Early warning analysis of electricity sales based on multi-factor correlation analysis

CHEN Hai1,*, CUI Yan - yan1, MA li1, LIU Wei1, WANG Jun2, YUAN Shuai2, Guo Ju2, HUANG Yi2 and Dong Xiao – jing3

1China Electric Power Research Institute, Beijing City, China
2State Grid Corporation of China, Beijing City, China
3Economic and Technical Research Institute of Ningxia Electric Power Company of State Grid, Ningxia Province, China

Abstract. Under the background of the slowdown in macroeconomic growth and the gradual liberalization of the power system reforming market, the competition pressure of power grid companies in the electricity sales market has intensified, and the growth of power sales is not optimistic. It is necessary to conduct research and analysis of electricity sales. This paper conducts the analysis with the following steps: first, determines the leading, coincident, lagging economic indicators based on multi-factor correlation analysis, then synthesizes early warning index, forecasts electricity sales, finally, achieves early warning of external risks to improve the company’s management quality of the electricity sales.

1 introduction

In March 2015, "Several Opinions of the Central Committee of the Communist Party of China and the State Council on Further Deepening the Reform of the Power System" (Zhongfa [2015] No. 9 Document) was issued, which opened the prelude to the reform of the power system. The "Implementation Opinions on Promoting the Reform of the Power-Sold Side" in the six supporting documents pointed out that the construction of the power market will be accelerated and market competition will be introduced. The National Development and Reform Commission and the National Energy Bureau will issue a notice on the pilot reform of the incremental power distribution business. The document of the National Development and Reform Commission [2016] No. 2480) pointed out that the placing of electricity business was released to social capital, the competition in the electricity sales market of power grid companies intensified, and the risk of loss of customers was large. In addition, the growth rate of macroeconomic slowed down in recent years. The rapid growth of electricity consumption is unsustainable, and the growth rate of electricity sales by power grid companies is limited. The growth of electricity sales of the company is not optimistic.

In order to actively respond to the impact of power system reform and external macroeconomic changes caused to the company's electricity sales market, it is necessary to establish a complete external risk early warning analysis system based on external risk indicators, on the one hand to achieve early warning of power sales risks [11] On the other hand, it provides managers with decision support to improve management quality. Based on the multi-factor correlation analysis between electricity sales and external macroeconomic indicators, this paper conducts early warning analysis of electricity sales through the synthesis of early warning indicators.

2 Association analysis theory

According to different attributes of different factors, the correlation analysis is mainly divided into classification, sequencing and distance correlation analysis.

(1) Correlation analysis of two categorical variables

A categorical variable is a nominal variable, which means that the value of the variable is the name or symbol of the research object. Each value represents a category, and there is no difference in size or order between these values, which is equal. For the gender variable, the values are male and female.

In equation 1: \( f_{ym} \) is the number of modes of y distributions for each class of x, and \( F_{ym} \) is the number of modes for each classification of the variable y; n is the total number of times. In general, the \( \lambda_y \) coefficient is between 0 and 1, and a larger value indicates a higher degree of correlation.

\[
\lambda_y = \frac{\sum f_{ym} - F_{ym}}{n - F_{ym}}
\]

(1)

In equation 1: \( f_{ym} \) is the number of modes of y distributions for each class of x, and \( F_{ym} \) is the number of modes for each classification of the variable y; n is the total number of times. In general, the \( \lambda_y \) coefficient is between 0 and 1, and a larger value indicates a higher degree of correlation.

(2) Correlation coefficient between two ordered variables

The ordering variable is the level variable, and the variable value has the meaning of order. In other words, the value has a grade or order. Spearman and Kendall
The gray correlation analysis calculation steps are as follows:

a. Identify reference series that reflect system behavior characteristics and comparison series that affect system behavior.

b. Dimensionless processing of the reference series and the comparison series.

c. Find the gray correlation coefficient between the reference series and the comparison series $\xi(X_i)$.

