

A short term load forecasting of integrated energy system based on CNN-LSTM

Xianjun Qi^{1,2}, Xiwei Zheng^{2*} and Qinghui Chen²

¹Anhui Provincial Laboratory of Renewable Energy Utilization and Energy Saving, Hefei University of Technology, Hefei, Anhui Province, 230009, China

²School of Electrical and Automatic Engineering, Hefei University of Technology, Hefei, Anhui Province, 230009, China

Abstract: The accurate forecast of integrated energy loads, which has important practical significance, is the premise of the design, operation, scheduling and management of integrated energy systems. In order to make full use of the coupling characteristics of electricity, cooling and heating loads which is difficult to deal with by traditional methods, this paper proposes a new forecast model of integrated energy system loads based on the combination of convolutional neural network (CNN) and long short term memory (LSTM). Firstly, the Pearson correlation coefficients among the electricity, cooling and heating load series of the integrated energy system are calculated, and the results show that there is a strong coupling relationship between the loads of an integrated energy system. Then, the CNN-LSTM composite model is constructed, and CNN is used to extract the characteristic quantity which reflects the load coupling characteristics of the integrated energy system. Then, the characteristic quantity is converted into the time series input to LSTM, and the excellent time series processing ability of LSTM is used for load forecasting. The results show that the CNN-LSTM composite model proposed in this paper has higher prediction accuracy than the wavelet neural network model, CNN model and LSTM model.

1 Introduction

Energy and environmental issues are the hotspots of today's society, and they affect human sustainable development at all times. However, the traditional fossil energy sources are increasingly depleted, and the existing energy supply methods have many disadvantages such as inefficient conversion, low concentration, and high costs. This makes the energy problem the primary factor affecting sustainable development. The integrated energy system is a system that integrates the supply of electricity, natural gas, heating and cooling energy [1]. It is an important trend in energy development and plays an important role in promoting the optimization of energy structure, inspiring renewable energy consumption, and improving energy utilization.

Recently, research about integrated energy load forecasting methods can be divided into two categories. The first type of research uses time series methods. For example, Reference [2] uses multiple linear regression model and initialized auto regressive model to predict integrated energy system loads. This method makes full use of the time sequence of the load data, but the disadvantage is that the model has insufficient nonlinear fitting ability. The second type of research uses machine learning. Reference [3] combines neural networks and support vector machines to propose a composite

prediction method for forecasting cooling and heating loads. The results of study in reference [4,5] show that the efficiency of heating and cooling load forecasting using the neural network is better than the results obtained by the logistic regression and random forest.

In recent years, with the development of artificial intelligence technology, deep learning algorithms such as convolutional neural network (CNN) and long short term memory network (LSTM) have gradually emerged [6]. CNN has the advantage of extracting local features from multi-dimensional data, and has been widely used in image recognition classification [7], face recognition [8], audio retrieval [9] and other fields. LSTM is a variant of recurrent neural network, which overcomes the problems of gradient disappearance and gradient explosion of traditional recurrent neural network [10]. and has the ability to process time series and nonlinear data simultaneously. It has a breakthrough performance in translation [11], speech [12] and other fields. At present, LSTM is gradually applied to the field of forecasting. For example, the study of reference [13] shows that the proposed LSTM approach outperforms the other listed rival algorithms in the task of short-term load forecasting for individual residential households. Reference [14] Proposes a hybrid model for short-term wind power forecasting, which consists of the variational mode decomposition the K-means clustering algorithm and LSTM network.

*Corresponding author's e-mail: 1227055262@qq.com

In order to fully utilize the coupling relationship between electricity, cooling and heating loads, and improve the accuracy of load forecasting of integrated energy systems, a short-term load forecasting method of integrated energy systems based on CNN-LSTM is proposed. First, the Pearson correlation coefficient is used to analyze the correlation between electricity, cooling, and heating loads, the intensity of load coupling relationship in integrated energy system is discussed. After that, a CNN-LSTM composite model is constructed, and CNN is used to extract the coupling features between electrical, cooling and heating loads, then the coupling features are input into the LSTM for load forecasting. Compared with the forecasting results of wavelet neural network model, CNN model and LSTM model, it can be verified that the combination model proposed in this article combines the advantages of CNN and LSTM, and can effectively extract the coupling characteristics between electric, cooling and heating loads, which improve the forecasting precision.

