

Article

A Frequency Decomposition-Based Hybrid Forecasting Algorithm for Short-Term Reactive Power

Jiabao Du ¹, Changxi Yue ², Ying Shi ^{1,*}, Jicheng Yu ^{2,*}, Fan Sun ³, Changjun Xie ¹ and Tao Su ¹

¹ School of Automation, Wuhan University of Technology, Wuhan 430070, China; jiabaodu@yeah.net (J.D.); jackxie@whut.edu.cn (C.X.); sutao_stan@163.com (T.S.)

² China Electric Power Research Institute, Wuhan 430070, China; yuechangxi@epri.sgcc.com.cn

³ Xinjiang Electric Power Research Institute of State Grid, Urumqi 830000, China; sfzp2012@163.com

* Correspondence: yingsh@whut.edu.cn (Y.S.); yujicheng@epri.sgcc.com.cn (J.Y.)

Abstract: This paper proposes a new frequency decomposition-based hybrid reactive power forecasting algorithm, EEMD-LSTM-RFR (ELR), which adopts a strategy of frequency decomposition prediction after ensemble empirical mode decomposition and then data reconstruction to improve the prediction ability of reactive power. This decomposition process can compress the high frequency of reactive power and benefits the following separate forecasting. Long short-term memory is proposed for the high-frequency feature of reactive power to deal with the forecasting difficulty caused by strong signal disturbance and randomness. In contrast, random forest regression is applied to the low-frequency part in order to speed up the forecasting. Four classical algorithms and four hybrid algorithms based on different signal decompositions are compared with the proposed algorithm, and the results show that the proposed algorithm outperforms those algorithms. The predicting index RMSE decreases to 0.687, while the fitting degree R^2 gradually approaches 1 with a step-by-step superposition of high-frequency signals, indicating that the proposed decomposition-predicting reconstruction strategy is effective.

Keywords: reactive power; forecasting algorithm; ensemble empirical mode decomposition; long short-term memory; random forest regression



Citation: Du, J.; Yue, C.; Shi, Y.; Yu, J.; Sun, F.; Xie, C.; Su, T. A Frequency Decomposition-Based Hybrid Forecasting Algorithm for Short-Term Reactive Power. *Energies* **2021**, *14*, 6606. <https://doi.org/10.3390/en14206606>

Academic Editor: Elyas Rakhshani

Received: 9 August 2021

Accepted: 28 September 2021

Published: 13 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the regular operation of power systems, various pieces of electrical equipment can exert their best performance under the rated voltage. However, with the frequent changes in the electrical load, the operating voltage also changes at the same time, and its quality depends on the balance of reactive power, the output of various reactive power sources in the system, which should be able to meet the requirements of the system load and network loss under the rated voltage; otherwise, the operating voltage will deviate from the rated value. Therefore, accurate forecasting of power load helps in maintaining the regular operation of the power system and the optimal management of energy resources [1]. In the last 40 years, short-term load forecasting has been widely studied. Many essential operations in the power system are closely related to reactive power, such as voltage/var optimization [2,3], power quality improvement [3], frequency control [4], and steady-state power flow analysis [5,6]. Accurate forecasting of short-term reactive power helps in the maintenance of the regular operation of the power system and the optimal management of energy resources [7], and helps to reduce the power loss of the power grid [8–10]. Nevertheless, existing forecasting algorithms for short-term reactive power struggle to reach the accuracy requirements in application due to the strong waveform randomness, noise, and local disturbance of reactive power.

These forecasting algorithms can be divided into statistical mathematical algorithms, machine learning-based algorithms, and hybrid forecasting algorithms. In the statistical

mathematical field, Nie et al. [11] applied multiple linear regression algorithms to medium-term and long-term power load forecasting. Wu et al. [12] introduced a random forest regression (RFR) method to improve the short-term power load forecasting. However, these algorithms based on mathematical statistics are poor in robustness with low accuracy for complex nonlinear systems. The classical representative algorithm based on machine learning is support vector machine (SVM) [13], which presents good results in the short-term forecasting of power load with strong periodicity [14]. However, it has a poor forecasting effect on power load data with significant random fluctuation. The neural network is widely used in nonlinear system forecasting for its strong nonlinear mapping ability. Gao et al. [15] applied the recurrent neural network (RNN) to power load forecasting. However, RNN cannot establish the dependencies between long time series data. Hochreiter et al. [16] designed the long short-term memory (LSTM) neural network to solve this problem, which alleviates gradient attenuation and better captures the dependencies between long time series data by controlling the flow of information. Ma et al. [17] proposed a power load forecasting algorithm that combined isolated forest (IForest) with LSTM neural network. However, the neural network has a slow convergence speed, quickly falling into the local minimum value and overfitting. Because the algorithms based on traditional mathematical and machine learning have their limitations, hybrid forecasting algorithms have emerged. The first type of algorithm consists of multiple forecasting algorithms, realizes the forecasting by means their weighted results, and obtains higher forecasting accuracy. Meanwhile, the second kind of algorithm uses a signal decomposition strategy. Fourier transform and wavelet transform are classical methods, which can extract multi-scale local features of load information. Wen et al. [18] decomposed the power signal by means of fast Fourier transform (FFT), which can accurately obtain the frequency domain features of the signal. Sun et al. [19] decomposed the power load into high-frequency and low-frequency by means of wavelet transform. This algorithm has advantages including dynamic time-frequency resolution and the adaptability of a wavelet basis. Huang et al. [20] proposed empirical mode decomposition (EMD), which decomposed a complex signal into a series of intrinsic mode functions (IMFs) and the remainder according to different time scales. This method has been widely used in signal processing. EMD does not need any pre-set basis functions except for the time scale features of the signal. Meanwhile, Kurbatskii et al. [21] proposed the two-stage adaptive approach for time series forecasting, which employed the EMD/HHT in feature extraction for forecasting the active/reactive power. Since EMD lacks decomposing load information with intermittent signals, Wu et al. [22] proposed ensemble empirical mode decomposition (EEMD). The noise-aided analysis is applied for promoting the anti-aliasing ability and solves the defect of mode aliasing caused by intermittent signals in EMD. EEMD is suitable for both stationary and non-stationary signals and does not need to pre-set the basis function. He et al. [23] decomposed the data before making predictions. Wu et al. [24] combined the EEMD algorithm and the LSTM algorithm. These prediction algorithms perform better than conventional forecasting algorithm. However, these studies only selected a suitable prediction algorithm for the signal as a whole. A single algorithm cannot extract the features of all signal signal decompositions well. Yang et al. [25] chose the suitable prediction algorithm for each signal decomposition. Although this method can further improve the prediction accuracy, it requires a unique structure for each signal, which is not universal. Compared with the conventional forecasting algorithm, the forecasting accuracy of the hybrid forecasting algorithm is improved, but the forecasting performance of reactive power with strong disturbance is still not good enough.

All in all, the difficulty of reactive power prediction lies in its strong randomness, high levels of noise, and violent local disturbance [26]. This paper proposes a short-term reactive power hybrid forecasting model based on EEMD-LSTM-RFR (ELR) to solve the above problems. Compared to previous studies, ELR adopts the hierarchical clustering algorithm, which can automatically classify signals without manually setting parameters. Then, ELR can extract the fine features of each class of signal decomposition, achieving

better prediction performance. Therefore, ELR has a certain versatility, which can be applied to various short-term signals. First, EEMD decomposes the reactive power data into several IMFs for deep mining data information, and the local features' representation at different time scales can remove the interference within different time scales. Subsequently, the LSTM neural network and the RFR algorithm are applied to predict different frequency components, respectively, and the prediction results are superimposed to reconstruct reactive power. In this paper, the reactive power data are noisy, with solid disturbance. The proposed algorithm realizes a better prediction and restoration of local details.

The reactive power forecasting algorithm based on ELR contributes to solving the two shortcomings of the existing prediction algorithms and provides a novel strategy for reactive power forecasting.

Aiming at combatting the difficulty in forecasting high-frequency components with strong randomness, the ELR algorithm adopts LSTM with a solid nonlinear fitting ability to improve the forecasting accuracy as much as possible. The local details of the reactive power can be effectively restored; therefore, the entire prediction result is closer to the actual value. To address the slow convergence and long training time of the neural network, the ELR algorithm directly uses the RFR algorithm to predict the low-frequency components. The case analysis proves that RFR accelerates the training speed and attains more accurate prediction results, which improves the training and prediction accuracy of the entire algorithm. For reactive power forecasting, the strategy of separately predicting after EEMD and signal reconstruction is proposed. The EEMD of reactive power obtains a high proportion of low frequency and a low proportion of high frequency, and this high-frequency compression benefits the following reactive power forecasting.

The rest of this paper is arranged as follows. Section 2 introduces the used dataset and proposes a reactive power multi-scale feature extraction algorithm. Section 3 introduces the theoretical background of related forecasting algorithms and proposes the structure of the hybrid reactive power forecasting framework. Section 4 presents the experiment and result analysis, while the conclusion is given in Section 5.

2. Analysis of Reactive Power and Feature Extraction Algorithm

Due to the weak periodicity and strong randomness of the reactive power data, choosing a suitable decomposition method to obtain in-depth information with regard to these data is meaningful. This paper analyzes the data features of the reactive power to select the appropriate decomposition method, and verifies its effectiveness.

2.1. Analysis of Reactive Power Features

A group of prefecture-level reactive power data for a city in East China are selected. The sampling interval is 15 min, and the total number of reactive power data is 10,340. Some of the data are depicted in Figure 1.

