Short-Term Combined Forecasting Method of Park Load Based on CEEMD-MLR-LSSVR-SBO

Bo Hu 1,2, Jian Xu 1,*, Zuoxia Xing 1, Pengfei Zhang 1, Jia Cui 1 and Jinglu Liu 1

1 School of Electrical Engineering, Shenyang University of Technology, Shenyang 110870, China; hubo@sut.edu.cn (B.H.); xingzuoxia@sut.edu.cn (Z.X.); zhangpeifei@sut.edu.cn (P.Z.); cuijayouxiang@sut.edu.cn (J.C.); liujinglu@sut.edu.cn (J.L.)
2 State Grid Liao Ning Electric Power Supply Co., Ltd., Shenyang 110004, China
* Correspondence: xujian@sut.edu.cn; Tel.: +86-134-7887-7588

Abstract: To improve the accuracy of park load forecasting, a combined forecasting method for short-term park load is proposed based on complementary ensemble empirical mode decomposition (CEEMD), sample entropy, the satin bower bird optimization algorithm (SBO), and the least squares support vector regression (LSSVR) model. Firstly, aiming at the random fluctuation of park load series, the modes with different characteristic scales are divided into low-frequency and high-frequency according to the calculation of sample entropy, which is based on the decomposition of historical park load data modes by CEEMD. The low-frequency is forecast by multiple linear regression (MLR), and the high-frequency component is the training input of the LSSVR forecasting model. Secondly, the SBO algorithm is adopted to optimize the regularization parameters and the kernel function width of LSSVR. Then, the park load forecasting model of each sequence component is built. The forecast output of each sequence component is superimposed to get the final park load forecast value. Finally, a case study of a park in Liaoning Province has been performed with the results proving that the proposed method significantly outperforms the state-of-art in reducing the difficulty and complexity of forecasting effectively, also eliminating the defect of large reconstruction error greatly through the decomposed original sequence by the ensemble empirical model.

Keywords: short-term park load forecasting; least squares support vector regression (LSSVR); complementary ensemble empirical mode decomposition (CEEMD); satin bower bird optimization algorithm (SBO); combination model; multiple linear regression

1. Introduction

With the rapid development of social economy and the rising power load of the park, the risk of heavy overload borne by the distribution transformer equipment in the park is becoming more and more serious, which lays a hidden danger for the safe operation of the power grid. If the power flow changing trend of distribution transformer could be forecast accurately in the short term, it is of great significance for power companies to forecast the heavy overload of distribution transformer accurately and take timely measures.

There is a great deal of researches on improving the forecasting accuracy, generalization ability, and adaptive ability of the short-term load forecasting model. The traditional load forecasting methods include ARIMA [1], the grey model method [2,3], the support vector machine [4,5], and the similar day method [6,7]. In recent years, with the continuous improvement of the degree of power grid information, the frequency and accuracy of power grid operation data collection by terminal devices such as on-line monitoring devices also continue to improve [8]. The artificial intelligence technology represented by deep learning has made the research of load forecasting develop rapidly. The most representative technologies, such as LSTM [8,9] and neural network [10,11], form a hybrid neural network by increasing the depth of the network, which further improves the ability of high-dimensional feature learning.
The existing researches on load forecasting with strong randomness and high complexity are mainly based on load decomposition and the multi-model fusion forecasting method. The widely used load decomposition methods include the filtering analysis method, the modal empirical decomposition method, and its derivative algorithm [12–14]. By introducing the load decomposition method into the forecasting, separating the local characteristic signals of the time series, and selecting the forecasting method separately, the learning ability of the local characteristic signals could be enhanced and the comprehensive forecasting accuracy could be improved. In [14], the variational mode decomposition is used to effectively reduce the instability of the original sequence, and then the depth gated network is used to model and forecast each sub-component respectively, so that the model has stronger time-series-dependent fitting ability. In [13], the load data are decomposed by the method of ensemble empirical mode decomposition, and different Elman recurrent neural network models are established for each time-series component, but the process of adjusting parameters of the model is complicated in actual forecasting. In [12], the GRU neural network and multiple linear regression method are used to forecast the time-series components of different frequencies respectively, and better forecasting results are obtained when the data contain more noise. A method for one-day-ahead forecast of electricity prices combines deterministic components and the residual component [15]. A hybrid algorithm for EPF combines the local forecasting paradigm, GRNN and has; this algorithm overcame the demerits of the global algorithms and the parameters were optimized by HAS [16]. A method is proposed to improve prediction accuracy by different machine learning algorithms: the deterministic component and the residual component were combined to get the specific properties of electricity prices [17]. A framework was proposed for forecasting electricity prices by using DNN, LSTM, GRU and the CNN model [18]. An approach was proposed for forecasting electricity demand by using linear regression-based models, spline function-based models, and traditional time-series models [19].

