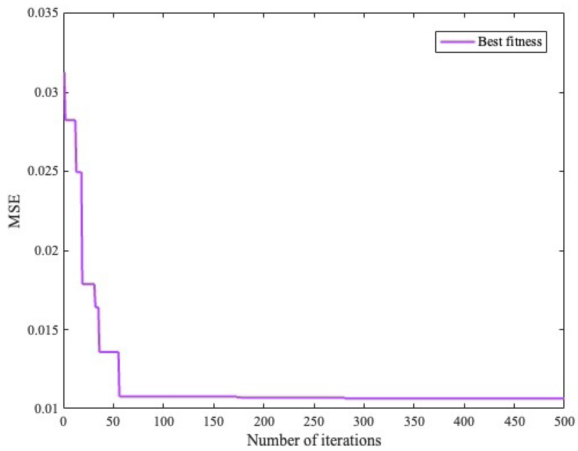
(a) DCS algorithm



(b) WOA algorithm



(c) PSO algorithm



(d) GWO algorithm

Fig. 3. Comparison of evolutionary curves of different algorithms.

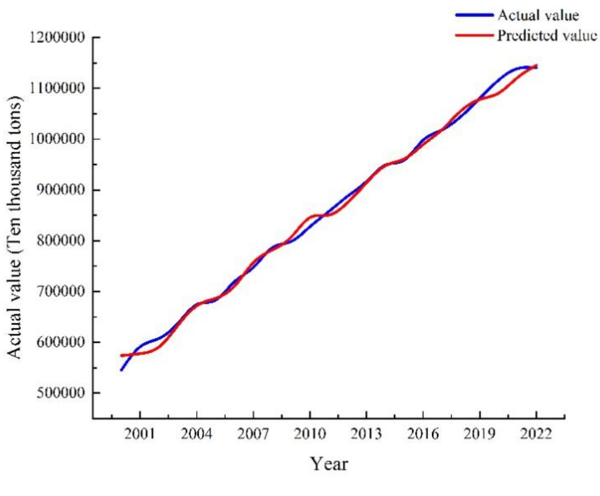


Fig. 4. Results of point projections of CO_2 emissions from 2000 to 2022.

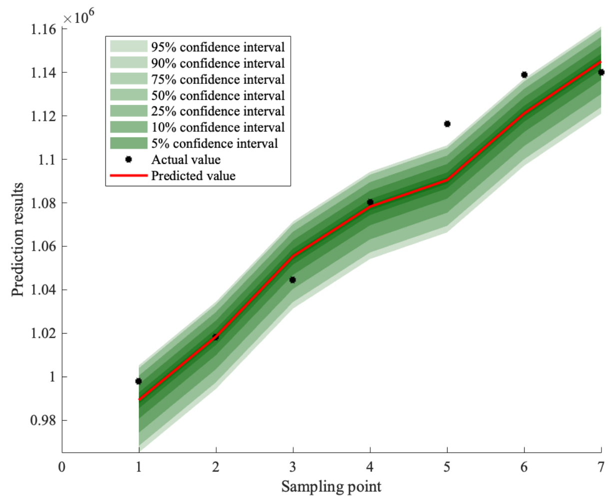


Fig. 5. Results of interval prediction of CO_2 emissions from 2016 to 2022.

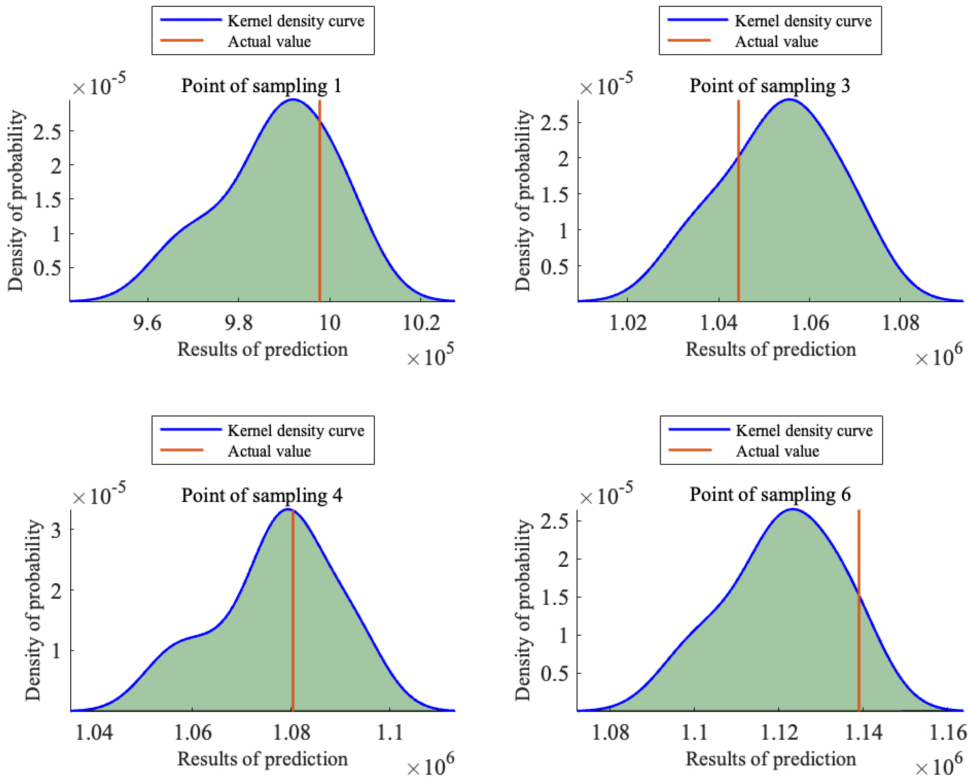


Fig. 6. Probability density function of partial points.