As can be seen in Figure 1, based on a certain periodicity, the reactive power data present strong local randomness in the red box, including a large number of disturbance signals (specifically a mixture of some intermittent signals) and noises, which reflect feature information under different time scales. Different scales information will be mixed, which makes the prediction algorithm poor. Therefore, we need to decompose the signal to extract multi-scale local features of the reactive power, making the prediction easier and more precise.

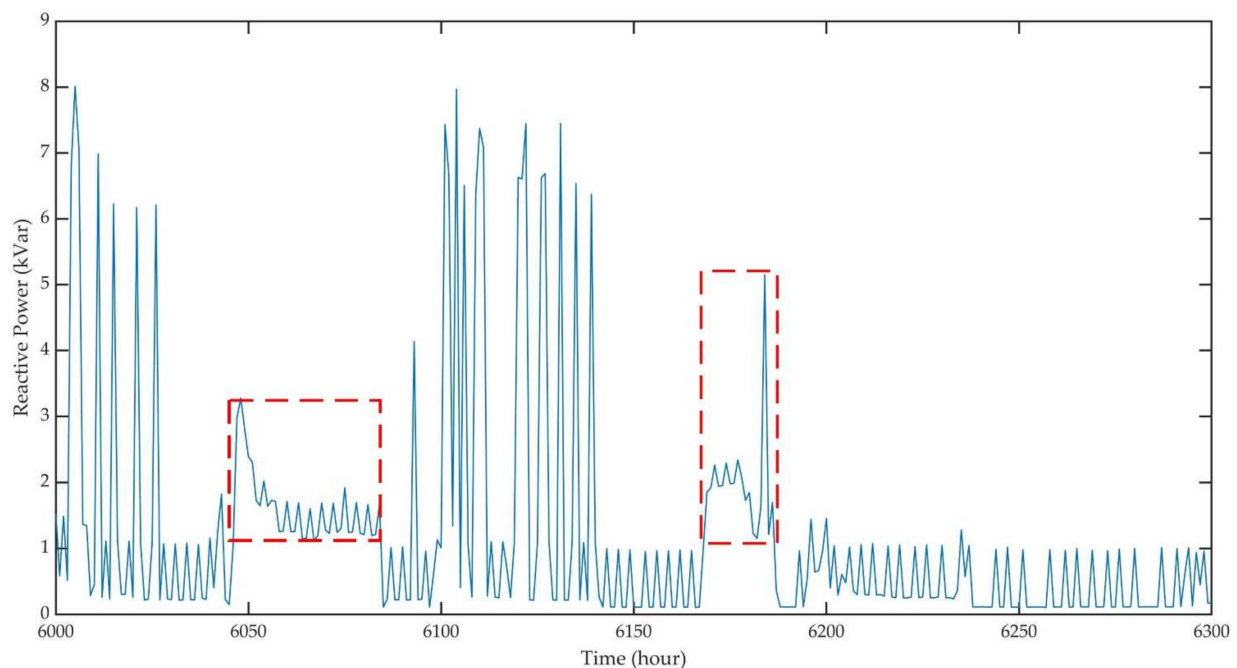


Figure 1. Data display of reactive power (part of the prefecture-level reactive power data of a city in East China).

2.2. Multi-Scale Feature Extraction Algorithm

There are many algorithms for signal decomposition. EMD is currently the most widely used algorithm, decomposing the complex signals into a linear combination of many IMFs with different time scales [27]:

$$X(t) = \sum IMF_k(t) + r(t) \quad (1)$$

where IMF_k is the k th intrinsic mode functions and r is the residual. The decomposed IMFs contain the local feature at different time scales, and the IMFs also need to meet the following two conditions:

- (1) The maximum difference between the number of extreme points and zero-crossing points is 1;
- (2) The average value of the local maximum and minimum is 0.

This algorithm decomposes the signal according to the time-scale characteristics of the data themselves. It does not need any basis function to be pre-set, making it better than Fourier decomposition and wavelet decomposition, which are based on the harmonic basis function and the wavelet basis function, respectively.

However, the distribution of the local extremum between intermittent signals and other signals is different, which causes the IMFs to fit the false envelopes. Since EMD is a non-difference algorithm, the IMFs must cover the other decomposed signals, resulting in mode mixing, and now these IMFs do not satisfy the uniqueness, which in turn brings about the difficulty of prediction. Thus, we have to find a way to deal with these intermittent signals, which is illustrated in Figure 2. It can be seen that IMF_8 and IMF_9 have similar time-scale signals, which shows typical mode mixing.

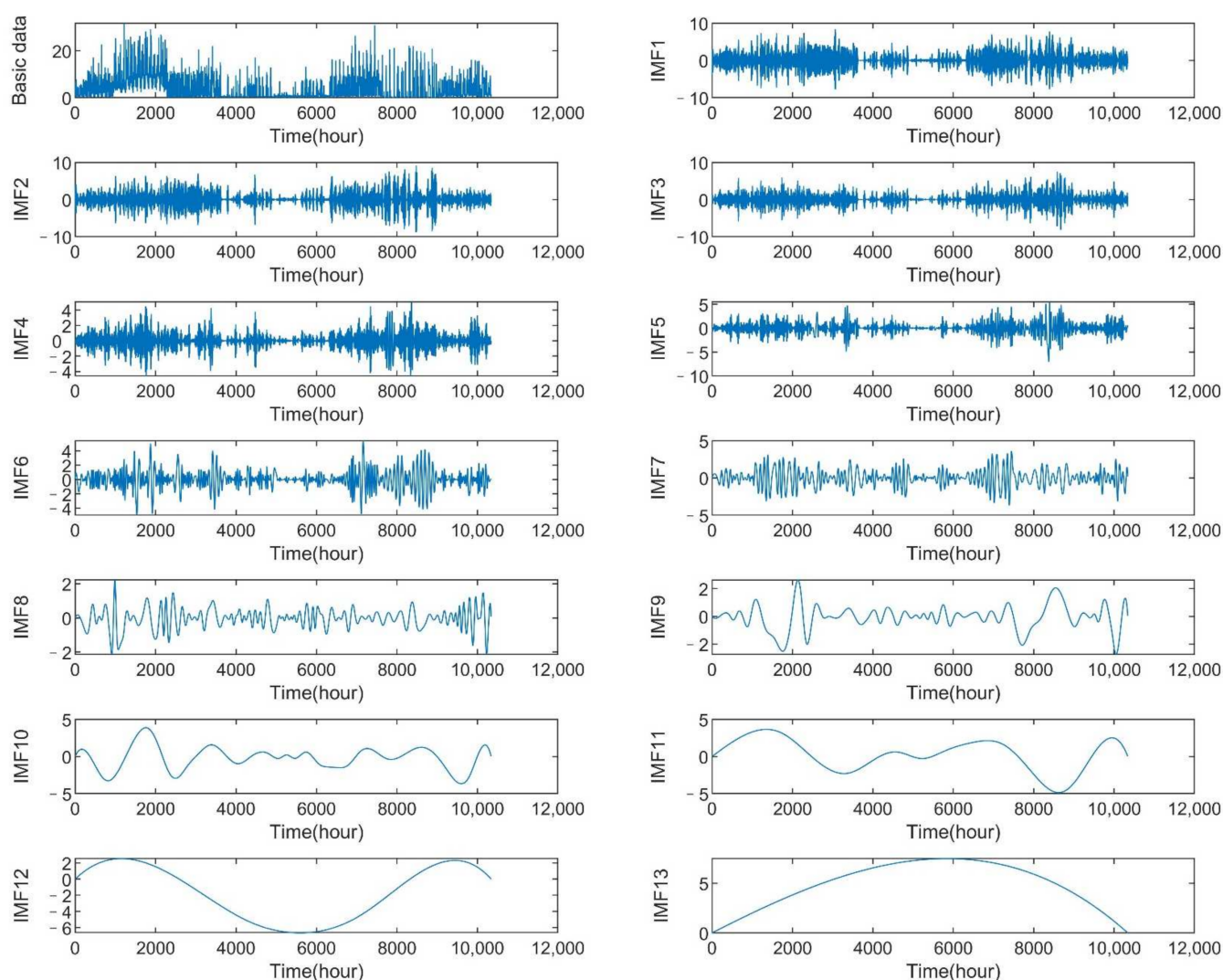


Figure 2. Empirical mode decomposition of reactive power.

EEMD enables EMD to decompose the reactive power with intermittent signals by adding white noise to the signal. White noise causes the local extremum to have a uniform distribution, which contributes to solving the mode mixing problem. From Figure 2, white noise will increase the signal-to-noise ratio of the reactive power data. Because the average value of white noise is 0, we use EMD to decompose the signal with white noise many times and take the average value; then, the white noise will be eliminated, and the predictable multi-scale local features of reactive power can be obtained in the end.

Therefore, EEMD is used to extract stable and effective multi-scale local features of reactive power. In this paper, EEMD has the following steps:

- (1) Add the white noise to the reactive power;
- (2) Confirm the maximum and minimum values in the target signal, and use cubic splines interpolation to fit the envelope. Moreover, record the mean of the maximum and minimum values as $m(t)$;
- (3) Calculate the residual value $r(t) = m(t) - x(t)$;
- (4) Repeat the above steps until the convergence condition is met.

