energies

Article

Ultra Short-Term Power Load Forecasting Based on Similar Day Clustering and Ensemble Empirical Mode Decomposition

Wenhui Zeng [1], Jiarui Li [2], Changchun Sun [2], Lin Cao [1], Xiaoping Tang [2,*], Shaolong Shu [1] and Junsheng Zheng [3]

1 School of Electronics and Information Engineering, Tongji University, Shanghai 201804, China
2 Huadian Electric Power Research Institute Co., Ltd., Hangzhou 310030, China
3 School of Automotive Studies, Tongji University, Shanghai 201804, China
* Correspondence: xiaoping-tang@chder.com

Abstract: With the increasing demand of the power industry for load forecasting, improving the accuracy of power load forecasting has become increasingly important. In this paper, we propose an ultra short-term power load forecasting method based on similar day clustering and EEMD (Ensemble Empirical Mode Decomposition). In detail, the K-means clustering algorithm was utilized to divide the historical data into different clusters. Through EEMD, the load data of each cluster were decomposed into several sub-sequences with different time scales. The LSTNet (Long- and Short-term Time-series Network) was adopted as the load forecasting model for these sub-sequences. The forecast results for different sub-sequences were combined as the expected result. The proposed method predicts the load in the next 4 h with an interval of 15 min. The experimental results show that the proposed method obtains higher prediction accuracy than other comparable forecasting models.

Keywords: cluster analysis; mode decomposition; LSTNet; ultra short-term load forecasting; non-stationary time series

check for updates

Citation: Zeng, W.; Li, J.; Sun, C.; Cao, L.; Tang, X.; Shu, S.; Zheng, J. Ultra Short-Term Power Load Forecasting Based on Similar Day Clustering and Ensemble Empirical Mode Decomposition. *Energies* **2023**, *16*, 1989. https://doi.org/10.3390/en16041989

Academic Editors: Ioannis Panapakidis, Aggelos S. Bouhouras and Minas Alexiadis

Received: 23 December 2022
Revised: 15 January 2023
Accepted: 1 February 2023
Published: 17 February 2023

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

1. Introduction

In recent years, with the rapid economic growth, the development of the power industry has been growing. Power load forecasting plays a key role in the operation of the power industry, and is the basis of economic dispatching, energy storage management, future energy contracts and plant maintenance plans [1]. The accuracy of power load forecasting has an extremely important impact on the operation of power systems, because the operation and control of power systems are very sensitive to prediction errors [2]. Overestimation of load will lead to excessive power production which results in energy waste. On the other hand, underestimation of load will lead to poor power production which does not satisfy the needs of living and industry.

Nowadays, there are many studies on power load forecasting. According to the prediction time horizon, it is mainly divided into ultra short-term power load forecasting (USTLF), short-term power load forecasting (STLF), medium-term power load forecasting (MTLF) and long-term power load forecasting (LTLF) [3,4]. USTLF refers to the load forecasting for one day or a shorter time; it is used for real-time dispatching and daytime dispatching of electric power [5]. STLF forecasts from one hour to one week for the daily operation of power systems, such as energy trading, load flow analysis and power system security research [6]. MTLF, used for fuel supply scheduling, includes forecasting from several weeks to one year, and LTLF is usually more than one year's forecasting for long-term power system planning [7,8].

Here, we focus on USTLF. Until now, there have been many studies on it. Traditional power load forecasting methods directly establish the correlation between the predicted values and the historical values, mainly including the regression analysis method [9], the time series method [10], and the exponential smoothing method [11]. The regression analysis method predicts the future load value by establishing a mathematical model reflecting the causal relationship. The time series method establishes a mathematical model by curve fitting and parameter estimation of historical load data. The exponential smoothing method does this through the exponential weighted combination. These methods can express the significant relationship between the independent variable and the dependent variable. However, they are limited in dealing with complex nonlinear systems and the performance of prediction is poor on power load, which is a non-stationary time series.

With the development of artificial intelligence, more modern load forecasting methods have been proposed, which mainly include SVM (Support Vector Machine) [12], FL (Fuzzy Logic) [13,14] and neural network methods [15,16]. Niu D et al. [17] proposed an SVM daily load forecasting model based on case-based reasoning. They showed that SVM has the advantages of high precision and high speed in power load forecasting. SVM has strong generalization ability, but lacks the ability to deal with uncertainty [18]. FL can deal with the uncertainty of load change well, but it is seriously dependent on expert experience.

In recent years, the neural network method has been widely used in load forecasting, such as CNN models (Convolutional Neural Network) [19], RNN models (Recurrent Neural Network) [20], LSTM model(Long Short-Term Memory) [21], etc. Dong et al. [22] used a joint CNN and K-means algorithm to predict hourly power load which can largely improve the prediction performance. Traditional neural networks have a strong nonlinear mapping ability, but are not good at processing time series data. Since power load data are a time series, power load forecasting should be considered a time series modeling problem [23]. RNN is an improved ANN that can solve the problem of processing time series. However, RNN is prone to gradient explosion and gradient vanishing. In order to solve this problem, the LSTM model [24] was proposed, and has been used for power load forecasting. It can deal with the time series property and nonlinearity of load data simultaneously and better learn the long-term dependence of time series [25] in virtue of its special memory ability and gate structure. In an LSTM, information is not lost with time. Hence, LSTM has higher accuracy when predicting the future load demand [26]. Zhao et al. [27] proposed an LSTM forecasting method, which proved that LSTM is better than Elman, BP and LSSVM. Zhang et al. [28] proposed a regional level ultra short-term load forecasting method based on deep LSTM to forecast the load in the next hour. However, as the sequence length increases, the gradient disappears and the prediction effect of LSTM decreases.

The above single model research has been relatively mature and the prediction effect is good. However, the single model also has some shortcomings, which hinder the improvement of prediction accuracy. Composite models are widely used and can effectively combine the advantages of multiple models, such as the CNN-LSTM model [29], the GRU-CNN model [30], etc. Some researchers use CNN to extract the power characteristic information of users, which shows unique advantages in extracting the nonlinearity of power load data [31]. S. H. Rafi et al. [32] proposed a short-term power load forecasting method based on the CNN-LSTM model. They showed that the proposed method had higher precision and accuracy compared with LSTM, RBFN and XGboost models. Wu, K. et al. [33] proposed an attention-based CNN combined with LSTM and BiLSTM (Bidirectional Long Short-Term Memory) model for short-term load forecasting, which had a better performance.

However, the above methods do not consider the long-term periodic characteristics of load and the long-term correlation between many other variables, such as weather conditions, temperature and so on [34,35]. Hence, LSTNet (Long- and Short-term Time-series Network) [36] was proposed, which can better learn the long-term correlation between multi-variables and extract the highly nonlinear long-term and short-term features and linear features in the data. Until now, LSTNet is the state-of-the-art model and achieves an excellent performance in multivariate time series. Liu, R. et al. [37] proposed a load forecasting model based on LSTNet, and proved that the accuracy and robustness are both better than those of other models. The comparison of different power load forecasting models is shown in Table 1.

Table 1. Comparison of different power load forecasting models.

Model	Advantage	Disadvantage
SVM	high precision, high speed, strong generalization ability	lacks the ability to deal with uncertainty
LSTM	deal with the time series property and nonlinearity of load data simultaneously	as the sequence length increases, the gradient disappears and the prediction effect decreases.
CNN-LSTM, CNN-BiLSTM	extract the power characteristic information	do not consider the long-term periodic characteristics of load and the long-term correlation between many other variables
LSTNet	better learn the long-term correlation between multi-variables and extract the highly nonlinear long-term and short-term features and linear features in the data	affected by nonstationarity of power load

Although LSTNet achieves a good performance in multivariate time series, the power load is a nonstationary time series [38], which obviously degrades the prediction performance. In order to reduce the effect of nonstationarity, we adopted similar day clustering and EEMD (Ensemble Empirical Mode Decomposition).

