sustainability

MDPI

Article

A Short-Term Power Load Forecasting Method of Based on the CEEMDAN-MVO-GRU

Taorong Jia, Lixiao Yao, Guoqing Yang * and Qi He

School of Electrical Engineering, Xi'an University of Technology, Xi'an 710048, China
* Correspondence: yanggq@xaut.edu.cn

Abstract: Given that the power load data are stochastic and it is difficult to obtain accurate forecasting results by a single algorithm. In this study, a combined forecasting method for short-term power load was proposed based on the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), Multiverse optimization algorithm (MVO), and the Gated Recurrent Unit (GRU) based on Rectified Adam (RAdam) optimizer. Firstly, the model uses the CEEMDAN algorithm to decompose the original electric load data into subsequences of different frequencies, and the dominant factors are extracted from the subsequences. Then, a GRU network based on the RAdam optimizer was built to perform the forecasting of the subsequences using the existing subsequences data and the associated influencing factors as the data set. Meanwhile, the parameters of the GRU network were optimized with the MVO optimization algorithm for the prediction problems of different subsequences. Finally, the prediction results of each subsequence were superimposed to obtain the final prediction results. The proposed combined prediction method was implemented in a case study of a substation in Weinan, China, and the prediction accuracy was compared with the traditional prediction method. The prediction accuracy index shows that the Root Mean Square Error of the prediction results of the proposed model is 80.18% lower than that of the traditional method, and the prediction accuracy error is controlled within 2%, indicating that the proposed model is better than the traditional method. This will have a favorable impact on the safe and stable operation of the power grid.

Keywords: short-term power load forecasting; CEEMDAN; RAdam; GRU; MVO

check for updates

Citation: Jia, T.; Yao, L.; Yang, G.; He, Q. A Short-Term Power Load Forecasting Method of Based on the CEEMDAN-MVO-GRU. *Sustainability* **2022**, *14*, 16460. https://doi.org/10.3390/su142416460

Academic Editor: Nicu Bizon

Received: 3 November 2022
Accepted: 2 December 2022
Published: 8 December 2022

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

Copyright: © 2022 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

The operation of the power grid becomes more complex with the large-scale access of distributed power sources and increasingly randomized electricity consumption patterns [1]. Addressing this issue requires carrying out high precision power load forecasting to accurately predict the load requirements of the power grid. It will help make reasonable adjustments in the operating processes of the power grid and improve the ability of the power grid in terms of continuous, safe, and stable operations. It is essential for optimizing the coordinated operations involving multiple sources of fire, wind, solar, and storage, building a modern digital and intelligent power grid, and further improving the operational mechanisms to meet the expectations and demands of the power sector [2].

Because of various factors, such as the user's power consumption behavior, meteorological factors, and economic policies, were affected the power load fluctuation, which makes it exhibiting both cyclical and short-term random characteristics [3,4]. Therefore, short-term power load forecasting is often affected by random factors and cannot accurately predict the power requirements at any point of time. At present, the commonly used short-term forecasting methods for power load are as follows: (1) time series methods based on statistics and (2) artificial intelligence methods with machine learning as the core [5–7]. The time series methods mainly include linear regression, Kalman filtering, and Exponential

smoothing techniques. These methods have certain advantages in terms of simple prediction models, simple data samples, and high computational efficiency [8,9]. However, these methods are not effective in nonlinear, and complicated data samples. Therefore, it is necessary to develop a model capable of extracting potential relationships among the time-domain features in the smart grid. Considering the complexity of the problem, the structure of the input data for load forecasting needs to be optimized so as to overcome the memory problem of multidimensional time series to obtain better forecasting accuracy.

As artificial intelligence network models have made significant strides in many fields, machine learning algorithms such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) have been used with good results in short-term load forecasting [10–13]. However, when a single prediction model encounters the power load prediction issues as a result of long sequence and multi-dimensional input, the loss of sequence memory, disordered data structure, and inability to deeply explore the potential characteristics of the sequence will occur, thus affecting the accuracy of load prediction. A study [14] used a combined prediction model comprising Convolutional Neural Networks (CNN) and LSTM algorithms to forecast the short-term load of regional power supply systems. In [15], a combined CNN-Bidirectional LSTM forecasting model based on a feature filtering technique has been proposed, to fully exploit the temporal characteristics in the electric load data and solved the problem of gradient disappearance, accelerated the training speed, and effectively analyzed and processed the potential information of power load data. Another study [16] made full use of the time-domain and frequency-domain features of load demand, used wavelet transform techniques to decompose the electric load time series into multiple subseries, and combined with deep neural networks for load forecasting, effectively improving the accuracy of short-term load forecasting. With the emergence of many new types of electricity loads (security systems, charging piles, etc.), a study designed a two-level intelligent feature engineering (IFE) and serial multi-timescale prediction framework to eliminate redundant and irrelevant features, balancing the stability of load trend prediction and the accuracy of load fluctuation detail prediction [17].

Meanwhile, to further improve the accuracy of short-term load forecasting, scholars combine signal decomposition algorithms with machine learning to further analyze and mine the potential patterns of load data changes [18,19]. A previous study [20] used the Variational Mode Decomposition (VMD) method to decompose the raw electric load data into multiple subseries, screened the influencing factors by constructing a composite variable selection algorithm, and then used Support Vector Regression (SVR) models for load forecasting to further explore the potential relationship between influencing factors and the frequency of load fluctuations. In [21], an integrated federated learning algorithm has been proposed by using VMD, the federated k-means clustering algorithm, and SecureBoost algorithm; the model used VMD to decompose the original data into several subsequences and used a clustering algorithm to reorganize the subsequences into feature clusters with common features, which enabled the model to effectively extract the implied features and simplified the complexity of the data, which in turn improved the prediction accuracy of the model. Two previous studies [22,23] used Empirical Mode Decomposition (EMD) or Ensemble Empirical Mode Decomposition (EEMD). EEMD decomposed load signals into multiple Intrinsic Mode Functions (IMFs) and established a prediction model respectively, which could better predict the fluctuation of load signals. The adaptive Gaussian white noise could be added to the decomposition process to eliminate the adverse effects of modal aliasing caused by EMD and EEMD on load prediction, and each modal component could be obtained by calculating the unique residual signal. Then, a complete EEMD with adaptive noise (CEEMDAN) algorithm was constructed [24]. When the decomposition process was finished by using this way, researchers find out that the reconstruction error was significantly low, and effectively re-solving the mode aliasing issue of EMD. Meanwhile, it overcame the problems of low decomposition efficiency and difficulty in eliminating noise while using previous algorithms [25,26]. Besides selecting a prediction model and analyzing signal decomposition, learning rate optimization has also become essential for

improving load prediction accuracy [27]. In [28], study revealed that applying traditional learning rate optimizers such as Root Mean Square Prop (RMSprop), Adaptive Moment Estimation (Adam), and Nesterov-accelerated Adaptive Moment Estimation (NAdam) in deep learning networks could easily to make the networklead training into local optimum, and the Rectified Adam (RAdam) optimizer was proposed. The convergence speed and calculation accuracy of the deep learning network were im-proved effectively. In [29], the scholars used deep convolutional neural networks and the RAdam optimizer for human-computer interaction for facial expression recognition. The results show that using the RAdam optimizer enables the network model to improve the recognition accuracy by 3% to 4% with better generalization. Through a comprehensive analysis of the above literature, we counted the main information of load prediction methods in the articles, and the statistical information is shown in Table 1.

Table 1. Statistical table of load forecasting methods.

Forecasting Methods		Representative Algorithms		Features
		Data Processing	Load Forecasting	
Traditional methods		-	Linear regression Kalman filtering Exponential smoothing techniques	Its simple and fast to compute. However, it cannot handle complex data samples.
Artificial intelligence methods	Single	-	RNN LSTM GRU CNN	It can handle more complex data. However, it has insufficient processing power and poor prediction accuracy when facing high dimensional data samples.
	Combined	VMD EMD EEMD K-Means		Its ability to mine potential features of data and achieve dimensionality reduction of data samples. However, a lot of work needs to be done in the future to find a more reasonable combination model.