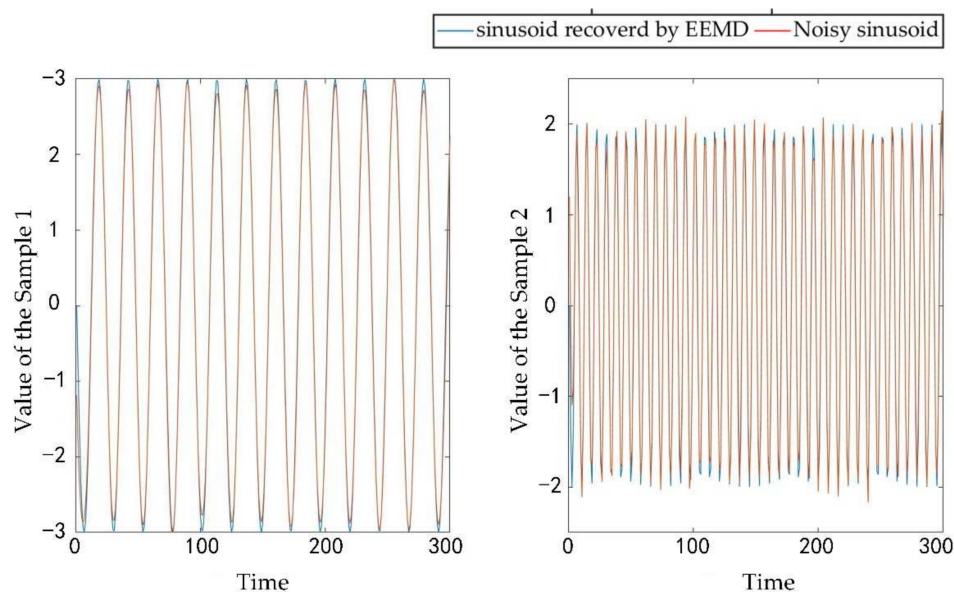
To verify the effectiveness of EEMD, this paper decomposes a signal of $N = 300$ samples consisting of a sinusoid with amplitude $H = 3$ and normalized frequency $\lambda_0 = 100$, and a sinusoid with amplitude $H = 2$ and normalized frequency $\lambda_1 = 300$, mixed with white Gaussian noise with a signal-to-noise ratio (SNR) equal to 20 dB [28]. EEMD obtains

this result and an unsupervised clustering algorithm, hierarchical clustering [29], and the number of classes is set to 2.

The quality of reconstruction factor (QRF) of an estimated component $\hat{x}$ relative to reference x is given by [29]:

$$QRF(\hat{x}, x) = 20 \log_{10} \left(\frac{\|x\|}{\|x - \hat{x}\|} \right) \quad (2)$$

As shown in Figure 3, EEMD allows us to recover the basic components, despite the presence of noise (the QRF of the component is higher or equal to the input SNR [20]).



(a) A sinusoid recovered by EEMD with QRF of 56 db

(b) A sinusoid recovered by EEMD with QRF of 43 db

Figure 3. Samples decomposed by EEMD.

EEMD decomposes the reactive power data to obtain n IMFs with frequencies from high to low and the remainder IMF_{13} . Figure 4 shows the decomposition components of reactive power data. The standard deviation of white noise is assigned as 0.2 times the signal standard deviation [20], and the value N is 300 [20].

It can be seen from Figure 4 that there is no obvious mode mixing phenomenon in each IMF component after EEMD, and each IMF has its unique feature at a specific stable frequency and disposes of the interaction between different time scales, which provides convenience for subsequent forecasts. Some IMFs reflect the overall trend of reactive power stability, and the others reflect the local random features with strong reactive power fluctuation. Therefore, this paper divides these IMFs into different categories according to their features, and then uses suitable prediction algorithms for each one. Here, we can see from Figure 4 that two categories are enough.

There are many algorithms for signal classification. Among them, some algorithms always show poor performance without manual feature and regularity analysis of given signals. Since the reactive power lacks regularity, we choose a hierarchical clustering algorithm, which can automatically classify signals without manually setting parameters.

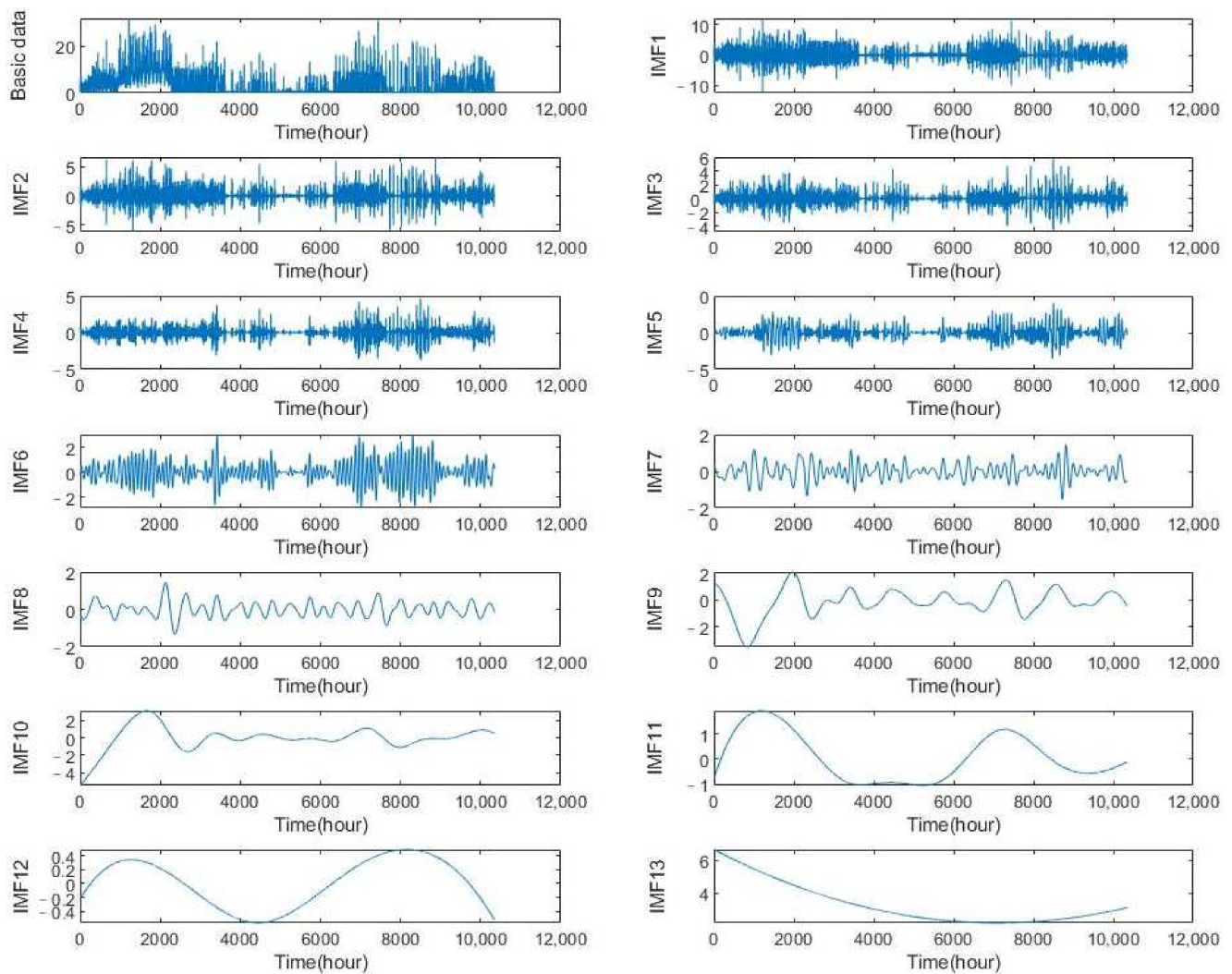


Figure 4. Ensemble empirical mode decomposition of reactive power.

In this paper, this algorithm divides IMFs into n sets. It then permits their reduction to $n-1$ mutually exclusive sets by considering the union of all possible $n(n-1)/2$ pairs and selecting a union with a maximal value for the functional relation or objective function [30]. By repeating this process until the number of categories we require is met, the complete hierarchical structure and the value of the similarity between every two sets can be obtained. The standard to evaluate the similarity between two IMFs is given by [28]:

$$d(x, y) = 1 - \frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|} \quad (3)$$

$$\text{With } \langle x, y \rangle = x^T y \text{ and } \|x\| = \sqrt{\langle x, x \rangle} \quad (4)$$

where x and y represent a time series.

After hierarchical clustering, IMF_{1-5} are classified as the high-frequency part, while IMF_{6-13} are classified as the low-frequency part. The high-frequency part presents reactive power's local randomness, disturbance, and noises, while the low-frequency part reflects its inherent periodicity.

To verify the advantages of EEMD in reactive power decomposition, this paper uses permutation entropy (PE) as a standard to evaluate the predictability of components decomposed by EMD and EEMD. The smaller the value of PE, the higher the predictability and the better effect of decomposition. We map the reactive power data into k sequences,

and then arrange each sequence in ascending order. There are $k!$ kinds of permutation methods with k sequences. The probability of the i th sequence occurring is P_i . Then, the value of PE can be calculated as [30]:

$$PE = -\sum_{j=1}^k P_j \ln(P_j) \quad (5)$$

The PE values of the high-frequency component and low-frequency component decomposed by EMD are 0.9824 and 0.7766, respectively, while those decomposed by EEMD are 0.9734 and 0.5429, respectively. This suggests that components decomposed by EEMD have better predictability. It can be concluded that using EEMD to decompose the reactive power has a better effect, as shown in Figure 5. This result is obtained by means of EMD/EEMD and hierarchical clustering, a kind of unsupervised clustering, with the number of classes set to 2. Compared with EEMD, more high-frequency signals are mixed in the low-frequency component of EMD, resulting in mode mixing. Because the mixed high-frequency signal will increase the difficulty of prediction, the low-frequency component of EMD has worse predictability. Therefore, we choose EEMD to decompose the reactive power.