The main contributions of this paper are the following: (1) we adopted the K-means clustering algorithm to divide the historical multivariate data into several different clusters based on the similarity of days, so as to improve the similarity and consistency of training data; (2) we used EEMD [39] to further decompose the historical load data sequence into different sub-sequences with different time scales for each cluster, so as to reduce the effect of nonstationarity; (3) LSTNet was used as the load forecasting model for these sub-sequences, which is much more stationary than the original power load time series. We combined these prediction results of these sub-sequences to obtain the final load forecasting value through superposition and reconstruction. Compared with other models, including LSTM, CNN-LSTM, CNN-BiLSTM and LSTNet, the experimental results show that the accuracy of the proposed method rises significantly.

The remainder of this paper is organized as follows. Section 2 introduces the research methods. In Section 3, experiments are undertaken to evaluate the performance of the proposed method. Finally, we conclude the paper in Section 4.

2. Research Methods

It is verified that the power load during work days is very different from that in the holidays. We also know that temperature, humidity, season, light intensity, precipitation, ground wind speed and relative air pressure affect the power load. Hence, we chose the historical data of the power load, holiday, temperature, humidity, season, light intensity, precipitation, ground wind speed and relative air pressure as the input variables.

2.1. Similar Day Clustering

We know the power load changes for different days. Here, we wanted to divide the historical power load data into several clusters so that each cluster of data is more similar. We did this based on the similarity of the eight characteristic factors (holiday, temperature, humidity, season, light intensity, precipitation, ground wind speed and relative air pressure) for each day. In this paper, the K-means clustering algorithm was used to cluster the historical data.

We chose the data of the eight characteristic factors for each day as a sample. The samples of historical data were denoted as

$$X = \{x_1, x_2, ..., x_n\}.$$

Note that each sample x_i is a matrix including eight characteristic vectors which are holiday, temperature, humidity, season, light intensity, precipitation, ground wind speed and relative air pressure.

Initially, we needed to determine the optimal K for the K-means clustering algorithm. In this paper, the CH (calinski_harabasz) index was used to evaluate the clustering score. The higher the score, the better the clustering result. The CH index is the ratio of the separation degree and the compactness of the data set X. The separation degree of the data set X is measured by the sum of the squares of the distances between the center points of various clusters and the center points of the data set X. The compactness is measured by the sum of the squares of the distances between the points in each cluster and the center point of the same cluster.

We determined the input and output of the algorithm. These samples and cluster number K were inputs. The output was the expected K clusters.

We divided the historical samples in X into K clusters through K-means. The process is as follows [40]:

Step 1. Select initial clustering center. Randomly select K samples in X as initial clustering centers. We denote the set of these centers as

$$C = \{c_1, c_2, ..., c_K\};$$

Step 2. Cluster division. Calculate the Euclidean distance between a sample in X and each of K cluster centers according to the following formula. The sample is then put in the cluster of which the center is nearest to the sample.

$$L_{ij} = \sqrt{\sum_{t=1}^{8} (x_{it} - c_{jt})^2}, 1 \le i \le n, 1 \le j \le K, 1 \le t \le 8,$$

where x_{it} is the tth element of sample x_i and c_{jt} is the tth element of center c_j. In this way, we place every sample into one cluster;

Step 3. Update the cluster center. Calculate the mean value of samples in each cluster according to the following formula as the updated cluster center as

$$c_j = \frac{\sum_{x_s^j \in C_j} x_s^j}{|C_j|}, 1 \le s \le |C_j|, 1 \le j \le K,$$

where C_j is the set of samples in the jth cluster;

Step 4. Repeat step 2 and step 3 until all cluster centers do not change any longer. The square error function E calculated according to the following formula will converge to a fixed value (minimum value).

$$E = \sum_{j=1}^{K} \sum_{x_s^j \in C_j} L_{sj}^2,$$

where x_s^j is the sth sample in the jth cluster; C_j is the set of samples in the jth cluster; L_{sj} is the distance between the sth sample in the jth cluster and the jth cluster center; K is the number of clusters.

In this way, we divided all the historical data of the eight characteristic factors into K different clusters. In each cluster, we could arrange the samples in the order of time to form a time series. For the load data with respect to samples in each cluster, we could also arrange them with the order of time to form a time series as shown in Figure 1.

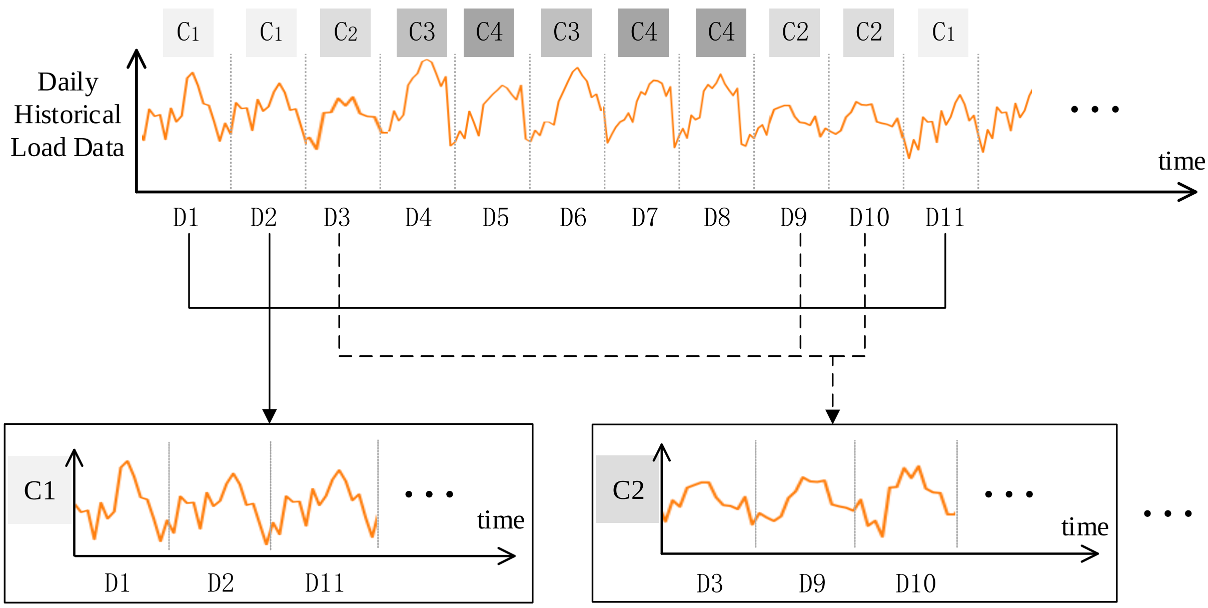

Figure 1. Construct time series of historical load data for each cluster.

Instead of considering the time series consisting of all the load data with respect to samples in X, we investigated each load time series with respect to these samples in each cluster, which is much more regular and easier to deal with.

2.2. EEMD Decomposition

For every load time series of each cluster, it is still non-stationary. Hence, we used EEMD to further decompose it into multiple hierarchically stable IMF (Intrinsic Mode Function) sub-sequences. The specific process of EEMD decomposition is as follows [21].

