ISSN: 1813-4890

Short-term power load prediction based on TCN-LSTM hybrid model

Yu Sun*, Zhen Tao
Anhui University of Finance and Economics, Bengbu 233000, China
*Corresponding author: Yu Sun (Email:3164832583@qq.com)

Abstract

Power load forecasting is a key link to ensure the safe, stable and economic operation of the power system. However, the non-stationarity and multivariate coupling characteristics of load data lead to the challenge of insufficient accuracy in traditional forecasting methods. To this end, this study proposes a hybrid TCN-LSTM model based on temporal convolutional networks (TCN) and long-short-term memory networks (LSTM), aiming at fusing local temporal feature extraction with long-term dependency modeling capabilities. The TCN model efficiently captures the short-term fluctuation patterns of load data through causal convolution and extended convolution, while the LSTM model further mines the time-series global trend features. The experiments use power load data with 15-minute sampling intervals from the 2016 Mathematical Modeling Competition for Electricians to validate the model performance by dividing the training set with the test set. The TCN-LSTM model has a mean absolute percentage error (MAPE) of 1.17% and a coefficient of determination (R^2) of 0.9950, which is significantly better than the other comparative models, TCN, LSTM, and CNN-LSTM models. The model provides a high-precision forecasting tool for power system scheduling and energy planning, helping to improve the efficiency of renewable energy utilization and carbon neutrality targets.

Keywords

Power Load Forecasting; LSTM Model; TCN Model; Hybrid Model; Deep Learning.

1. Introduction

Power load forecasting plays a pivotal role in smart grids and energy management systems, directly influencing the safety, economy and reliability of the power system [1]. Precise load forecasting not only facilitates power dispatching and reduces energy waste, but also enhances the utilization efficiency of renewable energy and promotes the realization of the carbon neutrality goal.

Traditional load forecasting methods mainly rely on time series analysis and statistical modeling. However, these methods are difficult to effectively capture the complex nonlinear load characteristics, especially when dealing with multiple variables such as weather, economic factors, and social behaviors, the prediction accuracy is limited [2].

In recent years, load forecasting methods based on machine learning and deep learning have gradually emerged. A feedback computing mechanism was designed by establishing a BiLSTM model, which solved the coupling problem of the time series information before and after the power load [3]. As the complexity, nonlinearity and non-stationarity of load data are increasingly prominent, a single model is often insufficient to comprehensively depict the changing patterns. Therefore, hybrid models have gradually gained widespread attention in power load forecasting. The hyperparameters of LSTM network are fine-tuned by using PSO algorithm to enhance the prediction performance of the model [4]. To handle complex nonlinear data, deep kernel learning and multi-kernel information bidirectional boosting are combined to form nonlinear composite kernels [2]. These methods have demonstrated outstanding performance in both short-term and medium-to-long-term load forecasting. Particularly, by integrating big data and cloud computing technologies, the power load forecasting system has been further optimized, enabling more efficient power dispatching and management [5].

The TCN algorithm has achieved good results in numerous prediction fields [6][7]. LSTM model is an improved variant of RNN model, specifically designed to handle and predict the long-term dependencies in sequential data, and has been widely applied in the field of power load forecasting [8][9][10]. The prediction error has been significantly reduced, which has remarkably enhanced the accuracy of power load forecasting. Given the outstanding performance of LSTM and TCN in the power load forecasting task, this paper combines the two to achieve high-precision forecasting of ultra-short-term power load. Due to the short time interval of ultra-short-term prediction, the influence of meteorological factors is relatively small [1], thus meteorological factors are not taken into consideration in this paper. The final experimental results demonstrate that the proposed hybrid model outperforms the single model in terms of prediction accuracy, thereby enhancing the reliability of power load forecasting.

