



# China's carbon dioxide emission forecast based on improved marine predator algorithm and multi-kernel support vector regression

Xiwen Qin<sup>1,2</sup> · Siqi Zhang<sup>1</sup> · Xiaogang Dong<sup>1</sup> · Yichang Zhan<sup>1</sup> · Rui Wang<sup>1</sup> · Dingxin Xu<sup>1</sup>

Received: 1 April 2022 / Accepted: 26 July 2022 / Published online: 18 August 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

Global warming has constituted a major global problem. Carbon dioxide emissions from the burning of fossil fuels are the main cause of global warming. Therefore, carbon dioxide emission forecasting has attracted widespread attention. Aiming at the problem of carbon dioxide emissions forecasting, this paper proposes a new hybrid forecasting model of carbon dioxide emissions, which combines the marine predator algorithm (MPA) and multi-kernel support vector regression. For further strengthening the prediction accuracy, a novel variant of MPA is proposed, called EGMPA, which introduces the elite opposition-based learning strategy and the golden sine algorithm into MPA. Algorithm test results show that EGMPA can effectively improve the convergence speed and optimization accuracy. The carbon dioxide emission data of China from 1965 to 2020 are taken as the research objects. Root-mean-square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are used to evaluate the performance of the proposed model. The proposed multi-kernel support vector regression model is used to forecast China's carbon dioxide emissions during the "14th Five-Year Plan" period. The results show that the proposed model has RMSE of 37.43 Mt, MAE of 30.63 Mt, and MAPE of 0.32%, which significantly improves the prediction accuracy and can accurately and effectively predict China's carbon dioxide emissions. During the "14th Five-Year Plan" period, China's carbon dioxide emissions will continue to show an increasing trend, but the growth rate will slow down significantly.

**Keywords** Carbon dioxide emissions · Multi-kernel support vector regression · Marine predator algorithm · Elite opposition-based learning strategy · Golden sine algorithm

## Introduction

Since the industrial revolution, in order to accelerate the global industrialization process, it is necessary to consume a large amount of coal and oil, and because of the rapid growth of population and deforestation for economic development, the greenhouse gases with strong heat absorption such as carbon dioxide emitted into the atmosphere by human beings

are increasing year by year; carbon dioxide emissions are the main cause of the greenhouse effect (Shakun et al. 2012). Therefore, the greenhouse effect of the atmosphere has also increased, causing a series of serious problems such as global warming, which has attracted the attention of all countries in the world. In order to collectively address the challenge of climate change and slow the global warming trend, in December 2015, nearly 200 parties jointly adopted The Paris Agreement, which sets out action arrangements for how the world will address climate change after 2020. As a carbon emitter and the world's second largest economy, China's advocacy on climate change has far-reaching influence. As a signatory to The Paris Agreement, China's carbon peak in 2030 and carbon neutrality in 2060 is China's commitment to the world. Therefore, the accurate prediction of China's carbon dioxide emissions has become a research hotspot.

There are many prediction models for carbon dioxide emissions in China, and the types of prediction models are

---

Responsible Editor: Marcus Schulz

✉ Xiwen Qin  
qinxiwen@ccut.edu.cn

<sup>1</sup> School of Mathematics and Statistics, Changchun University of Technology, No. 2055 Yan'an Street, Chaoyang District, Changchun 130012, China

<sup>2</sup> Graduate School, Changchun University of Technology, No. 2055 Yan'an Street, Chaoyang District, Changchun 130012, China

quite different, and the results obtained are also different. Li et al. using a hybrid approach combining the log-mean divisor index (LMDI) and decoupling index methods, the results suggest that China may struggle to achieve its goal of peaking CO<sub>2</sub> emissions by 2030 (Li and Qin 2019). Zhang et al. conducted a dynamic simulation and projection of China's carbon dioxide emissions from a system dynamics perspective. The results show that China will be able to achieve its emission reduction target by 2030 (Zhang et al. 2018). Qiao et al. proposed a hybrid model based on the improved lion swarm optimizer and LSSVM for predicting carbon dioxide emissions; the results show that global carbon emissions will continue to increase in the next few years, and only Japan is likely to achieve carbon emissions targets, while other countries, including China, are challenging to achieve the emissions targets (Qiao et al. 2020). Niu et al. proposed a prediction model combining improved fireworks algorithm (IFWA) optimization and generalized regression neural network (GRNN), and the results showed that there is still some pressures to achieve the carbon intensity reduction target under the current policy efforts in China (Niu et al. 2020). Although there are many models for forecasting carbon dioxide emissions in China, the types of models are quite different, and the predicted results are also different. Therefore, it is of great significance to establish an accurate and effective model for forecasting carbon dioxide emissions in China.

The main objective of this research is to use machine learning methods and meta-heuristic algorithms to establish a scientific and effective model for forecasting China's carbon dioxide emissions and accurately predict China's carbon dioxide emissions in the next five years. The main question of this research is how to build an accurate model for forecasting China's carbon dioxide emissions. The motivation of this research is to accurately predict the future trend of China's carbon dioxide emissions. Based on the above research goals, research questions, and research motivations, this study combined the multi-kernel support vector regression model with the improved marine predator algorithm (MPA) to establish a forecast model for China's carbon dioxide emissions. To make up for the limitations of traditional MPA, this paper integrates the elite opposition-based learning strategy and the Golden Sine Algorithm to improve the traditional MPA. Since carbon dioxide emission data are usually small sample data, this paper chooses multi-kernel support vector regression model as the prediction model of carbon dioxide emission. The multi-kernel support vector regression model selects different kernel functions for weighted combinations, which effectively enhances the learning ability and generalization ability of the model and improves the accuracy and stability of the prediction model. The main contributions of this paper are summarized as follows:

- (1) A novel variant of MPA-EGMPA is proposed, which introduces the elite opposition-based learning strategy and the golden sine algorithm into MPA, solving the problems of low convergence accuracy and the tendency to fall into local optimum of traditional MPA.
- (2) A new multi-kernel support vector regression is proposed. Multi-kernel function combines multiple kernel functions to replace single kernel function, which improves the learning effect of support vector regression model effectively, and EGMPA is used to optimize the main parameters of multi-kernel support vector regression, which solves the problem of parameter selection of multi-kernel support vector regression;
- (3) The hybrid model of EGMPA and multi-kernel support vector regression is selected to predict carbon dioxide emissions in China. The proposed hybrid model is applied to predict China's carbon dioxide emissions during the "14th Five-Year Plan" period to assess whether China will be able to achieve its carbon peak target by 2030.

The rest of this paper is organized as follows: "Literature review" is literature review; "Methods" briefly introduces the relevant theories and the hybrid forecasting model framework of this paper; "Algorithm test" uses 23 sets of benchmark functions to test EGMPA algorithms; "Results and discussion" uses the proposed model to forecast China's carbon dioxide emissions; "Conclusions" is the conclusion of this paper and future research directions.

## Literature review

### Meta-heuristic algorithms

Meta-heuristic algorithms are a method for solving optimal solutions to complex optimization problems based on computational intelligence mechanisms. Meta-heuristic algorithms can be divided into three categories: evolutionary algorithms, swarm intelligence algorithms, and physics-based algorithms.

Evolutionary algorithm is mainly through the simulation of the survival of the fittest evolutionary law in nature, to achieve the overall progress of the population, and finally complete the solution of the optimal solution. Evolutionary algorithms are mainly represented by genetic algorithm (GA) (Holland 1975) and differential evolution (DE) (Storn and Price 1997). Genetic algorithm (GA) is a random global search optimization method, which simulates the replication, cross-over, and mutation phenomena occurring in natural selection and heredity. Differential evolution algorithm (DE) is proposed based on genetic algorithm. Later, many other famous evolutionary algorithms were derived, such as

evolutionary strategy (ES) (Hansen et al. 2003) and biogeography-based optimization (BBO) (Simon 2008).

Swarm intelligence algorithm is an emerging optimization method that simulates various group behaviors of social animals and uses the information interaction and cooperation among individuals in a group to achieve optimization purposes. Particle swarm optimization (PSO) (Eberhart and Kennedy 1995) is a typical representative of swarm intelligence algorithms, which is a global stochastic search algorithm based on swarm intelligence by simulating the migration and clustering behavior of birds during foraging. In addition, many intelligent optimization algorithms simulate the behavior of other swarms. Artificial bee colony (ABC) (Karaboga and Basturk 2007) is an optimization method proposed to mimic the behavior of bees. It only requires comparing the quality of problem solutions. Grey wolf optimizer (GWO) (Mirjalili et al. 2014) is an optimization search method inspired by the hunting activities of grey wolves. Whale optimization algorithm (WOA) (Mirjalili and Lewis 2016) is proposed based on the behavior of whales to round up prey. The local search of the algorithm is realized by the mechanism of shrinking enclosing and spiraling updates, and the global search of the algorithm is realized by the random learning strategy. Salp swarm algorithm (SSA) (Mirjalili et al. 2017) simulates the swarm behavior of the salp chain. In each iteration, the leader guides followers to move towards food in a chain behavior. In the process of movement, leaders explore globally, while followers fully explore locally, which greatly reduces the situation of falling into local optimum.

The physics-based algorithms are mainly inspired by different physical laws in nature. Gravitational search algorithm (GSA) (Rashedi et al. 2009) is a novel optimization algorithm based on the laws of gravity and particle interactions in the universe. Multi-verse optimizer (MVO) (Mirjalili et al. 2016) is an optimization algorithm based on the multiverse theory in physics, which simulates the motion behavior of the multiverse population under the joint action of white holes, black holes, and wormholes. Sine cosine algorithm (SCA) (Mirjalili 2016) uses sine and cosine mathematical models to solve the optimization problem. Golden sine algorithm (Golden-SA) (Tanyildizi and Demir 2017) is inspired by the scanning within the unit circle of the sine function, which is similar to the spatial search of the solution to be optimized. The algorithm has strong local exploitation capabilities.

Marine predator algorithm (MPA) is a new meta-heuristic algorithm proposed by Faramarzi et al. in 2020. The main inspiration for MPA comes from a wide range of foraging strategies, namely, Lévy and Brownian movements of marine predators, and policies for optimal encounter rates for biological interactions between predators and prey (Faramarzi et al. 2020). MPA has been widely studied by scholars

for its powerful optimization ability. MPA has been used in many practical problems, such as rolling bearing fault diagnosis (Wang et al. 2021c), rotor angle stability of power systems (Yakout et al. 2021), sensor optimization design (Aly et al. 2021), and COVID-19 confirmed case prediction (Al-Qaness et al. 2020). In addition, in view of the disadvantages of the traditional MPA, such as slow convergence speed and easy to fall into local optimal, scholars have made various improvements to it. Zhong et al. combined MPA with teaching-based optimization algorithm and incorporated mutation and crossover strategies to effectively increase predator diversity and avoid premature convergence (Zhong et al. 2020). Elaziz et al. introduced the quantum theory to MPA and applied it to solve the problem of multi-threshold image segmentation (Elaziz et al. 2021). Oszust introduced local escaping operator (LEO) to balance the exploration and exploitation capabilities of MPA (Oszust 2021). Abdel et al. proposed a binary version of MPA to solve the 0-1 knapsack (KP01) problem (Abdel et al. 2020). Yang et al. proposed a multi-strategy MPA containing a chaotic opposition learning strategy, a neighborhood dimensional learning strategy, adaptive inertia weights, and adaptive step control factors, and applied it to logging oil formation identification (Yang et al. 2021). Wang et al. introduced the location update strategy of PSO to make up for the shortcomings of MPA in global search (Wang et al. 2021b). Houssein et al. introduced opposition learning strategy into MPA to improve the convergence speed of MPA (Houssein et al. 2021a, b). Aiming at the problems of slow convergence speed, weak local search ability, and easy to fall into local optimal solution of MPA, this paper combines elite opposition-based strategy and golden sine algorithm to improve MPA. Firstly, elite opposition-based learning strategy is introduced in the initialization stage to improve the diversity of the initialization population, increase the search space, and enhance the global search ability of the algorithm. Secondly, gold sine algorithm is integrated into the position update mode of MPA in phase 3 to enhance the optimization performance of MPA.

