Forecasting China's Renewable Energy Terminal Power Consumption Based on Empirical Mode Decomposition and an Improved Extreme Learning Machine Optimized by a Bacterial Foraging Algorithm

Peng Jiang *, Jun Dong and Hui Huang

School of Economics and Management, North China Electric Power University, Beijing 102206, China; dongjun624@126.com (J.D.); hh@ncepu.edu.cn (H.H.)

* Correspondence: 1172106016@ncepu.edu.cn; Tel.: +86-139-1151-3738

check for updates

Abstract: The energy consumption pattern dominated by traditional fossil energy has led to global energy resource constraints and the deterioration of the ecological environment. These challenges have become a major issue all over the world. At present, the Chinese government aims to significantly reduce the fossil energy consumption contribution in the terminal energy consumption. The development of renewable energy in the terminal energy and energy conversion links has significantly increased the proportion of clean low-carbon energy. In order to accurately get the proportion of renewable energy terminal power consumption, firstly, this paper selects a primary influencing-factors set including the gross GDP, fixed investment in renewable energy industry, total length of cross-provincial and cross-regional high-voltage transmission lines, etc. as influencing factors of China's electricity consumption fraction produced by renewable energy based on a multitude of papers. Secondly, from the perspective of signal decomposition, the data inevitably has a lot of interference and noise. This paper uses the empirical mode decomposition (EMD) algorithm to reduce the degree of signal distortion and decomposes the signal into natural modes including several intrinsic mode functions (IMFs) and a residual term (Res); afterwards, a new extreme learning machine (ELM) forecasting model optimized by an Inverse Square Root Linear Units (ISRLU) activation function is proposed, and the ISRLU function is used to replace the implicit layer activation function in the original ELM algorithm. Then, a new bacterial foraging algorithm (BFOA) is applied to optimize the parameters of the optimized ELM forecasting model. After multiple learning and training operations, the optimal parameters are obtained. Finally, we superimpose the output of each IMF and Res training task to get the amount of China's power consumption produced by renewable energy. Some statistical indicators including root mean squard error (RMSE) are applied to compare the accuracy of several intelligent machine forecasting algorithms. We prove that the proposed forecasting model has higher prediction accuracy and achieves faster training speed by an empirical analysis. Finally, the proposed combined forecasting algorithm is applied to predict China's renewable energy terminal power consumption from 2018 to 2030. According to the forecasting results, it is found that China's renewable energy terminal power consumption shows a gradual growth trend, and will exceeded 3300 billion kWh in 2030, which will represent a renewable energy terminal power ratio of about 38% in 2030.

Keywords: renewable energy terminal power consumption; energy forecasting; grey theory; empirical mode decomposition; bacterial foraging algorithm; extreme learning machine algorithm

1. Introduction

At present, China is formulating a national energy strategy calling for the use of a high proportion of renewable energy as the core means to achieve the national non-fossil energy development goals in 2020 and 2030 and to realize an energy production and consumption revolution. A high proportion of renewable energy terminal power consumption has become the basic energy structure layout for 2050. The China Energy Bureau announced that by 2050, China will form an integrated energy system based on renewable energy and a high proportion of renewable energy in the terminal energy consumption. The proportion should reach more than 60%, and the proportion of total renewable energy generation will reach more than 85% in the energy consumption layout, the electrification of the terminal energy consumption will be above 50%, the total electricity consumption will increase to 13.5~15 trillion kWh, and the per capita electricity consumption shall be 10,000~11,000 kWh. [1]. The central role of electricity in achieving a high proportion of renewable energy development is objectively determined by the characteristics of electricity, resource endowments and energy development. Regardless of the relationship between electricity and other energy sources, or from the aspects of ensuring energy security, optimizing energy structure, and promoting ecological civilization construction, studying the proportion of renewable energy consumption of electricity is crucial to achieving the government's strategic goals. As for significance, wind power, solar energy, and electricity have provided considerable macroeconomic and environmental benefits for achieving high proportions of renewable energy development. The high proportion of renewable energy development has also significantly replaced coal consumption. Through the high proportion of renewable energy development, the continuous reduction of the total emissions of major atmospheric pollutants (SO and NOx) will be ensured and controlled at 250 in 2050. Within 10,000 tons and 2.7 million tons, the emission of major pollutants (including heavy metal mercury, etc.) is equal to the emission level in 1980 [2], thus realizing China's responsibility for environmental protection in the world.

The energy consumption of the whole society is directly or indirectly affected by many factors. Up to now, many scholars have studied the factors affecting the energy consumption [3–6]. The problem of electricity consumption is a complex non-linear problem, for which so far scholars have proposed various prediction models, such as grey theory [7–9], multiple regression [10–12], and time series models [13–15]; In recent years, various intelligent algorithms have also been applied to power consumption prediction [16–22].

Meng et al. [3] proposed a three-dimensional decomposition model and a mixed trend extrapolation model to explore the factors driving the growth of household electricity consumption in China, predicting the future development trend before 2030, and summarizing some of the main influencing factors; Akay et al. [4] used the Grey Prediction and Rolling Mechanism (GPRM) method to predict Turkey's overall and industrial electricity consumption, and both social and economic factors were adopted to forecast power consumption. Castillo et al. [5] used a unified data set of 13 income and expenditure household surveys to assess changes in electrical and electricity consumption, taking into account income distribution, GDP, population, etc. as indicators of impact assessment; Pablo-Romero [6] analyzed the relationship between electricity consumption and tourism growth in hotels and restaurants in 11 EU countries between 2005 and 2012, and modelled energy use based on three variables: energy price, income and climate. The result showed that both income and climate have a significant impact on increasing electricity consumption, while energy prices have no effect on electricity consumption.

Meng et al. [7] proposed adding an improved grey model(1,1) ((GM(1,1)) into the method of residual correction and artificial neural network symbol estimation, and successfully predicted the power consumption in Taiwan. The example results showed that the improved grey prediction model had higher prediction accuracy. Wang et al. [8] considered that power consumption prediction stability is more important than accuracy. Therefore, they proposed a hybrid prediction model based on an improved grey prediction model optimized by a multi-objective ant colony optimization algorithm to improve the prediction stability. Chiang et al. [9] combined a neural network with grey theory to

predict electrical loads. The proposed grey correlation analysis can select high-efficiency influencing factors, which makes the results better than single scheme and statistical autoregressive methods.

Mikayilov et al. [10] used a time-varying coefficient cointegration method to study the correlation between electricity demand and the change of income and price in the time dimension, and proposed policy recommendations for the income and price of electricity consumption in Azerbaijan; Azadeh [11] estimated and predicted the power consumption in an uncertain environment by an algorithm of fuzzy regression analysis consisting of 16 fuzzy regression models, and experimented with historical data from Iran to prove the algorithm's superiority; Mohamed [12] not the author surname—check selected economic and demographic variables as influencing factors to analyze the changing characteristics of New Zealand's annual electricity consumption, and used a multiple linear regression analysis development model to predict the changes in New Zealand's electricity consumption. The experimental results proved that electricity consumption is effectively related to all variables.

Azadeh et al. [13] used a simulation-based comprehensive fuzzy regression time series model to estimate and predict the power demand for seasonal and monthly changes in power consumption in developing countries such as China and Iran. The results of the final example demonstrated the effectiveness and accuracy of the model. Kumar [14] used three time series models, in which the grey Markov model has been used to predict crude oil and oil consumption, the grey model uses rolling mechanisms to predict coal, utility electricity consumption and singularity spectral analysis (SSA) to predict natural gas consumption, and SSA predicts India's conventional energy consumption and compares the results with the Indian Planning Commission's predictions, indicating that these time series models can be considered as a viable alternative to energy consumption prediction; Hussain [15] applied the Holt-Winter and Autoregressive Integrated Moving Average models to time series secondary data from 1980 to 2011 to predict Pakistan's overall and component electricity consumption.

In recent years, many scholars in the field of energy research have studied a large number of predictive machine learning intelligent algorithms. Li et al. [16] used a new meta-heuristic algorithm, the *Drosophila* optimization algorithm (DOA), to determine the values of two parameters of the least squares support vector machine (LSSVM). Based on this, an annual power load was constructed. Meng et al. [17] considered the importance of monthly power consumption forecasting for planning power generation and distribution of electric utilities, using discrete wavelet transform to derive three relatively simple sequences, which are constructed in ascending trend and periodic wave respectively; Kandananond [18] forecasted power demand according to the population of Thailand, GDP, stock index, exports, etc. and compared the performance of the prediction models including autoregressive comprehensive moving average (ACMA), artificial neural network (ANN) and multiple linear regression (MLR). The results showed that the ACMA and MLR models are better than ANN due to their simple structure; Zhao [19] proposed a new hybrid power consumption prediction method, namely the grey model (1,1) (GM(1,1)), which was optimized by the moth flame optimization (MFO) algorithm with rolling mechanism. The example study proved the proposed better performance of the method. Ma et al. [20] proposed a new power load forecasting method based on fuzzy reasoning and artificial neural network, and verified that the method can improve the prediction accuracy. Liang [21] proposed a prediction model based on the improved fruit fly algorithm to optimize the parameters of the support vector machine (SVM) to improve the accuracy of the prediction. Wang et al. [22] used differential evolution algorithm-optimized support vector regression to predict power consumption.

Renewable energy grid-connected operation has been studied by many scholars. The influencing factors affecting the integration of renewable energy can be divided into technology, operational cost-effectiveness, and power system planning.

Eissa, et al. [23] considered various renewable energy sources, as well as information and communication technology components, and the grid will become more complex. Based on this, a wide area monitoring system (WAMS) based method was developed to solve technical difficulties in accessing renewable energy to smart grids; Dominguez-Navarro et al. [24] thought that renewable energy and storage systems can effectively increase the profitability of electric vehicle (EV) and

reduce the high energy required by the grid. The Monte Carlo method was used to simulate electric vehicle demand and renewable energy generation. Denholm et al. [25] argued that a large number of variable power generation (VG) resources can improve system flexibility by changing support technologies such as grid operation and deployment of energy storage, and simulated three different proportions of wind and solar power generation scenarios; Bornapour [26] proposed a stochastic model for coordinated scheduling of renewable heat units for renewable energy power dispatching, considering proton exchange membrane fuel cells, wind and photovoltaics, etc., and then using the improved teaching-learning-based optimization (MTLBO) algorithm to solve the problem; Emanuele et al. [27] believed that the integration of variable renewable energy (VRE) improved the flexibility and dispersion of power systems, and that electric vehicles (EVs) can increase the integration of VREs and capture the potential advantages of power systems; Angenendt et al. [28] considered the economics of grid-connected economics from the economics of residential photovoltaic cell energy storage. The strategy to evaluate operational strategies by simulating DC-coupled PV and battery systems was expected to reduce power leveling costs by 12%.

