An Integrated Forecasting Model for Electricity Demand in a Power System Considering Time Difference between Criteria

Huijuan Huo 1, *, Yijiang Liu 1, Tiannan Ma 2, Hongjin Wang 1, Yanchao Lu 1
1 State Grid Economic and Technological Research Institute co. ltd., Beijing, 102209, China
2 State Grid Sichuan Electric Power Company, Sichuan, 61000, China
*Corresponding author e-mail: huohuijuan@chinasperi.sgcc.com.cn

Abstract. As China's economy enters a stage of high-quality development, the trend of the power demand has changed and can affect the investment and construction of power industry. A series of key criteria impacting on power demand were sorter out to build a power demand index system currently. The time difference correlation analysis was applied to classify these criteria as precursor, consistency and lag index. The power demand index based on the three categories of criteria were fitted based on the multivariate statistical analysis. Then the comprehensive power demand index was assembled to forecast the electric power demand from 2020 to 2022 in China based on the current policy and economic situation.

Keywords: Electricity power demand, forecast model, Time difference correlation, power demand criteria classification.

1. Introduction

Electric power industry is closely related to various industrial sectors. For Electric power enterprises, it is of great significance to improve the accuracy and the scientific nature of regional electric power demand forecast. The macro-economic operation has a certain directional effect on power demand. Because of the new normal of China's economy, the growth rate of electricity consumption has slowed down. The implementations of the electricity substitution strategy and energy market reform can impact on the power demand structure. In order to mine information related to power demand changes from external macro and internal micro indicators, a lot of data mining technology based on statistical methods is widely used. Mohanad et al. [1] developed an artificial neural network (ANN) model combining the forecasts of multivariate adaptive regression spline (MARS), multiple linear regression (MLR), and autoregressive integrated moving average (ARIMA) models to forecast energy demand. Ranran et al. [2] proposed a novel hybrid model based Adaptive Fourier decomposition method to forecast the power demand. Sine cosine algorithm was introduced to eliminate seasonality influences. Wang T et al. [3] designed the improved BP wavelet neural network method (IBPWNN) to forecast the trend of electricity demand. The improved can be used to Improved prediction accuracy compared with the traditional BP neural network in the electricity demand prediction. Suhono S [4] developed a hybrid mode based on LEAP (Long-range Energy Alternative Planning System) and carried out a case study
of Sumatera systems. In addition, Stochastic error correlation model, cointegration theory, data mining techniques etc. was introduced to solve the power demand forecasting problems [5-7].

Power demand forecasting is a very important and complex work. On the one hand, power demand is affected by many macro factors such as economy, climate and population, and micro factors such as power grid infrastructure construction and power supply construction. On the other hand, there are many methods to be used in power demand forecasting, and choosing different forecasting methods according to different conditions will have a great impact on the accuracy of prediction results. In the literature above, the differences in time dimension of various influencing factors are ignored in these prediction methods.

Thus, the time difference between criteria was considered to improve prediction accuracy in the paper. A new integrated forecasting model considering the time difference was established to predict the electric power demand. In section 2, some criteria from macro and micro factors were collected to assemble a power demand forecasting index system. In section 3, time difference correlation analysis was introduced to divide the indicators into precursor indicators, consistency indicators and lag indicators. In section 4, a comprehensive power demand index model is established by multiple regression analysis to predict the electricity demand. In section 5, an example in China is implement to verify the integrated forecasting model. The last section made the conclusions.

2. An index system affecting power demand

According to four principles of comprehensiveness, practicality, convenience and sensitivity, a series of criteria affecting electricity demand for power system can be extracted from the external and internal aspects. Some experts were collected from macroeconomic research, policy analysis, power industry and other fields to screen out important criteria. Based on a Delphi method, important criteria impacting on power demand can be integrated into an index system, as shown in the Figure 1. The data of the criteria were sorted out based on the statistics yearbook of China and the data released by China Power Enterprise Federation from 2003 to 2018.