## Carbon dioxide emissions forecast

In recent years, many scholars have proposed different models to predict carbon dioxide emissions. According to the summary of the literature reading, the main carbon dioxide emission prediction models mainly include econometric and mathematical models, machine learning models and hybrid models. For econometrics and mathematical models, Wang et al. used a nonlinear grey multivariate model for predictive analysis of China's carbon dioxide emissions, and the proposed model has higher prediction accuracy (Wang and Ye 2017). Mga et al. introduced the environmental kuznets curve hypothesis and differential information principle in

a fractional grey Ricardian model, and the results suggest that the USA, China, and Japan will gradually reduce their carbon emissions in the future (Gao et al. 2020). Wang et al. investigated the drivers of carbon emissions from electricity generation in China based on a two-stage LMDI approach (Wang et al. 2021a). Song et al. used ST-LMDI model to identify potential dynamic trends of China's carbon emissions under different scenarios (Song et al., 2022b). For machine learning models, Abbas et al. proposed a multi-stage method of integrated clustering, machine learning, and dimensionality reduction to predict carbon dioxide emissions based on two important variables, economic growth, and energy consumption. The developed multi-stage method can accurately predict carbon dioxide emissions (Mardani et al. 2020). Lin et al. proposed an improved attention mechanism-based long short-term memory (LSTM) neural network (attention-LSTM) to predict the carbon dioxide emissions of different countries or regions (Lin et al. 2021). Sun et al. combined PCA and RELM to predict carbon dioxide emissions in China (Sun and Sun 2017). In order to make up for the shortcomings of a single model, many scholars used hybrid models to predict carbon dioxide emissions. Zhou et al. used support vector machines and improved particle swarm optimization (IPSO) to forecast carbon dioxide emissions in the Chinese power sector (Zhou et al. 2018). Qiao et al. proposed a hybrid model based on the improved lion swarm optimizer and LSSVM for predicting carbon dioxide emissions. (Qiao et al. 2020). Niu et al. proposed a prediction model combining improved fireworks algorithm (IFWA) optimization and generalized regression neural network (GRNN) (Niu et al. 2020). Wen and Yuan combined random forest, PSO, and BP neural network to establish a carbon dioxide hybrid prediction model (Wen and Yuan 2020). Some scholars have analyzed the impact of COVID-19 on air quality in India and the UK, and the results show that the blockade measures during the COVID-19 period cannot improve air quality in the long run (Shehzad et al. 2021 ; Ropkins and Tate 2021).

In addition, many scholars have analyzed and studied the factors influencing carbon dioxide emissions. Sangeetha and Amudha, Heydari et al. and Lin et al. selected historical data of energy sources such as oil, coal, natural gas, and primary energy sources to predict carbon dioxide emissions, and concluded that various energy consumptions play a crucial role in carbon dioxide emissions. (Sangeetha and Amudha 2018; Heydari et al. 2019; Lin et al. 2021). Song et al. investigated the characteristics of carbon intensity in various regions of China, and energy intensity was the main factor affecting the spatial differences and temporal dynamics (Song et al. 2019). Niu et al. considered seven categories of influencing factors for carbon dioxide emissions projections: economy, population, economic structure, industrial structure, energy

intensity, energy mix, and stringency of environmental policies (Niu et al. 2020). You et al. investigated the effect between coal-related carbon emissions (CCE) and economic growth, and showed that activity and energy intensity had the greatest effect on CCE (You et al. 2020). Song et al. investigated the temporal and spatial differences in per capita carbon emissions across Chinese provinces, and energy consumption was the main factor affecting per capita carbon emissions. (Song et al. 2022a). From the above literature on carbon dioxide emissions, it can be seen that traditional econometric models and machine learning models have been widely used in carbon dioxide emissions prediction. However, a single model may have problems such as poor robustness. Therefore, this paper combines the multi-kernel support vector regression model with the improved MPA algorithm to establish an accurate and effective hybrid model for forecasting China's carbon dioxide emissions.

## Methods

### Multi-kernel support vector regression

#### Support vector regression

Support vector regression (SVR) (Liang and Sun 2003) is the application of support vector machine (SVM) (Cortes and Vapnik 1995) to regression problems. Given the training dataset  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ , SVR is to construct the regression function with the following form:

$$f(x) = \omega^T x + b \quad (1)$$

where  $\omega$  is a coefficient vector, and  $b$  is a constant. The optimization problem of SVR can be expressed in Eq. (2):

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m \ell_{\varepsilon}(f(x_i) - y_i) \quad (2)$$

where  $C$  is the regularization constant and  $\ell_{\varepsilon}$  is the  $\varepsilon$ -insensitive loss function. By introducing non-negative slack variables  $\xi_i$  and  $\hat{\xi}_i$ , the constraint optimization problem can be expressed as follows:

$$\begin{aligned} & \min_{\omega, b, \xi_i, \hat{\xi}_i} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) \\ & s.t. \begin{cases} f(x_i) - y_i \leq \varepsilon + \xi_i, \\ y_i - f(x_i) \leq \varepsilon + \hat{\xi}_i, \\ \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, m. \end{cases} \end{aligned} \quad (3)$$

Through constructing a Lagrangian function, the optimization problem can be transformed into the dual problem and its solution is given by Eq. (4),

$$\begin{aligned} & \max_{\alpha, \hat{\alpha}} \sum_{i=1}^m y_i (\hat{\alpha}_i - \alpha_i) - \varepsilon (\hat{\alpha}_i + \alpha_i) - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\hat{\alpha}_i - \alpha_i) (\hat{\alpha}_j - \alpha_j) x_i^T x_j \\ & \text{s.t.} \begin{cases} \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) = 0, \\ 0 \leq \alpha_i, \hat{\alpha}_i \leq C. \end{cases} \end{aligned} \tag{4}$$

where  $\alpha_i$  and  $\hat{\alpha}_i$  are both Lagrange multiplier vectors. The solution of Eq. (1) is obtained by solving the dual problem:

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i^T x + b \tag{5}$$

Considering the feature mapping form, the corresponding SVR solution can be expressed as:

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \varphi(x_i)^T \varphi(x_j) + b \tag{6}$$

If  $\varphi(x_i)^T \varphi(x_j)$  is written in kernel function form, the final decision function can be written as:

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \kappa(x, x_i) + b \tag{7}$$

### Kernel function

Kernel functions have many forms, common kernel functions are the following:

(1) Linear kernel

$$\kappa_{Linear}(x_i, x_j) = x_i^T x_j \tag{8}$$

Linear kernel is the simplest kernel function and is mainly used for linearly separable cases. However, since most problems in practice are nonlinear, nonlinear kernel is more commonly used.

(2) Polynomial kernel

$$\kappa_{Poly}(x_i, x_j) = ((x_i^T x_j) + R)^d \tag{9}$$

where  $d \geq 1$  is the order of the polynomial.

(3) RBF kernel

$$\kappa_{RBF}(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{10}$$

where  $\sigma > 0$  is the bandwidth of the RBF kernel function, which determines the scope of action of the kernel function.

(4) Sigmoid kernel

$$\kappa_{Sigmoid}(x_i, x_j) = \tanh(\beta x_i^T x_j + \theta) \tag{11}$$

where  $\tanh$  is the hyperbolic tangent function,  $\beta$  is the slope, and  $\theta$  is the intercept.

### Multi-kernel support vector regression

In the process of SVR modeling, the selection of kernel function is very important. Single kernel cannot achieve good learning results, so multi-kernel learning has attracted extensive attention. Multi-kernel learning can effectively improve the learning effect of the model by weighting multiple kernel functions together instead of single kernel functions. A multi-kernel function can be expressed as:

$$\begin{aligned} \kappa_{multiple}(x_i, x_j) &= \sum_{\alpha=1}^N m_{\alpha} \kappa_{\alpha}(x_i, x_j) \\ \sum_{\alpha=1}^N m_{\alpha} &= 1, 0 < m_{\alpha} < 1, \alpha = 1, 2 \dots, N \end{aligned} \tag{12}$$

where  $m_{\alpha}$  is the weight of different kernel functions in multi-kernel functions, and the sum of the weights is 1. In this paper, multi-kernel functions are constructed based on linear kernel, RBF kernel, polynomial kernel, and Sigmoid kernel, and the multi-kernel functions are used in the multi-kernel support vector regression model.

### Marine predator algorithm

The marine predator algorithm (MPA) is a novel meta-heuristic optimization algorithm proposed by Afshin Faramarzi et al. in 2020, which is inspired by the survival of the fittest theory of the marine, where marine predators choose the best foraging strategy by choosing between Lévy or Brownian wandering.

According to Darwin’s theory of survival of the fittest, top predators in nature are more talented in foraging. Therefore, an Elite matrix is constructed based on top predators. This matrix supervises the search and finding of prey based on information about its location.

$$Elite = \begin{bmatrix} X_{1,1}^l & X_{1,2}^l & \dots & X_{1,d}^l \\ X_{2,1}^l & X_{2,2}^l & \dots & X_{2,d}^l \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1}^l & X_{n,2}^l & \dots & X_{n,d}^l \end{bmatrix}_{n \times d} \tag{13}$$

where  $\bar{X}^l$  represents the top predator vector, which is replicated  $n$  times to construct the Elite matrix.  $n$  is the number of search agents, and  $d$  is the dimension.

Another matrix with the same dimension as Elite is called Prey, and the predator positions are updated based on this matrix.

$$Prey = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,d} \\ X_{2,1} & X_{2,2} & \dots & X_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n,1} & X_{n,1} & \dots & X_{n,d} \end{bmatrix}_{n \times d} \tag{14}$$

where  $X_{i,j}$  represents the  $j$ th dimension of the  $i$ th prey.

MPA optimization process is described as follows:

*Step 1:* Initialization phase.

Similar to most of the meta-heuristic algorithms, MPA randomly initializes the prey position in the search space and initializes the Prey matrix with Eq. (15):

$$X_0 = X_{\min} + \text{rand}(X_{\max} - X_{\min}) \tag{15}$$

where  $X_{\max}$  and  $X_{\min}$  is the upper and lower bound of the search space, *rand* is a uniform random vector in the range of 0 to 1.

*Step 2:* MPA optimization phase.

*Phase 1:* This phase is mainly used for global search. When the predator is moving faster than the prey, the prey takes a Brownian motion and the predator does not move. The mathematical model of this rule is applied as:

$$\begin{cases} \text{stepsize} = R_B \otimes (Elite_i - R_B \otimes Prey_i) \\ Prey_i = Prey_i + P \cdot R \otimes \text{stepsize}_i \end{cases} \quad i = 1, 2, \dots, n \tag{16}$$

where  $R_B$  is a vector containing random numbers based on normal distribution representing the Brownian motion.  $P=0.5$  is a constant number.  $R$  is a vector of uniform random numbers in  $[0,1]$ .  $Iter$  is the current iteration, while  $Max\_Iter$  is the maximum one.

*Phase 2* In this phase, the global search is transformed into local search for the current optimal solution. Half the population is used for exploitation and the other half for exploration. When the predator and prey are at the same speed, prey is responsible for exploitation based on Lévy motion strategy. Predator is responsible for exploration based on Brownian motion strategy. The mathematical description of exploitation and exploration is as follows:

$$\begin{cases} \text{stepsize}_i = R_L \otimes (Elite_i - R_L \otimes Prey_i) \\ Prey_i = Prey_i + P \cdot R \otimes \text{stepsize}_i \end{cases} \quad i = 1, 2, \dots, n/2 \tag{17}$$

$$\begin{cases} \text{stepsize}_i = R_B \otimes (R_B \otimes Elite_i - Prey_i) \\ Prey_i = Elite_i + P \cdot CF \otimes \text{stepsize}_i \end{cases} \quad i = n/2 \dots, n \tag{18}$$

where  $R_L$  is a vector of random numbers based on Lévy distribution.  $CF = \left(1 - \frac{Iter}{Max\_Iter}\right)^{\left(2 \frac{Iter}{Max\_Iter}\right)}$  is considered as an adaptive parameter to control the step size for predator movement.

