Medium- and long-term electric power demand forecasting based on the big data of smart city

Zhanmeng Wei¹, Xiyuan Li¹, Xizhong Li¹, Qinghe Hu², Haiyang Zhang³ and Pengjie Cui⁴

¹ National Grid Yingkou Company, East 40, Bohai Street, Zhanqian District, Yingkou 115000, China.
² Corresponding Writer. Northeastern University, No. 3-11, Wenhua Road, Heping District, Shenyang 110819, China. Email: huqinghe@ise.neu.edu.cn
³ Northeastern University, No. 3-11, Wenhua Road, Heping District, Shenyang 110819, China.
⁴ Software College, Northeastern University, No.195, Chuangxin Road, Hunnan District, Shenyang, 110169, China.

Abstract. Based on the smart city, this paper proposed a new electric power demand forecasting model, which integrates external data such as meteorological information, geographic information, population information, enterprise information and economic information into the big database, and uses an improved algorithm to analyse the electric power demand and provide decision support for decision makers. The data mining technology is used to synthesize kinds of information, and the information of electric power customers is analysed optimally. The scientific forecasting is made based on the trend of electricity demand, and a smart city in north-eastern China is taken as a sample.

1. Introduction

Electric power demand forecasting plays an important role in the operation, planning and control of electric power system. It is an important part to ensure the safe operation of power system, to meet the scientific management and the same scheduling of the joint power grid, and to plan the commercial operation. However, the traditional prediction introduces fewer factors as the variables [1], by considering only the power system’s internal data, such as the maximum load, the average regional electricity consumption, and by using the traditional data mining and statistical methods, based on limited sampling data, proprietary data in the field, such as marketing systems, distribution management systems, production management systems, energy management systems, equipment inspection and monitoring systems, customer service systems and so on. Therefore, the prediction accuracy of power information is very limited, and it is urgent to propose a new approach to forecast power demand. That’s to say data preparation has become particularly important.

The development of smart city brings opportunities for cross-sectoral data acquisition. Powerful data analysis based on smart city, integrates both internal data and external data of power enterprises, especially government based public data. By exploring the data association and using as many categories of data as possible, analysis based on big data and a simple algorithm can be more effective than a small number of data and a complex algorithm. We included types of data, and brought meteorological information, geographic information [2], population information [3], business information, economic information [4-7] and other external data into the big database.
Based on the big data inside power supply companies and inside the smart city, we devised an improved algorithm to forecast the medium and long term power demand. Simulation results proved that the big data of smart city can lead to more accurate power demand forecasting.

2. Research Content

(1) Power Big Data Analysis Model

When electric power information is forecasted, the focus is shifted from complex algorithms to big data preparation. Big data and a simple algorithm is more effective than a small amount of data and a complex algorithm, so data preparation becomes particularly important, i.e. to absorb as many data indicators as possible, to extract the correlation between the data characteristics, and to predict the target data based on the associated data. The larger the total amount of data, the more categories of data, the more comprehensive the data feature, the more accurate the analysis is. For example, the annual electricity consumption, if we include the economic factors in to the short-term load forecasting to consider meteorological factors, the accuracy of prediction degree can be greatly improved. Table 1 shows the comparison of the two analysis approaches.

| Comparison Item | Analysis based on traditional power supply data [8] | Analysis based on big data of the smart city |
|-----------------|----------------------------------------------------|---------------------------------------------|
| Data amount     | very limited                                      | huge                                       |
| Data sources    | power enterprise production data only              | from the public basic database of both electric power enterprises and government departments |
| Variable        | power grid operation and equipment testing or monitoring data, power enterprise marketing data, etc. | The variables used by traditional data analysis methods, the type of regional population, weather meteorology, national policy, industrial legal persons, large customers, etc. |
| Algorithm       | traditional algorithms and intelligent algorithms | correlation analysis and feature extraction |
| Algorithm rate  | large convergence rate, long computation time     | mergers and acquisitions of the data can speed up rate |
| Algorithm accuracy | accuracy of the original data not high          | may fall into the local minimum             |

(2) By extracting the characteristic relationship between the total electricity consumption and the population, the GDP, the per capita GDP, the number of industrial enterprises above designated size and the total amount of the import and export, and so on. Combined with the internal data of the power enterprises, medium and long term electric demand forecasting model is established.

