

Article

Carbon Trading Price Prediction of Three Carbon Trading Markets in China Based on a Hybrid Model Combining CEEMDAN, SE, ISSA, and MKELM

Haoran Zhao ¹ and Sen Guo ^{2,*}

¹ School of Economics and Management, Beijing Information Science & Technology University, Beijing 100085, China; 20202461@bistu.edu.cn

² School of Economics and Management, North China Electric Power University, Beijing 102206, China

* Correspondence: guosen@ncepu.edu.cn; Tel.: +86-15811424568

Abstract: Carbon trading has been deemed as the most effective mechanism to mitigate carbon emissions. However, during carbon trading market operation, competition among market participants will inevitably occur; hence, the precise forecasting of the carbon trading price (CTP) has become a significant element in the formulation of competition strategies. This investigation has established a hybrid CTP forecasting framework combining complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), sample entropy (SE) method, improved salp swarm algorithm (ISSA), and multi-kernel extreme learning machine (MKELM) methods to improve forecasting accuracy. Firstly, the initial CTP data sequence is disintegrated into several intrinsic mode functions (IMFs) and a residual sequence by a CEEMDAN method. Secondly, to save calculation time, SE method has been utilized to reconstruct the IMFs and the residual sequence into new IMFs. Thirdly, the new IMFs are fed into the MKELM model, combining RBF and the poly kernel functions to utilize their superior learning and generalization abilities. The parameters of the MKELM model are optimized by ISSA, combining dynamic inertia weight and chaotic local searching method into the SSA to enhance the searching speed, convergence precision, as well as the global searching ability. CTP data in Guangdong, Shanghai, and Hubei are selected to prove the validity of the established CEEMDAN-SE-ISSA-MKELM model. Through a comparison analysis, the established CEEMDAN-SE-ISSA-MKELM model performs the best with the smallest MAPE and RMSE values and the highest R² value, which are 0.76%, 0.53, and 0.99, respectively, for Guangdong,. Thus, the presented model would be extensively applied in CTP forecasting in the future.

Keywords: carbon trading price prediction; CEEMDAN; sample entropy method; improved SSA; MKELM combining RBF kernel and poly kernel functions

MSC: 60G25



Citation: Zhao, H.; Guo, S. Carbon Trading Price Prediction of Three Carbon Trading Markets in China Based on a Hybrid Model Combining CEEMDAN, SE, ISSA, and MKELM. *Mathematics* **2023**, *11*, 2319. <https://doi.org/10.3390/math11102319>

Academic Editor: Steve Drekic

Received: 31 March 2023

Revised: 10 May 2023

Accepted: 11 May 2023

Published: 16 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the beginning of the 21st century, global warming has gradually become one of the major issues terrorizing humanity. Greenhouse gases (GHG) are critical factors contributing to global warming [1]. Excessive emissions of GHG have brought about many problems, such as glacier melting, sea level rising, species extinction, significant crop production decline, and the frequent occurrence of extreme weather [2]. It can therefore be seen that the harm caused by global warming influences every aspect of human society. Thus, it is urgent and necessary to control greenhouse gas emissions, especially carbon emissions [3].

The Kyoto protocol, which came into force in 2005, aimed at stabilizing the atmosphere's GHG content at a suitable level to combat climate change. In the same year, the Emission Trading Scheme (ETS) was issued by the European Union (EU) [4]. As the

largest carbon dioxide emitter, China should take responsibility for curbing the tension of global warming. At the 75th United Nations General Assembly, China made a solemn commitment that the peak of carbon emissions will be achieved by 2030 and the carbon neutrality will be realized by 2060 [5]. Carbon trading has been verified to be a significant way to propel carbon emission reduction [6]. Since 2011, China has successfully conducted eight carbon trading projects in Beijing, Shanghai, Guangdong, Tianjin, Shenzhen, Fujian, Chongqing, and Hubei, which contribute to reducing greenhouse gas emissions. During the operation of carbon trading markets, the competition among market participants will inevitably occur. Therefore, the precise forecasting of CTP is significant in the formulation of competition strategies. Governments and policy makers can draw up scientific policies and reasonable trading mechanisms in accordance with the changes in CTP so as to promote the sustainable development of carbon trading markets [7]. Enterprises need to assign an appropriate trading proportion of carbon emissions to participate in various market transactions based on CTP prediction results. Accurate CTP prediction can maximize revenue and decrease business risks. Therefore, it is particularly significant to increase the prediction accuracy of CTP.

Various methodologies have been employed to improve CTP prediction accuracy, which can be generally categorized into two kinds: econometric methods and artificial intelligence models. The econometric methods are simple, convenient, and fast, primarily containing the linear regression model, the autoregressive integrated moving average (ARIMA), the vector autoregression (VAR), the generalized autoregressive conditional heteroskedasticity (GARCH), etc. [4,8,9]. Lin and Peng established a new CTP prediction model, which conducted the linear and nonlinear forecasting, employing ARIMA and stochastic forest algorithm to the disintegrated CTP data [10]. Fu proposed the ARMA-EGARCH-SGED framework to estimate the impact elements of fluctuations, and the VAR method was employed to discuss results [11].

The artificial intelligence models are superior in self-learning, generalization, associative memory, and calculation speed than the econometric methods, which are more adaptable for solving the nonlinear nexus of CTP data sequences. The commonly used artificial intelligence models in CTP prediction include the backpropagation artificial neural network (BPANN) [2], the radial basis function neural network (RBFNN) [12], the multilayer perceptron network (MLP) [13], the extreme learning machine (ELM) (Seifert et al., 2007; Sun and Zhang, 2018), least-squares support vector machine (LSSVM) [14], long short-term memory (LSTM) [2], etc. These methods are also widely employed in climate change studies [15], crop sciences [16,17], GHGs studies [18], and renewable energy generation prediction studies [19].

By summarizing the previous literature related to prediction research, we found that the ELM method is widely employed in this field. The ELM model is proposed on the basis of the feedforward neuron network, which is adaptable to supervised and unsupervised learning issues. At the training stage, we can obtain the only optimal solution by supposing the amount of hidden layer neurons [20,21]. Compared with the BPANN, the ELM model is superior in learning speed and generalization. While the ELM feature mapping is statistical and sensitive to noise, the stability and generalization ability of the ELM are decreased. Considering this, the MKELM is put forward by leading into the multi-kernel approach and regularization coefficient into the ELM model, the impact of stochastic mapping on the ELM are then effectively reduced, and the forecasting precision can be enhanced [22]. Simultaneously, when employing the MKELM model, the parameter values in the model need to be determined. Setting the values of these parameters by subjective judgment always leads to low forecasting precision. Hence, this investigation employs a new bionic intelligent optimization algorithm, namely the SSA, to optimally select the parameters in the MKELM. To enhance the convergence ability of SSA, an ISSA is established in our research. ISSA brings the dynamic inertia weight and chaotic local searching method into the SSA to increase the searching speed and avert the plunging into the locally optimization.

Additionally, with the continuous exploration in CTP prediction, researchers found that the high complexity and nonlinear characteristics of CTP data sequences can greatly affect the prediction accuracy. Therefore, a few data decomposition methods are employed in preprocessing the initial CTP data sequence. The empirical mode decomposition (EMD) method was verified to be a valid method to decompose the initial CTP data with nonlinear and strong volatility features into several stable series, which can greatly improve the forecasting accuracy. Zhu preprocessed the initial CTP time sequence by EMD before forecasting, and the empirical results demonstrated that EMD could integrate the nonlinear and fluctuated initial CTP data into subsequences with high stability to improve the forecasting precision [23]. Nevertheless, although EMD can effectively enhance the forecasting precision, it can still be greatly improved. Variational mode decomposition (VMD), ensemble empirical mode decomposition (EEMD), and CEEMDAN are the improved methods based on EMD to further enhance the prediction precision.

On the basis of the previous research, this paper established a hybrid system integrating the CEEMDAN, ISSA, and MKELM model to predict CTP. Firstly, the initial CTP data sequence is decomposed and constructed into several series by CEEMDAN and the sample entropy (SE) method. Then, the decomposed series are, respectively, fed into the MKELM model optimized by ISSA. At last, the CTP forecasting results are obtained by summing up all the decomposed series forecasting results. The primary contributions of our investigation are as below:

- (1) Considering the complexity and fluctuation of the initial CTP data, the CEEMDAN and SE methods are applied to decompose and reconstruct the initial data in this investigation.
- (2) In order to handle the issues of local optimization and low convergence speed, the ISSA is put forward by bringing the inertia weight mechanism and chaotic local searching method into the SSA.
- (3) The MKELM is established by constituting the RBF and the poly kernel functions so as to combine the outstanding learning and the generalization abilities of them.
- (4) The established hybrid CTP prediction model CEEMDAN-SE-ISSA-MKELM can significantly improve the CTP forecasting accuracy compared with 16 selected comparison models.

The rest of our investigation is outlined below. The methods combined in this research, including CEEMDAN, SE, ISSA, and MKELM, are introduced in Section 2. Section 3 summarizes the process of the proposed CEEMDAN-SE-ISSA-MKELM model in CTP prediction. Results are obtained in Section 4. A comparison analysis is presented in Section 5 through comparing the prediction performance of 16 selected comparison models, and conclusions are discussed in Section 6.

2. Methodology

The proposed hybrid CTP prediction model consists of CEEMDAN, SE, ISSA, and MKELM methods. The principles of these methods are introduced in this section.

2.1. CEEMDAN and SE Method

In 1998, Huang et al. first established the EMD method (EMD), which is a self-adaptive orthogonal basis time-frequency signal processing method [24]. Aiming at preprocessing nonlinear and nonstable time series, EMD can directly resolve such data sequences into multiple IMFs. However, once there exists the intermittent signal in the actual signal, a modal aliasing phenomenon occurs. Specifically, there either exists multiple scale elements in one IMF, or one scale element exists in plural IMFs simultaneously [25]. To manage this problem, the EEMD method introducing the white Gaussian noise into the signal has been proposed [26].

