

Research on Influencing Factors and Analysis Model Construction of Distribution Network Investment Demand

Shun Ma¹, Ming Chen¹, Kai Hou¹, Jinlan Hu¹, Chao Chen^{2*}

¹ Power Grid Planning & Research Center, Guangdong Power Grid Co., Ltd., Guangzhou City, Guangdong, 510000, China

² School of Economics and Management, North China Electric Power University, Beijing 102206, China

Abstract: In recent years, as the scale of investment in the distribution network of power grid companies has continued to remain high, there have also been problems of heavy investment and light income to varying degrees, and investment waste has occurred from time to time. Therefore, it is urgent to strengthen the level of investment efficiency of the distribution network. Based on the systematic analysis of the influencing factors of the investment demand of the distribution network, this paper constructs an analysis model of the investment demand of the distribution network to provide reference and reference for the power grid enterprises to improve the investment efficiency and efficiency level.

1 Introduction

The investment and construction of the distribution network is an important public infrastructure guarantee in the decisive stage of the country's grand goal of building a well-off society in all respects [1]. As an important part of the company's power grid, the distribution network is both the end of the grid and the front end of the service[2]. The total amount of assets is large, and there are many people engaged in the operation of the distribution network. Internally, it is the main link that affects the company's safe production and corporate benefits. Externally, it represents the company's image and level of serving economic and social development, and its role and status are extremely important and prominent. Historically, there has been a tendency to "emphasize the main network, neglect the distribution network, and emphasize equipment and service" for a long time. Many uncoordinated and unsuitable contradictions and problems have been formed in the management of the distribution network. In recent years, with the gradual increase in the scale of investment in the distribution network, the construction and management of the distribution network has gradually been paid attention to[3]. Due to the current severe internal and external operating situation, severe challenges have been raised to the level of input and output and the level of refined management in the construction of the distribution network. In the future, we must continue to strengthen the standardization and standardized management mode of distribution network construction, and improve the quality, efficiency and benefit of distribution network project construction.

For this reason, under the premise of systematically identifying the influencing factors of the investment demand of the distribution network, this paper constructs an investment demand analysis model based on the multivariate linear-BP neural network, and verifies the validity of the model through empirical analysis. The research results can provide support and reference for improving the investment benefit and efficiency of the distribution network. The research ideas of this paper are shown in Figure 1 below:

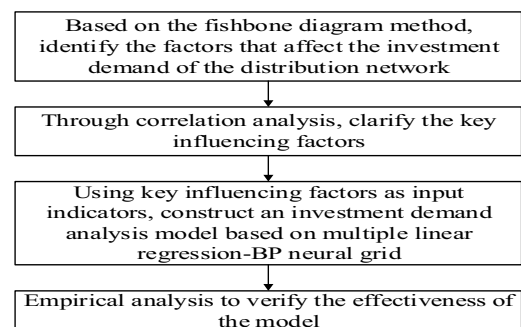


Figure 1. The main ideas of this paper.

2 Analysis of influencing factors of distribution network investment demand

2.1 Identification of factors affecting distribution network investment based on fishbone diagrams

Based on the theory of fishbone diagram, this paper conducts a systematic and comprehensive identification and analysis of internal and external factors affecting the

*Corresponding author's e-mail: 1182306191@ncepu.edu.cn

investment demand of the distribution network. The specific analysis is shown in Figure 2:

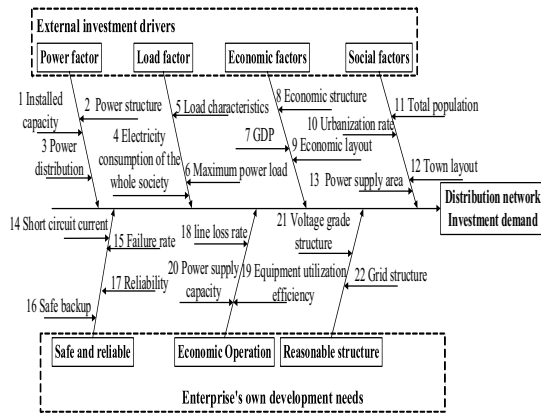


Figure 2. Factors affecting the investment demand of the distribution network.