Step 1. Set the number of original signal processing as N;

Step 2. Denote the load time series signal for the ith cluster as

$$Y_i(t) = [y_{i,1}, y_{i,2}, y_{i,3}, \cdots, y_{i,M_i}].$$

Add a noise signal $\omega(t)$ with standard normal distribution to $Y_i(t)$ to obtain a new signal $Y_i'(t)$ as

$$Y_i'(t) = Y_i(t) + \omega(t);$$

Step 3. Through EMD (Empirical Mode Decomposition), decompose the signal $Y_i'(t)$ to obtain all IMF components $c_1(t), c_2(t), \cdots, c_J(t)$ and the residual component $r(t)$ such that

$$Y_i'(t) = \sum_{j=1}^{J} c_j(t) + r(t).$$

Note that J represents the number of IMF components decomposed after adding white noise.

Step 4. Repeat steps (2) and (3) N times. Note that white noise is added with the same intensity and different sequences each time. Then obtain N sets of IMF components and the resident components as

$$c_1^1(t), c_2^1(t), \cdots, c_j^1(t), \cdots c_J^1(t), r^1(t)$$

$$c_1^2(t), c_2^2(t), \cdots, c_j^2(t), \cdots c_J^2(t), r^2(t)$$

$$\cdots$$

$$c_1^n(t), c_2^n(t), \cdots, c_j^n(t), \cdots c_J^n(t), r^n(t)$$

$$\cdots$$

$$c_1^N(t), c_2^N(t), \cdots, c_j^N(t), \cdots c_J^N(t), r^N(t),$$

where $c_j^n(t)$ is the jth IMF component obtained by the nth time and $r^n(t)$ is the residual component obtained by the nth time.

Step 5. Because the mean value of the white noise spectrum is zero, the final IMF component is the average of the N corresponding IMF components as

$$c_j(t) = \frac{1}{N}\sum_{n=1}^{N} c_j^n(t) \quad j = 1, 2, 3, \cdots, J,$$

and the final resident component is the average of the N residual components as

$$r(t) = \frac{1}{N}\sum_{n=1}^{N} r^n(t);$$

Step 6. Existing results show that some IMF components are weakly related to the original signal and should be eliminated in order to reduce the adverse effect on the prediction results. Here, compare the correlation coefficient between the decomposed IMF component and the original signal with a preset threshold to determine which IMF components should be eliminated. In order to avoid the small amplitude and real signal being eliminated, all IMF components are normalized with the original signal. The normalized correlation coefficient between the jth IMF component and the original signal is calculated as

$$NCC_j = \frac{\sum_{k=1}^{M_i}(y_{i,k} - \bar{y})(c_{j,k} - \bar{c})}{\sqrt{\sum_{k=1}^{M_i}(y_{i,k} - \bar{y})^2 \sum_{k=1}^{M_i}(c_{j,k} - \bar{c})^2}},$$

where

$c_{j,k}$ is the value of the kth point in the time series c_j;

$\bar{y} = \frac{1}{M_i}\sum_{k=1}^{M_i}(y_{i,k})$ is the mean value of all $y_{i,k}$;

$\bar{c} = \frac{1}{M_i}\sum_{k=1}^{M_i}(c_{j,k})$ is the mean value of all $c_{j,k}$;

M is the number of points of the original signal and all IMF components for the ith cluster.

The threshold TH is obtained from the standard deviation of the correlation coefficient as

$$TH = (\frac{1}{J-1}\sum_{j=1}^{J}(NCC_j - \overline{NCC})^2)^{1/2},$$

where $\overline{NCC} = \frac{1}{J}\sum_{j=1}^{J}(NCC_j)$ is the mean value of all NCC_j;

Step 7. If $NCC_j > TH$, the j-th IMF component is kept. Otherwise, eliminate it. In this way, we get the final IMF components as

$$c_{k_1}(t), c_{k_2}(t), \cdots, c_{k_l}(t).$$

Similarly, we can decompose the load time series signals for other clusters. Note that each IMF component is a stationary time series signal.

2.3. LSTNet Network

For each IMF component and residual component, we wanted to establish an LSTNet model for prediction. Note that in each cluster, we have eight characteristic factor time series. They are also the inputs of these LSTNet models.

LSTNet consists of a nonlinear part and a linear part [19]. The nonlinear part consists of a CNN layer, an RNN layer and a recurrent-skip layer. The linear part consists of an autoregressive linear layer. The model makes use of the advantages of CNN and RNN at the same time. CNN can extract the short-term local dependency patterns among the power load data. RNN and the recurrent-skip layer can capture the long-term patterns and simplify the optimization process based on the periodicity of the power load. Finally, the traditional autoregressive linear model was added to make the nonlinear deep learning model more robust to the power load time series. The final LSTNet forecast result was obtained by superimposing the results of the nonlinear part and the linear part. The LSTNet model is shown in Figure 2.

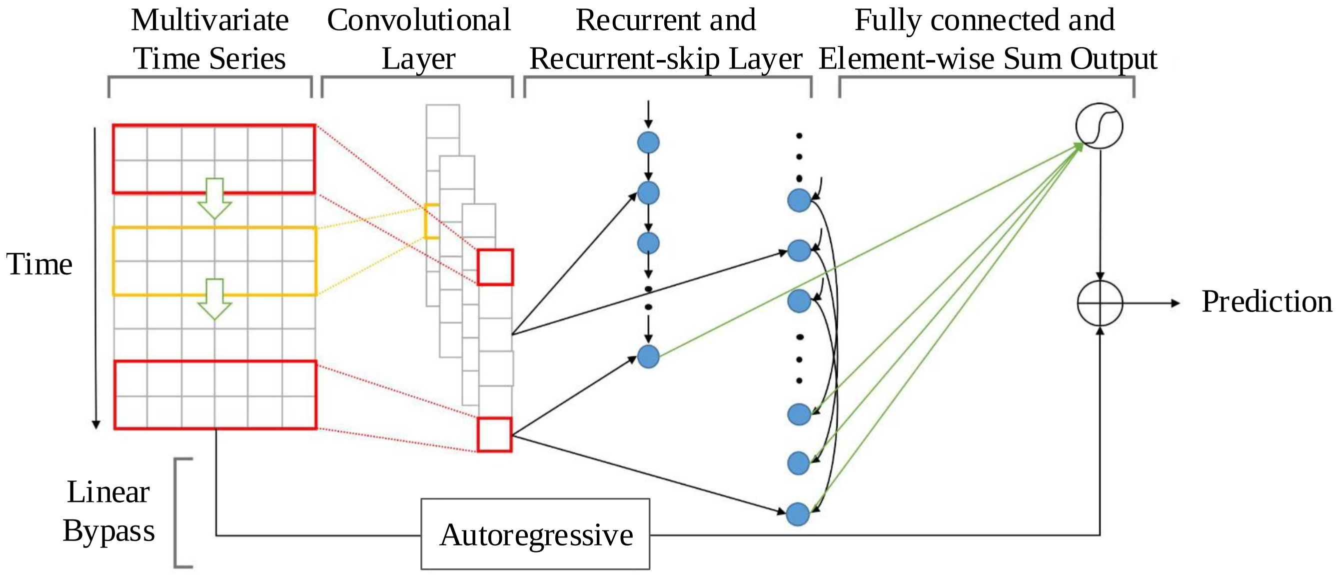

Figure 2. LSTNet model.

2.4. Load Forecasting

In this paper, the mentioned three methods were combined for ultra short-term power load forecasting. The ultra short-term power load forecasting method proposed in this paper takes 15 min as a unit, 16 h (64 data points) of similar day historical data as input, and outputs the load forecasting results in the next 4 h (16 data points).