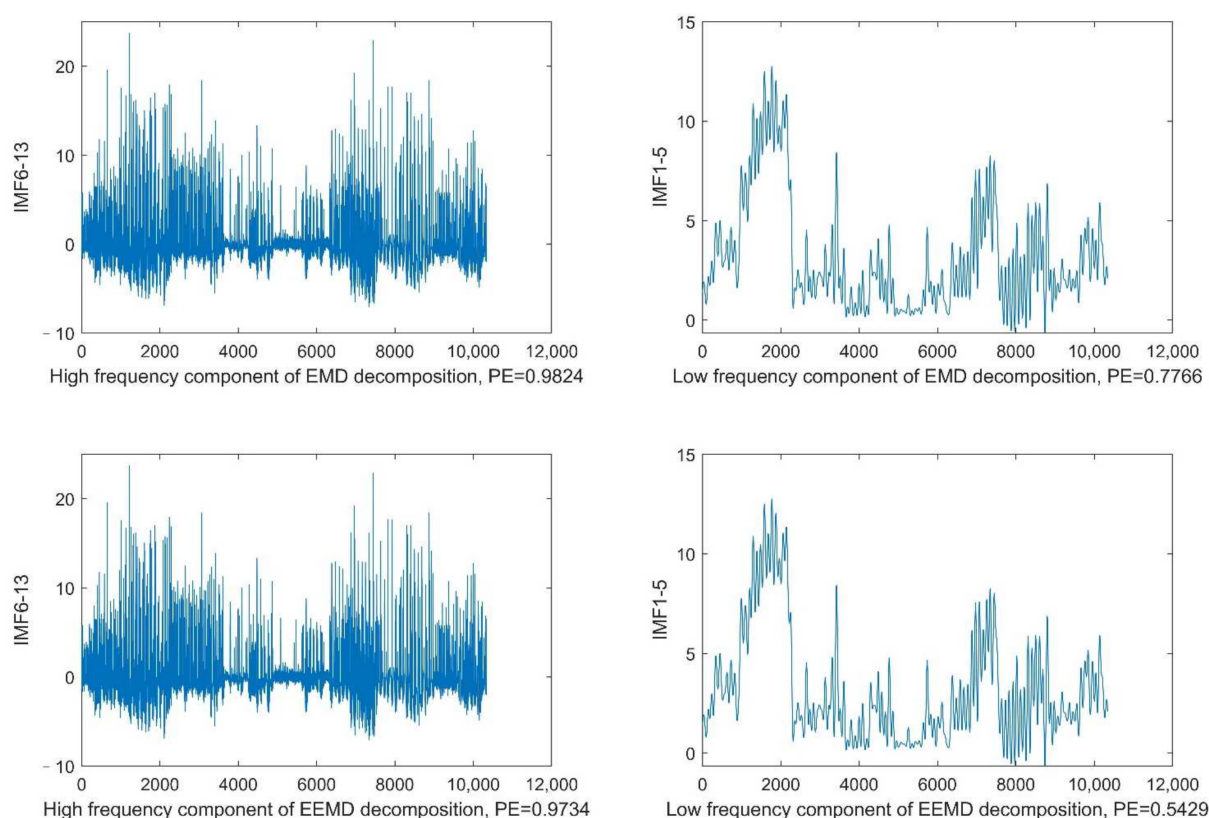


Figure 5. Components of the EMD/EEMD decomposition.

3. Hybrid Forecasting Algorithm of Reactive Power

EEMD decomposes one signal into dozens of IMFs, causing the data volume to increase and making the calculation of training data more difficult. To improve the training speed and forecasting accuracy, it is necessary to select the appropriate forecasting algorithm according to the time-frequency features of each feature component. In this paper, a hybrid forecasting algorithm of short-term reactive power based on multi-scale local feature extraction is proposed. Firstly, EEMD is used to decompose the reactive power data. Then, according to the different time-frequency features of each feature, the appropriate

algorithm is selected for forecasting. Except for the forecasting accuracy, high training speed is also a performance index of the algorithm. Finally, the forecasting results are superimposed to reconstruct the reactive power.

3.1. Research on the Different Time-Frequency Feature Forecasting Algorithm

This algorithm consists of RFR for the prediction of high-frequency signals and LSTM for the prediction of low-frequency signals. The machine learning-based methods consider fewer parameters, which have lower computational complexity and shorter training time. The deep learning-based methods consider more parameters, which have higher computational complexity and longer training time. In this research, a 1.2 GHz GPU with 16 GB memory and Windows10 system was used for all the experiments, and algorithms were programmed by Python3.6.

For the strong periodicity and high linearity of the low-frequency components, compared with deep learning algorithms, machine learning algorithms have higher training speed and stronger interpretability, while the accuracy can also meet the requirements in application [31].

Moreover, compared with mainstream machine learning algorithms such as support vector regression (SVR) and logistic regression (LR), RFR can more easily obtain an accurate forecasting value and avoids tedious parameter tuning [32]. It is an ensemble learning algorithm that can improve forecasting accuracy by integrating the regression tree (RT), showing a powerful generalization without overfitting phenomena usually occurring in neural networks.

The RFR is obtained by integrating RTs based on the loss minimization [12]:

$$MSE_{\min} = \min \left\{ \frac{1}{s} \sum_{i=1}^s \left(\frac{1}{n} \sum_{j=1}^n BRT_j(x_i) - y_i \right)^2 \right\} \quad (6)$$

where $BRT_j(x_i)$ is the IMF forecasting value of the j th RT, while x_i and y_i are the values of the i th feature input vector and the actual IMF value, respectively. The forecasting results of low-frequency features of reactive power are finally obtained by [12]:

$$Y_{predict} = \frac{1}{n} \sum_{p=1}^n C(p) BRT_p(E) \quad (7)$$

where the weight $C(p)$ is the sum of the feature correlation coefficients of the p th RT; $BRT_p(E)$ is the output value of the p th RT [33]. In this paper, the number of decision trees is set to 100.

The high-frequency components present strong randomness and volatility; traditional machine learning algorithms may not be able to accurately capture the intrinsic features, while neural networks are robust for nonlinear function fitting, such as back propagation neural network (BPNN), convolutional neural network (CNN), and RNN. RNN is the most mainstream algorithm for time series forecasting by constructing connections between adjacent neurons to capture time information [34].

Since the reactive power data length in practical applications can reach 40 or even longer, some algorithms such as RNN cannot establish the dependencies between long time series data. They will pass all the information to the adjacent neurons in forward propagation. The redundant information will accumulate in the process of propagation, which weakens the dependency information. However, the dependencies between long time series data are established based on the dependency information mentioned before; these algorithms are not suitable for predicting reactive power.

To solve this problem, LSTM adds a channel of cell state input and output and realizes the memory and transmission of data information through three gating units: forget gate, input gate and output gate, whose function is to control the flow of information in the

network, such as filtering out the redundant information. The function of the gating unit is presented as [35]:

$$g_t(x) = \sigma(x_t W_x + h_{t-1} W_h + b) \quad (8)$$

$$C_t = F_t \otimes C_{t-1} + I_t \otimes \tilde{C}_t \quad (9)$$

where W_x and W_h are the weights of the gating unit; x represents the input time series data; σ denotes the sigmoid function, which aims to add the nonlinear expression and enhance the fitting ability of the neural network; F_t and I_t represent the output of forget gate and input gate; $\otimes$ means to perform an element-wise multiplication; while C_t and $\tilde{C}_t$ in the formula are called the memory cell.

The forget gate controls the information in the memory cell of the previous time step, and decides whether the information is sent to the current time step. Moreover, the input gate decides whether input data will be used to train the network. If the value of the forget gate is close to 0 and the value of the input gate is close to 1, memory cells will discard invalid information of the previous time step. Then, the proportion of the dependency information will increase. Using this structure, LSTM can alleviate the problem of gradient attenuation and better capture the dependencies between long time series data.

The high-frequency component sequence of the reactive power is inputted into LSTM neurons for the forwarding calculation. After the output of LSTM is obtained, the error of each neuron can be calculated backward by the mean square error (MSE) equation as [35]:

$$V_{MSE} = \frac{1}{n} \sum_{t=1}^m (V_{actual} - V_{predict})^2 \quad (10)$$

where V_{actual} represents the actual value of the high-frequency component, $V_{predict}$ is its prediction value outputted by LSTM, and m denotes the data number. We build a neural network with two hidden layers of LSTM. The number of neurons in these two hidden layers is 128 and 32, respectively, the number of neurons in the input layer is 48, and the number of neurons in the output layer is 1.

After outputting the value of the neuron, the neuron weight is continuously optimized by Adam optimization toward the decrease in the error V_{MSE} . Through repeated iterations, the forecasting value of LSTM gradually approaches the actual value of the training set.

3.2. Validation of the Forecasting Algorithms

In this paper, RFR, LSTM, BPNN, and SVR are used to forecast a group of smooth time sequences and a group of strongly disturbed time sequences to verify the following algorithms' performance. The root mean square error (RMSE) and the determination of coefficient (R^2) are used to evaluate the forecasting accuracy. These two indexes show the predictive performance of the algorithm. The smaller the value of the root mean square error (RMSE) is, the more accurate the algorithm is in predicting the reactive power. In contrast, the closer the value of R^2 is to 1, the better the prediction model is.

