

## Research Article

# Microgrid Load Forecasting Based on Improved Long Short-Term Memory Network

Qiyue Huang <sup>1</sup>, Yuqing Zheng,<sup>2</sup> and Yuxuan Xu<sup>1</sup>

<sup>1</sup>Ningbo Polytechnic, Ningbo 315800, China

<sup>2</sup>Ningbo Center for CEEC Expo & Cooperation Promotion, Ningbo 315800, China

Correspondence should be addressed to Qiyue Huang; 839539199@qq.com

Received 15 January 2022; Revised 26 February 2022; Accepted 1 March 2022; Published 20 March 2022

Academic Editor: François Vallée

Copyright © 2022 Qiyue Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the fast growing of new energy technologies, the proportion of distributed renewable energy sources dominated by wind and light energy in the microgrid continues to increase. However, the uncertainty and randomness of energy itself bring challenges to the stable operation of the power system. Microgrid load forecasting with high accuracy is the key means to handle the above problems. It can provide help for power grid dispatching and decision-making, optimize resource allocation, reduce operation cost, and ensure system safety. In this paper, a load-forecasting algorithm for microgrid based on improved long short-term memory neural network (LSTM) is proposed. Firstly, the criticality analysis of load influencing factors is carried out, and the clustering classification and weight calculation are completed. Then, the input data is preprocessed to ensure the quality of database. Secondly, the LSTM gets improved from three aspects: multilayer convolution channel, lookahead optimizer, and AM weight. And a complete forecasting model is designed to accomplish the load forecasting. Finally, based on the data of a local microgrid in Zhejiang Province, China, simulation experiments are conducted. The results are quantitatively compared with other forecasting algorithms to verify the accuracy and superiority of the proposed algorithm.

## 1. Introduction

With the rapid progress of science and economy across the world, the environmental pollution caused by traditional ways of power generation is becoming more and more serious [1]. The new energy technology such as wind power and photovoltaic power has become the focus of research at home and abroad. However, renewable energy itself has considerable randomness and volatility, which will adversely affect the stability and reliability of power grid when used in power system widely [2, 3]. As a small, decentralized and independent system, microgrid is featured by self-control, protection, and management. The fast development of related technologies provides the possibility to promote the wide access of new energy sources and the reliable supply of multienergy forms [4]. The accurate prediction of microgrid load power and its inclusion in dispatching plan are important guarantees to promote the security and economy of modern power system [5].

Traditional load forecasting has been developed over the years and can be divided into time series methods represented by autoregressive models and deep learning algorithms represented by artificial neural networks. The paper [6] uses Deep Belief Networks for short-term load forecasting. The thesis [7] first performs feature extraction of load features by convolutional neural networks and then load forecasting by long short-term neural networks. Artificial intelligence has a very prominent data mining capability and complex function approximation capability, which is widely used in practical engineering. Its ability to autonomously mine the implicit features of load and its influencing factors has made it a mainstream method for probabilistic load prediction. In addition, gray theory, chaos theory, wavelet analysis, and other theories are also used for load prediction [8–10]. Many contributions have been made by domestic and foreign scholars to further improve the prediction accuracy and the generalization ability of the model. The article [11] used CNN to extract the nonlinear

invisible features between influencing factors and load in the short-term load forecasting model and improved the generalization ability of the model by using fuzzy time series. The paper [12] used CNN to extract spatially adjacent load-related features to improve the load forecasting effect. However, most of the above methods are targeted in traditional power systems and are not applicable to new microgrid systems with a large proportion of new energy access. The volatility of new energy sources puts higher demands on the accuracy of microgrid load forecasting, and the excessive information processing also leads to problems such as large computational workload and low efficiency. In addition, how to complete the data feature selection and improve the expression ability of neural network based on the big data information of microgrid is also an urgent problem to be solved.

To deal with these problems, a load forecasting algorithm for microgrid based on improved long short-term memory network is proposed in this paper. Firstly, the power load of microgrid is analyzed, and the key factors that influence the load value are found out. Then different types of key factor data are collected to complete characteristic analysis and feature extraction, forming the data set of key influencing factors. Secondly, the data are preprocessed. The missing data are supplemented, and the abnormal data are eliminated and smoothed to ensure the authenticity and reliability of the feature database. Then, traditional LSTM are optimized from three aspects: multilevel convolution channel, attention learning reinforcement, and lookahead optimizer to improve the prediction accuracy and convergence speed of the algorithm. Finally, a microgrid data set collected in Zhejiang Province, China, is simulated, and the load forecasting is completed. Compared with the traditional neural network, the accuracy and reliability of the new algorithm are verified.