2 Correlation analysis of integrated energy load

Integrated energy system has a unique structure, different types of loads can be coupled together through energy conversion devices. For example, electric boiler can transform electric loads into heating loads, electric refrigerator can transform electric loads into cooling loads, absorption refrigerator can transform heat loads into cooling loads, etc. Therefore, it is necessary to consider the coupling relationship among electric, cooling and heating loads. In order to intuitively analyze the coupling relationship among loads in integrated energy system, Pearson correlation coefficient is used to quantify the coupling relationship among electrical, cooling and heating loads. The calculation formula of Pearson correlation coefficient of time series $\{X_i\}$ and $\{Y_i\}$ is shown in formula (1).

$$\rho_{xy} = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}} \quad (1)$$

Where, \bar{X} and \bar{Y} represent the average value of time series $\{X_i\}$ and $\{Y_i\}$; N is the number of time series data. The greater the absolute value of ρ_{xy} , the stronger the correlation between $\{X_i\}$ and $\{Y_i\}$. In this paper, the correlation analysis is carried out by using the historical data of electricity, cooling and heating loads of the integrated energy system of an industrial area in 2018, and the results are shown in Table.1.

Table 1. Correlation coefficients among electricity, cooling and heating loads

	Electric load	Cooling load	Heating load
Electric loads	1	0.7298	0.7466
Cooling loads	0.7298	1	0.6558

Heating loads	0.7466	0.6558	1
---------------	--------	--------	---

From Table 1., it can be seen that the correlation coefficient between electric, cooling and heating loads is very large, and the correlation coefficient between each two loads is above 0.6, which shows that the integrated energy system has strong loads coupling characteristics. Therefore, in the load forecasting of integrated energy system, making full use of the coupling information between one loads and other loads will help to improve the forecasting accuracy, so as to obtain a better overall forecasting effect.

CNN has a strong ability of data feature extraction, which can effectively extract the coupling features among electrical, cooling and heating loads, and establish the corresponding feature quantity. However, CNN has a general performance in time series prediction, which is difficult to achieve accurate time series prediction. Therefore, this paper combines LSTM with CNN, which has strong timing processing ability, and inputs the coupling feature extracted by CNN into LSTM for load forecasting. The CNN-LSTM composite forecasting model has both advantages and it's very suitable for integrated energy system load forecasting.

3 CNN-LSTM composite prediction model

3.1 CNN model

CNN is a multilayer feedforward neural network, which is composed by convolution layer, pooling layer and full connection layer. Convolution layer is the key component of the whole convolution neural network. Each convolution layer contains multiple feature planes, each feature plane is composed of multiple neurons. When the data is input into the convolution layer, the convolution kernel on the neuron performs convolution operation along the data to extract the local features of the input data. The feature generated by convolution layer is input again and transferred into the pooling layer, which reduces the height and width of the feature and plays the role of secondary feature extraction. The full connection layer is fully connected with all neurons in the previous layer, which can integrate the local information with category differentiation in the convolution layer or pooling layer.

In this paper, the convolution neural network is used to process the electric, cooling and heating loads, and the output characteristic quantity includes the coupling characteristics of electric, cooling and heating loads. In front of the full connection layer, the LSTM network is added, and the extracted characteristic quantity is input into the LSTM for load forecasting.

3.2 LSTM network

LSTM is a special kind of cyclic neural network, which adds three gate structures: forget gate, input gate and output gate on the basis of cyclic neural network, and has

the ability to automatically store and delete time state information [15]. Therefore, LSTM can extract the complex feature relations of long and short time series, and has excellent performance in processing time series. The core of LSTM is memory unit, which is used to store current and past time state information. The LSTM adds or deletes information to the memory unit by controlling the input gate, forgetting gate and output gate. The structure diagram of LSTM is shown in Figure 1. The status update process of the memory unit is shown as follows:

(1) The forget gate deletes information from the memory unit according to the input of the current time x_t and the output of the previous time h_{t-1} :

$$f_t = \sigma(w_{fh} \cdot h_{t-1} + w_{fx} \cdot x_t + b_f) \quad (2)$$

Where, f_t is the state of forgetting gate, w_{fh} and w_{fx} are the weight matrix of forgetting gate, b_f is the offset of forgetting gate, and σ is the excitation function sigmoid.