RMSE is calculated as [36]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{actual} - Y_{predict})^2} \quad (11)$$

Additionally, R^2 is determined by [36]:

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_{actual} - Y_{predict})^2}{\sum_{i=1}^N (Y_{mean} - Y_{predict})^2} \quad (12)$$

where Y_{actual} and Y_{mean} are the actual value and its average value, respectively. $Y_{predict}$ is the prediction value; N is the total number of the forecasting data.

A portion of the forecasting results is shown in Figure 6. It can be seen that the prediction curves of RFR, LSTM, and BPNN follow the actual curve with small biases, while the prediction curve of SVR has a considerable bias.

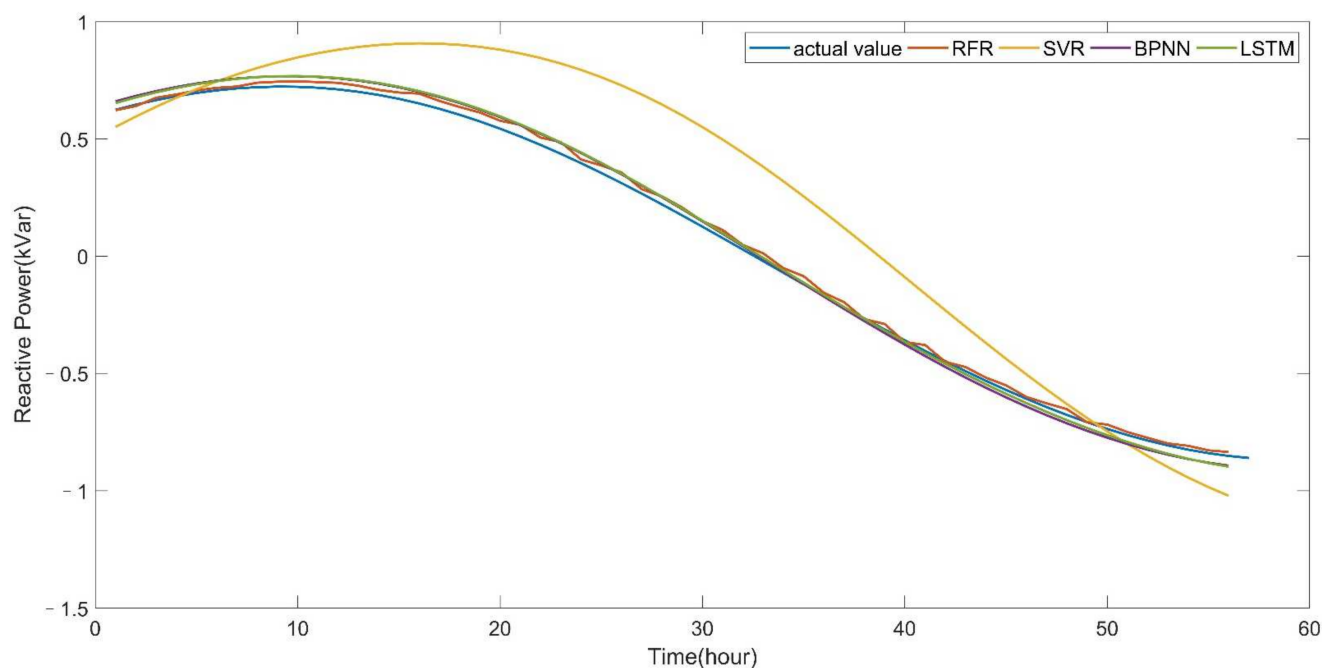


Figure 6. Forecasting on smooth sequences.

Table 1 shows the forecasting performance indexes of the four algorithms. RFR has an RMSE of only 0.024 and R^2 of 0.998, which shows that RFR outperforms SVR, BPNN, and LSTM. It has the highest forecasting accuracy and is very suitable for the forecasting of low-frequency component data.

Table 1. Forecasting index of each algorithm on smooth time sequences.

Algorithm	RFR	SVR	BPNN	LSTM
RMSE	0.024	0.264	0.042	0.036
R^2	0.998	0.780	0.994	0.996

In Table 2, LSTM has an RMSE of 0.762 and an R^2 of 0.419, indicating that LSTM is better than SVR, RFR, and BPNN, and the solid nonlinear fitting ability of LSTM can improve the forecasting accuracy of high-frequency components as much as possible.

Table 2. Forecasting index of each algorithm on strongly disturbed time sequences.

Algorithm	RFR	SVR	BPNN	LSTM
RMSE	0.778	1.026	0.786	0.762
R^2	0.396	−0.051	0.384	0.419

3.3. Research on the High-Frequency Feature Forecasting Algorithm Hybrid Reactive Power Forecasting Algorithm Based on EEMD-LSTM-RFR

This paper presents a hybrid forecasting algorithm for short-term reactive power based on EEMD-LSTM-RFR, which is illustrated in Figure 7.

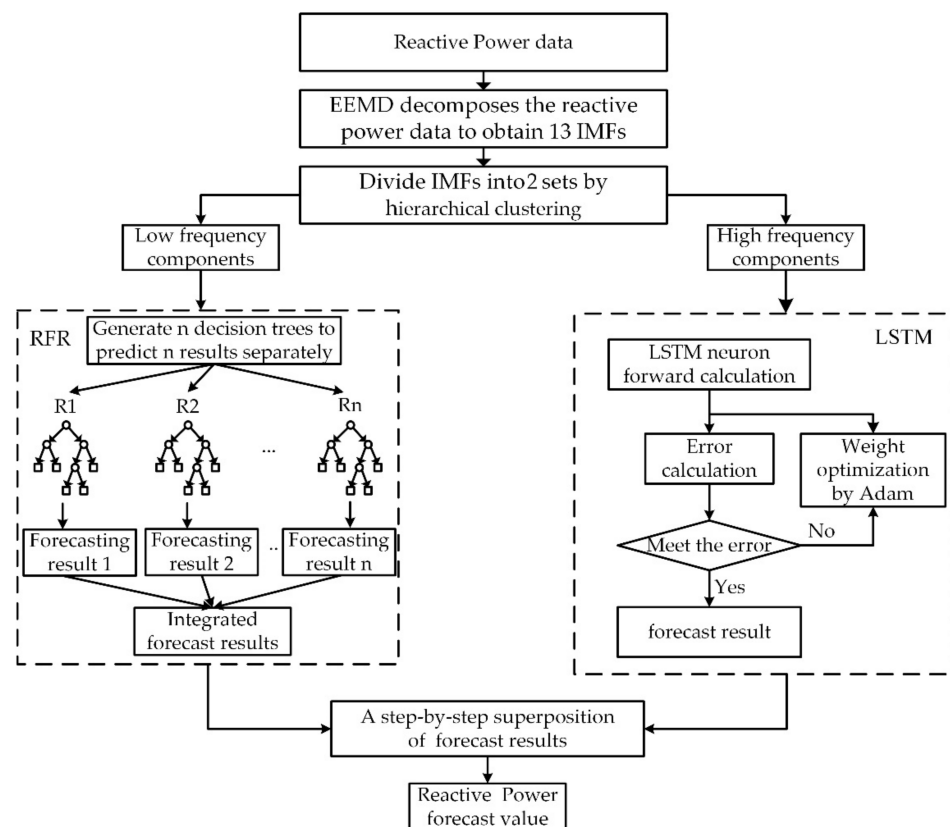


Figure 7. EEMD-LSTM-RFR hybrid forecasting framework.

EEMD is used to decompose a group of components with different time scales for the historical reactive power data, and each component's data are normalized to a range (0,1). After hierarchical clustering, the IMFs are classified as the high-frequency part and the low-frequency part. For the high-frequency component with strong randomness, LSTM is applied for its forecasting. For the low-frequency component with strong periodicity, RFR is used for rapid forecasting. The forecasting results are denormalized and superimposed to obtain the forecasting results of reactive power.

4. Case Analysis

This paper uses the data of reactive power in one city in China for experiments and analysis, recording 10,340 data in 29 days, which can be divided into three parts. The first part is the training set, recording 6204 data. The second part is the test set, recording 2068 data. The third part is the validation set, recording 2068 data. The proposed EEMD-LSTM-RFR hybrid forecasting algorithm is compared with four conventional prediction algorithms and four hybrid prediction algorithms based on signal decomposition. Finally, the effectiveness of the forecasting strategy and superposition reconstruction is verified.

4.1. EEMD-LSTM-RFR Hybrid Forecasting Results and Analysis

Figures 8 and 9 show the forecasting results of the EEMD-LSTM-RFR algorithm on this dataset. Figure 9 presents an enlarged view of the data in the black box of Figure 8. It can be seen that most of the forecasting points fit the real value better, and the forecasting effect at the peak decreases compared with that at the flat area, indicating that the forecasting of high-frequency signals with strong randomness is more difficult. The forecasting index R^2 reaches 0.943, and the RMSE error is 0.687, which indicates that the EEMD-LSTM-RFR algorithm can accurately predict reactive power.