## 2. Analysis of Influencing Factors and Data Processing

The output of microgrid load is affected by many aspects, such as climate, temperature, economic level, time, and special events. The change of load value shows randomness; however, it also has certain periodicity and regularity [13]. In this chapter, the load output of several typical microgrids is studied, and the possible load influencing factors are analyzed to find out the key influencing elements. The key data are preprocessed to improve the quality of input data of neural network and ensure the accuracy of microgrid load forecasting.

*2.1. Analysis of Key Influencing Factors.* Research shows that the load output of microgrid is affected not only by the structure and characteristics of the system itself, but also by the factors outside the system [14, 15]. From the economic aspect, the gross domestic product (GDP) is closely related to the load output. From an industrial point of view, the proportion of power consumption in regional factories has a greater impact on load. From meteorological factors,

extreme weather has large influence on load output. From the point of view of time, month, season, holidays will affect the load. From the angle of marketing, peak and valley power will act on the load. From the microgrid structure, the proportion of new energy will also play a role in load. In order to reveal the influence of different kinds of influencing factors on power grid output, this paper analyzes the characteristics of all factors.

The influencing factors of load are numerous; as a result, the amount of data is huge. In order to select key influencing factors conveniently and improve the efficiency of load forecasting, the influence weights of different factors are calculated separately. First, K-means clustering algorithm is used to cluster the original data and add corresponding labels. Then, a feature selection model is established, and the maximum sum of intervals is taken as the objective function to complete the feature selection. Euclidean distance is adopted for distance calculation of clustering algorithm, and the formula is as follows:

$$d_{mn} = \sqrt{\sum_{i=1}^N (x_{mi} - x_{ni})^2}, \quad (1)$$

where  $x_{mi}$  and  $x_{ni}$  are the data samples of different types of influencing factors.

Once clustering is complete, the sample data is mapped into a weight space  $W$  and intervals are calculated to reflect the data structure. The larger the interval, the better the data structure. Based on this, the data samples are trained to find the best feature weight. The formula is as follows:

$$D(\omega) = \|x - N_0(x)^T \omega\| - \|x - N_1(x)^T \omega\|, \quad (2)$$

where  $\omega$  denotes the weight vector,  $N_0(x)^T$  is the similar nearest neighbor of the sample  $x$ , and  $N_1(x)^T$  stands for the nonhomogeneous nearest neighbor of the sample  $x$ . After completing clustering analysis and feature weight calculation, the obtained results are shown in Table 1. In the long run, the price of electricity is the main factor affecting the load, and industrial power consumption is the most frequent factor. The proportion of new energy in this region is relatively high, and the electricity price of new energy also has certain impact on the load. In the short term, load output is greatly affected by time. The occurrence of special events also affects the load value of microgrid.

*2.2. Data Processing.* Analyzing influencing factors, there are many factors affecting the load prediction of microgrid. However, there are great differences between different types of collected data, which makes it impossible to compare and train directly, thus easily leading to neuron saturation. Therefore, it is necessary to normalize all data to ensure that they are within the same interval, so as to facilitate statistics calculation [16]. The normalization formula is as follows:

$$x_i^* = \frac{x_i - \bar{x}}{\sqrt{1/n - 1 \sum_{i=1}^n (x_i - \bar{x})^2}}, \quad (3)$$

TABLE 1: Analysis of influencing factors.

No.	Name	Description	Weight
01	Temperature	Current temperature of microgrid	0.74
02	Time	Current specific moment	0.82
03	Day	Working days, weekends, holidays	0.81
04	Month	Current month	0.76
05	Special events	Extreme high temperature, extreme low temperature, heavy snow, heavy rain, strong wind, epidemic, etc.	0.82
06	Proportion of heavy industry	Proportion of current heavy industry users	0.54
07	Proportion of light industry	Proportion of current light industry users	0.39
08	Proportion of charging pile	Proportion of current automobile charging piles	0.14
09	Proportion of new energy	Proportion of photovoltaic energy and wind energy	0.64
10	Civil electricity price	Current civil electricity price	0.11
11	Electricity price for industry and commerce	Current electricity price for industrial and commercial use	0.84
12	Electricity price of new energy	Electricity price of new energy generation	0.78
13	GDP	Current gross domestic product	0.44
14	Population	Total population of the current area	0.39

where  $x_i^*$  shows the data after normalization processing,  $x_i$  stands for the data before processing, and  $\bar{x}$  is the average of the initial data.

