Journal of Electrotechnology, Electrical Engineering and Management (2023) DOI: 10.23977/jeeem.2023.060303
Clausius Scientific Press, Canada ISSN 2560-6697 Vol. 6 Num. 3

# *Short-term load prediction based on Pearson-optimized CNN-LSTM hybrid neural network*

**Xu Chen[1], Xinying Liu[2,*]**

[1]*Department of Electrical and Mechanical Engineering, Xinjiang Agricultural University, Urumqi, 830000, China*

[2]*Department of Electrical and Control Engineering, Liaoning University of Engineering and Technology, Huludao, 125105, China*

[*]*Corresponding author: lxy66613998452017@163.com*

***Keywords:*** Load forecasting, Power systems, CNN, LSTM, Pearson correlation analysis

***Abstract:*** Under the demand of new power system construction, it is important to establish a solid and reliable power grid structure with stable operation by accelerating the construction of a "double high" strategy with the goal of "double carbon". Bus load can reflect the operation of the power grid, so bus load forecasting is important to maintain the safety and stability of the power system. To solve the problems of low accuracy and inefficiency of existing load forecasting methods for power systems, this paper adopts a combined CNN-LSTM load forecasting model with Pearson optimization, which is machine learning combined with deep learning. Firstly, Pearson correlation analysis is used for data processing to extract the main features of load data. Then three neural networks, CNN, LSTM, and CNN-LSTM, are used for training and load prediction, respectively. The experimental results show that the load prediction accuracy of the hybrid CNN-LSTM neural network prediction model based on Pearson optimization is higher than that of CNN and LSTM alone and matches with the actual value, which is a load prediction method with higher accuracy.

## 1. Introduction

With the introduction of the "double carbon" target and the ongoing promotion of the "double high" strategy [1], modern power systems are developing in the direction of intelligence, flexibility, and high interactivity, and the role of short-term load forecasting is becoming more and more important in the future power grid planning and operation [2]. Short-term load forecasting is an important part of power system scheduling and operation and planning, and its accuracy is of great significance for the economic, safe, and stable operation of power systems, which directly affects the balance of power supply and demand, power market transactions and the rational use of power resources. It also helps power technicians to make equipment maintenance plans to timely identify safety hazards in the power system, in addition to being of great significance to the promotion of national electricity sales side reform work [3]. Therefore, improving the accuracy of load forecasting has been a hot issue of research in the electric power field. Especially in the current environment of the new power market, to adapt to the more flexible and changing rhythm of the times, many scholars have done a lot of work and achieved many results in many aspects such as the research of method theory and the proposal and improvement of forecasting models.

To address the problem of low accuracy of short-term load prediction, Zhiwang Zhang used a back propagation neural network based on genetic algorithm optimization to avoid the algorithm falling into local extremes, which makes the prediction model have better accuracy and superiority [4]. At the same time, Wang Zhuyao and Gao Yaoyao proposed a GRU neural network load prediction model based on quantum weighted optimization and Adam algorithm optimization, respectively, to improve the prediction effect of the model by improving the structure of the neural network, which has better accuracy and stability for load prediction [5][6]. Based on this, Chunjie Yin established the LSTM neural network load forecasting model optimized by the Adam algorithm, which solves the problem that there are more studies for large grid load forecasting but fewer studies for microgrid load, and highlights the accuracy and rationality of this forecasting model in comparison with other forecasting methods [7]. Since the drawback of the load forecasting method of artificial neural networks is the reliance on the quality of the training samples, Jingfu Gan used wavelet threshold denoising and cluster analysis to process the samples, eliminate spikes and noisy data and select suitable training samples. This method achieves optimal processing of the data and more accurate prediction in imitation of real-time [8].

Traditional forecasting methods based on statistical learning methods are not effective due to the nonlinear and stochastic factors of load, so researchers have focused on the popular artificial intelligence fields in recent years, such as artificial neural networks, machine learning, and deep learning [9]. The artificial neural network has a strong autonomous learning capability, it can self-adjust and parameter optimize during the operation, and it is a strong nonlinear model, so it is suitable for application in load prediction.