*Phase 3:* In this phase, the local search for the current optimal solution location is carried out. When the predator is slower than the prey, the predator adopts the exploitation strategy based on Lévy motion. This phase is presented as:

$$\begin{cases} \text{stepsize}_i = R_L \otimes (R_L \otimes Elite_i - Prey_i) \\ Prey_i = Elite_i + P \cdot CF \otimes \text{stepsize}_i \end{cases} \quad i = 1, \dots, n \tag{19}$$

*Step 3:* FADs' effect and Eddy formation.

Fish aggregation devices (FADs) or eddy effects usually alter the foraging behavior of marine predators, a strategy that enables MPAs to overcome premature convergence problems and escape local extremes during the search for an optimum. Its mathematical description is as follows:

$$Prey_i = \begin{cases} Prey_i + CF[X_{\min} + R \otimes (X_{\max} - X_{\min})] \otimes U & \text{if } r \leq FADs \\ Prey_i + [FADs(1 - r) + r](Prey_{r1} - Prey_{r2}) & \text{if } r > FADs \end{cases} \tag{20}$$

where  $FADs=0.2$  is the probability of FAD effect on the optimization process.  $U$  is the binary vector with arrays including 0 and 1.  $r$  is the uniform random number in  $[0,1]$ .  $r1$  and  $r2$  subscripts denote random indexes of Prey matrix.

### Elite opposition-based learning strategy

The performance of the meta-heuristic algorithm is affected by the quality of the initial solution.

In order to further improve the population diversity of MPA, this paper applies the elite opposition-based learning strategy to the initial population stage of MPA. The elite individual most probably has more useful information than other individuals. The opposition solution of current elite individuals is generated, and the opposition solution of elite individuals is integrated with the current solution, and excellent individuals are selected to form a new population, which effectively increases the population diversity of MPA, improves the quality of population, and avoids the algorithm falling into local optimum.

The individual  $X_e = (X_{e,1}, X_{e,2}, \dots, X_{e,D})$  corresponding to the extreme point of fitness function in algorithm iteration is regarded as the elite individual of the population

$X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,D})$ . The elite opposition solution  $X'_i = (X'_{i,1}, X'_{i,2}, \dots, X'_{i,D})$  can be defined as:

$$X'_{i,j} = k(lb_j + ub_j) - X_{e,j}, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, D. \tag{21}$$

Where  $n$  is the population size,  $k$  is the random number of  $[0, 1]$ , and  $lb_j$  and  $ub_j$  are the dynamic boundary,  $lb_j = \min(X_{i,j})$ ,  $ub_j = \max(X_{i,j})$ . If the elite opposition solution exceeds the given boundary, it can be randomly generated according to Eq. (22):

$$X'_{i,j} = \text{rand}(lb_j, ub_j), \text{ if } X'_{i,j} < lb_j \text{ or } X'_{i,j} > ub_j \tag{22}$$

### Golden sine algorithm

Golden sine algorithm (Golden-SA) is a new meta-heuristic algorithm proposed by Tanyildizi et al. in 2017. Its inspiration comes from the spatial search of the solution to the problem to be optimized by scanning inside the unit circle of sine function. The golden ratio is used to narrow the search space to approximate the optimal solution. In the iterative process, the golden sine algorithm firstly generates the initial positions of  $N$  individuals, and updates the positions of each individual according to Eq. (23) :

$$X(t + 1) = X(t) * \left| \sin(R_1) \right| - R_2 * \sin(R_1) * |m_1 * X_{\text{best}}(t) - m_2 * X(t)| \tag{23}$$

where  $X_{\text{best}}(t)$  is the target position and it is the global optimal position in Gold-SA.  $R_1 \in [0, 2\pi]$ ,  $R_2 \in [0, \pi]$ ,  $m_1$ , and  $m_2$  are the coefficients obtained by the golden section method. These coefficients effectively narrow the search area and guide individuals to the optimal value. The values of  $a$  and  $b$  will change with the iteration of the algorithm; initial default values for  $a$  and  $b$  are considered to be  $-\pi$  and  $\pi$ , respectively.

$$m_1 = a * (1 - \tau) + b * \tau \tag{24}$$

$$m_2 = a * \tau + b * (1 - \tau) \tag{25}$$

$$\tau = \frac{\sqrt{5} - 1}{2} = 0.618033 \tag{26}$$

### MPA improvements

In order to improve the performance of MPA, this paper combines elite opposition-based learning strategy and golden sine algorithm to improve MPA. Elite opposition-based learning strategy can effectively improve the population quality of MPA. Firstly, the MPA

incorporating elite opposition-based learning strategy improves the diversity of the initial population and increases the search space; Secondly, in each iteration, the elite opposition-based learning strategy can generate an opposition solution far from the local extreme point, which leads the MPA to jump out of the local extreme point and enhances the global search ability of the algorithm. In addition, the elite opposition-based learning strategy uses the tracking search mode of dynamic boundary to locate individuals in the gradually reduced search area, which can effectively improve the convergence accuracy and speed of MPA.

In addition, based on retaining the traditional MPA's position update method in *phase1* and *phase2*, the golden sine algorithm is used to improve the position update method of the MPA in *phase3* to enhance the optimization performance and accuracy of the algorithm. The improved position update method as shown in Eq. (27):

$$Prey_i = Prey_i * \left| \sin(R_1) \right| - R_2 * \sin(R_1) * |m_1 * Elite_i - m_2 * Prey_i| \tag{27}$$

In order to simplify the algorithm and improve the operating efficiency, according to the research of Xie et al (Xie et al. 2019),  $m_1$  and  $m_2$  are set as constants. This approach guarantees that the search ability of the agents will not weaken, while increasing the stability of the algorithm. The specific formulas of  $m_1$  and  $m_2$  as shown in Eq.(28) and Eq.(29):

$$m_1 = -\pi + (1 - \tau) * \pi \tag{28}$$

$$m_2 = -\pi + \tau * \pi \tag{29}$$

Based on the above improvement of MPA, the pseudo-code of the proposed EGMPA is shown in the Table 1:

### Hybrid forecasting model framework

The specific flow of the hybrid prediction model in this paper is shown in Fig. 1, which consists of four steps: data preprocessing, forecasting model based on multi-kernel support vector regression, improved marine predator optimization model, and model evaluation criteria.

#### Step 1: Data preprocessing

To eliminate the adverse effects of large size differences in the raw data on the prediction, the raw data were preprocessed using normalization and restricted to the range of  $[0,1]$ .

#### Step 2: Forecasting Model

**Table 1** Pseudo-code of EGMP

Pseudo-code of EGMPA
Initialize algorithm parameters $Max\_Iter=100$ , $N=30$ , $P=0.5$ , $FADs=0.2$ .
Randomly initialize the prey matrix $Prey_i$ :
While $Iter < Max\_Iter$ :
for $i \leq N$ :
Calculate the fitness function value $fit_i$ according to $Prey_i$ ;
Apply Elite Opposition-Based Learning to generate the Elite Opposition solution $Prey_i'$ and add it to the population OP.
Calculate the fitness function value of the newly generated population OP;
Select individuals with excellent fitness values as the next generation population;
end for
if $Iter < Max\_Iter/3$ then:
Update the prey according to Eq.(16)
else if $Max\_Iter/3 < Iter < 2 * Max\_Iter/3$ then:
if $i < n/2$ then:
Update the prey according to Eq.(17)
else $i \geq n/2$
Update the prey according to Eq.(18)
end if
else if $Iter > 2 * Max\_Iter/3$ then:
Update the prey according to Golden Sine Algorithm Eq.(27)
end if
end if
end if
Accomplish memory saving and update Elite
Applying FADs effect and update based on Eq.(20)
$t=t+1$
end while

Multi-kernel support vector regression is used as the forecasting model of carbon dioxide emissions. Linear kernel, RBF kernel, polynomial kernel, and Sigmoid kernel are selected to make a weighted combination of the kernel functions, and they are used in the support vector regression model;

#### Step 3: Optimization Model

EGMPA is used to optimize the main parameters of multi-kernel support vector regression: penalty factor  $C$ , kernel function parameters  $g$ , and kernel function weight  $\omega$ . The fitness function selected in this paper is the root-mean-square error (RMSE) between the real value and the predicted value.

#### Step 4: Evaluation Criteria

MAE, RMSE, and MAPE are used as evaluation criteria to further evaluate the performance of the model.

#### Algorithm test

In this paper, 23 benchmark functions including seven unimodal, six multimodal, and ten fixed-dimension multimodal optimization problems are selected to verify the effectiveness of the proposed EGMPA. It is compared with eight classical meta-heuristics, namely, DE, PSO, CS, MVO, SCA, MFO, SSA, and GWO.



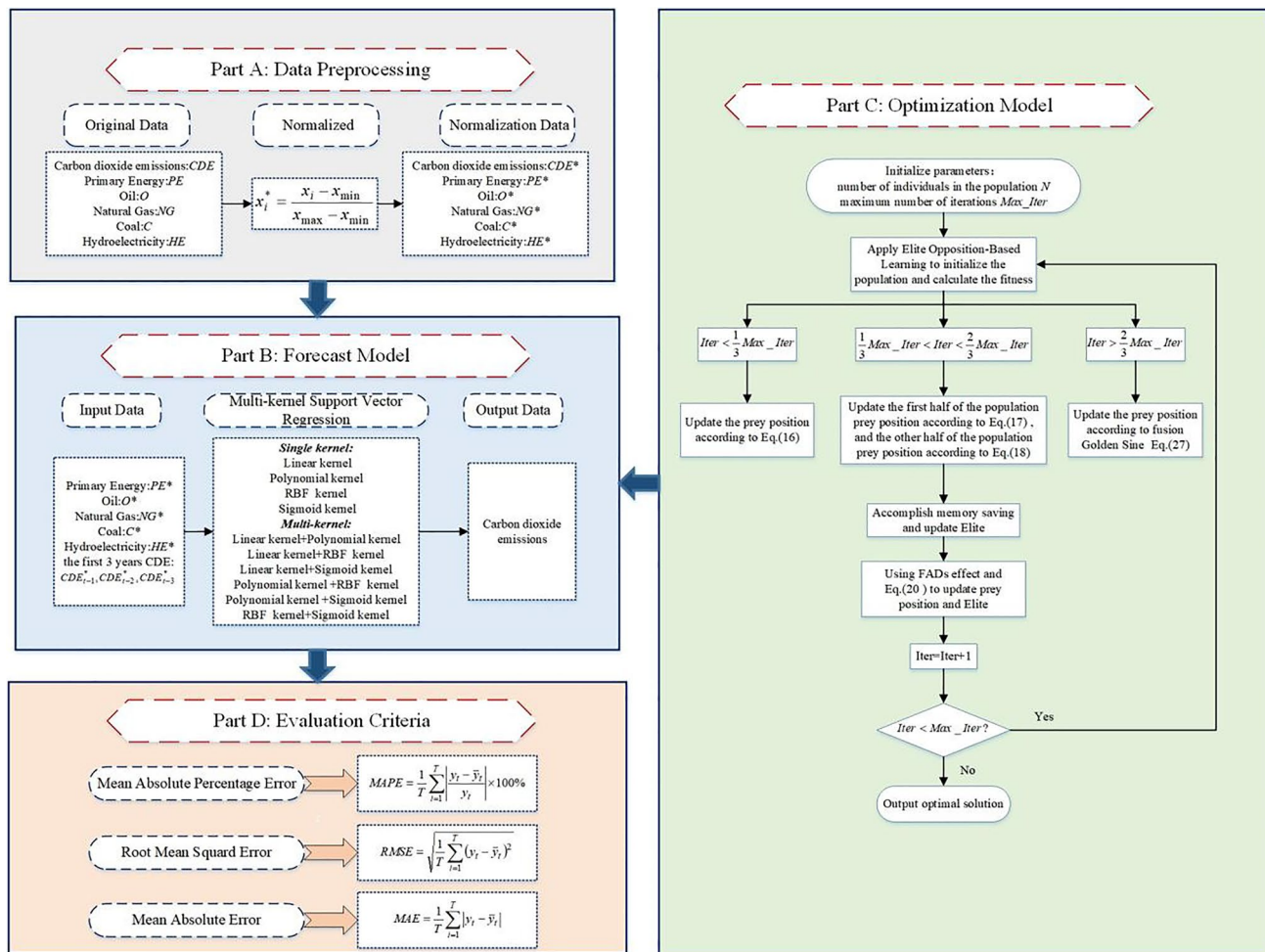


Fig. 1 Hybrid forecasting model framework

Each algorithm carries out 30 experiments to ensure fairness. In addition, the maximum iteration number is set to 500, the dimension is set to 50, and the population size is set to 30. The algorithm tests are conducted in Python3.8, under Windows 10 system with a 64-bit 2.30 GHz Intel(R) Core (TM) i7-10510U and 16.0GB of RAM.

