Differential Clustering Analysis of Power Grid Green Development Based on Hierarchical Cluster Method

Zhuonan Li*, Yanming Jin1, Jiujin Zhao2, Xue Tan1, Dongliang Zhang and Jia Wei3

1 State Grid Energy Research Institute CO., LTD. Beijing, 102209, China
2 State Grid Corporation of China, Beijing, China
3 Economic & Technology Research Institute, State Grid Shandong Electric Power Company, Jinan, China

*Corresponding author’s e-mail: lizhuonan@sgeri.sgcc.com.cn

Abstract. Under the call of ecological civilization construction, the internal provincial evaluation of power companies has also been gradually advanced. Taking into account the vast territory of my country, the large-scale power system, and the complex structure, power grids at all levels have differences in areas such as jurisdiction area, climatic conditions, power grid scale, and technical level. Under this premise, if a reasonable index system and evaluation method are not formulated, and a unified evaluation index system is adopted for each region, it may not be possible to achieve a relatively fair evaluation of the regional power grid. Under this condition, it is difficult to exert the guidance of the evaluation method on the green development of the power grid. In this paper, taking into account the differences that exist between all levels of power grids in our country, the hierarchical cluster method is used to cluster them.

1. Introduction
Under the background of ecological civilization construction and energy revolution strategy, the connotation and extension of green development of power grid will change. The green development function of power grid will extend from the power system to the entire energy system. It is necessary to explore and improve the theory of green development of power grid to adapt to the new situation. Moreover, under the premise that the national environmental supervision model is constantly changing, and all provinces and cities have adopted energy development plans and ecological environmental protection measures adapted to local conditions, the green development performance of the regional power grids also urgently need to be differentiated. Therefore, it is necessary to carry out a detailed classification of the provinces in the country based on the differentiated information of various regions.

2. Hierarchical cluster method

2.1. Overview
Hierarchical cluster method is a method for cluster analysis[1-3]. To perform the clustering process, it is necessary to regard each sample as a category at the beginning, and then group the closest samples (i.e., the group with the smallest distance) into small categories first. Then the aggregated small
categories according to the distance between the categories should be merged. After one group was clustered, the same process should be continued until all sub-categories were clustered into one big category. With the continuous attention and research of this method by experts and scholars from all over the world, this method has been applied in various fields and has made good contributions to statistics and clustering [4-6].

2.2. Calculation Process

Step 1: Suppose there are N samples in the initial pattern, and each sample forms its own category, that is, establish N categories,

\[ G_1^{(0)}, G_2^{(0)}, G_3^{(0)}, \ldots, G_n^{(0)} \]  

(1)

The distance between the various types (the initial distance is the distance between the samples) should be calculated, thus an N*N-dimensional distance matrix D(0) can be obtained. Here, the symbol (0) indicates the state before the clustering operation is started.

Step 2: Assuming that the distance matrix D(n) has been obtained in the previous step of clustering operation, and n is the number of successive clustering and merging, find the smallest element in D(n). If it is the distance between the two categories of G_i(n) and G_j(n), the two categories of G_i(n) and G_j(n) should be merged into one category.

\[ G_{ij}^{(n+1)} \]  

(2)

Thus a new category was created.

\[ G_1^{(n+1)}, G_2^{(n+1)}, G_3^{(n+1)}, \ldots, G_j^{(n+1)} \]  

(3)

Step 3: Calculate the distance between the new categories after merging, then D(n+1) was achieved. To calculate the distance between G_6^{(n+1)} and other G_i^{(n+1)}, G_2^{(n+1)}, G_3^{(n+1)}, \ldots, G_j^{(n+1)} that has not been merged, a variety of different distance calculation criteria can be used for calculation.

Step 4: Return to step 2, repeat the former calculation and merging until a satisfactory classification result is obtained (Such as: reaching the required number of clusters, or the smallest component in D(n) exceeds the given threshold D, etc.).

2.3. Internal Clustering

For the internal clustering process, the Ward method was adopted in this paper. In the calculation, the squared Euclidean distance was taken as the distance between the two categories, and each sample in the set was divided into one category, and the variance between the centers of gravity was calculated. Then, the two categories with the smallest increase in the sum of squared deviations was merged. Finally, all categories were thus merged level by level.

