Mathematical research on the Crucial Factors Affecting Substation Engineering

Lingyun Li \(^{a}\), Xiaohan Guo \(^{b}\), Xinju Guo \(^{c}\), Xuyang Li \(^{d}\)

State Grid Henan Economic Research Institute, Zhengzhou, China

\(^{a}\)454174972@qq.com, \(^{b}\)496451522@qq.com, \(^{c}\)lyzz1989@126.com, \(^{d}\)379523936@qq.com

Abstract. The design plan of the substation engineering is crucial to the construction and operation of the substation. Reasonable mathematical analysis is helpful for the comparison and decision-making of the substation engineering in the early stage. In this paper, based on the parameter index of the 2016-2018 substation engineering, 24 factors affecting the substation engineering are selected, and a mathematical model is constructed for research. The study found five key factors that affect the substation engineering. The research results can provide reference for the scheme comparison and scientific decision-making of substation engineering.

Key words: Substation engineering; mathematical model; scheme comparison.

1. Introduction

Power transformation project cost includes building engineering project, equipment purchase expense, installation cost and other expenses. Due to the characteristics of engineering cost and its valuation, the factors affecting the cost of power transformation project are very extensive, and the relationships among them are intricate and complex. Macro factors, such as the level of social and economic development, market supply and demand, technical level of power grid construction, government behavior, etc., and micro factors, such as the level of the professional technicians, electrical equipment selection, engineering geological and hydrological conditions, will directly or indirectly affect the cost of power transformation project \([1,2]\). How to rationally control the cost of substation project and help the company to make accurate investment has become an urgent problem to be solved.

At present, domestic studies on the influencing factors of power transformation project are mostly focus on qualitative analysis or quantitative analysis of a few samples \([3,4]\), and analyses based on a large amount of historical engineering data are relatively rare. Based on the final accounting data of the transformation project cost of a province from 2016 to 2018, this paper USES the principal component analysis method to analyze the correlation among the factors affecting the transformation project cost, and extracts the most critical factors affecting the cost from many factors, so as to provide direction for the company to control the transformation project cost reasonably.
2. Principal component analysis method

Principal Component Analysis (PCA) is a multivariate statistical method that examines the correlation among multiple variables, and studies how to reveal the internal structure among multiple variables through a few Principal components, that is, to derive a few Principal components from the original variables, so that they can retain as much information about the original variables as possible [5]. This method reduces the complex factors to several principal components while introducing multiple variables, simplifying the problem and obtaining more scientific and effective data information at the same time.

The main steps of principal component analysis method are as follows:

(1) Standardized processing of the original data:

\[ x_{ij} = \frac{a_{ij} - \mu_j}{s_j}, (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \]  

Type: \( \mu_j = \frac{1}{n} \sum_{i=1}^{n} a_{ij}, s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (a_{ij} - \mu_j)^2} \).

Accordingly, standardized indicator variables as follows:

\[ \widetilde{x}_j = \frac{x_j - \mu_j}{s_j}, (j = 1, 2, \ldots, m) \]  

(2) Calculate the correlation coefficient matrix R.

\[ R = (r_{ij})_{n \times m} \]  

\[ r_{ij} = \frac{\sum_{i=1}^{n} \widetilde{x}_{ij} \cdot \widetilde{a}_{ij}}{n-1}, (i, j = 1, 2, \ldots, m) \]  

Type: \( r_{ii} = 1, r_{ij} = r_{ji} \). \( r_{ij} \) is the correlation coefficient between the ith index and the jth index.

(3) Calculate the eigenvalues and eigenvectors of the correlation coefficient matrix.

The solution of the characteristic equation \( |\lambda I - R| = 0 \) gives the eigenvalue \( \lambda_i, (i = 1, 2, \ldots, m) \), \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0 \); and then find out the corresponding eigenvectors \( u_i, (i = 1, 2, \ldots, m) \) for the eigenvalues of \( \lambda_i \), among them: \( u_j = (u_{1j}, u_{2j}, \ldots, u_{mj})^T \), M new index variables composed of eigenvectors are:

\[ y_1 = u_{11} \overline{x}_1 + u_{12} \overline{x}_2 + \cdots + u_{1m} \overline{x}_m \\
y_2 = u_{21} \overline{x}_1 + u_{22} \overline{x}_2 + \cdots + u_{2m} \overline{x}_m \\
\vdots \\
y_m = u_{m1} \overline{x}_1 + u_{m2} \overline{x}_2 + \cdots + u_{mm} \overline{x}_m \]  

Type: \( y_1 \) is the 1st principal component, \( y_2 \) is the 2nd principal component, and \( y_m \) is the Mth principal component.