Even if the EEMD method can greatly decrease the emergence of mode mixing, it also leads to new problems. The EEMD with the white Gaussian noise should be repeatedly averaged, and the error mostly relies on the amount of integrations [27]. Hence, the

CEEMDAN method has been proposed by Torres et al., which can overcome this problem through adding adaptive white noise in finite times at each period [28]. The procedure of CEEMDAN is illustrated below.

Step 1: Add a sequence of self-adaptive white Gaussian noise into the initial sequence $x(t)$:

$$x^i(t) = x(t) + \omega_0 \varepsilon^i(t), i \in \{1, \dots, I\} \quad (1)$$

where $x(t)$ indicates the initial data sequence, ω_0 is the noise coefficient, $\varepsilon^i(t)$ demonstrates the i -th addition of white Gaussian noise, $x^i(t)$ implies the time series after adding the i -th white Gaussian noise, and I represents the integration's number.

Step 2: Decompose $x^i(t)$ by the EMD method and take the average value of the first IMF component c_1^i :

$$c_1(t) = \frac{1}{I} \sum_{i=1}^I c_1^i \quad (2)$$

where c_1^i represents the first IMF component.

Remove $c_1(t)$ from the initial series $x(t)$ to obtain the first residual sequence (RS):

$$r_1(t) = x(t) - c_1(t) \quad (3)$$

where $r_1(t)$ refers to the first residual sequence.

Step 3: Proceed to disintegrate $r_1(t) + \omega_1 E_1[\varepsilon^i(t)]$ by the EMD method to obtain the second RS:

$$c_2(t) = \frac{1}{I} \sum_{i=1}^I E_1 \{ r_1(t) + \omega_1 E_1[\varepsilon^i(t)] \} \quad (4)$$

where $E_j(\cdot)$ demonstrates the j -th IMF disintegrated by the EMD approach.

Step 4: Conduct the below process repeatedly to obtain the rest of the IMFs:

$$r_k(t) = r_{k-1}(t) - c_k(t), k = 2, 3, \dots, K \quad (5)$$

$$c_{k+1}(t) = \frac{1}{I} \sum_{i=1}^I E_1 \{ r_k(t) + \omega_k E_k[\varepsilon^i(t)] \} \quad (6)$$

where K indicates the total amount of modes.

The approach ends when the RS is unable to be further disintegrated. The final RS is written as:

$$R(t) = x(t) - \sum_{k=1}^K c_k(t) \quad (7)$$

Finally, several decomposed sequences can be obtained. However, if these decomposed sequences are directly fed into the forecasting model, it will be time consuming as they are trained one by one. Therefore, Pincus proposed the approximate entropy (ApEn) method in 1991 to measure the correlation degree of two distinct time series and reconstruct the disintegrated sequences into new sequences [29]. Nevertheless, the ApEn approach depends on the data length and the consistency of the result cannot be ensured [30]. Thus, a new method called sample entropy (SE) was proposed to solve the above problems [31]. The process of the SE methodology is described below.

Step 1: The series $\{x(n)\} = x(1), x(2), \dots, x(N)$ is formed into an m -dimensional vector represented as $X_m(1), \dots, X_m(N - m + 1)$:

$$X_m(i) = \{x(i), x(i+1), \dots, x(i+m-1)\}, 1 \leq i \leq N - m + 1 \quad (8)$$

where the vector demonstrates m consecutive values of x starting from point i .

Step 2: The distance between $X_m(i)$ and $X_m(j)$ is defined and calculated in:

$$d[X_m(i), X_m(j)] = \max_{0 \leq k \leq m-1} |x(i+k) - x(j+k)| \quad (9)$$

where $d[X_m(i), X_m(j)]$ means the distance between $X_m(i)$ and $X_m(j)$.

Step 3: The total number of $d[X_m(i), X_m(j)] < r$ for every i value is counted and the ratio with $N - m + 1$ is computed to attain $B_i^m(r)$. Afterward, the average value of $B_i^m(r)$ is computed as:

$$B_i^m(r) = \frac{1}{N - m} \text{num}\{d[X_m(i), X_m(j)] < r\} \quad (10)$$

$$B^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} B_i^m(r) \quad (11)$$

where $B_i^m(r)$ is the ratio of the sum amount of $d[X_m(i), X_m(j)] < r$ for every i value accounting for $N - m + 1$, and $B^m(r)$ is the mean value of $B_i^m(r)$.

Step 4: The dimension to $m + 1$ is increased and steps 1–3 are repeatedly conducted to obtain the average value of $B_i^{m+1}(r)$:

$$B^{m+1}(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^{m+1}(r) \quad (12)$$

where $B^{m+1}(r)$ is the average value of $B_i^{m+1}(r)$ after increasing the dimension to $m + 1$.

Step 5: The SE can be written as:

$$SE(m, r) = \lim_{N \rightarrow \infty} \left\{ -\ln\left(\frac{B^{m+1}(r)}{B^m(r)}\right) \right\} \quad (13)$$

The value of SE is calculated as below when the time series length is N :

$$SE(m, r, N) = -\ln\left(\frac{B^{m+1}(r)}{B^m(r)}\right) \quad (14)$$

In accordance with relevant references, m is supposed to be 2 and r is supposed at the 0.2 standard deviation range in this research.

2.2. Improved SSA Method

A new optimization algorithm named the SSA is established by Seyedali Mirjalili in 2017 [32]. This method is carried out by simulating the behavior of the salp chain group in searching food.

For the SSA method, the initial population is classified into leader and follower. The salp chain is led by the leader. The concrete procedure of the SSA is described below:

(1) Parameter setting

In the SSA, the original population amount (Agents_no), the maximum number of iteration (Max_iteration), the variables amount (dim), the lower bound of variables ($lb = [lb_1, lb_2, \dots]$), and the upper bound of variables ($ub = [ub_1, ub_2, \dots]$) need to be set.

(2) Population initialization

The location of the original population is set stochastically, and the location matrix of the salp swarm S is shown as:

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1d} \\ s_{21} & s_{22} & \cdots & s_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nd} \end{bmatrix} \quad (15)$$

where s_{ij} demonstrates the value of the j -th variable of the i -th salp, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, d$.

s_{ij} are computed by Equation (16) with random distribution:

$$S(i, j) = rand(i, j) \times (ub(i) - lb(i)) + lb(i) \quad (16)$$

where $S(i, j)$ is the value of the i -th row and j -th column in the matrix s_{ij} , $rand(i, j)$ demonstrates the random matrix, $ub(i)$ and $lb(i)$ are the upper and the lower bounds of the i -th salp, respectively.

(3) Fitness function establishment

The fitness function is established as follows:

$$OS = \begin{bmatrix} OS_1 \\ OS_2 \\ \vdots \\ OS_n \end{bmatrix} = \begin{cases} f \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1d} \end{bmatrix} \\ f \begin{bmatrix} s_{21} & s_{22} & \cdots & s_{2d} \end{bmatrix} \\ \vdots \\ f \begin{bmatrix} s_{n1} & s_{n2} & \cdots & s_{nd} \end{bmatrix} \end{cases} \quad (17)$$

In the matrix OS, the variable F will be assigned to the salp with the best fitness value.

(4) Iteration

In the SSA, aiming at performing global search and avoiding local optimum, all salps will change their locations. Leaders adjust the position of food sources according to the following equation:

$$x_j^1 = \begin{cases} F_j + c_1(ub_j - lb_j)c_2 + lb_j & c_3 \geq 0 \\ F_j - c_1(ub_j - lb_j)c_2 + lb_j & c_3 < 0 \end{cases} \quad (18)$$

where x_j^1 implies the location of the leader in the j -th dimension, F_j represents the food source location, ub_j and lb_j illustrate the upper and lower bounds, respectively, and c_1 , c_2 , and c_3 are random constants. c_2 and c_3 are randomly generated in the range of $[0, 1]$, and c_1 is a significant parameter which is depicted as:

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \quad (19)$$

where l is the current iterations and L is the maximum iterations.

The SSA utilizes Newton's law of motion to renovate the follower's location as:

$$x_j^i = \frac{1}{2}at^2 + v_0t = \frac{1}{2}\frac{v_{final}}{v_0}t^2 + v_0t \quad (20)$$

where x_j^i represents the location of the i -th follower in the j -th dimension, t indicates the time, and v_0 implies the initial velocity and $i \geq 2$.

To enhance the optimization results of the SSA, this investigation brings dynamic inertia weight and chaotic local searching mechanism into the SSA, called improved SSA (ISSA). Therefore, the optimization searching speed of SSA can be greatly improved.

The inertia weight method is added based on Equation (18), and the new equation is written as:

$$x_j^1 = \begin{cases} wF_j + c_1(ub_j - lb_j)c_2 + lb_j & c_3 \geq 0 \\ wF_j - c_1(ub_j - lb_j)c_2 + lb_j & c_3 < 0 \end{cases} \quad (21)$$

where w represents the inertia weight, controlling the impact of parent location on the new population, so that the searching trend can be adjusted and the searching scope can be expanded. Using a greater weight at the beginning of the iteration will be conducive to determine the boundary of the optimal solution quickly, and the convergence speed will be accelerated via decreasing the weight in the following procedure. The difficulties of making optimal adjustments in accordance with the overall optimization and iteration through

employing a fixed inertia weight resulted in the dynamic inertia weight with nonlinear time-varying renovating to be utilized in our investigation to enhance the SSA.

$$w = \left(\frac{2}{t}\right)^{0.3} \quad (22)$$

where t represents the iteration steps number of the current group.

With the increasing number of iterations, the inertia weight among individuals reduces from the original value in the following steps of optimization and expedites the overall convergence speed of the SSA.

When the SSA steps into the later stage, the group prefers to plunge into the local optimization. Chaotic motion is superior in regularity, which can traverse all states within a particular range without repeating. Hence, the chaotic local searching method is added to the SSA to decrease the disadvantage of plunging into local optimization and enhance the searching speed. The concrete procedure of adding the chaotic method into the SSA is shown below:

- (1) The s_{ij} is transferred to chaotic variables as:

$$\varphi_i^j = (s_{ij} - s_i^{\min}) / (s_i^{\max} - s_i^{\min}) \quad (23)$$

where φ_i^j represents the chaotic variable, and $s_i^{\min}$ and $s_i^{\max}$ are the minimum and maximum values of the j -th variable, respectively.

