Energy Proceedings
Vol 30, 2023
ISSN 2004-2965

Accurate Building Energy Consumption Prediction with Convolution Recurrent Deep Neural Networks#

Dingrong Dai[1], Huakun Huang[1,4*], Sihui Xue[1], Longtao Guo[1], Lingjun Zhao[3], Huijun Wu[2,4],

1 School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou China

2 School of Civil Engineering, Guangzhou University, Guangzhou, 510006, China

3 School of Software Engineering, Sun Yat-sen University, Zhuhai China

4 Guangdong Provincial Key Laboratory of Building Energy Efficiency and Application Technologies, Guangzhou University, Guangzhou 510006, China (* Corresponding Author: Huakun Huang)

ABSTRACT

Accurate energy consumption prediction is a prerequisite for effectively dispatching distributed power sources. For a building, due to the frequent fluctuations derived from many dynamic factors, the precise energy consumption prediction is still facing challenges. Existing methods usually only use common recurrent neural networks to predict building energy consumption, consider common recurrent neural networks model does not have the ability to extract spatial features and they have a long-term memory problem, so they have limitations to deal with long term task. To overcome these challenges, in this paper, we propose a hybrid model to predict the cooling consumption of a building.

Our hybrid model has the merits of convolutional neural network and gated recurrent unit in capturing spatial-temporal features. Experiment results show that our hybrid model has the best performance, compared with other methods. The result will benefits managers to make reasonable scheduling of power and equipments.

Keywords: Building energy consumption, Load forecasting, Prediction, Deep neural network, CNN, Recurrent

NONMENCLATURE

Abbreviations	
Conv	One-dimensional Convolution layer
Conv2D	two-dimensional convolution
TCN	Temporal Convolutional Network
R2	R-Square
RMSE	Root Mean Square Error
Symbols	
y_i, $\tilde{y}_i$	Actual values, Predicted values
n	Sample number

1. INTRODUCTION

With rapid industrial development, climate change has become one of the world's most relevant and urgent issues. Carbon emissions are an essential driver of climate change. Among the three largest carbon emitters, i.e., transportation, industry, and construction, construction is one of the main areas for carbon emissions[1,2].Thus, reducing carbon emissions from building sectors will significantly improve the current situation.

Building carbon emissions involve almost all aspects, e.g., construction, operation, and maintenance. The future power supply structure will contain a large proportion of renewable electricity, for example, building photovoltaic (PV) power. This enables buildings to consume zero-carbon power if we can predict building consumption and provide renewable power in advance.

However, the random fluctuation characteristic of scenery power generation makes the source side of the power system less controllable—the power balance changes from an unexpected single-side problem to a random double-side problem. At the same time, the large-scale PV source's access to the power grid may increase peak-to-valley difference. The overall shape of the grid load curve will be significantly changed. The short-time load fluctuation of the main power grid will be severe.

The new power balance is prone to supply and sales imbalance and damage to the grid. Predicting building energy consumption is one of the critical issues to alleviate this problem.

For building sectors, the main energy consumption occurs in the building operation phase[3,4].

In addition, building energy consumption is concerned with the Heating, Ventilation, and air conditioning (HVAC) system.

HVAC energy consumption accounts for the most significant part of electricity consumption[5]. It especially needs to be considered in most public buildings and residential buildings. Focusing on the energy consumption of HVAC systems can ensure accurate prediction of the main changes in building electricity consumption and avoid interference of too many irrelevant features.

The existing building energy consumption prediction methods can be divided into two categories, i.e., classical physic and data-based methods. The classical physic method is complex and suffers from limited precision. There are also problems in data acquisition and data quantify[6,7]. The data-based approach is a black-box model that only considers the relationship between historical data and future energy consumption. Time series models are typical representatives of black box models[8]. In addition, Kalman filter[9], and statistical regression are also typical representatives of such methods[10].

Artificial intelligence technologies have recently been widely introduced for building energy consumption prediction. Recurrent neural networks (RNNs) are most suitable for addressing such time-series regression tasks. Long and short-term memory neural network (LSTM) is a representative RNNs with powerful performance.

