

## Research on Power Demand Forecasting of Beijing Based on Regression Analysis

Shuyu Dai <sup>1,a \*</sup>, Yan Li <sup>1,b</sup> and Dongxiao Niu <sup>1,c</sup>

<sup>1</sup>School of Economics and Management, North China Electric Power University, Beijing, China

<sup>a</sup> DaiShuyu@ncepu.edu.cn, <sup>b</sup> ealinne13@ncepu.edu.cn, <sup>c</sup> niudx@126.com

\* the corresponding author

**Keyword:** Electricity demand forecast; Unitary regression; Multiple regression

**Abstract.** With the continuous development of the national economy, people's demand for electricity is increasing day by day, and the power industry plays an increasingly important role in the social development. Power demand forecasting is the basic work of power management, planning, programming and other management departments in power system. Improving the technical level of power demand forecasting can provide a strong basis for power system scheduling and long-term planning. This paper analyzes the influencing factors of electricity demand in Beijing, analyzes the influence of GDP, industrial structure, population and urbanization level on the electricity consumption of Beijing, and then uses the regression analysis method based on the collected economic data and power data to forecast the electricity consumption of the whole society in Beijing from 2017 to 2022.

### Introduction

Regression analysis is an effective forecasting method for medium and long term power demand forecasting. This prediction method has the advantages of simple principle and simple form, fast prediction and good extrapolation performance. Regression analysis using the principle of mathematical statistics to process a large amount of statistical data mathematically and determine the relationship between power consumption or power load and some independent variables such as population, national economic production value. A mathematical model with good correlation is established, namely regression equation. The regression equation is extrapolated to predict the future electricity consumption. According to the number of variables, regression analysis can be divided into unitary regression analysis and multiple regression analysis. In order to accurately predict the demand for electricity in Beijing, it is necessary to analyze the main factors affecting the electricity consumption of the whole society in Beijing. There are many factors that affect the electricity consumption; this paper mainly analyzes the impact of GDP, industrial structure, population, urbanization level and other factors on the electricity consumption in Beijing [1-3].

### Construction of Prediction Model

**Unitary Regression Model.** In unary linear regression, the independent variable is a variable that can be controlled or can be observed accurately and it is expressed by  $x$ . The dependent variable is a random variable that depends on  $x$  and it is represented by  $y$ . Suppose the relationship between  $x$  and  $y$  is

$$y = a + bx + \varepsilon \tag{1}$$

Where  $\varepsilon$  is a random error, also known as random interference, and it obeys the normal distribution  $N(0, \sigma^2)$ .  $a, b$  and  $\sigma^2$  are unknown parameters that do not depend on  $x$ . The relationship between  $x$  and  $y$  is called unary linear regression model. This model can also be remembered as:

$$y = a + bx + \varepsilon, \varepsilon \sim N(0, \sigma^2) \quad (2)$$

For fixed  $x$ ,  $y \sim N(a + bx, \sigma^2)$ , that is, the mathematical expectation of the random variable  $y$  is:

$$E_y = a + bx \quad (3)$$

This is because  $E(y_i) = E(a + bx_i + \varepsilon_i) = E(a + bx_i) + E(\varepsilon_i) = a + bx_i + 0$ . Obviously  $E_y$  is a function of  $x$ , which is called  $y$  regression on  $x$ . In practice,  $n$  times experimental observation was made on independent variable  $x$  and dependent variable  $y$  and the observation of  $y$  on each of the different values of  $x$  is independent of each other, then  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$  was called sample. If the unknown parameters  $a, b$ , can be estimated based on the sample, then denote valuation as  $\hat{a}, \hat{b}$  respectively and call the following formula:

$$\hat{y} = \hat{a} + \hat{b}x \quad (4)$$

As  $y$  on the linear regression equation of  $x$ .  $\hat{b}$  is the regression coefficient. The graph of regression equation is called regression line [4-5].