Because of the huge amount of data acquisition in the load forecasting of microgrid, the data loss is frequently caused by equipment failure or communication error. The lack of data will adversely affect the training quality of forecasting. In order to ensure the continuity and stability of load data, it is necessary to complete the missing data. To ensure the repair quality, the data shall be supplemented by the weighted processing method. The formula is as follows:

$$y_i = \lambda_1 y_{i-1} + \lambda_2 y_{i+1} + \lambda_3 y_i^*, \quad (4)$$

where  $y_i$  represents the missing data to be added;  $y_{i-1}$  and  $y_{i+1}$  are the data at the adjacent time,  $y_i^*$  indicates the similar data, and  $\lambda_{1-3}$  embodies the weight.

In addition to missing data, abnormal data may occur in data acquisition due to interference signal and other reasons. To ensure the prediction quality, it is necessary to eliminate and smooth the abnormal data. In this paper, the method of deviation identification is used to identify the abnormal data. When the identification is completed, the abnormal data are removed, and the missing data will be filled in. The anomaly identification formula is as follows:

$$\Delta x_i = \frac{|x_i - E(x)|}{\sqrt{D(x)}} \geq \varepsilon, \quad (5)$$

where  $E_x$  denotes the expectation of the data  $x$ ,  $D_x$  indicates the variance of the data  $x$ , and  $\varepsilon$  means the set identification threshold value.

### 3. Improvement of LSTM Algorithm

In order to improve the accuracy and convergence speed of load forecasting algorithm, LSTM is improved in this paper. Traditional algorithms are optimized from three aspects: multiconvolution channel cooperative optimization, attention learning enhancement, and lookahead optimizer optimization.

**3.1. Cooperative Optimization.** Traditional load forecasting algorithm uses single channel to perform convolution calculation on data. However, the prediction model of single channel is easily influenced by external noise, so the reliability of data acquisition is strictly required. Especially when the input types of load forecasting models are quite different, single channel cannot meet the needs of high prediction accuracy of microgrid.

Therefore, in this paper, a multiconvolution channel structure is constructed, and a cooperative neural network prediction model is adopted. At the same time of data training and load forecasting, data exchange and sharing are carried out to establish stable reconstruction loss. The accuracy of model prediction is improved by cross training and cooperative training. The structure of cooperative prediction is shown in Figure 1.

**3.2. Optimization of Attention.** Research shows that the characteristic signals in the key data feature set are easy to be submerged in a large amount of meaningless redundant data, which results in the difficulty and huge workload of forecasting. The attention mechanism is used to optimize the prediction model. By weighting the output of LSTM, the important features are highlighted, and the prediction efficiency is well guaranteed.

Attention mechanism is a kind of weighting function in essence. By integrating the information of multidimensional eigenvectors  $f_i$ , a new feature vector  $f_i^*$  can be obtained. This feature vector can reflect critical information of a plurality of features. The weighting formula is as follows:

$$f_i^* = \sum_{i=1}^n \alpha_i f_i, \quad (6)$$

where  $\alpha_i$  stands for the weight, which is also the key information in attention optimization.

Attention weight is mainly operated through the scoring function  $g$ . Each characteristic value is scored, with higher score indicating that the feature is more important. After scoring, the final weight is calculated using the SoftMax

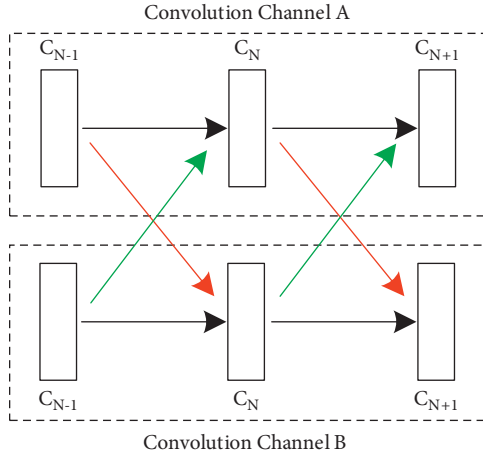


FIGURE 1: Structure diagram of cooperative training.

function. In this paper, attention optimization layer is added between LSTM layer and full connection layer, namely,

$$\alpha = \sigma(W_\alpha f + b), \quad (7)$$

where  $W_\alpha$  denotes the attention weight matrix,  $b$  shows the paranoid quantity, and  $\sigma$  indicates the sigmoid nonlinear function.

