

PAPER • OPEN ACCESS

## Improved time series electricity sales forecast based on economic prosperity method

To cite this article: Zhilong Yang *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **267** 062042

View the [article online](#) for updates and enhancements.

# Improved time series electricity sales forecast based on economic prosperity method

Zhilong Yang<sup>1\*</sup>, Yaming Liu<sup>1</sup>, Guanghua Wu<sup>2</sup>, Xuejun Li<sup>1</sup> and Qinghua Xu<sup>1</sup>

<sup>1</sup>State Grid Hebei Electric Power Co., Ltd. Handan Power Supply Branch, Handan City, Hebei Province, 056000, China

<sup>2</sup>State Grid Hebei Electric Power Co., Ltd. Electric Power Research Institute, Shijiazhuang City, Hebei Province, 050011, China

\*Corresponding author's e-mail: yyzllong@163.com

**Abstract.** With the construction of smart grids, a large amount of equipment operating data and power consumption data have been accumulated in the national grid system. However, these data have not been fully utilized nor has their potential value been exerted. Based on the customer's historical electricity consumption data, this paper analyzes the customer's electricity usage behavior from different dimensions, and predicts the sales of electricity in different dimensions, which is used as a supplement to customer category information such as customer industry, electricity properties, scale and voltage level. This will help with refined management. In this paper, the gray model and time series method are used to predict the sales of electricity. Combined with the economic index analysis method and the adjustment of holiday factors, the forecast results of electricity sales are revised.

## 1. Introduction

The electricity sales forecast is based on the known demand for electricity, considering the relevant factors of politics, economy, climate, etc., exploring the internal relationship between the power consumption and the main influencing factors and the law of their development and change, and predicting the future demand for electricity. Electricity sales forecast is an indispensable part of power system operating, controlling and planning. The results of power sales forecast have become the necessary basis for economic dispatch and implementation of power marketization.

One of the characteristics of electric energy is that it cannot be stored in a large amount, that is, the production, transportation, distribution and consumption of electric energy are carried out at the same time, so the available generating capacity in the system should meet the consumption demand at any time under normal operating conditions. The fact that electrical energy cannot be stored in large quantities makes it necessary for the power system to maintain a balance between supply and demand, and to provide reliable and standard-compliant power to all users of different type as economically as possible to meet their load requirements. Electricity is a sensitive commodity. Insufficient power supply will not only have a negative impact on the economy, but also affect social stability and the investment environment. Therefore, in order to ensure the safety and economic operation of the power system, it is necessary to grasp the law of changes in the sales of electricity and the future development trend. With the deepening of the power market reform, the so-called power companies in the power market should be based on the power market. All economic activities must focus on economic benefits, and take into account the supply and demand situation and development of the



power market as the basis of the company's business activities. Therefore, doing a good job of forecasting electricity sales is a necessary tool to grasp the pulse of the market and analyze the trend of future power demand. In the operation, control and planning management of power systems, the forecast of power sales determines the reasonable arrangement of power generation, transmission and distribution. It is not only an important part of power system planning, but also an important factor to improve the economic efficiency of power enterprises and promote the development of national economy.

According to the length of time, the electricity sales forecast can be divided into short-term (1 day to 1 week), medium-term (in the next few weeks to 5 years) and long-term (5 years or more) sales forecast. This paper studies the forecast of medium-term electricity sales. The medium-term electricity sales forecast can be divided into medium-short-term electricity sales forecast and medium-long-term electricity sales forecast. The short-medium-term electricity sales forecast is mainly used to determine fuel plan, reservoir dispatch, unit maintenance, etc., to ensure timely and qualified supply of electricity; The medium-long-term electricity sales forecast is conducive to planning the expansion and reconstruction of the future regional power grid in advance, so as to better promote the healthy, coordinated and rapid development of the local economy.