If the reference series is: $X_0=[X_0(1), X_0(2), ..., X_0(n)]$, the evaluation series is $X_i=[X_i(1), X_i(2), ..., X_i(n)]$, $i=1, 2, ..., N$, the correlation coefficient between the evaluated series $X_i$ and the reference series $X_0$ can be defined as: the average value of the correlation coefficient between the evaluation type and the reference type and the reference type is called the correlation between the comparison series $X_i$ and the reference series $X_0$, as shown in equation 4.

$$
\xi(i) = \frac{\min \max |x_i(k) - x_0(k)| + \xi \max \min |x_i(k) - x_0(k)|}{\min \max |x_i(k) - x_0(k)| + \xi \max \min |x_i(k) - x_0(k)|}
$$

The gray correlation analysis calculation steps are as follows:

a. Identify reference series that reflect system behavior characteristics and comparison series that affect system behavior.

b. Dimensionless processing of the reference series and the comparison series.

c. Find the gray correlation coefficient between the reference series and the comparison series $\xi(X_i)$.

In this paper, while analyzing the correlation between electricity sales and external macroeconomic indicators, the data characteristics are taken into account. The sequence and distance method are used for correlation analysis.

3 Establish an early warning model

Based on the theory of multi factor correlation analysis and macroeconomic environment prosperity analysis, this paper uses time difference correlation analysis, Granger causality test, impulse response function and other metering methods to analyse the relationship between the selected external economic indicators and electricity demand first, then establishes an external environment warning model. The electricity industry early warning index is used to reflect the fluctuation of demand in the future electricity market, and to provide early warning of external risks in the power grid industry, as shown in Figure 1.
4 Case Analysis

4.1 Select external macroeconomic indicators

From the perspective of the close relationship with electricity sales, this paper first selects 45 indicators such as GDP, industrial added value, and industrial product output for subsequent correlation analysis, as shown in Table 1.

4.2 Correlation analysis

The correlation analysis between the selected 45 external macroeconomic indicators and the electricity sales was carried out one by one [4-7], and the analysis results are shown in Tables 2 to 4.

Table 1. Primary selection external index

| External index | External index | External index |
|----------------|----------------|----------------|
| GDP            | Asphalt production | Ten non-ferrous metal production |

Table 2. Coincident indicator results

| Coincident indicator | Correlation coefficient | Delay period |
|----------------------|--------------------------|--------------|
| Steel production     | 0.8367                   | -2           |
| Aluminum production  | 0.8107                   | -2           |
| Synthetic detergent production | 0.7518           | 1            |
| Electric hand-held tool production | 0.7302       | 0            |
| New construction area of commercial housing | 0.7238    | 0            |
| Cloth production     | 0.7189                   | -1           |
| Room air conditioner production | 0.6713      | 1            |

Table 3. Leading indicator results

| Leading indicator | Correlation coefficient | Delay period |
|-------------------|-------------------------|--------------|
| Household freezer production | 0.7942           | -3           |
The results show that the coincident indicators of electricity sales include 7 items such as steel production, aluminum production, synthetic detergent production, and electric hand-held tool production, etc. The leading indicators include 6 items such as household freezer production, metal cutting machine tool production, cement production, and vehicle production, etc. The lagging indicators include 2 items such as the commodity retail price index and the consumer price index.

### 4.3 Synthesize warning index

Step 1: According to the results of the above-mentioned electricity sales first, coincident, and lagging indicators, 15 indicators including industrial added value, steel production, aluminum production, and synthetic detergent production are selected as early warning indicators, and the monthly growth rate of each indicator is calculated.

Step 2: Assume that the year-on-year growth rate of each indicator is in a normal distribution. According to the average and variance of the growth rate of each indicator, the data is calculated to fall in the two-tailed and one-tailed range of 90%, 85%, and 50%. Determine the range of overheating (red light), partial heat (yellow light), stable (green light), partial cold (light blue light), and overcooling (blue light) of each indicator, as shown in Table 5.

Step 3: Take the ratio of the correlation coefficient between each indicator and the electricity sales to the sum of the total correlation coefficients as the weight. The value of the early warning indicator is obtained by multiplying the annual growth rate of each indicator in each month with the corresponding weight, as shown in Table 6 and Table 7.

Step 4: Correspond to the obtained warning value of each indicator and the determined indicator early warning index threshold value (as shown in 8), and obtain the warning value $A_i$ of each indicator, as shown in Table 9.

Step 5: Synthesize the monthly warning index

$$\text{Warning index} = \sum_{i=1}^{15} A_i W_i / 3 \times 100$$

Figure 2, and finally calculate the average value of the monthly warning index to get the total early warning index [8-10], as shown in Table 11.