2. Methodology

2.1. Temporal Dependency Modeling with TCN Components

The TCN layer captures temporal dependencies and recognizes patterns in time series data through its internal structural features. It combines causal and extended convolution for temporal dependency learning. The mathematical expression for causal convolution is given in Equation (1), where y_t is the output of time step t, W is the convolution kernel weight, and in this study the gradient vanishing problem is solved using the ReLU activation function, and b is the bias term. In order to expand the sensory field, TCN uses extended convolution. The specific calculation formula is shown by Equation (2), where k is the convolution kernel size, d is the expansion factor, and $x_{t-d\cdot i}$ denotes the input value sampled at interval d.

$$y_t=f(W\cdot x_{t-k:t}+b). \tag{1}$$

$$yt=\sum W_i\cdot x_{t-d\cdot i}. \tag{2}$$

$$H(t)=F(x(t))+x(t). \tag{3}$$

Extended convolution inserts spacers into the convolutional kernel, allowing neurons at each layer to perceive a larger range of inputs without increasing the number of parameters. In addition, TCN also integrates residual joining to optimize the gradient flow, which is shown by Eq. (3), where H(t) is the output after residual joining and x(t) is the input data. F(x(t)) is the output of convolution operation. This alleviates the gradient vanishing problem and improves the training efficiency of the deep network.

2.2. Long-term dependency capture using LSTM components

The emergence of recurrent neural networks (RNN) solves the problem that traditional neural networks have difficulty in effectively extracting features from large amounts of data when dealing with time series prediction. When dealing with long sequence programs, RNNs suffer from gradient vanishing or gradient mutation. As one of the variants of RNN structure, the birth of LSTM solves the inherent problems of RNN to some extent. LSTM was first proposed by Schmidhuber et al (1997). Compared to RNN, LSTM adds input gates (i_t), forgetting gates (f_t), output gates (o_t), and storage units (c_t).The core of LSTM is the information about the state of the storage units, which can be added or deleted under the control of the three gates. The overall structure of the LSTM neural network is shown in Fig. 1.

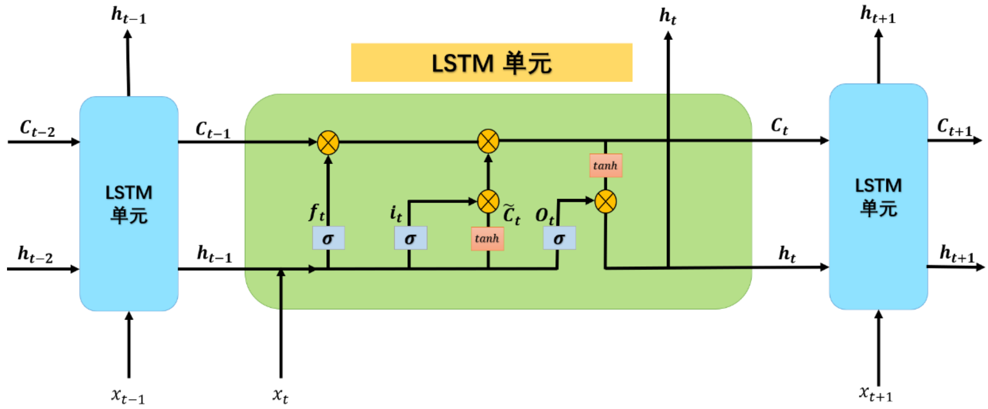

Fig. 1 LSTM internal structure diagram

The mechanism of the LSTM unit is described as follows:

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f). \tag{4}$$

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i). \tag{5}$$

$$o_t=\sigma(W_o\cdot[h_{t-1},x_t]+b_o). \tag{6}$$

$$c_t^{\sim}=\tanh(W_c\cdot[h_{t-1},x_t]+b_c). \tag{7}$$

$$c_t=(f_t*c_{t-1}+i_t*c_t^{\sim}). \tag{8}$$

$$h_t=o_t*\tanh(c_t). \tag{9}$$

In the first step, the LSTM model determines the information needed to forget the current cell state. This process is realized through the forgetting gate, whose main function is shown in equation (4),where W_f and b_f are the weight matrix and bias of the oblivious gate, respectively, x_t is the current input to the storage cell, h_{t-1} is the output, and c_t is the cell state. Second, the LSTM model determines what information needs to be added to the cell state, a process that is realized by inputting the gate and the state information of the candidate cell, $\tilde{c}_t$, whose main function is as in Eqs. (5)(7). Immediately after obtaining $\tilde{c}_t$ with i_t, the cell state needs to be updated. As shown in Equation (8), f_t*c_{t-1} indicates that the forgetting gate selectively forgets some of the state information of the old cell; $i_t*\tilde{c}_t$ indicates that the input gate adds some new information to the cell state. Finally, after updating the new cell state, the output depends not only on the new cell state but also on the output gate, as shown in Eqs. (6)(9).

2.3. Hybridization for short- and long-term forecasting

The optimized TCN component is responsible for extracting local temporal features and passing these features to the LSTM component to establish long-term dependencies before generating predictions through the fully connected layer.

$$\hat{y}(t+1)=f_{Dense}(h_t). \tag{10}$$

where h_t is the final hidden state of the LSTM at time t, and $\hat{y}(t+1)$ denotes the mapping of the output of the LSTM to the final predicted value through the fully connected layer.

3. Experimental Results and Discussion

3.1. Data sources

The experiment is based on the electrical loads from January 1, 2009 to January 10, 2015 provided by the 2016 Mathematical Modeling Competition for Electricians , and the sampling period of the data is 15 min. Data from January 1, 2009 to January 8, 2015 were used as model input data, and the training and test sets were divided at a scale of 0.8. In addition, these experiments were conducted in a python 3.11 environment running the Windows 11 operating system with a 64-bit AMD64 processor.

ISSN: 1813-4890

3.2. Performance measures

Four commonly used error detection methods, namely MAE, RMSE, MAPE, and R^2, are used in this study to compare the accuracy of experimental models.This study also uses these metrics as a basis for evaluating the predictive effectiveness of multiple prediction models. The formulas for the above four performance criteria are as follows:

$$MAE=1/n\cdot\sum|\ y_t-\hat{y_t}\ |. \tag{11}$$

$$RMSE=(1/n\cdot\sum(y_t-\ \hat{y_t}\ )^2)^{(1/2)}. \tag{12}$$

$$MAPE=1/n\cdot\sum(|\ y_t-\hat{y_t}\ |)/y_t. \tag{13}$$

$$R^2=1-\{1/n\cdot\sum(y_t-\hat{y_t}\ )^2\}/\{1/n\cdot\sum(y_t-\bar{y}_t)^2\}. \tag{14}$$

Where y_t is the actual power load data, $\hat{y_t}$ is the predicted power load data, and n is the number of power load data.

3.3. Analysis of results

The experiment is also compared with the prediction model of CNN-LSTM. The experiment shows the comparison between the hybrid model TCN-LSTM proposed in this paper and the benchmark model LSTM, TCN and CNN-LSTM . Table 1 lists the numerical results of the above predictive model performance metrics. Figure 2 visualizes the comparison between the true and predicted values of each model on January 8, 2015. From the images, it is found that the predicted curve of TCN-LSTM model is basically consistent with the trend of the actual curve. Figure 3 shows the RMSE and MAE values for each model. Based on the above data results, it is found that the TCN-LSTM model performs the best. Its RMSE value is 151.3114, which is lower than the other models.The predictive performance of TCN model and TCN-LSTM model are close to each other, and the predictive performance of LSTM model is at the bottom. For the R^2 value, the TCN-LSTM model predicts the best, with a model fit value of 0.9950.In conclusion, the TCN-LSTM model outperforms the other models and is able to achieve excellent predictions.