- (2) The chaotic sequence is obtained by logic mapping as:

$$\varphi_{i+1}^j = 4\varphi_i^j(1 - \varphi_i^j) \quad (24)$$

- (3) The chaotic sequence is transformed into the searching space of the target solution to produce a new location. A better value is then selected to update and finish the local searching procedure. This process is expressed as:

$$s_{i+1}^j = s_i^{\min} + \varphi_{i+1}^j(s_i^{\max} - s_i^{\min}) \quad (25)$$

Thus, the basis of SSA has been improved, which has a superior convergence precision and better optimization result. Additionally, it can cause the MKELM forecasting framework to be able to easily find the optimal parameters to enhance the prediction precision.

2.3. MKELM Method

The ELM, a machine learning approach, is presented by Huang in 2006 [33]. It is a single hidden-layer feedforward network and merely requires the assumption of the amount of hidden layer nodes without updating the bias values and input weights.

Suppose that there are N training samples, where (x_i, t_i) is the input value of the CTP samples, $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]$ is the expected output value of the CTP samples, and the number of hidden nodes of the ELM is L , then, the output can be calculated by:

$$f_L(x_j) = \sum_i^L \beta G(a_i, b_i, x_j) = h(x_j)\beta \quad (26)$$

where β is the output weight from the hidden node to output node, G illustrates the activation function, a_i and b_i demonstrate the parameters of the i -th hidden layer, and $h(x_j)$ is the output vector of the hidden layer.

Moreover, it is written by a matrix expression of linear system as:

$$H\beta = T \quad (27)$$

in Equation (27):

$$\left\{ \begin{array}{l} H(a_1, \dots, a_L; b_1, \dots, b_L; x_1, \dots, x_N) \\ = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{bmatrix}_{N \times L} \\ \beta = \begin{bmatrix} \beta_1^T \\ \beta_2^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times d}; T = \begin{bmatrix} t_1^T \\ t_2^T \\ \vdots \\ t_L^T \end{bmatrix}_{N \times d} \end{array} \right. \quad (28)$$

In the ELM, supposing that the activation function is infinitely derivable, it is unnecessary to change the bias value and input weight of the structure. Then, the ELM training issue is turned into handling with β . The least-square method is commonly utilized to compute the output weight of the ELM:

$$\hat{\beta} = H^+ T = H^T (HH^T)^{-1} T \quad (29)$$

where H^+ illustrates the Moore Penrose produced inverse of the hidden layer output matrix H .

When solving H^+ , a multicollinearity problem may exist in the sample dataset, which will lead to non-singularity and then exert a negative influence on the forecasting results. Taking this into account, $1/C$ is brought into diagonal matrix HH^T to make the characteristic root of HH^T diverge from zero. Thus, the weights β of the ELM is calculated via Formula (30):

$$\beta = H^T \left(\frac{1}{C} + HH^T \right)^{-1} T \quad (30)$$

The ELM model only needs to train a small amount of parameters in the training procedure, which can greatly prevent the issues of slow computation speed and the overfitting of the conventional neural network. Nevertheless, the procedure of choosing the amount of hidden nodes is complex and time-wasting, which requires countless experiments to choose the optimal amount of neurons, leading to the decrease in forecasting performance.

The KELM model is proposed to solve the above problems, which can avoid selecting a hidden layer. Hence, the output random fluctuation can be decreased, and the generalization ability and stability of the model can be enhanced [33].

The expression of the kernel matrix in the KELM model can be written as:

$$\Omega_{ELM} = HH^T : \Omega_{ELMij} = h(x_i) \cdot h(x_j) = K(x_i, x_j) \quad (31)$$

where Ω_{ELM} implies the kernel matrix of the KELM model and $K(x_i, x_j)$ represents the kernel function.

Then, Equation (30) is written as:

$$\beta = H^T \left(\frac{1}{C} + \Omega_{ELM} \right)^{-1} T \quad (32)$$

Thus, the output of the KELM model is:

$$f(x) = h(x) H^T \left(\frac{1}{C} + HH^T \right)^{-1} T = \begin{bmatrix} \text{and } K(x, x_1) \\ \text{and } \vdots \\ \text{and } K(x, x_N) \end{bmatrix} \left(\frac{1}{C} + \Omega_{ELM} \right)^{-1} T \quad (33)$$

The KELM model contains multiple kinds of kernel functions, including poly kernel, Gaussian RBF kernel, etc. Selecting an optimal kernel function can enhance the forecasting precision of the CTP. Therefore, this investigation combines various kernel functions and makes the best of the merits of these functions in learning and generalization. Thus, a MKELM model is constructed.

Kernel functions can be classified into local kernel (such as poly kernel function) and global kernel (such as RBF kernel function), which can be expressed as:

$$K_{RBF}(x, x_i) = \exp\left(\frac{-\|x - x_i\|^2}{\delta^2}\right) \quad (34)$$

$$K_{Poly}(x, x_i) = (x \cdot x_i + 1)^d \quad (35)$$

where δ and d represent the kernel parameters of the RBF and poly kernel function, respectively.

The RBF kernel function can make the near data influence the kernel value, and has a superior performance in local learning. The poly kernel function can distinguish the impact of remote data on the kernel value, and it has good performance in global generalization [33]. As a combination of multiple affecting factors, the CTP is a dynamic evolution procedure of multiple data characteristics containing trend, volatility, and chaotic features. Hence, it is difficult to employ a single kernel function to conduct a high-accuracy regression forecasting. Hence, this research integrates the RBF and the poly kernel functions to utilize their outstanding learning and generalization abilities. The integrated kernel function is established as:

$$K(x, x_i) = \rho K_{RBF}(x, x_i) + (1 - \rho) K_{Poly}(x, x_i) \quad (36)$$

where ρ and $(1 - \rho)$ imply the weight coefficient of $K_{RBF}(x, x_i)$ and $K_{Poly}(x, x_i)$ in the integrated kernel function, respectively.

Parameter values, including C , δ , d , and ρ , are determined by the ISSA optimization algorithm.

2.4. Evaluation Tests

To measure the forecasting results, several evaluation indicators are selected, which are the mean absolute percentage error (MAPE), the root-mean-square error (RMSE), and the coefficient of determination (R2). MAPE indicates the mean forecasting precision of the model, RMSE implies the divergence between forecasting values and actual values, and R2 demonstrates the degree of fitting to the actual values ranging from 0 to 1. Models with smaller MAPE and RMSE values have better forecasting performance. Additionally, the closer to 1 the R2 value, the more superior the prediction precision of the model. Their calculation equations are below:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} * 100\% \quad (37)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (38)$$

$$R2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (39)$$

where y_i , $\hat{y}_i$, and $\bar{y}_i$, respectively, represent the actual CTP, predictive CTP, and average CTP values, and n is the length of the prediction.

3. The Process of the CEEMDAN-SE-ISSA-MKELM Model

The calculation procedure of the CEEMDAN-SE-ISSA-MKELM model is depicted in Figure 1, which is conducted by MATLAB software. As can be seen from Figure 1, specific steps of the established model are as below:

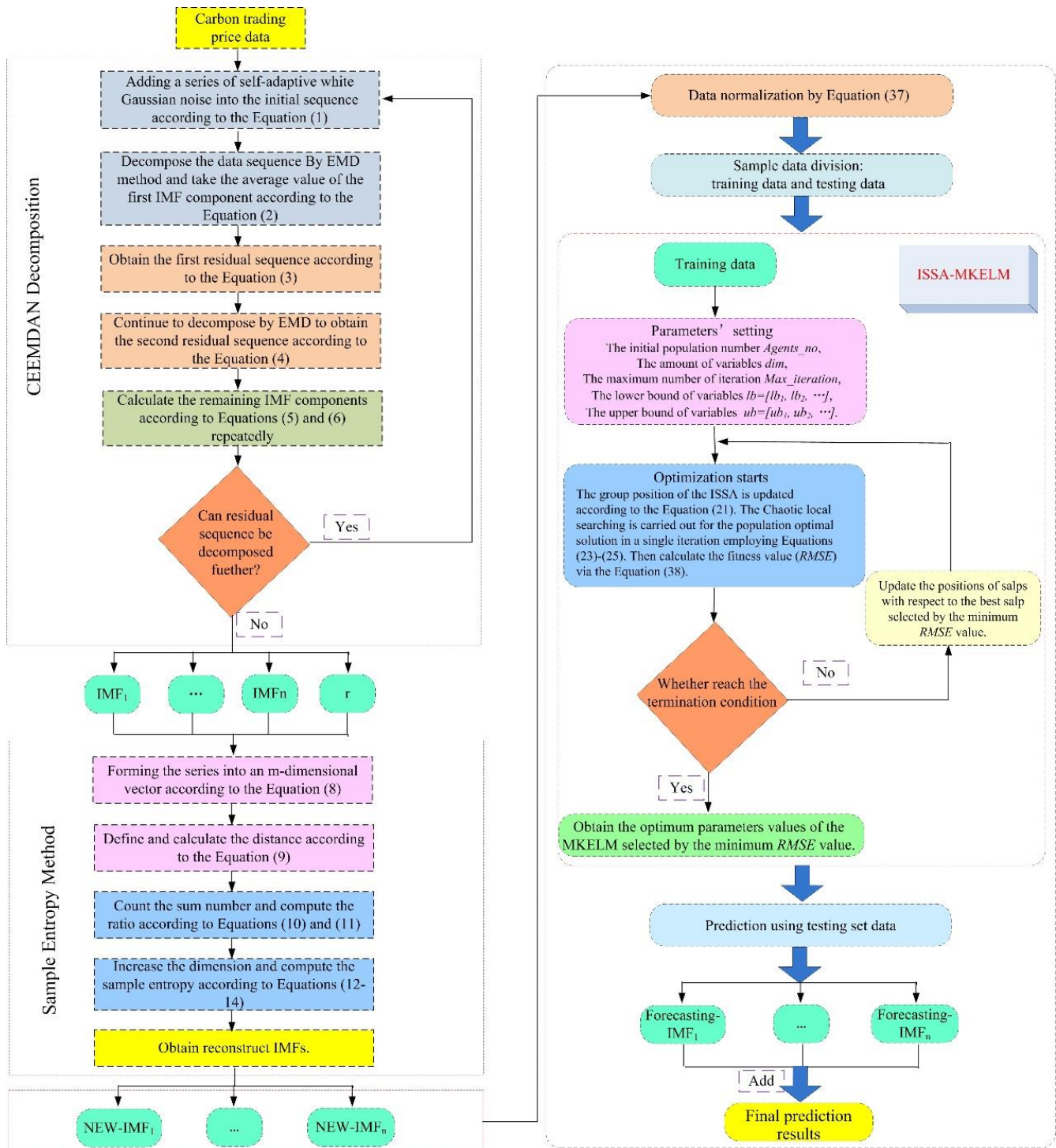


Figure 1. The calculation procedure of the proposed CEEMDAN-SE-ISSA-MKELM model.