In this paper, we propose a short-term power system load combination forecasting model based on Pearson correlation analysis combined with CNN and LSTM to address the current situation that load forecasting is stochastic and affected by various factors such as weather, temperature, and electricity price. First, the data are tested for outliers and the outliers are removed to fill in the missing values. Then the data are normalized using the principle of Pearson correlation analysis, and then the correlation between the data is judged. The neural network is made to read, train and learn the processed data, and the results of load prediction of the three neural network models are obtained. The experimental results show that the predicted values of the combined load prediction model proposed in this paper match the real values and have high prediction accuracy.

## 2. Theoretical basis

### 2.1 Pearson correlation analysis

Pearson correlation analysis is a statistical method used to measure the correlation between two continuous variables and is commonly used in data analysis and statistical research. The correlation coefficient is between -1 and 1, with a negative number representing a negative correlation and a positive number representing a positive correlation, and the higher the absolute value of the coefficient, the stronger the correlation. Its calculation formula is as follows:

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{1}$$

## 2.2 CNN neural network

CNN is a deep learning model widely used in the field of computer vision, which contains a convolutional layer, a pooling layer, and a fully-connected layer. The pooling layer is connected behind the convolutional layer, which can effectively reduce the number of parameters, improve the computational speed of the model and prevent overfitting; the fully-connected layer is the last layer, which can complete the classification task based on the extracted features. In addition, because of its excellent ability to extract features of continuous historical time-series data, CNN is often used in load prediction models [10], where load prediction tasks can be accomplished by extracting time-series features of load data. The structure diagram of CNN is shown in Figure.1:

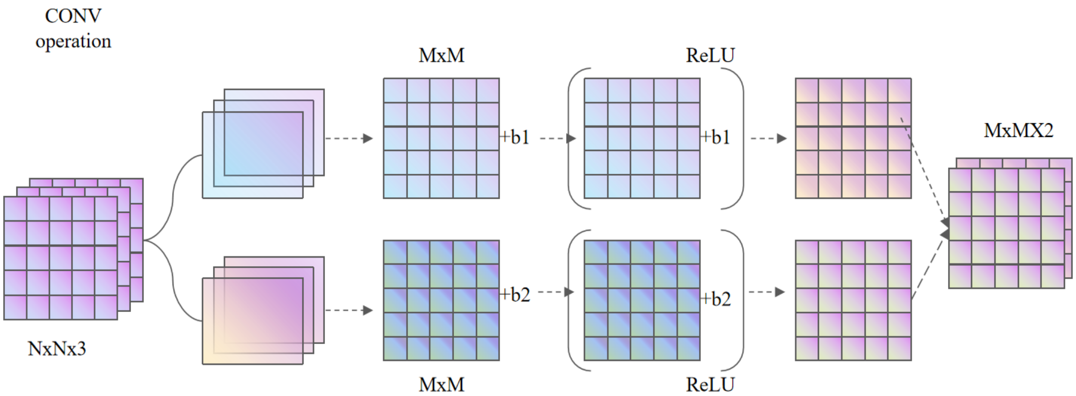


Figure 1 CNN structure schematic

## 2.3 LSTM neural network

LSTM neural network is a deep learning model for sequential data, which is commonly used in natural language processing, speech recognition, load prediction, etc. LSTM is improved from RNN and compared with traditional RNN, LSTM effectively solves the problem of gradient disappearance and gradient explosion, and therefore has better memory capability. LSTM controls the information flow by introducing a gating mechanism, which allows the network can retain the important information and ignore the irrelevant information. Its structure is shown in Figure.2:

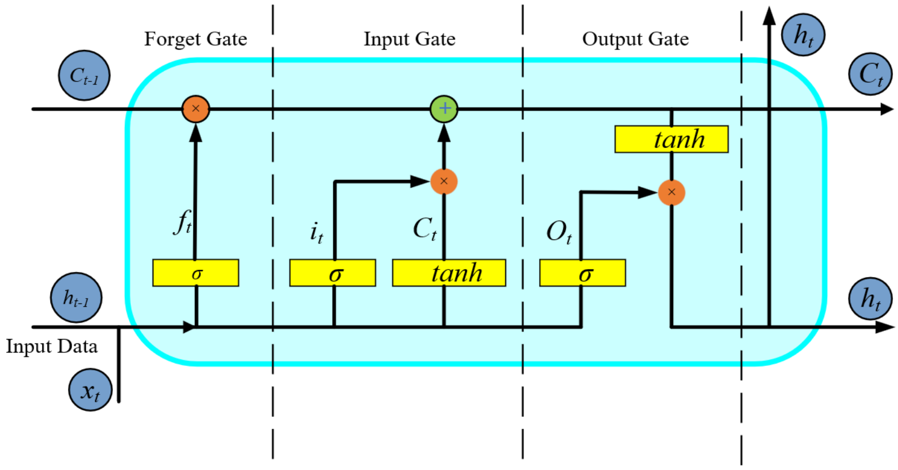


Figure 2 LSTM structure schematic

LSTM uses a gated output approach, i.e., three gates (input gate, output gate, forgetting gate) and two states (long time, short time). Its core is Cell State, which refers to the state of the Cell used for information propagation from $C_{t-1}$ to $C_t$. Memory Cell accepts two inputs, i.e., the output value $h_{t-1}$ of the previous moment and the input value $X_t$ of the current moment, from which the two parameters first enter the forgetting gate to get the information $f_t$ (i.e., information with less weight) that is decided to be discarded, and then enter the input gate to get the information it (i.e., information with less weight) that is decided to be the updated information it (i.e., the information with greater weight compared with the previous Cell) and the Cell state $\tilde{C}_t$ at the current moment (the candidate vector, which can be understood as an intermediate variable, stores the current Cell State information), and finally the output values (i.e., $f_t, i_t, \tilde{C}_t$) from these two gates (forgetting gate, output gate) are combined (the activation value ft of the previous Cell State $C_{t-1}$ multiples the information to be forgotten is superimposed with the activation value it of the current moment Cell State $\tilde{C}_t$ multiples the information to be remembered) to obtain the information of long-time $C_t$ and short-time $h_t$, respectively, and finally the storage operation and the input to the next neuron. The core expression of LSTM is as follows:

$$\begin{cases} f_t = \sigma\ (W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma\ (W_i \cdot [h_{t-1}, x_t] + b_i) \\ \widetilde{C_t} = tanh\ (W_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C_t} \\ o_t = \sigma\ (W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t \cdot tanh(C_t) \end{cases} \tag{2}$$

Where: $f$ is the activation function of the gate; $h$ is the activation function of the Cell output; $W$ refers to the parametric training matrix; $(z)$ is the Sigmoid activation function; $tanh(z)$ is the hyperbolic tangent activation function. The expressions of the Sigmoid activation function and hyperbolic tangent activation function are as follows:

$$\sigma(z) = \frac{1}{1+e^{-z}} = \frac{1+tanh(z/2)}{2} \tag{3}$$

$$\sigma^{'}(z) = \sigma(z)[1-\sigma(z)] \tag{4}$$

$$tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{5}$$

$$tanh^{'}(z) = 1 - tanh^2(z) \tag{6}$$

## 2.4 CNN-LSTM hybrid neural network model

The CNN-LSTM hybrid neural network is connected by connecting two kinds of neural networks, CNN and LSTM, in series. CNN is used to extract the features of the input load data, enhance the sequence features of the load and improve the prediction accuracy. In this paper, two convolutional layers are used, the first convolutional layer includes 32 3×1 convolutional kernels and relu activation layers, and the second convolutional layer includes 64 3×1 convolutional kernels and relu activation layers. An average pooling layer is added to compress the time series of the load features, and a fully connected layer is added to construct the SE attention mechanism. The output of the CNN is fed to the LSTM layer for load prediction, and the prediction results of the load are obtained. The structure diagram of the CNN-LSTM neural network is shown in Figure 3:

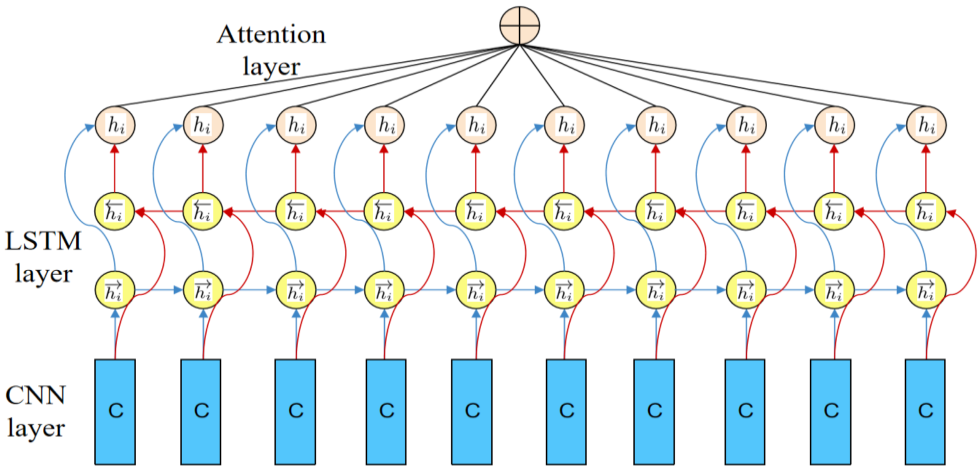


Figure 3 CNN-LSTM schematic diagram

## 2.5 Model evaluation metrics

To evaluate the accuracy of the established model more precisely, six indicators, R2 (R-squared), mean absolute error (MAE), mean deviation (MBE), mean absolute error (MAPE), and root mean square error (RMSE), are used in this paper to evaluate and analyze the prediction results of the model. The six indicators are calculated as follows:

$$R-squared = 1-\frac{\sum_{i-1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i-1}^{n}(\overline{y}_i - y_i)^2} \tag{7}$$

$$MAE = \frac{1}{n}\sum_{i-1}^{n}\left|\hat{y}_i - y_i\right| \tag{8}$$

$$MBE = \frac{1}{n}\sum_{i-1}^{n}(y_i - y_i) \tag{9}$$

$$MAPE = \frac{1}{n}\sum_{i-1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\% \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}(\hat{y}_i - y_i)^2} \tag{11}$$

R-squared is used to evaluate the fit of the model to the data, which represents the proportion of variance explained by the model to the total variance. MAE measures the mean absolute error between the predicted and true values, and the smaller the value, the more accurate the prediction is. RMSE is used to measure the root mean square error, again the smaller the value the higher the prediction accuracy.

## 3. Load forecasting model case study

This paper combines data analysis methods with machine learning and deep learning to propose a short-term bus load prediction model based on a hybrid CNN-LSTM neural network with Pearson optimization. First, the influencing factors of bus load are analyzed and selected. Considering the factors that bus load prediction is influenced by the sample data, the sample data are tested for outliers, the outliers are removed and the missing values are filled. Then the normalization of the data is done and the correlation between various influencing factors and the electric load is studied using Pearson correlation analysis. Then the processed data are read, and the CNN-LSTM neural network structure is defined and trained to learn and perform load prediction. Finally, the curves of actual and predicted values are obtained, and the correlation errors of predicted values are analyzed and compared with the results expected by CNN and LSTM alone.

### 3.1 Analysis of factors influencing busbar load

In practice, many factors affect the load of the power system. Meteorological factors, including temperature, humidity, and rainfall, can directly affect load fluctuations. For example, the use of air conditioners in summer can increase the load on the system, and thunderstorms can cause sudden changes in the load in a certain area. In general, the load on holidays is significantly lower than the load on normal working days, and the annual load curve has a similar trend in terms of longitudinal comparison for the same holiday, which can provide a corresponding basis for the load forecast on holidays. In addition, due to some irregularity factors, the load fluctuation itself has a certain randomness, which is a typical feature of this data. The objective function depends on the design variables, and the constraints are determined according to the factors affecting the load to determine the variables and the objective function further. The load-influencing factors selected in this paper are shown in Table.1:

Table 1 Load influencing factors

| Influencing Factors | Unit | Correlation (from strong to weak) |
|---|---|---|
| Time | h | 1 |
| Electricity price | KWh | 2 |
| Dry bulb temperature | ℃ | 3 |
| Wet bulb temperature | ℃ | 4 |
| Dew point temperature | ℃ | 5 |
| Humidity | RH | 6 |

