Evaluation of Low-Voltage Distribution Network Index Based on Improved Principal Component Analysis

Hanlu Fan 1, Suzhou Gao 2, Wenjie Fan 2, Yinfeng Zhong 2 and Lei Zhu 2

1School of Electrical Engineering, Southeast University, Nanjing 210000, China
2State Grid Suqian Power Supply Company, Suqian 223800, China

*Corresponding author e-mail: 2464701020@qq.com

Abstract. In order to evaluate the development level of the low-voltage distribution network objectively and scientifically, chromatography analysis method is utilized to construct evaluation index model of low-voltage distribution network. Based on the analysis of principal component and the characteristic of logarithmic distribution of the index data, a logarithmic centralization method is adopted to improve the principal component analysis algorithm. The algorithm can decorrelate and reduce the dimensions of the evaluation model and the comprehensive score has a better dispersion degree. The clustering method is adopted to analyse the comprehensive score because the comprehensive score of the courts is concentrated. Then the stratification evaluation of the courts is realized. An example is given to verify the objectivity and scientificity of the evaluation method.

1. Introduction

At present, the research on the evaluation methods of distribution network mainly focuses on the high and medium voltage distribution network, and the evaluation methods and cases of the low-voltage distribution network are still in the original stage [1]. Traditional grid index evaluation method and evaluation system are mainly based on the analytic hierarchy process [2], expert scoring method [3] and other subjective comprehensive evaluation method. This kind of approach can make full use of the actual work experience of the industry personnel, and it is easy to implement. Although the subjective comprehensive evaluation method is simple, it tends to ignore the characteristics of the data themself and the correlation between evaluation indexes. The objective comprehensive evaluation method mainly uses the characteristics of the data to evaluate the difference of the index data, which can better highlight the value of the data. In [4], the principal component analysis is used to screen out the main indexes in a series of indicators with the 10kV distribution network as the research object. It can objectively reflect the characteristics of the data and reduce the dimension of the evaluation index. In [5], TOPSIS method is used to analyze the cities with high level of distribution network construction. This paper constructs the evaluation model of low-voltage distribution area based on principal component analysis and uses the characteristics of principal component analysis to achieve the index reduction and correlation elimination. The improved algorithm is proposed based on the conventional principal component analysis combining with the distribution characteristics of the data of the courts, and the comprehensive score of the index is obtained. Finally, cluster analysis is conducted according to the comprehensive score and the stratification evaluation of the low-voltage district is realized.
2. Improved Principal Component Analysis

2.1. The Mathematical Model of Principal Component Analysis

Principal component analysis (PCA) is put forward by Lord Kelvin in 1901, and it is a kind of mathematical method of dimension reduction. Its principle is to find several synthetic variables to replace the many variables, so that these comprehensive variables can represent as much information of the original variables as possible, and they are not related to each other. For sample, the variable \( x_1, x_2, \ldots, x_p \) is observed.

\[
\begin{align*}
F_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p \\
F_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p \\
&\vdots \\
F_p &= a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pp}x_p
\end{align*}
\]  

(1)

The model is required to meet the following conditions:
1. \( F_1 \) and \( F_2 \) are unrelated.
2. The variance of \( F_1 \) is bigger than the variance of \( F_2 \), and so on.

Principal component analysis can eliminate related effects between evaluation index after forming the independent principal components based on the original index variable. Moreover, the higher the correlation between the indicators is, the better effect of this method is.

2.2. Improved Principal Component Analysis

The traditional principal component analysis ensures that all data are not dimensionalized by means of standardization, but the standardized method is not suitable for multi-index comprehensive evaluation [6]. The standardization leads to the loss of the variation degree of difference between index data. At the same time, the principal component analysis is a method of using linear transform to reduce the dimension, and the effect of the reduction effect will be affected by the nonlinear index data.

At present, the improvement of principal component analysis is mainly reflected in two points mentioned above. In [7], the average method is adopted to replace the standardized method. The correlation coefficient of each index in the correlation coefficient matrix of data remains unchanged after dimensionalization. Therefore, the mean method can effectively preserve the characteristics of data. In [8], the method of logarithmic transformation is used to linearize the original data, and the parameters of the principal component are calculated by using correlation coefficient matrix according to the distribution of scatter plots of data.

In this paper, the method of "logarithmic centralization" is adopted to centralize the original data,

\[
y_j = \ln \frac{x_j}{\left( \prod_{k=1}^{p} X_k \right)^{1/n}}
\]  

(2)

Where \( p \) is the number of indexes, and \( n \) is the number of distribution areas.