As in Figure 3, we adopted the K-means clustering algorithm to divide the historical days into K clusters based on the historical data of eight characteristic factors. We then obtained the historical data of the load and the eight characteristic factors for these days in each cluster. For the time series signal formed with the historical data of load in the given cluster C_i, we adopted EEMD to decompose it into J_i IMF components. For each IMF component $c_p(t)$, we constructed an LSTNet model for it by training. The IMF component $c_p(t)$ and the time series signal of eight characteristic factors formed with their historical data in the same cluster C_i were the input data for training. Finally, for each component in each cluster, we obtained a trained LSTNet model.

The load prediction was then implemented as follows. We first determined which cluster the forecast day belonged to. We adopted the prediction data of eight characteristic factors for the forecast day which could be purchased from commercial companies. We calculated the distance between the prediction data of the eight characteristic factors and the historical data in each cluster. It belongs to the cluster with the shortest distance. We then choose the the historical data of the latest 16 h in the same cluster for load prediction. We decomposed the historical load data of the latest 16 h into a series of IMF components. Each component and the historical data of the eight characteristic factors in the latest 16 h were then sent to the corresponding LSTNet model to output the 16 data points for the next 4 h. We then added all the outputs of all the LSTNet models. The sum is the expected load forecasting results for the next 4 h with an interval of 15 min.

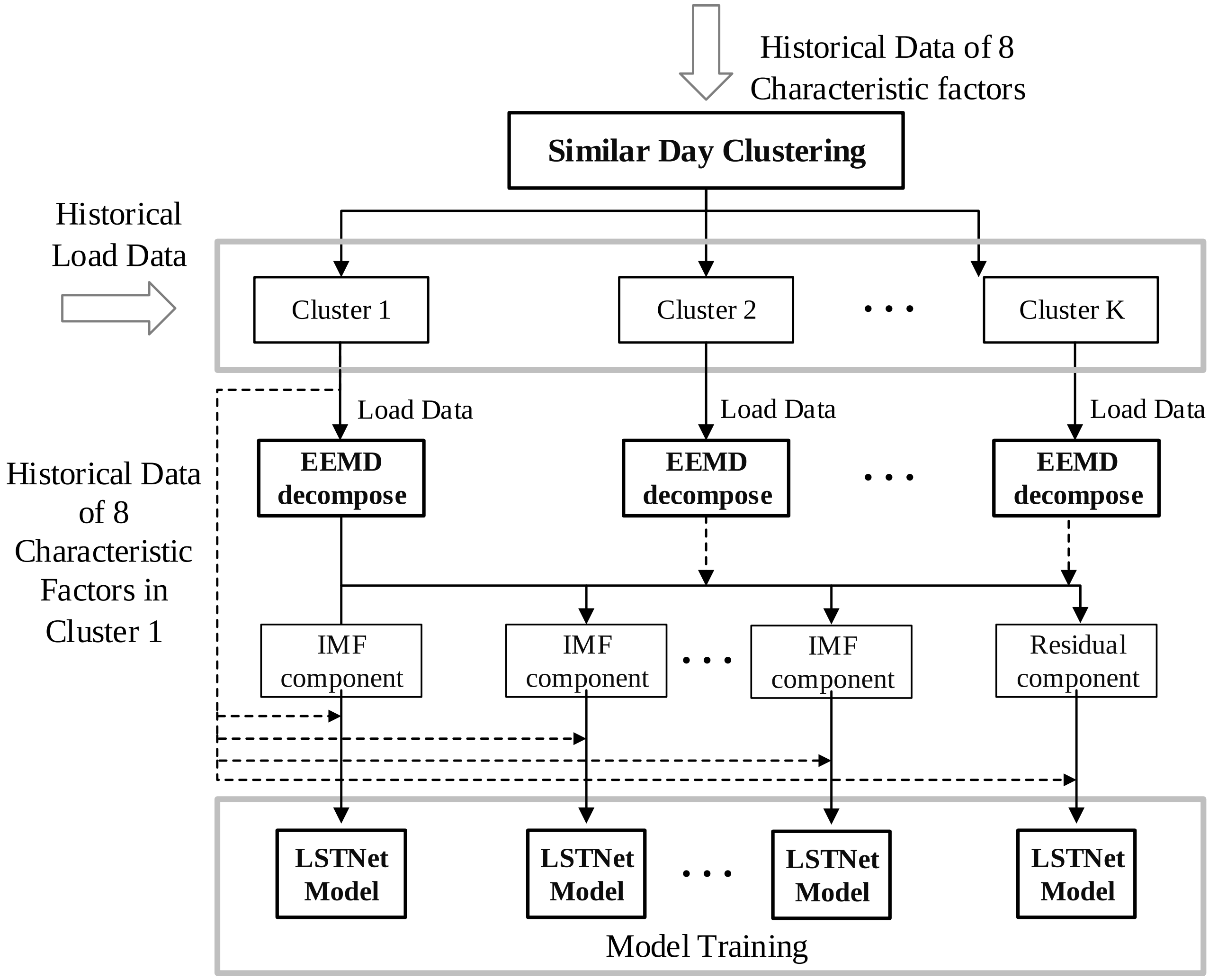

Figure 3. The process of the proposed load forecasting method.

3. Experimental Verification

The proposed method was verified by a photovoltaic energy storage system, which is located in the chemical laboratory building of Tsinghua University in Beijing. The real power load data of the laboratory from 1 August 2021 to 17 March 2022 were used as a training set (60%), a verification set (20%) and a test set (20%) to construct the forecast model. The model was applied to the system on 18 March 2022 for load prediction. The prediction results were used to compare with the real power load data of the day. This paper shows results of the similar day clustering, EEMD decomposition and LSTNet model prediction. A comparison experiment was also implemented between the proposed method and LSTM, CNN-LSTM, CNN-BiLSTM and LSTNet to verify the effectiveness of the proposed method.

The experiment used MAE and RMSE as error evaluation indexes. MAE represents the average value of the predicted value deviating from the actual value. MSE represents the mean value of the square of the error of the predicted value from the actual value. Since the square leads to dimensional changes, when the deviation becomes large, MSE will be amplified. In order to eliminate the dimensional influence, the square root of MSE is taken as RMSE. MAE and RMSE are calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - \hat{y}_i'|$$

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

where n represents the prediction data points; $\hat{y}_i$ represents the normalized actual load value for the ith prediction point and $\hat{y}_i'$ represents the normalized model predicted load value for the ith prediction point.

3.1. Model Training

We collected the historical data of power load, holiday, temperature, humidity, season, light intensity, precipitation, ground wind speed and relative air pressure from 1 August 2021 to 17 March 2022.

Based on the CH index, K-means clustering was performed on the collected daily characteristic factor data. For different numbers of clusters, the values of the CH index are shown in Figure 4. We can see that the optimal number of clusters should be five.