Table 3. Prediction evaluation results for different confidence intervals.

Confidence intervals	PICP	PINAW
95%	0.9021	0.2817
90%	0.8572	0.2515
75%	0.6429	0.1848
50%	0.4822	0.1121
25%	0.2411	0.0515
10%	0.1206	0.0212
5%	0.0603	0.0091

As illustrated in Fig. 6, when employing the DCS-CNN-BiLSTM-Attention-KDE model to forecast the probability density of CO_2 emissions, all actual values are encompassed within the bounds of the probability density function. Except for the sixth observed value, which is located in the tail region of the probability density curve, the remaining values are situated close to the mode of the curve. This finding suggests that the model proposed in this study is capable of capturing the uncertainty associated with changes in CO_2 emissions through the utilization of the probability density curve.

Model Comparative Analysis

To further validate the efficacy and practicality of the proposed model, this study compares the prediction results of the DCS-CNN-BiLSTM-Attention-KDE model with those of simple time series models (ARIMA, SARIMA), point prediction models (BiLSTM, CNN-LSTM, CNN-BiLSTM), and interval prediction models (QRBiLSTM, QRCNN-BiLSTM, QRTCN-BiLSTM). The comparison of point prediction results is presented in Fig. 7.

Combined with interval prediction, the prediction error indices calculated for various models are presented in Table 4. It can be seen from the data in Fig. 7 and Table 4 that the prediction effect of the DCS-CNN-BiLSTM-Attention-KDE model has been significantly improved compared with the simple time series models, point prediction models, and interval prediction models. The error indices MAE, RMSE, and MAPE were substantially decreased, and the fit was improved. When

Table 4. Prediction error metrics for different models.

Model type	Model name	MAE	RMSE	MAPE	PICP	PINAW
Simple time series model	ARIMA	11084.7485	13958.0357	0.0119	—	
	SARIMA	12261.1101	14573.9746	0.0137		
Point prediction model	BiLSTM	49054.8633	53144.6562	0.0448	—	
	CNN-LSTM	78021.3664	83081.5274	0.0699		
	CNN-BiLSTM	63091.2274	68606.2355	0.0564		
Interval prediction model	QR-BiLSTM	63074.0234	75767.9219	0.0591	0.8125	1.2744
	QRCNN-BiLSTM	60469.1953	67897.8359	0.0539	0.8444	0.3515
	QRTCN-BiLSTM	65018.3750	66370.7031	0.0573	0.8357	0.3831
Constructed model	DCS-CNN-BiLSTM-Attention-KDE	10063.0250	13100.8491	0.0092	0.9021	0.2817

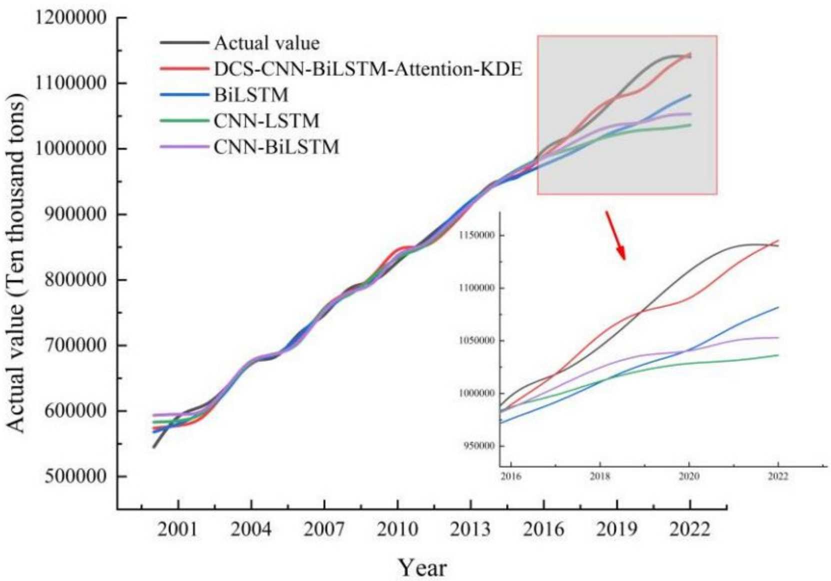


Fig. 7. Comparison of results for point prediction models.

compared to the BiLSTM, CNN-LSTM, and CNN-BiLSTM models, the DCS-CNN-BiLSTM-Attention-KDE model exhibited superior predictive performance and lower error index values for point prediction. Regarding the comparison of interval prediction models, as the PICP value increases and the PINAW value decreases, the feasibility and accuracy of the interval prediction model improve [30]. Thus, with 95% confidence, the PICP value of the DCS-CNN-BiLSTM-Attention-KDE model increased by 8.96%, 5.77%, and 6.64% when compared to QR-BiLSTM, QRCNN-BiLSTM, and QRTCN-BiLSTM, respectively. PINAW decreased by 0.9927, 0.0698, and 0.1014, suggesting that the DCS-CNN-BiLSTM-Attention-KDE model offers superior prediction accuracy and reduced uncertainty.