Table 2 shows the performance of nine different meta-heuristic algorithms on 23 benchmark functions. The convergence curves of the benchmark functions are shown in Figs. 2, 3 and 4.

As can be seen from Table 2. and Figs. 2, 3 and 4:

- (1) EGMPA performs significantly better than other meta-heuristic algorithms except F6 functions. SSA has the best performance on F6 functions. Although EGMPA is not the optimal algorithm, it is second only to SSA;
- (2) For unimodal functions F1–F7, EGMPA shows strong optimization ability on other benchmark functions except F6;

- (3) For multimodal functions F8–F13, EGMPA is significantly better than other algorithms on all multimodal functions, especially on F9 and F11 functions, MPA and EGMPA can get the actual optimal solution of the function;
- (4) For fixed-dimension multimodal functions F14–F23, almost all algorithms can find the optimal solution of functions F16–F19, but the standard deviation of EGMPA is significantly smaller than other algorithms. For F14, F15, and F20–F23 functions, the optimal solution obtained by EGMPA is closer to the actual optimal solution;
- (5) It can be seen from the convergence curves, EGMPA has a steeper convergence curve and higher convergence accuracy for unimodal functions F1, F2, F4, and F7 and multimodal functions F8, F10, F11, and F12. For fixed-dimension multimodal functions F15, F21, F22, and F23, several meta-heuristic algorithms can find the actual optimal solution, but EGMPA has a faster convergence speed.

**Table 2** Algorithm test results

Function	Metric	EGMPA	MPA	DE	CS	MVO	SCA	MFO	SSA	GWO
F1	Mean	<b>1.7867E-145</b>	2.7066E-20	8.4546E-01	3.0155E+02	1.0023E+01	7.9255E+02	9.7583E+03	2.0084E-03	3.3633E-22
	std	<b>7.0419E-145</b>	3.8950E-20	3.2500E-01	7.9224E+01	2.4010E+00	7.9320E+02	9.7701E+03	2.5280E-03	3.8196E-22
F2	Mean	<b>5.7936E-72</b>	6.2523E-13	7.5913E-01	2.3719E+01	3.3927E+06	7.2038E-01	8.7301E+01	4.8397E+00	1.0633E-13
	std	<b>3.1199E-71</b>	8.9277E-13	3.1451E-01	5.3689E+00	1.8265E+07	9.3436E-01	3.4223E+01	1.7140E+00	6.8230E-14
F3	Mean	<b>1.8399E-24</b>	4.7718E-01	9.0312E+04	1.1533E+04	5.4205E+03	5.1343E+04	6.2508E+04	4.5528E+03	2.7771E+00
	std	<b>9.4844E-24</b>	1.0525E+00	1.3770E+04	2.3197E+03	1.4183E+03	1.7159E+04	2.5124E+04	2.2174E+03	4.0771E+00
F4	Mean	<b>9.8209E-63</b>	5.1033E-08	8.3397E+01	1.8973E+01	1.8252E+01	6.7999E+01	7.9866E+01	1.4940E+01	3.6271E-03
	std	<b>5.2884E-62</b>	3.7707E-08	8.4655E+00	2.6893E+00	6.9134E+00	7.7610E+00	6.0818E+00	2.2841E+00	4.1828E-03
F5	Mean	<b>4.5973E+01</b>	4.7938E+01	7.4119E+02	2.7947E+04	8.5991E+02	5.7337E+06	1.3433E+07	2.8264E+02	4.7058E+01
	std	<b>5.3324E-01</b>	6.0414E-01	3.6819E+02	1.2108E+04	8.3026E+02	4.5562E+06	2.9796E+07	2.7547E+02	8.9326E-01
F6	Mean	3.7146E-01	3.9482E+00	8.8692E-01	2.6121E+02	1.0195E+01	1.1630E+03	1.0189E+04	<b>2.2248E-03</b>	2.5106E+00
	std	2.0046E-01	1.2235E+00	5.2839E-01	6.1619E+01	2.1046E+00	1.7287E+03	1.1475E+04	<b>4.0172E-03</b>	5.8982E-01
F7	Mean	<b>4.6076E-04</b>	1.7137E-03	1.5099E-01	3.3541E-01	1.1416E-01	3.1039E+00	1.6523E+01	3.3144E-01	4.3769E-03
	std	<b>2.9301E-04</b>	9.8806E-04	7.4727E-02	9.1199E-02	4.3560E-02	3.1833E+00	1.8465E+01	1.0482E-01	1.4902E-03
F8	Mean	<b>-1.5481E+04</b>	-7.4746E+03	-7.4755E+03	-1.2447E+04	-1.2190E+04	-4.8490E+03	-1.3073E+04	-1.2160E+04	-9.1989E+03
	std	3.6289E+03	2.9293E+03	3.8707E+02	1.7953E+03	1.0374E+03	1.9374E+03	<b>3.5673E+02</b>	9.1787E+02	1.6032E+03
F9	Mean	<b>0.0000E+00</b>	0.0000E+00	3.8899E+02	2.2321E+02	2.5984E+02	9.7425E+01	2.9847E+02	6.5535E+01	2.0489E+01
	std	<b>0.0000E+00</b>	0.0000E+00	2.0378E+01	1.7445E+01	4.0060E+01	5.5109E+01	4.7724E+01	2.1507E+01	1.7071E+01
F10	Mean	<b>4.4409E-16</b>	2.3755E-11	1.7698E+00	6.9413E+00	3.5344E+00	1.7347E+01	1.9795E+01	3.4209E+00	2.7074E-12
	std	<b>0.0000E+00</b>	1.4736E-11	4.4964E+00	8.0210E-01	3.0837E+00	6.2956E+00	4.5354E-01	9.3250E-01	1.9873E-12
F11	Mean	<b>0.0000E+00</b>	0.0000E+00	6.2329E-01	3.4330E+00	1.0870E+00	9.7817E+00	6.1242E+01	8.6846E-02	8.8463E-03
	std	<b>0.0000E+00</b>	0.0000E+00	1.8535E-01	6.4961E-01	1.9196E-02	8.4520E+00	1.0048E+02	3.8631E-02	1.1933E-02
F12	Mean	<b>9.4840E-03</b>	1.7958E-01	3.3652E+03	9.1418E+00	5.7869E+00	5.8620E+04	2.5664E+07	7.4525E+00	9.1315E-02
	std	<b>3.8594E-03</b>	5.7655E-02	1.4666E+04	2.8025E+00	2.2231E+00	2.7953E+05	7.6774E+07	2.4475E+00	3.6493E-02
F13	Mean	<b>3.3079E-01</b>	2.5936E+00	4.0673E+02	3.4170E+02	9.1750E+00	3.8310E+07	2.7397E+07	6.1173E+01	2.1747E+00
	std	<b>1.1121E-01</b>	5.7720E-01	1.8719E+03	4.1040E+02	1.1606E+01	5.7514E+07	1.0227E+08	1.4848E+01	3.3663E-01
F14	Mean	<b>9.9800E-01</b>	9.9800E-01	1.0641E+00	9.9800E-01	9.9800E-01	1.9191E+00	1.5911E+00	1.0973E+00	4.6527E+00
	std	<b>1.2360E-13</b>	3.1422E-10	3.5616E-01	1.8749E-13	2.1477E-11	1.8738E+00	1.3622E+00	3.9281E-01	4.2192E+00
F15	Mean	<b>3.0749E-04</b>	3.5342E-04	1.8325E-03	5.6345E-04	3.4330E-03	1.1097E-03	9.8582E-04	1.3584E-03	3.7318E-03
	std	<b>1.9684E-09</b>	6.7203E-05	4.9642E-03	1.0752E-04	6.6445E-03	3.8741E-04	3.3392E-04	3.5390E-03	7.4396E-03
F16	Mean	<b>-1.0316E+00</b>	-1.0316E+00	-1.0316E+00	-1.0316E+00	-1.0316E+00	-1.0316E+00	-1.0316E+00	-1.0316E+00	-1.0316E+00
	std	<b>3.1402E-16</b>	2.0220E-11	4.4409E-16	2.3076E-12	5.2691E-07	8.4468E-05	4.4409E-16	2.0564E-14	2.6679E-08
F17	Mean	<b>3.9789E-01</b>	3.9789E-01	3.9789E-01	3.9789E-01	3.9789E-01	3.9989E-01	3.9789E-01	3.9789E-01	3.9789E-01
	std	<b>0.0000E+00</b>	1.4455E-13	0.0000E+00	1.4089E-09	6.9177E-07	2.4534E-03	0.0000E+00	3.3348E-14	4.4005E-06
F18	Mean	<b>3.0000E+00</b>	3.0000E+00	3.0000E+00	3.0000E+00	3.0000E+00	3.0001E+00	3.0000E+00	3.0000E+00	3.0000E+00
	std	1.6175E-15	6.6526E-12	<b>1.3099E-15</b>	1.8435E-15	4.3220E-06	1.0260E-04	1.9677E-15	3.2073E-13	5.1974E-05
F19	Mean	<b>-3.8628E+00</b>	-3.8628E+00	-3.8628E+00	-3.8628E+00	-3.8628E+00	-3.8539E+00	3.8628E+00	-3.8628E+00	-3.8624E+00
	std	<b>3.8628E+00</b>	3.8628E+00	3.8628E+00	3.8628E+00	3.8628E+00	3.8628E+00	3.8628E+00	3.8628E+00	3.8628E+00

Table 2 (continued)

Function	Metric	EGMPA	MPA	DE	CS	MVO	SCA	MFO	SSA	GWO
F20	std	<b>6.8798E-16</b>	2.3903E-09	8.8818E-16	7.0217E-16	1.4393E-06	1.4231E-03	8.8818E-16	4.6251E-14	1.0553E-03
	Mean	<b>-3.3220E+00</b>	-3.3218E+00	-3.2491E+00	-3.3220E+00	-3.2576E+00	-2.8583E+00	-3.2299E+00	-3.2035E+00	-3.2568E+00
F21	std	<b>1.9167E-10</b>	7.0964E-04	6.0035E-02	6.9561E-06	6.0287E-02	3.5882E-01	5.8009E-02	3.3018E-02	7.2151E-02
	Mean	<b>-1.0153E+01</b>	-1.0153E+01	-8.1594E+00	-1.0153E+01	-7.1299E+00	-1.7803E+00	-7.3859E+00	-7.4807E+00	-8.3080E+00
F22	std	2.1378E-05	1.2281E-03	3.3064E+00	<b>1.2682E-06</b>	3.1395E+00	1.5039E+00	3.0634E+00	3.3521E+00	2.8937E+00
	Mean	<b>-1.0403E+01</b>	-1.0403E+01	-9.8938E+00	-1.0403E+01	-8.9324E+00	-3.3037E+00	-7.7405E+00	-9.3655E+00	-9.6195E+00
F23	std	<b>8.1209E-15</b>	2.0857E-04	1.9050E+00	2.7803E-05	2.7231E+00	1.7269E+00	3.3086E+00	2.3681E+00	2.0283E+00
	Mean	<b>-1.0536E+01</b>	-1.0534E+01	-1.0090E+01	-1.0536E+01	-8.3032E+00	-4.0406E+00	-8.5956E+00	-9.5912E+00	-1.0279E+01
	std	<b>4.3025E-15</b>	7.0450E-03	1.6715E+00	2.7353E-05	3.0181E+00	1.7606E+00	3.2502E+00	2.4372E+00	1.3757E+00

Note: Bold values in the table represent the best values

In summary, the convergence speed and convergence accuracy of EGMPA are significantly better than those of the eight meta-heuristics.