(4) Select \( p (p \leq m) \) principal components to calculate the comprehensive evaluation value.

Calculate the information contribution rate and cumulative contribution rate of the eigenvalue \( \lambda_j, (j = 1, 2, \ldots, m) \), and use \( b_j \) to represent the information contribution rate of the principal component \( y_m \), then there is:
Use $\rho_a$ represents the cumulative contribution rate of principal component $y_1, y_2, ..., y_p$, there is:

$$a_p = \frac{\sum_{k=1}^{m} \lambda_k}{\sum_{k=1}^{m} \lambda_k}$$

(7)

If $a_p$ is close to 1 (generally, the range of $a_p$ is 85-95%), then the former $p$ indicator variables $y_1, y_2, ..., y_p$ are used as $p$ principal components, instead of the original $m$ indicator variables, and then $p$ principal components are comprehensively analyzed.

(5) Calculate the comprehensive score.

Use $b_j$ to represent the information contribution rate of the $j$th principal component, then there is:

$$Z = \sum_{j=1}^{p} b_j y_j$$

(8)

According to the comprehensive score, evaluate the principal component of the cost.

3. Identification of key influencing factors of power transformation project cost based on PCA

3.1. Sorting out the factors affecting the cost of power transformation project

There are many factors affecting the cost of power transformation project. Based on the grid project cost data of a province from 2016 to 2018, 24 factors affecting the cost of power transformation project were obtained through preliminary sorting. There are voltage grade, whether for smart substation, topography, pollution severity, altitude, type of power transformation project, single main transformer capacity, unit price of main transformer, the number of outgoing circuit on the high voltage side, the number of outgoing circuit on the low voltage side, main transformer stations in current, high voltage side distribution pattern, high voltage side circuit breaker sets and unit price, low voltage side of distribution type, low voltage side circuit breaker sets and the unit price, high voltage side connection type, low voltage side connection type, quantity and price of low voltage capacitors, electrical secondary equipment such as measurement and control, Intelligent related equipment, the cost of leveling the site, the cost of treating the foundation and the cost of entering the station.

3.2. Selection of key influencing factors of power transformation project cost based on PCA

After preliminary sorting out the influencing factors of power transformation project cost, it was found in this paper that not all the indicators had a significant impact on power transformation project cost, and some indicators fluctuated within a specific range only had a small impact on the project cost. Therefore, principal component analysis was carried out by SPSS software, and the most significant and key influencing factors on power transformation project cost were extracted from many factors by processing the indicator data set of the past three years.

3.2.1. Calculate the correlation matrix of influencing factors. Firstly, standardized the sample data from 2016 to 2018. Then, the sample data were input into SPSS for principal component analysis, and calculate the correlation coefficients among the influencing factors to describe the degree of correlation among the influencing factors. The correlation matrix of the influencing factors obtained by calculation is shown in table 1:
Table 1. Correlation matrix of principal component analysis of influencing factors of power transformation project cost