![Figure 1. The index system affecting power demand](image-url)
3. Power Demand Criterion attribute classification considering time difference

3.1. Methods

Time difference correlation analysis is a common method to verify the precedence, consistency and lag of economic correlation series by correlation coefficient. The method can be introduced to analyze the attribute of these key criteria. A benchmark indicator should be chosen firstly to reflect current economic activity. Make the selected index lead or lag some periods, the correlation coefficient of two criteria can be calculated as the time difference correlation coefficient.

Suppose \( y \) is the benchmark index, and the value is \( \{y_1, y_2, y_3…y_n;\} \), \( x \) is a selected criterion, and the value is \( \{x_1,x_2, x_3…x_n\} \), the value of the time difference correlation coefficient is:

\[
r = \frac{\sum_{i=-L}^{L} (x_{i} - \bar{x})(y_{i} - \bar{y})}{\sqrt{\sum_{i=-L}^{L} (x_{i} - \bar{x})^2 \sum_{i=-L}^{L} (y_{i} - \bar{y})^2}}
\]

where \( i \) refers to the time difference or delay number, \( i \) is lead time for negative value and lag time for positive value. \( L \) refers to maximum delay number; and \( n \) refers to number of criterion data. \( i = 0, \pm 1, \pm 2L, \pm L \). The larger the \( r \) value, the stronger the correlation between two criteria.

3.2. Criterion attribute classification

The criteria in the index system can be divided into three categories: precursor indicators, consistency indicators and lag indicators. First, electricity consumption of the whole society in a region can be regarded as the benchmark indicator. Then the calculation results based on the time difference correlation analysis are listed in the table 1. The lag period corresponding to the maximum value of the correlation is the best.

**Table 1.** The calculations of the time difference correlation between electricity consumption of the whole society and key criteria

| Lag period | -4 | -3 | -2 | -1 | 0  | 1  | 2  | 3  | 4  | Best period |
|------------|----|----|----|----|----|----|----|----|----|-------------|
| GDP per capita | 0.482 | 0.617 | 0.761 | 0.883 | 0.994 | 0.88 | 0.724 | 0.537 | 0.419 | 0           |
| Tertiary industry added value | 0.518 | 0.634 | 0.764 | 0.874 | 0.98 | 0.865 | 0.702 | 0.497 | 0.372 | 0           |
| Proportion of tertiary industry added value | 0.307 | 0.394 | 0.486 | 0.532 | 0.602 | 0.692 | 0.776 | 0.891 | 0.775 | 3           |
| Residents’ disposable income | 0.631 | 0.766 | 0.88 | 0.986 | 0.871 | 0.791 | 0.712 | 0.601 | 0.514 | -1          |
| Regional total exports | 0.506 | 0.693 | 0.856 | 0.98 | 0.863 | 0.721 | 0.632 | 0.589 | 0.503 | -1          |
| Per capita electricity consumption | 0.507 | 0.58 | 0.746 | 0.882 | 0.999 | 0.887 | 0.747 | 0.578 | 0.421 | 0           |
| Carbon emissions | 0.414 | 0.632 | 0.808 | 0.934 | 0.849 | 0.753 | 0.65 | 0.571 | 0.376 | 1           |
| Length of grid line | 0.613 | 0.76 | 0.885 | 0.995 | 0.883 | 0.734 | 0.629 | 0.552 | 0.413 | -1          |
| Capacity of power generation equipment | 0.615 | 0.756 | 0.875 | 0.986 | 0.792 | 0.659 | 0.717 | 0.517 | 0.358 | -1          |
| Electricity sales | 0.392 | 0.56 | 0.732 | 0.874 | 0.995 | 0.715 | 0.739 | 0.582 | 0.472 | 0           |
| Asset liability ratio | -1.296 | -0.57 | 0.038 | 0.297 | 0.462 | 0.555 | 0.478 | 0.386 | 0.311 | 1           |
If the best lag time is 0, the criteria belongs to consistency indicators. If the best lag time is negative, the criteria belongs to precursory indicators. If the best lag time is positive, the criteria belongs to lag indicators. Thus, GDP per capita, tertiary industry added value, per capita electricity consumption, electricity sales belongs to consistency indicators. Residents' disposable income, regional total exports, length of grid line, capacity of power generation equipment are classified into precursory indicators. Proportion of tertiary industry added value, carbon emissions and asset liability ratio are classified into lag indicators.