## Results and discussion

### Data description

In this paper, China’s carbon dioxide emissions from 1965 to 2020 are used as the research object. According to previous studies (Sangeetha and Amudha 2018; Heydari et al. 2019; Lin et al. 2021), energy consumption is a major factor in carbon dioxide emissions, so in this paper, five important influencing factors of coal, oil, natural gas, hydroelectricity, and primary energy are selected to forecast carbon dioxide emissions (Table 3). Since historical carbon dioxide emissions affect carbon dioxide emissions, historical carbon dioxide emissions in China for the first 3 years are also used as input variables for prediction to improve the accuracy of prediction. The data comes from the “Statistical Review of World Energy 2021 | 70th edition” (<https://www.bp.com/>) released by BP in 2021. The data from 1965 to 2014 were used as the training set, and the data from 2015 to 2020 were used as the test set. The relevant descriptive statistical analysis is shown in Table 3. Due to the large variation in the size of the selected raw data, the normalization method is used in this paper to restrict the data to the range of [0,1]. Fig. 5 shows the normalized line graph of the five influencing factors and carbon dioxide emissions. It can be seen that the emissions of coal, oil, natural gas, hydroelectricity, primary energy, and carbon dioxide emissions are increasing year by year. Fig. 6 shows the annual growth rate of carbon dioxide emissions in China. It can be seen that the annual growth rate of carbon dioxide emissions shows a relatively large fluctuation, and even shows a negative growth. Since 2005, the annual growth rate of China’s carbon dioxide emissions has shown an overall downward trend, which shows that China is increasingly concerned about carbon emissions. The annual growth rate in 2019 and 2020 continued to decline, but its growth rate remained positive, indicating that COVID-19 has had an impact on China’s carbon dioxide emissions.

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{30}$$

where  $x_i$  is the original data,  $x_{\min}$  is the minimum value in  $x_i$ ,  $x_{\max}$  is the maximum value in  $x_i$ , and  $x_i^*$  is the normalized data.

### Evaluation criteria

In order to evaluate the performance of the carbon dioxide prediction model, three commonly used evaluation criteria,

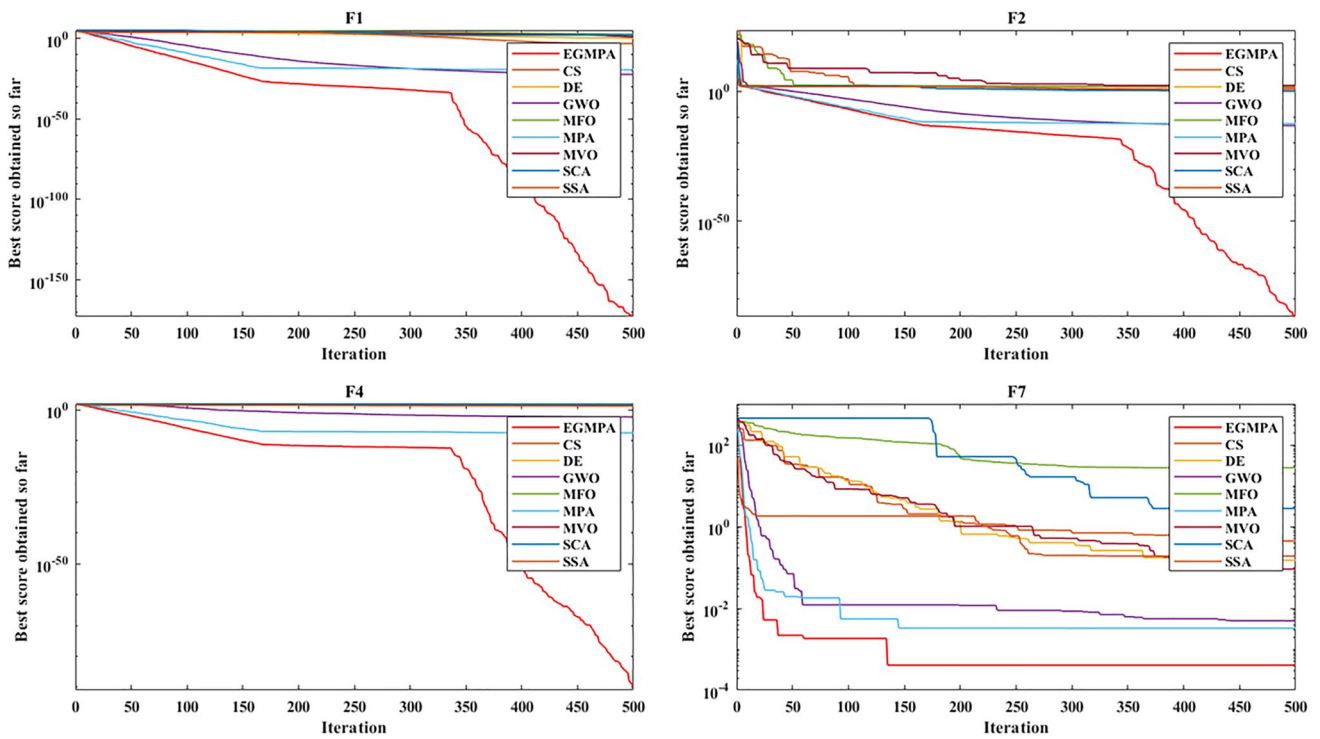


Fig. 2 Convergence curves of nine algorithms for unimodal test function (F1, F2, F4, F7)

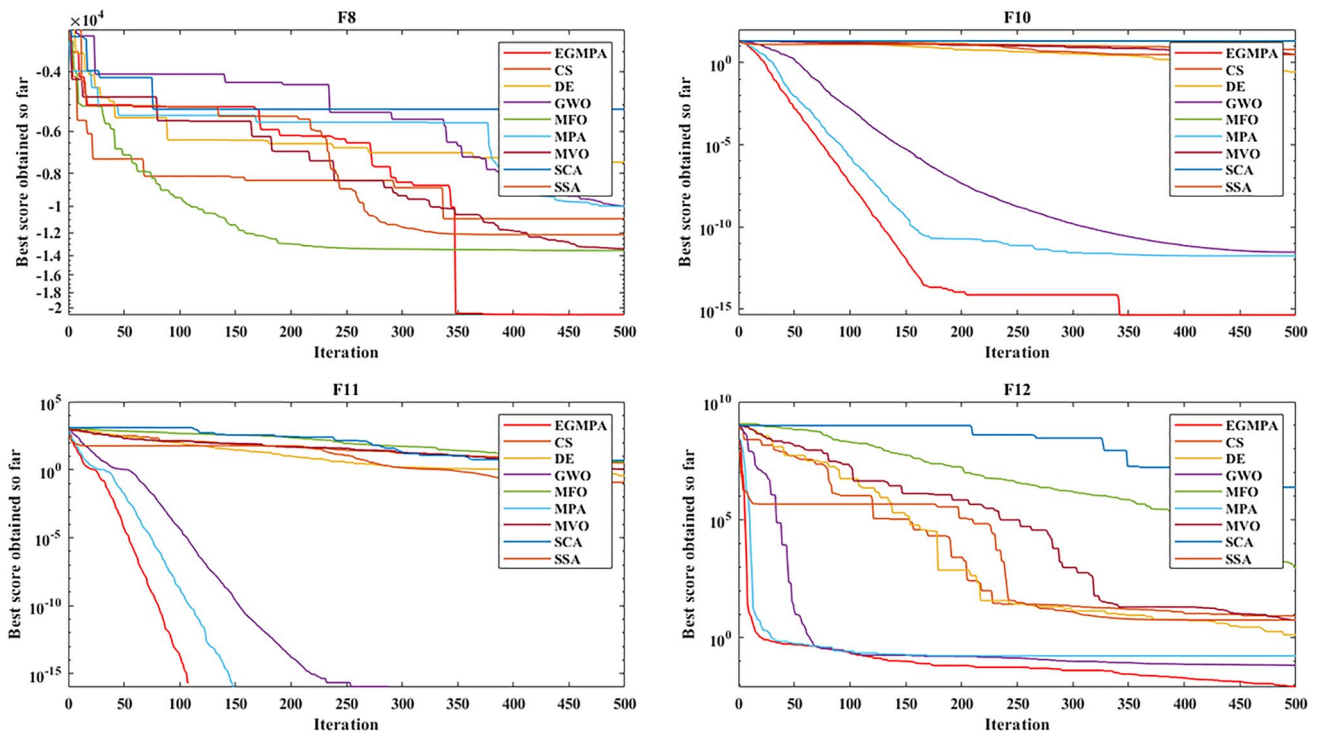


Fig. 3 Convergence curves of nine algorithms for multimodal test functions (F8, F10, F11, F12)

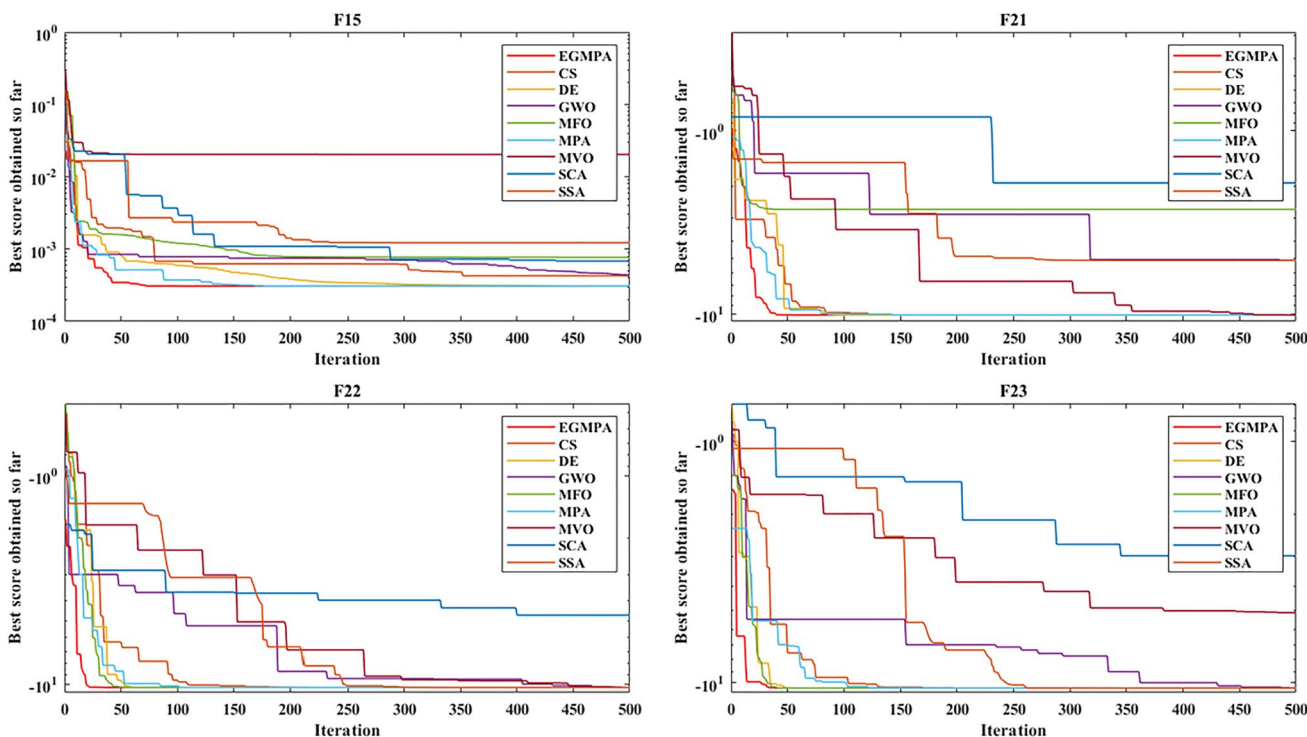


Fig. 4 Convergence curves of nine algorithms for fixed-dimension multimodal test functions (F15, F21, F22, F23)