In recent years, more and more scholars have applied the principle of signal decomposition to the fields of prediction and decision-making, and used the time series decomposition technique of signal science to decompose the original signal sequence to form several sub-sequences. Among them, empirical mode decomposition (EMD), wavelet signal decomposition etc. are commonly used by scholars. An et al. [29] used EMD to decompose wind farm power into several inherent mode function (IMF) components and a residual component, using different models to predict each component. The results showed that the decomposed results were more suitable for short-term wind farms; Kim et al. [30] used feature decomposition for deep learning to decompose the load profile into a weekly load profile and then trained the long-term short-term memory network model with three-step regularized three-dimensional input data to predict the demand side load. The experimental results show the validity of the proposed model; Pang et al. [31] analyzed the original vibration signal of the rotor by the improved singular spectral decomposition (ISSD) and Hilbert transform (HT) joint time-frequency method. Xie et al. [32] proposed a method based on improved set empirical mode decomposition (MEEMD) to decompose deformation time series into a series of subsequences with significantly different complexity, and then established an approximation for each new subsequence; Xiao et al. [33] obtained the eigenmode function (IMF) by improving empirical mode decomposition (IEMD). The Particle Swarm Optimization (PSO) algorithm was used to optimize the LSSVM algorithm to accurately identify the misalignment type of the large doubly-fed wind turbine (DFWT); Zhao et al. [34] used the correlation coefficient analysis method to calculate and determine three improved IMFs, so that they were close to the original signal, and then used the multi-scale fuzzy entropy to calculate the entropy of the IMF.

With the wide application of intelligent algorithms, more and more scholars apply intelligent algorithms to forecasting and decision making in various fields. Extreme learning machine (ELM) is one of the most widely used intelligent prediction algorithms with high accuracy and applicability. Aiming at the parameter optimization of ELM and the optimization of single hidden layer activation function, many scholars have conducted research. Li et al. [35] proposed using the kernel function in SVM instead of the connection weight matrix between the original hidden layer and the output layer in the ELM algorithm; li et al. [36] proposed a new type of Laplacian bipolar to learning machine (LapTELM), enabling LapTELM to fully exploit the benefits of large numbers of unlabeled samples while preserving the learning power and efficiency of the double extreme learning machine (TELM); Fang et al. [37] introduced a ELM's multimodal data hierarchical framework which demonstrated that ELM has better learning efficiency than gradient-based multimodal deep learning methods; shang et al. [38] developed a classification and regression tree (CART) based on A new predictive model of the Extreme Learning Machine (EELM) method, which improved the accuracy of PM2.5 concentration prediction per hour. Ming et al. [39] proposed two parallel changes of ELM including local data and

model parallel ELM (LDMP-ELM) and global data and model parallel ELM (GDMP-ELM), and used parallel technology to improve the parallelism and scalability of ELM.

In order to accurately predict the amount and proportion of China's renewable energy terminal power consumption, this paper proposes a combined forecasting model. We optimized original ELM model with Inverse Square Root Linear Units (ISRLU) activation function which named improved extreme learning machine (IELM) algorithm. Based on EMD and bacterial foraging algorithm (BFO), the combined EMD-BFO-IELM forecasting model is proposed to predict the amount and proportion of renewable energy power consumption in China. The main contents of the article are as follows: The second part introduces the mathematical principles of EMD, BFO and IELM algorithm and the flow chart of the overall forecasting model is put forward. In the third part, the proposed EMD-BFO-IELM model is applied to predict China's renewable energy terminal power consumption. By comparing with the IELM, BFO-IELM, the accuracy and training speed of EMD-BFO-IELM model has been proved better than others. Finally, we apply this model to predict China's renewable energy terminal power consumption from 2018 to 2030 and mining its change rule. The fourth part presents more discussions and forward-looking conclusions.

2. Forecasting Model Including Materials and Methods

2.1. E.mpirical Mode Decomposition

The empirical mode decomposition (EMD) algorithm is a form of converting an irregular frequency wave into a plurality of waves and residual waves of a single frequency. The basic principle of EMD is to determine the "instantaneous equilibrium position" by using the average of the upper and lower envelopes to extract the intrinsic eigenmode function (IMF), that is, to decompose a complex signal into a finite eigenmode function and margin, each IMF The component contains local characteristic signals of different time scales of the original signal, so as to preserve the characteristics of the original data as much as possible. IMF is orthogonal to each other, has good performance, and can express the original signal very well. The residual wave is also an extremely smooth trend sequence. Therefore, EMD can linearize and smooth the non-stationary data sequence.

The specific steps of the EMD algorithm are as follows:

(1) Firstly, determine all local maxima and minima points on the original signal $s_1(t)$; then, use the cubic spline interpolation function to determine the upper and lower envelopes $u_1(t)$, $v_1(t)$. Finally, calculate the average curve of the upper and lower envelopes as:

$$m_1(t) = \frac{u_1(t) + v_1(t)}{2} \tag{1}$$

Find the difference between the original signal and the envelope mean:

$$h_1(t) = s_1(t) - m_1(t) \tag{2}$$

In case $h_1(t)$ does not meet the two conditions of IMF, one needs to put $h_1(t)$ as the original signal, repeat the above steps to get:

$$h_{11}(t) = h_1(t) - m_{11}(t) \tag{3}$$

This step operates k times until $h_{1k}(t)$ becomes an IMF, called the first-order IMF, which is recorded as:

$$c_1(t) = h_{1k}(t) \tag{4}$$

(2) Subtracting $c_1(t)$ from the original signal yields the first-order residual signal $r_1(t)$. Considering that the first-order residual signal $r_1(t)$ still contains longer-period components, the same filtering is required for $r_1(t)$. Thus, the second order IMF, ..., the *n*-th order IMF and the second order

residual signal, ..., the *n*-th order residual signal are sequentially obtained as well. This process can be expressed as:

$$
\begin{cases}
r_1(t) - c_2(t) = r_2(t) \\
\quad\quad\quad . \\
\quad\quad\quad . \\
r_{n-1}(t) - c_n(t) = r_n(t)
\end{cases}
\tag{5}
$$

When $r_n(t)$ becomes a monotonous function, the filter ends. Then Equation (6) is obtained:

$$
s_1(t) = \sum_{i=1}^{n} c_i(t) + r_n(t)
\tag{6}
$$

In the formula, $r_n(t)$ represents the average trend of the signal, which means that the initial sequence is equal to the sum of several intrinsic mode functions and residual terms.

2.2. Improved Learning Function of theExtreme Learning Machine

ELM is a feedforward neural network learning algorithm. The algorithm has a good global search ability, and once the parameters of the algorithm are confirmed, no adjustment is needed during the training. Compared with other machine learning algorithms, ELM has the advantages of high learning efficiency and good generalization performance.

Training sample in this article (x_i, y_i) in the middle, let ELM have u input nodes, L hidden layer nodes, q output nodes, and the activation function is $g(x)$, then $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]$ the network output can be expressed as:

$$
\sum_{j=1}^{L} \beta_j g(\omega_j x_i + b_j) = y_i, i = 1, 2, \ldots, n
\tag{7}
$$

In the formula, ω_j a weight vector representing the *j*-th implicit node and the input node, b_j represents the threshold of the *j*-th hidden layer node, β_j represents the weight vector between the *j*-th implicit node and the output node

The activation function $g(x)$ is a key factor affecting network performance in ELM. Appropriate activation functions can improve the accuracy and generalization of ELM. In the current research, the Sigmoid function is commonly used as the traditional hidden layer activation function in ELM, which is a discriminant function using two-sided suppression. However, when the generalized Hop-world problem is encountered, the approximation value of the value function is monotonic, then the double-side suppression method will increase the waste operation [40]. At this time, unilateral suppression is needed to complete the value discrimination. In addition, a modified linear function is widely used in the field of deep learning as a new type of activation function [40], and its rectified linear unit (ReLU) is defined as:

$$
g(x) = \max(0, x)
\tag{8}
$$

The ReLU function is simple in form, fast in operation, and more generalized than Sigmoid, but the sparsity of the function will reduce the predictive ability of the function and reduce the average performance of the network. In this paper, an inverse square root linear units (ISRLU) is proposed as the activation function of the ELM algorithm, which is a nonlinear smooth representation of ReLU. The ISRLU function is nonlinear continuous and differentiable, and is closer to the biological activation model than the Sigmoid function, which can better avoid the forced sparsity of ReLU and improve the average performance of the network. In this study, the ISRLU function is selected as the activation function of ELM, and the function is defined as:

$$
g(x) = \begin{cases}
\frac{x}{\sqrt{1+\alpha x^2}}, & x < 0 \\
x, & x \geq 0
\end{cases}
\tag{9}
$$

In the formula, α is the parameter to the ISRLU function.

2.3. Bacterial Foraging Algorithm

The Bacterial Foraging Algorithm (BFOA) was proposed by the scholar Passino in 2002 to classify biopsies based on the foraging behavior of *E. coli* in the human large intestine. Passino mentioned in the initial publication of the algorithm that the algorithm can be used in the field of automatic control and adaptive control of automatic locomotive. After several years of research, the bacterial foraging algorithm is now applied to more fields: power system, control engineering, power forecasting, etc. The solution process for specific problems is generating the initial solution population, calculating the value of the evaluation function, using the interaction and mechanism of the group to iteratively optimize, and implementing the three main operators of chemotaxis, reproduction and migration to achieve the optimal solution. The general bacterial foraging algorithm is divided into four processes: chemotaxis operation, aggregation operation, copy operation, and migration operation.