4. Power Demand Index Construction and Forecast Analysis

4.1. Power demand index model

Based on the above classification, the indices of the precursory, consistent and lag for power demand can be constructed respectively by using the theory of multivariate statistical analysis.

There are four criteria in the precursory indicators. And residents' disposable income and regional total exports belongs to external factors, length of grid line and capacity of power generation equipment belongs to internal factors. According to the precursory period of the criteria, these data should be treated with lag. The multicollinearity of these precursory indicators and the correlation with the electricity consumption of the whole society were implemented by SPSS 22.0 software. The results of the D-W values are close to 2. Therefore, there is no Multicollinearity between the criteria. Additionally, the correlation test results show that there is a linear relationship between the factors and the electricity consumption. Then, the power demand based on external precursory factors: 

\[ Y^{pe} = 1.105X_1 + 0.193X_2 + 6053.7 \]  \( (2) \)

\[ Y^{pe} = 265.932X_3 + 0.181X_4 - 10004.369 \]  \( (3) \)

Where \( X_1 \) is residents' disposable income, \( X_2 \) is regional total exports, \( X_3 \) is length of grid line and \( X_4 \) is capacity of power generation equipment.

The contribution of macro-factors to power demand is significantly higher than that of internal micro-factors. Then, the \( Y^{pe} \) and \( Y^{pi} \) can be assembled according to 7:3 weight, which is determined by the experts. And the power demand based on the all precursory factors \( Y^p \) is:

\[ Y^p = 0.7735X_1 + 0.1351X_2 + 79.7796X_3 + 0.0543X_4 + 1236.2793 \]  \( (4) \)

The power demand based on the all consistent factors \( Y^c \) and the power demand based on the all lag factors \( Y^l \) can be fitted by repeating the above steps.

\[ Y^c = 0.392X_5 - 0.021X_6 + 6.307X_7 + 0.2985X_8 + 1191.4 \]  \( (5) \)

\[ Y^l = 1233.799X_9 + 484.125X_{10} + 3441.077X_{11} - 82706.791 \]  \( (6) \)

Where \( X_5 \) is GDP per capita, \( X_6 \) is tertiary industry added value, \( X_7 \) is per capita electricity consumption, \( X_8 \) is electricity sales, \( X_9 \) is proportion of tertiary industry added value, \( X_{10} \) is carbon emissions and \( X_{11} \) is asset liability ratio.

Based on the least squares, the power demand values based on the above three types can be fitted as a comprehensive power demand index, as is:

\[ Y = 0.15Y^{pe} + 2.117Y^{pi} + 1.381Y^c - 759.7 \]  \( (7) \)

4.2. Power demand forecast analysis

First, the error between the predicted value and the actual value for the demand from 2015 to 2018 is analyzed to test the prediction accuracy of the power demand index model, as shown in the Table 2. The fitting residual of the model is white noise. The trend of the fitting result is consistent with the actual value, as shown in Figure 1. The average deviation rate is 1.005%, which shows that the fitting effect of the model is very good. Thus, the power demand from 2020 to 2022 can be forecasted by using the model.
Table 2. Power demand forecast deviation from 2015 to 2018

| Year | 2015     | 2016     | 2017     | 2018     |
|------|----------|----------|----------|----------|
| The actual value (Billion Kwh) | 57930.12 | 61393.25 | 62155.54 | 69117.95 |
| The predicted value (Billion Kwh) | 56933    | 59747    | 63077    | 68449    |
| Deviation rate (%) | 1.75     | 2.76     | 2.47     | 0.98     |
| Average deviation rate (%) | 1.005    |          |          |          |

Figure 2. Prediction model fitting diagram

Then, based on the current policy and economic situation, the criteria data in the forecast model were predicted from 2020 to 2022 based on the trend extrapolation method without considering emergencies. Then, the power demand from 2020 to 2022 can be predicted by applying the power demand index model. The results are shown in the Table 3. The predictive power demand from 2020 to 2022 are 74745.37 billion kWh, 78076.74 billion kWh and 81413.31 billion kWh respectively.