|                   | Nominal Voltage | Intelligent substation | Substation type | The number of main transformers in this period | Transformer capacity | The number of outgoing circuit on the high voltage side | The number of outgoing circuit on the low voltage side | High voltage side distribution pattern |
|-------------------|-----------------|-----------------------|-----------------|-----------------------------------------------|---------------------|-----------------------------------------------------|--------------------------------------------------|--------------------------------------|
| Nominal Voltage   | 1.000           | .064                  | -.163           | .323                                          | .713                | .468                                                | -.332                                            | .085                                 |
| Intelligent substation | .064          | 1.000                | -.007           | -.073                                         | -.071               | -.049                                               | .048                                              | -.146                                |
| Substation type   | -.163           | -.007                 | 1.000           | .311                                          | -.157               | .025                                                | .503                                              | .745                                 |
| The number of main transformers in this period | .323           | .073                  | .311            | 1.000                                         | .244                | .295                                                | .320                                              | .257                                 |
| Transformer capacity | .713           | -.071                 | -.157           | .244                                          | 1.000               | .943                                                | .448                                              | -.362                                |
| The number of outgoing circuit on the high voltage side | .468           | -.049                 | .025            | .320                                          | .448                | .471                                                | 1.000                                             | -.019                                |
| The number of outgoing circuit on the low voltage side | -.332          | .048                  | .503            | .257                                          | -.362               | -.366                                               | -.019                                             | 1.000                                |
| High voltage side distribution pattern | .085           | -.146                 | .745            | .395                                          | .185                | .150                                                | .114                                              | .414                                 |

If the correlation coefficient of two factors is close to 1, the stronger the correlation between the two factors is. As can be seen from table 1, there is a strong correlation between many relevant factors, indicating that there is information overlap in the sample data. It is necessary to eliminate multicollinearity and reduce dimension for PCA.

3.2.2. Calculate the total variance of the interpretation. The corresponding eigenvalues of each principal component are obtained through calculation, and the eigenvalues are reordered in the order from large to small to obtain the gravel diagram, as shown in figure 1. If the eigenvalue is less than 1, it can be proved that the influence strength of the basic variable is greater than the principal component factor, and the principal component factor should be deleted to extract the principal component whose characteristic root is greater than 1.

The information interpretation of extracted principal components can be clearly obtained by the machete diagram of principal component analysis of influencing factors of power transformation project cost. Seven factors with characteristic values greater than 1 can be extracted from the influencing factors of transformation project cost from 2016 to 2018.

Using the orthogonal rotation method of variance maximization, the total variance of interpretation are obtained, as shown in table 2.
Table 2. The total variance of PCA interpretation affected by power transformation project cost

| Ingredients | Summation | The total variance of the interpretation | Extract the sum of squares and load | Rotation squared and loading |
|-------------|-----------|----------------------------------------|------------------------------------|-------------------------------|
|             | Summation | The percent of variance | cumulative percentage | Summation | The percent of variance | cumulative percentage | Summation | The percent of variance | cumulative percentage |
| 1           | 7.483     | 31.178                   | 31.178                          | 7.483     | 31.178                   | 31.178                          | 6.195     | 25.814                   | 25.814                          |
| 2           | 3.147     | 13.112                   | 44.29                           | 3.147     | 13.112                   | 44.29                           | 2.948     | 12.285                   | 38.099                          |
| 3           | 1.747     | 7.281                    | 51.571                          | 1.747     | 7.281                    | 51.571                          | 2.129     | 8.87                     | 46.969                          |
| 4           | 1.448     | 6.035                    | 57.606                          | 1.448     | 6.035                    | 57.606                          | 1.875     | 7.814                    | 54.783                          |
| 5           | 1.327     | 5.531                    | 63.136                          | 1.327     | 5.531                    | 63.136                          | 1.596     | 6.65                     | 61.433                          |
| 6           | 1.113     | 4.639                    | 67.776                          | 1.113     | 4.639                    | 67.776                          | 1.375     | 5.73                     | 67.163                          |
| 7           | 1.003     | 4.181                    | 71.957                          | 1.003     | 4.181                    | 71.957                          | 1.151     | 4.794                    | 71.957                          |
| 8           | 0.865     | 3.602                    | 75.559                          |           |                         |                                 |           |                           |                                 |
| 9           | 0.821     | 3.422                    | 78.981                          |           |                         |                                 |           |                           |                                 |
| 10          | 0.762     | 3.176                    | 82.156                          |           |                         |                                 |           |                           |                                 |
| 11          | 0.662     | 2.759                    | 84.916                          |           |                         |                                 |           |                           |                                 |
| 12          | 0.609     | 2.539                    | 87.455                          |           |                         |                                 |           |                           |                                 |
| 13          | 0.554     | 2.307                    | 89.762                          |           |                         |                                 |           |                           |                                 |
| 14          | 0.454     | 1.89                     | 91.652                          |           |                         |                                 |           |                           |                                 |
| 15          | 0.416     | 1.734                    | 93.386                          |           |                         |                                 |           |                           |                                 |
| 16          | 0.328     | 1.366                    | 94.752                          |           |                         |                                 |           |                           |                                 |
| 17          | 0.305     | 1.271                    | 96.022                          |           |                         |                                 |           |                           |                                 |
| 18          | 0.251     | 1.046                    | 97.068                          |           |                         |                                 |           |                           |                                 |
| 19          | 0.197     | 0.819                    | 97.887                          |           |                         |                                 |           |                           |                                 |
| 20          | 0.19      | 0.793                    | 98.681                          |           |                         |                                 |           |                           |                                 |
| 21          | 0.14      | 0.582                    | 99.263                          |           |                         |                                 |           |                           |                                 |
| 22          | 0.082     | 0.342                    | 99.604                          |           |                         |                                 |           |                           |                                 |
| 23          | 0.072     | 0.301                    | 99.905                          |           |                         |                                 |           |                           |                                 |
| 24          | 0.023     | 0.095                    | 100                             |           |                         |                                 |           |                           |                                 |