Step 1: Initial CTP data sequence decomposition.

The parameters of the CEEMDAN method are set as: $Maxiterations = 1000$, $Std = 0.2$ and the initial CTP data can be disintegrated into several data sequences.

Step 2: Decomposed data sequence reconstruction.

The SE value of each decomposed data sequence needs to be calculated. According to the SE values, some of the decomposed data sequences can be reconstructed into new series, which will be input into the ISSA-MKELM prediction model.

Step 3: Input data sequence normalization.

The ELM method is sensitive to data scale; thus, the input data sequences obtained based on SE values will be normalized as Equation (40). Therefore, they are calculated in the same dimension:

$$\hat{x}_i = \frac{x_i - \min_i(x)}{\max_i(x) - \min_i(x)} \quad (40)$$

Step 4: Parameter setting of the ISSA and the optimal value determination for the parameters in the MKELM.

Parameters in the ISSA need to be set, including the original population amount, $Agents_no$, the maximum number of iteration, $Max_iteration$, the variables amount, dim , the lower bound of variables, $lb = [lb_1, lb_2, \dots]$, and the upper bound of variables, $ub = [ub_1, ub_2, \dots]$. This paper sets $Agents_no = 100$, $Max_iteration = 1000$, $dim = 2$, $lb = [0.01, 0.001]$, and $ub = [100, 50]$. Then, the parameters, including C , δ , d and ρ , in the MKELM model can be optimally determined through employing the ISSA optimization algorithm. The RMSE values of the forecasting results are employed as the fitness function of the ISSA, which is expressed as Equation (38).

Additionally, the group position of the ISSA can be replaced based on Equation (21). If the termination condition is realized, the process is finished, and the optimal values of parameters are finally identified.

Step 5: CTP forecasting

The decomposed data sequences obtained according to SE values will be divided into training sample and testing sample. The training sample will be substituted into the MKELM model, and then the forecasting results can be obtained. Adding all the forecasting results can finally result in the CTP prediction values.

Step 6: Prediction performance evaluation.

After obtaining the prediction results, the tests for MAPE, RMSE, and R2 are employed to measure the forecasting performance, which can be calculated according to Equations (37)–(39).

4. Numerical Analysis

4.1. Data Sources

Up until now, eight pilot carbon markets have been established in Guangdong, Hubei, Tianjin, Shenzhen, Shanghai, Beijing, Chongqing, and Fujian. Among them, Guangdong has ranked first in China's carbon trading market five times since it was officially established in 2013. Figure 2 illustrates the trading turnover structure and trading volume structure of China's carbon market. As can be seen in Figure 2, the carbon trading turnovers of the carbon markets in Guangdong, Shanghai, and Hubei occupy about 60% and the carbon trading volumes account for approximately 70% of the total market up until 31 December 2019. Hence, these three carbon markets can better represent the circumstances of China's carbon market. Furthermore, as is shown in Figure 3, the trend of the transaction prices of Guangdong, Shanghai, and Hubei's carbon trading markets present a high degree of nonlinearity and fluctuation. Thus, we employ the transaction prices in the carbon trading markets in Guangdong, Shanghai, and Hubei to verify the application and feasibility of the established CEEMDAN-SE-ISSA-MKELM model in CTP prediction.

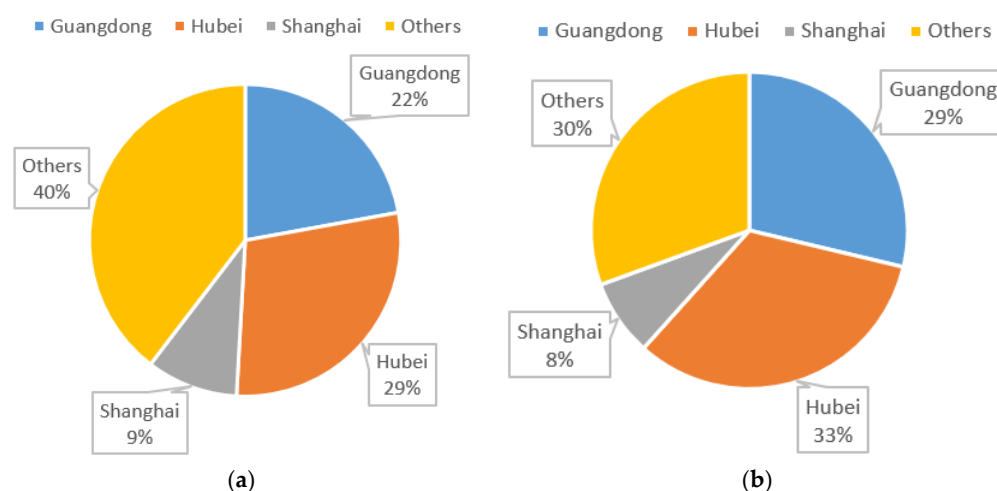


Figure 2. The trading turnover structure and trading volume structure of China's carbon market. (a) Trading turnover structure. (b) Trading volume structure.

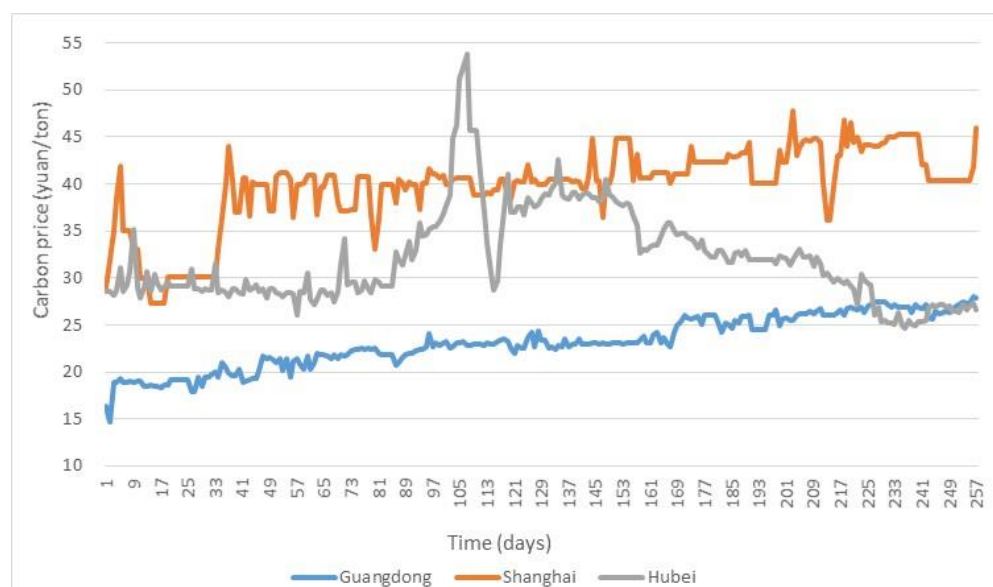


Figure 3. The original transaction prices of the carbon trading market in Guangdong, Shanghai, and Hubei.

Daily transaction prices of these markets are selected for numerical analysis from 7 January 2019 to 31 December 2019. The selected data are classified into two sets; one is the training set from 7 January 2019 to 11 October 2019 used to construct the prediction model. The other part is a testing set from 14 October 2019 to 31 December 2019 applied to prove the forecasting performance of the established model. All the relevant data are collected from <http://k.tanjiaoyi.com/> (accessed on 1 January 2021).

4.2. Data Decomposition and Reconstruction

Since the original CTP data sequences of Guangdong, Shanghai, and Hubei are non-linear and nonstationary, to improve prediction accuracy, they are decomposed into several series employing the CEEMDAN approach, and the decomposed series are reconstructed into several new sequences through the SE value of each sequence. The decomposed CTP data sequences of Guangdong, Shanghai, and Hubei are illustrated in Figure 4, which are obtained according to Equations (1)–(7). As is shown in Figure 4, each sequence after disintegration indicates strong regularity, and its nonstationarity is much more apparent, hence resulting in a large amount of IMFs after decomposing the original data sequence. If

the ISSA-MKELM model is employed to predict each decomposed sequence, the calculation scale will be increased and time-consuming. Aiming at predicting the CTP effectively and to save time, the principle of SE is employed to analyze the complexity of each IMF. Additionally, the SE values of each decomposed data sequence are listed in Figure 5, which are obtained in accordance with Equations (8)–(14).

As can be seen from Figure 5, according to the approximate SE values, the decomposed data sequence IMFs are reconstructed into new IMFs [34,35]. Taking the SE value of each decomposed data sequence of original CTP data in Guangdong as an example, IMF1–IMF3 and IMF4–IMF5 have similar complicity according to SE values, as shown in Figure 5, so that they are reconstructed as new IMFs. Similarly, the reconstruction of new IMFs are demonstrated in Table 1. Thus, the problem of over-decomposition and time-consuming computation can be solved, and the new reconstructed IMFs can be fed into the ISSA-MKELM prediction model.