Currently, there are two aspect deficiencies. Firstly, the above researches are usually oriented to the medium and short-term forecasting of wide-area total load, and few studies are specifically aimed at park distribution transformer load forecasting. The reason is that the wide-area total load is superimposed by the distribution transformer load of each park. High-frequency components and noise cancel each other internally, resulting in a smoother load curve, while the park distribution transformer load has obvious periodic characteristics in the long time scale, but there are a large number of high-frequency components and noises in the short time scale, and the random characteristics are more obvious, which greatly increases the fitting difficulty of the forecasting model. Therefore, it is necessary to seek a more accurate forecasting technology in line with the short-term load characteristics of the park.

Secondly, most of the high-frequency component forecasting algorithms in the above literature for high randomness are single, leading to a poor generalization of forecasting models, and boosting the comprehensive forecast ability to load different frequency feature components by multi-model fusion becomes an effective means. In view of the fact that the distribution transformer load in the park is affected by many complex factors, on the basis of its regular stationary characteristic load with strong regularity, the local part shows strong random characteristics, and the traditional single forecasting method makes it difficult to learn the rules accurately.

In this paper, a short-term park load combination forecasting method is proposed based on complementary ensemble empirical mode decomposition (CEEMD). The LSSVR parameters are optimized by SBO. The load data of the historical park is decomposed into series components by CEEMD, and the MLR algorithm is used in the obvious periodic low-frequency component to reduce the overall complexity of the forecasting model. The LSSVR load forecasting model is built for the high-frequency part, which is difficult to forecast. The regularization parameters and kernel function width of the LSSVR model are optimized by the SBO algorithm, and the forecast values of each component are superimposed to
get the final forecast value. The effectiveness of the proposed method is proved through simulation results.

A novel satin bower bird optimization algorithm is introduced, and the regularization parameters and kernel function width of the least squares support vector regression model are optimized by its excellent optimization performance. The short-term park load power forecasting simulation results show that the SBO algorithm with dynamic step size and mutation has better convergence and forecasting accuracy compared with the traditional algorithms.

In order to solve the problem of random fluctuation of park load power series, a combined forecasting method is proposed based on complementary ensemble empirical mode decomposition, the satin bower bird optimization algorithm (SBO), and least squares support vector regression (LSSVR) parameters. The original park load time series are decomposed into different time-scale series components by complementary ensemble empirical mode decomposition. As a result, the overall calculation of the model and the complex parameter adjustment work are reduced in the training process.

The simulation results of numerical examples show that the proposed method not only solves the defect of large reconstruction error of ensemble empirical mode decomposition and improves the forecasting accuracy effectively, but reduces the difficulty and complexity of forecasting greatly.

The remainder of this paper is organized as follows: Section 1 proposes a combined method forecasting strategy. Section 2 presents the algorithmic details of CEEMD and satin bower bird optimization algorithm. Section 3 describes the short-term forecasting model. Section 4 provides experimental results and discussion. Finally, the paper is concluded in Section 5.

2. Combined Forecasting Model and Method of Short-Term Park Load

In this paper, a combined model of short-term load forecasting of the park based on CEEMD-MLR-LSSVR-SBO is proposed. The nonlinear and unstable load time series of historical parks are decomposed into n IMF subseries components and 1 Re component, respectively, by CEEMD. According to the sample entropy of the obtained components, the components are divided into the high-frequency part with complex fluctuation characteristics and the low-frequency part with obviously smooth periodicity. The LSSVR high-frequency forecasting model of each component is established by multivariate linear regression low-frequency component forecasting and training set, respectively. Then the two parameters of each component corresponding to the LSSVR forecasting model are optimized by SBO, and the LSSVR forecasting model of each component is trained well. Then the components of the test set are used to forecast, and the final forecast value is obtained by superimposing the forecast results of all components. Finally, based on the distribution transformer load data of a commercial and residential mixed park in Liaoning Province, the effectiveness of EMD-Stacking-MLR’s short-term load forecasting method for distribution transformer load forecasting in the park is verified.

The forecasting flow chart of the combined method for this paper is shown in Figure 1.

The steps of forecasting short-term park load based on CEEMD-MLR-LSSVR-SBO combination model are as follows:

(a) The load output power of the original park is decomposed by the CEEMD method, which is divided into several groups of components on different time scales. According to the sample entropy of the components, the components are divided into the high-frequency part with complex fluctuation characteristics and the low-frequency part with smooth periodicity.

(b) On this basis, the reasonable forecasting step of the multivariate linear regression low-frequency component forecasting method is studied, and the LSSVR regression model is built for the decomposed high-frequency training set components. The satin bower bird optimization algorithm is used to optimize the regularization parameters
The decomposed signal of EMD is:

\[ x(t) = \sum_{i=1}^{n} C_i(t) + r_n(t) \]  

\[ c_j(t) = \frac{1}{2k} \sum_{j=1}^{k} (c_{ij} + c_{-ij}) \]  

\( \gamma \) and kernel function width \( \sigma \) of each regression model. In addition, the LSSVR forecasting models of each component are established respectively.

(c) Each component of the test-set decomposition is substituted into the corresponding forecasting model, and the forecast value of each component of the test sample is obtained.