### 3.2 Data processing analysis

The historical load data used for the simulations in this paper were obtained from the National Electricity City of Singapore [11], for the period from January 1, 2006, to January 1, 2011, with a data collection time of one hour. Firstly, methods such as box-line plots were used to test the data for outliers and eliminate the abnormal data, and the mean, plural, and median were used to replace the missing values. Then the normalization of the data was performed with the following formula:

$$x_{scale} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{12}$$

Where $x_{scale}$ is the normalized value, $x$ is the original data value, and $x_{min}$ and $x_{max}$ are the minimum and maximum values in the original data, respectively. Finally, Pearson correlation analysis is used to analyze the correlation between the elements. The thermodynamic diagram between the elements is shown in Figure.4:

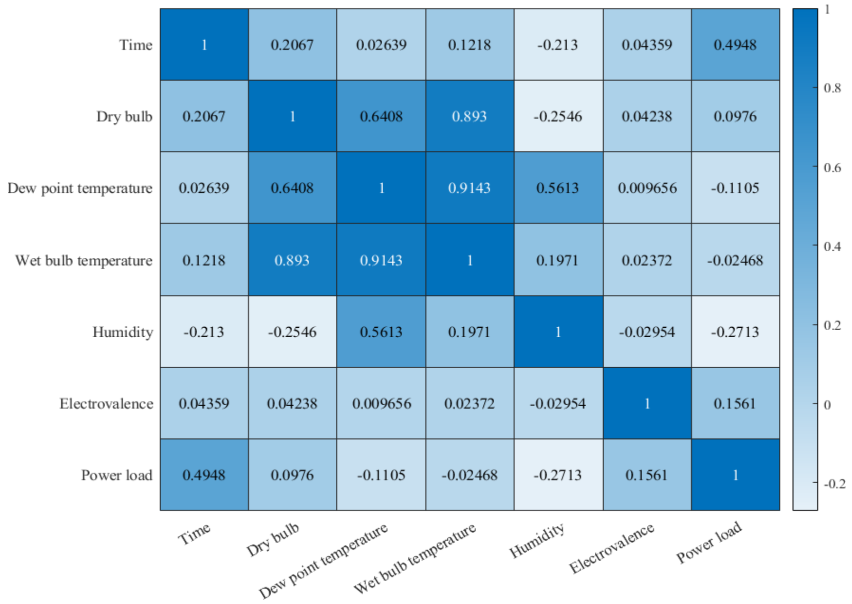


Figure 4 Heat map of elements

### 3.3 Load training and prediction using neural network models

CNN-LSTM hybrid neural network is chosen for load forecasting because CNN-LSTM has the advantages of both CNN and LSTM neural networks, which can effectively learn spatial and temporal features in time series data while avoiding the problem of gradient disappearance. CNN-LSTM is widely used in the field of load forecasting because of its superior performance, which can accurately and effectively predict future load demand.

First, the already processed Excel data, including electric, thermal, and cooling loads, are read. Next, the data set is divided, and the first 8000 data are classified as the training set, and the last 360 data are used as the test set. Three neural network models, CNN, LSTM, and CNN-LSTM, are defined in MATLAB to train the dataset and predict the test set results for each of the three loads. After training, the model output is shown in Figures 5, 6, and 7.

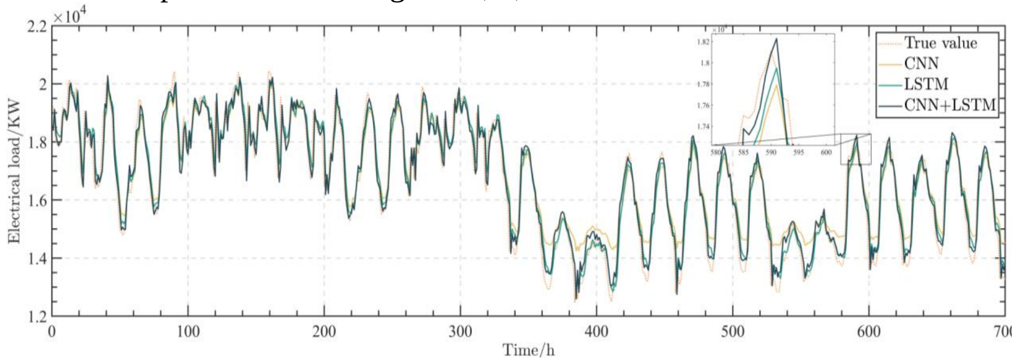


Figure 5 Comparison of training results of networks with different electrical loads