According to the above analysis, this paper proposes a new hybrid forecasting system that combines data preprocessing, deep learning prediction method, and intelligent optimization algorithm to further improve the short-term electrical load forecasting accuracy. The study used the CEEMDAN decomposition algorithm to decompose the original historical load data into modes to address the issues of significantly randomized short-term load fluctuation and difficulty in selecting influencing factors and explored the fluctuation rules of the power load data. Then, the gray relational analysis (GRA) was performed to calculate the degree of correlation of influencing factors for different modal components. Finally, the GRU algorithm based on the RAdam optimizer was used for the short-term prediction of modal components, and the final short-term power load prediction results were obtained by superposition. The main contributions and innovations of this research are as follows:

1. In order to deeply analyze the variation pattern of load and find out the inner connection between load components of different frequencies and external influencing factors, the article uses CEEMDAN algorithm to separate the frequencies of load data. The changing characteristics of the load components at different frequencies are analyzed, and the load fluctuation patterns and customers' electricity consumption habits in the area are summarized.
2. A combinatorial model for short-term load forecasting is proposed. Based on the original GRU algorithm, the algorithm optimization is carried out using the RAdam optimizer. Meanwhile, the optimization of the internal parameters of the GRU algorithm is carried out using the MVO algorithm. Based on this, the network was trained using the frequency components of the load and the main influencing factors, and the load was predicted for the next 24 h.

2. Materials and Methods

2.1. CEEMDAN and GRA

The CEEMDAN algorithm could adjust the reconstruction error of the decomposed signal close to 0 by adding normal distributed white noise ($n^i(t)$) for a limited number of times at each stage of decomposition and could effectively solve the problems of mode aliasing and low computational efficiency of EEMD and other algorithms [30,31]. Its decomposition process involved the following six steps [32]:

The load curve $s(t) + \varepsilon_0 n^i(t)$ was decomposed repeatedly for N times, and the first modal component was obtained using the mean value as

$$IMF_{s1}(t) = \frac{1}{N}\sum_{i=1}^{N} IMF_1^t(t) \tag{1}$$

where $s(t)$ is the original signal, $\varepsilon_0 n^i(t)$ is the noise signal, N is the total number of modal decompositions and IMF_1^t is obtained from the t_{th} decomposition.

1. The allowance signal $r_1(t)$ of the load curve to $IMF_{s1}(t)$ was calculated as

$$r_1(t) = s(t) - IMF_{s1}(t) \tag{2}$$

2. The load margin signal $r_1(t) + \varepsilon_1 E_1(n^i(t))$ was decomposed repeatedly for N times to obtain the second modal component *IMF* as

$$IMF_{s2}(t) = \frac{1}{N}\sum_{i=1}^{N} (E_1(r_1(t)) + \varepsilon_1 E_1(n^i(t))) \tag{3}$$

3. For k = 2, 3, . . . , K, the K_{th} residual signal was calculated as

$$r_k(t) = r_{k-1}(t) - IMF_{s2}(t) \tag{4}$$

4. Step 3 was repeated to obtain the k + 1 modal function as

$$IMF_{sk+1}(t) = \frac{1}{N}\sum_{i=1}^{N} (E_1(r_k(t)) + \varepsilon_k E_k(n^i(t))) \tag{5}$$

5. Steps 4 and 5 were then repeated until the decomposition termination condition was reached, and K modal components of the load curve could be obtained. The final residual signal was calculated as

$$R(t) = s(t) - \sum_{k=1}^{K} IMF_{sk}(t) \tag{6}$$

The GRA was performed to calculate the correlation between the potential influencing factors of load fluctuations and the components decomposed using the CEEMDAN algorithm [33]. This helped select the influencing factors with a strong correlation to specific modal functions [34]. We selected the obtained modal function using the CEEMDAN algorithm and combined it with the known influencing factor matrix U to form the association analysis matrix X as

$$X = \begin{bmatrix} IMF_i & U \end{bmatrix} \tag{7}$$

Subsequently, we divide each column in the matrix X by the mean value of the elements of each column simultaneously to eliminate the effect of the magnitude of the quantities between different physical quantities on the calculation. After that, the correlation degree *KSI* of each column element can be calculated as

$$X' = \begin{bmatrix} \frac{IMF_i}{IMF_{\text{mean}}} & \frac{U_1}{U_{1\text{mean}}} & \frac{U_2}{U_{2\text{mean}}} & \cdots & \frac{U_N}{U_{N\text{mean}}} \end{bmatrix} = \begin{bmatrix} X'_0 & X'_1 & X'_2 & \cdots & X'_N \end{bmatrix} \tag{8}$$

$$T(i) = abs(X'_{i=1\to N} - X'_0) \tag{9}$$

$$KSI = (\min(T) + rho \times \max(T))./(T + rho \times \max(T)) \tag{10}$$

where *rho* is the correlation coefficient and usually takes values between (0, 1).

With the *KSI* obtained, the degree of correlation (*M*) between the influencing factors and the modal components could be calculated as [35,36]

$$M = \text{mean}(KSI) \tag{11}$$

2.2. Multiverse Optimizer Algorithm

The Multiverse Optimizer (MVO) algorithm is a cluster intelligence optimization algorithm based on the multiverse theory in physics [37]. It builds a mathematical model based on the cosmological theories of white holes, black holes, and wormholes. Among other things, wormholes connect all universes to the current optimal universe. It is able to share the optimal information in the algorithm to the various groups involved in the calculation. While white holes and black holes are built between the various clusters and are used to exchange information with each other. It's conceptual model of MVO algorithm is shown in Figure 1.

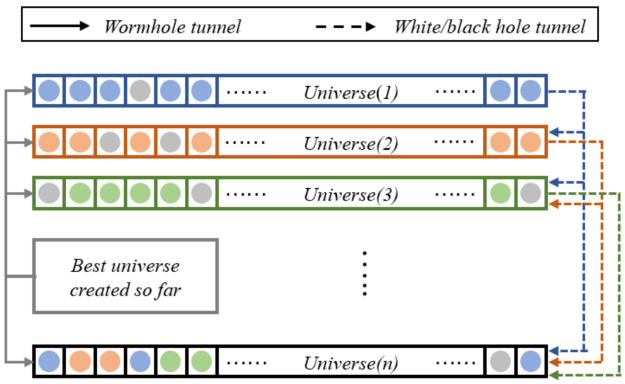


Figure 1. Conceptual model of MVO algorithm.

The algorithm defines the candidate solution as an initializing universe [38]. Assuming that the number of initializing population was *NP* and the vector dimension to be solved was *NQ*, the initial universe *U* could be expressed as

$$U = \begin{bmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_{NP} \end{bmatrix}^T \tag{12}$$

$$x_i{}^j = Q_{lj.\min} + r_1 * (Q_{lj.\max} - Q_{lj.\min}) \tag{13}$$

$$NI = \begin{bmatrix} NI_1 \\ NI_2 \\ \vdots \\ NI_{NP} \end{bmatrix} = \begin{bmatrix} f(x_1{}^1, x_1{}^2, \cdots x_1{}^{NQ}) \\ f(x_2{}^1, x_2{}^2, \cdots x_2{}^{NQ}) \\ \vdots \\ f(x_{NP}{}^1, x_{NP}{}^2, \cdots x_{NP}{}^{NQ}) \end{bmatrix} \tag{14}$$

where x_{ij} is the j_{th} dimension of the i_{th} solution vector, $Q_{li.\min}$ and $Q_{li.\max}$ are the lower and upper limits of the solution vector, and r_1 is a random number.

After the initializing universe (candidate solution set) was brought into the program to calculate the adaptive values of each solution set, the population updating and optimization process of the multiverse optimization algorithm could be carried out [39,40].

It is should be noted that the MVO algorithm had two essential parameters: Wormhole Existence Probability (WEP) and Travelling Distance Rate (TDR). During each iteration of

the calculation, the universe updates its own spatial position by the above parameters and gradually moves toward the optimal universe (optimal solution):

$$x_i^j = \begin{cases} \left.\begin{array}{l} x_{best.j} + TDR * ((Q_{lj.\max} - Q_{lj.\min}) * r_2) r_3 < 0.5 \\ x_{best.j} - TDR * ((Q_{lj.\max} - Q_{lj.\min}) * r_2) r_3 \geq 0.5 \end{array}\right\} r_4 < WEP \\ x_i^j \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad r_4 \geq WEP \end{cases} \tag{15}$$

where the $x_{best.\ j}$ is the element of the j_{th} dimension of the optimal solution and $r_{2\text{–}4}$ are random numbers.