(d) The forecast values of each component are superimposed to get the final load forecasting result of the park.

(e) The resulting forecast park load was analyzed in error with the actual park load data.

Figure 1. Combining method forecasting flowchart of CEEMD-MLR-LSSVR-SBO.

3. Complementary Ensemble Empirical Mode Decomposition (CEEMD) and the Satin Bower Bird Optimization Algorithm (SBO) Analysis

3.1. Analysis of CEEMD Principle

Empirical mode decomposition (EMD) is actually used to decompose the original time series into several intrinsic mode function (IMF) components and a residual component (Re) adaptively [14] that are independent of each other at different scales. The decomposed signal of EMD is:

\[ x(t) = \sum_{i=1}^{n} C_i(t) + r_n(t) \]  

In the formula, \( t \) is the time point, \( C_i(t) \) and \( r_n(t) \) are the IMF components, and the residual component \( n \) is the number of IMF.

While ensemble empirical mode decomposition (EEMD) is proposed for the existence of modal aliasing in EMD [20], the Gaussian white noise is repeatedly added to the whole time-frequency space for EMD decomposition, and multiple mean IMF components are obtained as the final decomposition results.

In this paper, positive and negative random Gaussian white noise is added in pairs to form a complementary ensemble empirical mode decomposition (CEEMD). Compared with EEMD, it not only solves the mode aliasing problem of EMD, but eliminates the residual auxiliary noise in the reconstructed signal after EEMD decomposition. After CEEMD decomposition, the IMF components and residual components are respectively:

\[ c_j(t) = \frac{1}{2k} \sum_{j=1}^{k} (c_{ij} + c_{-ij}) \]
\[ r_n(t) = \frac{1}{2k} \sum_{i=1}^{k} (r_i + r_{-i}) \]  
(3)

where \( c_{ij} \) and \( c_{-ij} \), respectively, means the \( j \)-th IMF component is obtained after the signal is decomposed by adding positive and negative white Gaussian noise for the \( i \)-th time, \( k \) is the maximum time of ensemble averages, \( r \) is the residual component after decomposition.

### 3.2. Sample Entropy

Sample entropy is one of the quantitative description tools to measure the complexity of the system. From the point of view of the complexity of time series, it is used to measure the probability of generating new patterns and to quantitatively describe the complexity and regularity of the system by stipulating window observation to compare the distance between the eigenvectors of two observations. The entropy value accurately reflects the complexity of the time series and the probability of the system generating new patterns. The larger the sample entropy is, the more complex the time series is, the greater the probability of the system generating new patterns is. On the contrary, the simpler the time series is, the smaller the probability is [21]. For a given time series \( X = [x_1, x_2, \cdots, x_n] \), the calculation steps of sample entropy (SampEn) are as follows:

(a) The sequence is composed into a set of \( m \)-dimensional vectors in order, \( X_m,1, \cdots, X_m,n-m+1 \) and \( X_m,i = [x_i, x_{i+1}, \cdots, x_{i+m-1}] \), \( 1 \leq i \leq n-m+1 \).

(b) The distance between vectors \( X_m,i \) and \( X_m,j \) is defined as the absolute value of the maximum difference between the corresponding elements. The formula is as follows:

\[ d[X_m,i, X_m,j] = \max_{k \in [0, m-1]} |x_{i+k} - x_{j+k}| \] 
(4)

(c) According to the given distance threshold \( r \), count the number of \( d[X_m,i, X_m,j] < r \), record it as \( N_m,i \), and define: \( B_m,i = \frac{N_m,i}{n-m+1} \), \( 1 \leq i \leq n-m \) and \( i \neq j \).

(d) The average \( B_m \) of \( B_m,i \) is determined, and the formula is as follows:

\[ B_m = \frac{\sum_{i=1}^{n-m+1} B_m,i}{n-m+1} \] 
(5)

(e) The above steps (a)–(d) are repeated to get \( B_{m+1} \).

(f) When \( n \) is finite, the sample entropy (SampEn) formula is as follows:

\[ \text{SampEn}(m, r, n) = - \ln \left( \frac{B_{m+1}}{B_m} \right) \] 
(6)

where the parameter \( m \) is the observation window dimension and \( r \) is the distance threshold. Because this paper targets the IMF component time series \( c_i(t) \), the sample entropy calculation only focuses on its changing trend, while the changing trend of sample entropy is not affected by \( m \) and \( r \), so the parameter is set as a regular numerical value, \( m = 2, r = 0.2 \text{ Std} \), and the IMF component with sample entropy value close to 0 is divided into low-frequency component, while the rest is divided into high-frequency component.

### 3.3. Analysis of the Satin Bower Bird Optimization Algorithm (SBO) Principle

The satin bower bird optimization algorithm (SBO) was a novel meta heuristic proposed in 2017, which imitated the courtship mechanism of adult male satin bower birds attracting female satin bower bird through building nests [22], and the SBO optimization process was as follows:
(a) Population initialization. The initial $B(t) = \{X^t_i = x^t_{i1}, x^t_{i2}, \ldots, x^t_{iD}\}$ of $N$ nests is randomly generated, where $i = 1, 2, \ldots, N$, $D$ is the location dimension of the nest and $t$ is the current number of iterations.

