Investment Demand Forecast for Grid Companies Based on Multiple Linear Regression Analysis

Yizheng Li[1], Junxian Ma[2], Lang Zhao[1], Xue Feng[2*], Dong Peng[1], Haiqiong Yi[1]

lyztg@2022@163.com, 410176153@qq.com, zhaolang@chinasperi.sgcc.com.cn, ysfvab@126.com, pengdong@chinasperi.sgcc.com.cn, yihaiqong@chinasperi.sgcc.com.cn

State Grid Economic and Technological Research Institute Co., 5th and 6th floors, Building A, No. 18 Binhe Avenue, Future Science and Technology City, Changping District, Beijing, 102200[1]
Economic and Technical Research Institute of State Grid Ningxia Electric Power Co., State Grid CBD Building, Houhai Road and Fengmai Street, Jinfeng District, Yinchuan City, Ningxia Hui Autonomous Region, 750011[2]

Abstract. Grid investment demand refers to the investment made by provincial power companies to ensure that strategic projects and decisions are steadily promoted, while building a safe and reliable grid according to the needs of social and economic development. First, the role of population, GDP, total energy consumption and total social electricity consumption on the investment demand of grid companies is analyzed. Secondly, the correlation between each influencing factor and grid investment demand is analyzed by Pearson correlation coefficient analysis, and then the key factors affecting grid investment demand are selected. Finally, combining the degree of influence of key factors, the future investment demand of grid companies is predicted by multiple regression prediction model.

Keywords: Grid investment; investment demand forecast; investment impact factors; multiple linear regression analysis

1 Introduction

In the new electricity reform situation, power grid enterprises need to develop investment strategies to adapt to the new environment and invest limited funds in projects that can best enhance the benefits of the power grid, so as to achieve accurate and effective investment. The prerequisite for accurate and effective investment is a scientific and reasonable forecast of grid investment demand. Grid investment demand refers to the investment made by provincial power companies to ensure that strategic projects and decisions are steadily promoted, while building a safe and reliable grid according to the needs of social and economic development. Grid investment is generally huge and has many influencing factors, so the current domestic research on investment demand mainly focuses on two aspects of investment demand influencing factors and measurement models. The literature [1] reflects the demand for grid infrastructure from three aspects: socio-economic level, grid scale and load consumption level, and the specific influencing factors include power supply population, GDP, substation capacity, social electricity consumption and regional maximum load. The literature [2] analyzes the market and regulatory factors affecting smart grid investments in Europe and encourages regulators to adopt innovative mechanisms to stimulate pilot investments. The literature [3] investigates the

ICBBEM 2023, May 19-21, Hangzhou, People's Republic of China

DOI 10.4108/eai.19-5-2023.2334303

sensitivity of grid investment to factors such as interest rate changes and equity financing in the capital market. In terms of grid investment demand model construction, literature [4] proposed a grid investment demand measurement model based on cointegration theory and error correction model, and constructed a model of the short-term regulatory relationship between the two based on the long-term equilibrium model through the error correction model to improve the short-term forecasting accuracy. In the literature [5], an adaptive bell curve with environmental factors is proposed to describe the grid investment demand based on the relationship between the grid investment demand and the grid development stage. The literature [6] used cointegration theory to analyze the main influencing factors of private electricity demand and obtain a short-term equilibrium model of private electricity demand. The perspective and methods of foreign research on grid investment demand are slightly different compared with domestic research. Few foreign studies consider combining the influencing factors of power supply construction with investment demand, and focus more on forecasting methods and models.

2 Factors influencing the demand for power grid investment

2.1 Population factor

Population size is an important parameter affecting electricity demand and power load, directly affecting the final value of electricity consumption of the whole society, while the population size and structure are closely related to the level of per capita electricity consumption. The number of population will constitute a certain influence on the power grid, in general, the more the number of population, the faster the power grid development. Population has a direct impact on macro economy, national economic development, GDP per capita, total energy consumption, etc., which will eventually affect the grid investment demand.

2.2 GDP

GDP (Gross Domestic Product) refers to the sum of the value of all final goods and services produced by all resident units of a country (or region) in a certain period of time, and is often regarded as an indicator to measure the economic status of a country (or region). GDP has a direct impact on macroeconomy, national economic development, per capita GDP, total energy consumption, three industries, grid capital construction, power load, etc., which will affect the grid investment demand.

2.3 Total energy consumption

Total energy consumption refers to the sum of all kinds of energy consumed by material production sectors, non-material production sectors and domestic consumption in a certain period of time, and is an aggregate indicator to observe the level, composition and growth rate of energy consumption. Total energy consumption includes raw coal and crude oil and its products, natural gas, electricity, excluding the use of low calorific value fuels, biomass and solar energy, etc. Energy consumption directly affects the demand for grid investment, the growth of energy consumption drives the growth of electricity consumption and clean energy consumption, corresponding to the growth in demand for grid investment.