Table 3 Descriptive statistics of input and output variables

Variable name	Unit	Min	Max	Mean	Std
Oil	Exajoules	0.46	28.5	9.51	8.37
Natural gas		0.04	11.9	2.08	3.04
Coal		4.52	82.49	34.63	27.89
Hydroelectricity		0.19	11.74	3.09	3.53
Primary energy		5.39	145.46	50.48	44.31
CO <sub>2</sub> emissions	Mt	475.9	9893.5	3896.16	3200.44

including MAE, RMSE, and MAPE, are used, and the calculation formulas are as follows:

$$MAE = \frac{1}{T} \sum_{i=1}^T |y_i - \bar{y}_i| \tag{31}$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (y_i - \bar{y}_i)^2} \tag{32}$$

$$MAPE = 100 \times \frac{1}{T} \sum_{i=1}^T \left| \frac{y_i - \bar{y}_i}{y_i} \right| \tag{33}$$

where  $T$  represents the number of samples,  $y_i$  represents the real value of data at time  $t$ , and  $\bar{y}_i$  represents the predicted value of the model at time  $t$ .

### Forecasting results comparison

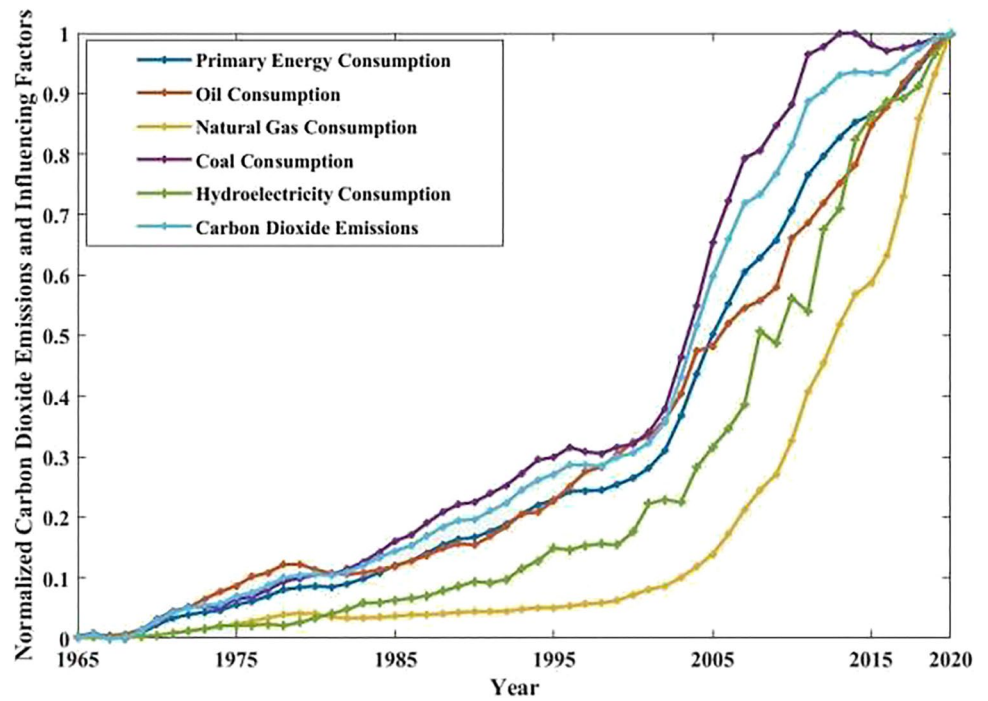
#### Multi-kernel support vector regression comparison

This paper adopts the support vector regression model suitable for small sample modeling as the prediction model of carbon dioxide emissions. Meanwhile, four different kernel functions, linear kernel, polynomial kernel, RBF kernel, and Sigmoid kernel, are weighted and combined to construct the multi-kernel support vector model. The main parameters of multi-kernel support vector regression including penalty factor  $C$ , kernel function parameters  $g$ , and kernel function weight  $\omega$  were optimized by EGMPA. The maximum number of iterations of the EGMPA algorithm is set to 100, and the number of populations is set to 30. The optimization range of penalty factor  $C$  and kernel function parameters  $g$  is set to [0.01,100], and kernel function weight  $\omega$  are set to [0.001,0.999]. Table 4 shows the error analysis of different multi-kernel functions on the test set, and Fig. 6 shows the prediction results of support vector regression for different multi-kernel functions on the test set.

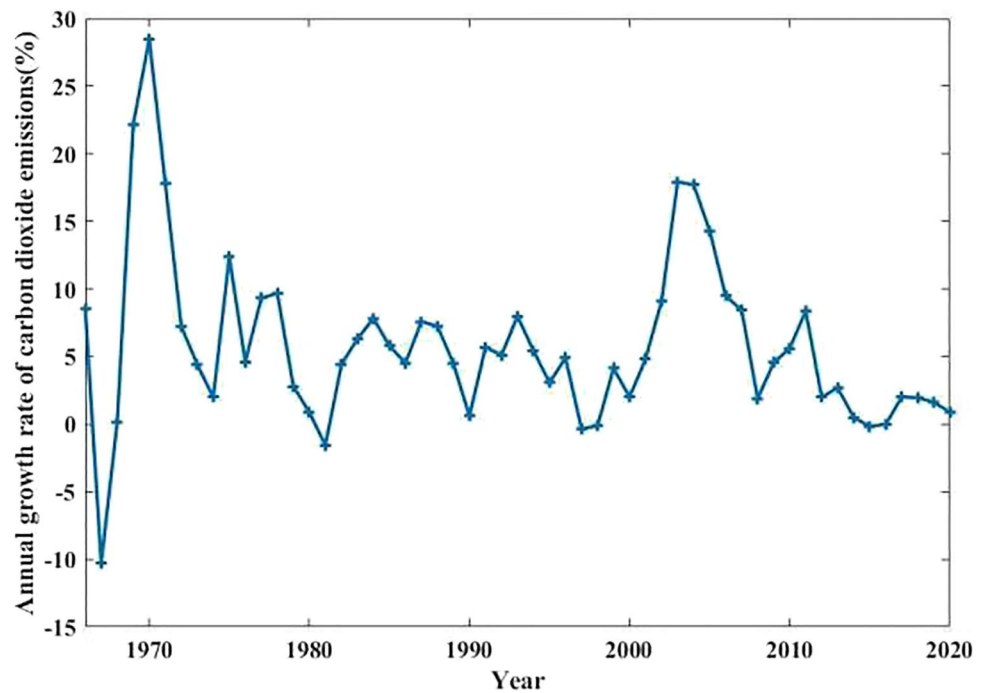
It can be seen from Table 4 and Fig. 7.

RMSE, MAE, and MAPE of EGMPA optimized single kernel support vector regression are significantly smaller

**Fig. 5** Normalized carbon dioxide emissions and influencing factors



**Fig. 6** Annual growth rate of carbon dioxide emissions



than those of the unoptimized model, and linear kernel has the best performance in single kernel support vector regression. The error of multi-kernel support vector regression is obviously smaller than that of single kernel support vector regression, which indicates that multi-kernel support vector regression can obtain better learning effect. Among the 10 models, the best model is EGMPA-RBF-Sigmoid, whose

RMSE is 37.43, MAE is 30.63, and MAPE is 0.32, which is obviously better than other models.

This is because the RBF kernel is a typical local kernel with strong learning ability but weak generalization ability, while the Sigmoid kernel is a global kernel with weak learning ability but strong generalization ability. The weighted combination of the two kernel functions

**Table 4** Error analysis results of different multi-kernel support vector regression models

Kernel function	RMSE	MAE	MAPE (%)
Linear	93.88	79.59	0.84
RBF	878.03	859.19	8.95
Poly	751.88	728.42	7.58
Sigmoid	1697.99	1692.36	17.68
EGMPA-Linear	93.88	79.59	0.84
EGMPA-RBF	135.64	121.22	1.28
EGMPA-Poly	97.98	79.22	0.83
EGMPA-Sigmoid	130.91	115.84	1.22
EGMPA-Linear-Poly	89.92	74.48	0.78
EGMPA-Linear-RBF	93.88	79.59	0.84
EGMPA-Linear-Sigmoid	87.78	72.26	0.76
EGMPA-Poly-RBF	57.27	46.75	0.49
EGMPA-Poly-Sigmoid	89.95	74.48	0.78
EGMPA-RBF-Sigmoid	<b>37.43</b>	<b>30.63</b>	<b>0.32</b>

Note: Bold values in the table represent the best values

can meet the requirements of strong learning ability and strong generalization ability at the same time, and a better learning effect can be obtained. The prediction results of different multi-kernel support vector regression on the test data show that the EGMPA-RBF-Sigmoid model and the

EGMPA-Poly-RBF model are closer to the true value than the other models, and the difference with the true value is smaller.

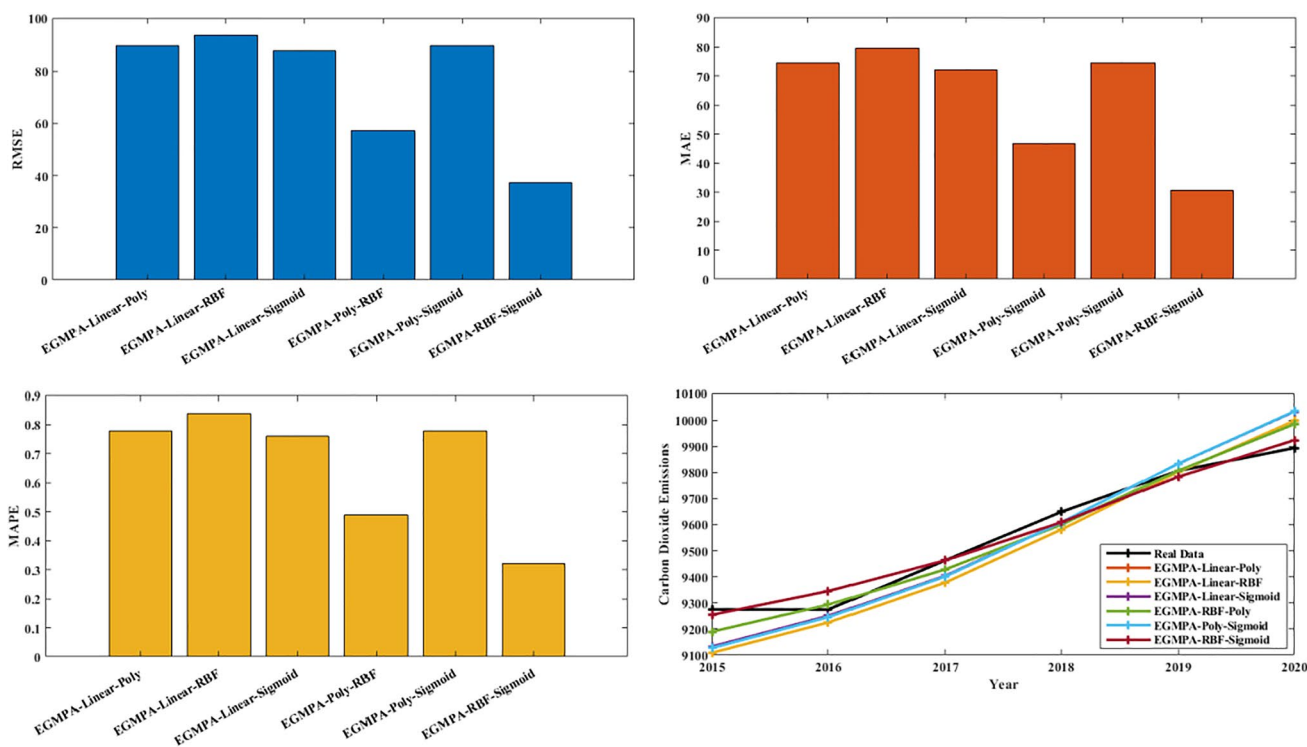
In summary, the optimized support vector regression of EGMPA is better than the unoptimized model. The error result of multi-kernel support vector regression is significantly smaller than that of single-kernel support vector regression. In multi-kernel support vector regression, RBF-Sigmoid-EGMPA has the best learning ability and generalization ability.