Extraction method: principal component analysis.

As can be seen from table 2, the cumulative variance of the first seven principal components reach 71.957%, which can explain most of the information about the factors affecting the cost of power transformation project.

3.2.3. Selection of key factors. Seven principal components were obtained through the aggregate diagram and the total variance of the interpretation, and the information contained therein could represent the influencing factors of the transformation project cost in the past three years. The rotating component matrix was further obtained through SPSS software, as shown in table 3. The values in the component rotation matrix are the correlation coefficient between each variable and each principal component. By comparing the correlation coefficient between the first seven principal components and each variable, the main variables with large correlation coefficients can be selected, and the key influencing factors of the transformation project cost can be obtained.
Table 3. Rotating component matrix of principal component analysis on the influence of substation engineering cost

| Ingredients                                           | Rotation component matrix<sup>a</sup> | 1  | 2  | 3  | 4  | 5  | 6  | 7  |
|------------------------------------------------------|--------------------------------------|----|----|----|----|----|----|----|
| Nominal Voltage                                      |                                      | 0.65 | -0.136 | 0.347 | 0.33 | 0.086 | 0.206 | 0.034 |
| Intelligent substation                               |                                      | -0.095 | 0.008 | -0.004 | -0.038 | 0.732 | 0.073 | 0.053 |
| Substation type                                       |                                      | -0.085 | 0.834 | 0.07 | -0.111 | 0.008 | -0.021 | 0.113 |
| The number of main transformers in this period        |                                      | 0.195 | 0.469 | 0.257 | 0.574 | -0.102 | 0.089 | -0.026 |
| Transformer capacity                                  |                                      | 0.932 | -0.087 | 0.27 | 0.066 | 0.015 | -0.019 | 0.034 |
| Unit price of main transformer                        |                                      | 0.86 | -0.12 | 0.32 | 0.207 | 0.091 | 0.044 | 0.05 |
| The number of outgoing circuit on the high voltage side|                                      | 0.267 | 0.064 | 0.829 | 0.157 | 0.026 | -0.027 | -0.112 |
| The number of outgoing circuit on the low voltage side|                                      | -0.341 | 0.754 | -0.033 | -0.002 | 0.051 | 0.02 | -0.177 |
| High voltage side distribution pattern                |                                      | 0.287 | 0.839 | 0.003 | 0.029 | -0.156 | 0.017 | 0.132 |
| Number of circuit breakers at high voltage side       |                                      | 0.426 | 0.091 | 0.774 | 0.242 | 0.003 | -0.096 | 0.013 |
| Unit price of circuit breaker at high voltage side    |                                      | 0.818 | 0.182 | -0.106 | 0.205 | -0.293 | 0.026 | 0.046 |
| Low voltage side distribution pattern                 |                                      | -0.717 | 0.02 | -0.22 | 0.207 | 0.282 | -0.032 | -0.024 |
| Number of circuit breakers on low-voltage side        |                                      | -0.172 | 0.753 | 0.009 | 0.244 | 0.023 | -0.067 | -0.109 |
| Unit price of circuit breaker at low voltage side     |                                      | 0.655 | 0.083 | 0.123 | 0.047 | 0.263 | 0.096 | -0.085 |
| Number of low voltage capacitors                     |                                      | 0.192 | 0.086 | 0.329 | 0.821 | -0.042 | 0.066 | 0.026 |
| Unit price of low voltage capacitors                  |                                      | 0.01 | 0.02 | 0.102 | 0.093 | -0.132 | 0.096 | 0.072 |
| Unit price of secondary equipment                     |                                      | 0.745 | -0.129 | 0.41 | 0.117 | 0.121 | 0.774 | -0.024 |
| Unit price of intelligent related equipment           |                                      | 0.252 | -0.113 | -0.053 | 0.584 | 0.536 | 0.003 | -0.029 |
| The cost of leveling the site                         |                                      | 0.625 | -0.187 | 0.184 | 0.091 | 0.315 | -0.136 | 0.041 |
| The cost of treating the foundation                   |                                      | 0.505 | -0.062 | 0.128 | 0.012 | 0.491 | -0.131 | -0.008 |
| The cost of the road into the station                 |                                      | 0.839 | -0.082 | -0.14 | 0.185 | -0.133 | -0.063 | -0.046 |
| Altitude                                             |                                      | -0.071 | 0.086 | -0.114 | -0.198 | 0.219 | 0.237 | 0.673 |
| Topography                                            |                                      | 0.052 | -0.061 | -0.166 | -0.006 | 0.176 | 0.724 | -0.119 |
| Pollution severity                                    |                                      | 0.063 | -0.081 | 0.007 | 0.159 | -0.113 | -0.238 | 0.754 |