To evaluate whether the performance improvement of the model is statistically significant, this study conducted a paired-sample t-test on the prediction error of the test set and used Bootstrap sampling (repeated 1000 times) to calculate the 95% confidence intervals for the improvement magnitudes of PICP and PINAW. The results show that the DCS-CNN-BiLSTM-Attention-KDE model is significantly superior to the comparison models in terms of MAE, RMSE, and MAPE. The Bootstrap confidence intervals were [4.12%, 10.31%], [3.85%, 9.28%], and [4.50%, 9.87%], respectively, and none of them crossed the zero value, indicating that the improvement was statistically significant. However, due to the limited sample size (n = 23), the power of statistical tests is relatively low, and further verification in large sample data is necessary in the future.

Table 5. The development rate of each factor under different scenarios.

Factor	Scenario	2023-2025 (%)	2026-2030 (%)	2031-2035 (%)
Per capita GDP	Baseline	7.5	7.5	7.5
	Green and low-CO_2	7.5	7.5	7.5
	Strong low-CO_2	7.5	7.5	7.5
The proportion of the secondary industry in GDP	Baseline	-2.0	-1.5	-1.0
	Green and low-CO_2	-2.5	-2.0	-1.5
	Strong low-CO_2	-3.0	-2.5	-2.0
Investment in the fixed assets of the whole society	Baseline	6.5	3.5	-1.0
	Green and low-CO_2	5.5	2.5	-2.0
	Strong low-CO_2	4.0	1.0	-3.5
Forest coverage rate	Baseline	0.6	0.6	0.6
	Green and low-CO_2	0.8	0.8	0.8
	Strong low-CO_2	1.0	1.0	1.0
Energy intensity	Baseline	-2.6	-3.1	-3.6
	Green and low-CO_2	-3.1	-3.6	-4.1
	Strong low-CO_2	-3.6	-4.1	-4.6
Energy mix	Baseline	-0.4	-0.6	-0.8
	Green and low-CO_2	-0.7	-0.9	-1.1
	Strong low-CO_2	-1.0	-1.2	-1.4
Research and development intensity	Baseline	7.0	8.0	9.0
	Green and low-CO_2	8.0	9.0	10.0
	Strong low-CO_2	8.5	9.5	10.5

Scenario Prediction Analysis

Scenario Parameter Setting

Scenario analysis is one of the most widely utilized methodologies for forecasting the trends in CO_2 emissions [1, 4]. Consequently, in order to scientifically analyze the changes in China's future carbon emissions under different policies and development models, this paper determines the values of the intermediate years based on the historical changing trends of each influencing factor, relevant literature, policy documents, and the average growth rates of the base year and the ending year, and sets the baseline, green low-CO_2, and stringent low-CO_2 scenarios.

In the baseline scenario, it is assumed that China's energy conservation and emission reduction policies have been successfully implemented, and environmental governance has yielded satisfactory outcomes, thereby providing insight into the potential future trend of CO_2 emissions under the present policy framework. The green low-CO_2 scenario suggests gradual progress in innovative green technologies, along with a reduction in energy use required to meet the existing targets for economic development planning. The strong low-CO_2 scenario suggests that China's future will continue to prioritize economic development, with energy consumption and coal usage expected to sustain a significant growth trajectory.

(1) Per capita GDP

Lin et al. [31] forecasted that China's per capita GDP would exhibit an average annual growth potential of approximately 8% between 2020 and 2035, with the basic realization of socialist modernization anticipated by 2035. Zhang and Wang [32] investigated the influence of three key factors: accelerated population aging, the impact of COVID-19, and the technological decoupling between China and the United States, on China's potential rate of economic growth. Within the baseline scenario, China's per capita GDP is projected to be 22,700 by 2035. Cui et al. [33] adopted the relative per capita GDP benchmark method to predict that China's per capita GDP would maintain a growth rate of more than 6.7% before 2035. Meanwhile, the "Blue Book of Chinese Modernization: China's Modernization Development Report (2024)" pointed out that by 2035, China's per capita GDP is expected to reach about \$30,000. Consequently, this paper posits that China's economy will exhibit steady development across three scenarios, with China's per capita GDP reaching $30,000 by 2035, accompanied by an average annual growth rate of 7.49%.

(2) The proportion of the secondary industry in GDP

In recent years, China's industrial structure has undergone a transformation, shifting its focus from the secondary sector to the tertiary sector. Nonetheless, the proportion of the tertiary industry in China still lags behind that of developed countries, suggesting considerable room for optimization in China's industrial structure. Fan and Zhang [34] projected that the proportion of China's secondary industry will fall below 30% by 2035. Consequently, drawing on the research by Fan and Zhang [34] and Liu and Qian [1], this study establishes a baseline scenario where the growth rate of the secondary industry is set at -2.0% for the period 2023-2025 and -1.5% for 2026-2030. For the years 2031 to 2035, the growth rate is projected to be -1.0%. Under the green and low-CO_2 scenario, the growth rate of the added value of the secondary industry is adjusted downwards by 0.5% compared to the baseline scenario. Additionally, considering the industrial structures of developed countries and China's manufacturing industry development goals, the strong low-CO_2 scenario entails a further 0.5% reduction in the added value growth rate of the secondary industry, based on the green and low-CO_2 scenario.