### Different model comparison

The prediction model is compared with BPNN, LSTM, RNN, and GRU models and used to highlight the advantages of the proposed model in predicting carbon dioxide emissions. Table 5 shows the error analysis of different models on the test set, and Fig. 8 shows the prediction results of different models on the test set.

It can be seen from Table 5 and Fig. 8:

RMSE, MAE, and MAPE of the EGMPA-RBF-Sigmoid model were 37.43, 30.63, and 0.32, respectively, which were significantly lower than those of other comparable models. According to the evaluation criteria, the prediction accuracy of the model is ranked from high to low: EGMPA-RBF-Sigmoid > BPNN > RNN > LSTM > GRU. It can be seen from the predicted values of China’s carbon dioxide emissions



**Fig. 7** Error analysis results of different multi-kernel support vector regression models

**Table 5** Error analysis results of different models

Model	RMSE	MAE	MAPE (%)
EGMPA-RBF-Sigmoid	<b>37.43</b>	<b>30.63</b>	<b>0.32</b>
BPNN	212.34	187.13	1.93
LSTM	519.08	515.24	5.41
RNN	433.72	372.34	3.85
GRU	553.07	549.36	5.76

Note: Bold values in the table represent the best values

by different models that the predicted values of EGMPA-RBF-Sigmoid are closest to the real predicted values, the predicted results of GRU and BPNN are significantly higher than the real predicted values, while the predicted results of LSTM and RNN are lower than the real predicted values.

In summary, the prediction accuracy of EGMPA-RBF-Sigmoid on carbon dioxide emissions in China is significantly higher than that of BPNN, RNN, LSTM, and GRU models, which further demonstrates the superiority of the proposed model.

#### *China's carbon dioxide emissions forecast during the "14th Five-Year Plan" period*

In this paper, the EGMPA-RBF-Sigmoid model is used to forecast China's carbon dioxide emissions during the "14th Five-Year Plan" period (2021–2025). According to the "Ten Year Outlook for China's Energy Revolution (2021–2030)," it is expected that during the "14th Five-Year Plan" period, the growth rate of primary energy demand will decline steadily and slightly, with an average annual growth rate of about 2.5%. In 2025, the total energy consumption will exceed 5.5 billion tons of standard coal; after hard work, the total proportion of clean energy such as non-fossil energy and natural gas will exceed 30%; the oil demand will be about 700 million tons, accounting for nearly 18%, and the proportion of coal will decrease. To less than 50%, according to the above policies, the predicted values of influencing factors of China's carbon dioxide emissions during the "14th Five-Year Plan" period are determined, as shown in Table 6.

According to the predicted values of the influencing factors, the EGMPA-RBF-Sigmoid model can be used to predict the carbon dioxide emissions in China during the "14th Five-Year Plan" period. It can be seen from the Fig. 9.

The predicted carbon dioxide emission in 2021 is 9949.14 Mt, and then increases year by year. The predicted carbon dioxide emissions in 2025 will reach 10,258.95 Mt. Although China's carbon dioxide emissions have gradually increased, the increase has gradually decreased, indicating that China is getting closer to its goal of "Carbon Peak." The above predicted results are consistent with the studies of Niu et al. (2020) and Qiao et al. (2019). Niu et al. predicted that

**Table 6** Predicted values of influencing factors

Influencing factors	2021	2022	2023	2024	2025
Primary energy	149.10	152.82	156.64	160.56	164.57
Coal	74.55	76.41	78.32	80.28	82.29
Natural gas	22.36	22.92	23.50	24.08	24.69
Oil	26.84	27.51	28.20	28.90	29.62
Hydroelectricity	22.36	22.92	23.50	24.08	24.69

China's carbon dioxide emissions will continue to increase from 2020 to 2025 under the basic as usual, policy tightening and market allocation scenarios. Qiao et al. predict that China's carbon dioxide emissions will reach 11521.21 MT in 2025.

The "14th Five-Year Plan" is a critical period and a window period for the "Carbon Peak" target. In order to achieve the "Carbon Peak" by 2030 and the "Carbon Neutrality" by 2060, the Chinese government has taken a series of measures in different fields to implement energy Green and low-carbon transformation action, industrial carbon peaking action, transportation green and low-carbon action, green and low-carbon technological innovation action, etc. China is also continuing to promote the energy supply revolution, the energy technology revolution and the energy system revolution, and intensify efforts to develop green energy.

Carbon dioxide emissions are closely related to economic development. Accurately predicting China's carbon dioxide emissions can provide an important basis for future economic development. Green and circular development will be one of China's important economic growth drivers in the future, as well as an important direction for developing emerging industries. At the same time, the accurate prediction of carbon dioxide emissions also has certain guiding significance for China's future low-carbon emission reduction and economic sustainable development.

In summary, during the "14th Five-Year Plan" period, China's carbon dioxide emissions will continue to grow, but the growth rate will slow down significantly, and China is approaching the goal of "Carbon Peak."

## Conclusions

In this paper, MPA combining elite opposition-based learning strategy and golden sine algorithm is proposed to optimize the parameters of the multi-kernel support vector regression model, and it is used to predict carbon dioxide emissions in China. The following conclusions can be drawn from the comparison between the algorithm test and the carbon dioxide emissions prediction results in China:



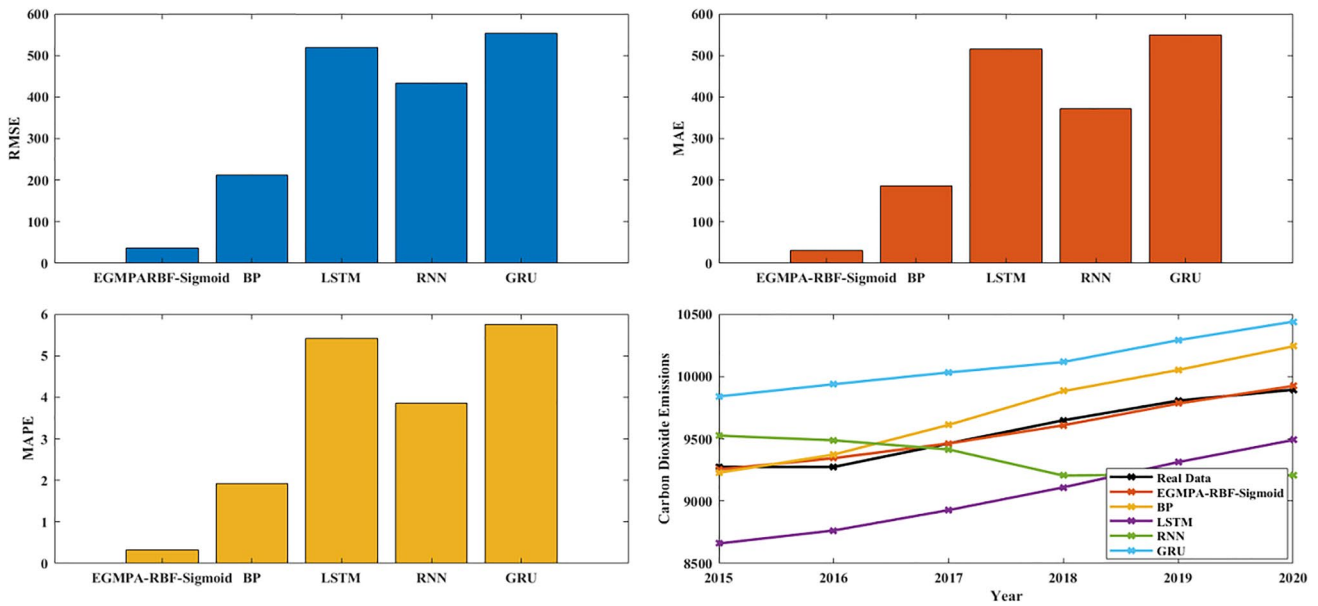


Fig. 8 Error analysis results of different models

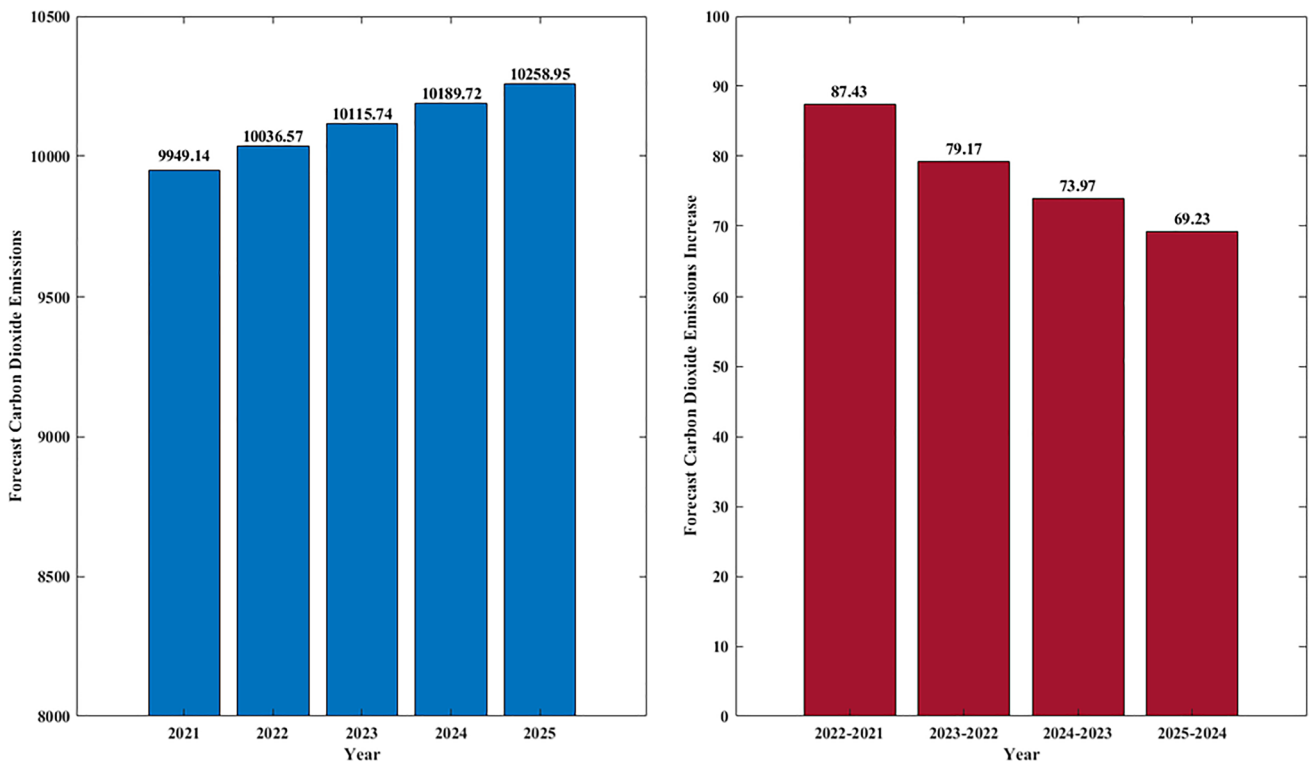


Fig. 9 Forecast results of China's carbon dioxide emissions during the “14th Five-Year Plan” period