(2) The input gate determines the information stored in the memory unit:

$$i_t = \sigma(w_{ih} \cdot h_{t-1} + w_{ix} \cdot x_t + b_i) \quad (3)$$

$$\tilde{c}_t = \tanh(w_{ch} \cdot h_{t-1} + w_{cx} \cdot x_t + b_c) \quad (4)$$

Where, i_t is the state of input gate, \tilde{c}_t is the state of candidate element, w_{ih} , w_{ix} and w_{ch} are the weight matrix of input gate, b_i and b_c are the offset of input gate, and \tanh is the excitation function Tanh.

(3) Update the unit status C_t according to the status of forgetting gate and input gate:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{c}_t \quad (5)$$

(4) Update the status of the output door O_t in the current unit status:

$$o_t = \sigma(w_{oh} \cdot h_{t-1} + w_{ox} \cdot x_t + b_o) \quad (6)$$

Where, w_{oh} and w_{ox} are the weight matrix of the output gate, b_o is the offset of the output gate.

(5) The final output of the output gate is:

$$h_t = o_t \cdot \tanh(C_t) \quad (7)$$

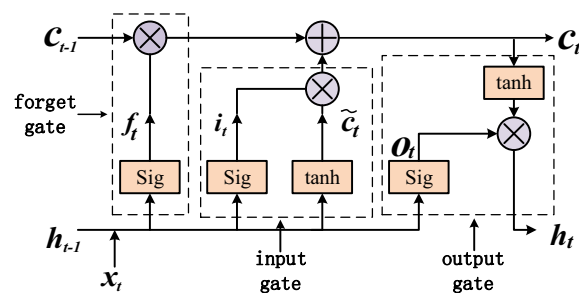


Figure 1. Basic structure of LSTM model

3.3 CNN-LSTM composite model

In this paper, CNN-LSTM composite model is used to predict the electric, cooling and heating load of the integrated energy system. Considering the characteristics of CNN and LSTM networks, the input matrix X is composed of the historical data of electric, cooling and heating loads from the starting time to the $T-1$ time, as shown in equation (8).

$$X = \begin{bmatrix} X_e \\ X_c \\ X_h \end{bmatrix} = \begin{bmatrix} x_e(0) & x_e(1) & \cdots & x_e(t-1) \\ x_c(0) & x_c(1) & \cdots & x_c(t-1) \\ x_h(0) & x_h(1) & \cdots & x_h(t-1) \end{bmatrix} \quad (8)$$

Where, X_e , X_c and X_h represent electrical load, cooling load and heating load respectively.

In the training and prediction of the model, it is input to CNN-LSTM composite model by sliding window along the matrix X value in a certain step to carry out feature extraction and timing prediction. Considering the long and short memory function of LSTM, the size of sliding window is set to 24 and the step size is set to 1. That is to say, the historical load data of electric cooling and heating from $t-25$ to $t-1$ are used to predict the load values of electric cooling and heating at t time.

The structure of CNN-LSTM composite model is shown in Figure 2. The first half of the model is CNN network, which is responsible for extracting the coupling characteristics of electricity, cooling and heating loads. The second half is LSTM network, which is responsible for the time sequence prediction of electricity, cooling and heating loads. Taking the prediction of electrical, cooling and heating loads at time t as an example, the 3×24 input matrix is convoluted with 3×3 convolution kernels in the convolution layer to form a $3 \times 3 \times K$ array. The three-dimensional array reduces the size after the maximum pooling layer. The input data is convoluted and pooled many times to form a $2 \times 2 \times 24$ feature array, and then flattened through the flattening layer to form a 1×96 time series. Then the time series is input into the LSTM network for load forecasting. Finally, the electric, cooling and heating loads forecasting values at moment t with dimension of 1×3 are output through the full connection layer.