2.3.1. Chemotaxis Operation

The chemotaxis operation consists of two basic actions: flipping and swimming. When the bacteria encounters a favorable area with good nutrition, they will continue to swim. If the area where the adverse concentration is not as good as the previous step, it will flip and change the direction of swimming. Each set of bacteria moves to a new area to represent a set of optimization parameters. Calculate the individual fitness at this point to derive the value $j(i, j, k, l)$ which is used as an indicator of the next move formula. i represents the number of individuals, and j represents the chemotaxis operation, k represents the copy operation, l represents the migration operation. The i-th bacterial trending operation is expressed as follows:

$$\theta^i(j+1, k, l) = \theta^i(j, k, l) + c(i)\frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}} \tag{10}$$

In the formula, $c(i)$ indicates the step size of the bacteria, $\Delta(i)$ is the direction vector of the random direction of the element which value is a random number of [–1,1].

2.3.2. Aggregation Operations

In the process of searching for food in the bacillus community, there is an interaction force between a bacillus and other bacilli, that is, gravity and repulsion, and the gravitation makes the individual be "between". The behavior of the group is held, and the repulsion allows the individual to have a position, gain energy, and maintain life. There is an attractive function in the bacterial foraging algorithm to describe this aggregation operation, whose definition function is:

$$
\begin{aligned}
j_{cc}(\theta, p(j, k, l)) &= \sum_{i=1}^{s} j_{cc}^i(\theta, \theta^i(j, k, l)) \\
&= \sum_{i=1}^{s} [-d_{attract} \exp(-\omega_{attract} \sum_{m=1}^{p} (\theta_m - \theta_m^i)^2)] + \sum_{i=1}^{s} [h_{repellant} \exp(-\omega_{repellant} \sum_{m=1}^{p} (\theta_m - \theta_m^i)^2)]
\end{aligned} \tag{11}
$$

In the formula: $d_{attract}$ is the depth at which the bacteria release the substance, $\omega_{attract}$ is a measure of the width of the substance that attracts bacteria. $h_{repellant}$ is the height of rejection, $\omega_{repellant}$ is the width of the exclusion. These parameters are mainly selected according to the characteristics of food richness. This aggregation behavior is only to accelerate the convergence rate of the bacterial foraging algorithm, but the application process is more complicated. Passino introduced the description when publishing the algorithm. This step can be omitted. Learning and researchers who later explored the bacterial foraging algorithm rarely applied this step to the algorithm, so there is no discussion of what principles these parameters should follow in this paper.

2.3.3. Copy Operation

The process of biological evolution has a rule of survival of the fittest. In the process of bacterial foraging, the adaptability is strong, and the weak will be eliminated. After the chemotaxis operation is completed, it is concluded that all individual health values (sum of all function values) and better healthy bacterial positions represent better optimization parameters. In order to speed up the search, bacteria need to search in these good positions, then the difference is poor. The location will be eliminated, this is the copy operation given by k with l for $i = (1, 2, \ldots s)$:

$$j_{health}^i = \sum_{j=1}^{N_c+1} j(i, j, k, l) \tag{12}$$

The formula calculates each individual's health value which is ranked from large to small, half of the health is better, so half of the poor health is eliminated. The formula is $s_r = \frac{s}{2}$. The surviving bacteria split into two at the same position, so that the total number of bacteria can be kept constant, and the optimal position of high nutrition can be found more quickly, which improves the efficiency of bacterial convergence.

2.3.4. Migration Operation

The migration operation is based on a certain set probability, and each bacterium will randomly generate a random number *rand*(). If the probability of a given bacillus' migration is greater than the random number, the bacillus will be eliminated, and the bacillus will randomly generate a new bacillus in the solution area to keep the total bacterial population unchanged. Randomly generated individuals may be closer to the global optimal position, solving the situation of entering premature and local optimal stagnation in the chemotaxis operation.

Due to space limitations, only the flow chart of the migration operation in the bacterial foraging algorithm is shown in Figure 1:
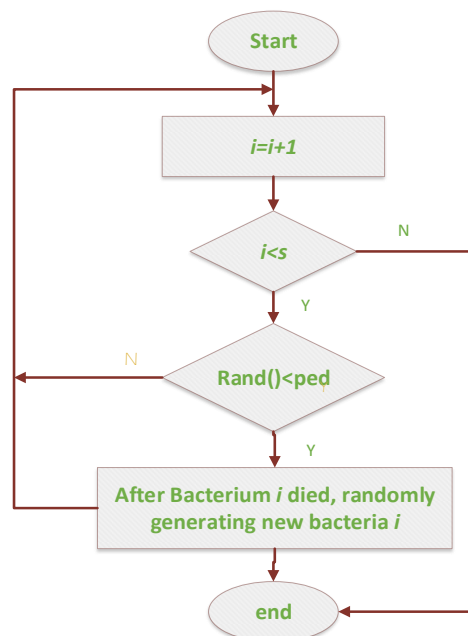

Figure 1. Migration operation flow chart.

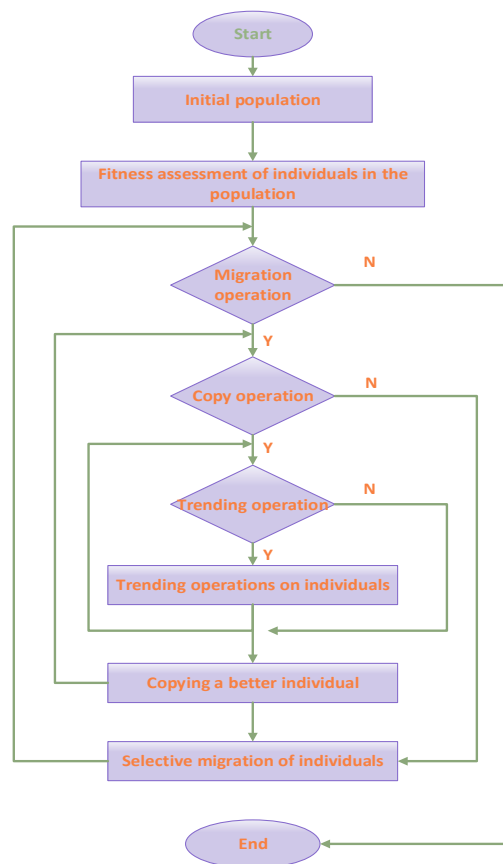The flow chart of the bacterial foraging process is shown in Figure 2:

Figure 2. Flow chart of the bacterial foraging process.

2.4. Renewable Energy Power Consumption Forecasting Model Design Process

Because the initial parameters of the traditional ELM network model are random, and the suitability of the activation function is not considered, the bacterial foraging algorithm can solve the problem of optimal chattering and precocity, and determine the optimal weight and threshold. The ISRLU function improves the generalization of ELM. The original data for the influencing factors is easy to homogenize, and the different scale information features of the data cannot fully discover the time-frequency characteristics of the time series data and affect the performance of the forecasting model. Based on the above reasons, this paper combines the three algorithms of EMD, BFO and improved IELM to propose a new forecasting model of renewable energy terminal power consumption. The overall forecasting steps are as follows:

(1) Time series data decomposition. Decompose x(t) to obtain IMF components and one residual r_n with EMD.

(2) Construct training and test sample sets. In each IMF component, the input and output of each component training sample set and test sample set are constructed.

(3) Construct an optimized limit learning machine training and forecasting model for each component. In the bacterial foraging algorithm, the fitness function in the bacterial foraging algorithm is calculated; the initial population size and the maximum evolution algebra maxgen are set, and the genetic operations such as selection, improved crossover and mutation are performed on the individuals in the population, and finally the global excellent fitness; use optimal fitness to obtain optimal weight a_{best}, and threshold b_{best};

(4) Set the activation function of the ELM network to the ISRLU function, and then calculate the output matrix h and output weight of ELM β with a_{best} and b_{best}. Determine the IELM network structure; use the BFO algorithm to iteratively optimize parameters of each IELM model. The IELM

fitting prediction model of the optimal parameters is established in each IMF component and remainder r_n to obtain the forecasting results of each component:

$$f = \frac{1}{n}\sum_{t=1}^{N}\left(y_i(t) - \widehat{y_i}(t)\right)^2 \tag{13}$$

$$H(a_1,\ldots,a_l,b_1,\ldots,b_l,x_1,\ldots,x_n) = \begin{bmatrix} g(a_1x_1 + b_1) & \ldots & g(a_1x_1 + b_1) \\ . & \ldots & . \\ . & \ldots & . \\ g(a_1x_n + b_1) & \cdots & g(a_1x_n + b_1) \end{bmatrix}_{n\times l} \tag{14}$$

$$\widetilde{\beta} = H^+Y \tag{15}$$

In the formula,$y_i(t)$ is the actual value at time t, $\widehat{y_i}(t)$ is the predicted value at time t, H^+ is the generalized inverse matrix of the output matrix H.

(5) Output of prediction results. The predicted results of each IMF component and remainder r_n are summed to obtain the final forecasting result of China's renewable energy terminal power consumption.

The algorithm flow of the renewable energy terminal power consumption based on EMD-BFO-IELM proposed in this paper is shown in Figure 3.
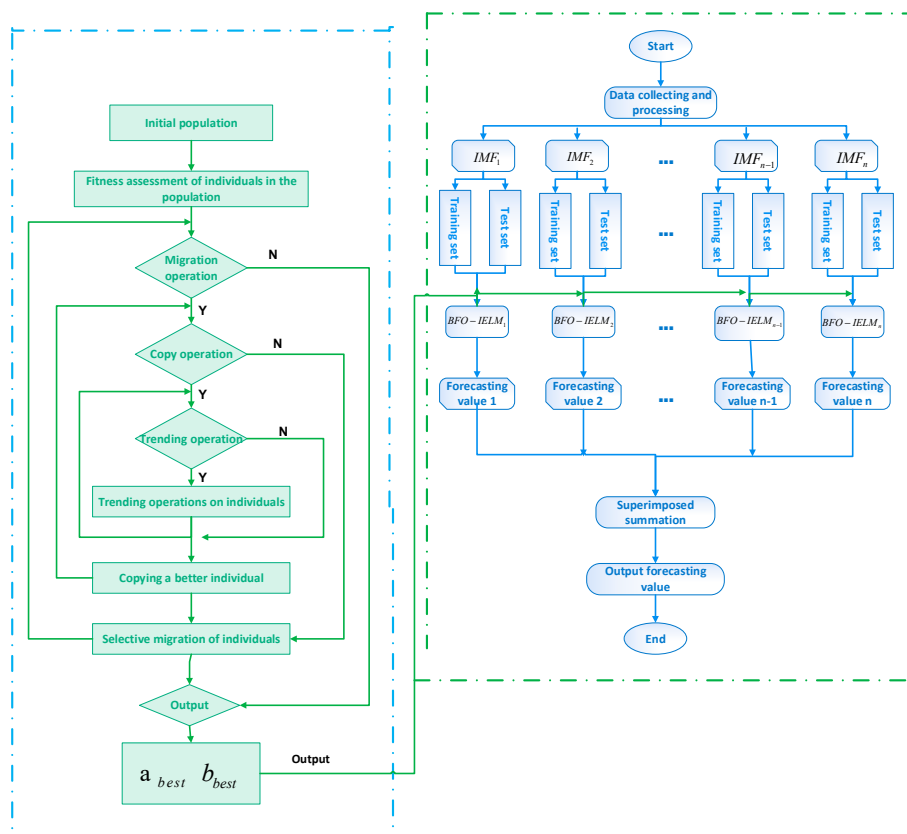

Figure 3. EMD-BFO-IELM algorithm flow chart.