It can be seen from Table 10 and Figure 2 that the monthly warning index of electricity sales fluctuated greatly. From September 2008 to December 2010, the early warning indicators changed from cold to overheating. After January 2012, the fluctuations tend to be flat, and the early warning index is in a stable state, but there is still a tendency for fluctuations to decline, and entered a colder state at the end of 15 years. The total synthetic early warning index is 100.5, which is in the stable green light interval. It shows that since January 2007, the average growth rate of electricity sales has been stable, and many places are in the 4-5% growth rate range.

### 4.4 Electricity sales warning result

According to the results of the above-mentioned early warning index and macroeconomic forecast, the forecast results of the early warning index on electricity sales will show a volatility trend in the sales of electricity in late 2017. The electricity consumption situation is relatively stable, that is, the risk of electricity sales is relatively small. As shown in figure 3.

### 5 Conclusion suggestion

In this paper, multi-factor correlation analysis is used to determine 15 macroeconomic indicators with high correlation coefficient with electricity sales, and then the early warning index is used to realize the early warning of electricity sales. From the early warning structure, the external macroeconomic indicators have been better since 2017, the electricity consumption situation is more optimistic, and the risk of selling electricity is small.

In view of the fact that the growth rate of electricity sales is the most important indicator for grid companies to input and output, the changes in the future economic environment will inevitably affect the results of the electricity sales index. In addition, the power system reform has been comprehensively promoted in recent years, and there are certain shortcomings in the sale of electricity warnings only from macroeconomic indicators. Therefore, it is recommended that grid companies use the analysis results of the early warning index of electricity sales, and need to establish a set of electricity sales adjustment analysis system based on power system reform, comprehensively assess the risk of electricity sales, and improve the company’s risk prevention capabilities.
### Table 5. Early-warning lamp setting

| State                        | Overheating | Partial heat | Stable       | Partial cold | Overcooling |
|------------------------------|-------------|--------------|--------------|--------------|-------------|
| Early-warning lamp           | Red light   | Yellow light | Green light  | Light blue light | Blue light  |
| Early warning value (A)      | 5           | 4            | 3            | 2            | 1           |
| Probability                  | 10%         | 15%          | 50%          | 15%          | 10%         |

### Table 6. Calculation of weights of each indicator

| Early warning indicator                  | Indicator type   | Mean (X) | Standard deviation (ε) | Correlation coefficient | Weights |
|------------------------------------------|------------------|----------|------------------------|-------------------------|---------|
| Industrial added value                   | Coincident indicator | 109.25   | 4.24                   | 0.85                    | 0.08    |
| Steel production                         | Coincident indicator | 115.02   | 17.12                  | 0.84                    | 0.08    |
| Aluminum production                      | Coincident indicator | 118.37   | 17.01                  | 0.81                    | 0.07    |
| Household freezer production             | Leading indicator | 110.12   | 22.79                  | 0.79                    | 0.07    |
| Metal cutting machine tool output        | Leading indicator | 109.49   | 32.17                  | 0.77                    | 0.07    |
| Synthetic detergent production           | Coincident indicator | 107.18   | 10.36                  | 0.75                    | 0.07    |
| Commodity retail price index             | Lagging indicator | 102.41   | 2.72                   | 0.75                    | 0.07    |
| Electric hand-held tool output           | Coincident indicator | 99.51    | 16.95                  | 0.73                    | 0.07    |
| New construction area of commercial housing | Coincident indicator | 108.60   | 28.57                  | 0.72                    | 0.07    |
| Cloth production                         | Coincident indicator | 106.06   | 11.70                  | 0.72                    | 0.06    |
| Cement production                        | Leading indicator | 102.05   | 6.14                   | 0.70                    | 0.06    |
| Asphalt production                       | Leading indicator | 125.56   | 55.18                  | 0.68                    | 0.06    |
| Room air conditioner production          | Coincident indicator | 114.55   | 29.33                  | 0.67                    | 0.06    |
| Chemical fiber production                | Leading indicator | 111.54   | 5.59                   | 0.66                    | 0.06    |