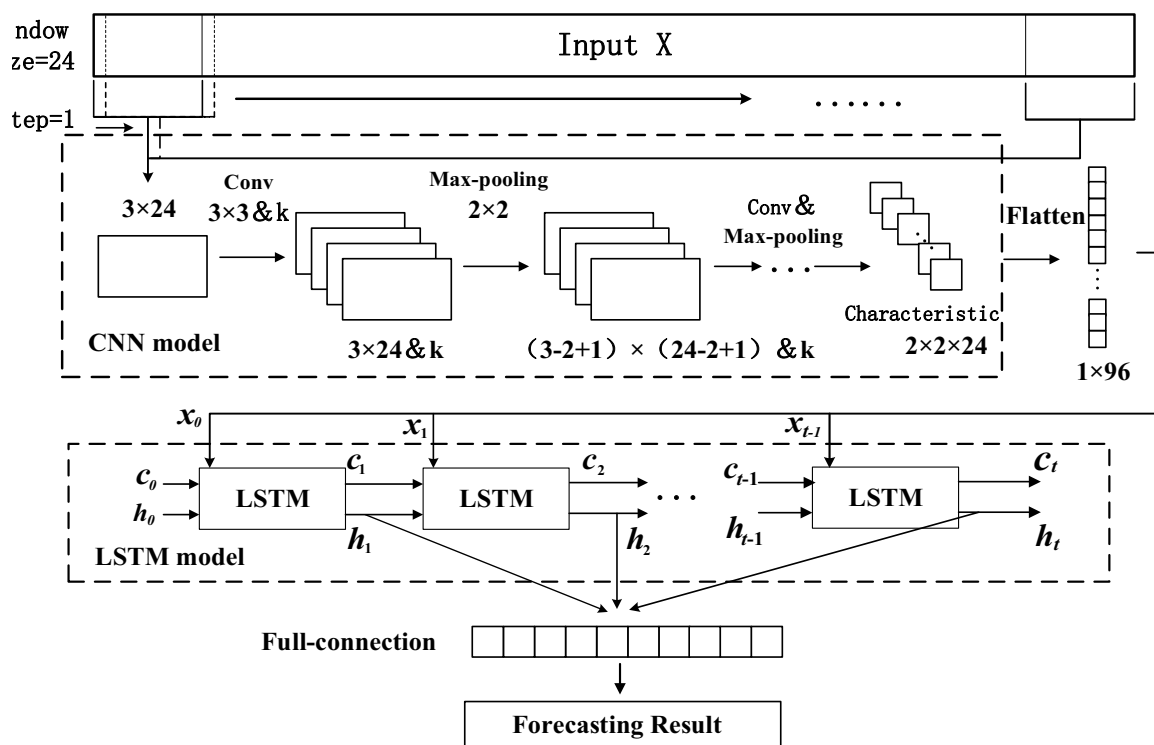


Figure 2. Structure diagram of CNN-LSTM composite model

3.4 evaluation index

In this paper, mean absolute percentage error (MAPE) and weighted mean accuracy (WMA) are selected as the evaluation indexes of model prediction accuracy. MAPE and WMA are calculated as follows:

$$M_{APE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y'(t) - y(t)}{y'(t)} \right| \times 100\% \quad (9)$$

$$M_A = 1 - M_{APE} \quad (10)$$

$$W_{MA} = \alpha_e M_{Ac} + \alpha_c M_{Ac} + \alpha_h M_{Ah} \quad (11)$$

Where: $y'(t)$ and $y(t)$ are the actual value and predicted value at time t , n is the total number, α_e , α_c and α_h are the weights of electric, cooling and heating loads respectively. Considering the dominance and importance of power resources in the integrated energy system, the weight coefficients of electric, cooling and heating loads are set as 0.6, 0.2 and 0.2 [5-6].

4 example analysis

The example data comes from the integrated energy system of an industrial area in China, which is a combined electric, cooling and heating supply system. The power, cooling and heating load data of the park in March and April 2018 are used as the training data set for the training prediction model. The sampling interval is 1 hour. The

data from March 1 to April 29 are used as the training data set. The data from April 30 is used as the test data set.

4.1 data processing

Firstly, the input data is normalized and mapped to the range of 0 ~ 1 to meet the input requirements of the neural network model. The normalization method is shown in equation (12):

$$x'_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (12)$$

Where: x'_i is the input data after normalization, $\min(x)$ and $\max(x)$ are the minimum value and maximum value in the data sequence respectively, x_i is the data before processing.

4.2 model parameter setting

The structure of neural network has a great influence on the result of load forecasting. The more layers CNN and LSTM modules are, the stronger the learning ability of the model is, and the higher the prediction accuracy is. However, too many layers are likely to lead to over fitting of the model. Therefore, the control variable method is used to gradually increase the number of layers of CNN and LSTM network, test the prediction effect of different structures of the model, select the optimal combination of parameters to build the load prediction model, and the test results are shown in Table 2.