(3) Investment in fixed assets of the whole society

Due to the significant allocation of social fixed asset investment towards China's construction and manufacturing sectors, coupled with the secondary industry's contribution to GDP and the diminishing marginal returns, China's social fixed asset investment is projected to transition from growth to decline by approximately 2035 [1]. Accordingly, under the baseline scenario established in this study, the growth rates of total social fixed asset investment for the periods 2023-2025, 2026-2030, and 2030-2035 are forecasted to be 6.5%, 3.5%, and -1.0%, respectively. In the context of the green and low-CO_2 scenario, the rate of change in total social fixed asset investment from 2023 to 2035 is anticipated to be 1% lower compared to the baseline scenario. Under the strong low-CO_2 scenario, the decline in production capacity of high-CO_2 industries will lead to a reduction in the fixed assets of related sectors, while simultaneously accelerating the process of green investment. Consequently, within the strong low-CO_2 scenario, the growth rate of total social fixed asset investment is expected to decrease by an additional 1.5% relative to the green and low-CO_2 scenario.

(4) Forest coverage rate

The document titled "Opinions on Comprehensively Promoting the Construction of a Beautiful China" mandates the implementation of comprehensive protection and systematic oversight of mountains, rivers, forests, fields, lakes, grasslands, and deserts. We aim to strengthen and enhance the carbon sequestration capacity of ecosystems. It is projected that by 2035, the forest coverage rate in the country will reach 26%. Consequently, aligning with national directives, we establish a baseline scenario where the forest coverage rate is expected to attain 26% by 2035, with an average annual growth rate from 2023 to 2035 is 0.6%. The annual growth rates for the green and low-CO_2 scenario and the strong low-CO_2 scenario are projected to be 0.2% and 0.4% higher, respectively, compared to the baseline scenario.

(5) Energy intensity

Energy intensity is an indicator of the efficiency of energy use. A decline in energy intensity signifies an increase in energy efficiency, typically accompanied by technological advancements and reduced CO_2 emissions in energy-intensive industries. According to Li et al. [35], it is expected that the decrease in energy consumption intensity in China throughout the "14th Five-Year Plan" will be less pronounced compared to what was seen during the "13th Five-Year Plan". As a result, this research utilizes the metrics set forth during the "13th Five-Year Plan". The decline rate observed from 2018 to 2022 serves as the baseline decline rate for 2023 to 2025. Furthermore, the average annual decline rates for the periods 2026-2030 and 2031-2035 are set to increase by 0.5% and 1%, respectively, compared to the baseline. Under the green and low-CO_2 scenario, the rate of decline in energy intensity is projected to be 0.5% higher than that in the baseline scenario. Under the strong low-CO_2 scenario, the rate of decline in energy intensity is expected to further increase by 0.5%, building upon the green and low-CO_2 scenario.

(6) Energy structure

This study employs the proportion of total energy consumption derived from coal as an indicator of energy structure. The substantial proportion of coal has consistently been a pivotal factor in the escalation of CO_2 emissions. The "Strategy for the Revolution of Energy Production and Consumption (2016-2030)" highlights that China is anticipated to enter the later stage of industrialization between 2020 and 2030, with certain provinces progressively advancing towards regional coal-free transitions. Consequently, drawing upon the research conducted by Zhang et al. [36], this study adopts the annual average change rate of energy structure observed between 2018 and 2022 as the baseline decline rate for the period 2023 to 2025, with the annual average change rate decreasing by 0.2% every five years thereafter. An additional 0.3% reduction is applied to the baseline decline rate to represent the decline in the green and low-CO_2 scenario. Furthermore, an extra 0.3% reduction is added to the green and low-CO_2 scenario to define the reduction in the strong low-CO_2 scenario.

(7) Research and development intensity

According to China's "14th Five-Year Plan", the targeted average annual growth rate of R&D investment across society is set at above 7%. Consequently, drawing upon the research by Zhang et al. [36], this study establishes the rate of change in R&D intensity for

Table 6. Results of the carbon peak.

Scenario	Time of carbon peak	Peak value (Ten thousand tons)
Baseline	2033	1551882.69
Green and low-CO_2	2030	1346070.50
Strong low-CO_2	2027	1336442.31

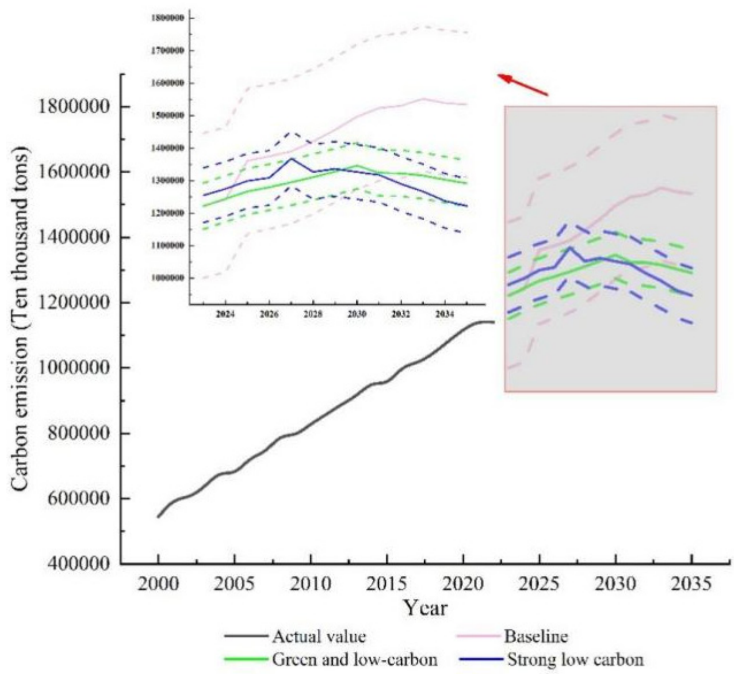


Fig. 8. Forecast results of CO2 emissions from 2023 to 2035.

the period 2023-2025, aligned with the baseline scenario outlined in the "14th Five-Year Plan". The average annual growth rate of R&D investment for the period 2026-2035 is projected to incrementally rise by 1% compared to the 2023-2025 baseline. Under the green and low-CO_2 scenario, as well as the strong low-CO_2 scenario, the growth rates are anticipated to be 1% and 1.5% higher, respectively, than those in the baseline scenario.