Data type is a big feature of the big data. Power consumption is affected by multiple factors, so we took as many data sources and data categories as possible as associated objects, and forecast power information by extracting associated features. We made full use of the big data from the smart city basic database outside power enterprises, and did the demand forecasting and evaluation basing on the variation of the associated objects and features.

3. Algorithm Implementation

3.1. Functional Structure Modeling

According to this paper, a model of the functional structure which relies on basic database of smart city is established. Firstly, a data database is built, then the relevant data are extracted, the data cleaning and format conversion is carried out. Finally, the mining results are analysed. The effectiveness of customer analysis depends heavily on the selection and measurement of customer evaluation criteria. The mining method aiming for the actual needs is the key of the whole model.
Power consumption is affected by many factors. This paper not only takes into account the internal data of electric power enterprises but also makes full use of the information provided by the smart city database. The specific names of the data indicators are as follows [9].

1. The speed of economic development (GDP)
2. Changes in industrial structure (adjustment of industrial structure, the proportion of high energy consuming industries in industry)
3. Population and living standards (total population, living standards of residents)
4. Temperature and climate change (temperature, precipitation, natural disaster and so on)
5. Demand management measures (shift peak load Valley, energy storage equipment and so on)
6. Prices of other alternative energy (It mainly refers to the price of natural gas)
7. Policy factors (macro industrial structure adjustment, energy saving)
8. Supply factors (power supply capacity and wind power integration)
9. Large number of complex data information.

The architecture of the forecasting model is shown in Fig. 1.

3.2. Elman Neural Network Structure
We include an undertake layer to record the input of middle layer (i.e. hidden layer) to increase the sensitivity to historic data on purpose. The adding of internal feedback network increases the network’s ability of self-processing dynamic information. Thus, a dynamic model is achieved.

The proposed Elman neural network structure is as shown in Fig. 2.
3.3. Improved Algorithms

3.3.1. Improved Learning Algorithm. In the process of learning, the traditional Elman neural network often shocks and converges slowly, and falls into local optimal solution. In order to solve the above problems, the nonlinear damping least square method is used to optimize it. The nonlinear damping least square method is an improved method based on quasi-Newton method. The equation of weights and thresholds of the network is as follows.

\[ W(k+1) = W(k) - [J^T J + \mu I]^{-1} J^T e \]  \hspace{1cm} (1)

Wherein, \( W \) is the network parameter set that will be adjusted, and \( k \) is the times of iteration. Where \( J \) is the Jacobi matrix of the error function and contains the weights and threshold derivative. Wherein, \( \mu \) is the learning rate, and \( I \) is the unit matrix. And \( e \) is the error vector of network. When \( \mu \) is 0, equation (2) is quasi-Newton method. When \( \mu \) is larger, it is gradient descent with smaller stride. The quasi-Newton method can approach the minimum error quickly and accurately. And \( \mu \) decreases continuously after each iteration. Then the algorithm is close to Quasi-Newton. If error performance increases after iteration, \( \mu \) should be enlarged. Therefore, the performance of the iterative error has been reduced.

3.3.2. Improved Excitation Function. The excitation function [10] of Elman neural network model is generally a sigmoid function \( f(x) = \frac{1}{1 + \exp(-h(x + n))} \). However, it is easy to have the model to converge slowly and to fall into local minimum. The improved function is as follows. Its function is as equation (2).

\[ f(x) = m + \frac{1}{1 + \exp(-h(x + n))} \]  \hspace{1cm} (2)

The derived function of equation (2) is as equation (3).

\[ f'(x) = \frac{h[\exp(-h(x + n))]}{1 + \exp(-h(x + n))^2} \]
\[ = h \left[ \frac{1}{1 + \exp(-h(x + n))} - \frac{1}{1 + \exp(-h(x + n))^2} \right] \]
\[ = -h \left[ \frac{1}{1 + \exp(-h(x + n))} - \frac{1}{2} \right] + \frac{h}{4} \]  \hspace{1cm} (3)

The authors get equation (4) from equation (2).

\[ \frac{1}{1 + \exp(-h(x + n))} = f(x) - m \]  \hspace{1cm} (4)

By substituting equation (4) into equation (3), the authors get equation (5).
Wherein, \( m, n \) is a constant and \( h \) is a slope. Set \( m, n \) value, so that the function moves vertically and horizontally. In the model, the values of \( m, n \) and \( h \) are corrected with the error function. When the function value \( f(x) \) is closer to \( m+1/2 \), the larger the derivative value is, the faster the function converges. In addition, the value of \( f'(x) \) is positively related to the learning rate. And the higher the value is, the faster the learning rate is. By adjusting the \( m, n \) and \( h \) values, the convergence speed and prediction accuracy of the model can be optimized.