Gated recurrent unit (GRU) neural network streamlines the structure of LSTM but is more efficient than LSTM. Most studies use GRU networks to deal with such problems.[11] However, both LSTM and GRU are deficient for long-term memory prediction. Whereas building energy consumption data are often sampled with long time steps, so we hope to have a way to shorten the time step while ensuring that feature information is not lost. At the same time, the traditional rnn model cannot extract spatial features, and we also hope we can get spatial feature information for more accurate predictions.[12]

In this work, we propose a hybrid prediction model for accurately predicting building energy consumption. In our model, a convolutional GRU (Conv GRU) neural network is exploited to address the shortcomings of RNNs. The Conv GRU model has the merits of CNN and RNN in capturing spatial-temporal features. It is expected to achieve considerable performance in the long-time-step consumption prediction of a building.

2. METHODOLOGY

2.1 Overview of the proposed approach

The purpose of this paper is to investigate a commonly used energy consumption prediction model for public buildings.

The performance of the neural network model depends mainly on the structure of the network and the adjustment of various parameters, so the workflow of the experimental method is shown in Figure 1.

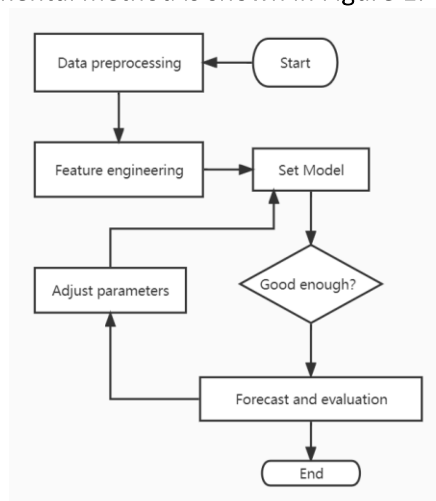

Fig. 1 Workflow of the proposed methods

The pre-processing data stage includes missing data processing, outlier processing, and normalization.

In feature engineering, the variation of building energy consumption is influenced by many factors, which can be divided into internal and external factors. We selected the most relevant features through correlation analysis to improve the prediction accuracy[13].In terms of model selection, this paper iterates each hyperparameter of all models several times. It builds improved models based on convolutional and RNN to make predictions at different time steps.

The evaluation functions in this paper are selected as RMSE and R2, and their formulas are shown as follows.

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

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

2.2 RNN design for extracting temporal features

The neural network used in deep learning comprises the input layer, output layer, and hidden layer. The complex parameter space is fitted by combining a variety of linear activation functions and hidden layers in each layer. To complete the task, the network is continuously trained to achieve a better solution through backpropagation optimization parameters. Many researchers have applied Artificial neural networks (ANN) to building energy consumption prediction and achieved good results[14,15]. All ANNs are feedforward neural networks. RNN has a reverse connection, and the generated input is sent back to itself. The same network is trained and shares weight parameters at each time step, which gives the network the possibility of "memory."

The LSTM is a practical improvement to the simple RNN, preserving the recursive nature of the RNN through the interaction of forgetting gates, input gates, and output gates and avoiding, to a considerable extent, the gradient disappearance and gradient explosion problems commonly encountered in recurrent networks.

The gated loop unit is a simplified version of the LSTM, which combines the functions of the LSTM forgetting gate and input gate, reducing the network parameters and speeding up the training while the model still performs similarly to the LSTM.

2.3 CNN model design for spatial feature extraction

Convolutional networks are networks that specialize in processing grid-like data. Time series data (1D dimension) and image data (2D dimension) are data structures that are well-suited for convolutional network applications. In general, convolutional operations can learn the features of the data and aggregate the low-level basic features into higher-level semantic features to achieve outstanding recognition results.

The convolution operation of time series can aggregate the features in long time steps in shorter time steps by setting the convolution kernel and step size. This method can effectively reduce the length of time steps and help GRU networks perform better. It is even possible to go one step further and discard the recurrent layer entirely and let the network use all convolutional layers, using the method of inflated convolution so that the convolutional layer has the same effect as RNN. This network is called the Temporal convolutional network (TCN) [16] and is shown in Figure 2.