Table 1 Prediction performance for each model

Model	MAPE	RMSE	MAE	R2
CNN-LSTM	0.0206	280.7940	156.2568	0.9828
LSTM	0.0270	314.4159	216.6398	0.9756
TCN	0.0225	253.4598	157.8595	0.9859
TCN-LSTM	0.0117	151.3114	88.8020	0.9950

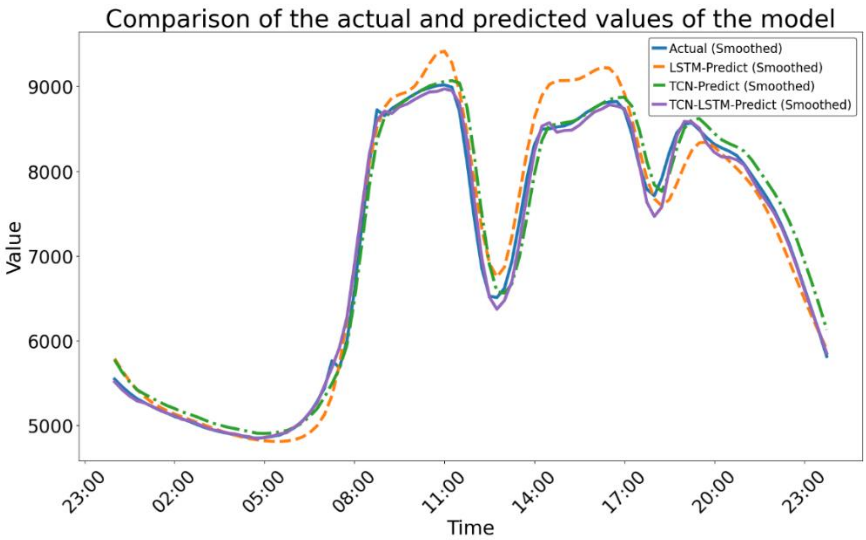

Figure 2 Comparison of the model prediction value with the true value

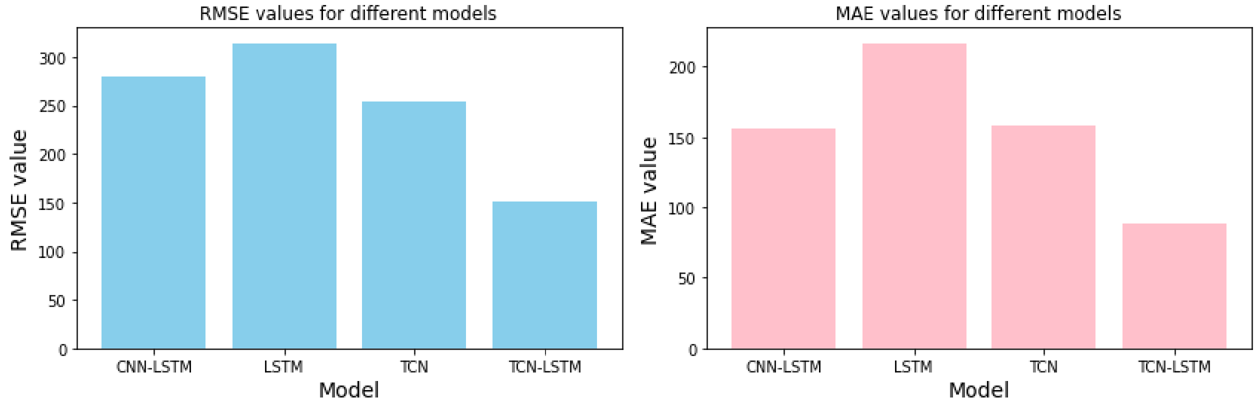

Figure 3 Comparison of RMSE and MAE values for different models

4. Conclusion

Electricity load data is characterized by non-stationarity, non-linearity and complexity. Research on power load forecasting has increased extensively in recent years. In order to obtain highly accurate and robust prediction results for PV power series, this study uses the 2016 Electrician's Cup mathematical modeling data as an example of the proposed hybrid combination model based on TCN and LSTM, in which the TCN model extracts the local features of the time series to capture the short-term patterns in the time series.The LSTM receives the features extracted by the TCN and efficiently captures the long-term dependencies for a better understanding of the overall trend of the time series. Two benchmark models, LSTM and TCN with CNN-LSTM, are used as comparison models in the experimental design, and the results show that the TCN-LSTM model proposed in this paper is better in prediction by comparing the performance metrics of model evaluation.