Table 2. Test results of model structure

CNN layers	LSTM layers	Number of iterations	MAPE
2	2	200	0.0311
	3	200	0.0258
	4	200	0.0297
3	3	200	0.0305
	4	200	0.0344

It can be seen from Table 2 that the model fitting error of the combination of two layers CNN with three layers LSTM is the smallest. Therefore, this structure is used to build the prediction model, and the number of convolution kernels in two layers CNN are set to 8 and 24, the size of convolution kernels are set to 3×3 and 2×2 , the number of neurons in three layers LSTM are set to 16, 16 and 32, and the number of nodes in the full connection layer is set to 10. The initial learning rate is set to 0.01, and the optimization algorithm is set to Adam optimization algorithm. In order to improve the operation speed of the composite model and prevent the model from over fitting,

the dropout method is adopted for each layer of LSTM, and the dropout is set to 0.5.

4.3 analysis of calculation results

In the experiment, Tensorflow 2.0, which was a learning framework developed by Google, and python 3.5 were used as the programming language to complete the construction and training of CNN-LSTM composite model in the development environment of Pycharm and Anaconda.

In order to verify the validity of the combination model, four groups of prediction models are set up: model 1 adopts wavelet neural network (WNN) model; model 2 adopts single CNN model; model 3 adopts single LSTM model; model 4 adopts CNN-LSTM composite model. The above structural parameters are used in model 2 and model 3, and the same training and test data sets are used in 4 groups of models. The load prediction curve of each model and the prediction accuracy are shown in Figure 3.

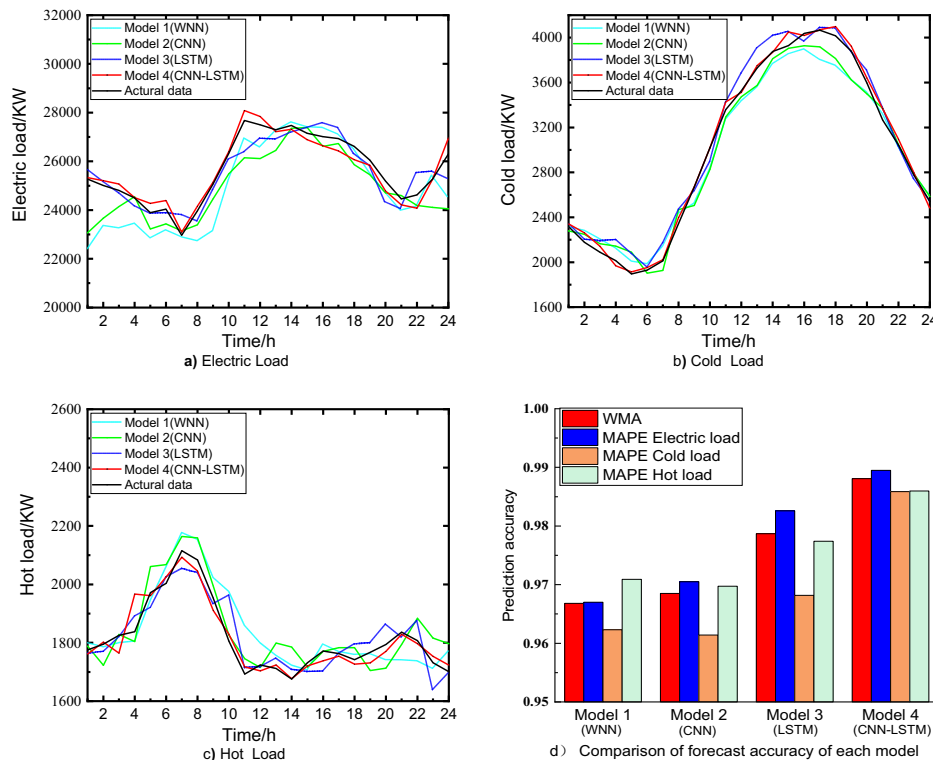


Figure 3. Forecast results of each model

It can be seen from Figure 3 that:

(1) From Figure 3 a), b) and c), it can be seen that the predicted load curve of WNN and CNN model does not fit the actual load curve well, while the predicted curve of LSTM model and CNN-LSTM combination model fits the actual curve well.

