Power System Load Forecasting Method Based on LSTM Network

CuiYun Luo¹, WuNeng Ling¹, Dong Mo¹, *, YiMing Li¹, QiuWen Li¹, ZhenCheng Liang¹ and Pan Zhang¹

¹Guangxi Power Grid Corporation, Dispatch and Control Center, Nanning, Guangxi, China

* Corresponding author: liujiawen@orait.cn

Abstract. Scientific load forecasting methods and accurate forecasting results are an important basis for the power system planning department work, as well as the basis and guarantee for correct decision-making on investment and construction. This paper has thoroughly studied the basic principles of recurrent neural networks and their shortcomings that they cannot solve long-term dependence problems, and a power system load forecasting method based on LSTM networks is proposed. Through the built-in two memory state lines of long term and short term, the problems that it is difficult to capture the long-term development trend and short-term fluctuation characteristics of data in time series prediction were solved. The results of case analysis show that this method can effectively improve the accuracy of power system load prediction.

1. Introduction

Power system load forecasting refers to exploring the influence of power load’s own historical data changes on future loads based on historical data such as power load, economy, society, and weather, and seeking the internal connection between power load and various related factors, and making advance estimates and speculations on the future power demand based on the future development trends of economic, meteorological and other factors. According to the time scale, the medium- and long-term load forecasting carried out between one month and several years in advance is called medium- and long-term load forecasting. The forecast object is usually electricity consumption or maximum and minimum load, which is mainly used to formulate power system planning, and provide power grid planning and power supply in the region. The basis provided by planning is an important foundation for power system planning and grid construction. Scientific medium and long-term load forecasting methods and accurate forecasting results are an important basis for the work of power system planning departments, as well as the basis and guarantee for correct decision-making on investment and construction.

In terms of medium- and long-term load forecasting, some forecasting methods have been put forward theoretically to consider the influence of domestic and foreign macroeconomic situation, electricity demand trend, industry expansion and installation situation, climate change and other factors on power demand. In particular, unit consumption method, elastic coefficient method, regression method and analogy method are commonly used for forecasting [1,2,3,4]. However, in the actual forecasting process, due to method defects and weak data foundation, it is still generally dependent on the subjective judgment of experts. Subjective empirical judgments account for a large proportion of the forecasting process, which leads to relatively extensive forecasting results. The specific problems are as follows:
Frist, the related mechanism of load, economy and weather is not clear, and the deep law of various influencing factors such as load and economic and financial indicators, extreme weather, production characteristics of various industries, and the living habits of thousands of households has not been implemented. Systematic mining; Second, the forecasting methods still use traditional linear models such as regression/extrapolation/multiplying ratios. These simple methods are not ideal for the non-linear correlation between current load and economic and meteorological factors.

Therefore, non-analytic methods that consider multiple factors such as gray theory and neural networks are constantly emerging. Zhang et al. [5] proposed a power load forecasting method based on lifting wavelet feature extraction, and using the feature data after noise reduction to do time series analysis of residential community load, but the simple data dimensionality reduction and forecasting process can easily lead to the adaptability of the model Low, Yang and W.Z et al. [6,7] proposed an improved gray forecasting model to achieve higher mid- and long-term load forecasting accuracy than traditional gray prediction models. Y.S et al. [8] proposed a long-term load forecasting model based on an improved genetic neural network. Xinran L et al. [9] proposed a composite load model based on Elman artificial neural network, which has a simple structure and is convenient for engineering applications. Although many achievements have been made in the field of medium and long-term load forecasting at home and abroad, as the factors affecting the medium and long-term load appear more and more time scales, the medium and long-term load forecasting needs more considerations of the dependence between data at multiple scales, to solve the problem that the long-term development trend of data and short-term fluctuation characteristics are difficult to capture simultaneously.

To this end, based on the above requirements, this paper has thoroughly studied the basic principles of recurrent neural networks and their inability to solve the long-term dependence problem, and proposed a power system load forecasting method based on the LSTM network. This method solves the problem that it is difficult to capture the long-term development trend and short-term fluctuation characteristics of data simultaneously in time series prediction by built-in long-term and short-term memory state lines, and proves the effectiveness and accuracy of the method proposed in this paper in medium- and long-term load prediction through case analysis.

2. Algorithm principle

2.1. Recurrent Neural Networks (RNNs)

Traditional neural networks cannot record the continuity of time series, which makes them often not ideal for time series forecasting. But an improved recurrent neural network can do it. In the network of RNNs, there is a cyclic operation that allows them to retain the time series content learned before.

In the above network structure, for the part of the rectangular block A, by inputting xt (the feature vector at time t), it will output a result ht (the state or output at time t). The cyclic structure in the network enables the state at a certain moment to be transmitted to the next moment.
The structure of these loops makes RNNs seem a bit difficult to understand. Here, RNNs can be regarded as a common network that has been duplicated multiple times and then superimposed together. Each network will pass its output to the next network. Here, RNNs can be expanded in time steps, and the following figure is obtained:

![RNNs expanded network structure](image)

From the chain structure of RNNs, it is easy to understand that it is related to sequence information. This structure seems to be born to solve sequence-related problems. RNNs can achieve such success, mainly due to the use of LSTMs. This is a special kind of RNNs, and for many tasks, it is much better than ordinary RNNs. Basically, the recurrent neural networks in use now use LSTMs.