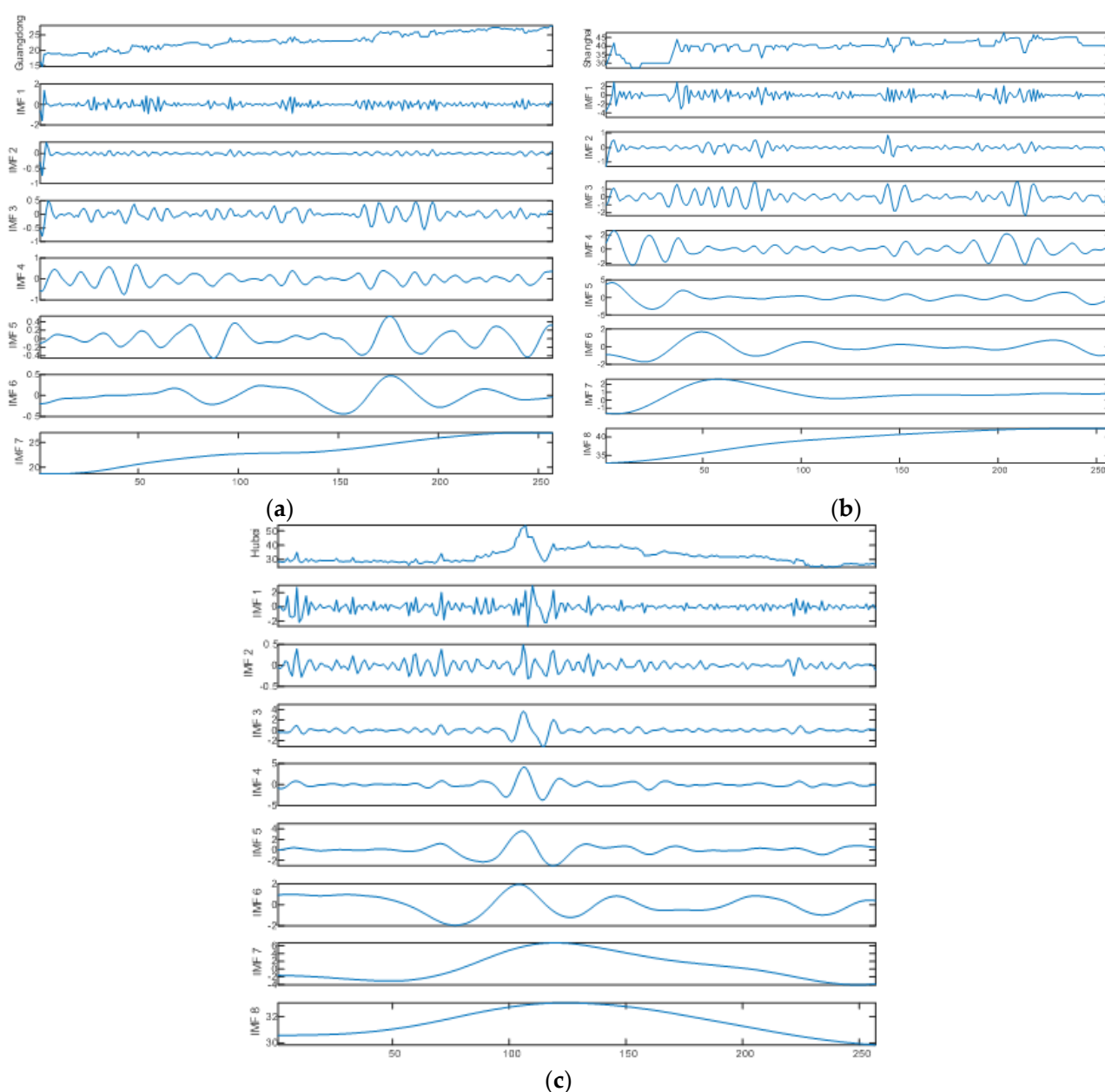


Figure 4. The decomposed CTP data in Guangdong, Shanghai, and Hubei. (a) Decomposed CTP data in Guangdong. (b) Decomposed CTP data in Shanghai. (c) Decomposed CTP data in Hubei.

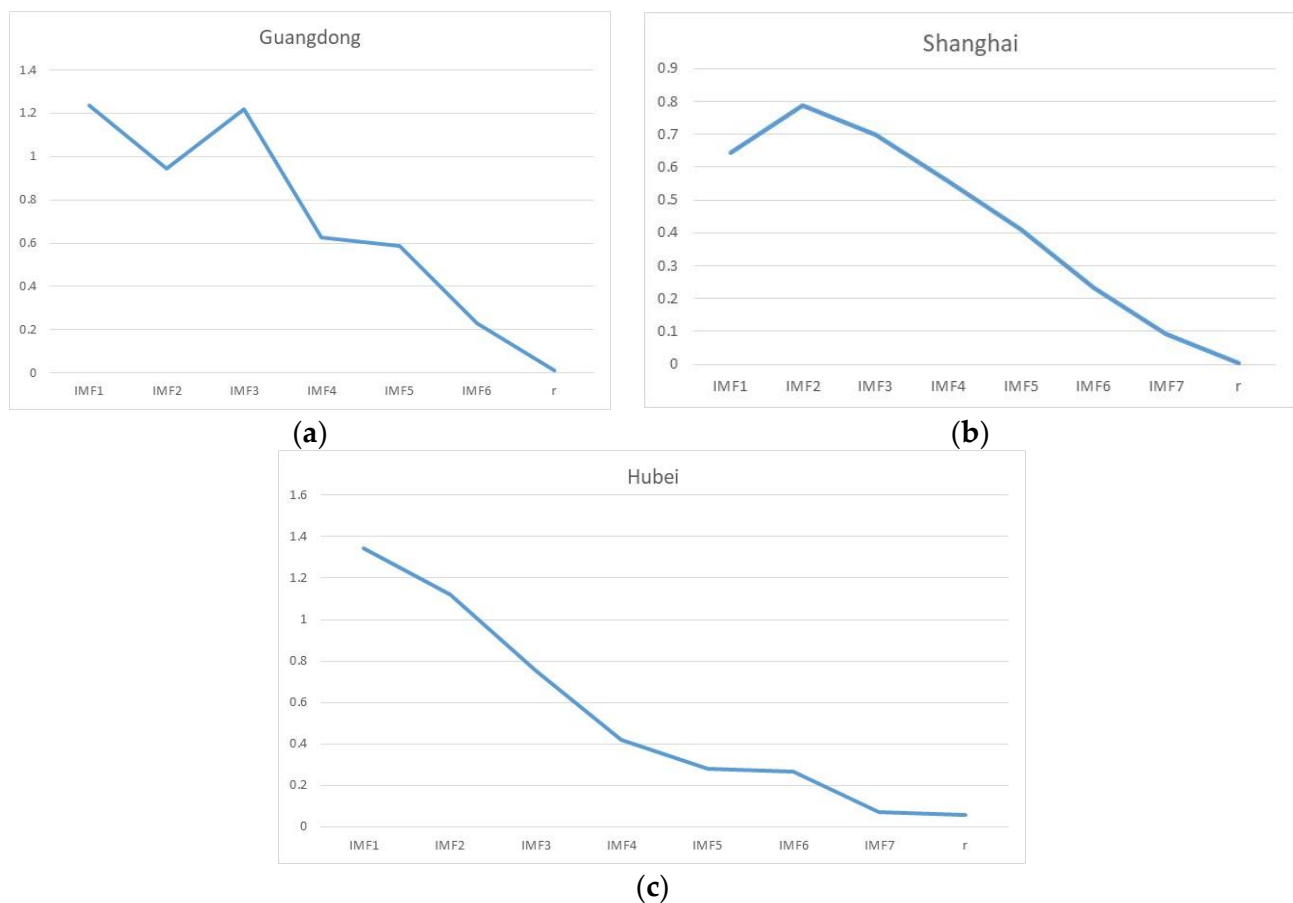


Figure 5. The SE values of each decomposed data sequence. (a) The SE values in Guangdong. (b) The SE values in Shanghai. (c) The SE values in Hubei.

Table 1. The reconstructed new data sequences of decomposed CTP data for Guangdong, Shanghai, and Hubei.

Province	NEW-IMF1	NEW-IMF2	NEW-IMF3	NEW-IMF4
Guangdong	IMF1 + IMF2 + IMF3	IMF4 + IMF5	IMF6	r
Shanghai	IMF1 + IMF2 + IMF3 + IMF4	IMF5	IMF6	IMF7 + r
Hubei	IMF1 + IMF2	IMF3	IMF4 + IMF5 + IMF6	IMF7 + r

4.3. Forecasting Results of the Established CEEMDAN-SE-ISSA-MKELM

After obtaining the reconstructed new IMFs, the input data sequences into the ISSA-MKELM can be determined. Then, the forecasting results from 14 October 2019 to 31 December 2019 can be obtained. To verify the forecasting performance of the established CEEMDAN-SE-ISSA-MKELM model, the forecasting CTP results should be compared with the actual CTP data. Evaluation indicators, including MAPE, RMSE, and R2, are selected to verify the prediction accuracy, which are obtained via Equations (37)–(39), and the results are illustrated in Table 2. As listed in Table 2, the values of R2 for Guangdong, Shanghai, and Hubei's carbon markets are all 0.99, which implies that the established forecasting model has a superior prediction performance. The values of MAPE of Guangdong, Shanghai, and Hubei's carbon markets are 0.76%, 1.65%, and 1.33%, respectively. According to the measuring criterion of MAPE [36], as the values of the MAPE of three carbon markets are all smaller than 10%, and the performance of the established model in CTP forecasting is excellent. The values of RMSE of Guangdong, Shanghai, and Hubei carbon markets are 0.53, 0.77, and 0.67, respectively, which are relatively high prediction accuracy values.

Table 2. The forecasting performance of the established CEEMDAN-SE-ISSA-MKELM model.

Province	MAPE	RMSE	R2
Guangdong	0.76%	0.53	0.99
Shanghai	1.65%	0.77	0.99
Hubei	1.33%	0.67	0.99

5. Comparison Analysis

To further verify the validity of the established CEEMDAN-SE-ISSA-MKELM model in CTP forecasting, the single model, including Markov forecasting model (MFM) [37], ARIMA [38], ELM [39], LSSVM [40], LSTM [38], poly kernel ELM (pELM), RBF kernel ELM (rELM), and MKELM (combining poly kernel and RBF kernel), as well as the combined model, including CEEMDAN-MKELM, EMD-MKELM [41], CEEMD-MKELM [42], CEEMDAN-SE-MKELM, EMD-SE-MKELM, CEEMD-SE-MKELM, CEEMDAN-SE-FOA-MKELM [43], CEEMDAN-SE-PSO-MKELM [44], and CEEMDAN-SE-SSA-MKELM, are selected as the comparison model. The values of MAPE, RMSE, and R2 for the forecasted results in Guangdong's carbon market are illustrated in Figure 6. Moreover, the values of MAPE, RMSE, and R2 for the forecasted results in Shanghai's carbon market are illustrated in Figure 7. Values of MAPE, RMSE, and R2 for the forecasted results in Hubei's carbon market are depicted in Figure 8.