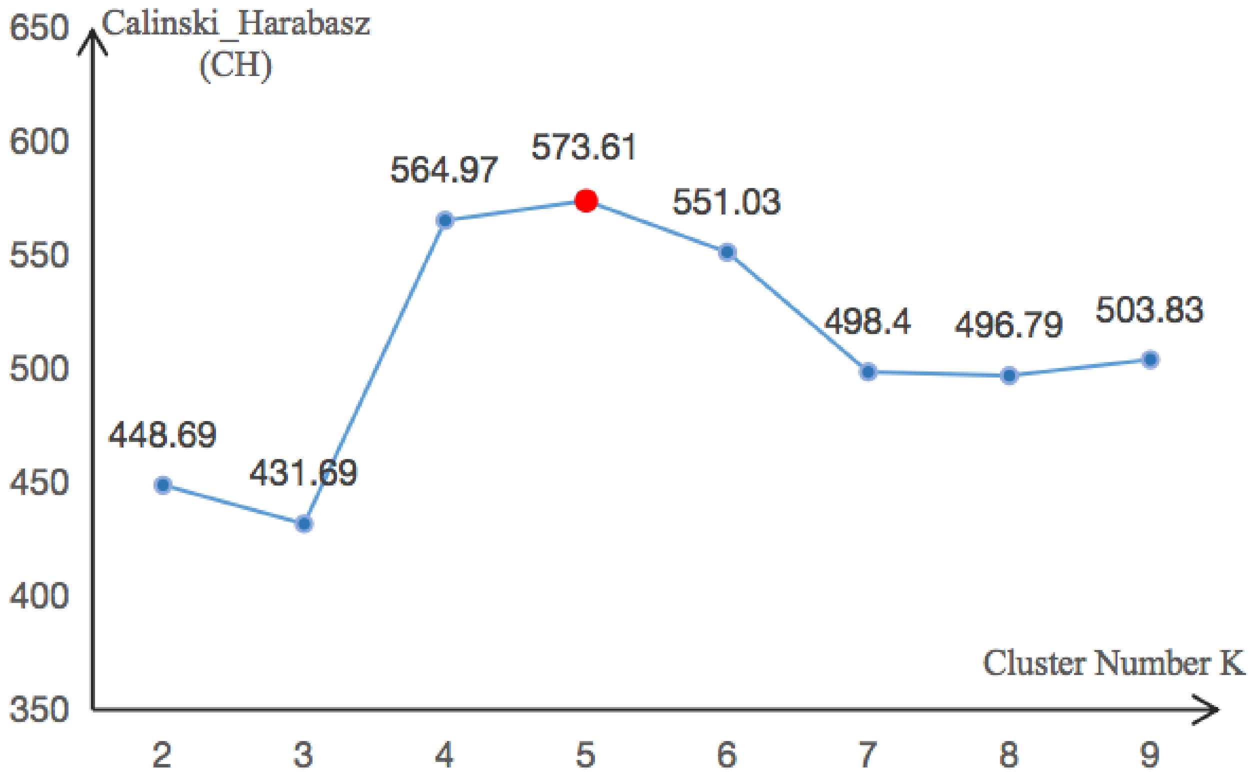

Figure 4. The CH values for different numbers of clusters.

We divided the days to be considered into five clusters. After down sampling, five clusters of similar day historical data with an interval of 15 min were obtained. We denoted them as K_1, K_2, K_3, K_4 and K_5.

For each cluster K_i, its load data were decomposed to load sub-sequences with EEMD. We calculated the correlation coefficient NCC_i and the standard deviation TH. We then eliminated the sub-sequences whose correlation coefficients NCC_i were less than the standard deviation TH.

For example, after EEMD decomposition for K_3, eight stationary sub-sequences and one residual sub-sequence were obtained as shown in Figure 5.

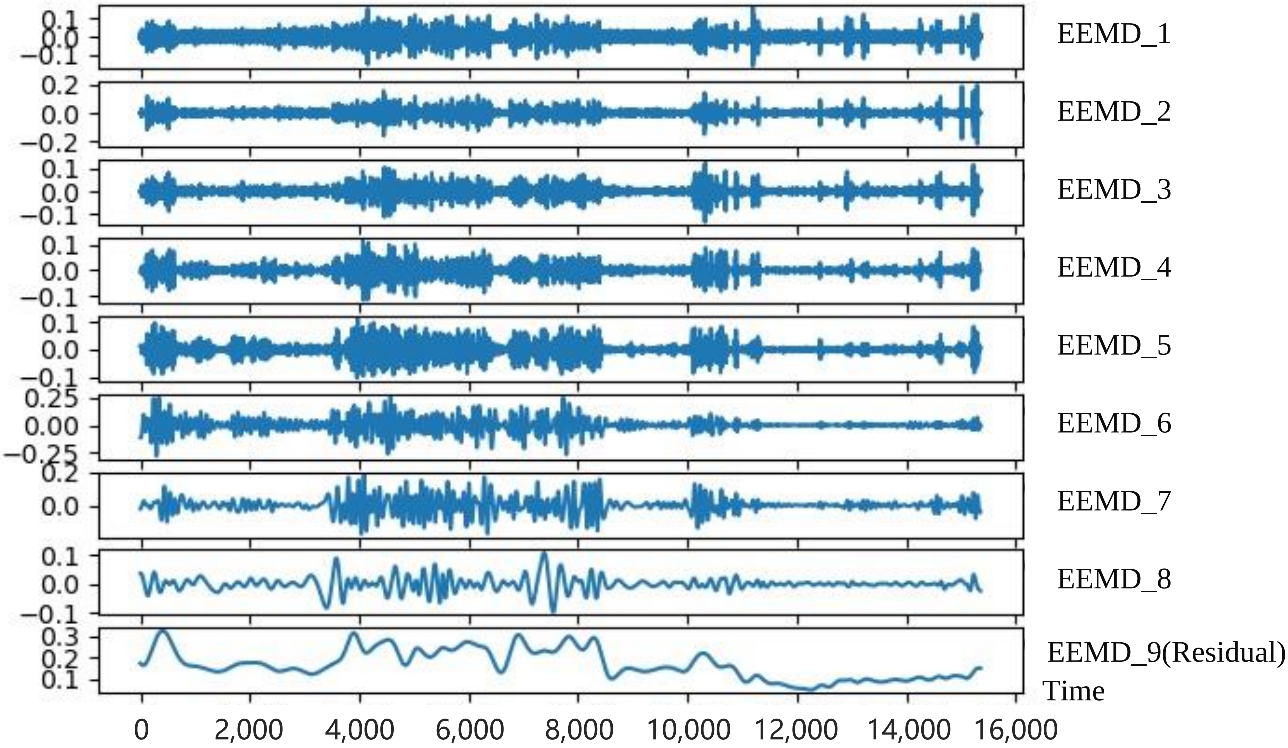

Figure 5. EEMD decomposition for K_3.

The correlation coefficients NCC_i for each subsequence and TH are shown in Table 2. Since the correlation coefficients of the 1, 2, and 3 sub-sequences were less than the standard deviation TH, they were eliminated.

Table 2. The correlation coefficient NCC_3 and the standard deviation TH for sub-sequences of K_3.

K_3 Sub-Sequences	1	2	3	4	5	6	7	8	9
NCC3	0.087	0.141	0.143	0.204	0.270	0.548	0.544	0.318	0.594
TH					0.186				

Each similar day historical sub-sequence was put into a different LSTNet network for training. For example, the remaining six load sub-sequences of K_3 were combined with the characteristic factor sequence of the corresponding time to form six similar day historical sub-sequences. Six different LSTNet load forecasting sub-models were obtained by training, respectively. These sub-models forecast the power load for the next 4 h (16 points) with an interval of 15 min together. The results of the model training are shown in Figure 6.

The evaluation indexes of six LSTNet load forecasting sub-models and the load forecasting performance are shown in the following Table 3.

3.2. Load Forecasting

The proposed method was applied to the photovoltaic energy storage system in Tsinghua University on 18 March 2022. The data of eight characteristic factors for one point on 18 March 2022 are shown in the following Table 4.

We first calculated the mean distance of the data of the eight characteristic factors on 18 March 2022 with the data of the eight characteristic factors in each cluster and put them in the cluster with the shortest distance. This belongs to K_3. The latest 30 days of data in K_3 were used for EEMD decomposition. After EEMD decomposition, the data of the last 16 h were selected as the input to the six trained LSTNet sub-models to obtain six forecasting sub-outputs. They were added to form a load forecasting output with 16 points for the next 4 h.