2.4 Electricity consumption of the whole society

The electricity consumption of the whole society is used for economic statistics, referring to the total consumption of electricity in all fields of primary, secondary and tertiary industries, including industrial electricity, agricultural electricity, commercial electricity, residential electricity, public facilities electricity and other electricity consumption. The electricity consumption of the whole society has a direct impact on power grid capital construction, power load, power supply capacity, power balance grid operation, etc., which will eventually affect the power grid investment demand.

3 Power grid investment demand forecast

3.1 Analysis of key influencing factors

In this paper, Pearson correlation coefficient analysis is used to analyze the correlation between each influencing factor and the grid investment demand, and then select the key factors affecting the grid investment demand. The Pearson correlation coefficient analysis method is calculated as follows:

(1) Calculate the standard deviation of each variable itself

$$S_x = \sqrt{(n-1)^{-1}\left(\sum\left(x_i - \overline{x}\right)^2\right)} \tag{1}$$

$$S_y = \sqrt{(n-1)^{-1}\left(\sum\left(y_i - \overline{y}\right)^2\right)} \tag{2}$$

(2) Calculate the covariance between variables

$$S_{xy} = \left(\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)\right)\cdot(n-1)^{-1} \tag{3}$$

(3) Calculate the correlation coefficient between variables

$$r_{xy} = S_{xy}\left(S_x \cdot S_y\right)^{-1} \tag{4}$$

Using SPSS software, the correlations between the investment demand of power grid and the influencing factors in 25 regions of China were analyzed, and the calculation results are shown in Table 1.

Table 1. Correlation of investment demand influencing factors

Relevance	Demographic factors	GDP	Total energy consumption	Electricity consumption of the whole society
Investment Demand	0.731	0.856	0.397	0.883

According to the correlation matrix, the correlation coefficients of grid investment demand with population, GDP, total energy consumption and total social electricity consumption are 0.731, 0.856, 0.397 and 0.883, respectively.According to the correlation classification table, the correlation between grid investment demand and GDP, total social According to the correlation table, the demand for grid investment is highly correlated with GDP and social electricity consumption. Population is the next most significant correlation. Therefore, the key influencing factors of investment demand are selected in this paper as GDP, social electricity consumption and population.

3.2 Multiple linear regression model

Regression analysis is a statistical analysis method to study the relationship between the action of variables. If there are two or more independent variables, it is called multiple regression, and the regression analysis method to describe the quantitative relationship between the dependent variable and two or more independent variables is called multiple linear regression analysis. In practical applications, a phenomenon is often associated with more than one factor, so multiple regression prediction models are more realistic and effective. Multiple regression is based on the principle of least squares to build the model, and its solution steps are:

(1) Establish the regression equationt

Suppose a certain dependent variable y is influenced by k independent variables $x_1, x_2, ..., x_k$ with n sets of observations （$y_a, x_{1a}, x_{2a}, ..., x_{ka}$）, $a = 1,2,...,n$. Then, the structure of the multiple linear regression model takes the form of

$$y_a = \beta_0 + \beta_1 x_{1a} + \beta_2 x_{2a} + ... + \beta_k x_{ka} + \varepsilon_a \tag{5}$$

Where: $\beta_0, \beta_1, ..., \beta_k$ is the parameter to be determined; ε_a is the random variable.

If $b_0, b_1, ..., b_k$ is the fitted value of $\beta_0, \beta_1, \beta_2 ..., \beta_k$, respectively, the regression equation is

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k \tag{6}$$

Where: b_0 is a constant; $b_1, b_2, ..., b_k$ is called the partial regression coefficient, the meaning is the value of the dependent variable y that changes on average for each unit change in the independent variable x_i when all other independent variables x_j ($j \neq i$) are fixed.

(2) Solving regression parameters by least squares method

According to the principle of least squares, the estimate of β_i b_i ($i = 0,1,2,..., k$) should be such that:

$$Q = \sum_{a=1}^{n} \left(y_a - \hat{y}_a \right)^2 = \sum_{a=1}^{n} \left[y_a - \left(b_0 + b_1 x_{1a} + b_2 x_{2a} + ... + b_k x_{ka} \right) \right]^2 \to \min \tag{7}$$

The necessary conditions for the extreme value are given by

$$\begin{cases} \dfrac{\partial Q}{\partial b_0} = -2\sum_{a=1}^{n}\left(y_a - \hat{y}_a\right) = 0 \\ \dfrac{\partial Q}{\partial b_j} = -2\sum_{a=1}^{n}\left(y_a - \hat{y}_a\right)x_{ja} = 0(j = 1,2,...,k) \end{cases} \tag{8}$$