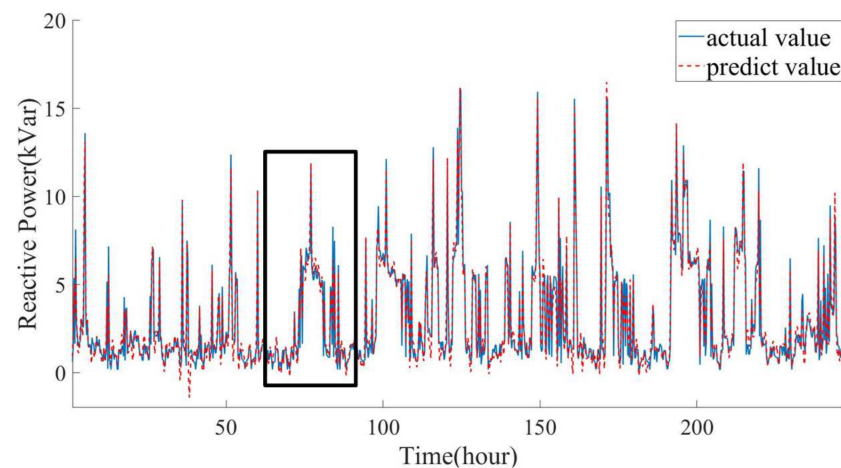


Figure 8. Forecasting results of EEMD-LSTM-RFR.

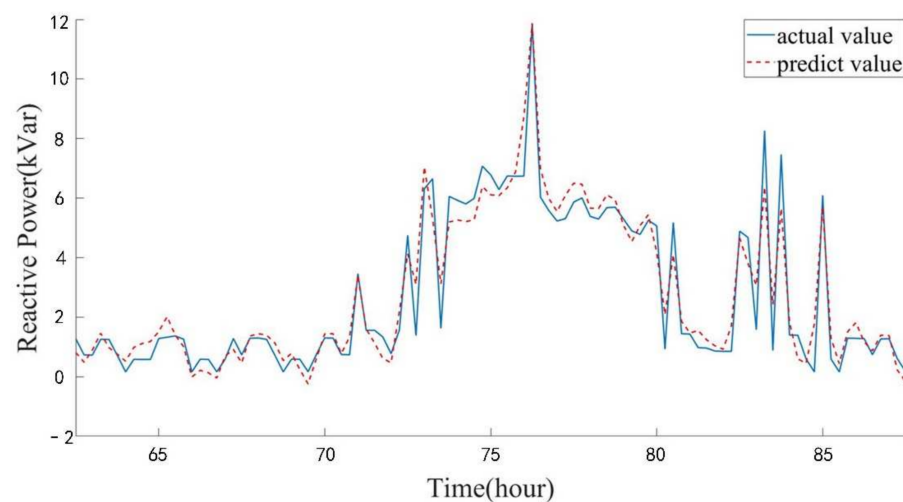


Figure 9. Partial enlarged drawing.

4.2. The Comparative Experiment of the Forecasting Algorithm

To verify the advantages of EEMD-LSTM-RFR in reactive power forecasting, four groups of mainstream conventional forecasting algorithms and four groups of hybrid forecasting algorithms are used, and 12 h of data are predicted—the results are shown in Figures 10a,b and 11a,b. For short, EEMD-LSTM-RFR is denoted as ELR.

4.2.1. Contrast Experiment with a Conventional Forecasting Algorithm

The conventional forecasting algorithms are compared with BPNN, LSTM, SVR, and RFR.

As shown in Figure 10a, the forecasting curve of the ELR algorithm proposed in this paper always follows the trend of the actual data curve, and the forecasting result is far better than the other four conventional forecasting algorithms. Figure 10b shows that the forecasting RMSE and R^2 of this algorithm are less than 0.7 and close to 1, respectively, while other algorithms are higher than 2 and less than 0.5, respectively, which shows the highest forecasting accuracy of the ELR method. The proposed algorithm experiences the lowest RMSE and the highest R^2 , maintaining its best forecasting performance.

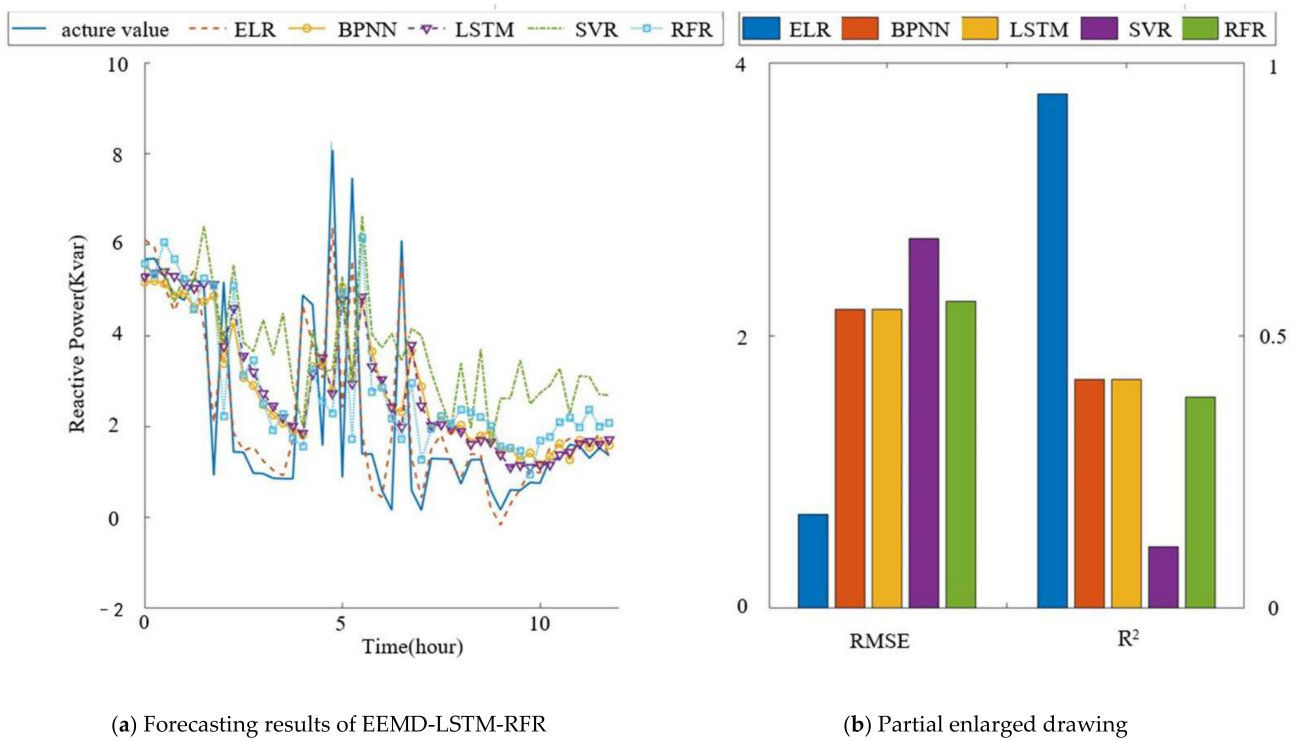


Figure 10. Results of four groups of conventional forecasting algorithms.

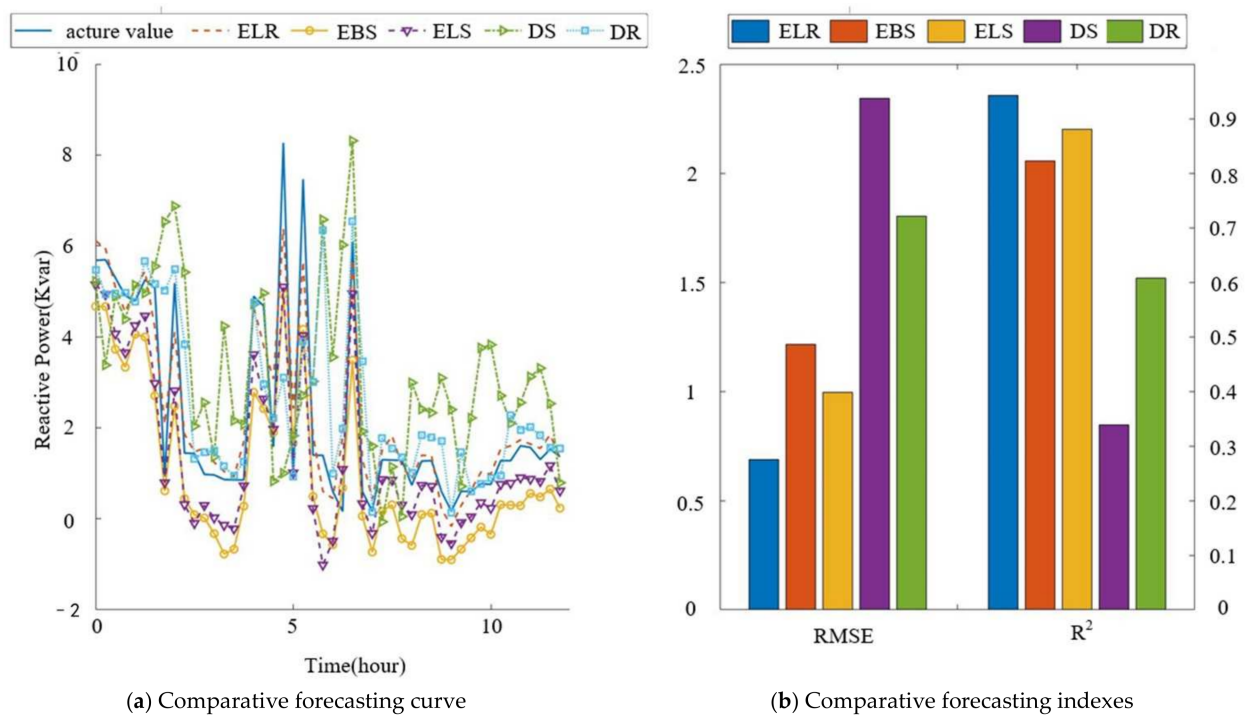


Figure 11. Results of four groups of hybrids forecasting algorithms.