With the attention optimization mechanism, the data processing of the prediction model can be focused on the key features and ignore the irrelevant features, thus ensuring the prediction efficiency and accuracy.

**3.3. Lookahead Optimizer.** The weight optimization of LSTM in load forecasting has always been the focus of neural learning. High-quality optimization measures can guarantee the convergence speed and stability of the model. In this paper, a forward-looking called lookahead optimizer is used to determine the search direction by looking ahead at a fast weight sequence generated by another optimizer. Combined with the conventional Adam optimizer, this optimizer can improve the fitting speed and generalization ability and guarantee the robustness of self-learning.

The fast weight part of the optimizer means that the traditional Adam optimizer updates the model  $k$  times full-time and saves the sequence weights. The update expression is as follows:

$$\theta_{j,i+1} = \theta_{j,i} + A_d(L, \theta_{j,i-1}, x), \quad (8)$$

where  $A_d$  is the Adam optimizer function,  $L$  represents the model objective function,  $\theta$  shows the update weight, and  $x$  is the input data.

The slow weight part updates the weight of training model through exponential weighted average algorithm. The formula is as follows:

$$\omega_{j+1} = \omega_j + \alpha(\theta_{j,k} - \omega_j), \quad (9)$$

where  $\omega_j$  denotes the model weight, and  $\alpha$  shows the learning rate.

In the initialization phase, the initial value of training  $\theta_{j+1,0} = \omega_{j+1}$ .

The working steps of the lookahead optimizer are shown in Figure 2.

## 4. Improve LSTM Prediction Model

**4.1. The Basic Principles of CNN.** In order to fully excavate the deep feature of the image, convolution neural network is used in this paper to extract the original feature information effectively [17, 18]. Convolutional neural networks can reduce the number of parameters inside the network by special connection between adjacent layers. The latter convolution layer can extract the data features of the previous convolution layer and convert the feature information from low level to high level. It mainly consists of convolution layer, pooling layer, active layer, and full connection layer.

Convolution layer is the most important link in convolution neural network, which uses filter to obtain the characteristic information of previous layer data. Generally, after the convolution layer, the pooling layer resamples and obtains new feature information by using the correlation characteristics of the input data. Its main function is to process the dimension of input data and improve the efficiency of information processing. The activation layer improves the convergence speed through the activation function and avoids the problem of gradient disappearance. The full-connection layer is usually located at the end of the network, which integrates and outputs the data to obtain the eigenvector reflecting the characteristics of the data.

**4.2. The Basic Principles of LSTM.** LSTM neural network belongs to cyclic neural network [19]. It can solve the problem of long term dependence through the inner gating mechanism of neural unit. Its basic internal structure is as shown in Figure 3. Each cell is circularly connected to form a traditional RNN structure, and the internal unit adopts the gating mechanism with LSTM core.

**4.3. Improved LSTM Algorithm.** The improved LSTM neural network proposed in this paper takes the data set of the critical factors affecting the load output of microgrid as the input of the network. The feature extraction is carried out through two-layer convolution neural network channel. The high-value information contained in the data is further mined by cross-training between the two convolution layers. Output features of the convolutional neural network are then used as inputs to the LSTM. Its unique 3 gates and memory units can further screen the features. At the same time, lookahead optimizer is designed to inverse optimize the convolution neural network. At last, the output of LSTM is weighted by AM, and the load power is predicted. The detailed model structure is shown in Figure 4.

The structural design of the improved LSTM is related to the input data setting, the design of the convolutional kernel, and the choice of the activation function. The specific design of the proposed scheme is presented below in terms of each of these three aspects.

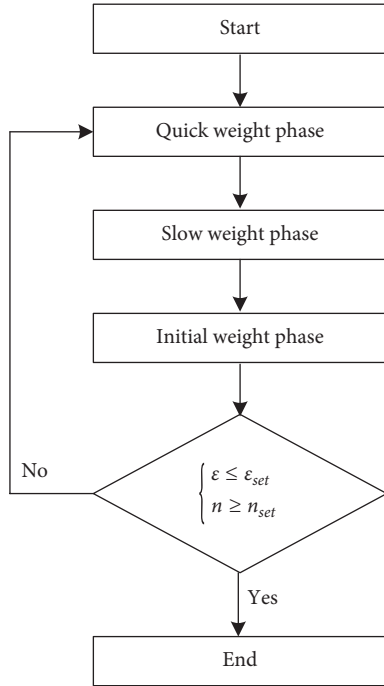


FIGURE 2: Optimization logic diagram.