2.3. GRU Algorithm Based on the RAdam Optimizer

The GRU model is a variant structure simplified from the traditional LSTM prediction model [41]. It combines the original forgetting gate and input gate of LSTM into a single update gate, which reduces one gate function compared with the original LSTM structure and has fewer parameters than LSTM, but it can achieve the same function as LSTM and has higher computational performance [42,43]. The schematic diagram of the GRU model is shown in Figure 2.

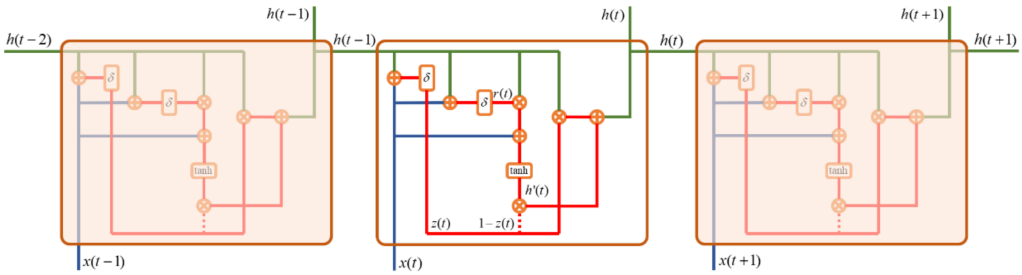


Figure 2. GRU algorithm framework.

In general, in the reverse update process of the load prediction algorithm, the fixed learning rate, the learning rate updated with the number of iterations, and the Adam optimizer are used to update the weights and bias values of the network model. In this study, the RAdam optimizer was selected to further optimize the learning rate of the algorithm. The application flowchart of the RAdam optimizer is shown in Figure 3.

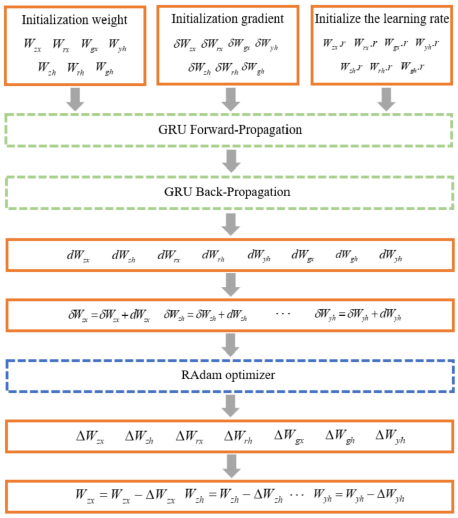


Figure 3. Application flow of RAdam optimizer.

In the reverse update process of the GRU algorithm, the error terms of each weight were obtained by taking the partial derivative of the total error function. Subsequently, the

error terms were brought into the RAdam optimizer to obtain the updated terms of weight parameters. Furthermore, the weight parameters were updated along the direction of the negative gradient until the total iteration limit was met (Figure 3).

3. Hybrid Forecasting System

This study used the CEEMDAN algorithm to decompose the historical load data into modal functions of different frequencies and residual functions. Then, the external dominant influencing factors with a higher degree of influence were selected for different frequency components using the GRA. Further, the specific modal functions and the corresponding dominant influencing factors were transferred to the MVO-GRU algorithm, and the number of super parameters in the GRU model was selected based on the optimization algorithm. Finally, the prediction results of different modal functions were recombined to obtain the final load prediction results. The CEEMDAN-MVO-GRU algorithm flow chart is shown in Figure 4.

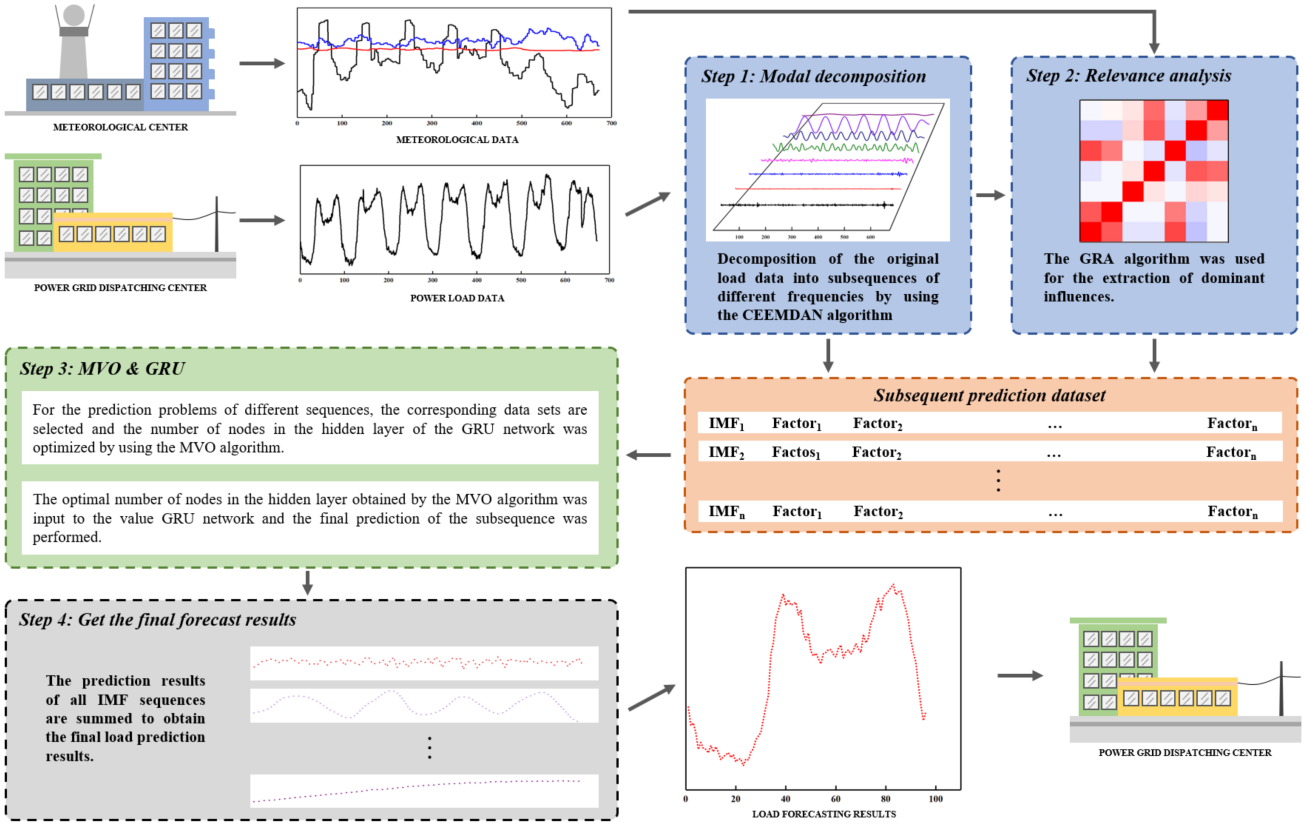


Figure 4. CEEMDAN-MVO-GRU algorithm flow.

In this hybrid forecasting system, the optimization search process of the MVO algorithm used the Root Mean Square Error (RMSE) and the similarity (R) of the prediction curves to obtain the fitness value (*NI*) to determine the number of nodes in the implicit layer in the prediction model. In addition, the Root Mean Square (MAE), Mean Absolute Percentage Error (MAPE) and Nash–Sutcliffe efficiency coefficient (NSE) were selected, which together with RMSE and R to evaluate the predictive performance of the proposed model [44–46]. The parametric optimization of the MVO algorithmic process is shown in Figure 5.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - f(x_i)\right)^2} \tag{16}$$

$$R = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{A_i - \mu_A}{\delta_A}\right)\left(\frac{B_i - \mu_B}{\delta_B}\right) \tag{17}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Y_i - f(x_i)| \tag{18}$$

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{Y_i - f(x_i)}{Y_i}\right| \tag{19}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}(Y_i - f(x_i))^2}{\sum_{i=1}^{N}(Y_i - \mu_Y)^2} \tag{20}$$

$$NI = \alpha \times RMSE + \beta \times R \tag{21}$$

where N is the number of samples, $f(x_i)$ is the predicted value, Y_i is the true value, μ is the mean value of the matrix, δ is the standard deviation of the matrix, and α and β are weight coefficients.