The emergence of RNNs is mainly because they can link the previous information to the present, thereby solving the current problems. For example, using the previous time sequence can predict the subsequent time sequence, and using the previous picture can help understand the content of the current picture here.

Sometimes when dealing with the current task, we only need to look at some recent information. For example, load forecasting only needs the recent load. In this case, the interval between the content to be predicted and the relevant information is very small. In this case, RNNs can be easily implemented using past information.

![Short-term dependence](image)

However, as the interval between prediction information and related information increases, it is difficult for RNNs to correlate them.

![Long-term dependence](image)

Theoretically speaking, by selecting appropriate parameters, RNNs can indeed connect such long-term dependencies ("long-term dependencies") and solve such problems. Unfortunately, in practice, RNNs cannot solve this problem.

2.2. LSTM network structure

Long Short-Term Memory networks—usually called "LSTMs"—are a special type of RNN. Proposed by Hochreiter & Schmidhuber (1997), it is very popular, and it has been improved and adjusted by many people since then.

Specifically, the design of LSTMs is mainly to avoid the aforementioned long-term dependency problem. Their essence is to be able to remember information over a long period of time.
All recurrent neural network structures are copied from (neural network) modules with exactly the same structure. In ordinary RNNs, this module structure is very simple, such as only a single tanh layer.

![Figure 5 Internal structure of ordinary RNNs](image)

LSTMs also have a similar structure (the only difference is the middle part). But they are no longer just using a single tanh layer, but using four interacting layers.

![Figure 6 LSTM internal structure](image)

In the network structure diagram, each line transmits a vector, which is output from one node, and then input to another node. The pink circle represents point-by-point operations, such as vector addition; the yellow rectangle represents a neural network layer (that is, a lot of neural nodes); the merged line represents the combination of the vectors carried on the two lines (such as one with \( h_{t-1} \) and the other with \( x_t \)), then the combined output is \( [h_{t-1}, x_t] \); the separate lines mean that a copy of the vector passed on the line is passed to both places.

The most critical part of LSTMs is the state of the cell (the entire green box is a cell) and the horizontal line on the structure diagram.

The transmission of the cell state is like a conveyor belt. The vector passes through the entire cell with only a few linear operations. This structure can easily realize that information can pass through the entire cell without changing it.

![Figure 8 Conveyor belt structure](image)
If there is only the horizontal line above, there is no way to add or delete information. It is achieved through a structure called gates. The gate can selectively let information through, mainly through a sigmoid neural layer and a point-by-point multiplication operation.

![Figure 9 Gate structure (sigmoid layer)](image)

Each element of the output of the sigmoid layer (which is a vector) is a real number between 0 and 1, which represents the weight (or proportion) that allows the corresponding information to pass. For example, 0 means ‘don't let any information pass’, and 1 means ‘let all information pass’.

Each LSTM has three such gate structures to realize protection and control information, namely ‘forget gate layer’, ‘input gate layer’, and ‘output gate layer’.

1. Forget gate

The first is that LSTM decides to let that information continue to pass through this cell, which is achieved through a sigmoid neural layer called "forget gate layer". Its input is $h_{t-1}$ and $x_t$, and the output is a vector with a value between 0 and 1 (the length of the vector is the same as the state of the cell $C_{t-1}$), which denotes that the proportion of passed information of each part of $C_{t-1}$. 0 means ‘don't let any information pass’, and 1 means ‘let all information pass’.

Back to the load forecasting model, here we need to predict the load for the next time period based on all recent load information. In this case, the current load information should be included in the state of each cell. But when forecasting the load on Saturday, it should forget the load on Friday, because the load pattern on Friday and Saturday is very different, and the load pattern on Saturday is similar to that of last Saturday.

![Figure 10 Forget gates](image)

2. Input gates

The next step is to decide how much new information to add to the cell state. This requires two steps: First, a sigmoid layer called ‘input gate layer’ determines which information needs to be updated; a tanh layer generates a vector, which is the alternative content for updating, $C_t$. In the next step, the two parts are combined here to update the state of the cell.

![Figure 11 Input gates](image)
In the load forecasting example here, we want to add new Saturday information to the cell state to replace the old state information. With the above structure, the cell state can be updated here, that is, \(C_{t-1}\) is updated to \(C_t\). It should be clear from the structure diagram: First, multiply the old state \(C_{t-1}\) with \(f_t\), and forget some information that it don't need to keep; then add it with \(i_t\)\(\tilde{C}_t\). This part of the information is the new content to be added here.

**Figure 12** Update cell status

3. Output gates

Finally, here is the need to decide what value to output. This output mainly depends on the state of the cell \(C_t\), but not only on \(C_t\), which needs to go through a filtering process. First of all, a sigmoid layer is still used (calculated) to determine which part of the \(C_t\) information will be output. Next, \(C_t\) pass through a tanh layer (return the values to between -1 and 1), and then multiply the output of the tanh layer and the weight calculated by the sigmoid layer to get the final output result.