The data input selected in this scheme is a multidimensional data matrix. After completing the data selection, the critical data with the highest relevance are used as the input to the prediction model. Since different types of data have different sensitivities to time scales, some data characteristics require longer time periods to be reflected in the data, while others are the opposite. It is not appropriate to use the same time window for data extraction. Therefore, for different types of data information, three different time resolutions, high, medium, and low, are used for data collection to ensure the accuracy and efficiency of power prediction. The traditional convolutional kernel design is prone to inconsistent weight matching factors. This leads to the need to identify the data in the feature extraction process of the convolution kernel, which increases the difficulty of feature extraction. In this paper, the problem of confusing weight matching is avoided by setting the lateral size of the convolutional kernel equal to the number of critical factors to ensure the convolutional direction. To ensure the nonlinear approximation ability of the neural network, and the suppression ability of useless signals, the ReLU function is used as the activation function. The function expression is as follows:

$$f(x) = \max(0, x). \quad (10)$$

## 5. Experimental Results and Analysis

In this paper, the microgrid load and related data sets of Zhejiang Province of China over the past 5 years are used for model training and simulation verification. Among them, 50,000 samples are selected for testing, 70% of which are

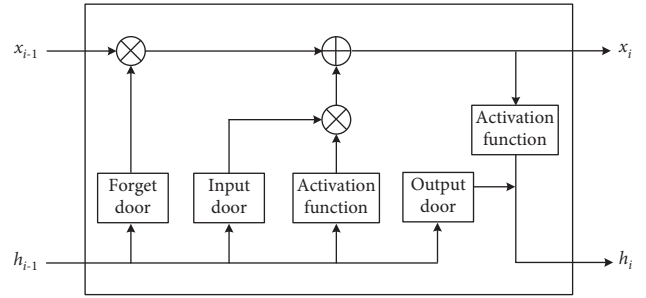


FIGURE 3: Structural drawing of LSTM unit.

taken as training set and 30% as verification set. This article uses Python as the programming language.

**5.1. Evaluation Index.** Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are used to measure the prediction error in order to quantitatively evaluate the accuracy of the improved LSTM model. The specific formula is as follows:

$$e_{\text{mape}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{z_i^* - z_i}{z_i} \right| \quad (11)$$

$$e_{\text{rmse}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (z_i^* - z_i)^2},$$

where  $z_i^*$  indicates the forecast data and  $z_i$  indicates the real data.

**5.2. Identification of Characteristics.** It can be seen from Table 1 that there are 14 factors influencing the load, and their critical weights are different. As the input of neural network, the data of key influencing factors will directly affect the accuracy and speed of load forecasting. Small amount of data will lead to poor prediction accuracy, while large amount of data will slow down the calculation speed. Therefore, it is necessary to select the number of influencing factors reasonably.

Firstly, the data of all influencing factors of load are taken as model input for load forecasting, and the prediction precision and optimal convergence speed are calculated, respectively. The key weight values are deleted one by one from low to high, and the prediction accuracy and optimal convergence speed are recalculated. The conclusions obtained are shown in Figure 5.

It can be seen that when the number of influencing factors is more than 7, the prediction error is minimum, and the prediction accuracy is the highest. The speed of convergence reaches the fastest at 8, and only about 0.56 s running time is needed to complete the accurate prediction. Therefore, in the selection of influencing factors, 8 kinds of data with the highest critical weight are selected as the input of the model. And the improved neural network with 3 layers and 250 rounds of period is used for training.

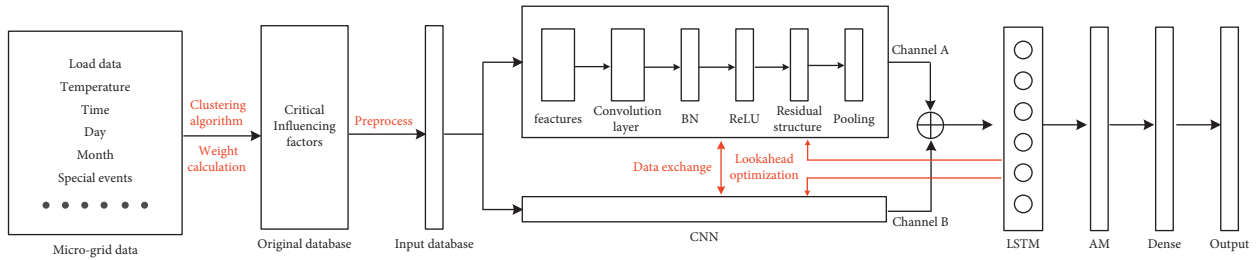


FIGURE 4: Model structure.