In the subsequent research work, it should focus on the load forecasting problem in specific segments. Expanding the data analysis method to multiple specific fields such as industry, agriculture, commerce, etc. can better combine the actual scenarios and dig out more valuable information, thus effectively improving the accuracy of the prediction. In future power load forecasting, the advantages of the big data platform should be fully utilized, and the massive data should be efficiently processed with the help of parallel artificial intelligence in order to speed up the data analysis, thus significantly improving the real-time and accuracy of forecasting.

Acknowledgements

The successful completion of this research and thesis was made possible by the support and assistance of many people. First of all, I would like to sincerely thank my supervisor, Associate Professor Du Juntao. From the selection of the research direction, the design of the model framework to the optimization of the experimental scheme, Mr. Du always gave me meticulous guidance with rigorous academic attitude and profound professionalism. His unique insights in the field of energy resources have provided important inspiration for innovative breakthroughs in this research. We would like to thank the Economic Statistics Program of Anhui University of Finance and Economics for providing high quality academic resources and research environment. The open academic atmosphere and perfect experimental platform of the college laid a solid foundation for the data processing and model validation of this paper. We also thank the organizing committee of the 2016 Mathematical Modeling Competition for Electrical Engineering for making the electric load data publicly available, which provided reliable data support for this study. In addition, I would like to thank my fellow laboratory students for their discussions and suggestions during the research process, and their valuable comments helped me to improve the experimental design and analysis of the results. Finally, I would like to express my deepest gratitude to my family, whose understanding and encouragement have been a lasting motivation for me to complete my studies and research. In the future, I will continue to uphold the motto of Anhui University of Finance and Economics, “Integrity and erudition, unity of knowledge and action”, to explore the field of economic statistics, and to contribute more to the research of smart grid and energy management.

Supported by the National Innovation and Entrepreneurship Training Program of Anhui University of Finance and Economics (202410378113).

Anhui University of Finance and Economics Undergraduate Research and Innovation Fund Project Grant(XSKY25053ZD)

References

[1] Liu, Jingrui, Zhiwen Hou, and Tianxiang Yin. "Short-term power load forecast using OOA bidirectional long short-term memory network with spectral attention for the frequency domain." Energy Reports 12 (2024): 4891-4908.

[2] Zhang, Yan, et al. "Power Load Forecasting System of Iron and Steel Enterprises Based on Deep Kernel–Multiple Kernel Joint Learning." Processes 13.2 (2025): 584.

[3] Cai, Changchun, et al. "Short-term load forecasting based on deep learning bidirectional lstm neural network." Applied Sciences 11.17 (2021): 8129.

[4] Jiao, Dongxiang, et al. "Research on a short-term forecast of industry load based on PSO-LSTM." Journal of Physics: Conference Series. Vol. 2814. No. 1. IOP Publishing, 2024.

[5] Wang, Limeng, et al. "Optimal allocation of customer energy storage based on power big data and improved LSTM load forecasting." Energy Reports 11 (2024): 3902-3913.

[6] Zhao, Jiwei, et al. "Forecasting Gate-Front Water Levels Using a Coupled GRU–TCN–Transformer Model and Permutation Entropy Algorithm." Water 16.22 (2024): 3310.

[7] Liu, Yang, et al. "Research on runoff prediction based on Time2Vec-TCN-Transformer driven by multi-source data." Electronics 13.14 (2024): 2681.

[8] Abumohsen, Mobarak, Amani Yousef Owda, and Majdi Owda. "Electrical load forecasting using LSTM, GRU, and RNN algorithms." Energies 16.5 (2023): 2283.

[9] Bareth, Rashmi, et al. "Daily average load demand forecasting using LSTM model based on historical load trends." IET Generation, Transmission & Distribution 18.5 (2024): 952-962.

[10] Wang, Kun, et al. "Long-term power load forecasting using LSTM-informer with ensemble learning." Electronics 12.10 (2023): 2175.