3. Design of Evaluation Algorithm for Low-Voltage Distribution Network

3.1. Evaluation Index of Low-Voltage Distribution Network

The purpose of this paper is to evaluate the comprehensive efficiency of the low-voltage distribution network. This paper will not consider the superior power supply office in the distribution network, and only consider the characteristics of the courts. Refer to related documents issued by the state grid
corporation, including the inherent characteristics and operation management of low-voltage distribution network.

The main characteristics indexes of the low-voltage distribution network are the capacity, the power supply radius and the insulation rate. The operation and management indexes of low-voltage distribution network are mainly reflected in the quality of power supply. The quality of power supply includes power supply reliability, voltage qualification rate, the average load rate, maximum load rate, three-phase imbalance degree, average power factor, the line loss rate, power supply and electricity sales.

3.2. The Evaluation Algorithm Flow of Low-Voltage Distribution Network

Combining with low-voltage distribution network index evaluation system, the design theory of improved principal component analysis and its evaluation model, this paper proposes the evaluation scheme of low-voltage distribution network based on improved principal component analysis. Eventually, establish a comprehensive evaluation process of low-voltage distribution network, as shown in Figure 1.

![Comprehensive evaluation algorithm flow of low-voltage distribution network.](image)

The specific evaluation algorithm process is as follows. To begin with, select targeted evaluation index set combining with concrete examination field of the low-voltage distribution network, and collect the basic data of the low-voltage distribution network in accordance with the established index set. The original evaluation data are preprocessed and the logarithmic centralization is performed on the basis of dimensionless. Then enter the first round of judgment, and carry out the feasibility test of the principal component analysis. If the analysis of principal component analysis is not feasible, it is necessary to reselect the evaluation index system and repeat the previous process. If the validation is feasible, the principal component analysis model is adopted. Establish correlation coefficient matrix, solve eigenvalue and characteristic root and generate principal component expression, etc. Besides, the contribution rate is calculated to determine the number of principal component evaluation indexes, and the following comprehensive evaluation is conducted in two parts. Firstly, the comprehensive scores of the main components are calculated and the cluster analysis is carried out according to the comprehensive scores of the courts. So the courts are divided into different levels for evaluation. Secondly, cluster analysis is carried out based on the principal component factor loading matrix. The original evaluation index system is reclassified according to the categories according to the category,
and the actual physical meaning of the indexes is given, reflecting the multiple dimensions of the intelligent grid evaluation architecture.

4. Case Analysis

4.1. The Original Data

The overall situation of construction and development of 51 courts in a city are evaluated according to the algorithm design process of low-voltage distribution network characteristic evaluation, combining with the national power grid development planning strategy and construction target.

4.2. Calculation Results and Analysis

Data is attached as the research object and forms the initial matrix according to the low-voltage distribution network evaluation algorithm. Structure improved principal component analysis mathematical model and uses SPSS software to calculate. The principal component eigenvalue and variance contribution rate of 51 regional indexes are obtained, and the results are compared with the conventional principal component analysis, as shown in table 1.

| Principal Component | Conventional Principal Component Analysis | Improved Principal Component Analysis |
|---------------------|-------------------------------------------|--------------------------------------|
|                     | Eigenvalue | Variance Contribution Rate | Accumulated Contribution Rate | Eigenvalue | Variance Contribution Rate | Accumulated Contribution Rate |
| 1                   | 4.872      | 40.601                     | 40.601                           | 5.717      | 47.641                      | 47.641                           |
| 2                   | 1.395      | 11.627                     | 52.228                           | 2.896      | 24.134                      | 71.775                           |
| 3                   | 1.304      | 10.868                     | 63.096                           | 2.561      | 21.341                      | 93.117                           |
| 4                   | 1.130      | 9.416                      | 72.512                           |            |                            |                                  |
| 5                   | 1.051      | 8.754                      | 81.266                           |            |                            |                                  |
| 6                   | 0.998      | 8.320                      | 89.587                           |            |                            |                                  |

From table 1, the first six principal components can replace the original indexes. The cumulative contribution rate of the first three eigenvalues has reached 93.117%, indicating that the first three main components basically contain all the information of the indexes. The improved method of principal component analysis is more obvious. The corresponding principal component evaluation is calculated, which shown in table 2.

| Courts | Conventional Principal Component Analysis Score | Improved Principal Component Analysis Score |
|--------|-----------------------------------------------|---------------------------------------------|
| 1      | -0.169                                        | -0.134                                      |
| 2      | -0.667                                        | -0.728                                      |
| 50     | -0.469                                        | -0.467                                      |
| 51     | -0.168                                        | -0.201                                      |
The evaluation value of the comprehensive principal component reflects the relative level of the courts. The mean value of all object evaluation results is zero. Combining with table 2, it can be found that the conventional evaluation value of the main component analysis method is consistent with the improvement of the comprehensive evaluation value of principal component analysis. However, the improvement of the comprehensive evaluation value of the principal component analysis can clearly divide the data.