The development rate of each factor under different scenarios is shown in Table 5.

Scenario Simulation and Analysis

Utilizing the growth rate parameters for the seven factors presented in Table 5, this study computes the changes in each factor for China over the period from 2023 to 2035, considering three different scenarios. The previously trained DCS-CNN-BiLSTM-Attention-KDE model is employed to predict China's CO_2 emissions for the years 2023 to 2035. The results pertaining to peak CO_2 emissions and the predictive outcomes are presented in Table 6 and Fig. 8, respectively.

The dashed line in Fig. 8 shows the upper and lower bounds of the interval prediction at the 95% confidence level. The following results can be obtained from Fig. 8 and Table 6. Firstly, from the carbon peak perspective, the baseline scenario does not achieve the carbon peak before 2030. The other two scenarios reach the carbon peak target either on schedule or ahead of it, demonstrating that the low-CO_2 development strategy plays a more effective role in advancing the achievement of the "double carbon" objective. In addition, compared with the baseline scenario, the peak times of the green low-CO_2 scenario and the strong low-CO_2 scenario are advanced by 3 years and 6 years, respectively. In the green and low-CO_2 scenario, the possible variation range of the CO_2 peak value is [1275307.94, 1416833.06] ten thousand tons. Compared with the baseline scenario, the predicted peak value of the point is 205812.19 ten thousand tons lower. In the strong low-CO_2 scenario, the possible variation range of the CO_2 peak value is [1283866.05, 1452046.58] ten thousand tons. In addition, the predicted peak value is 215440.38 ten thousand tons lower than the baseline scenario.

Secondly, under both green low-CO_2 and strong low-CO_2 scenarios, China's future CO_2 emissions show a trend of first increasing to the peak and then slowly decreasing after a period of plateau, which is consistent with the existing research conclusions [1, 37]. Meanwhile, the comparison of the two scenarios shows that in 2023-2029, although the CO_2 emission of the strong low-CO_2 scenario is higher than that of the green low-CO_2 scenario, the emission reduction effect of the strong low-CO_2 scenario is better.

Finally, from the perspective of the development trend of CO_2 emissions, with the requirement of the net zero emission goal, the downward trend of CO_2 emissions will continue to slow down. This is mainly because the earlier the CO_2 peak time and the lower the peak value, the easier it is for a country to achieve carbon neutrality smoothly through policy adjustment [1].

Discussion

From the perspective of CO_2 emission quantification, in comparison to traditional methodologies such as direct emission measurement [22], carbon footprint analysis [23], and life-cycle assessment [24], the method of delineating emission scopes essentially encompasses corresponding approaches across diverse boundaries. Consequently, this approach enables more precise quantification of CO_2 emissions and offers a more valuable reference for analyzing trends in CO_2 emission changes.

In terms of research methodologies, the output of traditional point prediction yields solely a single predicted value, which fails to quantitatively characterize the uncertainty associated with the prediction outcome. The DCS-CNN-BiLSTM-Attention-KDE interval prediction model proposed in this study is capable of ascertaining the potential range within which the predicted values are likely to occur. Meanwhile, the accuracy of the CO_2 emission prediction model developed in this research has been enhanced in comparison with those reported in existing studies [4, 17, 38]. For policymakers, the two core metrics defining the model's interval prediction results have clear and actionable interpretability for practical decision-making. In practical policy design, policymakers can interpret these two metrics in tandem: a high-PICP and low-PINAW combination is the optimal outcome, providing both reliable and precise emission range references. Based on the research findings, the predicted trend of CO_2 emission changes presented in this paper aligns with the existing research conclusions [1, 37], with both demonstrating a trend of initial increase, followed by a plateau phase, and subsequent decline. It is imperative to explicitly emphasize that all the prediction results in this paper are predicated on scenario assumptions and the accuracy of emission accounting. The prediction results exhibit sensitivity and are contingent upon the accuracy and consistency of CO_2 emission accounting data. Meanwhile, the model developed in this research is unable to furnish driving factor elasticity coefficients that are as clear and interpretable as those provided by the STIRPAT model. Should there be significant alterations in the actual policy intensity, technological breakthroughs, or data statistical methods in the future, the prediction path may necessitate corresponding adjustments.

The carbon peaking time and peak value forecasted in this study not only mirror the technical performance of the model but also signify the evolution of the economic and energy structure under diverse development scenarios. In comparison to international institutions such as the IEA (2023) and CNRS, along with authoritative domestic research, the baseline scenario predictions in this study exhibit a high degree of consistency with the conclusions of most policy continuation scenarios. The earlier carbon peaking time and lower peak value under the green, low-CO_2, and strong low-CO_2 scenarios align with the IEA's accelerated transformation scenario and the research trend of the 1.5°C pathway proposed domestically, thereby suggesting that proactive emission reduction policies can effectively facilitate the carbon peaking process.