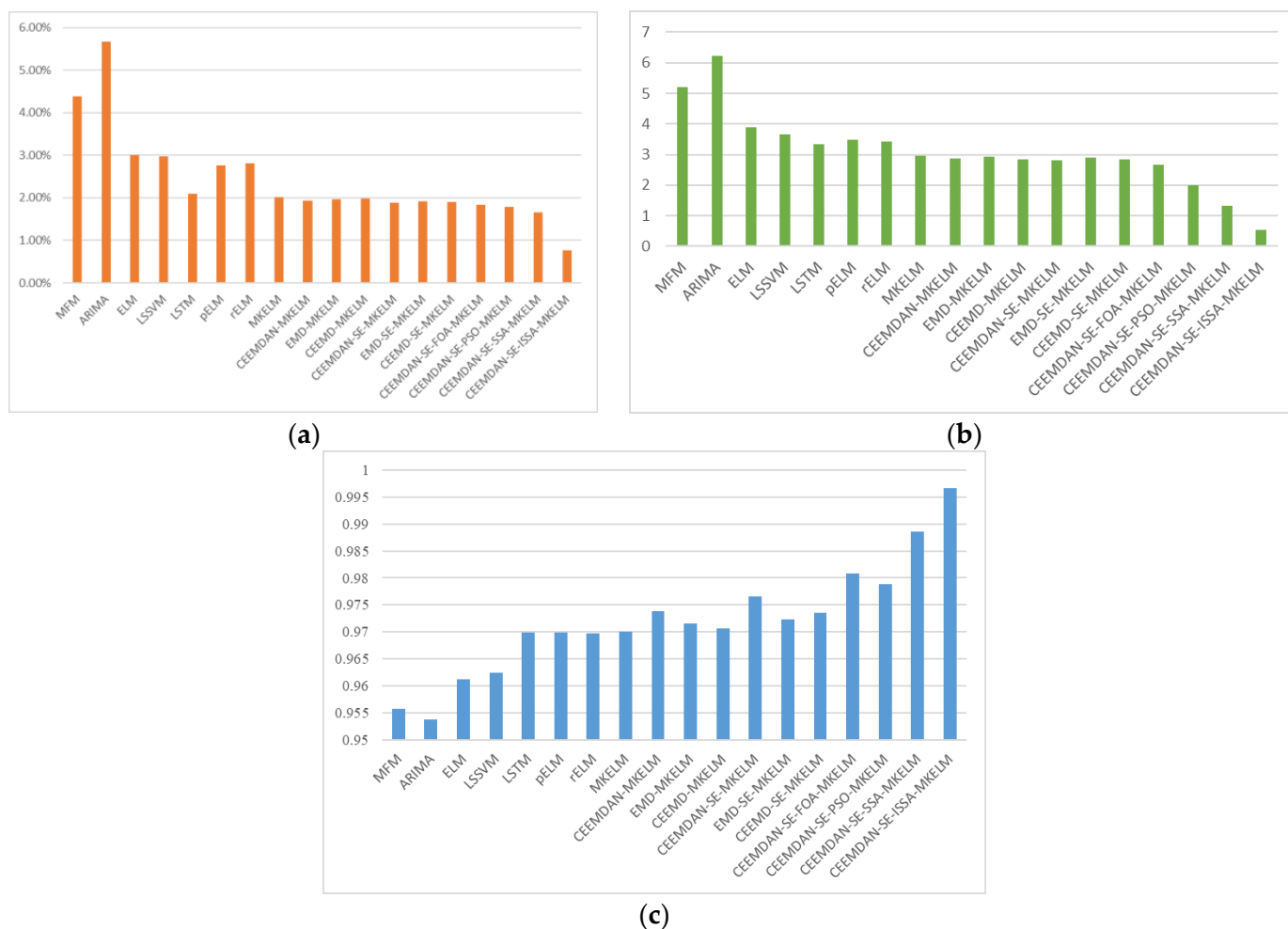


Figure 6. The MAPE, RMSE, and R2 values for the forecasted results in Guangdong's carbon market. (a) MAPE (%). (b) RMSE. (c) R2.

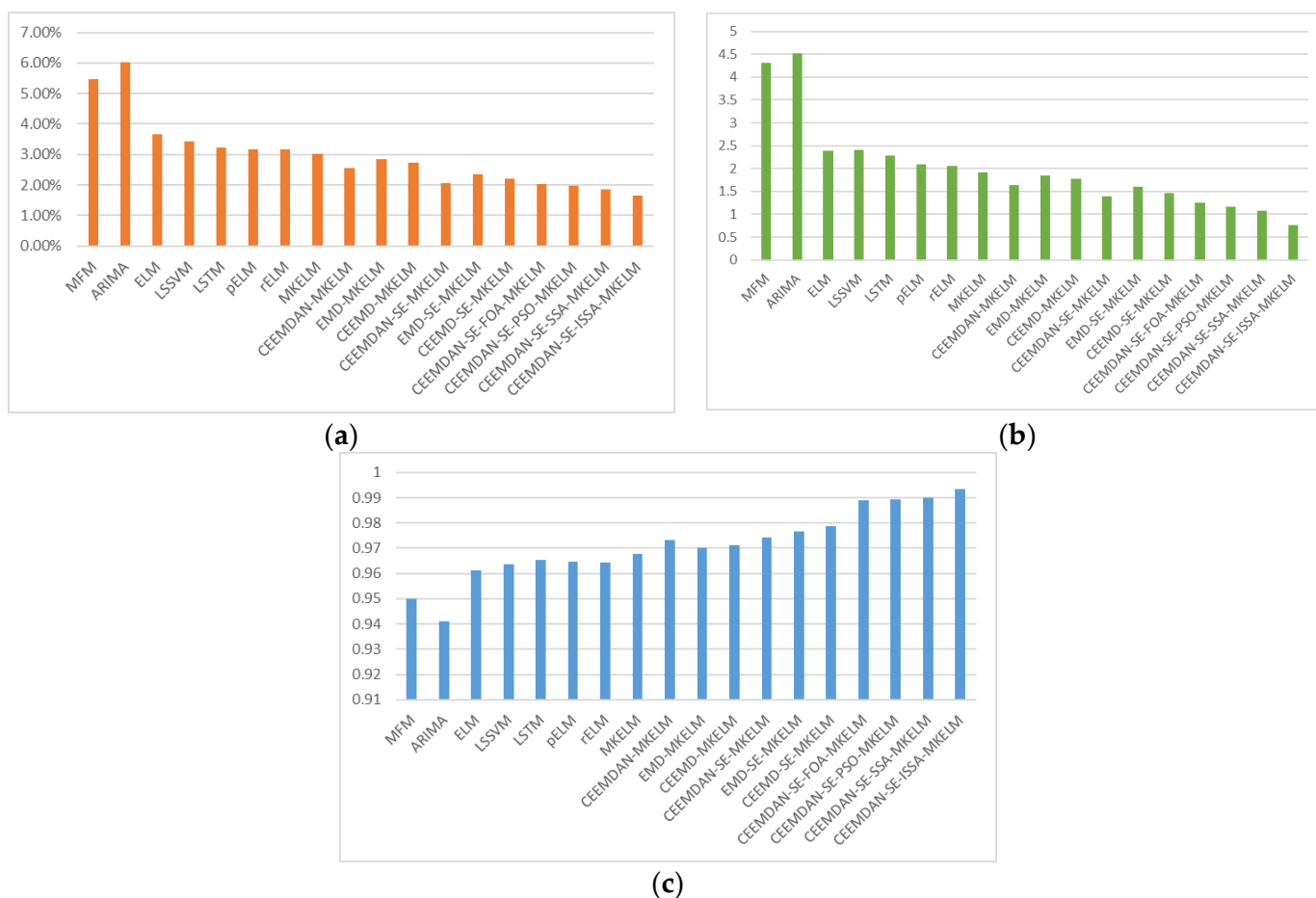


Figure 7. The MAPE, RMSE, and R2 values for the forecasted results in Shanghai's carbon market. (a) MAPE (%). (b) RMSE. (c) R2.

With the aim of comparing the role of CEEMDAN and the SE method in improving the forecasting accuracy of CTP, a comparison analysis among MKELM, CEEMDAN-MKELM, EMD-MKELM, and CEEMD-MKELM has been discussed. Moreover, the results imply that the performance of the decomposed data sequences are better than those of undecomposed data. Furthermore, in order to simplify the input sequences and save prediction time, the SE method is applied to reconstruct the decomposed data sequences to obtain new IMFs, which will be fed into the prediction model. Through a comparison of the performance of CEEMDAN-MKELM, EMD-MKELM, CEEMD-MKELM, CEEMDAN-SE-MKELM, EMD-SE-MKELM, and CEEMD-SE-MKELM, the superiority of the SE approach can be verified, as the combined models with SE approach have relatively low MAPE and RMSE values and high R2 values. Therefore, it is necessary to employ the disintegration method to preprocess the initial CTP data sequences to enhance the prediction precision and actual fitting ability. Additionally, the CEEMDAN-SE approach can greatly improve the prediction precision.

By comparing the prediction performance of single models, including MFM, ARIMA, ELM, LSSVM, LSTM, pELM, rELM, and MKELM, it can be found that the forecasting performance of the MKELM model is superior than other single models, of which the values of MAPE, RMSE, and R2 in Guangdong's carbon market are 2.01%, 2.96, and 0.9701, respectively. Thus, the MKELM is selected as the basic model instead of other models in this investigation.