The above equation is expanded and organized to obtain the regular system of equations:

$$\begin{cases} nb_0 + (\sum_{a=1}^{n} x_{1a})b_1 + (\sum_{a=1}^{n} x_{2a})b_2 + ... + (\sum_{a=1}^{n} x_{ka})b_k = \sum_{a=1}^{n} y_a \\ (\sum_{a=1}^{n} x_{1a})b_0 + (\sum_{a=1}^{n} x_{1a}^2)b_1 + (\sum_{a=1}^{n} x_{1a}x_{2a})b_2 + ... + (\sum_{a=1}^{n} x_{1a}x_{ka})b_k = \sum_{a=1}^{n} x_{1a}y_a \\ (\sum_{a=1}^{n} x_{2a})b_0 + (\sum_{a=1}^{n} x_{1a}x_{2a})b_1 + (\sum_{a=1}^{n} x_{2a}^2)b_2 + ... + (\sum_{a=1}^{n} x_{2a}x_{ka})b_k = \sum_{a=1}^{n} x_{2a}y_a \\ \cdots \\ (\sum_{a=1}^{n} x_{ka})b_0 + (\sum_{a=1}^{n} x_{1a}x_{ka})b_1 + (\sum_{a=1}^{n} x_{2a}x_{ka})b_2 + ... + (\sum_{a=1}^{n} x_{ka}^2)b_k = \sum_{a=1}^{n} x_{ka}y_a \end{cases} \tag{9}$$

The above equation is further written in matrix form:

$$A = X^T X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{pmatrix} \begin{pmatrix} 1 & x_{11} & \cdots & x_{k1} \\ 1 & x_{12} & \cdots & x_{k2} \\ 1 & \cdots & \cdots & \cdots \\ 1 & x_{1n} & \cdots & x_{kn} \end{pmatrix}$$

$$B = X^T Y = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} \tag{10}$$

Solving the above equation yields: $b = A^{-1}B = (X^T X)^{-1} X^T Y$, which is the regression parameter.

(3) Significance test

Significance checks were performed on the regression equation and regression parameters to test the effect of fitting the equation to the actual data into. Each independent variable is tested to determine that the linear effect of each independent variable on the dependent variable is significant.

(4) Regression diagnosis

The basic idea of residual analysis is to use the residual e that can be calculated as an estimate of the random error ε. The characteristics of the residual are used to examine the reasonableness

of the original model assumptions, i.e., to test the reasonableness of the normal distribution of the error term, independence and equal variance assumptions.

(5) Using multiple linear regression equations for forecasting

The multiple linear regression model has better efficiency and accuracy in measuring the linear relationship between multiple independent variables and dependent variables under the condition of studying linear correlation. By fitting the linear function relationship between multiple independent variables and the dependent variable through regression analysis of historical data, it is a very widely used and effective forecasting model. The size of grid investment is affected by the joint action of multiple influencing factors, and the future grid investment can be effectively predicted by multiple linear regression modeling of multiple influencing factors and grid investment size with good significance test and fitting effect. The specific implementation process is shown in Figure 1.

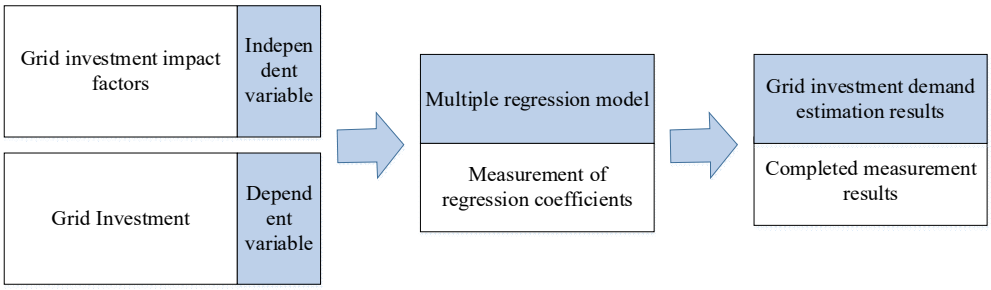

Fig. 1. Grid investment multiple linear regression model implementation process

Acknowledgments. This paper is supported by "the Science and Technology Project of State Grid Ningxia Electric Power Co.".

References

[1] Bechu Hu, Gang Hu, Chaohua Hu, et al. A grey prediction-based model for measuring power grid infrastructure investment[J]. Journal of University of Electronic Science and Technology,2013,42(6):890-894.

[2] Cambini C, Meletiou A, Bompard E, et al. Market and regulatory factors influencing smart-grid investment in Europe: Evidence from pilot projects and implications for reform[J]. Utilities Policy,2016,40:36-47.

[3] Lewis T M. The construction industry in the economy of Trinidad& Tobago[J]. Construction Management&Economics,2004,22(5):541-549.

[4] Huiru Zhao, Lu Yang, Chunjie Li, et al. Research on grid investment demand forecasting based on co-integration theory and error correction model[J]. Power Grid Technology,2011,35(9):193-198.

[5] Ying Zhou. Analysis of power network investment requirements for different development stages[J]. Journal of Power Systems and their Automation.2015,27(8):8-13

[6] Xueyan Mao, Liying Sun, Junv Ge, et al. Research on private economy electricity demand forecast based on co-integration theory[J]. East China Electric Power,2013,41(1):28-32.