Concurrently, the CO_2 emission path postulated in this research exhibits intrinsic consistency with macro-level variables, including economic growth and the reduction in energy intensity. Particularly in the green and strongly low-CO_2 scenarios, by implementing structural measures such as the accelerated decrease in the proportion of the secondary industry, the optimization of the energy structure, and the enhancement of research and development intensity, a rapid reduction in CO_2 emission intensity has been attained. This reduction is adequate to counterbalance the rise in emissions resulting from economic growth, thereby presenting a well-defined pathway for decoupling economic growth from CO_2 emissions. This finding aligns with the strategic orientation of "promoting high-quality development through an energy consumption revolution" outlined in China's 14th Five-Year Plan. Moreover, it offers quantitative support and a pathway foundation for achieving the goal of reaching the CO_2 emission peak before 2030. Future research can further integrate dynamic Computable General Equilibrium (CGE) models or energy system models to conduct more systematic policy simulations regarding the interaction between CO_2 emission paths and variables, including macroeconomics, employment, and trade, with the aim of enhancing the systematic nature and decision-making reference value of the prediction results.

This study nonetheless presents several limitations. On the one hand, in contrast to existing research that relies on provincial-level predictions [4, 10, 17], this manuscript fails to probe deeply into the carbon peaking and carbon neutrality of various provinces and regions, nor is it capable of predicting the carbon peaking scenario while accounting for regional disparities. The aggregation of data at the national level has the potential to obscure substantial regional heterogeneity. The national carbon peaking time presented in this study ought to be comprehended as the integrated outcome following the coordination of regional pathways. Future research intends to construct provincial panel data models and integrate them with spatial econometric approaches to unveil the synergy, transfer, and spillover effects among regions, thereby offering more refined decision-making support for the development of differentiated climate policies.

Furthermore, given that the CO_2 emission data utilized constitutes a small sample, it is infeasible to conduct training with a large-scale dataset. Hence, in future research, remote sensing data can be taken into account and integrated with statistical data from various regions and provinces to validate the efficacy of the generalization capacity of the interval prediction model and the DCS algorithm under large-sample conditions. Moreover, the principal risk associated with deep models trained on small samples is the comparatively substantial variance in parameter estimation, which may result in conclusions being highly sensitive to the training samples. To bolster robustness, future research might consider employing transfer learning and performing pre-training using large-area panel data with analogous physical mechanisms. In the meantime, the scenario analysis carried out in this study presented the peak years under three scenarios, yet failed to quantify the extent to which different policy levers exerted the greatest impact on peak time or peak magnitude. In future research, control variable methods or explainable machine learning techniques can be adopted for in-depth exploration. Structural mechanisms, such as physical energy balance constraints and technology diffusion models, should be incorporated into the machine learning framework to improve the physical consistency of prediction results and the dynamic characteristics of the system.

Conclusions

Research Conclusions

Utilizing relevant data on China's CO_2 emissions spanning from 2000 to 2022, this study initially estimated China's CO_2 emissions by categorizing emission sources. Subsequently, an index system encompassing economic, social, environmental, energy, and technological factors influencing CO_2 emissions was established. Then, Lasso regression was employed to identify seven key factors that significantly impact CO_2 emission changes. Building on these findings, a DCS-CNN-BiLSTM-Attention-KDE interval prediction model was developed, and its prediction results were benchmarked against other point and interval prediction models. Lastly, the trained DCS-CNN-BiLSTM-Attention-KDE model, integrated with scenario analysis, was utilized to forecast China's CO_2 emissions for the period 2023 to 2035.

(1) According to the results of the Lasso regression analysis, essential elements affecting CO_2 emissions were recognized as GDP per capita, the share of the secondary industry in GDP, societal expenditure on fixed assets, the rate of forest coverage, energy intensity levels, the composition of energy sources, and the intensity of R&D.

(2) The predicted values of the DCS-CNN-BiLSTM-Attention-KDE model demonstrated a high degree of fit with the actual values. The average prediction error for the test set was 0.92%, indicating a strong predictive performance. Meanwhile, as the confidence level increased, the average width of the model's prediction interval expanded continuously, resulting in improved coverage of CO_2 emission data. Furthermore, when the DCS-CNN-BiLSTM-Attention-KDE model was applied to predict the probability density of CO_2 emissions, all actual values fell within the predicted probability density function, suggesting that the model presented in this study can effectively capture the uncertainty associated with CO_2 emission changes through the probability density curve.

(3) Compared to the BiLSTM, CNN-LSTM, and CNN-BiLSTM models, the DCS-CNN-BiLSTM-Attention-KDE model exhibited superior prediction performance and a lower point prediction error index. Regarding the comparison of interval prediction models, under a 95% confidence level, the PICP value of the DCS-CNN-BiLSTM-Attention-KDE model was 8.96%, 5.77%, and 6.64% higher than that of the QR-BiLSTM, QRCNN-BiLSTM, and QRCNN-BiLSTM models, respectively. The PINAW of the DCS-CNN-BiLSTM-Attention-KDE model was reduced by 0.9927, 0.0698, and 0.1014 compared to the QR-BiLSTM, QRCNN-BiLSTM, and QRCNN-BiLSTM models, respectively, suggesting that the proposed model offers improved prediction accuracy and reduced uncertainty.

(4) The simulation and prediction analysis of the baseline, green low-CO_2, and strong low-CO_2 scenarios indicate that the CO_2 emissions peak will occur in 2033, 2030, and 2027, respectively. The possible ranges of peak value are [1328963.56, 1774801.82] ten thousand tons, [1275307.94, 1416833.06] ten thousand tons, and [1283866.05, 1452046.58] ten thousand tons, respectively. Apart from the baseline scenario, both the green low-CO_2 and strong low-CO_2 scenarios have met the “carbon peaking” target.