(b) To determine the fitness value $f_{it}$ of each nest and the probability $p_i$ of being selected in the population, the expressions are as follows:

$$f_{it} = \begin{cases} \frac{1}{1 + f(x_i)} & f(x_i) \geq 0 \\ 1 + |f(x_i)| & f(x_i) < 0 \end{cases}$$

$$p_i = \frac{f_{it}}{\sum_{n=1}^{NB} f_{in}}$$

where $f(x_i)$ is the objective function of the $i$-th nest.

(c) Population renewal. The male bird updates the nest position through continuous communication and learning, that is, the male dynamically updates the population according to the current random search of the best nest $x_{jk}$, the optimal nest $x_{elite}$, in the whole population and the step size factor $\lambda_k$ determined by the target nest selection probability $p_j$, as shown in Formulas (9) and (10), respectively:

$$x^{t+1}_{ik} = x^t_{ik} + \lambda_k \left(\frac{x_{jk} + x_{elite}}{2} - x^t_{ik}\right)$$

$$\lambda_k = \frac{\alpha}{1 + p_j}$$

In the formula, $k$ is the corresponding dimension of each component, $j$ is obtained by roulette selection mechanism, and $\alpha$ is the upper limit of step size.

(d) Individual variation. Usually, strong males rob the decorations of other males’ nests, so the nests are randomly mutated in the form of the probability distribution, as shown in Formulas (11)–(13):

$$x^{t+1}_{ik} \sim N \left(x^t_{ik}, \sigma^2 \right)$$

$$N \left(x^t_{ik}, \sigma^2 \right) = x^t_{ik} + (\sigma \ast N(0, 1))$$

$$\sigma = z \ast (var_{max} - var_{min})$$

where $\sigma$ is the standard deviation, $z$ is the proportional coefficient, $var_{max}$ and $var_{min}$ are the upper and lower boundaries of the nest location, respectively.

Finally, all the populations are combined to obtain the optimal nest position as the parameter value of optimal selection.

4. Short-Term Forecasting Model of Park Load Based on CEEMD-MLR-LSSVR-SBO

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, and the experimental conclusions that can be drawn.

4.1. Least Squares Support Vector Regression (LSSVR) Model

LSSVR is suitable for the training and forecasting of small sample data. Its basic idea is to use the nonlinear mapping function $\phi(\cdot)$ to map the given sample set $(x_i, x_j)$ to the high-dimensional feature space for fitting. According to the structural risk principle, the error quadratic term is selected as the experience loss of the training set, and the
minimized objective function is built. LSSVR regression model is an optimization problem with constraints:

\[
\begin{align*}
\min f(\omega, e) &= \frac{1}{2} \| \omega \|^2 + \frac{1}{2} \gamma \sum_{i=1}^{l} e_i^2 \\
\text{s.t.} \quad y &= \omega^T \varphi(x_i) + b + e_i
\end{align*}
\]  

where \( s.t. \) represents the equality constraints on the objective function, \( \omega \) is the weight vector, \( e \) is the error variable, \( \gamma \) is the regularization parameter, and \( b \) is the bias vector.

The mathematical derivation of LSSVR can be found in the Ref. [23], and the final regression function is as follows:

\[
y(x) = \sum_{i=1}^{l} a_i k(x, x_i) + b
\]

where \( a_i \) is the Lagrange multiplier, and \( k(x_i, x_j) \) is the kernel function. In this paper, the radial basis kernel function is selected as follows:

\[
k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)
\]

where \( \sigma \) is the width of the kernel function.

The regularization parameter \( \gamma \) is used to measure the smoothness of the fitting curve and minimize the fitting error [24]. If its value is too large or too small, the generalization ability of the model would become worse. The value of kernel function width \( \sigma \) is too small, which can easily lead to overfitting; when the value of the parameter \( \sigma \) is too large, it will lead to the deterioration of the generalization ability of the model [25]. Regularization parameter \( \gamma \) and kernel function width \( \sigma \) have a great influence on the learning ability and forecasting accuracy of the LSSVR regression model. In order to improve the forecasting effect of LSSVR, it is necessary to find two suitable parameters for it. The traditional parameter selection method is the cross-validation method.