Considering the number of the courts is great, and the comprehensive score difference is not obvious. The method is equivalent to ward, the distance is the square of Euclidean distance. The clustering algorithm gathers the courts into three categories, and the category with the lowest scores needs to be improved. Then the original evaluation index is analyzed by using the principal component factor load matrix of the grid. The factor loading matrix is shown in table 3.

| No. | Evaluation Index          | First Factor Load | Second Factor Load | Third Factor Load |
|-----|--------------------------|-------------------|--------------------|-------------------|
| 1   | Capacity                 | 0.689             | -0.016             | 0.150             |
| 2   | Power Supply Radius      | 0.761             | -0.146             | 0.112             |
| 3   | Insulation Rate          | 0.919             | -0.123             | 0.284             |
| 4   | Power Supply Reliability | -0.480            | 0.733              | -0.342            |
| 5   | Voltage Qualification Rate | -0.505          | 0.760              | -0.379            |
| 6   | Average Load Rate        | -0.739            | 0.666              | 0.021             |
| 7   | Maximum Load Rate        | -0.739            | 0.665              | 0.021             |
| 8   | Three-phase Imbalance Degree | 0.233         | -0.001             | 0.948             |
| 9   | Average Power Factor     | -0.170            | -0.870             | 0.659             |
| 10  | Line Loss Rate           | 0.930             | -0.191             | 0.262             |
| 11  | Power Supply Rate        | 0.931             | -0.177             | 0.241             |
| 12  | Electricity Sales        | 0.930             | -0.186             | 0.263             |

According to the classification criteria of \( \text{cov}(F_i, z_j) > 0.5 \), the index 1~3 and index 10~12 can be divided into a class by the factor load of the first principal component. It is not difficult to find that these
six indexes reflect the economic properties of the low-voltage distribution network. In the same way, the indexes 4~8 can be divided into a class by the factor load of the second principal component, reflecting the power supply ability of the court. And the indexes 9~10 can be divided into a class by the factor load of the third principal component, reflecting the power supply quality. The hierarchical design of the original evaluation index system is shown in Figure 2.

5. Conclusion
This paper aims at the evaluation method of low-voltage distribution network. The method of "logarithmic centralization" is proposed to improve the conventional principal component analysis. The improved principal component analysis is combined with the clustering analysis method, applied to the index evaluation of low-voltage distribution network. The procedure of evaluating the indexes of low-voltage power distribution network is planned and a scientific and reasonable evaluation model is constructed. The evaluation method does not need to set the index weight artificially, so the final evaluation result is avoided because of subjective confirmation. The correlation between indicators is eliminated, making the evaluation more accurate. At the same time, the results show that the improved principal component analysis method not only has better health effect, but also can better divide the index comprehensive score level.

References
[1] LUO Guodong, MEI Min, LIN Zhenmin. Evaluation for Distribution Network Planning Program and Completed Project[J]. Shanxi Electric Power, 2011, 39(6): 16-20.
[2] WAN Wei, WANG Chun, CHENG Hong, et al. Preliminary frame of index system for evaluating power network [J]. Power System Protection and Control, 2008, 36(24): 14-18.
[3] LI Jianxiu, LI Lisheng, ZUO Xinbin, et, al. Application of energy efficiency evaluation model of distribution network based on improved AHP[J]. Electrical Applications, 2013(s2): 514-519.
[4] LI Jiahui, XU Qun, MA Jun, et, al. Evaluation of 10kV distribution network based on principal component analysis[J]. Science and Technology Information, 2016(03): 17-19.
[5] ZHAO Peipei. Distribution Network Management Evaluation and Optimization Based on TOPSIS Method[D]. Beijing: North China Electric Power University, 2014.
[6] JI Rongfang. Improvement of data processing method in principal component analysis[J]. Journal of Shandong uni-versity of Science and Technology, 2007, 26(5): 95-98.
[7] SUN D I H, ABE S, SHOULTS RR. Calculation of energy losses in a distribution system[J]. IEEE Transactions on Power Apparatus and Systems, 1980, PAS-99(4): 1347-1356.
[8] ZHANG Tao, CHENG Zhiyou, LIANG Dong, et al. The application of nonlinear PCA in synthetic evaluation of power quality[J]. Electrical Measurement & Instrumentation, 2008, 45(510): 5-9.