## 2. Material and Methods

Based on a large number of existing literatures, this paper selects the following electricity sales forecasting methods, including time series and gray model, combined with time difference correlation analysis, holiday factor adjustment, to revise prediction model.

### 2.1 Holt-winters

A time series (or a dynamic sequence) is a sequence in which the values of the same statistical indicator are arranged in chronological order in which they occur. The main purpose of time series analysis is to predict the future based on existing historical data.

Components: long-term trends, seasonal changes, cyclical changes, irregular changes.

The general trend of change in the long-term trend (T) phenomenon that is affected by some fundamental factor over a longer period of time;

The seasonal variation (S) phenomenon is the regularly and periodically change of seasons within one year;

The cyclical variation (C) phenomenon has a regular variation of the undulating form of the undulation;

Irregular change (I) includes two types of random changes, the strict random change and irregular sudden changes with great impact.

In this paper, the Holt-winters seasonal prediction model is used to do the prediction. The Holt-winters seasonal prediction model is also called the three exponential smoothing prediction methods in time series. The Holt-Winters seasonal prediction model is composed of a prediction function and a cubic smoothing function - one is the horizontal function  $L_t$ , one is the trend function  $b_t$ , one is the seasonal component function  $S_t$ , and the smoothing parameters  $\alpha$ ,  $\beta$  and  $\gamma$ , finally the prediction model function is  $F_{t+k}$  means.

$$L_t = \alpha(y_t - s_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (1)$$

$$b_t = (L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (2)$$

$$S_t = \gamma(y_t - L_t) + (1 - \gamma)S_{t-s} \quad (3)$$

$$F_{t+k} = L_t + kb_t + S_{t+k-s} \quad (4)$$

### 2.2 Gray model

The gray model is a differential equation established after the original data is generated and processed. Because the system is polluted by noise, the data presents a chaotic situation. The chaotic sequence is

a gray series, or gray process. The model established for the gray process becomes a gray model, which reveals the process of continuous development and change of things inside the system.

Grey models are gray-predictive models based on a small amount of incomplete information, and a long-term description of the fuzzy development of the law of the development of things (a branch of predictions with perfect theory and methods in the field of fuzzy prediction).

The general steps for a grey model prediction are:

Let the raw data be:

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) \quad (5)$$

(1) Raw data is accumulated, weakening volatility and randomness

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \quad (6)$$

$$x^{(1)}(t) = \sum_{k=1}^t x^{(0)}(k), t = 1, 2, \dots, n \quad (7)$$

(2) Accumulate the generated data to make the mean generation matrix A, B

$$A = (x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))^T \quad (8)$$

$$B = \begin{bmatrix} -0.5(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -0.5(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -0.5(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix} \quad (9)$$

(3) The parameters of the gray model are a, b

$$\begin{pmatrix} a \\ b \end{pmatrix} = (B^T B)^{-1} B^T A \quad (10)$$

(4) Calculate the fitted value

$$\tilde{X}^{(1)}(t+1) = \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-at} + \frac{b}{a} \quad (11)$$

(5) Discretize the fitted values to obtain predicted values

$$\tilde{X}^{(0)}(t+1) = \tilde{X}^{(1)}(t+1) - \tilde{X}^{(1)}(t) \quad (12)$$

### 2.3 Time difference correlation analysis

Time difference correlation analysis is often used in the selection of indicators for economic sentiment analysis. According to the time difference correlation analysis, the leading, consistent, and lagging indicators are determined, and the prediction model is revised. In this paper, the results of time difference analysis are used to divide the external environmental indicators into leading, consistent, and lagging indicators. According to the rising and falling trend of the leading indicators and the forecast results of power sales is revised weightedly.