4.2. Regression Model of Optimizing LSSVR Parameters Based on SBO

In this paper, the SBO optimization algorithm is introduced to optimize two parameters of the LSSVR model. LSSVR modeling usually needs to construct the mapping relationship \( f : R^m \in R \) between input \( x_t = \{x_{t-1}, x_{t-2}, \ldots, x_{t-m}\} \) and output \( y_t = (x_t) \), where \( m \) is the embedded dimension. Before the regression model training, the original park load data \( y \) is normalized, as shown in Formula (17).

\[
y_i = \frac{y - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}}
\]

The normalized processed data were divided into training and test sets. The training set learning was utilized within a set maximum number of iterations, resulting in the best combination of parameters. The objective function for the two-parameter optimization was as follows:

\[
\begin{align*}
\min f(\gamma, \sigma) &= \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \\
\text{s.t.} \quad \gamma &\in [\gamma_{\text{min}}, \gamma_{\text{max}}] \\
\sigma &\in [\sigma_{\text{min}}, \sigma_{\text{max}}]
\end{align*}
\]

where \( y_{\text{max}} \) and \( y_{\text{min}} \) are the park load Max and Min, respectively, in the selected data set; \( y_i \) and \( \hat{y}_i \) are the output power true and forecast values, respectively, for the \( i \)-th sample.

According to the selected training data set and the set objective function, a regression model based on SBO to optimize LSSVR parameters is constructed, and the flow chart is shown in Figure 2.
In order to verify the scientific nature and reliability of the method proposed in this paper, the experimental samples in this paper are the distribution transformer load data of a 10 kV park in a commercial and residential mixed area in Liaoning Province. The time span is from 1 January 2020 to 31 December 2020, and the sampling interval is 15 min (Figure 3). From the local magnification diagram of the load curve in Figure 3, it can be seen that the load curve of the distribution transformer in the park has an obvious long periodicity but fluctuates greatly in a short time between the peak and valley of electricity consumption. In this experiment, the data of the first 11 months are divided into training sets, and the data of December are divided into test sets. This method can be widely applied to any region of the world, and we simply use this case in China to demonstrate its implementation.
5. Numerical Example Analysis

In order to verify the scientific nature and reliability of the method proposed in this paper, the experimental samples in this paper are the distribution transformer load data of a 10 kV park in a commercial and residential mixed area in Liaoning Province. The time span is from 1 January 2020 to 31 December 2020, and the sampling interval is 15 min (Figure 3). From the local magnification diagram of the load curve in Figure 3, it can be seen that the load curve of the distribution transformer in the park has an obvious long periodicity but fluctuates greatly in a short time between the peak and valley of electricity consumption. In this experiment, the data of the first 11 months are divided into training sets, and the data of December are divided into test sets. This method can be widely applied to any region of the world, and we simply use this case in China to demonstrate its implementation.

Firstly, the long-period characteristics and short-term fluctuation characteristics of the load curve are separated by empirical mode decomposition, and the appropriate forecasting method is selected according to the complexity of different components. Then the multiple linear regression is proposed to forecast the low-frequency IMF curve, and the CEEMD-MLR-LSSVR-SBO ensemble learning forecasting model is used to forecast the high-frequency IMF curve. Finally, the low-frequency component and high-frequency component are superimposed and reconstructed, and the final load forecasting curve of the distribution transformer in the park is obtained. Relative mean square error (RMSE) and mean absolute percentage error (MAPE) are used in the forecasting and evaluation indicators, as follows:

\[ \text{err}_{\text{RMSE}} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2} \]  
\[ \text{err}_{\text{MAPE}} = \frac{100}{m} \sum_{i=1}^{m} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \]

5.1. Empirical Mode Decomposition

Fourteen groups of IMF with frequency from high to low are obtained by empirical mode decomposition of distribution and transformer load data, as shown in Figure 4. In this paper, the reconstruction curve is obtained by superposing each IMF component, and the reconstruction error of EMD load decomposition is calculated by using absolute error (AE) and mean absolute error (MAE). The formula is as follows:

\[ \text{err}_{\text{AE}} = |y_i - \hat{y}_i| \]
\[ \text{err}_{\text{MAE}} = \frac{\sum_{i=1}^{m} |y_i - \hat{y}_i|}{m} \]
Fourteen groups of IMF with frequency from high to low are obtained by empirical mode decomposition of distribution transformer load. The CEEMD-MLR-LSSVR-SBO forecasting model is used to mine the rules to the maximum extent, so as to improve the forecasting accuracy. Although IMF1 to IMF5 are high-frequency components, the numerical variation range is small and the proportion of the original load is small, that is, the superposition of the component forecasting results has little impact on the accuracy of the final forecasting results. The other components have a great influence on the accuracy of the final forecasting results, among which IMF6, IMF7, and IMF14 have the characteristics of a large proportion of the original load, so it is necessary to improve the forecasting accuracy of the above components as far as possible.
multiple linear regression to forecast them. IMF1~IMF8 are defined as high-frequency components. Therefore, the unified forecasting step size of each algorithm is set as 30 min in this paper.

Because the factors affecting the forecast time point component data in the MLR model are only historical component data, the forecast performance of the model for different time spans needs to be contrasted according to its forecast properties. In this paper, single-step and multi-step forecasting experiments are done for IMF9 and IMF14, respectively, and the RMSE and MAPE are calculated for the forecast outcomes at different forecast step sizes. The results are shown in Table 1. Among them, the single-step forecasting is forecast 15 min in advance, and the multi-step forecasting is forecast 30 min, 1 h, and 1 day in advance. It was found by the experiment results that when the component frequencies were high, the MLR forecasting method is affected by the forecasting step greatly. For IMF9, forecasting more than 30 min in advance would lead to a large decrease in forecasting accuracy and cannot meet the requirement of MAPE less than 10 for the result of low-frequency, single-component forecasting. Therefore, the unified forecasting step size of each algorithm is set as 30 min in this paper.