3. Empirical Analysis Including Results and Discussions

3.1. Influencing Factors Screening for Model Input

According to literature research and the China Statistical Yearbook "Energy Consumption" section data analysis, this paper selects gross national product GDP, total population, primary industry GDP, secondary industry GDP, tertiary industry GDP, urbanization rate, fixed investment in renewable energy industry, renewable energy consumption, total installed capacity of power generation, total installed capacity of wind power generation, total installed capacity of solar power generation, total investment in fixed assets of power transmission and distribution, total length of cross-provincial and cross-regional high-voltage transmission lines are used as the primary set of factors affecting the renewable energy terminal power consumption. The grey relational analysis (GRA) is used to select 10 main influencing factors with grey correlation degrees greater than 0.5 (the data source is the 2018 China Statistical Yearbook). Since there is no officially published data on renewable energy terminal power consumption in China, the renewable energy terminal power consumption in this paper is approximated by the difference between the total consumption of social energy and the amount of main fossil fuel:

$$\sum E^r = \sum E^t - \sum E^f \tag{16}$$

$\sum E^r$ represents renewable energy terminal power consumption, $\sum E^t$ represents the total power consumption, $\sum E^f$ indicates the main fossil fuel power consumption.

3.2. Grey Relational Analysis (GRA)

Through the GRA [41–45], this paper selects 10 main influencing factors that are related to the renewable energy terminal power consumption, which aims to reduce the amount of information input of the EMD-BFO-IELM model, and improve the accuracy of prediction. The calculation steps of GRA has been introduced in Appendix B part. Finally, a set of factors have been obtained, as shown in Table 1:

Table 1. Factors affecting renewable energy terminal power consumption.

Influencing Factor	Grey Relational Value
Gross national product GDP(A1)	0.845
Total population(A2)	0.754
Secondary industry gross production(A3)	0.712
Total renewable energy consumption(A4)	0.694
Total installed capacity(A5)	0.644
Total installed capacity of grid-connected thermal power generation(A6)	0.638
Total investment in fixed assets of power transmission and distribution(A7)	0.619
total length of cross-provincial and cross-regional high-voltage transmission lines(A8)	0.601
Grid-connected wind power generation capacity(A9)	0.532
Grid-connected photovoltaic power generation total installed capacity(A10)	0.519

3.3. Data Normalization

In order to eliminate the dimensional difference between the data of different metrics, and due to the fact that some data is missing in some years, China only started reporting installed wind power capacity statistics in 2011, and photovoltaic installed capacity statistics start from 2014, therefore, for data availability first, we must make some rational adjustments to supplement the data. We normalize all the corrected data, and use the Z-score data normalization method to standardize the N sets of data in the data sets of the M indicators:

$$Z_{mn} = \frac{x_{mn} - \bar{x}_n}{\sigma}, (m = 1, 2, \ldots, j; n = 1, 2, \ldots, j) \tag{17}$$

Z_{mn} is the normalized data, x_{mn} is the corrected data, $\overline{x}_n$ is the average of x_{mn}, σ is the standard deviation of x_{mn}. Since the installed capacity of thermal power generation is a negative index, in order to achieve the forwardization of data, the inverse index forward processing method is adopted:

$$X_{mn} = \frac{Z_{mn} - \min}{\max - \min} \tag{18}$$

X_{mn} refers to the data after the forwardization, max is the maximum value of the normalized data. min refers to the minimum value of the normalized data. The numerical values of the main influencing factors set data normalized are shown in Table 2. Due to space limitations, only the data for 2007–2017 is displayed:

Table 2. Influencing factors after normalization.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
2007	−1.37995	−1.561225	−1.4862261	−1.72559	−1.58673	0.91	−1.18795	0.572503	−1.19164	−0.9308
2008	−1.10982	−1.232771	−1.13817201	−1.58573	−1.26339	0.906028	−0.15134	0.932666	−1.07049	−0.88273
2009	−0.94777	−0.951238	−0.9857299	−1.13508	−0.91337	0.817376	−0.11702	1.093559	−0.94934	−0.78659
2010	−0.59727	−0.622784	−0.51627398	−0.56257	−0.62003	0.719858	−0.98715	1.750708	−0.75551	−0.73852
2011	−0.93837	−0.341251	−0.94795029	0.282709	−0.32668	0.625887	−0.82583	−0.65618	−0.56168	−0.49817
2012	0.100671	−0.012797	0.27485777	0.528628	−0.01333	0.521277	−0.39848	−1.48869	−0.19824	−0.16169
2013	0.401454	0.2687355	0.533220369	0.821751	0.330014	0.443262	−0.27663	0.661363	0.116741	−0.06555
2014	0.668542	0.5971901	0.766256784	0.8608	0.706697	0.349291	−0.02606	−0.35684	0.601325	0.222867
2015	0.915616	0.9256446	0.832941161	0.829738	0.756699	0.218085	0.064905	−1.2227	0.625554	0.270936
2016	1.214515	1.3010212	1.049438179	0.746129	1.25672	0.10461	1.832636	−0.70489	1.449346	0.943907
2017	1.67238	1.6294758	1.617638017	0.939212	1.673404	0.10213	2.07291	−0.58149	1.93393	2.626333

3.4. Forecasting Renewable Energy Terminal Power Consumption in China Based on EMD-BFO-IELM Model

This paper decomposes the original historical renewable energy terminal power consumption sequence by EMD. The data information of 1990–2017 was input into the EMD model, and four IMFs and the Rseidual are obtained. The decomposition results are shown in Figure 4:
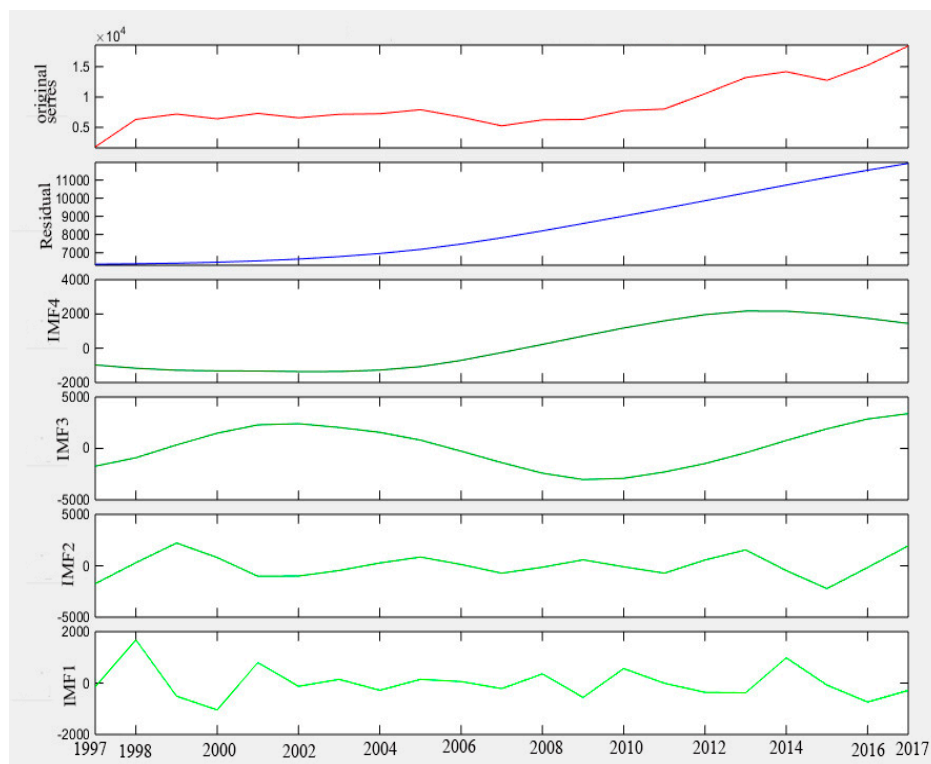

Figure 4. Decomposition results of renewable energy terminal power consumption.

From Figure 4, it can be found that the time series of renewable energy terminal power consumption has obvious multi-scale features, and the four IMF components present information of high and low variations with different fluctuation scales. Among them, the frequency of IMF1 is relatively high, which can reflect the random noise information of the original time series data, and the residual frequency is low, the change is stable, and the trend information of the renewable energy terminal power consumption timing can be reflected, showing the renewable energy terminal power consumption. Regarding overall change characteristics, the number of IMFs obtained by EMD decomposition is related to the characteristics of the original data time series itself. Through the nature of these IMF components, we can understand the practical significance of EMD decomposition. The original renewable energy terminal power consumption sequence and each of the IMF components decomposed from the original data sequence are independent of each other, and the respective IMF components are mutually orthogonal. The last term obtained by decomposition is the remainder of the original sequence after all the IMF components have been removed. In fact, it represents the average trend of the original renewable energy terminal power consumption. The various IMF components are trained and tested with res. The main parameters of the BFO-IELM model are shown in Table 3.

Table 3. Main parameters of the BFO-IELM model.

Parameter	Value
regularization coefficient C	7
number of node in hidden layer	150
a_{best}	0.82
b_{best}	0.91

The normalized values of the main influencing factors derived from 3.3 and the values of each series are used as inputs to the model training set (1997–2017), and the predicted results of the test sets are shown in Figure 5.