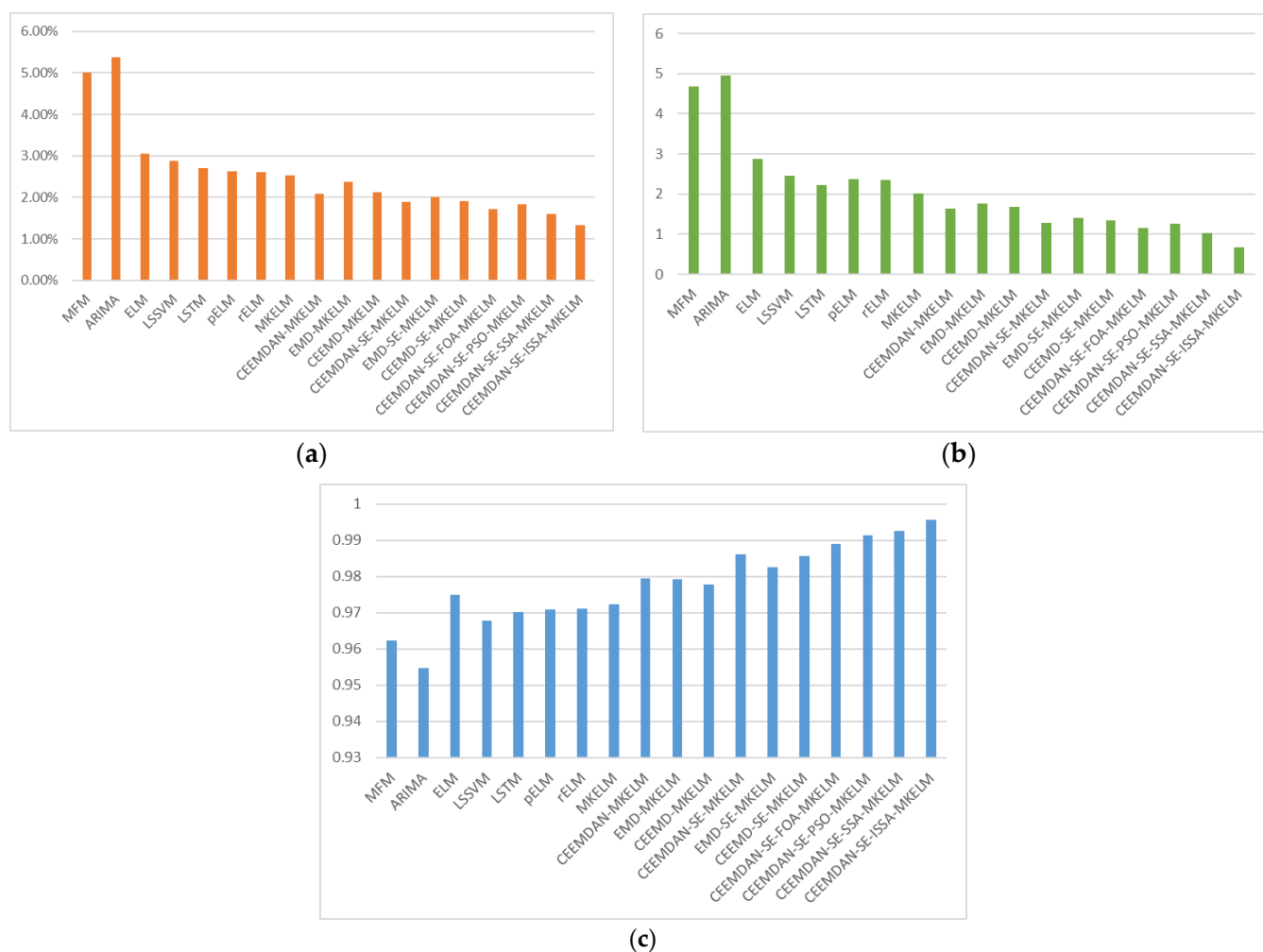


Figure 8. The MAPE, RMSE, and R2 values for the forecasted results in Hubei's carbon market. (a) MAPE (%). (b) RMSE. (c) R2.

Additionally, the parameters of single MKELM model can also be optimally determined by optimization algorithms. In this investigation, by comparing the forecasting performance among CEEMDAN-SE-MKELM, CEEMDAN-SE-FOA-MKELM, CEEMDAN-SE-PSO-MKELM, and CEEMDAN-SE-SSA-MKELM models, it can be discovered that the combined models with optimization algorithms have a higher prediction precision than that without the optimization algorithm. Specifically, the MAPE, RMSE, and R2 values of the CEEMDAN-SE-SSA-MKELM model in Guangdong's carbon market are 1.65%, 1.33, and 0.9886, respectively, while the MAPE, RMSE, and R2 values of CEEMDAN-SE-PSO-MKELM model in Guangdong's carbon market are 1.79%, 2.01, and 0.9789, respectively, and the MAPE, RMSE, and R2 values of the CEEMDAN-SE-MKELM model in Guangdong's carbon market are 1.88%, 2.81, and 0.9766, respectively. Furthermore, aiming at further improving the optimization performance of the basic SSA, the chaotic local searching strategy and inertia weight mechanism are added to the SSA to establish the ISSA model. Through a comparison of the forecasting performance of CEEMDAN-SE-SSA-MKELM and CEEMDAN-SE-ISSA-MKELM models, we can verify that the established model with ISSA (for Guangdong's carbon market, MAPE = 0.76%, RMSE = 0.53, and R2 = 0.9967) is superior than that with the basic SSA method. Therefore, the ISSA method proposed in this investigation can further improve the forecasting performance in CTP.

For the forecasting results of Guangdong, Shanghai, and Hubei's carbon markets, the established CEEMDAN-SE-ISSA-MKELM model has the best prediction precision, with the least MAPE values of 0.76%, 1.65%, and 1.33%, respectively, the smallest RMSE values

of 0.53, 0.77, and 0.67, respectively, and the largest R2 values of 0.9967, 0.9934, and 0.9956, respectively. This demonstrates that the complicated hybrid model established in this investigation has high-prediction accuracy, is feasible, and is superior in CTP forecasting. However, the performance of the ARIMA model is the worst, with the highest MAPE values of 5.67%, 6.02%, and 5.38%, respectively, the highest RMSE values of 6.23, 4.52, and 4.96, respectively, and the lowest R2 values of 0.9537, 0.9412, and 0.9547, respectively.

Through multi-angles comparison, the respective contributions of CEEMDAN, SE, ISSA, and MKELM models in CTP prediction are verified. Additionally, the established CEEMDAN-SE-ISSA-MKELM model has a superior feasibility in CTP forecasting in different carbon trading markets.

6. Conclusions

Since China has promised to realize the peak of carbon emissions by 2030 and carbon neutrality by 2060, carbon trading attracts much attention as it has been verified as a valid way to promote carbon emission reduction. Considering that the competition among market participants will inevitably occur through the operation of carbon trading, the accurate prediction of CTP has become a primary factor in the competition strategy formulation by all market participants. Therefore, this investigation proposed a hybrid CTP prediction model combining CEEMDAN, SE, ISSA, and MKELM methods to enhance the prediction precision. By analyzing the initial CTP data sequences, it can be determined that the CTP data are volatile and unstable. Hence, the CEEMDAN method has been employed to disintegrate the initial CTP data sequences into several IMFs and a residual sequence. In order to save calculation time, the SE method has been utilized to reconstruct the IMFs and the residual sequence into new IMFs. Then, the new IMFs are fed into the MKELM model, combining RBF and the poly kernel functions to utilize the outstanding learning and generalization abilities of them. The parameters of the MKELM model are optimized by ISSA, which combines dynamic inertia weight and chaotic local searching mechanism into the SSA to enhance the searching speed, convergence precision, and global searching ability of the basic SSA.

CTP data in Guangdong, Shanghai, and Hubei are selected to prove the robustness of the established CEEMDAN-SE-ISSA-MKELM model, as the carbon trading turnovers and volumes of the carbon markets in Guangdong, Shanghai, and Hubei occupy more than 60% of the total market up until the end of December 2019, which is representative of these three carbon markets. The single model, including MFM, ARIMA, ELM, LSSVM, LSTM, pELM, rELM, and MKELM (combining poly kernel and RBF kernel), as well as the combined model, including CEEMDAN-MKELM, EMD-MKELM, CEEMD-MKELM, CEEMDAN-SE-MKELM, EMD-SE-MKELM, CEEMD-SE-MKELM, CEEMDAN-SE-FOA-MKELM, CEEMDAN-SE-PSO-MKELM, and CEEMDAN-SE-SSA-MKELM, are selected as the multi-angle comparison models. Additionally, by comparing the values of MAPE, RMSE, and R2, the following conclusions are drawn:

- (1) The established CEEMDAN-SE-ISSA-MKELM model performs the best among all other comparison models, which demonstrates that the integration of these methods can remarkably increase the forecasting accuracy of CTP.
- (2) The forecasting performance of the MKELM model is superior than other single models. Thus, the MKELM is selected as the basic model instead of other models in this investigation.
- (3) The performance of the decomposed data sequences are better than those of undecomposed data. Therefore, it is necessary to employ a decomposition approach to preprocess the initial data with high uncertainty and volatility. Moreover, the comparison analysis has proven that the performance of the CEEMDAN approach exceeds other decomposition methods. Additionally, the SE method is applied to rebuild the decomposed data sequences, which can save calculation time and improve prediction accuracy.

- (4) The optimization performance of the SSA method is better than other optimization algorithms, including FOA and PSO. Furthermore, the ISSA has great improvement in searching speed, convergence precision, and global searching ability, which further improves the forecasting accuracy. Therefore, the ISSA proposed in this research is of great significance.
- (5) The carbon markets selected for the experiment herein represent the circumstances of China's current carbon trading markets; the established CEEMDAN-SE-ISSA-MKELM model has high universality and generality in CTP prediction. The established CTP prediction model can not only be widely applied in CTP prediction, but also offer reference and guidance for decision-makers in reality.

However, this research only forecasted future CTP by historical data. CTP can be influenced by other factors, including oil price, coal price, gas price, related policy, etc. In future study, a further investigation will be conducted to research the influences of these factors on the CTP. Moreover, the development of an improved Markov forecasting method to better predict the carbon trading price is another avenue of research which we plan to explore in the future.