(2) The CNN-LSTM composite model has both the feature extraction ability of CNN model and the timing fitting ability of LSTM model. The coupling relationship among electricity, cooling and heating loads is fully utilized to improve the prediction accuracy. As can be seen from Figure 3 d), the CNN-LSTM composite model has

the highest prediction accuracy among the four groups of models. Compared with the single LSTM model, the prediction accuracy of electric, cooling and heating loads of the CNN-LSTM composite model has increased by 0.7%, 1.7% and 0.8% respectively, and the weight accuracy has reached 0.988.

5 Conclusions

In order to make full use of the coupling relationship among the electric, cooling and heating loads of the integrated energy system and improve the overall

accuracy of the load forecasting of the integrated energy system, a load forecasting model of the integrated energy system based on CNN-LSTM combined neural network is proposed in this paper. The main advantages are as follows:

(1) By using the feature extraction ability of CNN network, the data of electricity, cooling and heating loads are input into CNN network, and the coupling features of electricity, cooling and heating loads are deeply mined.

(2) The features including the coupling characteristics of electricity, cooling and heating loads are input into the LSTM, and the memory function and time sequence processing ability of the LSTM network are used for prediction.

(3) Compared with wavelet neural network model, single CNN model and LSTM model, CNN-LSTM composite prediction model has the highest prediction accuracy.

Acknowledgments

This work was supported by Institute of Economy and Technology, State Grid Anhui Electric Power Co., Ltd. (Foundation No. B3440818K005)

References

1. Shen F., Li Z., Shi X.(2020) Singular Perturbation for the Dynamic Modeling of Integrated Energy Systems. *IEEE Transactions On Power Systems*, 35:1718-1728.
2. Vaghefi A., Jafari M., Emmanuel B. (2014) Modeling and forecasting of cooling and electricity load demand. *Applied Energy*, 136:186-196.
3. Kody M., Akshay S., Wesley J.(2014) Heating cooling and electrical load forecasting for a large-scale district energy system. *Energy*, 74: 877-885.
4. Vladimir V.M.(2019) Prediction of Heating Load and Cooling Load of Buildings Using Neural Network. *International Ural Conference on Electrical Power Engineering. Chelyabinskaya Oblast, Russia*. pp. 417-421.
5. Nnamdi I.N. (2017) An artificial neural network model for predicting building heating and cooling loads. *International Artificial Intelligence and Data Processing Symposium (IDAP). Malatya, Turkey*. pp.17-22.
6. Chourasia V.S., Mitra A.K.(2010)Wavelet-based denoising of fetal phonocardiographic signals. *International Journal of Medical Engineering & Informatics.*, 2 :139-150.
7. Szegdy C., Liu W., Jia Y. (2015) Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA*. pp. 1-9.
8. Chen L. (2014) A Fair Comparison Should Be Based on the Same Protocol—Comments on “Trainable Convolution Filters and Their Application to Face Recognition”. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 36: 622-623.
9. Ossama A., Mohamed A., Jiang H.(2012) Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. *Proceeding of the IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP). Kyoto, Japan*. pp. 4227-4280.
10. Hochreiter S., Bengio Y., Frasconi P., Schmidhuber J.(2001) Gradient flow in recurrent nets: the difficulty of learning long-term dependencies: a field guide to dynamical recurrent neural networks. *IEEE Press*. 10:237-243.
11. Lore K.G., Stoecklein D., Davies M. (2017)A deep learning framework for causal shape transformation. *Neural Networks*.98:305-317.
12. Medennikov I., Bulusheva A. (2016)LSTM-Based Language Models for Spontaneous Speech Recognition. *International Conference on Speech & Computer. Okayama, Japan*. pp. 469-475.
13. Kong W., Zhao Y., David J.H.(2019) Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Transactions On Smart Grid*. 13: 2062-2069.
14. Sun Z., Zhao S., Zhang J.(2019) Short-Term Wind Power Forecasting on Multiple Scales Using VMD Decomposition, K-Means Clustering and LSTM Principal Computing. *IEEE ACCESS*.7: 166917-166929.
15. Gers F., Schmidhuber J.(1997) Cummins F. Long short-term memory. *Neural Computation*. 9: 1735-1780.