In order to verify that the proposed model has high prediction accuracy, we also input the sample data into IELM and BFO-IELM respectively, and the predicted results are shown in Figure 6 and the detailed data are shown in Table A2 of Appendix A:
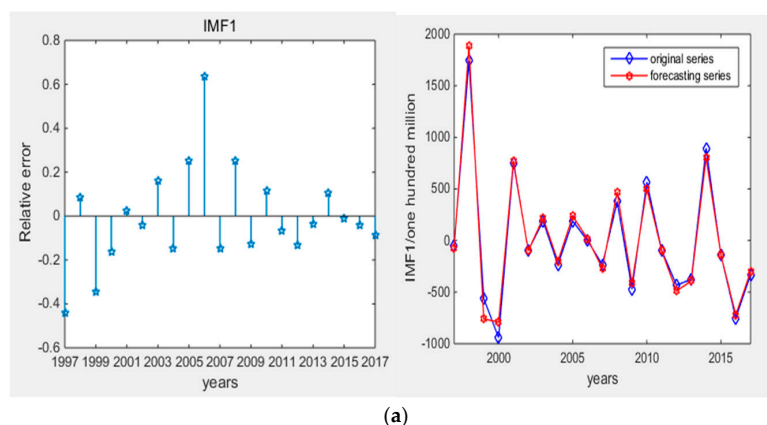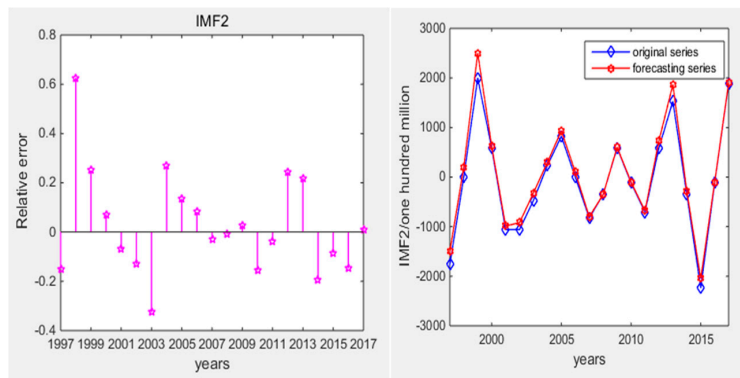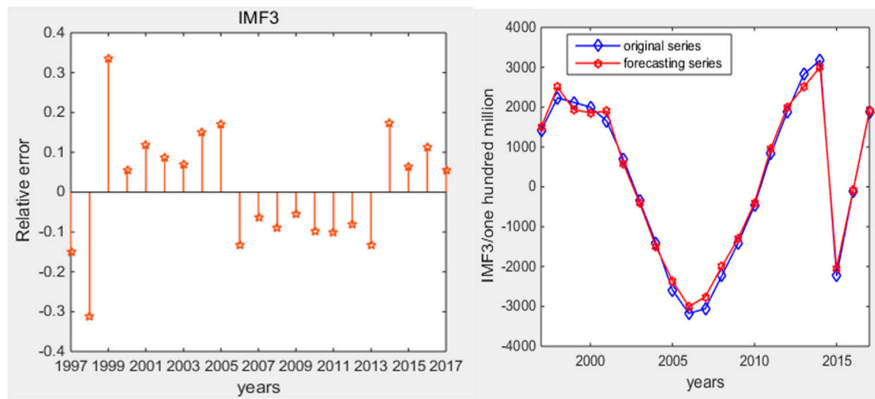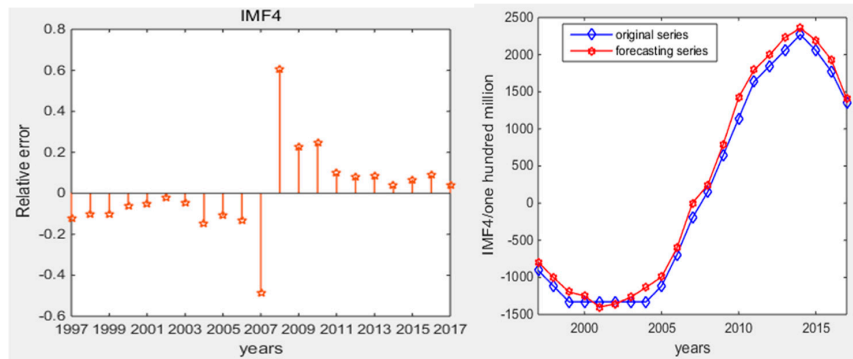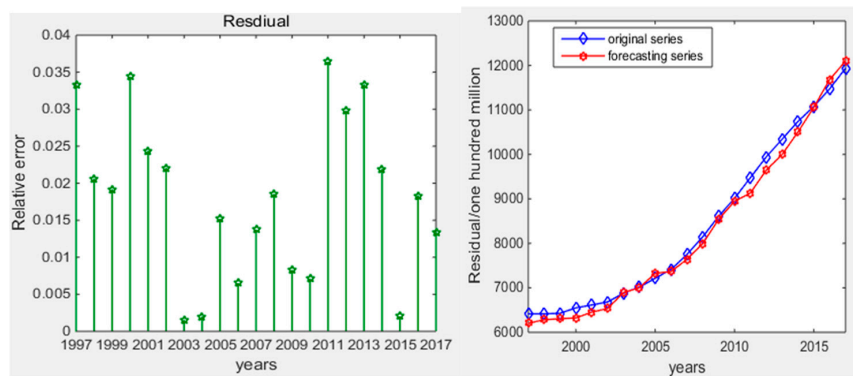

(**a**)

Figure 5. *Cont*.

Figure 5. The predicted results of the test sets: (**a**) the predicted results of IMF1; (**b**) the predicted results of IMF2; (**c**) the predicted results of IMF3; (**d**) the predicted results of IMF4; (**e**) the predicted results of Resdiual.
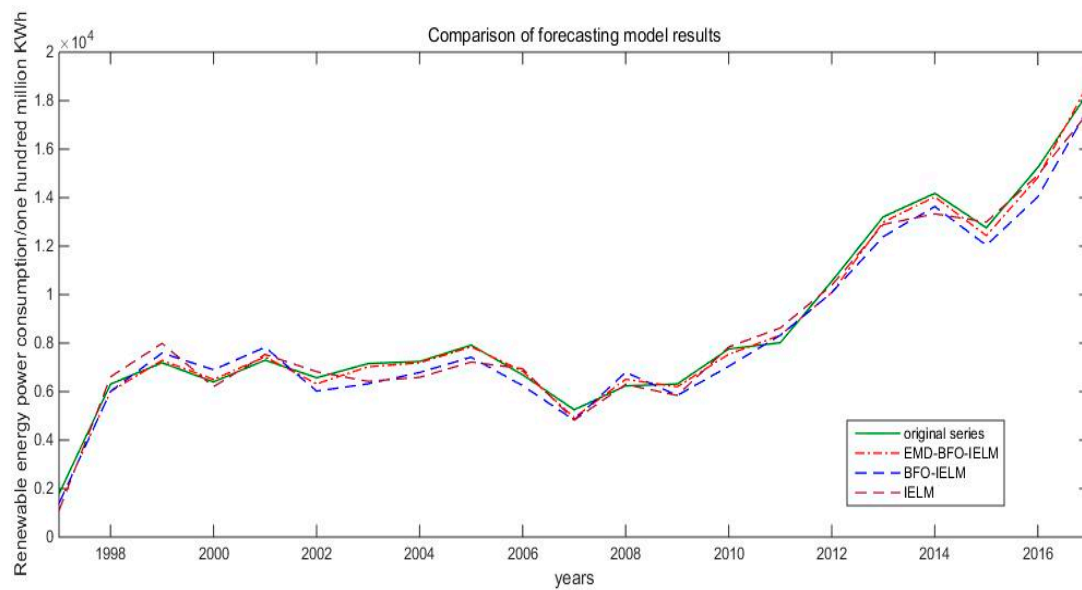
Figure 6. Comparison of prediction results for the three models.

In order to objectively compare the accuracy of a variety of models, common statistical indicators including RMSE, r^2, and mre are adopted and the index calculation formula is as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{q}_i - q_i)^2} \tag{19}$$

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{q}_i - \bar{q}_i)^2}{\sum_{i=1}^{n}(q_i - \bar{q})^2} \tag{20}$$

$$\text{mre} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{q}_i - q_i|}{q_i} \times 100\% \tag{21}$$

Among them, $\hat{q}_i$ is a predicted value, $\bar{q}_i$ is a sample mean, n is a sample number.

The calculation results for the three models are compared as follows:

As can be seen from Table 4 and Figures 6 and 7, this paper proposes that the EMD-BFO-IELM model is reliable and can achieve a good predictive effect on renewable energy terminal power consumption. Through comparative analysis, we can find:

Table 4. The calculation results of the three models.

Model	RMSE	r^2 (100%)	mre
IELM	0.5968	98.3	7.42
BFO-IELM	0.5051	99.2	6.68
EMD-BFO-ELM	0.2588	99.8	3.54

(1) The fitting effect of the nonlinear time series combined prediction model after EMD noise reduction is obviously better than the fitting effect of single mode prediction.

(2) As a complex multi-combination prediction model, EMD-BFO-IELM can realize the complementary advantages of different algorithms.The EMD-BFO-IELM model proposed in this paper, through the non-stationary time series of decomposition and denoising, makes more timing rules eliminate the confusion of the original data.The model continuously optimizes the parameters of the IELM algorithm through BFO, which improves the prediction accuracy and achieves a reasonable choice of parameters.Therefore, the prediction model proposed in this paper combines multiple

prediction methods to make up for the shortcomings of other single algorithm models and shows strong generalization ability and stability.
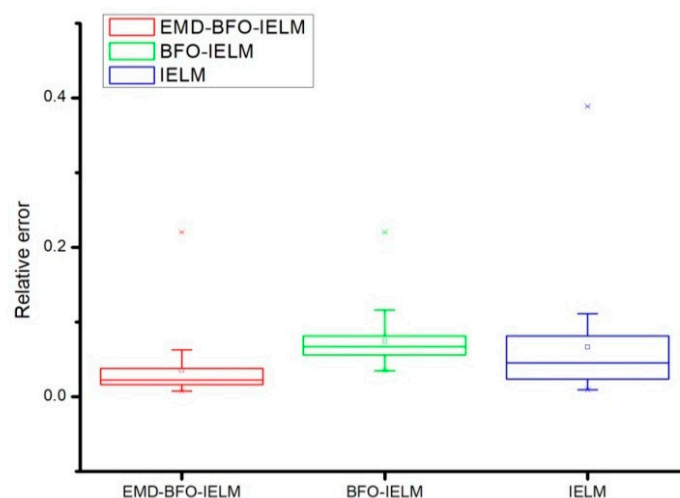

Figure 7. Relative error box plots for the three models.