**Multiple Regression Model.** Assume that  $x_1, x_2, \dots, x_p$  are ( $p > 1$ ) linearly independent controllable variables and  $y$  is a random variable. The relationship between them is:

$$\left. \begin{aligned} y &= b_0 + b_1x_1 + \dots + b_px_p + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2) \end{aligned} \right\} \quad (5)$$

Where:  $b_0, b_1, \dots, b_p, \sigma^2$  are unknown parameters independent of  $x_1, x_2, \dots, x_p$ , and  $\varepsilon$  is random error (or random disturbance), which is the  $p$ -ary linear regression model. Taking the mathematical expectations at both ends of Eq.5 and then can obtain:

$$E_y = b_0 + b_1x_1 + \dots + b_px_p \quad (6)$$

Obviously,  $E_y$  is a function of  $x_1, x_2, \dots, x_p$ . Eq.6 is called the regression plane equation, where  $b_1, b_2, \dots, b_p$  are called the regression coefficients [6-7].

## Empirical Analysis

**Unary Regression Prediction.** Electricity consumption is closely related to the size of GDP. Therefore, we establish a linear regression model for electricity consumption and GDP in Beijing from 2000 to 2016. This paper takes GDP in Beijing during 2000-2016 years independent variable, total social electricity consumption as dependent variable and uses Eviews software to do unary linear regression. The confidence interval width is set to 95% and then gets unary linear equation:

$$EC = 285.8741GDP + 3455162$$

Where: EC is electricity consumption, GDP is regional gross domestic product in Beijing.

In order to forecast the electricity consumption in Beijing from 2017 to 2022, this paper makes a polynomial fitting prediction of Beijing GDP from 2017 to 2022, and the fitting result is shown in Fig. 1:

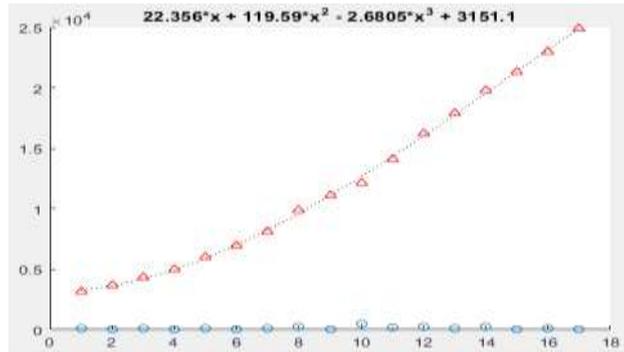


Figure 1. Polynomial fitting results

Thus, we can get the predictive value of GDP in Beijing in 2017-2022 years, and set it as the common situation, and the high and low situation takes 5% of the common situation, and the specific forecast data is shown in Table 1.

Table 1 Prediction value of GDP in Beijing city in 2017-2022 years

Unit: [100 million yuan]						
Year	2017	2018	2019	2020	2021	2022
High situation	28001.39	29780.42	31489.73	33112.44	34631.65	36030.49
Common situation	26667.99	28362.3	29990.22	31535.66	32982.53	34314.75
Low situation	25334.59	26944.19	28490.71	29958.87	31333.4	32599.02

According to the three situation of GDP, we make prediction of three schemes (high, common and low) for electricity consumption, and the result of prediction are shown in Table 2:

Table 2 Forecast value of electricity consumption in unitary linear regression

Unit: [10MW]						
Year	2017	2018	2019	2020	2021	2022
High situation	11460035	11968613	12457261	12921150	13355455	13755347
Common situation	11078850	11563210	12028589	12470389	12884013	13264862
Low situation	10697666	11157808	11599918	12019628	12412570	12774377

Fig. 2 shows the fitting and forecast of electricity for unary linear regression:

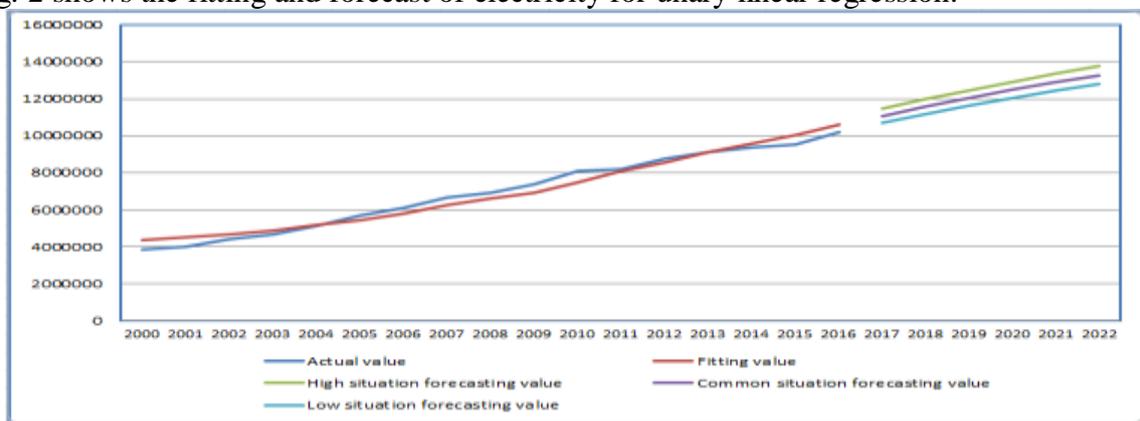


Figure 2. The fitting and forecasting of electricity for unary linear regression

**Multiple Regression Prediction.** From the analysis of the above factors, consumption is not

only affected by the GDP, but also affected by the industrial structure, population, urbanization level and other factors. So we build a multiple linear regression model of electricity consumption of Beijing in 2012-2016. This paper takes the 2000-2016 GDP in Beijing City, industrial structure, population and urbanization level as independent variables (among them, the industrial structure is represented by the ratio of the added value of the third industry to the GDP, and the level of urbanization is expressed by the ratio of the urban population to the total population of the region) and uses Eviews software to do multiple linear regression. The confidence interval width is set to 95% and then gets multiple linear equation:

$$EC = 151.2702GDP + 1018137ST + 1617.163PP + 15022689UR - 11065461$$

Where: EC is electricity consumption, GDP is regional gross domestic product in Beijing, ST is the industrial structure, PP is the population, UR is the urbanization level

The industrial structure and the population did not pass the t-test, so it should be removed from the model and multiple regression should be done again. Then we obtain the multiple linear equation:

$$EC = 207.9554GDP + 19144806UR - 11575308$$

Where: EC is the electricity consumption, GDP is regional gross domestic product in Beijing, UR is the level of urbanization.

In order to predict the electricity consumption in Beijing from 2017 to 2022, it is necessary to predict the GDP and urbanization level of Beijing in 2017-2022 years by polynomial fitting. Since GDP has been predicted above, only the polynomial fitting prediction of urbanization level is carried out here, and the fitting result is shown in Fig. 3.

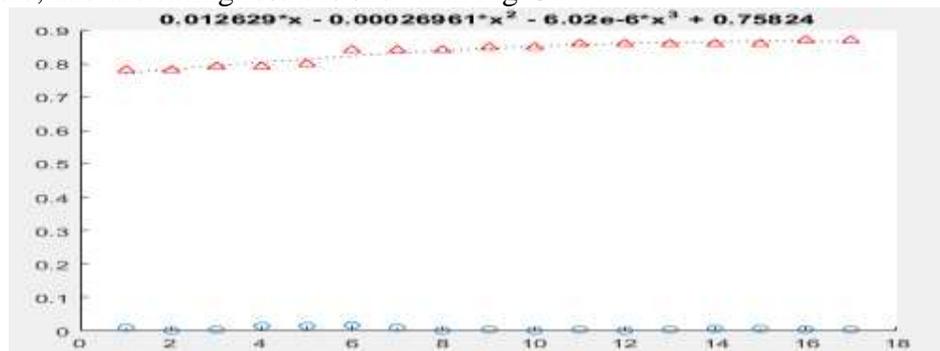


Figure 3. Polynomial fitting results

Thus, we can get the predictive value of urbanization level of Beijing in 2017-2022 years, and set it as the common situation, and the high and low situation takes 1% of the common situation, and the specific forecast data is shown in Table 3.