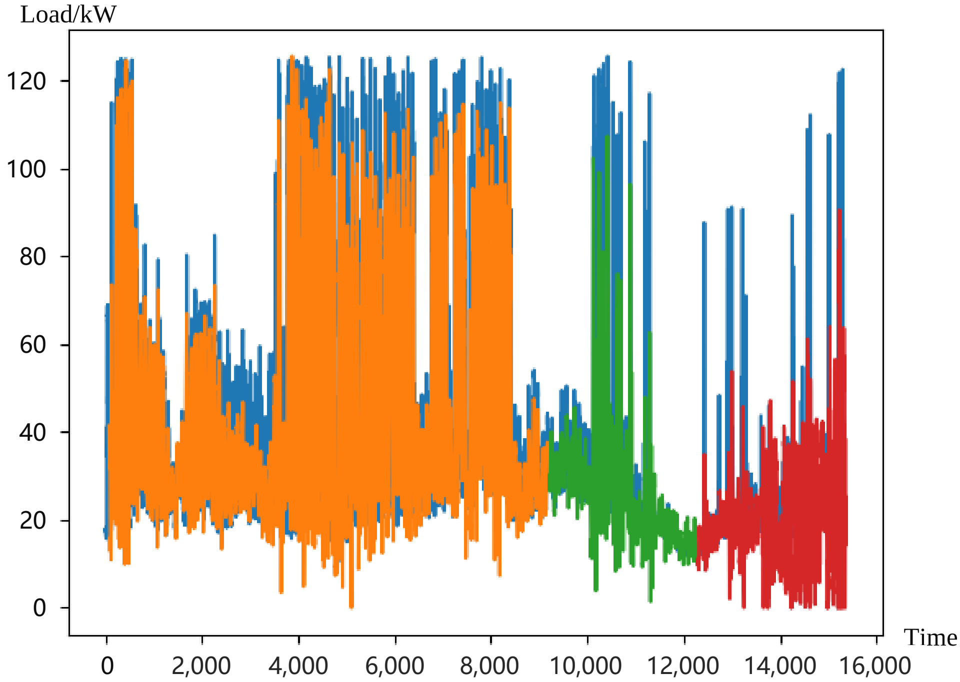

Figure 6. Results of model training. Blue is the original data, yellow is the training set, green is the verification set and red is the test set.

Table 3. The evaluation indexes of model training for load forecasting.

K_2 Sub-Sequences	Error Index	Training Set	Verification Set	Test Set
EEMD_4	MAE RMSE	0.009 0.015	0.006 0.012	0.007 0.011
EEMD_5	MAE RMSE	0.008 0.014	0.005 0.009	0.007 0.012
EEMD_6	MAE RMSE	0.008 0.014	0.004 0.007	0.004 0.007
EEMD_7	MAE RMSE	0.006 0.010	0.005 0.009	0.004 0.006
EEMD_8	MAE RMSE	0.001 0.002	0.002 0.004	0.002 0.003
EEMD_9	MAE RMSE	0.005 0.008	0.005 0.006	0.004 0.006
K-means_EEMD_LSTNet	MAE RMSE	0.024 0.038	0.016 0.028	0.015 0.026

Table 4. The data of 8 characteristic factors for one point on 18 March 2022, where 0 represents working days and 1 represents holidays for the factor of holiday; 0, 1, 2 and 3 represent spring, summer, autumn and winter for the factor of season, respectively.

Time	Holiday	Temperature (°C)	Humidity (%)	Season	Light Intensity (W/m^2)	Precipitation (mm/h)	Ground Wind Speed (m/s)	Relative Air Pressure (hPa)
2022-03-18 00:00:00	0	−1.226	87.7	0	95.916	0.184	1.607	1014.044

The proposed method outputs the load forecasting results for the next 4 h each time and updates it every 4 h. Each update needs to add the load and characteristic factor data of the latest 4 h to the historical data.

Comparing the load forecasting results with the actual load on 18 March 2022, as shown in Figure 7, the results show that the load prediction value is very close to the actual load value at each point.

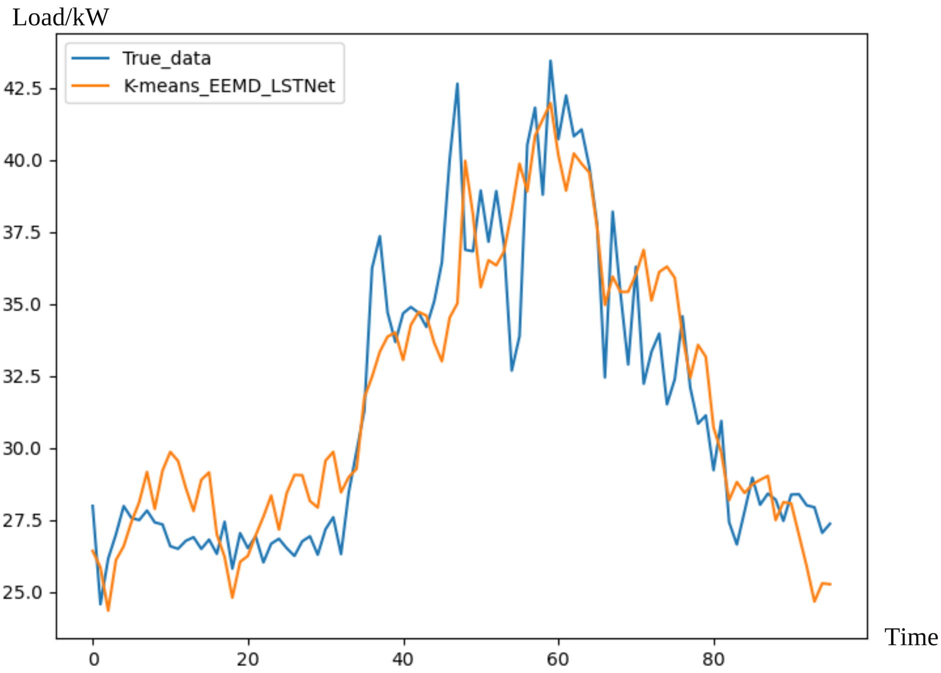

Figure 7. Comparison between the load forecasting results and the actual load on 18 March 2022.

3.3. Comparison Experiment

In order to verify the effectiveness of our method, we compared it with other models including LSTM, CNN-LSTM, CNN-BiLSTM and LSTNet. The historical data from 1 August 2021 to 15 March 2022 were directly put into these models for training. After training, we obtained corresponding load forecasting models which could be directly used for load prediction.

The comparison of the prediction indexes is shown in Table 5. It shows that the performance of the proposed method in this paper is significantly better.

Compared with LSTNet, MAE of K-means-EEMD-LSTNet descends to 0.015 from 0.045 and RMSE descends to 0.026 from 0.070, respectively. Therefore, similar day clustering and EEMD are helpful for improving the accuracy of load forecasting.

The load prediction results on 18 March 2022 of these models are shown in Figures 8 and 9. The results show that the load forecasting method combining K-means, EEMD and LSTNet can effectively improve the accuracy of power load forecasting.

Table 5. The prediction index comparison of LSTM, CNN-LSTM, CNN-BiLSTM, LSTNet and K-means-EEMD-LSTNet load forecasting models.