Through the above analysis, the example proves that the EMD-BFO-IELM renewable energy terminal power consumption prediction model proposed in this paper is both practical and effective.

The time spent on training and prediction is also one of the important indicators for measuring intelligent machine algorithms. The training speeds of the three models (average time of training 100 times) are shown in Table 5:

Table 5. Comparison of training and prediction time for the three predictive models.

Forecasting Model	Average Training Time (s)	Average Forecasting Time (s)
IELM	38	20
BFO-IELM	54	35
EMD-BFO-ELM	84	38

Since EMD-BFO-IELM requires more resources and more computation, it takes longer to calculate during the training process. This is weaker than the other two prediction models, but after the training is completed, the forecasting speed is faster than other two forecasting models, so the weaknesses in training time and speed are almost negligible, and the accuracy is much higher than the other two models, which is practical and superior.

Finally, we applied grey model(1,1)(GM(1,1)) to predict the gross national product GDP, total population, secondary industry GDP, total renewable energy consumption, total installed power capacity, and total grid-connected thermal power generation Installed capacity, total investment in fixed assets of power transmission and distribution, total length of cross-provincial and cross-regional high-voltage transmission lines, total installed capacity of grid-connected wind power generation, total installed capacity of grid-connected photovoltaic power generation from 2018 to 2030, which has been shown in Table A1 of Appendix A. The data were used as input data of EMD-BFO-IELM forecasting model. Meanwhile, the calculation steps to GM(1,1) also has been shown in Appendix B. Finally, we calculate the amount of China's renewable energy terminal power consumption from 2018 to 2030, which has been shown in Figure 8:

This paper comprehensively uses the predicted values of China's total power consumption from 2018–2030 obtained in [46,47], combined with the predicted renewable energy terminal power consumption, and finally obtains China's future terminal power renewable energy terminal proportion, as shown in Figure 9:
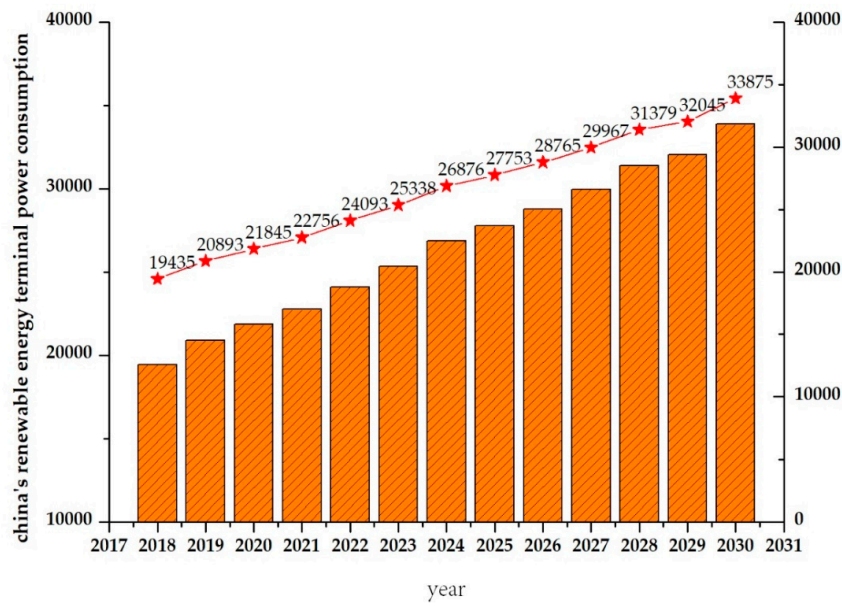
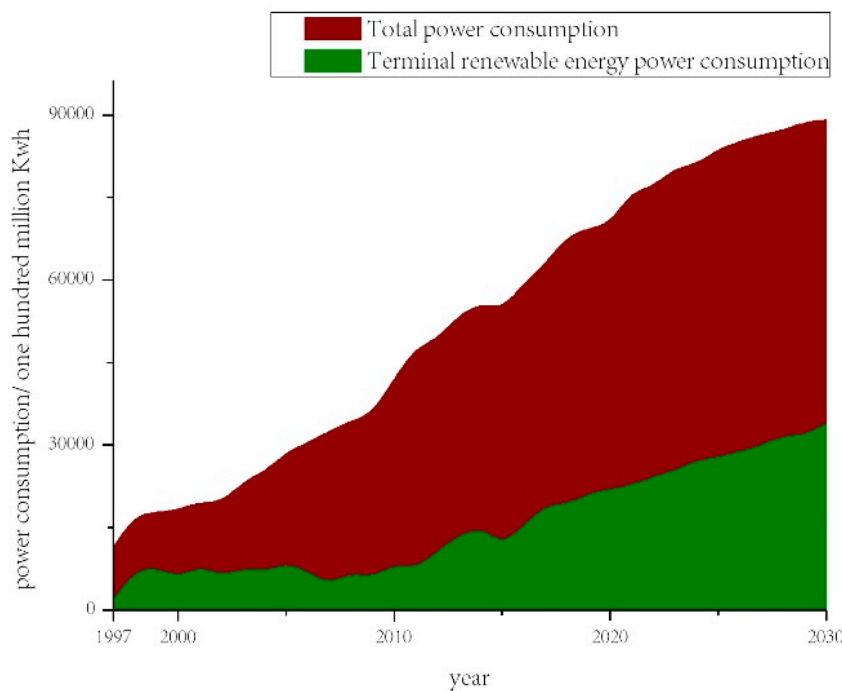**Figure 8.** Output result of the EMD-BFO-IELM forecasting model.

Figure 9. China's renewable energy terminal power consumption ratio chart.

According to the trend in the figure, China's renewable energy terminal power is gradually increasing, and is expected to break through 38% in 2030.

4. Conclusions

In this paper, the grey relational analysis (GRA) theory is applied to screen the influencing factors affecting China's renewable energy terminal power consumption. On this basis, a new EMD-BFO-IELM renewable energy terminal power consumption forecasting model is proposed. Firstly, we use EMD to decompose and denoise the data of the original historical renewable energy terminal power consumption in series, and remove the noise sequence, which improved the quality of the original data and successfully increased the data of training sets and test sets. Therefore, the data series

met the data-level requirements of the machine intelligence algorithm, and successfully realized the realization of the machine intelligent algorithm prediction of China's renewable energy terminal power consumption. Then, we use BFO algorithm to optimize the parameters of the ELM algorithm including optimal weight a_{best}, and threshold b_{best}. This novel BFO-IELM forecasting model is applied to predict the sub-sequences after EMD denoising. Finally, we reconstruct the prediction series and superimpose the predicted values of each subsequence to obtain the prediction results of renewable energy terminal power consumption. In order to show the effectiveness of the proposed forecasting model, some commonly used statistical indicators are used to compare the accuracy of IELM, BFO-IELM and EMD-BFO-IELM models. The comparison results verify that the EMD-BFO-IELM forecasting model proposed in this paper is far better than the others. The generalization ability and robustness are proved by empirical analysis. After analysis, the main reason for the improvement of prediction accuracy is that China's renewable energy terminal power consumption is a complex non-linear prediction problem, and it is a new research field. The lack of historical data makes the traditional methods useless. Because the training process of any intelligent algorithm needs a multitude of data, this paper introduces signal decomposition in the issue of renewable energy terminal power consumption, and the original data volume has been upgraded, increasing the reliability of the training process, so the proposed combined model enables machine intelligence algorithms to be applied to the issue of China's renewable energy terminal power consumption, which is the main advantages of the proposed prediction model. Although some of the computational speed advantages are lost, the novel forecasting model can ultimately achieve higher prediction accuracy. Finally, the EMD-BFO-IELM forecasting model proposed in this paper is applied to predict the amount of renewable energy terminal power consumption in China from 2018 to 2030. The results show that China will realize 3.30 billion kWh of renewable energy terminal power consumption in 2030, and China's renewable energy terminal power consumption ratio will exceed 38%, which indicates that China has great potential of renewable energy terminal power consumption, and can fulfill non-fossil energy development goals in 2030, and achieve the goal of energy production and consumption revolution. A high proportion of renewable energy terminal energy consumption can transform China's current unsustainable energy consume and supply mode, and stop relying on heavy energy consumption of fossil energy. A high proportion of renewable energy terminal energy consumption mode brings China pressures on cost-benefit costs to a certain extent, including grid-connected infrastructure for renewable energy generation, renewable energy generation, and energy storage technology upgrades, which all require large investments. This will lead to an increase in China's overall average cost of power generation in the short term, but the cost will also bring high external benefits, including upgrading and transformation of the power and energy industries, and reducing environmental pollution. From an economic perspective, the transformation of investment in terminal power and energy indicates that a large number of employment opportunities will be created in the future, thus making up for the current reduction of employment opportunities in China's traditional coal industry supply chain. In generally, all results proposed in the paper are in line with China's current active energy innovation strategy.

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

Appendix A

Table A1. The predicted results of the test sets by the proposed model.