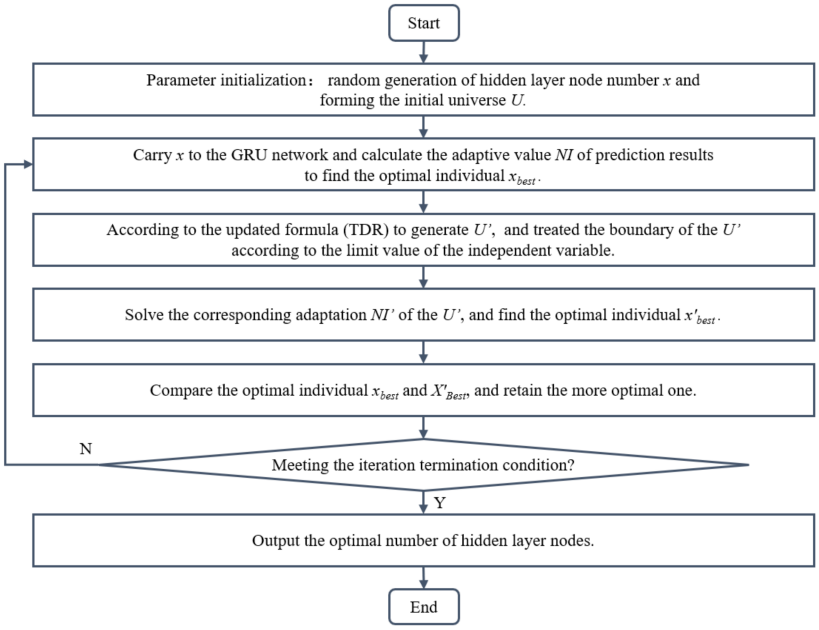


Figure 5. Parametric optimization of the MVO algorithm process.

4. Data Sets

4.1. Data Preprocessing

In short-term power load forecasting, the trend of load changes is often closely related to the load data of the recent period. Therefore, mainly the meteorological data of the load forecasting site and the load characteristics data were used to constitute the data set for short-term load forecasting to verify the forecasting method proposed in this study.

Among the load characteristics data, the actual load data of the DG Substation in Weinan city from 1 January 2021 to 31 January 2021 were considered for this study. The data sampling interval considered for this substation was 15 min, and the data from 2976 points were collected. Meteorological data is the hourly temperature and humidity of the area and other 9 collection types, a total of 744 points. The load characteristic data mainly consisted of the maximum, minimum, and average loads calculated by rolling historical load data every hour. The changes in sample data curves of some factors are shown in Figure 6.

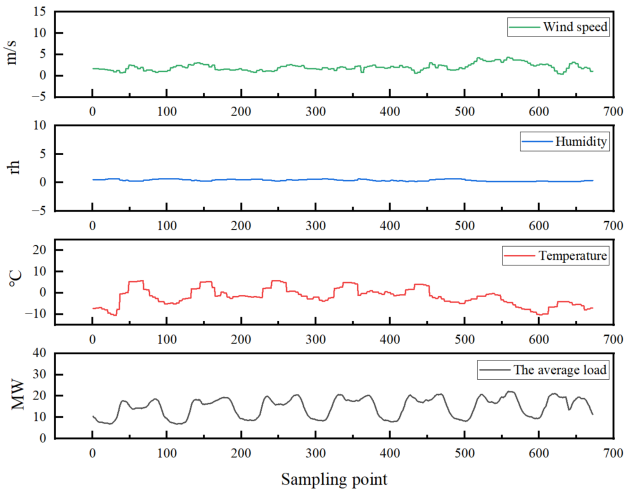


Figure 6. Sample data curve (First week of January 2021).

4.2. Substation Load Curve Decomposition Based on the CEEMDAN Algorithm

This study considered the daily load curve of the station from 1 January 2021 to 7 January 2021 to obtain the load curve to examine the changes in the load characteristics of the power substation (Figure 7). As discussed in Section 2.1, the adaptive Gaussian white noise could be added at each stage of the original power system load curve decomposition process, following which each modal component of the power load curve based on the CEEMDAN algorithm could be obtained. The modal function and residual load curves of the station in the first week of January 2021 are shown in Figure 8.

The load curve of the station consisted of a large number of high-frequency components, indicating that the power consumption behavior of consumers was highly random. However, the amplitude of the high-frequency component was smaller than that of the total power consumption at the same time, indicating that the load change was mainly affected by the long-term stable power consumers.

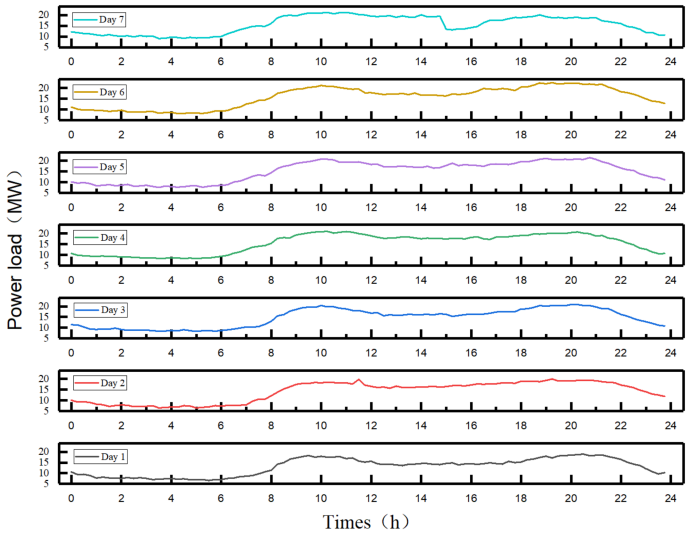


Figure 7. Load curve of DG substation in the first week of January.

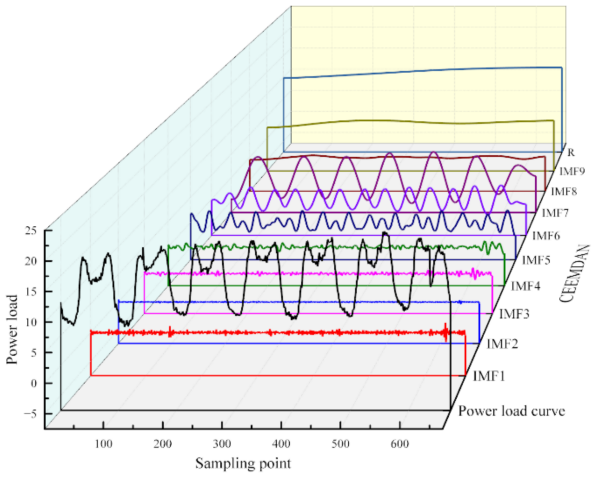


Figure 8. Modal component curve.

5. Results and Discussion

5.1. Analysis of the Effect of RAdam Optimizer and MVO Algorithm Application

From previous research results, the RAdam optimizer can greatly improve the training speed and computational accuracy of deep network models. Based on the above theory, we have performed a validation calculation and analysis. We used the load data shown in Figure 7, selected the GRU network models based on the RAdam and Adam optimizers, respectively, to train the GRU network and to calculate the load prediction for day 7. The variation curves of the loss function of the network training using the two optimizers are shown in Figure 9, and the prediction results are shown in Figure 10.

The comparison of the computational results of the two models shows that the GRU network based on the RAdam optimizer has higher computational accuracy in the training process with the same external parameters. In particular, the maximum computational error of the RAdam-GRU model is also less than half of the computational result of the Adam-GRU model when the number of computations lies in the interval of 10,000 to 20,000 (Figure 9). In addition, in the process of predicting the load trough, the RAdam-GRU model is more flexible enough to respond to the actual load changes, while the Adam-GRU model has poorer prediction results (Figure 10). All the above computational results demonstrate the excellent performance of the RAdam optimizer and show the feasibility of choosing the RAdam optimizer for updating the parameters of the weight matrix of the GRU network in the paper.