Author Contributions: Conceptualization, H.Z. and S.G.; data curation, S.G.; formal analysis, H.Z.; investigation, H.Z.; methodology, H.Z.; software, H.Z.; validation, S.G.; writing—original draft, H.Z.; writing—review and editing, S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is granted by Qin Xin Talents Cultivation Program, Beijing Information Science and Technology University, under grant no. QXTCPC202113.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the relevant data are collected from <http://k.tanjiaoyi.com/> (accessed on 1 January 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hashim, H.; Ramlan, M.R.; Shiun, L.J.; Siong, H.C.; Kamyab, H.; Abd Majid, M.Z.; Lee, C.T. An Integrated Carbon Accounting and Mitigation Framework for Greening the Industry. *Energy Procedia* **2015**, *75*, 2993–2998. [\[CrossRef\]](#)
2. Sun, W.; Huang, C. A novel carbon price prediction model combines the secondary decomposition algorithm and the long short-term memory network. *Energy* **2020**, *207*, 118294. [\[CrossRef\]](#)
3. Saleh, T.A.; Fadillah, G. Efficient detection of CO₂ by nanocomposites: Environmental and energy technologies. *Trends Environ. Anal. Chem.* **2021**, *32*, e00142. [\[CrossRef\]](#)
4. Zhang, J.; Li, D.; Hao, Y.; Tan, Z. A hybrid model using signal processing technology, econometric models and neural network for carbon spot price forecasting. *J. Clean. Prod.* **2018**, *204*, 958–964. [\[CrossRef\]](#)
5. Jiang, K.; Ashworth, P. The development of Carbon Capture Utilization and Storage (CCUS) research in China: A bibliometric perspective. *Renew. Sustain. Energy Rev.* **2021**, *138*, 110521. [\[CrossRef\]](#)
6. Zhang, J. Estimating the impacts of emissions trading scheme on low-carbon development. *J. Clean. Prod.* **2019**, *238*, 117913. [\[CrossRef\]](#)
7. Huang, Y.; He, Z. Carbon price forecasting with optimization prediction method based on unstructured combination. *Sci. Total Environ.* **2020**, *725*, 138350. [\[CrossRef\]](#)
8. Li, S.; Wang, Q. India's dependence on foreign oil will exceed 90% around 2025—The forecasting results based on two hybridized NMGM-ARIMA and NMGM-BP models. *J. Clean. Prod.* **2019**, *232*, 137–153. [\[CrossRef\]](#)
9. Zhu, B.; Wei, Y. Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology. *Omega* **2013**, *41*, 517–524. [\[CrossRef\]](#)
10. Jiang, L.; Wu, P. International Carbon Market Price Forecasting Research Based on ARIMA-RF Model. In Proceedings of the 2016 International Conference on Strategic Management, Sichuan, China, 10 March 2016; pp. 1089–1092.
11. Fu, Y.; Zheng, Z. Volatility modeling and the asymmetric effect for China's carbon trading pilot market. *Phys. A Stat. Mech. Its Appl.* **2020**, *542*, 123401. [\[CrossRef\]](#)
12. Wang, S.; Li, S.G. A Novel Hybrid Carbon Price Forecasting Model Based on Radial Basis Function Neural Network. *Acta Phys. Pol. A* **2019**, *135*, 368–374. [\[CrossRef\]](#)

13. Fan, X.; Li, S.; Tian, L. Chaotic characteristic identification for carbon price and an multi-layer perceptron network prediction model—ScienceDirect. *Expert Syst. Appl.* **2015**, *42*, 3945–3952. [\[CrossRef\]](#)
14. Tian, C.; Hao, Y. Point and interval forecasting for carbon price based on an improved analysis-forecast system. *Appl. Math. Model.* **2020**, *79*, 126–144. [\[CrossRef\]](#)
15. Salehnia, N.; Ahn, J. Modelling and reconstructing tree ring growth index with climate variables through artificial intelligence and statistical methods. *Ecol. Indic.* **2022**, *134*, 108496. [\[CrossRef\]](#)
16. Soltis, P.S.; Nelson, G.; Zare, A.; Meineke, E.K. Plants meet machines: Prospects in machine learning for plant biology. *Appl. Plant Sci.* **2020**, *8*, 11371. [\[CrossRef\]](#)
17. Dijk, A.; Kootstra, G.; Kruijer, W.; Ridder, D.D. Machine learning in plant science and plant breeding. *iScience* **2020**, *24*, 101890. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Moazami, S.; Noori, R.; Amiri, B.J.; Yeganeh, B.; Safavi, S. Reliable prediction of carbon monoxide using developed support vector machine. *Atmos. Pollut. Res.* **2016**, *7*, 412–418. [\[CrossRef\]](#)
19. Zfla, B.; Llla, B.; Mltc, D.; Ming, K. Prediction short-term photovoltaic power using improved chicken swarm optimizer—Extreme learning machine model. *J. Clean. Prod.* **2020**, *248*, 119272.
20. Chen, H.; Zhang, Q.; Luo, J.; Xu, Y.; Zhang, X. An enhanced Bacterial Foraging Optimization and its application for training kernel extreme learning machine. *Appl. Soft Comput.* **2019**, *86*, 105884. [\[CrossRef\]](#)
21. Liu, X.; Lei, W.; Huang, G.B.; Jian, Z.; Yin, J. Multiple kernel extreme learning machine. *Neurocomputing* **2015**, *149*, 253–264. [\[CrossRef\]](#)
22. Wang, M.; Chen, H.; Li, H.; Cai, Z.; Zhao, X. Grey wolf optimization evolving kernel extreme learning machine: Application to bankruptcy prediction. *Eng. Appl. Artif. Intell.* **2017**, *63*, 54–68. [\[CrossRef\]](#)
23. Zhu, B.; Han, D.; Wang, P.; Wu, Z.; Zhang, T.; Wei, Y.M. Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression. *Appl. Energy* **2017**, *191*, 521–530. [\[CrossRef\]](#)
24. Nikkhoo, A.; Karegar, H.; Mohammadi, R.K.; Hejazi, F. Improving the performance of the autoregressive method in modal identification of output-only systems using the empirical mode decomposition. *Structures* **2020**, *27*, 1165–1173. [\[CrossRef\]](#)
25. Jia, Y.; Li, G.; Dong, X.; He, K. A novel denoising method for vibration signal of hob spindle based on EEMD and grey theory. *Measurement* **2021**, *169*, 108490. [\[CrossRef\]](#)
26. Zhaohua, W.U.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2011**, *1*, 1–41.
27. Gao, B.; Huang, X.; Shi, J.; Tai, Y.; Zhang, J. Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks. *Renew. Energy* **2020**, *162*, 1665–1683. [\[CrossRef\]](#)
28. Torres, M.E.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 4144–4147.
29. Pincus, M.S. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [\[CrossRef\]](#)
30. Wei, S.; Yuwei, W. Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network. *Energy Convers. Manag.* **2018**, *157*, 1–12.
31. Richman, J.S.; Randall, M.J. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, 2039. [\[CrossRef\]](#)
32. Mirjalili, S.; Gandomi, A.H.; Mirjalili, S.Z.; Saremi, S.; Faris, H.; Mirjalili, S.M. Salp swarm algorithm: A bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* **2017**, *114*, 163–191. [\[CrossRef\]](#)
33. Huang, G.B.; Zhou, H.; Ding, X.; Rui, Z. Extreme Learning Machine for Regression and Multiclass Classification. *IEEE Trans. Syst. Man Cybern. Part B* **2012**, *42*, 513–529. [\[CrossRef\]](#)
34. Wu, Q.; Lin, H. Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and LSTM neural network. *Sustain. Cities Soc.* **2019**, *50*, 101657. [\[CrossRef\]](#)
35. Wang, J.; Cao, J.; Yuan, S.; Cheng, M. Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the CEEMDAN-SE-and the PSO-ALS-optimized GRU network. *Energy* **2021**, *233*, 121082. [\[CrossRef\]](#)
36. Zhao, Z.; Wang, J.; Zhao, J.; Su, Z. Using a Grey model optimized by Differential Evolution algorithm to forecast the per capita annual net income of rural households in China. *Omega* **2012**, *40*, 525–532. [\[CrossRef\]](#)
37. Baum, L.E.; Petrie, T. Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *Ann. Math. Stat.* **1966**, *37*, 1554–1563. [\[CrossRef\]](#)
38. Jin, Y.; Wang, R.; Zhuang, X.; Wang, K.; Wang, H.; Wang, C.; Wang, X. Prediction of COVID-19 Data Using an ARIMA-LSTM Hybrid Forecast Model. *Mathematics* **2022**, *10*, 4001. [\[CrossRef\]](#)
39. Qiu, J.; Yin, X.; Pan, Y.; Wang, X.; Zhang, M. Prediction of Uniaxial Compressive Strength in Rocks Based on Extreme Learning Machine Improved with Metaheuristic Algorithm. *Mathematics* **2022**, *10*, 3490. [\[CrossRef\]](#)
40. Wang, G.; Wang, X.; Wang, Z.; Ma, C.; Song, Z. A VMD-CISSA-LSSVM Based Electricity Load Forecasting Model. *Mathematics* **2022**, *10*, 28. [\[CrossRef\]](#)
41. Mwata-Velu, T.; Avina-Cervantes, J.G.; Cruz-Duarte, J.M.; Rostro-Gonzalez, H.; Ruiz-Pinales, J. Imaginary Finger Movements Decoding Using Empirical Mode Decomposition and a Stacked BiLSTM Architecture. *Mathematics* **2021**, *9*, 3297. [\[CrossRef\]](#)

42. Valtierra-Rodriguez, M.; Amezcua-Sanchez, J.P.; Garcia-Perez, A.; Camarena-Martinez, D. Complete Ensemble Empirical Mode Decomposition on FPGA for Condition Monitoring of Broken Bars in Induction Motors. *Mathematics* **2019**, *7*, 783. [[CrossRef](#)]
43. Zainel, Q.M.; Darwish, S.M.; Khorsheed, M.B. Employing Quantum Fruit Fly Optimization Algorithm for Solving Three-Dimensional Chaotic Equations. *Mathematics* **2022**, *10*, 4147. [[CrossRef](#)]
44. Berg, S.S. Utility of Particle Swarm Optimization in Statistical Population Reconstruction. *Mathematics* **2023**, *11*, 827. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.