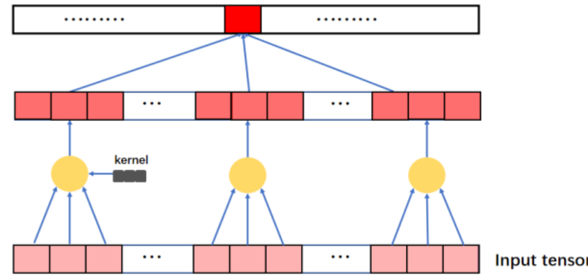

Fig. 2 Outline diagram of temporal convolutional network (TCN)

3. EXPERIMENT SETUP AND DATA PROCESSING

3.1 Data description

The dataset used in the pilot study is the historical operational data of a building water cooling system in a commercial facility, ranging from April 1, 2020, at 0:00 to June 30, 2021, at 24:00. The data is sampled every 15 minutes during this period. The dataset will be divided into three groups in the ratio of 7:2:1, with 70% as the training set, 20% as the validation set, and 10% as the test set, containing 30643, 8755, and 4377 sample sizes, respectively. We set the prediction range to half an hour and 2 hours.

To facilitate understanding of the data distribution, Figure 4 shows a heat map of each data. The purpose of using a heat map is to show the data distribution in the dataset and to give the researcher an intuitive global view of the data. The vertical coordinates in the figure represent all sampling times, and the horizontal coordinates do not define real meaning.

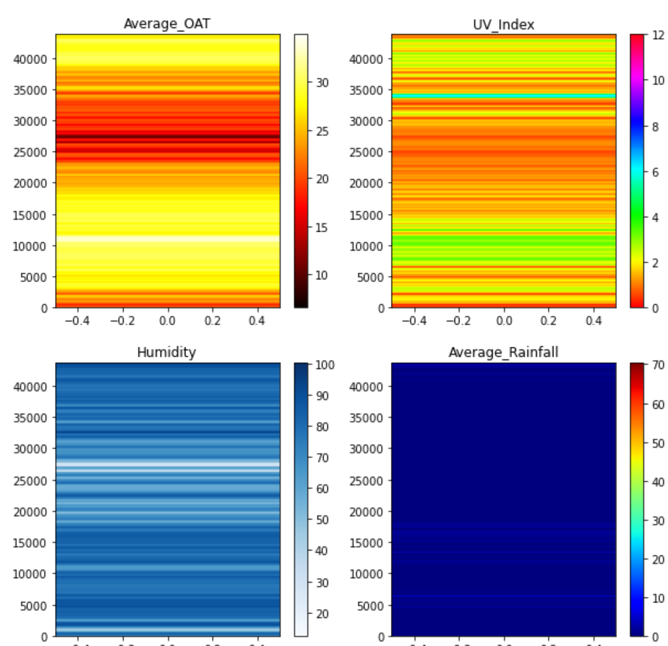

Fig. 3 Heat map of dataset features

3.2 Data preprocessing

To speed up network training and convergence, it is also necessary to reshape the data into the dimensions needed for RNN. We need to pre-process the data. First, for the missing values, the experiments took the nearest valid value to fill the way. Secondly for outliers, thanks to the powerful ability of neural networks, outliers do not bring too many negative impacts on the correct results as traditional statistical analysis[17], in addition, after the analysis of the data set, it is found that the main problem of this data set is more missing values, there are no outliers, so no special treatment of outliers is done in this paper. Third, to avoid the effect of the order of magnitude, the maximum-minimum method is used to scale all data in the range of 0 to 1. The formula is as follows.

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{3}$$

The inverse normalization is needed to recalculate the error after completing the prediction of the data output from the network, and the inverse normalization formula is as follows.

$$X = (X_{max} - X_{min})X_n + X_{min} \tag{4}$$

Finally, we set up sliding windows with window sizes of 96. sliced the entire data into window-sized sequential data and used the data two times after the window size as labels.

3.3 Model setup

In this paper, based on the background work, a standard GRU-based model, a hybrid model of Convolutional Gate recurrent unit (Conv GRU), a hybrid model of 2D convolution and fully connected layers(Conv2D Dense), a hybrid model of 2D convolution and GRU(Conv 2D GRU), and a TCN model are developed. All models used in the experiments were designed using Keras. Figure 4 and Figure 5 are two main model structures.