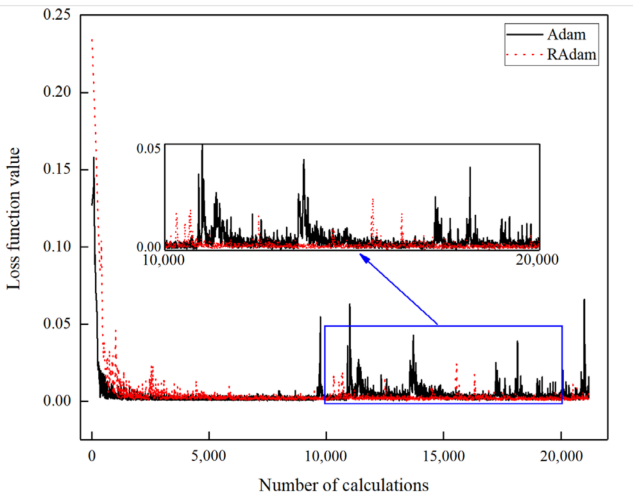


Figure 9. The variation curves of the loss function.

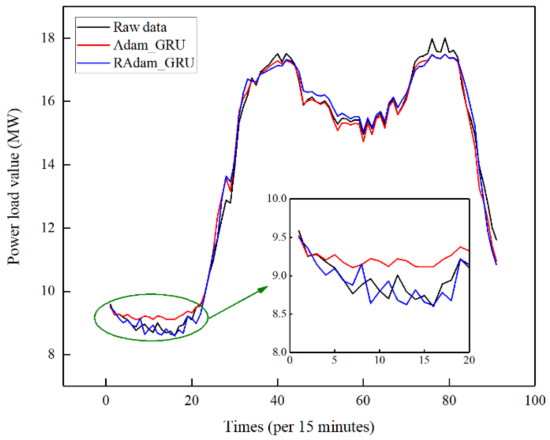


Figure 10. Comparison curves of the two models.

By using the MVO optimization algorithm, it helps us to find the optimal number of nodes in the hidden layer of the neural network. In this subsection, we set the number of nodes in the hidden layer in the range of 10~300; based on the MVO algorithm, the number of populations *NP* was 50 and the maximum number of iterations was 1000, and finally the optimal number of nodes was calculated (Figure 11a). The optimal number of hidden layer nodes per 200 iterations is brought into the GRU model for network training, and it can be seen from the training results that the GRU model training metric (RMSE) becomes more stable as the number of MVO iterations increases. In Figure 11b, the maximum and minimum values of RMSE curves with the number of nodes in the hidden layer of 135 were used as the bounds, and it can be seen that the other five curves crossed the limits in the interval from 100 to 200. It indicates that the optimization algorithm is used for the number of nodes of the hidden layer in the GRU network to find the optimal number of nodes, which is beneficial for the improvement of the computational performance of the GRU model.

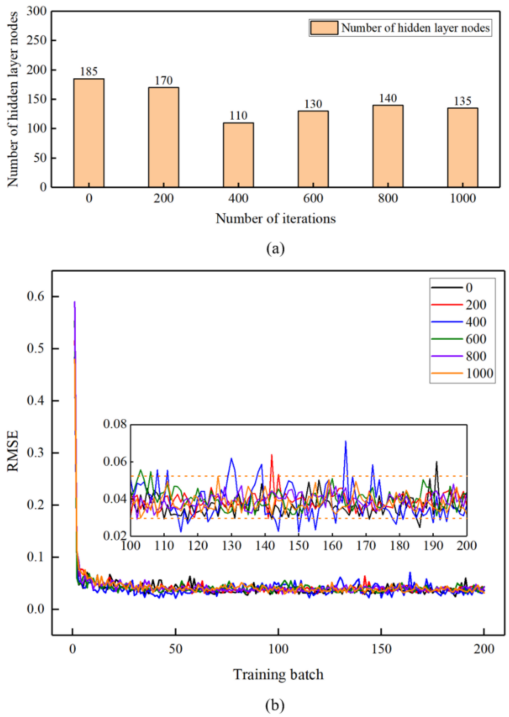


Figure 11. (**a**) Optimal number of hidden layer nodes; (**b**) Network training process.

5.2. Selection of Leading Influencing Factors for Load Forecasting

The GRA was performed to calculate the degree of correlation between the historical load data of the predicted site in January 2021 and the selected influencing factors, and the thermal diagram of the degree of correlation was drawn (Figure 12). The degree of correlation between the selected sequence and other sequences could be obtained by drawing the thermal diagram of the correlation between each influencing factor. The first column in the figure shows the degree of correlation between the load value and the selected influencing factors. They had the highest correlation with themselves. The load characteristic data is presented next. The degree of correlation with temperature was the lowest, only 0.7813, indicating that the substation load change was less affected by temperature.

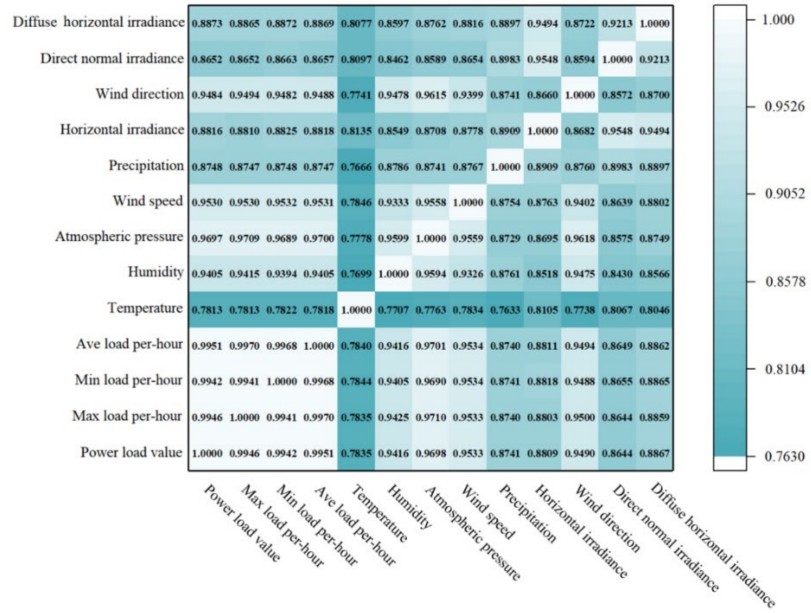


Figure 12. Heat map of correlation analysis of influencing factors of load forecasting.

The CEEMDAN algorithm was used to decompose the load data of January 2021. The level of influence of each factor on the load variation component could be obtained by calculating the degree of correlation between each modal function and the influencing factors. The calculation results of the degree of correlation and the influencing factors are presented in Table 2.

Table 2. Calculation of influencing factors and the degree of correlation.

	IMF 1	IMF 2	IMF 3	R
Max load per-hour	0.8563	0.8555	0.8539	0.9746
Min load per-hour	0.8556	0.8547	0.8532	0.9619
Ave load per-hour	0.8559	0.8551	0.8536	0.9634
Temperature	0.8483	0.8470	0.8457	0.8603
Humidity	0.8525	0.8535	0.8508	0.9480
Atmospheric pressure	0.8565	0.8569	0.8548	0.9904
Wind speed	0.8425	0.8435	0.8409	0.9452
Precipitation	0.8575	0.8587	0.8563	0.8487
Horizontal irradiance	0.8238	0.8209	0.8224	0.8435
Wind direction	0.8439	0.8443	0.8402	0.9509
Direct normal irradiance	0.8218	0.8203	0.8203	0.8300
Diffuse horizontal irradiance	0.8248	0.8223	0.8235	0.8495

The degree of correlation varies between different influencing factors and the decomposed modal functions (Table 2). The degree of correlation between air temperature and IMF1–3 was 0.8483, 0.8470, and 0.8457, respectively, indicating that the high-frequency variation in the load curve was less affected by temperature. However, the degree of correlation increased with the increase in the modal function frequency. At the same time, the degree of residual correlation with the trend of the response load curve was about 0.8603, indicating that the load change had a significant correlation with the local long-term temperature change. The load variation of a place was affected by both the local long-term temperature and the load fluctuation caused by short-term temperature mutation, which was consistent with the actual situation and indicated the effectiveness of the calculation of the degree of correlation.

The calculation results of the degree of correlation of IMF influencing factors are presented in Table 2. This study selected the main influencing factors of load prediction with a degree of correlation greater than 0.85 for each modal component to simplify the subsequent load prediction calculation. The selection of some of the main influencing factors of IMF is shown in Figure 13.