In the load forecasting example, suppose that the model here has just touched the information on Saturday, and then the load pattern on Saturday will be output. Then it is need to add all the information it just learned about Saturday to the cell state to make correct predictions.

**Figure 13** cell output

Aiming at the problem of medium and long-term load forecasting, the number of input unit layers of LSTM is designed here as 5 (load data of the same month in the past 5 years), and the number of output unit layers is 1 (forecasting the load of the year and month).

3. Case analysis

Using the electricity consumption data of the whole society from January 2004 to December 2018 in a province as input, the trained model is used to predict the electricity consumption data from January to December 2019. The prediction results are shown in the following figure:

**Figure 14** Forecast results of electricity consumption in a province from January to December 2019 (Long- and Short-Term Memory Neural Network, LSTM)
The comparison between forecast data and actual data is shown in the following table:

Table 1  Forecast results of electricity consumption in a province from January to December 2019
(Long- and Short-Term Memory Neural Network, LSTM)

| Time     | Predicted value (10^4 kWh) | Actual value (10^4 kWh) | Forecast accuracy (%) |
|----------|----------------------------|------------------------|-----------------------|
| 2019-12  | 1835836.36                 | 2160191                | 84.98                 |
| 2019-11  | 1732316.24                 | 1950864                | 88.8                  |
| 2019-10  | 1809389.97                 | 1781873                | 98.46                 |
| 2019-09  | 1930019.96                 | 1728784                | 88.36                 |
| 2019-08  | 1948104.41                 | 1874086                | 96.05                 |
| 2019-07  | 1794701.48                 | 1972117                | 91                    |
| 2019-06  | 1718393.95                 | 1865217                | 92.13                 |
| 2019-05  | 1715204.64                 | 1848608                | 92.78                 |
| 2019-04  | 1834867.53                 | 1800463                | 98.09                 |
| 2019-03  | 1973810.81                 | 1840647                | 92.77                 |
| 2019-02  | 2106842.59                 | 1715390                | 77.18                 |
| 2019-01  | 2008961.47                 | 2080795                | 96.55                 |
| Average value             |                          |                        | 91.43                 |
| Cumulative value          | 22408449.41              | 22619035               | 99.07                 |

Analyzing the above data, the cumulative annual forecast accuracy of this method is as high as 99.07%. This result shows that this method has a good grasp of the total development trend from January to December, and solves the problem that a single-layer neural network model cannot capture the long-term trend of the data and short-term fluctuation characteristics at the same time, therefore it achieves a higher accuracy in macro forecast results.

4. Conclusion
This paper has thoroughly studied the basic principles of recurrent neural networks and their shortcomings that they cannot solve the problem of long-term dependence, and proposes a power system load forecasting method based on LSTM networks. The results of case analysis show that, by using two built-in long-term and short-term memory state lines, this method solves the problem of the single-layer neural network model in the time series forecasting problem that it is difficult to capture the long-term development trend and short-term fluctuation characteristics of the data simultaneously, which effectively improves the forecasting accuracy of power system load.

Acknowledgements
This work was supported by Guangxi Power Grid Dispatch and Control Center (Grant No.: 046000KK52200025).

References
[1]  Li Jin, Liu Jinpeng, Wang Jianjun. Medium and long-term load forecasting method using support vector machine and simulated annealing algorithm. Proceedings of the Chinese Society of Electrical Engineering 31, 63-66 (2011)
[2]  Zhou Huan, Ren Haijun, Li Jian, Zhang Yun, Zhou Yongyong, Sun Caixin, et al. Medium and long-term power load variable weight combined forecasting method under hierarchical structure. Proceedings of the Chinese Society of Electrical Engineering 30, 47-52 (2010)
[3] Wang Ning, Xie Min, Deng Jialiang, Liu Mingbo, Li Jialong, Wang Yi, et al. Medium and long-term cooling load forecast based on support vector machine regression combined model. Power system protection and control 44, 92-97 (2016)

[4] Wu Yaowu, Lou Suhua, Lu Siyu, Qiao Hui, Kang Futian. Medium and long-term load forecasting method based on improved D-S evidence theory. Acta Electrotechnical Society 27, 157-62 (2012)

[5] Zhang Fan, Zhang Feng, Zhang Shiwen. Power load forecasting based on time series analysis method of lifting wavelet. Electrical Automation 39, 72-76 (2017)

[6] Yang Nan, Li Hongsheng, Yuan Jingyan, Li Suoya, Wang Xuan. Medium and long-term load gray forecasting method considering gray correlation analysis. Journal of Electric Power System and Automation 30, 108-14 (2018)

[7] W. Z, D. N. A mid-long term load forecasting model based on improved grey theory, 2010 2nd IEEE International Conference on Information Management and Engineering pp, 633-35 (2010)

[8] Y. S, H. Y, Y. D, N. P. Research on Long Term Load Forecasting Based on Improved Genetic Neural Network, 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application pp, 80-84 (2008)

[9] Xinran L, Lide W, Peiqiang L. The study on composite load model structure of artificial neural network, 2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies pp, 1564-70 (2008)