4.2.2. Comparison with Hybrid Forecasting Algorithm Based on Signal Decomposition

The comparative hybrid forecasting algorithms based on signal decomposition are as follows:

- (1) After using EEMD to decompose the reactive power data, SVR is used for the low-frequency part and BPNN is used for the high-frequency part, abbreviated as EEMD-BPNN-SVR (EBS).
- (2) After using EEMD to decompose the reactive power data, SVR is used for the low-frequency part and LSTM is used for the high-frequency part, abbreviated as EEMD-LSTM-SVR (ELS).
- (3) After using discrete wavelet transform (DWT) to decompose the reactive power data, SVR is used to predict the result, abbreviated as DWT-SVR (DS).
- (4) After using DWT to decompose reactive power data, the result of forecasting using the RFR algorithm is abbreviated as DWT-RFR (DR).

Figure 11a,b show the forecasting curves, RMSE error, and determination coefficient R^2 of the five hybrid forecasting algorithms.

Figure 11a shows that the RMSE and R^2 of the ELR algorithm are still significantly better than those of EBS, ELS, DS, and DR. Figure 11b shows that ELR has the lowest RMSE and the highest R^2 , indicating that compared with other hybrid forecasting algorithms, ELR still has the best forecasting performance, which confirms that the signal decomposition and prediction method selected in this paper has better quality. The average RMSE of the four hybrid forecasting algorithms is 1.589, and the average R^2 is 0.662, which is much better than the average RMSE (2.337) and the average R^2 (0.335) of the four conventional forecasting algorithms, which indicates that the hybrid forecasting algorithm performs better in reactive power forecasting. The RMSE of the forecasting algorithm based on EEMD is lower than 1.3, while R^2 is higher than 0.8. Comparison with the DWT-based forecasting algorithm (where the RMSE is higher than 1.8 and R^2 is lower than 0.7) indicates that EEMD has a better effect on improving forecasting accuracy for the reactive power.

4.3. Verification of Superposition Reconstruction Effect

To further demonstrate and analyze the forecasting process of the ELR algorithm, the partial superposition reconstruction process of forecasting results is given as follows. In this paper, the high-frequency components IMF_5 , IMF_4 , IMF_3 , IMF_2 , and IMF_1 are gradually superimposed on the low-frequency components IMF_{6-13} , as shown in Figure 12. Figure 13 shows the RMSE error and the coefficient of determination R^2 after step-by-step superposition, where $IMF_{5-13} = IMF_{6-13} + IMF_5$, $IMF_{4-13} = IMF_{5-13} + IMF_4$ and so on.

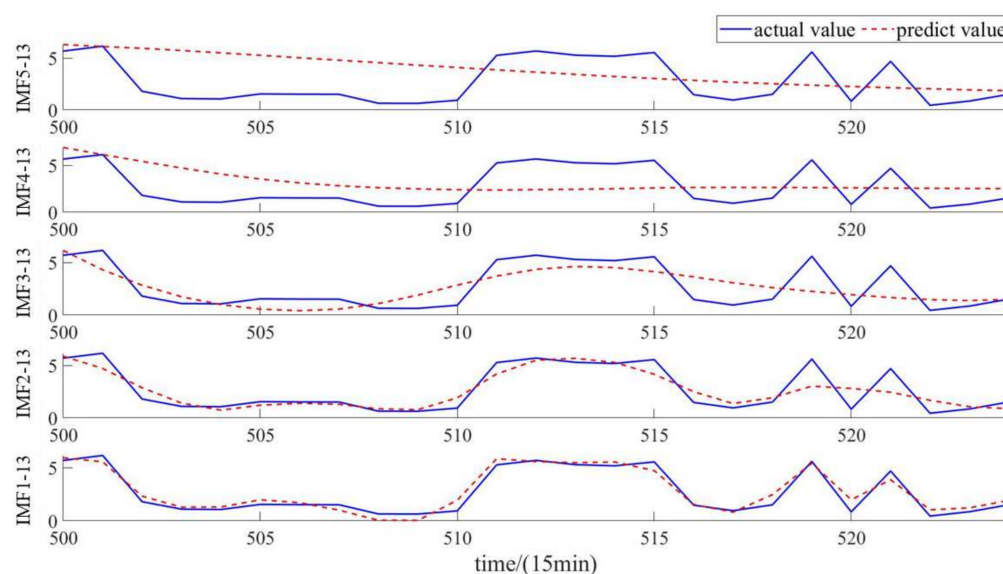


Figure 12. Process of reactive power signal reconstruction by feature superposition ($IMF_{5-13} = IMF_{6-13} + IMF_5$, $IMF_{4-13} = IMF_{5-13} + IMF_4$ and so on.).

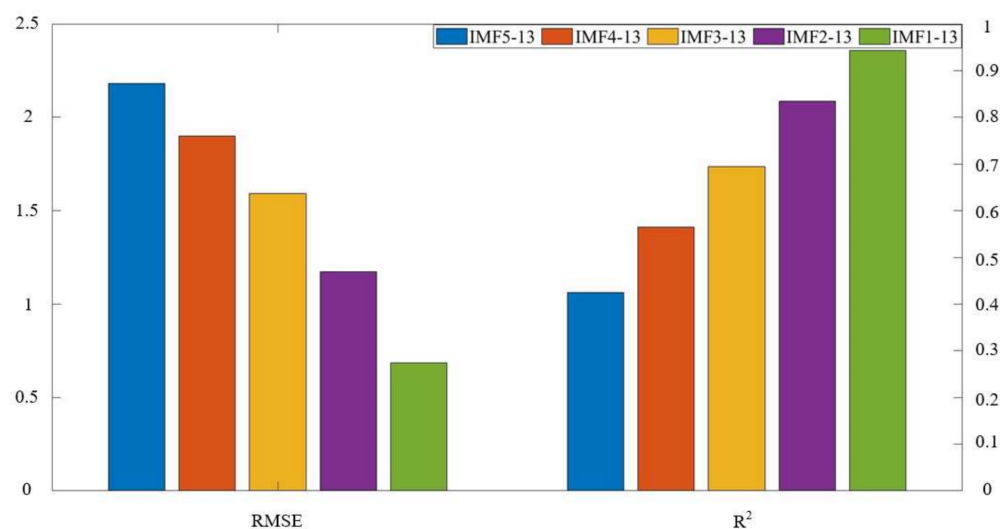


Figure 13. Forecasting index of the reactive power signal.

Figure 12 shows that the local forecasting details are gradually enriched with the increase in superimposed high-frequency signals. As shown in Figure 13, the value of RMSE decreases from 2.182 of IMF_{5-13} to 0.687 of IMF_{1-13} . IMF_{5-13} mainly includes the overall trend, while IMF_{1-13} includes more local details. With the continuous enhancement of high-frequency signals, the value of RMSE is on the decline. Additionally, the coefficient of determination R^2 gradually increases to 1 from 0.425, indicating that the forecasting curve is gradually approaching the actual value. The strategy of forecasting separately after the EEMD decomposition and then superimposing is effective for reactive power forecasting.

5. Conclusions

This paper proposes a hybrid forecasting algorithm based on EEMD-LSTM-RFR to analyze the forecasting of reactive power. Firstly, the reactive power data are decomposed into several IMFs by EEMD, which improves the subsequent reactive power forecasting. After hierarchical clustering, the IMFs are classified as the high-frequency part and the low-frequency part. Secondly, different algorithms are used to predict the IMFs of the high-frequency and low-frequency parts, respectively. Finally, the forecasting results are superimposed to reconstruct the predicted value of reactive power. The forecasting results of the proposed algorithm show that the hybrid forecasting index R^2 is 0.943, the RMSE is 0.687, and in total, the forecasting curve matches with the real value well.

Compared with four conventional forecasting algorithms, the results show that the RMSE of the proposed algorithm is lower than that of the others, with its R^2 closer to 1, indicating the highest forecasting accuracy. Compared with four hybrid forecasting algorithms based on signal decomposition, the results show that the proposed algorithm has the lowest RMSE and the highest R^2 , and still has the best forecasting performance.

Finally, the signal reconstruction results show that the RMSE decreases from 2.182 of IMF_{5-13} to 0.687 of IMF_{1-13} with the step-by-step superimposed high-frequency signals. At the same time, the fitting degree R^2 gradually increases to 1 from 0.425, which indicates that the forecasting curve gradually approaches the actual value. It can be concluded that the strategy of separately predicting after EEMD and signal reconstruction is effective for reactive power forecasting.

The accurate forecasting of reactive power can optimize the power flow calculation of smart grids and microgrids, assist the power market to test the technical feasibility of the energy path from the power plant to the load, realize the optimal management of energy resources, and effectively improve the operation performance of the power grid. However, the algorithm cannot be widely applied to the prediction of long-term data. In our future work, we will focus on improving the ability of the algorithm to transfer learning and its

forecast accuracy regarding long-term data. We will also use cross-validation in the next step to test its generalization and universality.