Model	MAE	RMSE
K-means_EEMD_LSTNet	0.015	0.028
LSTNet	0.045	0.070
CNN-BiLSTM	0.087	0.132
CNN-LSTM	0.073	0.117
LSTM	0.091	0.139

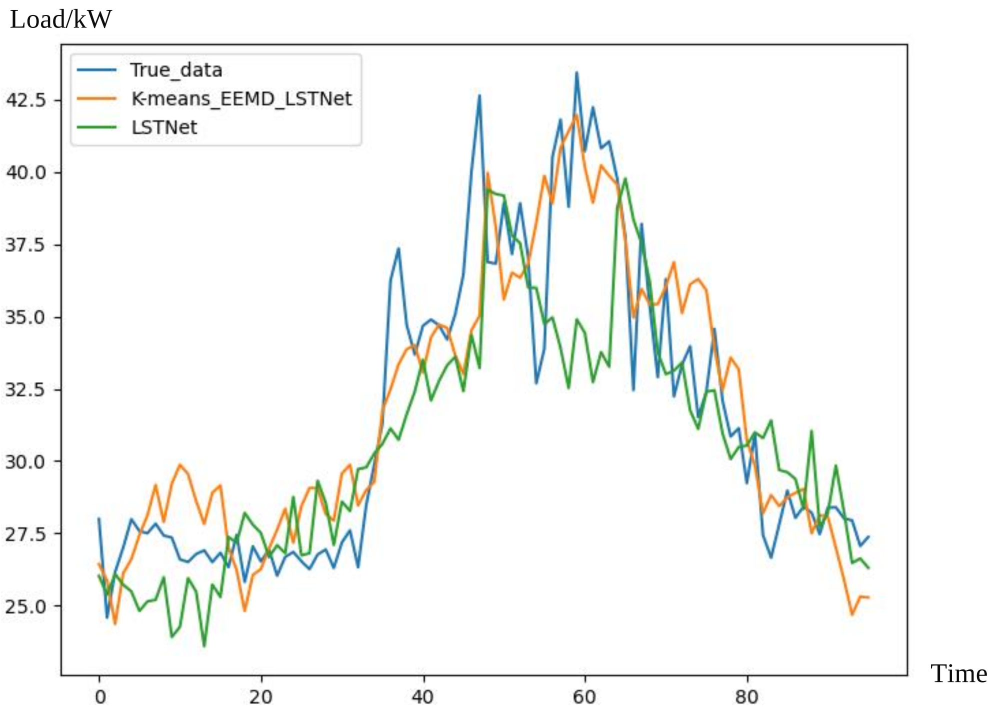

Figure 8. Comparison of predicted results for LSTNet and K-means-EEMD-LSTNet load forecasting models.

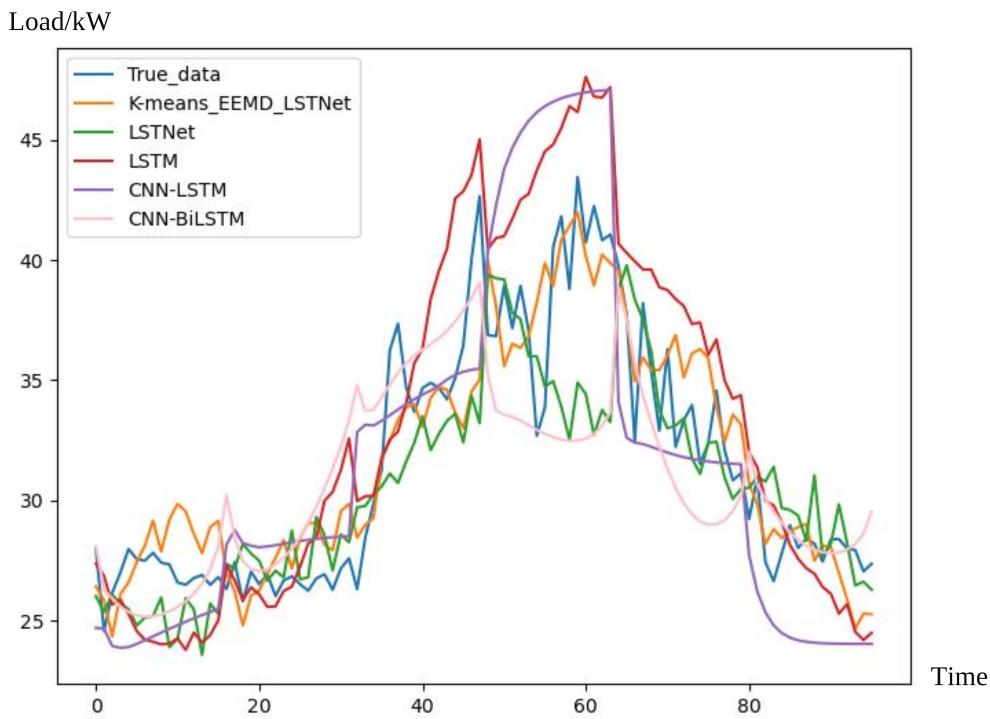

Figure 9. Comparison of predicted results for LSTM, CNN-LSTM, CNN-BiLSTM, LSTNet and K-means-EEMD-LSTNet load forecasting models.

4. Summary and Prospect

In this paper, similar day clustering and EEMD decomposition were proposed to combine with LSTNet for ultra short-term power load forecasting. Since the similar day and EEMD decomposition can decompose the original non-stationary historical data sequence into stationary sub-sequences, which is very suitable for LSTNet, the performance of load prediction has been improved significantly. As shown in the comparison experiment, the proposed method improves the accuracy of power load forecasting effectively.

In the proposed method, the data of the eight characteristic factors for the day to be predicted are needed. However, they cannot be directly obtained and what we can acquire is their prediction values. One possible work in the future is to determine the type of the next day using the data of the past days.

Author Contributions: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Visualization, W.Z., L.C. and S.S.; Resources, C.S., X.T. and J.L.; Data Curation, J.L.; Writing—Original Draft Preparation, W.Z.; Writing—Review and Editing, S.S. and J.Z.; Supervision, C.S. and X.T.; Project Administration, X.T. and C.S.; Funding Acquisition, X.T. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially funded by Science Project of China Huadian Co., Ltd., under Grant CHDKJ21-02-77.

Data Availability Statement: Not applicable.