Extraction method: principal component.
Rotation method: with Kaiser's standardized orthogonal rotation method. A. Rotation converges after 8 iterations.

As can be seen from table 3, in the column one, the coefficient of main transformer capacity is the largest, shows that the first principal component is the main transformer capacity, the second column, high pressure coefficient of lateral distribution pattern is the largest, shows that the second principal component is high voltage side distribution pattern and so forth, to get the third principal component is the current high side line number, the fourth principal component is the number of low voltage capacitor, the fifth principal component is intelligent substation, the sixth principal component is the secondary equipment unit price, the seventh main component is the pollution severity.
To sum up, the key influencing factors of substation project cost selected through principal component analysis are: main transformer capacity, high-voltage side distribution pattern, number of high-voltage side outlet in this period, number of low-voltage capacitors, intelligent substation, unit price of secondary equipment and pollution severity.

Verification of key influencing factors of substation project cost

In this paper, multiple linear regression method is used to verify whether the above seven factors will be the main factors affecting the cost of power transformation project in the past three years. The seven key influencing factors obtained from the principal component analysis were taken as independent variables, and the transformation project cost was taken as the dependent variable for the regression analysis. The regression analysis results were shown in table 4-6.

Table 4. Fitting degree of multiple linear regression analysis of power transformation project

| Model | R    | R square | Adjust the R square | The standard estimate error |
|-------|------|----------|---------------------|-----------------------------|
| 1     | 0.978a | 0.957    | 0.956               | 2836.2596260               |

Table 5. Significance test results of multiple linear regression analysis of power transformation project

| Model        | Sum of squares | df | Mean square | F      | Sig.  |
|--------------|----------------|----|-------------|--------|-------|
| Regression   | 41784018046.043 | 7  | 5969145435.149 | 742.028 | .000a |
| Residual     | 1858249161.875  | 231| 8044368.666  |        |       |
| Aggregate    | 43642267207.918 | 238|             |        |       |