4. Case Analysis

In order to verify the effectiveness of the proposed methods of load forecasting, the power grid from 2006 to 2015 in a certain region of the historical annual electricity sales data is used, as shown in Table 2. It shows that over the past years the electricity sales data has an overall growth trend, but a fluctuation in 2009 which has a great influence on the load forecasting precision. This paper mainly considers the load forecasting of the coming month to the coming several years which is medium- and long-term electric power demand forecasting for dispatching department and planning department. In Table 2, years from 2006 to 2015 are numbered by 1~10, and the corresponding sale data is taken as the original input of the model [11].

Based on the traditional data of the internal data of the electric power, the Elman neural network prediction model of the large and medium data of the smart city is selected and the improved data of the internal and external data of the smart city are obtained by using the Mathlab with raw data from year 2006 to 2015 as the input. Elman neural network prediction model is used to fit the load series from year 2006 to 2015, and the three models are used to predict the data of year 2014 and 2015, and the historical sales volume in year 2014 and 2015 is taken as the model testing samples.

The load model of year 2006 to 2015 is normalized by the established prediction model. The fitting results and the relative error ratios are shown in Table 3.

| Serial number | Particular year | Electricity sales/(TWh) | Serial number | Particular year | Electricity sales/(TWh) |
|---------------|----------------|-------------------------|---------------|----------------|-------------------------|
| 1             | 2006           | 26                      | 6             | 2011           | 37                      |
| 2             | 2007           | 27                      | 7             | 2012           | 46                      |
| 3             | 2008           | 31                      | 8             | 2013           | 52                      |
| 4             | 2009           | 32                      | 9             | 2014           | 54                      |
| 5             | 2010           | 34                      | 10            | 2015           | 56                      |

The results of the analysis show that the maximum prediction error ratio of the traditional prediction model based on the data of internal power system is 12.91%, and the minimum error ratio is 4.01%. As for Elman neural network model based on big data of smart city, the maximum error ratio is 11.78%, and the minimum error ratio is 1.93%. As for the proposed Elman neural network model based on big data of smart city, the maximum error ratio is 6.70%, the minimum error ratio is 1%. It can be concluded that it is possible to achieve better power forecasting by referring to the big data inside and outside the power company, and the improved Elman neural network model has higher prediction accuracy.
Table 3. Comparison between the forecasting data and actual data

| Serial number | Actual charge value (TWh) | Traditional forecasting model based on internal data of power company | Elman neural network model based on big data of smart city | Improved Elman neural network model based on big data of smart city |
|---------------|---------------------------|---------------------------------------------------------------------|-----------------------------------------------------------|-----------------------------------------------------------------|
|               |                           | calculated value (TWh) | relative error ratio (%) | calculated value (TWh) | relative error ratio (%) | calculated value (TWh) | relative error ratio (%) |
| 1             | 26                        | 26.1558                | 3.13                     | 26.5483                | 1.67                     | 26.5483                | 1.67                     |
| 2             | 27                        | 24.7765                | 8.24                     | 26.1558                | 3.13                     | 26.5483                | 1.67                     |
| 3             | 31                        | 26.8980                | 12.91                    | 27.3493                | 11.78                    | 28.9244                | 6.70                     |
| 4             | 32                        | 34.3442                | 7.32                     | 34.1335                | 6.67                     | 34.1335                | 6.67                     |
| 5             | 34                        | 35.0225                | 3.01                     | 34.8756                | 2.58                     | 34.8756                | 2.58                     |
| 6             | 37                        | 39.4652                | 4.50                     | 39.8243                | 5.67                     | 39.8243                | 5.67                     |
| 7             | 46                        | 46.8871                | 1.93                     | 46.4554                | 1.00                     | 46.4554                | 1.00                     |
| 8             | 52                        | 53.1062                | 2.13                     | 52.9876                | 1.90                     | 52.9876                | 1.90                     |