Table 1. Evaluation of MLR forecasting results at different time spans.

| Low-Frequency Component | Evaluating Indicator | 1 d | 1 h | 30 min | 15 min |
|-------------------------|----------------------|-----|-----|--------|--------|
| IMF9                    | RMSE                 | 19.83 | 0.98 | 0.45   | 0.24   |
|                         | MAPE                 | 180.74 | 19.5 | 9.84   | 4.97   |
| IMF14                   | RMSE                 | 0.46  | 0.02 | 0.02   | 0.00   |
|                         | MAPE                 | 0.37  | 0.03 | 0.04   | 0.00   |

5.3. Studies of Different Forecast Methods for High-Frequency IMFs

In order to optimize the performance of CEEMD-MLR-LSSVR-SBO, it is necessary to analyze the forecasting effect of each high-frequency component of each model based on a common forecasting model. Firstly, this paper designs an experiment to compare and analyze the individual forecasting results of each common forecasting model (LSTM, BP, SVM, XGBoost algorithm) on the high-frequency components from IMF1 to IMF8. Then the error difference degree of each model of each high-frequency component is calculated by using the Pearson correlation coefficient.

In the paper, the component data IMF1~IMF8 are brought into the algorithm training of each forecasting model to comprise the effects of forecasting individual forecasting models, and the forecasting accuracy is evaluated by MAPE. The results are shown in Table 2.
Table 2. Forecasting Error MAPE/% of Algorithms.

| Component Serial Number | Independent Forecasting Models | CEEMD-MLR-LSSVR-SBO |
|-------------------------|-------------------------------|----------------------|
| IMF1                    | 118.02 127.28 101.96 102.40 | 115.24 104.71 |
| IMF2                    | 133.74 145.67 133.5 141.63 | 137.43 127.23 |
| IMF3                    | 249.24 311.37 269.07 315.85 | 284.37 265.62 |
| IMF4                    | 219.11 233.54 243.51 236.91 | 234.83 231.52 |
| IMF5                    | 48.42 48.99 47.55 48.97 | 49.02 47.91 |
| IMF6                    | 43.54 44.19 42.98 44.02 | 43.05 43.12 |
| IMF7                    | 9.98 12.02 12.85 9.54 | 11.23 13.01 |
| IMF8                    | 15.76 18.01 11.24 9.92 | 14.08 11.99 |

Table 2 shows that when each algorithm carries on the forecasting separately, the IMF7 and IMF8 data in the high-frequency component have obvious continuity characteristics, so compared with other algorithms, the LSTM algorithm uses the cumulative information of previous training effectively, and the forecasting performance is also superior.

As the change frequency from IMF1 to IMF6 is higher and the complexity of the time series increases, the forecasting effect of the algorithm decreases obviously. Among them, the individual forecasting error of XGBoost for the above high-frequency components is lower generally than that of other algorithms, because XGBoost is good at mining and learning deep-feature relationships, which makes the model training fuller. For the high-frequency part of IMF, according to the evaluation index MAPE, CEEMD-MLR-LSSVR-SBO model compared with other general forecasting methods has a forecasting accuracy that is not the best. However, the MAPE value is generally lower than the MAPE average forecast by the individual basic learner. Therefore, it is proved that the CEEMD-MLR-LSSVR-SBO model has higher general applicability than the individual model for different data sets.

5.4. Forecasting Effect of Load Curve

By adding the low-frequency components of IMF9~14 directly forecast by multiple linear regression and the high-frequency components of IMF1~8 forecast by the CEEMD-MLR-LSSVR-SBO ensemble learning model, the final load forecasting curve of the distribution transformer is obtained. In order to continue to verify whether the CEEMD-MLR-LSSVR-SBO forecasting model is able to maintain a high prediction accuracy for consecutive days, six models are used to forecast the load situation for consecutive seven days, respectively. The error analysis is shown in Table 3.

Comparing the data in Table 3, it can be seen that the average error of SVR and XGBoost forecasting results for one consecutive week is still lower than that of LSTM and BP. The error of SVR and XGBoost decreases by 0.29% and 0.81% compared with LSTM and BP, while XGBoost decreases by 0.07% and 0.59%. It can be seen that the XGBoost forecasting model is still better than the traditional heuristic algorithm forecasting model. The mean error of SVR also decreased by 0.22% compared to XGBoost. CEEMD-MLR-LSSVR-SBO has the highest accuracy in the continuous week forecasting, and the error fluctuation is not large. The mean value of weekly error decreases by 0.16%, 0.31%, 0.04%, 0.13% and 0.26% compared with the other five models, respectively. The mean root errors also decreased by 0.49%, 0.58, 0.52%, 0.41% and 0.5%, indicating that the CEEMD-MLR-LSSVR-SBO combination method can improve the accuracy and stability of model forecasting.