Author Contributions: Conceptualization, J.D. and Y.S.; methodology, J.D., C.Y. and F.S.; software, J.D.; validation, J.D., C.Y. and F.S.; formal analysis, J.Y. and C.X.; investigation, J.D. and Y.S.; data curation, J.D. and J.Y.; writing—original draft preparation, J.D.; writing—review and editing, J.D., C.X. and T.S.; visualization, J.D. and T.S.; supervision, J.D. and Y.S.; project administration, Y.S. and J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the State Grid Corporation of China Headquarter Science and Technology Project, grant number 5600-201918181A-0-0-00.

Data Availability Statement: The datasets generated during this study are available from the corresponding author on reasonable request.

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

Nomenclature

BPNN	Back Propagation Neural Network
CNN	Convolutional Neural Networks
DR	DWT-RFR
DS	DWT-SVR
DWT	Discrete Wavelet Transform
EBS	EEMD-BPNN-SVR
EEMD	Ensemble Empirical Mode Decomposition
ELR	EEMD-LSTM-RFR
ELS	EEMD-LSTM-SVR
EMD	Empirical Mode Decomposition
FFT	Fast Fourier Transform
IMF	Intrinsic Mode Function
LR	Logistic Regression
LSTM	Long Short-Term Memory
MSE	Mean Square Error
PE	Permutation Entropy
QRF	Quality of Reconstruction Factor
R2	determination of coefficient
RFR	Random Forest Regression
RMSE	Root Mean Square Error
RT	Regression Tree
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
SVR	Support Vector Regression

References

1. Rashid, K. Design, Economics, and Real-Time Optimization of a Solar/Natural Gas Hybrid Power Plant. Ph.D. Dissertation, The University of Utah, Salt Lake City, UT, USA, 2019.
2. Xu, Y.; Dong, Z.; Zhang, R.; Hill, D. Multi-Timescale Coordinated Voltage/Var Control of High Renewable-Penetrated Distribution Systems. *IEEE Trans. Power Syst.* **2017**, *32*, 4398–4408. [\[CrossRef\]](#)
3. Kim, Y.J.; Kirtley, J.L.; Norford, L.K. Reactive Power Ancillary Service of Synchronous DGs in Coordination with Voltage Control Devices. *IEEE Trans. Smart Grid* **2015**, *8*, 515–527. [\[CrossRef\]](#)
4. Shen, C.; Kaufmann, P.; Hachmann, C.; Braun, M. Three-stage power system restoration methodology considering renewable energies. *Int. J. Electr. Power Energy Syst.* **2018**, *94*, 287–299. [\[CrossRef\]](#)
5. Bracale, A.; Caramia, P.; Carpinelli, G.; Fazio, A.R. A Bayesian-Based Approach for a Short-Term Steady-State Forecast of a Smart Grid. *IEEE Trans. Smart Grid* **2013**, *4*, 1760–1771. [\[CrossRef\]](#)
6. Bracale, A.; Carpinelli, G.; De, P.; Hong, T. Short-term industrial reactive power forecasting. *Int. J. Electr. Power Energy Syst.* **2019**, *107*, 177–185. [\[CrossRef\]](#)
7. Qin, W.; Wang, P.; Han, X.; Meng, F. Risk analysis of power systems for both real and reactive power. *J. Mod. Power Syst. Clean Energy* **2013**, *1*, 150–158. [\[CrossRef\]](#)

8. Ren, H.; Wang, Z.; Sun, H.; Huang, S. Reactive power optimization planning of high-voltage distribution networks with technical standards being considered. *Power Syst. Technol.* **2020**, *44*, 1463–1472.
9. Tao, W.; Zhong, J.; Huang, H. Selecting and Optimization Method of Reactive Power Compensation Node in Power System. *Power Capacit. React. Power Compens.* **2018**, *39*, 116–121.
10. Peng, W.; Zhou, Q.; Chen, Z.; Zhang, N.; Chen, J. Practical and Simplified Calculation Method and Case Analysis on Reactive Power Balance Allocation for 500 kV Substation. *Power Capacit. React. Power Compens.* **2018**, *39*, 1–6.
11. Nie, H.; Wang, F.; Zhao, Y. Power load prediction based on multiple linear regression model. *Bol. Tec. Tech. Bull.* **2017**, *55*, 390–397.
12. Wu, X.; He, J.; Zhang, P.; Hu, J. Power System Short-term Load Forecasting Based on Improved Random Forest with Grey Relation Projection. *Autom. Electr. Power Syst.* **2015**, *39*, 50–55.
13. Chapelle, O.; Haffner, P.; Vapnik, V.N. Support vector machines for histogram-based image classification. *IEEE Trans. Neural Netw.* **1999**, *10*, 1055–1064. [[CrossRef](#)] [[PubMed](#)]
14. Bo, L.; Cheng, X.; Liu, X.; Zheng, H.; Hao, J. The Forecasting Model of Reactive Power Based on SVM. In Proceedings of the 4th IEEE Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqin, China, 12–14 June 2020.
15. Gao, X.; Ying, W.; Yang, G.; Sun, C.; Yue, Y. Shortterm Load Forecasting Model of GRU Network Based on Deep Learning Framework. In Proceedings of the 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Xi'an, China, 20–22 October 2018.
16. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
17. Ma, Y.; Zhang, Q.; Ding, J.; Wang, Q.; Ma, J. Short Term Load Forecasting Based on iForest-LSTM. In Proceedings of the 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), Xi'an, China, 19–21 June 2019.
18. Wen, H.; Guo, S.; Teng, Z.; Li, F.; Yang, Y. Frequency estimation of distorted and noisy signals in power systems by fft-based approach. *IEEE Trans. Power Syst.* **2014**, *29*, 765–774. [[CrossRef](#)]
19. Sun, W.; Bai, Y. Short-term load forecasting based on wavelet transform and BP neural network. In Proceedings of the 2nd International Conference on Mechanic Automation and Control Engineering IEEE, Shanghai, China, 15–17 July 2011.
20. Huang, N.; Shen, Z.; Long, S.; Wu, M.; Shih, H.; Zheng, Q.; Yen, N.; Tung, C.; Liu, C. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
21. Kurbatskii, V.G.; Sidorov, D.N.; Spiryaev, V.A.; Tomin, N.V. On the neural network approach for forecasting of nonstationary time series on the basis of the hilbert-huang transform. *Autom. Remote Control* **2011**, *72*, 1405–1414. [[CrossRef](#)]
22. Wu, Z.; Huang, N. Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [[CrossRef](#)]
23. He, F.; Zhou, J.; Feng, Z.K.; Liu, G.; Yang, Y. A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with bayesian optimization algorithm. *Appl. Energy* **2019**, *237*, 103–116. [[CrossRef](#)]
24. Wu, Y.X.; Wu, Q.B.; Zhu, J.Q. Improved eemd-based crude oil price forecasting using lstm networks. *Phys. A Stat. Mech. Its Appl.* **2019**, *516*, 114–124. [[CrossRef](#)]
25. Yang, Y.; Yang, Y. Hybrid Method for Short-Term Time Series Forecasting Based on EEMD. *IEEE Access* **2020**, *8*, 61915–61928. [[CrossRef](#)]
26. Rashid, K.; Sheha, M.N.; Powell, K.M. Real-time optimization of a solar-natural gas hybrid power plant to enhance solar power utilization. In Proceedings of the Annual American Control Conference (ACC), Milwaukee, WI, USA, 27–29 June 2018.
27. Wu, Z.; Huang, N.E. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. Math. Phys. Eng. Sci.* **2004**, *460*, 1597–1611. [[CrossRef](#)]
28. Harmouche, J.; Fourer, D.; Auger, F.; Borgnat, P.; Flandrin, P. The Sliding Singular Spectrum Analysis: A Data-Driven Nonstationary Signal Decomposition Tool. *IEEE Trans. Signal. Process.* **2018**, *66*, 251–263. [[CrossRef](#)]
29. Ward, J.H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. [[CrossRef](#)]
30. Cao, Y.; Tung, W.W.; Gao, J.B.; Protopopescu, V.A.; Hively, L.M. Detecting dynamical changes in time series using the permutation entropy. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* **2004**, *70*, 046217. [[CrossRef](#)] [[PubMed](#)]
31. Peyman, Z.; Saeid, P.; Alireza, A.; Brenton, S.M.; Siamack, A.S.; Brett, A.M. Random Forest regression prediction of solid particle Erosion in elbows. *Powder Technol.* **2018**, *338*, 983–992.
32. Wu, X.; He, J.; Yip, T.; Lu, J.; Lu, N. A Two-Stage Random Forest Method for Short-Term Load Forecasting; IEEE Power and Energy Society General Meeting: Boston, MA, USA, 2016.
33. Huang, N.; Lu, G.; Xu, D. A Permutation Importance-Based Feature Selection Method for Short-Term Electricity Load Forecasting Using Random Forest. *Energies* **2016**, *9*, 767. [[CrossRef](#)]
34. Khargharia, H.S.; Santana, R.; Shakya, S.; Ainslie, R.; Owusu, G. Investigating RNNs for Vehicle Volume Forecasting in Service Stations; IEEE Symposium Series on Computational Intelligence: Canberra, Australia, 2020.
35. Jin, Y.; Guo, H.; Wang, J.; Song, A. A Hybrid System Based on LSTM for Short-Term Power Load Forecasting. *Energies* **2020**, *13*, 6214. [[CrossRef](#)]
36. Wang, Z.; Xin, R.; Bai, T.; Zhao, J.; Wei, M.; Li, J.; Zhuang, L. The Power of Short-term Load Algorithm Based on LSTM. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *453*, 012056. [[CrossRef](#)]