2.2 Correlation analysis

Correlation analysis is a commonly used statistical method to study the correlation between multiple random variables. This method quantifies the closeness of the relationship between objective things by calculating the principle of correlation coefficient. There are many correlation analysis methods. This article takes the Pearson correlation coefficient analysis method as an example to quantitatively analyze the key factors of the investment demand of the distribution network. The main calculation principles of Pearson's correlation coefficient analysis method are as follows:

(1) Calculate the standard deviation of each variable itself, namely:

$$s_x = \sqrt{(n-1)^{-1} \sum (x_i - \bar{x})^2} \quad (1)$$

$$s_y = \sqrt{(n-1)^{-1} \sum (y_i - \bar{y})^2} \quad (2)$$

(2) Calculate the variance between variables, namely:

$$s_{xy} = \left[\sum (X_i - \bar{X}) (Y_i - \bar{Y}) \right] \cdot (n-1)^{-1} \quad (3)$$

(3) Calculate the correlation coefficient between variables, namely:

$$r_{xy} = s_{xy} (s_x \cdot s_y)^{-1} \quad (4)$$

(4) Judge the correlation of variables according to the correlation coefficient. The value range of the correlation coefficient is between 1 and -1, -1 means a complete negative correlation, 1 means a complete positive correlation, the closer the absolute value is to 1, the stronger the correlation, and 0 means the two variables are not correlated. Refer to Table 1 below for the correlation criterion:

Table 1. Correlation Judgment Table.

Correlation coefficient (absolute value)	Classification	Correlation coefficient (absolute value)	Classification
<0.3	Weak correlation	0.5~0.8	Significant correlation
0.3~0.5	Moderately related	0.8~1	Strong correlation

Combining the analysis conclusions of the factors affecting distribution network investment in Chapter 2.1, and through correlation calculations, the key factors affecting the investment demand of the distribution network include: GDP, total population, electricity consumption of the whole society, (power) installed capacity, and urbanization rate 5 key factors.

3 Construction of a distribution network investment demand analysis model based on multiple linear regression-BP neural network

3.1 Basic principles of multiple linear regression models

Regression Analysis refers to the use of statistical data to analyze the quantitative changes of various variables, and to reflect and describe this relationship in the form of regression equations. If there are multiple independent variables involved in the regression analysis, it is a multiple linear regression analysis, and the result obtained is called a multiple linear regression equation. Due to the diversity of factors affecting the cost of power transmission and transformation projects, multiple linear regression analysis is often used in factor analysis. The mathematical model of multiple linear regression is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (5)$$

In the above two formulas, ε refers to the change part of y caused by other random factors; $\beta_1, \beta_2, \dots, \beta_p$ is the regression coefficient of x_1, x_2, \dots, x_p . Because the regression coefficient is the partial derivative of y with respect to x, it is also called partial regression coefficient.

Based on multiple linear regression distribution network investment demand analysis steps are as follows:(1)Determine the independent variable in the regression equation, that is, the influencing factor. The dependent variable y is the cost level. Collect sample data.(2)Estimate each parameter in the model under certain statistical fitting criteria, and obtain the regression equation.(3)Test the regression coefficient. Analyze the results.

The SPSS software can be used for linear regression analysis. Select the "Analysis"->"Regression"->"Linear" command. Input the independent variable and the dependent variable to get the regression coefficient table. The table gives the non-standard coefficient, standard coefficient, T statistic, significance, and 95% confidence interval of the non-standard coefficient. Generally, the significance is less than 0.3, and it can be considered that x_i has a significant effect on y. Regression coefficients and correlation coefficients are both methods used to measure the relationship between variables, their

meanings are similar, and the coefficients must be tested for significance.

3.2 Basic Principles of BP Neural Network

The basic principle of the BP neural network model is: the network level of the neural network includes input layer, output layer, single or multiple hidden layers. The upper and lower layers are fully connected, and the neurons in each layer are independent of each other, and no connection is generated. The input signal x_i enters from the input layer, passes through the nodes in the intermediate hidden layer, acts on the output node, and undergoes a non-linear transformation to produce the output signal y_j . Each sample of the network training includes the input vector x_i and the expected output Y , and the network output value y_j . The deviation between y_j and the expected output value Y , according to the error, adjust the connection value w_{ij} between the input layer node and the hidden layer node, the connection t_{ij} between the hidden layer node and the output node, and the threshold, so that the error decreases in a gradient direction. Repeat Learn and train, determine the network parameters (weights and thresholds) corresponding to the minimum error, and then the training will stop. At this time, the trained neural network can process the non-linearly transformed information with the smallest output error on the input information of similar samples.