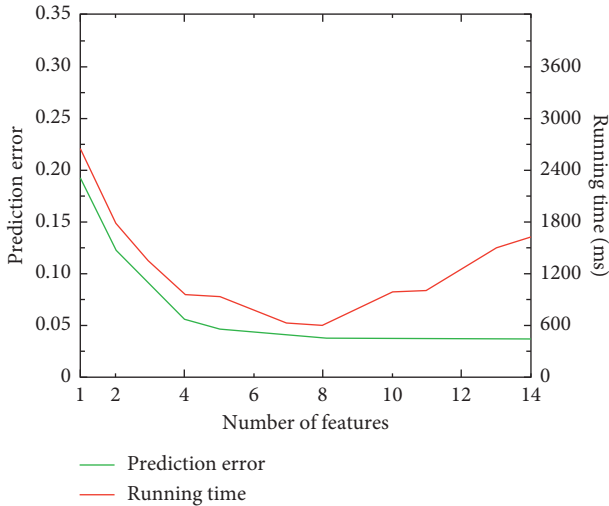


FIGURE 5: Prediction results and running time under different number of features.

5.3. *Predictive Performance Test.* In order to verify the prediction accuracy of the proposed algorithm, the prediction performance is tested after the neural network training is fully completed. Eight groups of data under different conditions are, respectively, selected for load prediction, and prediction error is calculated. Then, the load prediction is completed using the traditional LSTM load prediction algorithm for comparison experiments. The power error comparison can be obtained as shown in Table 2. It can be seen that the basic error of the improved LSTM algorithm is within 4% and about 1 kW. Experiments show that this algorithm can complete the task of microgrid load forecasting. Compared with traditional LSTM, the accuracy gets greatly improved.

100 prediction experiments on power are performed using the same data, and the average error of prediction is calculated separately for different scenarios. The results are shown in Figure 6, which shows the daily comparison of load power forecasts. It can be seen that the load shows a certain daily cycle. It is lower at night and higher during the day, and there is a certain peak-to-valley characteristic. The experiment proves that the algorithm can accomplish the task of microgrid load forecasting. Compared with the traditional LSTM, the prediction accuracy has been improved.

TABLE 2: Power prediction error.

No.	Improved LSTM		Traditional LSTM	
	$e_{\text{mape}}$	$e_{\text{rmse}}/W$	$e_{\text{mape}}$	$e_{\text{rmse}}/W$
01	0.0332	1016.68	0.0847	3520.34
02	0.0304	1049.83	0.0796	2788.22
03	0.0295	907.07	0.0782	2498.53
04	0.0351	1071.98	0.0983	3748.87
05	0.0347	1057.44	0.0924	3951.08
06	0.0286	894.37	0.0617	2302.14
07	0.0362	1143.23	0.0925	3980.83
08	0.0317	981.71	0.0749	2718.10

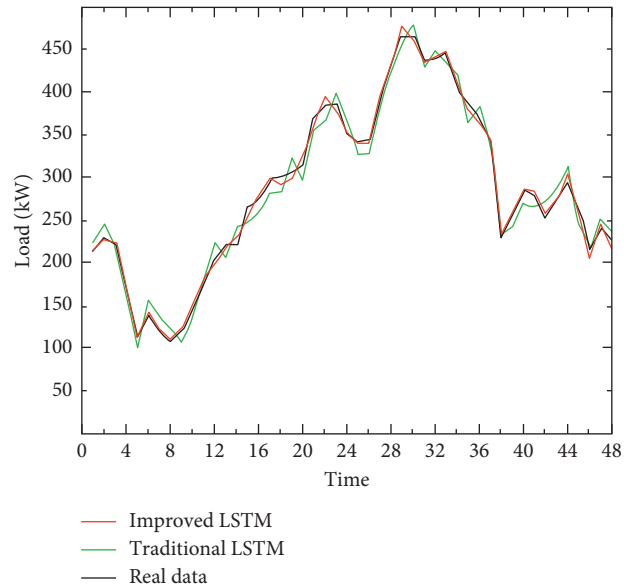


FIGURE 6: Load power prediction.