Policy Recommendations

Drawing on the research findings, with the objective of achieving carbon peaking prior to 2030 and aspiring to an even earlier peaking time, the subsequent policy recommendations are proposed:

(1) Prioritize the advancement of the clean energy structure. Comparative scenario analyses indicate that

the transformation of the energy structure serves as the most pivotal factor influencing the peak time. It is recommended that, in subsequent planning, explicit total control targets for coal consumption be established for each province, and the clean energy substitution roadmap for key industries be actively promoted. Simultaneously, the flexibility transformation and capacity compensation mechanism for coal-fired power plants ought to be enhanced to guarantee that the proportion of non-fossil energy rises to exceed 22%.

(2) Expedite the low-carbon transition of the industrial structure. The reduction in the share of the secondary industry has made a substantial contribution to emission mitigation. It is recommended that high-carbon industries be directed to undergo an orderly transition through measures including green taxation, subsidies for capacity replacement, and the promotion of low-carbon technologies, with the objective of maintaining the share of the secondary industry in GDP below 30% by 2030.

(3) Reinforce scientific and technological innovation and green investment initiatives. The augmentation of R&D intensity exerts a sustained promoting influence on medium- and long-term emission reduction. It is recommended that investment in low-carbon technology R&D be augmented, a provincial green technology transformation fund be set up, and social capital be directed to invest in renewable energy, carbon capture and storage (CCS), and other relevant sectors, with the objective of attaining an R&D intensity of 3.2% by 2030.

Acknowledgements

The authors thank all contributors for their assistance with this research.

Conflict of Interest

The authors declare no conflict of interest.

References

1. LIU C.M., QIAN X.Y. Prediction of carbon emissions from energy consumption in China under the "dual carbon" goal. Resources Science, **10**, 1931, **2023** [In Chinese].
2. CHANG H., ZHAO Y. The impact of carbon trading on the "quantity" and "quality" of green technology innovation: A dynamic QCA analysis based on carbon trading pilot areas. Heliyon, **10** (3), e25668, **2024**.
3. SARKAR B., KAR S., BASU K., GUCHHAIT R. A sustainable managerial decision-making problem for a substitutable product in a dual-channel under carbon tax policy. Computers & Industrial Engineering, **172**, 108635, **2022**.
4. LIU B., CHANG H., LI Y., ZHAO Y. Carbon emissions predicting and decoupling analysis based on the PSO-ELM combined prediction model: evidence from Chongqing Municipality, China. Environmental Science and Pollution Research, **30** (32), 78849, **2023**.
5. CONG J.H., LIU X.M., ZHAO X.R. Demarcation Problems and the Corresponding Measurement Methods of the Urban Carbon Accounting. China Population, Resources and Environment, **4**, 19, **2014** [In Chinese].
6. ZHANG G.X., SU Z.X. Analysis of Influencing Factors and Scenario Prediction of Transportation Carbon Emissions in the Yellow River Basin. Management Review, **32** (12), 283, **2020** [In Chinese].
7. HUANG X.Y., WU J.Y., LIN W.H., WU Q.X. Forecast of carbon emission of Jiangsu Province based on GM(1, 1) model. Heilongjiang Science, **13** (18), 26, **2022** [In Chinese].
8. HU Z., GONG X., LIU H. Prediction of household consumption carbon emission in western cities based on BP model: Case of Xi'an city. Journal of Arid Land Resources and Environment, **34** (7), 82, **2020**.
9. HU Y. Bidirectional analysis model of green investment and carbon emission based on LSTM neural network. Thermal Science, **27** (2 Part B), 1405, **2023**.
10. WEN T., LIU Y., BAI Y., LIU H. Modeling and forecasting CO2 emissions in China and its regions using a novel ARIMA-LSTM model. Heliyon, **9** (11), **2023**.
11. WEI X., XU Y. Research on carbon emission prediction and economic policy based on TCN-LSTM combined with attention mechanism. Frontiers in Ecology and Evolution, **11**, 1270248, **2023**.
12. GAN Z., LI C., ZHOU J., TANG G. Temporal convolutional networks interval prediction model for wind speed forecasting. Electric Power Systems Research, **191**, 106865, **2021**.
13. LI C., TANG G., XUE X., SAEED A., HU X. Short-term wind speed interval prediction based on ensemble GRU model. IEEE Transactions on Sustainable Energy, **11** (3), 1370, **2020**.
14. PENG X., WANG H., LANG J., LI W., XU Q., ZHANG Z., CAI T., DUAN S., LIU F., LI C. EALSTM-QR: Interval wind-power prediction model based on numerical weather prediction and deep learning. Energy, **220**, 119692, **2021**.
15. SHAN B.G., YAO L., ZHANG C.L., TAN X.D. Carbon Peak Prediction of Energy Consumption in China: Based on STIRPAT Extended Model. Environmental Science, **46** (7), 4052, **2025**.
16. LI G.Z., HUANG Q.H. Prediction of Carbon Peak in Beijing-Tianjin-Hebei Region Based on Lasso-GRNN Neural Network Model. Environmental Science, **46** (2), 636, **2025** [In Chinese].
17. LIU B., CHANG H., LI Y., ZHAO Y. Analysis of influencing factors of industrial green and low-carbon transformation under the background of "double carbon": evidence from Sichuan province, China. Frontiers in Environmental Science, **11**, 1257855, **2023**.
18. CHEN J., ZHANG Y., WANG M. Digital intelligence technology innovation, energy transformation and urban carbon emission efficiency. Chinese Journal of Population, Resources and Environment, **23** (2), 205, **2025**.
19. CAI B.F. Advance and Review of City Carbon Dioxide Emission Inventory Research. China Population, Resources and Environment, **10**, 72, **2013**.
20. WANG H.K., ZHANG R.R., BI J. Carbon accuonting for Chinese cities: A case of Wuxi city. Chinese Environmental Science, **31**, 1029, **2011**.
21. ZHANG Z.L., HOU Y.Z., SUN H.H. Calculation of carbon emissions and the difference of low-carbon development efficiency on city territorial space. Journal of Natural Resources, **38** (6), 1464, **2023**.