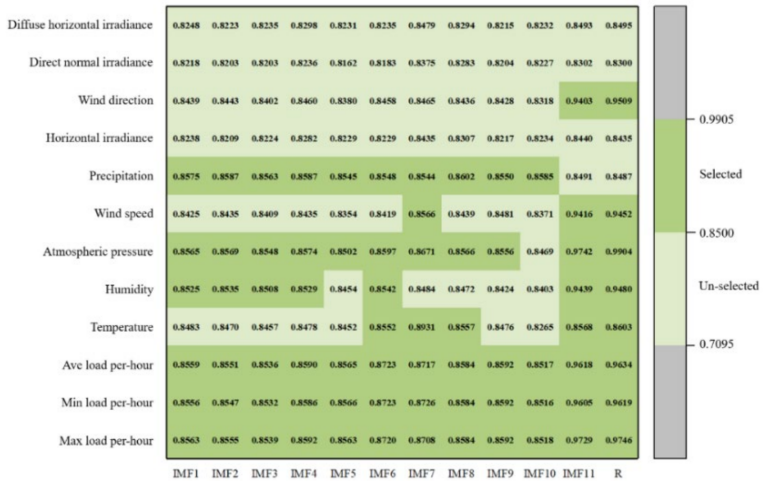


Figure 13. The selection results of the main influencing factors of each IMF.

5.3. Load Prediction and Result Analysis Based on the CEEMDAN-MVO-GRU Algorithm

This study used a three-layer GRU prediction model and the RAdam algorithm as neural network optimizers. In the GRU model, the aforementioned MVO algorithm was used to select the optimal super parameters of the modal function prediction model for different modal functions and the corresponding influencing factors.

In the RAdam-GRU model, the initialized learning rate was set to 0.01 considering that only the first operation was involved using the RAdam optimizer; furthermore, the minimum training batch size for the GRU model was set to 20 and the maximum training batch size was set to 1000. The MVO algorithm was used to select the number of hidden layer nodes when different modal functions were predicted. In the MVO algorithm, the population size *NP* was set to 100, and the maximum iteration number was set to 500. The convergence curves of the adaptive values for each IMF to predict the hyperparameter preference are shown in Figure 14, and finally, we obtain the complete architecture of the algorithm proposed in the paper, and the complete architecture is shown in Table 3.

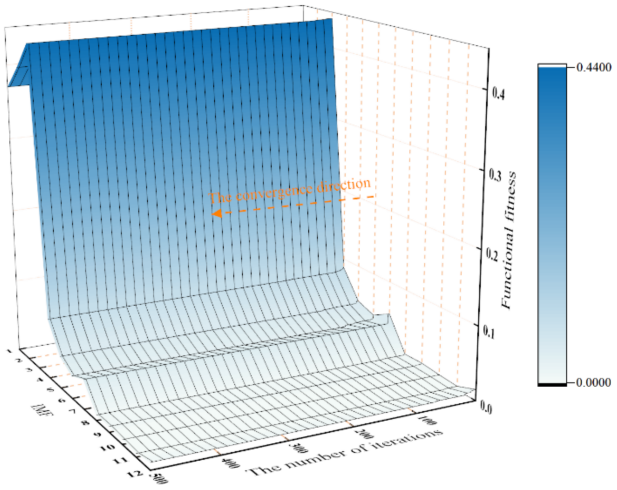


Figure 14. Fitness value convergence curve for each IMF.

Table 3. Complete architecture of the proposed model.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10	IMF11	R
Number of training data	2880	2880	2880	2880	2880	2880	2880	2880	2880	2880	2880	2880
Number of data to predict	96	96	96	96	96	96	96	96	96	96	96	96
Number of inputs	6	6	6	6	5	6	7	7	5	4	8	8
Number of nodes in hidden layer	50	70	40	100	90	100	80	100	70	80	90	70
Number of outputs	1	1	1	1	1	1	1	1	1	1	1	1

After the number of nodes in the hidden layer of each modal function was obtained, the parameters were put into the GRU prediction model to predict each modal function. The prediction results are shown in Figure 15.

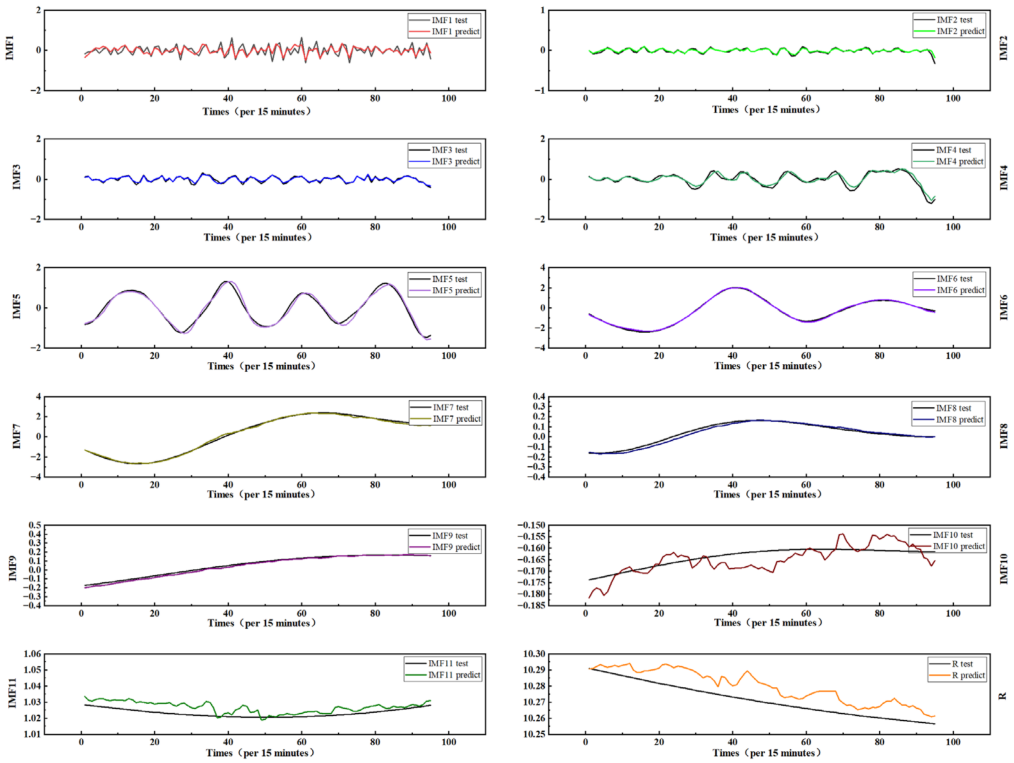
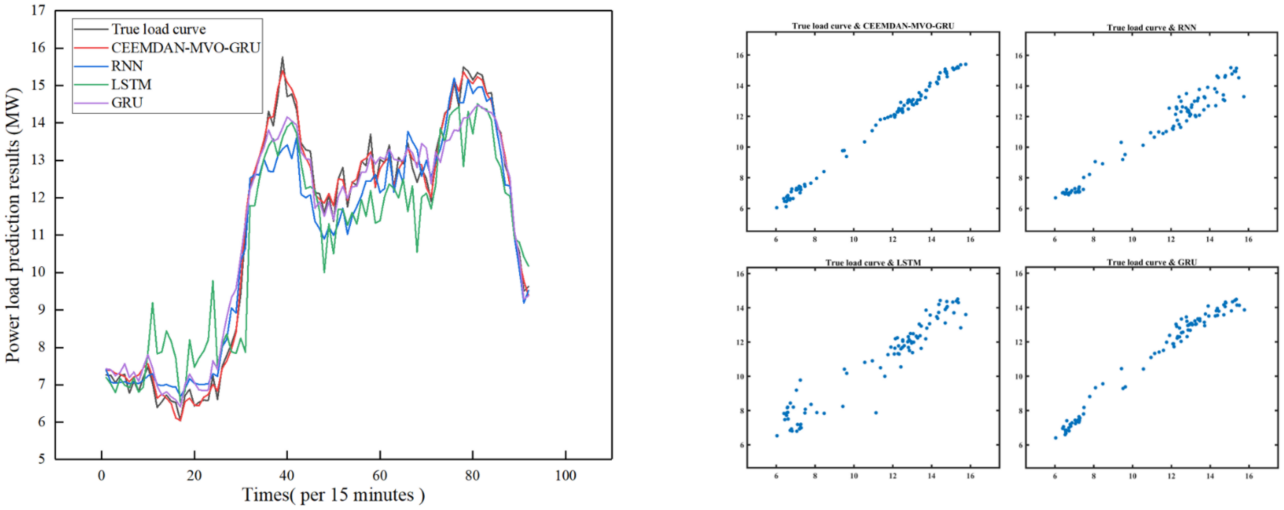


Figure 15. Prediction results of IMF.