Among them, the EMD superposition forecasting method selected well by each component model is to select the forecasting model with the best forecasting effect of each IMF component according to the individual forecasting error of the previous basic learner. Finally, the high-frequency and low-frequency components are superimposed to form the load curve of the forecasting period. In the experiment of the simple forecasting method, the historical original load data are used as input, and the load curve of the output forecasting period is forecast directly. The experimental results show that the forecasting results of simple forecasting methods have the problem of peak lag to a certain extent, because the forecasting results are greatly affected by the historical load of input. The forecasting
accuracy based on CEEMD-MLR-LSSVR-SBO is obviously higher than that of the simple forecasting method, and the problem of peak lag is obviously improved. Although its forecasting accuracy is slightly lower than that of the EMD superposition forecasting method, the EMD superposition forecasting method based on each component model selects the component forecasting model according to the forecasting error, that is, it is difficult to realize in practical application, while the forecasting based on CEEMD-MLR-LSSVR-SBO overcomes the shortcomings of re-selection of each component forecasting model and has higher practicability.

Table 3. Evaluation index of load curve forecasting results of six forecasting models in one week of December 2020.

| Day   | LSTM [8] | BP [13] | SVR [2] | XGBoost [14] | Model Preferred | CEEMD-MLR-LSSVR-SBO |
|-------|----------|---------|---------|--------------|-----------------|---------------------|
|       | MAPE/\% | RMSE    | MAPE/\% | RMSE         | MAPE/\%         | RMSE                |
| Monday| 3.51     | 7.76    | 2.49    | 6.80         | 2.31            | 7.52                |
| Tuesday| 2.84    | 5.16    | 2.89    | 7.71         | 2.36            | 7.16                |
| Wednesday| 1.49    | 5.15    | 2.69    | 6.51         | 2.01            | 7.08                |
| Thursday| 1.52    | 4.27    | 1.67    | 7.48         | 1.18            | 3.96                |
| Friday | 1.65     | 4.06    | 3.82    | 7.19         | 1.10            | 3.92                |
| Saturday| 2.92    | 8.15    | 3.96    | 8.05         | 2.73            | 7.98                |
| Sunday | 2.97     | 8.59    | 3.05    | 8.20         | 2.71            | 7.88                |
| The average value| 2.41 | 6.16    | 2.93    | 7.42         | 2.12            | 6.50                |

6. Conclusions

A short-term load forecasting model of CEEMD-MLR-LSSVR-SBO that contains high-frequency local components and noise aiming at the park load is proposed in the paper. Through the case study and the simulation results of numerical examples, it is demonstrated that the proposed method not only solves the defect of large reconstruction error of ensemble empirical mode decomposition and improves the forecasting accuracy effectively, but reduces the difficulty and complexity of forecasting greatly.

In order to solve the problem of random fluctuation of park load power series, a combined forecasting method is proposed based on complementary ensemble empirical mode decomposition, the satin bower bird optimization algorithm (SBO), and least squares support vector regression (LSSVR) parameters. The original park load time series are decomposed into different time-scale series components by complementary ensemble empirical mode decomposition. As a result, the overall calculation of the model and the complex parameter adjustment work are reduced in the training process. Meanwhile, a novel satin bower bird optimization algorithm is introduced. And the regularization parameters and kernel function width of the least squares support vector regression model are optimized by its excellent optimization performance. The short-term park load power forecasting simulation results show that the SBO algorithm with dynamic step size and mutation has 15.23–48.38% improved forecasting accuracy compared with the traditional algorithms.

Author Contributions: Conceptualization, B.H.; methodology, B.H.; software, B.H. and J.C.; validation, J.X. and J.C.; formal analysis, J.X.; writing—original draft preparation, J.X.; visualization, Z.X.; supervision, P.Z.; writing—review and editing, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Xing Liao Talent Plan. No. XLYC1902090.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.
References

1. Bessani, M.; Massignan, J.A.D.; Santos, T.M.O.; London, J.B.A.; Maciel, C.D. Multiple households very short-term load forecasting using bayesian networks. *Electr. Power Syst. Res.* 2020, 189, 106733. [CrossRef]

2. Luo, J.; Hong, T.; Fang, S.C. Robust Regression Models for Load Forecasting. *IEEE Trans. Smart Grid* 2018, 10, 5397–5404. [CrossRef]

3. Zhang, W.; Chen, Q.; Yan, J.; Zhang, S.; Xu, J. A novel asynchronous deep reinforcement learning model with adaptive early forecasting method and reward incentive mechanism for short-term load forecasting. *Energy* 2021, 236, 121492. [CrossRef]