Table 3. The prediction of power demand and relative criteria from 2020 to 2022

| Criteria | 2020     | 2021     | 2022     |
|----------|----------|----------|----------|
| GDP per capita (yuan) | 69021.9  | 72667.4  | 76312.9  |
| Regional total exports (Billion yuan) | 163261.48 | 165221.22 | 166587  |
| Tertiary industry added value (Billion yuan) | 398340.55 | 419643.58 | 441061.2 |
| Proportion of tertiary industry added value (%) | 35.9      | 35.8     | 35.7     |
| Residents' disposable income (yuan) | 33388.98 | 36013.59 | 38742.1  |
| Electricity sales (Billion kWh) | 46592.7  | 48658    | 50723.3  |
| Per capita electricity consumption (kWh) | 5412.68  | 5642.34  | 5872     |
| Carbon emissions (Billion tons) | 94.119   | 90.613   | 86.379   |
| Capacity of power generation equipment | 169174.36 | 180940.99 | 193115  |
| Length of grid line | 112.5784 | 117.5847 | 122.591  |
| Asset liability ratio (%) | 56.3      | 55.8     | 55.2     |
| Prediction results | Power demand (Billion kWh) | 74745.37 | 78076.74 | 81413.31 |
All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

5. Conclusion
Power demand prediction is affected by a variety of macrocosm and microcosm factors, and is important to power grid planning and investment construction. A series of influencing criteria were sorted out and classified as precursor index, consistency index and lag index. Then the power demand based on the precursor index, consistent index and lag index were conducted respectively by multivariate statistical analysis. And a comprehensive power demand index was assembled based on least squares method to forecast the power demand from 2020 to 2022 in China. Based on the current policy and economic situation and ignoring emergencies, the predictions of power demand from 2020 to 2022 are 74745.37 billion kWh, 78076.74 billion kWh and 81413.31 billion kWh.

Acknowledgments
This study is supported by the science and technology project of State Grid corporation of China (An Empirical Study on Key Technologies of Power Grid Investment Strategy and "One province, one policy" under the Situation of Development and Reform, B3441018K003)

References
[1] Mohanad S. AL-Musaylh, Ravinesh C. Deo, Jan F. Adamowski, Yan Li. Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF Reanalysis atmospheric predictors in southeast Queensland, Australia[J]. Renewable and Sustainable Energy Reviews, 2019, 113.
[2] Ranran Li, Ping Jiang, Hufang Yang, Chen Li. A novel hybrid forecasting scheme for electricity demand time series[J]. Sustainable Cities and Society, 2020, 55.
[3] Wang T, Wang H, Zhang B, et al. Electricity demand forecasting based on improved wavelet neural network method[J]. JOURNAL OF INFORMATION & COMPUTATIONAL SCIENCE, 2010, 7(13): 2855-2861.
[4] Suhono S. Long-term electricity demand forecasting of Sumatera system based on electricity consumption intensity and Indonesia population projection 2010-2035[J]. Energy Procedia, 2015, 68: 455-462.
[5] Verdejo H, Awerkin A, Saavedra E, et al. Stochastic modeling to represent wind power generation and demand in electric power system based on real data[J]. Applied Energy, 2016, 173: 283-295.
[6] Zhao H, Yang L, Li C, et al. Research on prediction to investment demand of power grid based on co-integration theory and error correction model[J]. Power System Technology, 2011, 35(9): 193-198.
[7] Fan C, Xiao F, Wang S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques[J]. Applied Energy, 2014, 127: 1-10.