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

References

1. Lu, J.C.; Zhang, X.P.; Sun, W. A Real-time Adaptive Forecasting Algorithm for Electric Power Load. In Proceedings of the Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES, Dalian, China, 18 August 2005.
2. Alfares, H.K.; Nazeeruddin, M. Electric load forecasting: Literature survey and classification of methods. *Int. J. Syst. Sci.* **2002**, *33*, 23–34. [CrossRef]
3. Kyriakides, E.; Polycarpou, M. Short term electric load forecasting: A tutorial. In Proceedings of the 1st International Conference on Natural Computation (ICNC 2005), Changsha, China, 27–29 August 2005; p. 391.
4. Yang, J. Power System Short-Term Load Forecasting Power System. In Proceedings of the 2017 5th International Conference on Machinery, Materials and Computing Technology ICMMCT 2017, Beijing, China, 25–26 March 2017; pp. 1–7.
5. Loewenstem, Y.; Katzir, L.; Shmilovitz, D. The Effect of System Characteristics on Very-Short-Term Load Forecasting. In Proceedings of the 12th International School on Nonsinusoidal Currents and Compensation (ISNCC), Lagow, Poland, 15–18 June 2015.
6. Haq, M.R.; Ni, Z. A New Hybrid Model for Short-Term Electricity Load Forecasting. *IEEE Access* **2019**, *7*, 125413–125423. [CrossRef]
7. Abu-Shikhah, N.; Elkarmi, F.; Aloquili, O.M. Medium-Term Electric Load Forecasting Using Multivariable Linear and Non-Linear Regression. *Smart Grid Renew. Energy* **2011**,*2*, 126–135. [CrossRef]
8. Ye, S.; Zhu, G.; Zhi, X. Long Term Load Forecasting and Recommendations for China Based on Support Vector Regression. *Energy Power Eng.* **2012**, *4*, 380–385. [CrossRef]
9. Karthika, S.; Margaret, V.; Balaraman, K. Hybrid Short Term Load Forecasting using ARIMA-SVM. In Proceedings of the Conference on Innovations in Power and Advanced Computing Technologies (i-PACT), Vellore, India, 21–22 April 2017.
10. Wu, M.H.; Yang, J.W. Analysis and comparison of several electric load forecasting methods. *Guangdong Electr. Power* **2004**, *1*, 17–21.
11. Taylor, J.W. Short-Term Load Forecasting With Exponentially Weighted Methods. *IEEE Trans. Power Syst.* **2012**, *27*, 458–464. [CrossRef]
12. Alade, I.O.; Abd Rahman, M.A.; Saleh, T.A. Predicting the specific heat capacity of alumina/ethylene glycol nanofluids using support vector regression model optimized with Bayesian algorithm. *Sol. Energy* **2019**, *183*, 74–82. [CrossRef]
13. Cevik, H.H.; Cunkas, M. Short-term load forecasting using fuzzy logic and ANFIS. *Neural Comput. Appl.* **2015**, *26*, 1355–1367. [CrossRef]
14. Cerne, G.; Dovzan, D.; Skrjanc, I. Short-Term Load Forecasting by Separating Daily Profiles and Using a Single Fuzzy Model Across the Entire Domain. *IEEE Trans. Ind. Electron.* **2018**, *65*, 7406–7415. [CrossRef]
15. Kwon, B.S.; Park, R.J.; Song, K.B. Short-Term Load Forecasting Based on Deep Neural Networks Using LSTM Layer. *J. Electr. Eng. Technol.* **2020**, *15*, 1501–1509. [CrossRef]
16. Nalcaci, G.; Ozmen, A.; Weber, G.W. Long-term load forecasting: Models based on MARS, ANN and LR methods. *Cent. Eur. J. Oper. Res.* **2019**, *27*, 1033–1049. [CrossRef]
17. Niu, D.; Li, J.; Li, J.; Wang, Q. Daily load forecasting using support vector machine and case-based reasoning. In Proceedings of the 2nd IEEE Conference on Industrial Electronics and Applications (ICIEA 2007), Harbin, China, 23–25 May 2007; pp. 1271–1274.
18. Zhang, Z.C.; Hong, W.C. Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads. *Knowl.-Based Syst.* **2021**, *228*, 107297. [CrossRef]
19. Li, L.Z.; Ota, K.; Dong, M.X. Everything is Image: CNN-based Short-term Electrical Load Forecasting for Smart Grid. In Proceedings of the 14th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN) / 11th International Conference on Frontier of Computer Science and Technology (FCST) / 3rd International Symposium of Creative Computing (ISCC), Exeter, UK, 21–23 June 2017; pp. 344–351.

20. Sehovac, L.; Grolinger, K. Deep Learning for Load Forecasting: Sequence to Sequence Recurrent Neural Networks With Attention. *IEEE Access* **2020**, *8*, 36411–36426. [CrossRef]
21. Han, L.Y.; Peng, Y.X.; Li, Y.H.; Yong, B.B.; Zhou, Q.G.; Shu, L. Enhanced Deep Networks for short-Term and Medium-Term Load Forecasting. *IEEE Access* **2019**,*7*, 4045–4055. [CrossRef]
22. Dong, X.; Qian, L.; Huang, L. Short-Term Load Forecasting in Smart Grid: A Combined CNN and K-Means Clustering Approach. In Proceedings of the IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, Republic of Korea, 13–16 February 2017; pp. 119–125.
23. Box, G.; Jenkins, G.M.; Reinsel, G.C. Time series analysis forecasting and control. *J. Time* **1976**, *31*, 238–242.
24. Gensler, A.; Henze, J.; Sick, B.; Raabe, N. Deep Learning for Solar Power Forecasting - An Approach Using Autoencoder and LSTM Neural Networks. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; pp. 2858–2865.
25. Zheng, J.; Xu, C.C.; Zhang, Z.; Li, X.H. Electric Load Forecasting in Smart Grids Using Long-Short-Term-Memory based Recurrent Neural Network. In Proceedings of the 51st Annual Conference on Information Systems and Sciences (CISS), Baltimore, MD, USA, 22–24 March 2017; pp. 1–6.
26. Narayan, A.; Hipel, K.W. Long Short Term Memory Networks for Short-Term Electric Load Forecasting. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 2573–2578.
27. Zhao, J.; Huang, J. Application and research of short-term Power load forecasting based on LSTM. In Proceedings of the 2020 International Conference on Information Science, Parallel and Distributed Systems (ISPDS), Xi'an, China, 14–16 August 2020; pp. 212–215.
28. Zhang, Y.; Ai, Q.; Lin, L.; Yuan, S.; Li, Z. A Very Short-term Load Forecasting Method Based on Deep LSTM RNN at Zone Level. *Power Syst. Technol.* **2019**, *43*, 1884–1892.
29. Zhuang, Y.N.; Chen, M.; Pan, F.F.; Feng, L.; Liang, Q.H. Research on Short Term Power Load Forecasting Combining CNN and LSTM Networks. In Proceedings of the 14th International Conference on Intelligent Robotics and Applications (ICIRA), Yantai, China, 22–25 October 2021; pp. 628–638.
30. Wu, L.Z.; Kong, C.; Hao, X.H.; Chen, W. A Short-Term Load Forecasting Method Based on GRU-CNN Hybrid Neural Network Model. *Math. Probl. Eng.* **2020**, *2020*, 1428104. [CrossRef]
31. Kim, T.Y.; Cho, S.B. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* **2019**, *182*, 72–81. [CrossRef]
32. Rafi, S.H.; Nahid Al, M.; Deeba, S.R.; Hossain, E. A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network. *IEEE Access* **2021**, *9*, 32436–32448. [CrossRef]
33. Wu, K.; Wu, J.; Feng, L.; Yang, B.; Liang, R.; Yang, S.; Zhao, R. An attention-based CNN-LSTM-BiLSTM model for shortterm electric load forecasting in integrated energy system. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e12637. [CrossRef]
34. Liao, L.; Xin, J.H.; Zhai, H.Q.; Wei, Z.H. Short Term Load Forecasting Model and the Influencing Factors. *J. Shanghai Jiaotong Univ.* **2004**, *38*, 1544–1547.
35. Raza, M.Q.; Khosravi, A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renew. Sustain. Energy Rev.* **2015**, *50*, 1352–1372. [CrossRef]
36. Lai, G.; Chang, W.-C.; Yang, Y.; Liu, H. Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks. In Proceedings of the 41st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), Ann Arbor, MI, USA, 8–12 July 2018; pp. 95–104.
37. Liu, R.; Chen, L.; Hu, W.H.; Huang, Q. Short-term load forecasting based on LSTNet in power system. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e13164. [CrossRef]
38. Liao, N.H.; Hu, Z.H.; Ma, Y.Y.; Lu, W.Y. Review of the short-term load forecasting methods of electric power system. *Power Syst. Prot. Control* **2011**, *39*, 147–152.
39. Wu, Z.H.; Huang, N.E. Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Adv. Adapt. Data Anal.* **2011**, *1*, 41. [CrossRef]
40. Wong, J.A.H.A. Algorithm AS 136: A K-Means Clustering Algorithm. *J. R. Stat. Soc.* **1979**, *28*, 100–108.

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