4. Cui, J.; Pan, J.; Wang, S.; Okoye, M.O.; Yang, J.; Li, Y.; Wang, H. Improved normal-boundary intersection algorithm: A method for energy optimization strategy in smart buildings. *Build. Environ.* 2022, 212, 108846. [CrossRef]

5. Buzna, L.; de Falco, P.; Ferruzzi, G.; Khormali, S.; Proto, D.; Refa, N.; Straka, M.; van der Poel, G. An ensemble methodology for hierarchical probabilistic electric vehicle load forecasting at regular charging stations. *Appl. Energy* 2021, 283, 116337. [CrossRef]

6. Park, R.J.; bin Song, K.; Kwon, B.S. Short-term load forecasting algorithm using a similar day selection method based on reinforcement learning. *Energies* 2020, 13, 2640. [CrossRef]

7. Mohan, N.; Soman, K.P.; Sachin Kumar, S. A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model. *Appl. Energy* 2018, 232, 229–244. [CrossRef]

8. Zang, H.; Xu, R.; Cheng, L.; Ding, T.; Liu, L.; Wei, Z.; Sun, G. Residential load forecasting based on LSTM fusing self-attention mechanism with pooling. *Energy* 2021, 229, 120682. [CrossRef]

9. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Single and multi-sequence deep learning models for short and medium term electric load forecasting. *Energies* 2019, 12, 149. [CrossRef]

10. Tian, C.; Ma, J.; Zhang, C.; Zhan, P. A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. *Energies* 2018, 11, 3493. [CrossRef]

11. Tang, X.; Dai, Y.; Wang, T.; Chen, Y. Short-term power load forecasting based on multi-layer bidirectional recurrent neural network. *IET Gener. Transm. Distrib.* 2019, 13, 3847–3854. [CrossRef]

12. Wang, S.; Zhang, Z. Short-term multiple load forecasting model of regional integrated energy system based on qwgru-mtl. *Energies* 2021, 14, 6555. [CrossRef]

13. Sideratos, G.; Ikonomopoulos, A.; Hatzigiagriou, N.D. A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks. *Electr. Power Syst. Res.* 2020, 178, 106025. [CrossRef]

14. Wang, Y.; Sun, S.; Chen, X.; Zeng, X.; Kong, Y.; Chen, J.; Guo, Y.; Wang, T. Short-term load forecasting of industrial customers based on SVMD and XGBoost. *Int. J. Electr. Power Energy Syst.* 2021, 129, 106830. [CrossRef]

15. Shah, I.; Bibi, H.; Ali, S.; Wang, L.; Yue, Z. Forecasting One-Day-Ahead Electricity Prices for Italian Electricity Market Using Parametric and Nonparametric Approaches. *IEEE Access* 2020, 8, 123104–123113. [CrossRef]

16. Elattar, E.E.; Elsayed, S.K.; Farrag, T.A. Hybrid Local General Regression Neural Network and Harmony Search Algorithm for Electricity Price Forecasting. *IEEE Access* 2019, 7, 2044–2054. [CrossRef]

17. Bibi, N.; Shah, I.; Alsabie, A.; Ali, S.; Lone, S.A. Electricity Spot Prices Forecasting Based on Ensemble Learning. *IEEE Access* 2021, 9, 150984–150992. [CrossRef]

18. Lago, J.; de Ridder, F.; de Schutter, B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Appl. Energy* 2018, 221, 386–405. [CrossRef]

19. Shah, I.; Ifitikhar, H.; Ali, S.; Wang, D. Short-term electricity demand forecasting using components estimation technique. *Energies* 2019, 12, 2532. [CrossRef]

20. Aslam, S.; Herodotou, H.; Mohsin, S.M.; Javaid, N.; Ashraf, N.; Aslam, S. A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. *Renew. Sustain. Energy Rev.* 2021, 144, 110992. [CrossRef]

21. Cui, J.; Yu, R.; Zhao, D.; Yang, J.; Ge, W.; Zhou, X. Intelligent load pattern modeling and denoising using improved variational mode decomposition for various calendar periods. *Appl. Energy* 2019, 247, 480–491. [CrossRef]

22. Moosavi, S.H.S.; Bardsi, V.K. Satin bowerbird optimizer: A new optimization algorithm to optimize ANFIS for software development effort estimation. *Eng. Appl. Artif. Intell.* 2017, 60, 1–15. [CrossRef]

23. Hafeez, G.; Khan, I.; Jan, S.; Shah, I.A.; Khan, F.A.; Derhab, A. A novel hybrid load forecasting framework with intelligent feature engineering and optimization algorithm in smart grid. *Appl. Energy* 2021, 299, 117178. [CrossRef]

24. Imani, M.; Ghasseman, H. Residential load forecasting using wavelet and collaborative representation transforms. *Appl. Energy* 2019, 253, 113505. [CrossRef]

25. Zhou, D.; Ma, S.; Hao, J.; Han, D.; Huang, D.; Yan, S.; Li, T. An electricity load forecasting model for Integrated Energy System based on BiGAN and transfer learning. *Energy Rep.* 2020, 6, 3446–3461. [CrossRef]