Year	IMF1			IMF2			IMF3			IMF4			Res		
	o	f	er	o	f	er	o	f	er	o	f	er	o	f	er
1997	−48.05	−69.3	−0.44225	−1764.5	−1500.3	−0.14973	−1647.3	−1400.3	−0.14994	−905.2	−793.4	−0.12351	6415.6	6202.2	0.033263
1998	1741	1890	0.085583	123	200	0.626016	−941.1	−1234.5	−0.31176	−1117.3	−1003.4	−0.10194	6415.7	6283.4	0.020621
1999	−564.7	−760.3	−0.34638	2000	2500	0.25	235.2	156.3	0.335459	−1329.4	−1196.5	−0.09997	6426.6	6303.4	0.01917
2000	−941.1	−790.2	−0.16034	588.2	630.3	0.071574	1411.3	1490.4	0.056048	−1329	−1249.6	−0.05974	6548.6	6323.3	0.034404
2001	752.9	772.2	0.025634	−1058.4	−984.3	−0.07001	2235.6	2500.4	0.118447	−1329	−1399.4	−0.05297	6615.6	6454.3	0.024382
2002	−96.1	−100	−0.04058	−1058.2	−920.3	−0.13032	2117.3	1934.6	0.086289	−1329	−1356.4	−0.02062	6681.8	6534.3	0.022075
2003	188.2	218.2	0.159405	−470.5	−317.4	−0.3254	2000	1859.5	0.07025	−1329	−1267.5	−0.04628	6881.8	6892.4	0.00154
2004	−235.2	−200.3	−0.14838	235.2	298.5	0.269133	1647.8	1896.5	0.150929	−1329	−1134.5	−0.14635	7014.5	7000.4	0.00201
2005	188.2	235.3	0.250266	823.5	934.5	0.134791	705.8	586.3	0.169311	−1117	−994.5	−0.10967	7214.7	7324.4	0.015205
2006	11	18	0.636364	111	102	0.081081	−352.9	−399.3	−0.13148	−694.1	−603.5	−0.13053	7413.6	7365.4	0.006502
2007	−235.2	−270.4	−0.14966	−823.5	−800.3	−0.02817	−1411.8	−1503.4	−0.06488	−200	−103	−0.485	7746.8	7640.3	0.013748
2008	376.4	470.9	0.251063	−352.9	−350.1	−0.00793	−2588.7	−2356.6	−0.08966	152.9	245.5	0.605625	8145.7	7994.5	0.018562
2009	−470.5	−410.5	−0.12752	588.2	604.5	0.027712	−3176.3	−3002.4	−0.05475	647	794.5	0.227975	8610.8	8539.5	0.00828
2010	564.7	500.3	0.114043	−117.6	−99.3	−0.15561	−3058.7	−2760.5	−0.09749	1141	1423.4	0.247502	9009.5	8945.4	0.007115
2011	−94.11	−100.4	−0.06684	−705.8	−680.3	−0.03613	−2235.6	−2006.7	−0.10239	1635	1794.5	0.097554	9475.7	9130.5	0.03643
2012	−432.5	−490.2	−0.13341	588.2	732.4	0.245155	−1411.6	−1295.6	−0.08218	1847	1996.5	0.080942	9941.7	9645.5	0.029794
2013	−376.4	−390.5	−0.03746	1529.3	1864.4	0.21912	−470.5	−408.4	−0.13199	2058	2230.5	0.083819	10340	9995.3	0.033337
2014	894.1	800.2	0.105022	−352.9	−284.5	−0.19382	823.5	965.4	0.172313	2270	2356.6	0.03815	10739	10503.4	0.021939
2015	−141.1	−139.4	−0.01205	−2235.5	−2045.3	−0.08508	1882.5	2004.5	0.064807	2058	2194.5	0.066327	11079	11056.6	0.002022
2016	−752.1	−720.4	−0.04215	−117.6	−100.3	−0.14711	2823.2	2507.7	0.111753	1776	1934.5	0.089245	11475	11685.4	0.018336
2017	−329.4	−300.4	−0.08804	1882.4	1903	0.010943	3179.4	3004.4	0.055042	1352	1404.5	0.038831	11936	12095.6	0.013371

Table A2. The forecasting results of 3 intelligent machine algorithms.

Year	Original Data	IELM	BFO-IELM	EMD-BFO-IELM
1997	1781	1089	1389	1389
1998	6311	6610	6010	6010
1999	7191	7991	7591	7291
2000	6400	6201	6901	6501
2001	7299	7531	7831	7431
2002	6571	6821	6021	6321
2003	7155	6423	6323	7023
2004	7245	6589	6789	7189
2005	7916	7220	7420	7847
2006	6693	6943	6243	6843
2007	5251	4823	4823	4923
2008	6238	6298	6798	6508
2009	6313	5842	5842	6199
2010	7757	7843	7043	7543
2011	8019	8623	8323	8323
2012	10549	10385	10085	10085
2013	13210	12894	12394	12994
2014	14179	13332	13632	14032
2015	12757	12994	12034	12434
2016	15240	14933	14033	14833
2017	18404	17564	17764	18764

Appendix B

The specific calculation steps of the grey correlation analysis are as follows:

Step 1: Determine the analysis series

Determine the reference and the comparison series that reflects the behavior of the system series. A sequence of data that reflects the behavioral characteristics of the system, called the reference sequence. A sequence of data that is a component of factors affecting system behavior, called a comparison series. Assume that the reference number is $Y = \{y(k)|k = 1, 2, \ldots n\}$, and compare series as

$$X_i = \{x_i(k)|k = 1, 2, \ldots n\}, i = 1, 2, \ldots, m \tag{A1}$$

Step 2: Dimensionlessness of variables

Since the data in the various factor columns in the system may be different in dimension, it is not convenient to compare or it is difficult to get a correct conclusion when comparing. Therefore, in the gray correlation analysis, the data is generally dimensionless:

$$x_i(k) = \frac{X_i(k)}{X_i(l)}, k = 1, 2, \ldots n; i = 1, 2, \ldots, m \tag{A2}$$

Step 3: Calculate correlation coefficient

Calculate correlation coefficient of $x_0(k)$ and $x_i(k)$:

$$\xi_i(k) = \frac{\min_i \min_k |y(k) - x_i(k)| + \rho \max_i \max_k |y(k) - x_i(k)|}{|y(k) - x_i(k)| + \rho \max_i \max_k |y(k) - x_i(k)|} \tag{A3}$$

Let $\Delta_i(k) = |y(k) - x_i(k)|$, and then

$$\xi_i(k) = \frac{\min_i \min_k \Delta_i(k) + \rho \max_i \max_k \Delta_i(k)}{\Delta_i(k) + \rho \max_i \max_k \Delta_i(k)} \tag{A4}$$

$\rho \in (0, \infty)$ is called the resolution coefficient. The smaller ρ, the greater the resolution. Generally, the value range of ρ is (0, 1), and the specific value may depend on the situation. When $\rho \leq 0.5463$, the resolution is best, and we take $\rho = 0.5$ in this paper.

Step 4: Calculate relevance

Because the correlation coefficient is the degree of correlation between the comparison series and the reference sequence at each moment (i.e., each point in the curve), it has more than one number, and the information is too scattered to facilitate the overall comparison. Therefore, it is necessary to concentrate the correlation coefficients at various moments (i.e., points in the curve) into one value, that is, to find the average value thereof, as the quantity representation of the degree of correlation between the comparison series and the reference series. The formula of relevance r_i is as follows:

$$r_i = \frac{1}{n} \sum_{k=1}^{n} \xi_i(k), k = 1, 2, \dots, n \tag{A5}$$

Step 5: Relevance ranking

Sort the relevance by size. If $r_1 < r_2$, the reference sequence y is similar to the comparison sequence X_2. After calculating the correlation coefficient between the $X_i(k)$ sequence and the $Y(k)$ sequence, the average value of each type of correlation coefficient is calculated, and the average value r_i is called the degree of association between $Y(k)$ and $X_i(k)$. The mathematical model of GM(1,1) forecasting model is shown:

The specific calculation steps of the grey model(1,1) are as follows:

Assuming a raw sequence:

$$x^0 = (x_1^0, x_2^0, \dots x_n^0) \tag{A6}$$

where $x_k^0 \geq 0, k = 1, 2, \dots, n$

Adding raw arrays:

$$x^1 = (x_1^1, x_2^1, \dots, x_n^1) \tag{A7}$$

where $x_k^1 = \sum_{i=1}^{k} x_i^0$

Make an equal-valued equal-weight generation sequence of x^1:

$$z^1 = (z_1^1, z_2^1, \dots, z_n^1) \tag{A8}$$

where $z_k^1 = \frac{1}{2}(x_k^1 + x_{(k-1)}^1)$

Therefore, the original differential equation of GM(1,1) forecasting model is:

$$x_k^0 + az_k^1 = b \tag{A9}$$

where a is a development coefficient and b is ash action, so the whitening differential equation of Equation (A9) is:

$$\frac{dx^1}{dt} + ax^1 = b \tag{A10}$$

The least-squares parameter estimates for a and b are:

$$\hat{a} = (a, b)^T = (B^T, B)^{-1} B^T Y \tag{A11}$$

where:

$$
B = \begin{bmatrix} -z_2^1 & 1 \\ -z_3^1 & 1 \\ \vdots & \vdots \\ -z_n^1 & 1 \end{bmatrix}, Y = \begin{bmatrix} x_2^0 \\ x_3^0 \\ \vdots \\ x_n^0 \end{bmatrix}
$$

The time response sequence of the differential Equation (A9) in the GM(1,1) model is:

$$
\hat{x}_{k+1}^1 = (x_1^1 - \frac{b}{a})e^{(-ak)} + \frac{b}{a} \tag{A12}
$$

Similarly, the whitening response function of the whitening equation is:

$$
\hat{x}_{t+1}^1 = (x_1^1 - \frac{b}{a})e^{(-at)} + \frac{b}{a} \tag{A13}
$$

Substituting initial conditions $\hat{x}_1^1 = x_1^1 = x_1^0$ to get:

$$
\hat{x}_{k+1}^1 = (x_0^1 - \frac{b}{a})e^{-ak} + \frac{b}{a} \tag{A14}
$$

The GM(1,1) forecasting model of the original sequence x^0 can be obtained according to the cumulative reduction by Equation (A13):