Table 1. Time difference analysis indicator type

Index type	Meaning	Selection criteria	example
Leading indicator	The indicator that changes first before the overall economic growth or recession has yet to come can be used to estimate the trend of economic fluctuations.	From the economic point of view, there is a clear and affirmative advance relationship; Compared with the peak of the benchmark cycle, its peak is at least 3 months ahead of the curve, and in the last three consecutive cycle fluctuations, at least two stay ahead, leading by more than three months.	Inflation rate; Investment price index; Consumer price index

consistent	The indicator that changes with the fluctuation of the economy, also known as the consistent indicator, the time when the peaks and valleys appear coincides with the time when the economic running peaks and valleys appear, which comprehensively reflects the state of the economy as a whole.	From the economic point of view, the synchronization index should have obvious synchronization characteristics with its reference cycle; It is close to the peak of the reference cycle, and the peak difference is within 2 months.	Commodity sales; Labor unemployment rate
Lagging indicator	An indicator that shows the effect of economic fluctuations. It is a confirmation of the peaks and valleys that have emerged in the overall economic operation. It can verify the signals displayed by the leading indicators.	From an economic point of view, there should be a positive lag relationship with its baseline cycle; The peak is delayed by more than 3 months compared to the reference cycle.	Urban and rural residents' savings; Commodity stocks; Employee wage

The general steps for determining the time difference correlation coefficient are as follows:

The screening of the primary selection indicators includes: using the Pearson correlation coefficient method to screen the primary selection indicators, and using the primary selection index with a Pearson correlation coefficient higher than a preset threshold as an economic indicator.

Determine the time difference correlation coefficient as follows:

$$p_i = \frac{\sum_{k=1}^n (x_{i(k+L_i)} - \bar{x}_i)(x_{0(k)} - \bar{x}_0)}{\sqrt{\sum_{k=1}^n (x_{i(k+L_i)} - \bar{x}_i)^2} \sqrt{\sum_{k=1}^n (x_{0(k)} - \bar{x}_0)^2}} \quad (13)$$

Where  $p_i$  is the Pearson correlation coefficient of the  $i$ -th economic indicator,  $L_i$  is the time difference coefficient of the  $i$ -th economic indicator,  $k$  is the number of periods in which the economic indicator is located,  $n$  is the total number of periods,  $x_{i(k+L_i)}$  is the value of the  $i$ -th economic indicator in the  $k+L_i$  period,  $\bar{x}_i$  is the average value of the  $i$ -th economic indicator in the total period,  $x_{0(k)}$  is the sales volume of the  $k$ -th period, and  $\bar{x}_0$  is the sale average value of electricity;

Find the time difference coefficient  $L_i$  at the maximum value of  $p_i$ .

#### 2.4 Holiday factor adjustment

Under normal circumstances, in the electricity sales forecast, the electricity forecast in January and February is often a large error and needs to be corrected. This is mainly caused by two reasons: the impact of the Spring Festival and the impact of the Jubilee. The Spring Festival is the most important festival in China. The holiday time is long and the impact time is long. The Spring Festival is calculated according to the lunar calendar, and the month is divided according to the solar calendar. Therefore, the date of the solar calendar in the Spring Festival is different every year, and the month is different. This affects the accuracy of the electricity sales forecast in January and February.

It is assumed that the average daily electricity consumption during the normal period of January and February in the  $n$ th year of the forecast is  $X_0$ , and the average daily electricity consumption during the Spring Festival period is  $X_1$ . The ratio of the average daily electricity consumption in the influence period of the Spring Festival in the  $n$ th year to the average daily electricity consumption in the normal period is  $\lambda_n$ , and the weighted average value of the data over the years can be obtained.  $D_1$  is the number of days affected by the Spring Festival in January,  $D_2$  is the number of days affected by the Spring Festival in February, and  $D$  is the total number of days in February ( $D=28$  in normal years and  $D=29$  in leap years). It is generally believed that the influence period of the Spring Festival is: 2 days before the Spring Festival and 7 days after the Spring Festival.