3.3 Empirical analysis

This article is based on the historical data of the distribution network investment scale of my country's Z city from 2010 to 2019 for predictive analysis. As shown in the figure below, with the continuous economic development of City Z, its distribution network investment level has shown a steady growth trend.

Table 2. Basic data table.

years	Investment (100 million yuan)	GDP (100 million yuan)	Electricity consumption of the whole society (100 million kWh)	Installed capacity (Ten thousand kilowatts)	Urbanization rate (%)
2010	2.89	149.56	17.73	249.74	46.98%
2011	3.89	160.17	18.25	276.12	48.33%
2012	3.45	191.63	20.95	303.58	49.96%
2013	3.69	227.05	23.50	327.65	51.37%
2014	3.67	245.61	25.67	359.34	52.57%
2015	3.86	277.95	26.17	391.48	53.74%
2016	4.12	308.06	27.85	386.87	54.78%
2017	4.64	346.15	28.45	388.78	56.06%
2018	5.43	384.22	29.88	390.81	57.35%
2019	5.32	427.03	31.52	392.75	58.50%

This article first uses the 2010-2017 data as the basic data to establish a multiple linear regression model (the independent variable is the absolute value of each factor without considering the unit, and the dependent variable is the distribution network investment scale), and the 2018-2019 data is used as the reference data to verify the model. Effectiveness. Under the 10% significance check level, the average error between the predicted value and the actual value through the multiple linear regression

model is 25.36%, indicating that the prediction error is large and the model needs to be corrected.

Table 3. Model fitting coefficients and significance test results.

variable	coefficient	P value
Constant term	6.012	0.1764
GDP	0.1825	0.0687
Total population	-0.0992	0.5370
Urbanization rate	195.7684	0.0060
Installed capacity	0.1097	0.0003
Electricity consumption of the whole society	-0.1980	0.0089

In order to further improve the prediction accuracy of the model, the BP neural network is used to correct the error of the fitting result of the multiple linear regression model. The residuals between the predicted value of grid investment calculated by the multiple linear regression model and the actual value are brought into the BP neural network model for fitting, and the error correction of the prediction result of the multiple linear regression model is obtained. After correction, the average error is reduced to 8.02%, indicating that the corrected model has better prediction accuracy. The revised model is used to forecast the distribution network investment demand from 2020 to 2022. The forecast results are shown in the following table 4:

Table 4. Forecast of investment scale in the next three years.

Years	2020	2021	2022
Predictive value (100 million yuan)	5.60	5.64	5.87

4 Conclusion

Based on the fishbone diagram method, this paper systematically identifies the factors that affect the investment demand of the distribution network from the two dimensions of the external environment of the power grid enterprise and the enterprise's own development. On this basis, a distribution network investment demand analysis model based on multiple linear regression-BP neural network is constructed. The research results can provide reference for power grid companies to make scientific and reasonable distribution network investment decisions.

References

1. Zeng Ming, Yu Zhuangzhuang, Wang Yuqing, Tian Shi, Cui Kai, Zhao Juan, Cui Rongjing. Distribution network investment scale optimization model based on transmission and distribution price constraints[J]. Smart Power, 2020, 48(11) :1-8.
2. Wang Ling, Wu Hongliang, Peng Daoxin. Grid investment scale prediction under the new round of transmission and distribution price supervision

cycle—Based on optimized support vector machine model[J]. *Technoeconomics*, 2020, 39(10): 45-53 .

3. Pu Di, Xu Xiaomin, Niu Dongxiao. Research on the investment strategy of complex grid under the promotion of smart grid [J]. *Complexity Science Management*, 2020(02): 51-64.
4. Luo Bin, Su Wenjun, Zhang Xiaoxuan, Luan Fengkui. Research on key influencing factors and transmission paths of power grid investment[J]. *Price Theory and Practice*, 2019 (05): 104-108+133.
5. Xu Xiaomin, Niu Dongxiao, Qin Honghao, Wu Han, Zhu Guorong. Research on the Influencing Mechanism of Internal and External Factors in the Investment Capacity of Power Grid Enterprises[J]. *Management Review*, 2017, 29(03): 74-89.