Table 6. Coefficients of multiple linear regression equation in power transformation project

| Model                               | Coefficienta Unstandardized coefficients | Standardized coefficients | t | Sig.  | 60.0% confidence interval for B |
|-------------------------------------|------------------------------------------|---------------------------|---|-------|--------------------------------|
| (Constant)                          | -1710.577                                | 1175.671                  | -1.455 | .147 | 2701.879 -719.276             |
| The capacity of the main transformer| 54.397                                   | .971                      | .875 | 56.032 | 53.579 55.216                |
| High voltage side distribution pattern | 333.826                                | 208.360                   | .023 | 1.602 | 158.140 509.511             |
| Number of low voltage capacitors    | 1712.045                                | 136.681                   | .189 | 12.526 | 1596.798 1827.292            |
| Intelligent substation              | 189.603                                 | 526.647                   | .005 | .360  | -254.455 633.661            |
| Unit price of intelligent related equipment | 195.170                              | 72.587                    | .037 | 2.689 | 133.966 256.374             |
| Number of circuit breakers at high voltage side | 594.724                              | 225.747                   | .042 | 2.634 | 404.379 785.070             |
| Pollution severity                  | -172.046                                | 186.410                   | -.013 | -.923 | -329.223 -14.868           |

As can be seen from table 4, the adjusted R2 is 0.956, indicating that the seven key influencing factors obtained through principal component analysis can explain the cost level of the power transformation
project of 95.6% in the recent three years, with a high degree of fitting. At the same time, it can be seen from table 5 that the results of the multiple regression model are significant, that is, there is a significant relationship between the cost of the transformer project in the past three years and the main transformer capacity, the distribution pattern of the high-voltage side, the number of outgoing circuit on the high voltage side in the current period, the number of low-voltage capacitors, intelligent substation, unit price of secondary equipment, pollution severity and other indicators.

The coefficient of multiple linear regression equation shows the significance of each influence factor on the cost of power transformation project. The smaller the Sig. Value is, the correlation between the influence factor and the cost of substation engineering exists, that is, the greater the influence of the influence factor on the cost is. The coefficient of multiple linear regression equation shows the significance of each influence factor on the cost of power transformation project. The smaller the Sig. Value is, the correlation between the influence factor and the cost of substation engineering exists, that is, the greater the influence of the influence factor on the cost is. As can be seen from table 6, for the cost of power transformation project in the past three years, the main transformer capacity, distribution pattern at high voltage side, the number of capacitors, unit price of secondary equipment, and the number of outgoing circuit on the high voltage side in this period are obviously small (Sig. Value), which is very close to zero, indicating that these five factors have a significant impact on the cost of power transformation project. However, the Sig. Value of intelligent substation and pollution level is greater than 0.3. At the confidence severity, it can be considered that these two factors have no significant influence on the transformation project cost, so they are not considered as the key influencing factors.

According to the multiple linear regression analysis, the paper further screened out five key factors affecting the substation project cost, namely the main transformer capacity, the distribution pattern of the high-voltage side, the number of outgoing circuit on the high voltage side in the current period, the number of low-voltage capacitors, and the unit price of secondary equipment.

4. Conclusions and Suggestions

Based on a large number of final accounting data of power transformation project in a province in the past three years, this paper uses principal component analysis and multiple regression analysis to obtain five key influencing factors of power transformation project cost, namely the main transformer capacity, the distribution pattern of the high-voltage side, the number of outgoing circuit on the high voltage side in the current period, the number of low-voltage capacitors, and the unit price of secondary equipment. Managers should focus on these key influencing factors and demonstrate such parameters as fully as possible in the early stage, so as to control the cost of power transformation project reasonably and realize accurate investment.

References

[1] Wang Xiaojian, Zhu tinghan, Luo Yongchang, et al. Cost evaluation of power transmission and transformation project based on artificial immune optimization neural network [J]. Zhejiang Electric Power, 2018 (37). 62-67.

[2] Yu chi. Analysis and evaluation model research on influencing factors of power grid project cost [D]. Beijing: North China Electric Power University, 2016-12-23.

[3] Yu Xiaoyan, Cai Zhixiong, Shangguan Linlin, et al. Technical study on the influence factors of transmission and transformation cost based on whole-process management [J]. Technology and Industry, 2014 (14).79-82.

[4] Wang Jiao. Analysis of the main influencing factors on the cost of 500 kv substation engineering [J]. Journal of Engineering Management, 2012 (6). 67-69.

[5] Gao Ting, Liu Fang. Application of SPSS in project cost estimation [J]. Guangxi Urban Construction, 2007.101-103.