$$\lambda_n = X_1/X_0 \quad (14)$$

$$E'_{(n,1)} = (31 - D_1)X_0 + D_1X_1 \quad (15)$$

$$E'_{(n,2)} = (D - D_2)X_0 + D_2X_1 \quad (16)$$

$$E'_{(n,1)} + E'_{(n,2)} = E_{(n,1)} + E_{(n,2)} \quad (17)$$

According to the above formula, the predicted value of the electricity sales in the first year of January,  $E_{(n,1)}$  and the predicted value of the electricity sold in February,  $E_{(n,2)}$  are known, and  $D$ ,  $D_1$ ,  $D_2$  are also known. The quantity, so that the predicted correction value  $E'_{(n,1)}$ ,  $E'_{(n,2)}$  of January and February of the  $n$ th year can be obtained.

### 3. Results

The forecasting process is as follows. Firstly, the data is corrected by the holiday factor method. Then, the Holt-winters model is used to predict the electricity sales in the next year. Finally, the results of the correlation analysis are referenced and substituted into the gray model to correct the prediction results of Holt-winters. The forecast flow chart is as follows.

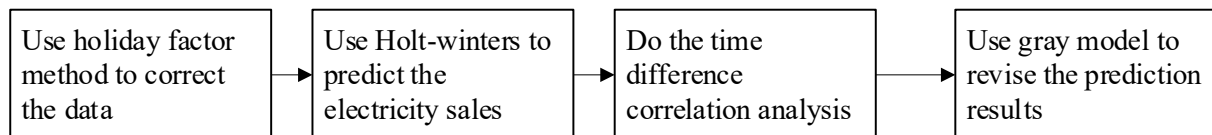


Figure 1. Forecast flow chart

Time difference correlation analysis result:

Table 2. Time difference index and the corresponding correlation coefficient

Index type	Index name	Lag period	Correlation coefficient
leading index	Revenue	-2	0.9030
	Plate volume	-2	0.9092
Consistent index	Coal mine production	0	0.7937
Lagging index	Real estate purchase area	1	0.7471
	Secondary metallurgical coke price	4	0.8006

Forecast correction results:

Table 3. Time difference index and correlation coefficient

Area/error	2018/1	2018/2	2018/3	2018/4	2018/5	2018/6	average error
A1	7.10%	2.20%	11.05%	9.57%	-0.92%	3.46%	5.72%
A2	3.84%	-7.96%	4.92%	0.06%	1.24%	-2.87%	3.48%
A3	4.97%	-1.66%	-1.07%	3.95%	-1.89%	8.71%	3.70%
A4	2.27%	-1.28%	2.05%	6.54%	5.11%	-6.23%	3.91%
A5	-7.07%	-3.97%	7.69%	0.14%	-6.99%	-7.12%	5.49%
A6	1.31%	1.52%	1.23%	5.38%	5.01%	9.26%	3.95%
A7	4.67%	-3.59%	3.64%	-2.05%	9.57%	-5.12%	4.77%
A8	-5.30%	1.11%	6.61%	-0.51%	-4.11%	-4.05%	3.61%
A9	-3.48%	-1.05%	8.12%	1.38%	3.63%	0.11%	2.96%
A10	-12.03%	-5.23%	-4.14%	-1.72%	-2.10%	3.15%	4.73%
A11	9.07%	10.60%	-6.02%	6.75%	3.00%	5.48%	6.82%
A12	-0.27%	1.12%	0.52%	1.98%	5.43%	-9.83%	3.19%
A13	-3.83%	6.77%	11.88%	1.66%	5.40%	-11.63%	6.86%
A14	-2.88%	3.35%	-1.06%	2.57%	5.43%	-3.61%	3.15%
A15	-6.12%	-2.36%	-8.70%	-8.34%	3.81%	0.76%	5.01%
A16	-7.11%	0.22%	2.62%	3.47%	-5.05%	-6.29%	4.13%

#### 4. Discussion

Judging from the composition of the eight major power sales categories in the A city, the proportion of large industrial sales in the past years accounted for about 64%, which is the main driving force for the growth of power consumption in the a city, but the proportion of large industrial sales is decreasing by year.