A better fitting effect was achieved by using the prediction method proposed in this study to fit each modal function (Figure 15). However, when the high-frequency modal component was fitted, the fitting result could not clearly reflect the high-frequency variation trend of the actual curve. At the same time, a small abnormal fluctuation was observed in the fitting results of low-frequency curves. Based on the algorithm proposed in this study, the final load prediction results of the DG Substation could be obtained by summing up the prediction results of the aforementioned modal components. At the same time, this study also used the RNN, LSTM, and GRU algorithms for load prediction to verify the prediction accuracy of the proposed algorithm. The load prediction results of various algorithms are shown in Figure 16.



(**a**) (**b**)

Figure 16. (**a**) Load Forecast Results; (**b**) Linear regression scatter plot.

The comparison algorithms used in this study exhibited poor fitting accuracy for the load curve. Among these comparison algorithms, the GRU algorithm had better fitting accuracy than the RNN and LSTM, but it could not accurately change the trend of the load curve and exhibited a large fitting error in the load peak-to-peak output (Figure 16a). In addition, it can be further seen from the linear regression scatter plots that the scatter plots of the prediction results using the method proposed in the paper for load prediction tend to be more linear; the scatter plots of the prediction results of the other three control algorithms all have different degrees of dispersion problems (Figure 16b). The original load curve was decomposed using the CEEMDAN algorithm and then the load prediction was performed, which exhibited a higher fitting consistency. Table 4 shows the calculation results of the evaluation indexes of the calculation accuracy of the four load prediction models.

Table 4. Evaluation index of calculation results.

Algorithm	Evaluation Index				
	RMSE	R	MAE	MAPE (%)	NSE
Proposed model	0.2109	0.9977	0.1734	1.6883	0.9978
RNN	0.7020	0.9813	0.5448	4.8801	0.4563
LSTM	1.0643	0.9593	0.8599	8.3985	−0.1474
GRU	0.5612	0.9887	0.4237	3.9586	0.9899

Among the four algorithms, the LSTM algorithm has the largest Root Mean Square Error of 1.0643 and the prediction accuracy of the model was only 91.60% (Table 4). The prediction accuracy of the GRU algorithm and RNN algorithm was significantly improved

compared with the LSTM algorithm. The RMSE of the GRU algorithm was 0.5612, which was 47.27% lower than that of the LSTM algorithm. At the same time, the prediction accuracy of the GRU algorithm was improved to 96.04%. The CEEMDAN algorithm was used to carry out modal segmentation and predict the high-, medium-, and low-frequency components of the load curve, which could better perceive the potential change characteristics of the load curve. The RMSE of fitting could be reduced from 1.0643 of the LSTM algorithm to 0.2109 using the algorithm proposed in this study, which was a significant decrease of 80.18%. The similarity of fitting results also increased from 0.9887 to 0.9977, and the prediction accuracy of the model was further improved to 98.31%. It indicated that the proposed method could effectively mine the potential information in load forecasting and improve the accuracy of load forecasting.

6. Conclusions

Considering that the short-term load fluctuation is strong and the load prediction accuracy is not high due to many factors, this study established a short-term load prediction model based on the CEEMDAN, GRA, MVO, and GRU algorithms.

The CEEMDAN algorithm was used to decompose the original load curve into modal functions of different frequencies, which could extract the hidden information related to load fluctuation effectively. The key influencing factors of different frequency modal functions could be selected to effectively reduce the dimension of the input data of the prediction model based on the GRA. Finally, the MVO algorithm was used to optimize the hyperparameters of the GRU model, which could further improve the prediction accuracy of the model.

The results showed that, compared with the control prediction model, the proposed prediction model could accurately reflect the load change situation, and had a higher fitting consistency with the original load curve. The RMSE of fitting decreased from 1.0643 to 0.2109 (80.18%) in the control group. The similarity of fitting results also increased from 0.9887 to 0.9977. Hence, this study confirms that the proposed algorithm has higher prediction accuracy. And the operation optimization of grid power, energy storage, and load can be performed based on this prediction result in the subsequent study.

Author Contributions: Conceptualization, T.J.; methodology, L.Y.; visualization, Q.H.; writing—review and editing, G.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financially supported by the Youth Program of the National Natural Foundation of China (No. 51507134).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Son, J.; Cha, J.; Kim, H.; Wi, Y.-M. Day-Ahead Short-Term Load Forecasting for Holidays Based on Modification of Similar Days' Load Profiles. *IEEE Access* **2022**, *10*, 17864–17880. [CrossRef]
2. Fan, G.-F.; Liu, Y.-R.; Wei, H.-Z.; Yu, M.; Li, Y.-H. The new hybrid approaches to forecasting short-term electricity load. *Electr. Power Syst. Res.* **2022**, *213*, 108759. [CrossRef]
3. Sina, A.; Kaur, D. An Accurate Hybrid Approach for Electric Short-Term Load Forecasting. *IETE J. Res.* **2021**, 1–16. [CrossRef]
4. Cao, Z.J.; Wan, C.; Zhang, Z.J.; Li, F.R.; Song, Y.H. Hybrid Ensemble Deep Learning for Deterministic and Probabilistic Low-Voltage Load Forecasting. *IEEE Trans. Power Syst.* **2020**, *35*, 1881–1897. [CrossRef]
5. Kozak, D.; Holladay, S.; Fasshauer, G.E. Intraday Load Forecasts with Uncertainty. *Energies* **2019**, *12*, 1833. [CrossRef]
6. Butt, F.M.; Hussain, L.; Mahmood, A.; Lone, K.J. Artificial Intelligence based accurately load forecasting system to forecast short and medium-term load demands. *Math. Biosci. Eng.* **2020**, *18*, 400–425. [CrossRef]
7. Guo, Y.X.; Li, Y.; Qiao, X.B.; Zhang, Z.Y.; Zhou, W.F.; Mei, Y.J.; Lin, J.J.; Zhou, Y.C.; Nakanishi, Y. BiLSTM Multitask Learning-Based Combined Load Forecasting Considering the Loads Coupling Relationship for Multienergy System. *IEEE Trans. Smart Grid* **2022**, *13*, 3481–3492. [CrossRef]