Electric load is the amount of electricity consumed per unit of time in the power system, which can reflect important information such as load situation and load change pattern in the power system. As can be seen from Fig. 5, compared with the other two neural network models, the prediction effect of CNN-LSTM thermal load is closer to the real value and overlaps with the real value curve. The CNN, on the other hand, deviates significantly from the true value curve at the load peaks and valleys and fails to achieve a good prediction effect; the LSTM has a lower prediction accuracy than the

CNN-LSTM.

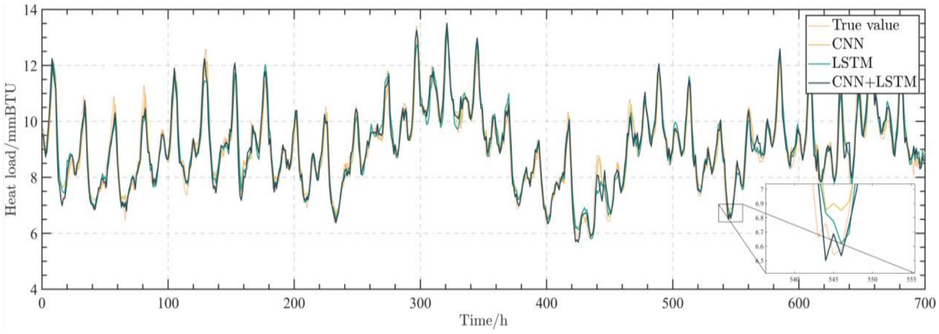


Figure 6 Comparison of training results of networks with different thermal loads

The heat load is the required heating capacity in a certain period, which is an indicator of the heating demand and is usually related to the temperature, the number of residents, and the building properties. As can be seen from Fig. 6, the predicted heat load values of CNN and LSTM are significantly lower than the true values at the peaks and valleys, so the demand for power supply from the power system is not reached. The shape of the prediction curve of CNN-LSTM is closer to the true value curve, and the prediction accuracy is higher.

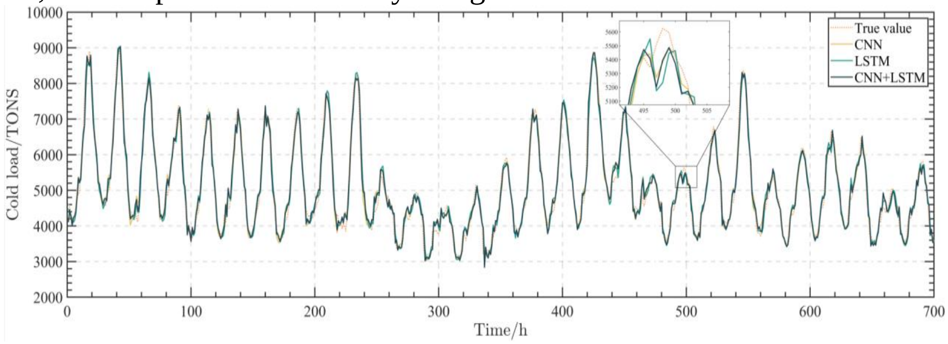


Figure 7 Comparison of training results of different networks with cold load

The cold load is the cooling capacity that the system needs to provide in a certain period, which is an indicator of the cooling demand. As can be seen from Figure 7, compared with the other two models, the CNN-LSTM has a better tracking ability and the predicted value is closer to the true value when the load undergoes more significant fluctuation changes. Based on the prediction results, three neural network model load prediction error indicators can be plotted as shown in Fig.8.