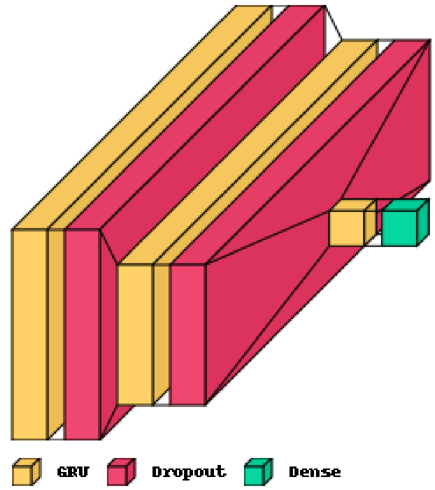

Fig. 4 Framework of GRU model

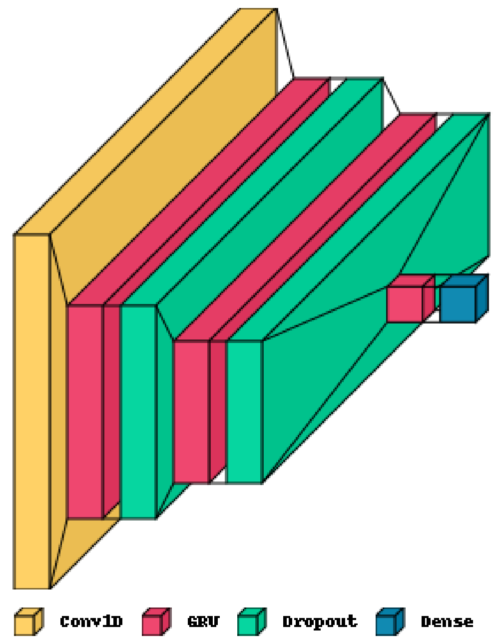

Fig. 5 Framework of the proposed Conv GRU mode

Moreover, to fully demonstrate the model performance, 17 cases were designed in this paper to test the performance of four different models at different time steps for the same feature data.

4. RESULT

The experimental results of the hybrid model with the best effect of Conv GRU are shown in Figure 6 and 7. The red line represents the real energy consumption data, and the blue line represents the predicted data. Considering that the total of more than 4000-time steps is too long, to avoid images masking each other, each type of network is compared with the absolute data values separately. We also took the 1000th to 2000thtime step and enlarged it for a display to show the details of the prediction accuracy.