22. IBRAHIM N., SUGAR L., HOORNWEG D., KENNEDY C. Greenhouse gas emissions from cities: comparison of international inventory frameworks. Local Environment, **17** (2), 223, **2012**.
23. KENNEDY S., SGOURIDIS S. Rigorous classification and carbon accounting principles for low and Zero Carbon Cities. Energy Policy, **39** (9), 5259, **2011**.
24. ZHOU X.H., HU P.C. CHENG P.F. Carbon emission accounting and carbon peak prediction of China's construction industry based on the perspective of full life cycle. Environmental Science, 1, **2024**.
25. DUANKHAN P., SUNAT K., CHIEWCHANWATTANA S., NASA-NGIUM P. The Differentiated Creative search (DCS): Leveraging Differentiated knowledge-acquisition and Creative realism to address complex optimization problems. Expert Systems with Applications, **252**, 123734. **2024**.
26. LU W., LI J., WANG J., QIN, L. A CNN-BiLSTM-AM method for stock price prediction. Neural Computing and Applications, **33** (10), 4741, **2021**.
27. RAI A., SHRIVASTAVA A., JANA K.C. A CNN-BiLSTM based deep learning model for mid-term solar radiation prediction. International Transactions on Electrical Energy Systems, **31** (9), e12664, **2021**.
28. SHAN L., LIU Y., TANG M., YANG M., BAI X. CNN-BiLSTM hybrid neural networks with attention mechanism for well log prediction. Journal of Petroleum Science and Engineering, **205**, 108838, **2021**.
29. MENDEZ M., MERAYO M.G., NUNEA M. Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model. Engineering Applications of Artificial Intelligence, **121**, 106041, **2023**.
30. LI Q., XU S.A. Power load interval prediction based on GSABO-BP and Bootstrap. Modern Electronics Technique, **10**, 28, **2024**.
31. LIN Y.F., WEN Y.H., GU Y.W. Prospect of China’s Economic Growth Potential: 2020-2035, 2035-2050. Finance Forum, **27** (6), 3, **2022** [In Chinese].
32. ZHANG X.J., WANG Y. Prospects for economic growth under the long-range objectives of socialist modernization: Based on the measurement of the potential economic growth rat. Social Sciences in China, **4**, 4, **2023** [In Chinese].
33. CUI J.F., CHEN J.W., MIAO J.J. Research on the potential of China's economic growth under the goal of Chinese-style modernization. Lanzhou Academic Journal, **4**, 43, **2024** [In Chinese].
34. FAN D.C., ZHANG X.F. Scenario prediction of China’s carbon emissions based on PSO-BP Neural Network model and research on low carbon development path. Sino-Global Energy, **26** (8), 11, **2021**.
35. LI J.F., GU A.L., ZHANG C.L., SUN Z.Q. China’s provincial economic development, energy demand and carbon emission prospects during the 14th Five-Year Plan period: an analysis based on CMRCGE model. Climate Change Research, **15** (6), 649, **2019**.
36. ZHANG G.M., WANG T., LOU Y.Y., GUAN Z.C., ZHENG H.J., LI Q., WU J.Q. Research on China’s Provincial Carbon Emission Peak Path Based on a LSTM Neural Network Approach. Chinese Journal of Management Science, 1, **2024**.
37. ZHANG X., ZHANG D. Urban carbon emission scenario prediction and multi-objective land use optimization strategy under carbon emission constraints. Journal of Cleaner Production, **430**, 139684, **2023**.
38. SUN W., HUANG C. Predictions of carbon emission intensity based on factor analysis and an improved extreme learning machine from the perspective of carbon emission efficiency. Journal of Cleaner Production, **338**, 130414, **2022**.

Appendix

Appendix 1. Abbreviation explanation and description.

Abbreviation	Full name
AKDE	Adaptive Kernel Density Estimation
Attention	Attention Mechanism
BiLSTM	Bidirectional Long Short-Term Memory Network
CCS	Computable General Equilibrium
CNN	Convolutional Neural Network
CV	Cross-Validation
DCS	Differentiated Creative Search
DKA	Differentiated Knowledge Acquisition
IEA	International Energy Agency
KDE	Kernel Density Estimation
LSTM	Long Short-Term Memory Network
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NBS	National Bureau of Statistics
PICP	PI Coverage Probability
PINAW	PI Normalized Average Width
PSO	Particle Swarm Optimization
QR	Quantile Regression
QRCNN	Quantile Regression in Convolutional Neural Networks
RA	Retrospective Assessment
RMSE	Root Mean Square Error
TCN	Temporal Convolutional Neural Networks
GWO	Grey Wolf Optimizer
VIF	Variance Inflation Factor
WOA	Whale Optimization Algorithm