Table 3 The predictive value of urbanization level of Beijing in 2017-2022 years

Year	2017	2018	2019	2020	2021	2022	Unit: [%]
High situation	87.17	86.82	86.34	85.73	84.99	84.12	
Common situation	86.31	85.96	85.48	84.88	84.15	83.28	
Low situation	85.45	85.10	84.63	84.03	83.31	82.45	

According to the three situation of GDP and urbanization level, we make prediction of three schemes (high, common and low) for electricity consumption, and the result of prediction are shown in Table 4:

Table 4 Forecast value of electricity consumption in multiple linear regression

	Unit: [10MW]					
Year	2017	2018	2019	2020	2021	2022
High situation	10936848	11238567	11502091	11723210	11897715	12021394
Common situation	10494322	10779099	11026607	11232809	11393669	11505152
Low situation	10051795	10319631	10551122	10742407	10889624	10988910

Fig. 4 shows the fitting and forecast of electricity for multiple linear regression:

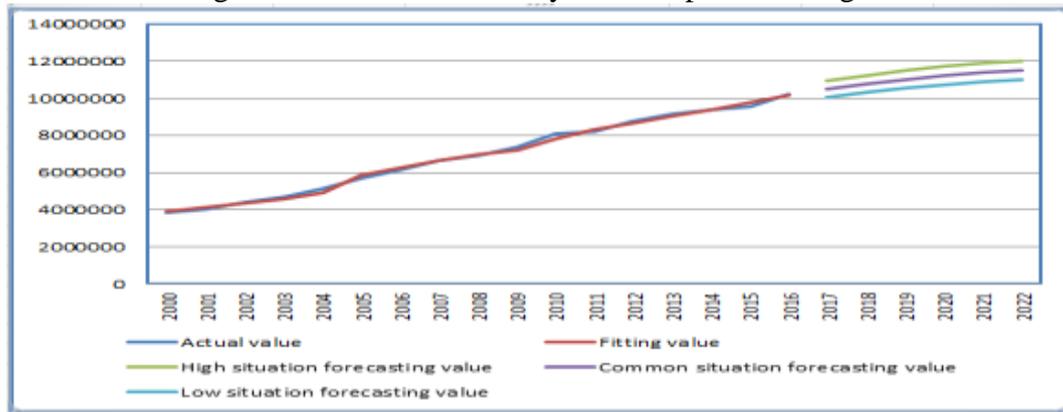


Figure 4. The fitting and forecast of electricity for multiple linear regression

**Conclusions**

This paper analyzes the influence factors of power demand in Beijing, mainly analyzes the impact of GDP, industrial structure, population, urbanization level and other factors on the electricity consumption in Beijing. Then, according to the collected economic data and power data, using the regression analysis method to predict the whole society electricity consumption of Beijing in 2017-2022 years. It can be seen from the forecast results that the demand for electricity in Beijing will continue to rise in the coming years. The forecast results provide a reliable basis for decision-making of power system planning and operation in Beijing.

**References**

- [1] K.G. Boroojeni, M.H. Amini, S. Bahrami, et al: Electric Power Systems Research, Vol. 142 (2017) No.1, p. 58-73.
- [2] M.D. Felice, A. Alessandri and F. Catalano: Applied Energy, Vol. 137 (2015) No.1, p. 435-444.
- [3] K. Berk and A. Müller: Journal of Energy Markets, Vol. 9 (2016) No.2, p. 1-20.
- [4] J.C.Mi: Coal Quality Technology, Vol. 1 (2014) No.1, p.21-22.
- [5] Y. Peng: Science Mosaic, Vol. 7 (2015) No.7, p.39-42.
- [6] G.T. Knofczynski and D. Mundfrom: Educational & Psychological Measurement, Vol. 68 (2008) No.3, p. 431-442.
- [7] T. Catalina, V. Iordache and B. Caracaleanu: Energy & Buildings, Vol. 57 (2013) No.57, p. 302-312.