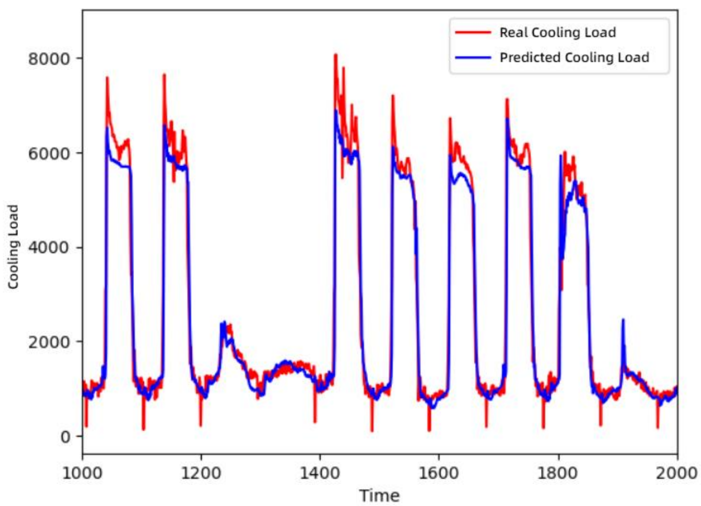

Fig. 6 Performance of GRU on Test Set at 96 Time Steps

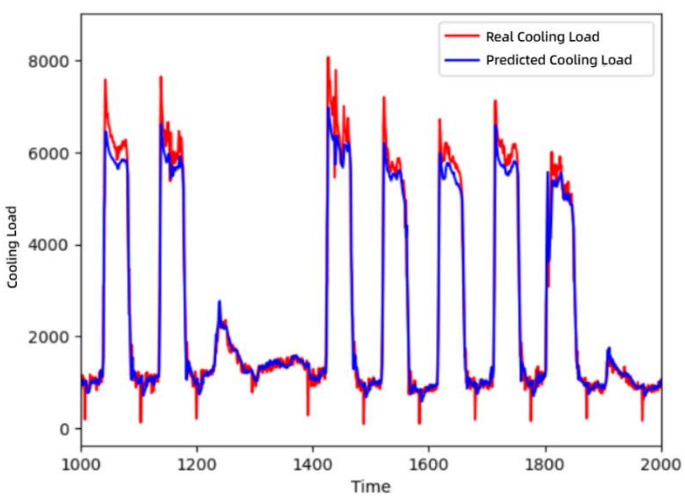

Fig. 7 Performance of Conv GRU on Test Set at 96 Time Steps

Finally, even under the condition of shorter time steps (96-time steps are used as an example in this paper) using the two models GRU and Conv GRU hybrid models for testing, The model performance still has a significant drop in predicting eight-time steps in the future. The R2 values are around 0.7, with an average R2 decrease of about 20% compared to the task type with 2 future time steps, and the RMSE values are about 1000, with an average RMSE increase of 2.5 times compared to the task type with 2 future time steps.

5. DISCUSSION

The current work aims to predict energy consumption in a commercial building using different machine-learning techniques.

From the results, we can conclude that predicting the energy consumption beyond longer future time steps will significantly reduce model performance. The first reason is that the data for this task is sampled every 15 minutes, and the selected features are highly correlated with the building energy consumption, this leads to a decline in model performance. So if we want to predict the average energy consumption of, for example, a day or a week in the future. In that case, we need to reprocess the original data separately. However, it does not meet the purpose of this paper, and the prediction for a longer time in the future can hardly be a practical help for monitoring grid security.

From the aspects of the performance of each model, it can be analyzed that firstly, the Conv GRU model has the best performance with the R2 value of over 0.95 and the RMSE value of only 413. This is because the convolutional layer effectively aggregates the feature information while shortening the time step, which solves the problem of RNN's difficulty in remembering long-time information and is the most suitable network model for such tasks.

Table 1 Comparison of models perform

		GRU	Conv GRU	Conv2D GRU	Conv2d Dense	TCN
96 times step	R2	0.932	**<u>0.959</u>**	0.947	0.951	0.935
	RMSE	512.5	**<u>413.9</u>**	496.4	447.7	498.7

Both from Table 1, and Figure7, it can be concluded that the Conv GRU has the best performance at all time steps, and the version is improved compared to the ordinary GRU model, and the predicted value curves fit the actual value curves more closely. In addition, it is worth mentioning that although the hybrid network of Conv2D GRU does not achieve the best results, this network has the least fluctuation and the most stable model performance under all conditions.

The fluctuation of the Conv2D GRU model is minimal under different conditions because its convolutional kernel can move between time steps and features, which has one more dimension than one-dimensional convolution. Feature information is extracted more fully, but it also increases the time step length, which leads to the degradation of model performance. This model is suitable for task types with a large variety of features and a small amount of information contained in a single feature.

Considering that the above RNN is incapable of handling long time steps and the two-dimensional convolution must increase the value of the data time step dimension, this paper is supplemented by designing a comparison experiment of the Conv2D Dense model. According to the results, it can be concluded that the overall effect of the model of Conv2D Dense is better than that of the model of Conv2D GRU. The RMSE value can be decreased by about 50. Still, the number of parameters of the Conv2D Dense model will be about 100 times more than that of the model of 2D convolution plus GRU, so it will be discussed only a little.

Finally, The reason for the poor performance of the TCN model is that the inflated convolution approach used by the TCN model expands the parametric field of view of the model but also loses essential information between the areas of view. At the same time, the task studied in this paper is that the vital information is hidden in each similar time step data. TCN is more suitable for the job of audio recognition.

In summary, Conv GRU is suitable for tasks with a small number of features, a strong correlation between features task.

Conv2D GRU is suitable for many kinds of features, and the correlation between features is not clear. TCN is suitable for tasks with particularly long time steps, such as audio tasks.

6. CONCLUSIONS

This paper aims to overcome several drawbacks of traditional rnn and predict building energy consumption more accurately while dealing with complex energy consumption conditions. For overall consideration, we performed experiments under various time-step conditions. Predicting energy consumption would play a vital role in early response to peak loading (peak loading), power, or equipment regulation to ensure grid security. Moreover, suitable situations for each model are also analyzed to provide a reference for other application scenarios.