5.4. *Optimization Performance Test.* The optimization speed of load forecasting model is the key index to evaluate the forecasting performance. The model uses less iteration times and achieves higher test accuracy, which is the necessary condition for excellent prediction model. The improved LSTM method is compared with the traditional LSTM method and other load forecasting algorithms. The following test performance is obtained in Figure 7.

It can be seen from the figure that the improved LSTM method converges faster than the traditional method. It also has certain advantages over other methods: the proposed

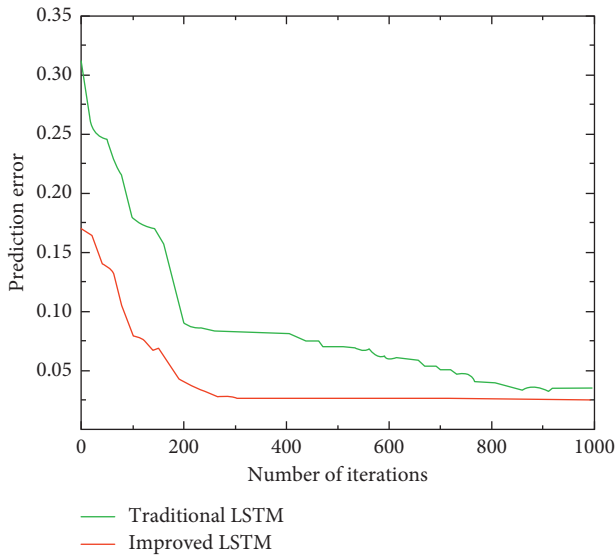


FIGURE 7: Iterative optimization evaluation.

TABLE 3: Error comparison.

Program	$e_{\text{mape}}$	$e_{\text{rmse}}/w$
Improved LSTM	0.0332	1016.68
CNN-LSTM	0.0837	2908.43
Bi-LSTM	0.0673	2042.15
GAN-LSTM	0.0542	1743.18

method can achieve the best performance after about 250 iterations, and the prediction error is generally less than 4%.

**5.5. Comparison with Other Schemes.** In order to further compare the forecasting performance of the proposed scheme, this paper compares the proposed scheme with the traditional load forecasting methods, such as CNN-LSTM, Bi-LSTM, and GAN-LSTM. Among them, CNN-LSTM is a data feature proposed after the power signal is recognized by CNN as the input of LSTM, and the power prediction is completed after important feature capture. Three convolutional layers are used, and the LSTM neurons are 64, where Bi-LSTM is a bidirectional LSTM considering feature selection. Power accurate prediction is achieved by sequential and inverse order information exchange, the number of network layers is 3, and the training period is 250 rounds. The GAN-LSTM is to expand the number of samples and improve the training quality by data generation. Two convolutional layers with  $3 \times 3$  kernels and 64 feature maps are used, followed by batch-normalization layers and ReLU as the activation function. The comparison results obtained are shown in Table 3.

It can be seen that, compared with other load forecasting algorithms, the improved LSTM algorithm has certain advantages in accuracy and convergence speed. Experiments show that this scheme is feasible in microgrid load forecasting.

## 6. Conclusion

An improved LSTM algorithm for load power prediction of microgrid is proposed in this paper. Firstly, the analysis of influencing factors and data processing are completed, and high-quality model input database is established. Secondly, the traditional LSTM load forecasting algorithm is optimized in three aspects: multi-layer convolution neural network optimization, lookahead optimizer, and attention mechanism optimization to improve prediction accuracy, generalization ability, and convergence speed. At last, the validity and reliability of the algorithm are proved by simulation experiments. The main contributions of this paper can be summarized as follows:

- (1) Comprehensive analysis on load influencing factors is carried out, and the key influencing factors are found out by cluster analysis and characteristic weight calculation. Then, the collected data are preprocessed to ensure the quality of input data.
- (2) The multilayer convolution channel is used to extract the feature, and the deep mining of the feature is completed by means of data exchange and cross-training between channels of different layers. Lookahead optimizer is also adopted to optimize the forward-looking weights to improve the fitting speed and generalization ability of the model and ensure robustness. The output is weighted by AM attention mechanism to further ensure the accuracy of prediction.
- (3) Taking a microgrid model in Zhejiang Province as an example, the simulation experiment is carried out. Experiments show that the 8 key factors can be used as model input to guarantee the prediction speed and accuracy. Compared with the traditional LSTM algorithm and other algorithms, the proposed scheme has higher prediction accuracy and convergence speed.