8. Cai, G.W.; Wang, W.J.; Lu, J.H. A Novel Hybrid Short Term Load Forecasting Model Considering the Error of Numerical Weather Prediction. *Energies* **2016**, *9*, 994. [CrossRef]
9. Paparoditis, E.; Sapatinas, T. Short-Term Load Forecasting: The Similar Shape Functional Time-Series Predictor. *IEEE Trans. Power Syst.* **2013**, *28*, 3818–3825. [CrossRef]
10. Jin, Y.; Guo, H.G.; Wang, J.Z.; Song, A.Y. A Hybrid System Based on LSTM for Short-Term Power Load Forecasting. *Energies* **2020**, *13*, 6241. [CrossRef]
11. Chen, T.; Huang, W.; Wu, R.J.; Ouyang, H.B. Short Term Load Forecasting Based on SBiGRU and CEEMDAN-SBiGRU Combined Model. *IEEE Access* **2021**, *9*, 89311–89324. [CrossRef]
12. Rafi, S.H.; Deeba, S.R.; Hossain, E. A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network. *IEEE Access* **2021**, *9*, 32436–32448. [CrossRef]
13. Cai, C.C.; Li, Y.J.; Su, Z.H.; Zhu, T.Q.; He, Y.Y. Short-Term Electrical Load Forecasting Based on VMD and GRU-TCN Hybrid Network. *Appl. Sci.* **2022**, *12*, 6647. [CrossRef]
14. Shang, C.; Gao, J.W.; Liu, H.B.; Liu, F.Z. Short-Term Load Forecasting Based on PSO-KFCM Daily Load Curve Clustering and CNN-LSTM Model. *IEEE Access* **2021**, *9*, 50344–50357. [CrossRef]
15. Javed, U.; Ijaz, K.; Jawad, M.; Khosa, I.; Ansari, E.A.; Zaidi, K.S.; Rafiq, M.N.; Shabbir, N. A novel short receptive field based dilated causal convolutional network integrated with Bidirectional LSTM for short-term load forecasting. *Expert Syst. Appl.* **2022**, *205*, 117689. [CrossRef]
16. Zhang, X.Y.; Kuenzel, S.; Colombo, N.; Watkins, C. Hybrid Short-term Load Forecasting Method Based on Empirical Wavelet Transform and Bidirectional Long Short-term Memory Neural Networks. *J. Mod. Power Syst. Clean Energy* **2022**, *10*, 1216–1228. [CrossRef]
17. Yu, B.B.; Li, J.J.; Liu, C.; Sun, B. A novel short-term electrical load forecasting framework with intelligent feature engineering. *Appl. Energy* **2022**, *327*, 120089. [CrossRef]
18. Yang, Y.; Wang, Z.J.; Gao, Y.C.; Wu, J.R.; Zhao, S.R.; Ding, Z. An effective dimensionality reduction approach for short-term load forecasting. *Electr. Power Syst. Res.* **2022**, *210*, 108067. [CrossRef]
19. Yang, Y.; Zhou, H.; Wu, J.R.; Liu, C.J.; Wang, Y.G. A novel decompose-cluster-feedback algorithm for load forecasting with hierarchical structure. *Int. J. Elec. Power* **2022**, *142*, 108249. [CrossRef]
20. Zhou, M.R.; Hu, T.Y.; Bian, K.; Lai, W.H.; Hu, F.; Hamrani, O.; Zhu, Z.W. Short-Term Electric Load Forecasting Based on Variational Mode Decomposition and Grey Wolf Optimization. *Energies* **2021**, *14*, 4089. [CrossRef]
21. Yang, Y.; Wang, Z.; Zhao, S.; Wu, J. An integrated federated learning algorithm for short-term load forecasting. *Electr. Power Syst. Res.* **2022**, *214*, 108830. [CrossRef]
22. Semero, Y.K.; Zhang, J.H.; Zheng, D.H. EMD-PSO-ANFIS-based hybrid approach for short-term load forecasting in microgrids. *IET Gener. Transm. Distrib.* **2020**, *14*, 470–475. [CrossRef]
23. Liu, T.X.; Jin, Y.; Gao, Y.Y. A New Hybrid Approach for Short-Term Electric Load Forecasting Applying Support Vector Machine with Ensemble Empirical Mode Decomposition and Whale Optimization. *Energies* **2019**, *12*, 1520. [CrossRef]
24. Zhou, F.T.; Huang, Z.H.; Zhang, C.H. Carbon price forecasting based on CEEMDAN and LSTM. *Appl. Energy* **2022**, *311*, 118601. [CrossRef]
25. Huang, N.T.; Qi, J.J.; Li, F.Q.; Yang, D.F.; Cai, G.W.; Huang, G.L.; Zheng, J.; Li, Z.X. Short-Circuit Fault Detection and Classification Using Empirical Wavelet Transform and Local Energy for Electric Transmission Line. *Sensors* **2017**, *17*, 2133. [CrossRef]
26. Hu, L.; Wang, L.G.; Chen, Y.L.; Hu, N.I.; Jiang, Y. Bearing Fault Diagnosis Using Piecewise Aggregate Approximation and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise. *Sensors* **2022**, *22*, 6599. [CrossRef]
27. Halgamuge, M.N.; Daminda, E.; Nirmalathas, A. Best optimizer selection for predicting bushfire occurrences using deep learning. *Nat. Hazards* **2020**, *103*, 845–860. [CrossRef]
28. Liu, L.; Jiang, H.; He, P.; Chen, W.; Liu, X.; Gao, J.; Learning, J.H.J. On the Variance of the Adaptive Learning Rate and Beyond. *arXiv Preprint* **2019**, arXiv:1908.03265.
29. Melinte, D.O.; Vladareanu, L. Facial Expressions Recognition for Human-Robot Interaction Using Deep Convolutional Neural Networks with Rectified Adam Optimizer. *Sensors* **2020**, *20*, 2393. [CrossRef]
30. Colominas, M.A.; Schlotthauer, G.; Torres, M.E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomed. Signal Process. Control.* **2014**, *14*, 19–29. [CrossRef]
31. Humeau-Heurtier, A.; Abraham, P.; Mahe, G. Analysis of Laser Speckle Contrast Images Variability Using a Novel Empirical Mode Decomposition: Comparison of Results with Laser Doppler Flowmetry Signals Variability. *IEEE Trans. Med Imaging* **2015**, *34*, 618–627. [CrossRef]
32. Yang, S.M.; Yuan, A.J.; Yu, Z.Q. A novel model based on CEEMDAN, IWOA, and LSTM for ultra-short-term wind power forecasting. *Environ. Sci. Pollut. Res.* **2022**, 1–17. [CrossRef]
33. Dong, W.J.; Liu, S.F.; Fang, Z.G. On modeling mechanisms and applicable ranges of grey incidence analysis models. *Grey Syst. Theory Appl.* **2018**, *8*, 448–461. [CrossRef]
34. Zhang, B.; Zhao, X.H.; Dou, Z.H.; Liu, L.X. A New Medium and Long-Term Power Load Forecasting Method Considering Policy Factors. *IEEE Access* **2021**, *9*, 160021–160034. [CrossRef]
35. Liu, S.F.; Zhang, H.Y.; Yang, Y.J. Explanation of terms of grey incidence analysis models. *Grey Syst. Theory Appl.* **2017**, *7*, 136–142. [CrossRef]

36. Mi, J.W.; Fan, L.B.; Duan, X.C.; Qiu, Y.Y. Short-Term Power Load Forecasting Method Based on Improved Exponential Smoothing Grey Model. *Math. Probl. Eng.* **2018**, *2018*, 1–11. [CrossRef]
37. Mirjalili, S.; Mirjalili, S.M.; Hatamlou, A. Multi-Verse Optimizer: A nature-inspired algorithm for global optimization. *Neural Comput. Appl.* **2016**, *27*, 495–513. [CrossRef]
38. Mishra, S.; Nayak, P.C.; Prusty, R.C.; Panda, S. Modified multiverse optimizer technique-based two degree of freedom fuzzy PID controller for frequency control of microgrid systems with hydrogen aqua electrolyzer fuel cell unit. *Neural Comput. Appl.* **2022**, *34*, 1–17. [CrossRef]
39. Sundaram, A. Multiobjective multi verse optimization algorithm to solve dynamic economic emission dispatch problem with transmission loss prediction by an artificial neural network. *Appl. Soft Comput.* **2022**, *124*, 109021. [CrossRef]
40. Sulaiman, M.; Ahmad, S.; Iqbal, J.; Khan, A.; Khan, R. Optimal Operation of the Hybrid Electricity Generation System Using Multiverse Optimization Algorithm. *Comput. Intel. Neurosci.* **2019**, *2019*, 1–12. [CrossRef]
41. Mahjoub, S.; Chrifi-Alaoui, L.; Marhic, B.; Delahoche, L. Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks. *Sensors* **2022**, *22*, 4062. [CrossRef]
42. Jia, P.T.; Liu, H.D.; Wang, S.J.; Wang, P. Research on a Mine Gas Concentration Forecasting Model Based on a GRU Network. *IEEE Access* **2020**, *8*, 38023–38031. [CrossRef]
43. Ding, G.R.; Wang, W.B.; Zhu, T. Remaining Useful Life Prediction for Lithium-Ion Batteries Based on CS-VMD and GRU. *IEEE Access* **2022**, *10*, 89402–89413. [CrossRef]
44. Zhao, Y.; Guo, N.; Chen, W.; Zhang, H.; Guo, B.; Shen, J.; Tian, Z. Multi-step ahead forecasting for electric power load using an ensemble model. *Expert Syst. Appl.* **2022**, *211*, 118649. [CrossRef]
45. Kim, N.; Park, H.; Lee, J.; Choi, J.K. Short-Term Electrical Load Forecasting With Multidimensional Feature Extraction. *IEEE Trans. Smart Grid* **2022**, *13*, 2999–3013. [CrossRef]
46. Hu, H.W.; Xia, X.; Luo, Y.L.; Zhang, C.; Nazir, M.S.; Peng, T. Development and application of an evolutionary deep learning framework of LSTM based on improved grasshopper optimization algorithm for short-term load forecasting. *J. Build. Eng.* **2022**, *57*, 104975. [CrossRef]