### Table 7. Values of early warning indicators for each indicator (increased rate)

| Early warning indicator value | 2007-1-1 | 2007-2-1 | 2007-3-1 | ...... | 2016-8-1 | 2016-9-1 | 2016-10-1 |
|------------------------------|----------|----------|----------|-------|----------|----------|-----------|
| Industrial added value       | 112.89   | 113.83   | 114.86   | ...... | 107.56   | 107.43   | 107.32    |
| Steel production             | 141.36   | 140.61   | 140.07   | ...... | 94.48    | 93.91    | 93.21     |
| Aluminum production          | 127.39   | 132.00   | 136.28   | ...... | 118.22   | 120.01   | 121.80    |
| Household freezer production | 137.94   | 139.31   | 141.83   | ...... | 99.69    | 97.66    | 96.26     |
| Metal cutting machine tool production | 192.85 | 186.68   | 177.62   | ...... | 100.20   | 101.25   | 101.38    |
| Early warning indicator                | Mean (X) | Standard deviation (σ) | Overheating (red light) | Partial heat (yellow light) | Stable (green light) | Partial cold (light blue light) | Overcooling (blue light) |
|---------------------------------------|----------|------------------------|-------------------------|-----------------------------|----------------------|-------------------------------|------------------------|
|                                       | Spot probability | Spot probability | Upper limit (X+1.282σ) | Spot probability | Upper limit (X+0.674σ) | Lower limit (X-0.674σ) | Spot probability |
|                                       |           |                       |                         |                             |                      |                               |                        |
| Industrial added value                | 109.25   | 4.24                   | 10%                     | 15%                         | 114.68               | 50%                           | 106.40                 | 15%                     | 103.82                | 10%                     |
| Steel production                      | 115.02   | 17.12                  | 10%                     | 15%                         | 136.96               | 50%                           | 126.56                 | 15%                     | 93.08                 | 10%                     |
| Aluminum production                   | 118.37   | 17.01                  | 10%                     | 15%                         | 140.17               | 50%                           | 129.83                 | 15%                     | 96.57                 | 10%                     |
| Household freezer production          | 110.12   | 22.79                  | 10%                     | 15%                         | 139.34               | 50%                           | 125.48                 | 15%                     | 80.90                 | 10%                     |
| Metal cutting machine tool production | 109.49   | 32.17                  | 10%                     | 15%                         | 150.73               | 50%                           | 131.17                 | 15%                     | 68.24                 | 10%                     |
| Synthetic detergent production        | 107.18   | 10.36                  | 10%                     | 15%                         | 120.46               | 50%                           | 114.16                 | 15%                     | 93.90                 | 10%                     |
| Commodity retail price index          | 102.41   | 2.72                   | 10%                     | 15%                         | 105.89               | 50%                           | 104.24                 | 15%                     | 98.93                 | 10%                     |
| Electric hand-held tool production    | 99.51    | 16.95                  | 10%                     | 15%                         | 121.24               | 50%                           | 110.93                 | 15%                     | 77.79                 | 10%                     |
| New construction area of commercial housing | 108.60   | 28.57                  | 10%                     | 15%                         | 145.23               | 50%                           | 127.86                 | 15%                     | 71.97                 | 10%                     |
| Cloth production                      | 106.06   | 11.70                  | 10%                     | 15%                         | 121.05               | 50%                           | 113.94                 | 15%                     | 91.06                 | 10%                     |
| Cement production                     | 102.05   | 6.14                   | 10%                     | 15%                         | 109.93               | 50%                           | 106.20                 | 15%                     | 94.18                 | 10%                     |

Table 8. Threshold values of early warning indicators for each indicator
Table 9. Monthly warning indicators for each indicator

| Early warning indicator value | 2007-2-1 | 2007-3-1 | 2007-4-1 | ...... | 2016-8-1 | 2016-9-1 | 2016-10-1 |
|-------------------------------|----------|----------|----------|-------|-----------|-----------|-----------|
| Industrial added value        | 4        | 5        | 5        | ......| 3         | 3         | 3         |
| Steel production              | 5        | 5        | 5        | ......| 2         | 2         | 2         |
| Aluminum production           | 4        | 4        | 4        | ......| 3         | 3         | 3         |
| Household freezer production  | 4        | 5        | 5        | ......| 3         | 3         | 3         |
| Metal cutting machine tool production | 5        | 5        | 5        | ......| 3         | 3         | 3         |
| Synthetic detergent production| 2        | 2        | 3        | ......| 3         | 4         | 4         |
| Commodity retail price index  | 3        | 3        | 3        | ......| 2         | 2         | 2         |
| Electric hand-held tool production | 3        | 4        | 5        | ......| 3         | 3         | 3         |
| New construction area of commercial housing | 3        | 3        | 3        | ......| 4         | 4         | 4         |
| Cloth production              | 3        | 3        | 3        | ......| 3         | 3         | 3         |
| Cement production             | 5        | 5        | 4        | ......| 5         | 5         | 5         |
| Asphalt production            | 3        | 3        | 3        | ......| 3         | 3         | 3         |
| Room air conditioner production| 1        | 2        | 3        | ......| 3         | 3         | 3         |
| Chemical fiber production     | 5        | 5        | 5        | ......| 3         | 3         | 3         |
| Vehicle production            | 3        | 3        | 3        | ......| 3         | 3         | 3         |