- (1) Compared with the other eight meta-heuristic algorithms, EGMPA, which combines elite opposition-based learning strategy and golden sine algorithm, has stronger global search ability, faster convergence speed and higher accuracy;
- (2) Based on the prediction results of multi-kernel vector regression models with different kernel functions, it can be seen that RMSE, MAE, and MAPE of single-kernel support vector regression after EGMPA

optimization are significantly smaller than those of the unoptimized model, and the error of multi-kernel support vector regression model is significantly smaller than that of single-kernel support vector regression model. It shows that the multi-kernel support vector regression model can obtain a better learning effect;

- (3) Among the 10 different kernel support vector regression models, the EGMPA-RBF-Sigmoid model has the best performance, which is because RBF kernel is a classical local kernel, while Sigmoid kernel is a global kernel. The weighted combination of the two kernels can simultaneously meet the requirements of strong learning ability and strong generalization ability of kernel function, get better learning results;
- (4) The EGMPA-RBF-Sigmoid model is superior to BPNN, LSTM, RNN, and GRU in RMSE, MAE, and MAPE, which further demonstrates the accuracy and effectiveness of the proposed model in carbon dioxide emissions prediction;
- (5) The EGMPA-RBF-Sigmoid model, which has the highest prediction accuracy, is used to forecast China's carbon dioxide emissions during the "14th Five-Year Plan" period. According to the prediction results, China's carbon dioxide emissions will continue to increase during the "14th Five-year Plan" period, but the growth rate will significantly slow down, indicating that China is gradually approaching the goal of "Carbon Peak."

In the future, the carbon dioxide emission prediction model constructed in this paper can be used in more countries and provide a forecasting tool for national policy makers and relevant researchers. At the same time, more factors, such as economic factors, social factors, and energy prices, can be added in future studies to explore the dynamic changes of carbon dioxide emissions.

**Author contribution** All authors contributed to the study conception and design. X. Q. and S. Z. conceived and designed the study; X. D., R. W., and Y. Z. framed the methodology; S. Z., R. W., and Y. Z. performed the modeling experiments; S. Z., X. Q., and X. D. analyzed the results; S. Z., D. X., and X. Q. wrote the original draft of the article. All authors reviewed and edited the article. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** This work was supported by Education Department of Jilin Province project (JJKH20210716KJ); National Natural Science Foundation of China (12026430); Department of Science and Technology of Jilin Province project (20210101149JC, 20200403182SF).

**Data availability** The datasets analyzed during the current study are available in the "Statistical Review of World Energy 2021/70th edition" repository, <https://www.bp.com/>.

## Declarations

**Ethical approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

## References

- Abdel BM, Mohamed R, Chakraborty RK, Ryan M, Mirjalili S (2020) New binary marine predators optimization algorithms for 0–1 knapsack problems. *Comput Ind Eng* 151:106949
- Al-Qaness MAA, Ewees AA, Fan H et al (2020) Marine predators algorithm for forecasting confirmed cases of COVID-19 in Italy, USA, Iran and Korea. *Int J Environ Res Public Health* 17(10):3520
- Aly M, Ahmed EM, Rezk H et al (2021) Marine predators algorithm optimized reduced sensor fuzzy-logic based maximum power point tracking of fuel cell-battery standalone applications. *IEEE Access* 9:27987–28000
- Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297
- Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: *Proceedings of the Sixth International Symposium on Micro Machine and Human Science MHS*, 95th edn. IEEE, Nagoya, pp 39–43
- Elaziz MA, Mohammadi D, Oliva D, Salimifard K (2021) Quantum marine predators algorithm for addressing multilevel image segmentation. *Appl Soft Comput* 110:107598
- Faramarzi A, Heidarnejad M, Mirjalili S et al (2020) Marine predators algorithm: a nature-inspired metaheuristic. *Expert Syst Appl* 152:113377
- Gao M, Yang H, Xiao Q, Goh M et al (2020) A novel fractional grey Riccati model for carbon emission prediction. *J Clean Prod* 282:124471
- Hansen N, Müller SD, Koumoutsakos P (2003) Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol Comput* 11(1):1–18
- Heydari A, Garcia DA, Keynia F et al (2019) Renewable energies generation and carbon dioxide emission forecasting in microgrids and national grids using GRNN-GWO methodology. *Energy Procedia* 159:154–159
- Holland JH (1975) *Adaptation in natural and artificial systems*. University of Michigan Press, Michigan
- Houssein EH, Hussain K, Abualigah L et al (2021a) An improved opposition-based marine predators algorithm for global optimization and multilevel thresholding image segmentation. *Knowl-Based Syst* 229:107348
- Houssein EH, Mahdy MA, Fathy A et al (2021b) A modified marine predator algorithm based on opposition based learning for tracking the global MPP of shaded PV system. *Expert Syst Appl* 183:115253
- Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J Glob Optim* 39(3):459–471
- Li H, Qin Q (2019) Challenges for China's carbon emissions peaking in 2030: A decomposition and decoupling analysis. *J Clean Prod* 207:857–865
- Liang YC, Sun YF (2003) An improved method of support vector machine and its applications to financial time series forecasting. *Prog Nat Sci* 13:696–700

- Lin X, Zhu X, Feng M et al (2021) Economy and carbon emissions optimization of different countries or areas in the world using an improved attention mechanism based long short term memory neural network. *Sci Total Environ* 792:148444
- Mardani A, Liao H, Nilashi M, Alrasheedi M, Cavallaro F (2020) A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques. *J Clean Prod* 275:122942
- Mirjalili S (2016) SCA: a sine cosine algorithm for solving optimization problems. *Knowl-Based Syst* 96:120–133
- Mirjalili S, Lewis A (2016) The whale optimization algorithm. *Adv Eng Softw* 95:51–67
- Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. *Adv Eng Softw* 69(Supplement C):46–61
- Mirjalili S, Mirjalili SM, Hatamlou A (2016) Multi-verse optimizer: a nature-inspired algorithm for global optimization. *Neural Comput & Applic* 27(2):495–513
- Mirjalili S, Gandomi AH, Mirjalili SZ, Saremi S, Faris H, Mirjalili SM (2017) Salp swarm algorithm: a bio-inspired optimizer for engineering design problems. *Adv Eng Softw* 114:163–191
- Niu D, Wang K, Wu J et al (2020) Can China achieve its 2030 carbon emissions commitment? Scenario analysis based on an improved general regression neural network. *J Clean Prod* 243(C):118558
- Oszust M (2021) Enhanced marine predators algorithm with local escaping operator for global optimization. *Knowl-Based Syst* 232:107467
- Qiao W, Lu H, Zhou G, Azimi M et al (2020) A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer. *J Clean Prod* 244:118612
- Rashedi E, Nezamabadi-Pour H, Saryazdi S (2009) GSA: a gravitational search algorithm. *Inf Sci* 179(13):2232–2248
- Ropkins K, Tate JE (2021) Early observations on the impact of the COVID-19 lockdown on air quality trends across the UK. *Sci Total Environ* 754:142374
- Sangeetha A, Amudha T (2018) A novel bio-inspired framework for CO<sub>2</sub> emission forecast in India. *Procedia Comp Sci* 125:367–375
- Shakun JD, Clark PU, He F, Marcott SA, Mix AC et al (2012) Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* 484(7392):49–54
- Shehzad K, Xiaoxing L, Ahmad M et al (2021) Does air pollution upsurge in megacities after Covid-19 lockdown? A spatial approach. *Environ Res* 197:111052
- Simon D (2008) Biogeography-based optimization. *IEEE Trans Evol Comput* 12(6):702–713
- Song C, Zhao T, Wang J (2019) Spatial-temporal analysis of China's regional carbon intensity based on ST-IDA from 2000 to 2015. *J Clean Prod* 238:117874
- Song C, Zhao T, Xiao Y (2022a) Temporal dynamics and spatial differences of household carbon emissions per capita of China's provinces during 2000–2019. *Environ Sci Pollut Res* 29:31198–31216
- Song C, Zhao T, Wang J (2022b) Analyzing driving forces of China's carbon emissions from 1997 to 2040 and the potential emission reduction path: through decomposition and scenario analysis. *Clean Techn Environ Policy* 24(4):1219–1240
- Storn R, Price K (1997) Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. *J Glob Optim* 11(4):341–359
- Sun W, Sun J (2017) Prediction of carbon dioxide emissions based on principal component analysis with regularized extreme learning machine: the case of China. *Environ Eng Res* 22(3):302–311
- Tanyildizi E, Demir G (2017) Golden sine algorithm: a novel math-inspired algorithm. *Adv Electr Comp Eng* 17(2):71–78
- Wang ZX, Ye DJ (2017) Forecasting Chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models. *J Clean Prod* 142:600–612
- Wang J, Song C, Yuan R (2021a) CO<sub>2</sub> emissions from electricity generation in China during 1997–2040: The roles of energy transition and thermal power generation efficiency. *Sci Total Environ* 773:145026
- Wang N, Wang JS, Zhu LF et al (2021b) A novel dynamic clustering method by integrating marine predators algorithm and particle swarm optimization algorithm. *IEEE Access* 9:3557–3569
- Wang Z, Yao L, Chen G et al (2021c) Modified multiscale weighted permutation entropy and optimized support vector machine method for rolling bearing fault diagnosis with complex signals. *ISA Trans* 114(12):470–484
- Wen L, Yuan X (2020) Forecasting CO<sub>2</sub> emissions in China's commercial department, through BP neural network based on random forest and PSO. *Sci Total Environ* 718:137194.1–137194.14
- Xie W, Wang JS, Tao Y (2019) Improved black hole algorithm based on golden sine operator and Levy flight operator. *IEEE Access* 9:161459–161486
- Yakout A, Sabry W, Hasanien HM (2021) Enhancing rotor angle stability of power systems using marine predator algorithm based cascaded PID control. *Ain Shams Eng J* 3:1849–1857
- Yang W, Xia K, Li T, Xie M, Song F (2021) A multi-strategy marine predator algorithm and its application in joint regularization semi-supervised ELM. *Mathematics* 9(3):291
- You Z, Zhao T, Song C et al (2020) Analyzing China's coal-related carbon emissions from economic growth perspective: through decoupling and decomposition model. *Environ Sci Pollut Res* 28(3):3703–3718
- Zhang L, Jiang Z, Liu R et al (2018) Can China achieve its CO<sub>2</sub> emission mitigation target in 2030: a system dynamics perspective. *Pol J Environ Stud* 27(6):2861–2871
- Zhong K, Luo Q, Zhou Y et al (2020) TLMPA: Teaching-learning-based Marine Predators algorithm. *AIMS Math* 6(2):1395–1442
- Zhou J, Guang F, Tang R (2018) Scenario analysis of carbon emissions of China's power industry based on the improved particle swarm optimization-support vector machine model. *Pol J Environ Stud* 27(1):439–449

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.