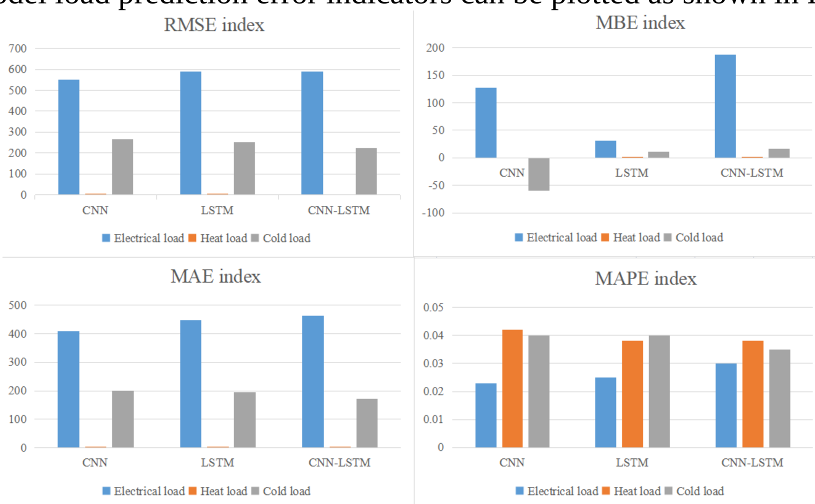


Figure 8 Error indicators

## 4. Conclusion

(1) For the short-term load prediction problem, this paper proposes a CNN-LSTM model based on Pearson optimization, analyzes the prediction principle as well as the advantages of the model, and uses MATLAB software to train and predict the electric load, thermal load, and cold load, respectively. The final results of the case study are found to be consistent with the theoretical analysis.

(2) Based on the analysis of the prediction results and error indicators of the three networks, it can be concluded that the hybrid CNN-LSTM neural network short-term load prediction model based on Pearson optimization proposed in this paper has high accuracy, well combines the advantages of the two networks, and the prediction effect is better than CNN and LSTM, which has better practical value.

## References


*[1] Zhu Haohao, Zhu Jizhong, Li Shenglin, et al. Fractional-order power system current calculation and voltage analysis [J]. Journal of Electrical Machinery and Control, 2022, 26(04):38-46.*

*[2] Weicong Kong, Zhao Yang Dong, Youwei Jia, et al. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. [J]. IEEE Trans. Smart Grid,2019,10(1).*

*[3] Sheng Hongying, Zhao Weiguo, Chen Yang, et al. Short-term load forecasting for power systems based on deep data mining[J]. Journal of Jilin University (Information Science Edition), 2023, 41(01):131-137.*

*[4] Zhang Zhiwang. Application of BP network optimized by genetic algorithm in power system load forecasting [J]. Electronic World, 2021(17): 168-169.*

*[5] Wang Zhuyao, Zhang Zhisheng. Short-term load forecasting for power systems based on quantum-weighted GRU neural networks [J]. Journal of Power Systems and Automation, 2022, 34(01):1-7.*

*[6] Gao Ao, Li Guoyu, Password Aoyang,et al. Short-term load forecasting model based on optimized GRU neural network with Adam algorithm [J]. Electronic Design Engineering, 2022, 30(9): 180-183, 188.*

*[7] Yin Chunjie, Xiao Dafang, Li Pengfei, Zhao Qin. Short-term load forecasting for regional microgrid based on LSTM neural network [J]. Computers and Modernization,2022(04):7-11.*

*[8] Gan Jingfu, Yan Kun, Ma Minghan, He Pengkang, Tian Xincheng. Research on short-term load prediction by artificial neural network based on improved clustering algorithm [J]. New Technology of Electric Power,2022,41(09):40-46.*

*[9] Meng S, Shao JF. A review of deep learning-based electric load forecasting methods [J]. TV Technology, 2022, 46(01): 44-51.*

*[10] Fan Jiangchuan, Yu Haozheng, Liu Huiting, et al. Short-term electric load forecasting based on the multi-branch gated residual convolutional neural network [J]. China Electric Power,2022,55(11):155-162+174.*

*[11] Shang C, Gao J, Liu H, et al. Short-term Load Forecasting Based on PSO-KFCM Daily Load Curve Clustering and CNN-LSTM Model[J]. IEEE Access, 2021, PP(99):1-1.*