## Data Availability

The data used to support the findings of this study were supplied by YUQING ZHENG under license and so cannot be made freely available. Requests for access to these data should be made to YUQING ZHENG (zhengyq@cceceexpo.org)

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Project of Ningbo Polytechnic in 2022 "Basic research on renewable energy power consumption based on big data" (NZ22003).

## References

- [1] S. A. Vargas, G. R. T. Esteves, P. M. Maçaira, B. Q. Bastos, F. L. Cyrino Oliveira, and R. C. Souza, "Wind power generation: a review and a research agenda," *Journal of Cleaner Production*, vol. 218, pp. 850–870, 2019.
- [2] Z. Zonghe Gao, J. Jian Geng, K. Kaifeng Zhang et al., "Wind power dispatch supporting technologies and its implementation," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1684–1691, 2013.
- [3] C. Zhang, *Research on some issues of short-term wind speed forecasting for wind farms*, Ph.D. dissertation, 2017.
- [4] L. Chun, F. Gaofeng, and W. Weisheng, "Combined prediction model of wind farm output power," *Power System Technology*, vol. 33, no. 13, pp. 74–79, 2009.
- [5] W. Xie, P. Zhang, R. Chen, and Z. Zhou, "A nonparametric bayesian framework for short-term wind power probabilistic forecast," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 371–379, 2019.
- [6] K. O. N. G. Xiangyu, Z. H. E. N. G. Feng, and E. Zhijun, "Short-term load forecasting based on deep belief network[J]," *Automation of Electric Power Systems*, vol. 42, no. 5, pp. 133–139, 2018.
- [7] L. U. Jixiang, Z. H. A. N. G. Qipei, and Y. A. N. G. Zhihong, "Short-term load forecasting method based on CNN-LSTM hybrid neural network mode," *Automation of Electric Power Systems*, vol. 43, no. 8, pp. 131–137, 2019.
- [8] J. Lago, B. De Schutter and B. D. Schutter, Forecasting spot electricity prices: deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386–405, 2018.
- [9] H. Shubhi and G. Volker, "One dimensional convolutional neural network architectures for wind prediction," *Energy Conversion and Management*, vol. 195, pp. 70–75, 2019.
- [10] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction intervals for short-term wind farm power generation forecasts," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 3, pp. 602–610, 2013.
- [11] J. S. Hossein, C. L. S. Petronio, and G. G. Frederico, "Short-term forecasting by using a combined method of convolutional neural networks and fuzzy time series," *Energy*, vol. 175, pp. 365–377, 2019.
- [12] D. A. G. Vieira, B. E. Silva, and T. V. Menezes, "Large scale spatial electric load forecasting framework based on spatial convolution[J]," *International Journal of Electrical Power & Energy Systems*, vol. 117, Article ID 105582, 2019.
- [13] F. Shahid, A. Khan, A. Zameer, J. Arshad, and K. Safdar, "Wind power prediction using a three stage genetic ensemble and auxiliary predictor," *Applied Soft Computing*, vol. 90, Article ID 106151, 2020.
- [14] Z. Haiyang and A. Kusiak, "Prediction of wind farm power ramp rates: a data-mining approach," *Journal of Solar Energy Engineering*, vol. 3, pp. 376–385, 2009.
- [15] H. Zareipour, D. Huang, and W. Rosehart, "Wind power ramp events classification and forecasting : a data mining approach," in *Proceedings of the 2011 IEEE Power & Energy Society General Meeting San Diego:IEEE*, Detroit, MI, USA, July, 2011.
- [16] M. Cui, J. Zhang, and A. R. Florita, "An optimized swinging door algorithm for identifying wind ramping events," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 150–162, 2015.
- [17] C. Mingjian, S. Yunzhang, and K. Deping, "Prediction of wind power ramp based on sparse decomposition of atoms and BP neural network," *Automation of Electric Power Systems*, vol. 38, no. 12, pp. 6–11, 2014.
- [18] T. Ouyang, X. Zha, and Q. Liang, "Wind power ramp events forecast method based on similarity correction," *Proceeding of the CSEE*, vol. 37, no. 2, pp. 572–580, 2017.
- [19] L. Han, R. Zhang, X. Wang, A. Bao, and H. Jing, "Multi-step wind power forecast based on VMD-LSTM," *IET Renewable Power Generation*, vol. 13, no. 10, pp. 1690–1700, 2019.