From the perspective of seasonal characteristics, different types of electricity sales have different capacity utilization characteristics, which can be divided into three types: those closely related to temperature, including commercial, residential, and non-residential, reflecting the characteristics of high in summer and winter, low in spring and autumn. Closely related to precipitation is reflected in the obvious irrigation power characteristics of agriculture, with summer as the highest peak; seasonal characteristics are not obvious including non-Pro-Industry.

From the perspective of annual characteristics, there are the following types: in recent years, the sales volume has not fluctuated such as commercial and non-residential; the relative stable increments such as residents have been maintained year by year; the precipitation related to the current year, such as agriculture, the wet year and the peaceful water year should be Different treatment; the rise and fall of non-general industries depends on districts and counties, such as the rising and falling trend in A1 and A2 counties respectively. From the perspective of economic prosperity, large industries are greatly affected by the economic environment.

#### 5. Conclusions

In this paper, the electricity sales forecast is carried out in each district and county of A city, and the electricity sales forecasting model based on time series, gray model, time difference correlation analysis and holiday factor adjustment are established. The prediction results can be used for the operation, control and planning management of power systems, as well as operational management decisions. However, due to the large random disturbance factors of small-scale electricity sales, there is still room for improvement in prediction accuracy.

For the improved method, it can be considered to analyze the internal business indicators of the grid enterprise, such as line loss rate, capacity ratio, comprehensive voltage qualification rate, power supply reliability rate, and incorporate the prediction model for more accurate analysis. It has been shown in the literature that in economically developed regions, the power supply of distribution networks is consistent with the speed of improvement of power quality and reliability. In economically underdeveloped regions, the growth of electricity sales lags behind the growth of reliability and power quality. The line loss rate and the integrated voltage pass rate are consistent and synchronized in time, and the power supply reliability rate and the integrated voltage pass rate are also kept in time synchronization. The sales volume represents a high correlation between economic efficiency and power supply reliability rate and comprehensive voltage qualification rate, indicating that the company's reliability and voltage qualification rate are a guarantee for electricity sales, and also indicates the development of power consumption in recent years. It is an improvement in power quality and power supply capability, and is a high-quality growth that combines the economic benefits based on safety reliability. In the economically developed regions, the sales volume and the reliability of power supply and the operating efficiency maintain a certain degree of synchronization in the time dimension of output. In economically underdeveloped regions, more performance is provided in the improvement of power supply quality and operational efficiency, while the magnitude of electricity increased is relatively lagging behind.

#### Acknowledgements

This research was financially supported by the State Grid Handan Electric Power Supply Company's "Research on the Dynamic Evaluation Model of Substation Economic Benefits Based on Electricity Reform Investment Constraint" (Grant No. kj2018-055).

**References**

- [1] Shi Y.Y, Ling W.J, Guo X.Y, etc. (2016) Electricity sales analysis and forecasting research. J. Manager, 18.
- [2] Zhao P.C, Yan Y.J. (2009) Application of Grey Prediction Technology in Forecasting Annual Electricity Consumption. J. Heilongjiang Province Science and Technology Information, 30: 56-66.
- [3] Kong L.Y, Lu J.P, Yan W, etc. (2008) Electricity sales classification prediction model and its software development. J. Journal of Power System and Automation, 20: 51-55.
- [4] Pan X.H, Liu L.P, Li Y. (2013) A new method to improve the accuracy of monthly electricity sales forecast. J. Power demand side management, 3: 11-15.
- [5] Liu Q.H, Chen J, Gan H.Q. (2009) Forecast analysis of electricity sales based on improved grey model. J. Statistics and Information Forum, 11: 17-21.
- [6] Wang T, Zhang W, W M.M, etc. (2010) Research on Forecast Method of Electricity Consumption of Regional Power Grid Based on Time Series. J. Electrical Technology, 11: 9-12.