The number of neurons in the input and output layers of the neural network is determined by the problem itself. In this paper, the number of neurons in the Elman neural network is set as 8. And in the experiment, the neurons in the hidden layer are gradually increased and decreased. The count of the units in hidden layer is 9, and the count of neurons in the output layer is 1. With the original load data as the output data $T$, the output layer uses Logsig as the output function, the training function is Trainlm, and the training is 1000-round. The predicted load data by the forecasting model is input into the training-ready network, and the simulation result is just the forecasting result, as shown in Table 4.

Table 4. Prediction results

| Year | Actual charge value (TWh) | Traditional forecasting model based on internal data of power company | Elman neural network model based on big data of smart city | Improved Elman neural network model based on big data of smart city |
|------|---------------------------|---------------------------------------------------------------------|-----------------------------------------------------------|-----------------------------------------------------------------|
|      |                           | calculated value (TWh) | relative error (TWh) | calculated value (TWh) | relative error (TWh) | calculated value (TWh) | relative error (TWh) |
| 2014 | 54                        | 59.8034                | 10.75                  | 57.9767                | 7.36                     | 55.3216                | 2.45                     |
| 2015 | 56                        | 61.0043                | 8.94                   | 58.1062                | 3.76                     | 57.4061                | 2.51                     |

5. Conclusion
Based on the big data of smart city, the authors include more indicators and consider the relationship among the data. Compared with traditional forecasting, the proposed algorithm is better on prediction accuracy and stability. And, by using improved Elman neural network prediction model, the proposed model has better network computing, time-varying adaptability, error controlling and network learning ability. The forecasting results of a practical case proved both the effectiveness of the smart city data for medium- and long-term power demand forecasting and the availability and superiority of the improved algorithm.

6. References
[1] Muhammad Qamar Raza, Zuhairi Baharudin, Perumal Nallagownden, et al. A Comparative Analysis of PSO and LM Based NN Short Term Load Forecast with Exogenous Variables for Smart Power Generation [C]. 2014 5th International Conference on Intelligent and Advanced Systems (ICIAS), 2014, 1-6.
[2] Franklin L. Quilumba, Wei-Jen Lee, Heng Huang, et al. Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behavior Similarities [J], IEEE Transactions on Smart Grid, 2014, 6(2): 911-918.
[3] Shu Fan, Methapraynoon K, Wei-Jen Lee. Multiregion Load Forecasting for System with Large Geographical Area [J]. Industry Applications, IEEE Transactions on, 2009, 45(4): 1452-1459.
[4] Lee Chien-Chiang, Chang Chun-Ping. Structural Breaks. Energy Consumption and Economic Growth Revisited: Evidence from Taiwan [J]. Energy Economics, 2005, 27(6): 857-872.

[5] Paresh Kumar Narayan, Russell Smyth. Energy Consumption and Real GDP in G7 Countries: New Evidence from Panel Co-Integration with Structural Breaks [J]. Energy Economics, 2008, 30(5): 2331-2341.

[6] Nicholas Apergis, James E. Payne. Energy Consumption and Economic Growth in Central America: Evidence from a Panel Co-Integration and Error Correction model [J]. Energy Economics, 2009, 31(2): 211-216.

[7] Hwang D B K, Gum B. The Casual Relationship between Energy and GNP: the Case of Taiwan [J]. The Journal of Energy and Development, 1992, 16: 219-226.

[8] Wu Jinwei. Big Data Technology in the Application of Electricity Demand Analysis [D]. Jiangsu: Southeast University, 1-65 (2016).

[9] ZHANG Ai-hui. Study on Medium and Long Term Electric Demand Forecasting Based on Multi-index Model [D]. Hunan: Changsha University of Science and Technology, 1-52 (2013).

[10] Chen Ya, Li Ping. Study on Short-Term Load Forecasting Based on Neural Network [J]. Electrical Technology, (1): 26-29 (2017).

[11] Zhang Jianmei, Zhou Shuxiang, Lin Nan, Zhang Qin, Chen Jie. Gray Elman Neural Network Medium and Long Term Load Forecast [J], 25 (4): 146-149 (2013).