Table 10. Monthly warning index

| Early warning indicator value | 2007-1-1 | 2007-2-1 | 2007-3-1 | ...... | 2016-8-1 | 2016-9-1 | 2016-10-1 |
|-------------------------------|----------|----------|----------|-------|-----------|-----------|-----------|
| Industrial added value        | 4        | 4        | 5        | ......| 3         | 3         | 3         |
| Steel production              | 5        | 5        | 5        | ......| 2         | 2         | 2         |
| Aluminum production           | 3        | 4        | 4        | ......| 3         | 3         | 3         |
| Household freezer production  | 4        | 4        | 5        | ......| 3         | 3         | 3         |
| Metal cutting machine tool production | 5        | 5        | 5        | ......| 3         | 3         | 3         |
Table 1. Total Synthetic Warning Index

| Early warning indicator | Indicator type | Mean (X) | Standard deviation (ε) | Overheating (red light) | Partial heat (yellow light) | Stable (green light) | Partial cold (light blue light) | Overcooling (blue light) |
|-------------------------|----------------|----------|------------------------|-------------------------|-----------------------------|----------------------|-------------------------------|--------------------------|
|                         | Spot probability | Spot probability | Upper limit (X+1.28ε) | Spot probability | Upper limit (X+0.674ε) | Lower limit (X-0.674ε) | Spot probability | Lower limit (X-1.28ε) | Spot probability |
| Early warning index      | 100.50          | 22.67    | 10%                    | 15%                    | 129.56                      | 50%                  | 115.78                       | 85.21                    | 15%                      | 71.43                   | 10%                    |

Fig. 2. Monthly warning index
Fig. 3. Electricity sales warning results

Reference

1. Z.L. Dong, Association Analysis of Industrial Convergence and Industrial Structure Optimization and Upgrading. Cooperative Economy and Technology, 16 (2017)

2. Y. Zhang et al., Gray correlation analysis combined with support vector machine for near-infrared spectroscopy. Spectroscopy and Spectral Analysis, 2 (2013)

3. Y.C. Liu, et al., China's Macroeconomic Analysis and Forecast 2014-2015 – China's Macroeconomics in the New Normal Period. Economic Theory and Economic Management, 3 (2015)

4. J.G. Hua, Advance Indicators and Macroeconomic Fluctuation Prediction. Systems Engineering Theory and Practice, 10 (2014)

5. F.H. Lu, First, Synchronous Synthesis Index. Data, 4 (2011)

6. H. Zhang, Analysis of the Dynamic Relationship between Economic Advance Index and China's Macroeconomic Growth. Economic Issues, 11 (2012)

7. Z. Zhang, A problem that should be paid attention to when using time difference correlation analysis to determine leading, coincident and lagging indicators. Forecast, 3 (1991)

8. Y.N. Shi, China's Prosperity Forecast Based on Diffusion Index. Finance and Economics, 4 (2014)

9. L.Y. Chen, Li Renbo, Li Chunfeng, Analysis of the Synthetic Index of Current Economic Trends in China. Contemporary Economic Research, 2 (2014)

10. X.L. Kong, Zhang Tongbin, Gao Tiemei, Study on the Periodic Fluctuation Characteristics and the Characteristics of Current Volatility of China's Industrial Economy Based on Prosperity Index. Mathematics Practice and Cognition, 7 (2012)

11. J.M Wang, Research on the Early Warning Function of China's Leading Index on Economic Fluctuation. Jiangsu Social Science, 3 (2014)