In the best case, the Conv GRU has an RMSE value of around 400 and an R2 value of around 0.96. Compared with the cold consumpution value of 6000+ in most cases, the RMSE value of 400 is small enough, which means the performance of Conv GRU is good enough for practical use. In addition, the Conv GRU could achieve more accurate predictions than both GRU model and TCN model, which is 10% and 30% improvement, respectively.

In summary, the Conv GRU could consistently achieve better performance in building cold consumption prediction than the other compared models, which could be the optimized alternative to a standard RNN. Additionally, it can be expected that the Conv GRU will outperform LSTM and GRU on most consumption prediction tasks with long time steps.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 62001126, in part by the Natural Science Foundation of Guangdong Province of China (Grant No. 2022A1515010163), and in part by the Science and Technology Program of Guangzhou (Project No. 202201010348).

REFERENCE

[1] Guangpeng Zhang,Huijun Wu,Jia Liu,Jianming Yang,Huakun Huang,Yujie Ding,and Lei Xie.Dynamic performance and energy efficinency of reflective and insulative composite coating on building exterior wall. In Building Simulation, pages 1 15.Springer, 2022

[2] Huang H, Zhou Y, Huang R, et al. Optimum insulation thicknesses and energy conservation of building thermal insulation materials in Chinese zone of humid subtropical climate[J]. Sustainable Cities and Society, 2020, 52: 101840.

[3] Hong T, Wang Z, Luo X, et al. State-of-the-art on research and applications of machine learning in the building life cycle[J]. Energy and Buildings, 2020, 212: 109831.

[4] Huang H, Zeng X, Zhao L, et al. Fusion of building information modeling and blockchain for metaverse: a survey[J]. IEEE Open Journal of the Computer Society, 2022, 3: 195-207.

[5] Yu X, Ergan S, Dedemen G. A data-driven approach to extract operational signatures of HVAC systems and analyze impact on electricity consumption[J]. Applied Energy, 2019, 253: 113497.

[6] Fumo N, Mago P, Luck R. Methodology to estimate building energy consumption using EnergyPlus Benchmark Models[J]. Energy and Buildings, 2010, 42(12): 2331-2337.

[7] Prakash A K, Xu S, Rajagopal R, et al. Robust building energy load forecasting using physically-based kernel models[J]. Energies, 2018, 11(4): 862.

[8] Amasyali K, El-Gohary N M. A review of data-driven building energy consumption prediction

studies[J]. Renewable and Sustainable Energy Reviews, 2018, 81: 1192-1205.

[9] Kuo P H, Huang C J. A high precision artificial neural networks model for short-term energy load forecasting[J]. Energies, 2018, 11(1): 213.

[10] Ahmad A, Javaid N, Mateen A, et al. Short-term load forecasting in smart grids: An intelligent modular approach[J]. Energies, 2019, 12(1): 164.

[11] Liu B, Fu C, Bielefield A, et al. Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network[J]. Energies, 2017, 10(10): 1453.

[12] Fan C, Wang J, Gang W, et al. Assessment of deep recurrent neural network-based strategies for short-term building energy predictions[J]. Applied energy, 2019, 236: 700-710.

[13] Zhang X, Wang F. Hybrid input-output analysis for life-cycle energy consumption and carbon emissions of China's building sector[J]. Building and Environment, 2016, 104: 188-197.

[14] Acheampong A O, Boateng E B. Modelling carbon emission intensity: Application of artificial neural network[J]. Journal of Cleaner Production, 2019, 225: 833-856.

[15] Huang H, Ogbodo M, Wang Z, et al. Smart energy management system based on reconfigurable AI chip and electrical vehicles[C]//2021 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2021: 233-238.

[16] Fan J, Zhang K, Huang Y, et al. Parallel spatio-temporal attention-based TCN for multivariate time series prediction[J]. Neural Computing and Applications, 2021: 1-10.

[17] Liu H, Chen C. Data processing strategies in wind energy forecasting models and applications: A comprehensive review[J]. Applied Energy, 2019, 249: 392-408.