$$
\hat{x}_k^0 = \hat{x}_k^1 - \hat{x}_{k-1}^1 \tag{A15}
$$

References

1. China-2050-High-Renewable-Energy-Penetration-Scenario-and-Roadmap-Study-Executive-Summary. Available online: http://www.cnrec.org.cn/go/AttachmentDownload.aspx?id=\protect\T1\textbraceleft314648b5-6dcd-43f8-9903-db1ac953cc55\protect\T1\textbraceright (accessed on 30 October 2018).
2. Promoting the Development of High Proportion of Renewable Energy. Available online: http://www.creei.cn/portal/article/index/id/23793/cid/6.html (accessed on 28 October 2018).
3. Meng, M.; Wang, L.; Shang, W. Decomposition and forecasting analysis of China's household electricity consumption using three-dimensional decomposition and hybrid trend extrapolation models. *Energy* **2018**, *165*, 143–152. [CrossRef]
4. Akay, D.; Atak, M. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. *Energy* **2007**, *32*, 1670–1675. [CrossRef]
5. Escoto Castillo, A.; Sánchez Peña, L. Diffusion of Electricity Consumption Practices in Mexico. *Soc. Sci.* **2017**, *6*, 144. [CrossRef]
6. Pablo-Romero, M.; Sánchez-Braza, A.; Sánchez-Rivas, J. Relationships between Hotel and Restaurant Electricity Consumption and Tourism in 11 European Union Countries. *Sustainability* **2017**, *9*, 2109. [CrossRef]
7. Meng, F.; Liu, Y.; Liu, L.; Li, Y.; Wang, F. Studies on Mathematical Models of Wet Adhesion and Lifetime Prediction of Organic Coating/Steel by Grey System Theory. *Materials* **2017**, *10*, 715. [CrossRef]
8. Zhou, W.; Pei, D.; Yan, L.; Yong, Y.; Tong, N. An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. *Appl. Soft Comput.* **2018**, *72*, 321–337.
9. Chiang, C.; Ho, M.; Chen, J. A hybrid approach of neural networks and grey modeling for adaptive electricity load forecasting. *Neural Comput. Appl.* **2006**, *15*, 328–338. [CrossRef]
10. Mikayilov, J.I.; Hasanov, F.J.; Bollino, C.A.; Mahmudlu, C. Modeling of Electricity Demand for Azerbaijan: Time-Varying Coefficient Cointegration Approach. *Energies* **2017**, *10*, 1918. [CrossRef]
11. Azadeh, A.; Seraj, O.; Saberi, M. An integrated fuzzy regression-analysis of variance algorithm for improvement of electricity consumption estimation in uncertain environments. *Int. J. Adv. Manuf. Technol.* **2011**, *53*, 645–660. [CrossRef]

12. Mohamed, Z.; Bodger, P. Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy* **2005**, *30*, 1833–1843. [CrossRef]

13. Azadeh, A.; Saberi, M.; Gitiforouz, A. An integrated simulation-based fuzzy regression-time series algorithm for electricity consumption estimation with non-stationary data. *J. Chin. Inst. Eng.* **2011**, *34*, 1047–1066. [CrossRef]

14. Kumar, U.; Jain, V.K. Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. *Energy* **2010**, *35*, 1709–1716. [CrossRef]

15. Hussain, A.; Rahman, M.; Memon, J.A. Forecasting electricity consumption in Pakistan: The way forward. *Energy Policy* **2016**, *90*, 73–80. [CrossRef]

16. Li, H.; Guo, S.; Zhao, H.; Su, C.; Wang, B. Annual Electric Load Forecasting by a Least Squares Support Vector Machine with a Fruit Fly Optimization Algorithm. *Energies* **2012**, *5*, 4430–4445. [CrossRef]

17. Meng, M.; Niu, D.; Sun, W. Forecasting Monthly Electric Energy Consumption Using Feature Extraction. *Energies* **2011**, *4*, 1495–1507. [CrossRef]

18. Kandananond, K. Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach. *Energies* **2011**, *4*, 1246–1257. [CrossRef]

19. Zhao, H.; Zhao, H.; Guo, S. Using GM (1,1) Optimized by MFO with Rolling Mechanism to Forecast the Electricity Consumption of Inner Mongolia. *Appl. Sci.* **2016**, *6*, 20. [CrossRef]

20. Ma, W.; Bai, X.; Mu, L. Short term load forecasting using artificial neuron network and fuzzy inference. *Power Syst. Technol.* **2003**, *5*, 29–32.

21. Liang, N. *Short-Term Load Forecasting Based on Modified Fruit Fly Algorithm and Support Vector Machine*; Guangxi University: Nanning, China, 2014.

22. Wang, J.; Li, L.; Niu, D.; Tan, Z. An annual load forecasting model based on support vector regression with differential evolution algorithm. *Appl. Energy* **2012**, *94*, 65–70. [CrossRef]

23. Eissa, M.M. Challenges and novel solution for wide-area protection due to renewable sources integration into smart grid: An extensive review. *IET Renew. Power Gener.* **2018**, *12*, 1843–1853. [CrossRef]

24. Navarro, J.A.D.; López, R.D.; Loyo, J.M.Y.; Sevil, J.S.A.; Agustín, J.L.B. Design of an electric vehicle fast-charging station with integration of renewable energy and storage systems International. *J. Electr. Power Energy Syst.* **2019**, *105*, 46–58. [CrossRef]

25. Denholm, P.; Mai, T. Timescales of energy storage needed for reducing renewable energy curtailment. *Renew. Energy* **2019**, *130*, 388–399. [CrossRef]

26. Bornapour, M.; Hooshmand, R.A.; Parastegar, M. An efficient scenario-based stochastic programming method for optimal scheduling of CHP-PEMFC, WT, PV and hydrogen storage units in micro grids. *Renew. Energy* **2019**, *130*, 1049–1066. [CrossRef]

27. Taibi, E.; del Valle, C.F.; Howells, M. Strategies for solar and wind integration by leveraging flexibility from electric vehicles: The Barbados case study. *Energy* **2018**, *164*, 65–78. [CrossRef]

28. Angenendt, G.; Zurmühlen, S.; Axelsen, H.; UweSauer, D. Comparison of different operation strategies for PV battery home storage systems including forecast-based operation strategies. *Appl. Energy* **2018**, *229*, 884–899. [CrossRef]

29. An, X.; Jiang, D.; Zhao, M.; Liu, C. Short-term prediction of wind power using EMD and chaotic theory Communications in Nonlinear. *Sci. Numer. Simul.* **2012**, *17*, 1036–1042. [CrossRef]

30. Kim, S.H.; Lee, G.; Kwon, G.Y.; Kim, D.I.; Shin, Y.J. Deep Learning Based on Multi-Decomposition for Short-Term Load Forecasting. *Energies* **2018**, *11*, 3433. [CrossRef]

31. Pang, B.; Tang, G.; Zhou, C.; Tian, T. Rotor Fault Diagnosis Based on Characteristic Frequency Band Energy Entropy and Support Vector Machine. *Entropy* **2018**, *20*, 932. [CrossRef]

32. Xie, S.; Liang, Y.; Zheng, Z.; Liu, H. Combined Forecasting Method of Landslide Deformation Based on MEEMD, Approximate Entropy, and WLS-SVM. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 5. [CrossRef]

33. Xiao, Y.; Kang, N.; Hong, Y.; Zhang, G. Misalignment Fault Diagnosis of DFWT Based on IEMD Energy Entropy and PSO-SVM. *Entropy* **2017**, *19*, 6. [CrossRef]

34. Zhao, H.; Sun, M.; Deng, W.; Yang, X. A New Feature Extraction Method Based on EEMD and Multi-Scale Fuzzy Entropy for Motor Bearing. *Entropy* **2017**, *19*, 14. [CrossRef]

35. Li, M.; Wang, W.; De, G.; Ji, X.; Tan, Z. Forecasting Carbon Emissions Related to Energy Consumption in Beijing-Tianjin-Hebei Region Based on Grey Prediction Theory and Extreme Learning Machine Optimized by Support Vector Machine Algorithm. *Energies* **2018**, *11*, 2475. [CrossRef]

36. Li, S.; Song, S.; Wan, Y. Laplacian twin extreme learning machine for semi-supervised classification. *Neurocomputing* **2018**, *321*, 17–27. [CrossRef]

37. Du, F.; Zhang, J.; Ji, N.; Shi, G.; Zhang, C. An effective hierarchical extreme learning machine based multimodal fusion framework. *Neurocomputing* **2018**, *322*, 141–150. [CrossRef]

38. Shang, Z.; Deng, T.; He, J.; Duan, X. A novel model for hourly PM2.5 concentration prediction based on CART and EELM. *Sci. Total Environ.* **2019**, *651*, 3043–3052. [CrossRef]

39. Ming, Y.; Zhu, E.; Wang, M.; Ye, Y.; Liu, X.; Yin, J. DMP-ELMs: Data and model parallel extreme learning machines for large-scale learning tasks. *Neurocomputing* **2018**, *320*, 85–97. [CrossRef]

40. Shi, P.; Yuan, Y.; Kuang, L.; Li, G.; Zhang, H. Water Temperature Prediction in Pond Aquaculture Based on EMD-IGA-SELM Neural Network. *Trans. Chin. Soc. Agric. Mach.* **2018**, 10.

41. Liu, Y.; Jiang, Z.; Liu, C. Optimization of Processing Parameters for a Reverse Drawing–Flanging Combined Process for a B550CL High-Strength Steel Spoke Based on Grey Relational Analysis. *Matels* **2018**, *8*, 7.

42. Hua, Y.; Wang, S.; Chai, C.; Liu, Z.; Liu, X.; Zou, L.; Wu, Q.; Zhao, H.; Ying, Y. Quality Evaluation of Pseudostellariae Radix Based on Simultaneous Determination of Multiple Bioactive Components Combined with Grey Relational Analysis. *Molecules* **2017**, *22*, 13. [CrossRef]

43. Rajesh, R.; Ravi, V. Supplier selection in resilient supply chains: A grey relational analysis approach. *J. Clean. Prod.* **2015**, *86*, 343–359. [CrossRef]

44. Hashemi, S.H.; Karimi, A.; Tavana, M. An integrated green supplier selection approach with analytic network process and improved Grey relational analysis. *Int. J. Prod. Econ.* **2015**, *159*, 178–191. [CrossRef]

45. Sarıkaya, M.; Güllü, A. Multi-response optimization of minimum quantity lubrication parameters using Taguchi-based grey relational analysis in turning of difficult-to-cut alloy Haynes 25. *J. Clean. Prod.* **2015**, *91*, 347–357. [CrossRef]

46. Yuan, J.; Lei, Q.; Wang, Y. Prospect of China electricity demand and Analysis of Coal and Power demand in the new normal economy. *Energy China* **2015**, *7*, 21–27.

47. Que, B.; Zhao, Y.; Xia, Q. Research on the trend of Power demand growth in the medium and long term in China. *Energy China* **2018**, *11*, 21–27.