For example, there are n regional samples divided into k categories, and X ij represents the j-th sample in G i, n i represents the number of samples in G i, and X'i is the center of gravity of G i, so the squared Euclidean distance of the samples in G i is:

\[ S_i = \sum_{i=1}^{n_i} (X_i - X')^t (X_i - X') \]  

(4)

Then the internal squared Euclidean distance in k clusters is
3. Calculation case

According to the Measures for Evaluation and Assessment of Ecological Civilization Construction Goals issued by the General Office of the Central Committee of the Communist Party of China and the General Office of the State Council, the 26 provinces in the State Grid Corporation’s operating area were clustered to extract resource utilization, environmental governance, environmental quality, ecological protection, and growth. Six important indicators of quality and green life, as well as the scores of each province, form an initial matrix [7].

Table 1. Provincial scores

| Province | Resource utilization | Environmental governance | Environmental Quality | Ecological Protection | Growth quality | Green Life |
|----------|----------------------|--------------------------|----------------------|----------------------|---------------|------------|
| Beijing  | 82.92                | 98.36                    | 78.75                | 70.86                | 93.91         | 83.15      |
| Tianjin  | 84.4                 | 83.1                     | 67.13                | 64.81                | 81.96         | 75.02      |
| Hebei    | 83.34                | 87.49                    | 77.31                | 72.48                | 70.45         | 70.28      |
| Shanxi   | 78.87                | 80.55                    | 77.51                | 70.66                | 71.18         | 78.34      |
| Inner Mongolia | 79.99       | 78.79                    | 84.6                 | 72.35                | 70.87         | 72.52      |
| Liaoning | 76.69                | 81.11                    | 85.01                | 71.46                | 68.37         | 67.79      |
| Jilin    | 86.13                | 76.1                     | 85.05                | 73.44                | 71.2          | 73.05      |
| Heilongjiang | 81.3           | 74.43                    | 86.51                | 73.21                | 72.04         | 72.79      |
| Shanghai | 84.98                | 86.87                    | 81.28                | 66.22                | 93.2          | 80.52      |
| Jiangsu  | 86.89                | 81.64                    | 84.04                | 62.84                | 82.1          | 79.71      |
| Zhejiang | 85.87                | 84.84                    | 87.23                | 72.19                | 82.33         | 77.48      |
| Anhui    | 83.19                | 81.13                    | 84.25                | 70.46                | 76.03         | 69.29      |
| Fujian   | 90.32                | 80.12                    | 92.84                | 74.78                | 74.55         | 73.65      |
| Jiangxi  | 82.95                | 74.51                    | 88.09                | 74.61                | 72.93         | 72.43      |
| Shandong | 82.66                | 84.36                    | 82.35                | 68.23                | 75.68         | 74.47      |
| Henan    | 83.87                | 80.83                    | 79.6                 | 69.34                | 72.18         | 73.22      |
| Hubei    | 86.07                | 82.28                    | 86.86                | 71.97                | 73.48         | 70.73      |
| Hunan    | 83.7                 | 80.84                    | 88.27                | 73.33                | 77.38         | 69.1       |
| Chongqing| 84.49                | 79.95                    | 89.31                | 77.68                | 78.49         | 70.05      |
| Sichuan  | 84.4                 | 75.87                    | 86.25                | 75.48                | 72.97         | 68.92      |
| Xizang   | 75.43                | 62.91                    | 94.39                | 75.22                | 70.08         | 63.16      |
| Shaanxi  | 82.84                | 78.69                    | 82.41                | 69.95                | 74.41         | 69.5       |
| Gansu    | 85.74                | 75.38                    | 90.27                | 68.83                | 70.65         | 69.29      |
| Qinghai  | 82.32                | 67.9                     | 91.42                | 70.65                | 68.23         | 65.18      |
| Ningxia  | 83.37                | 74.09                    | 79.48                | 66.13                | 70.91         | 71.43      |
| Xinjiang | 80.27                | 68.85                    | 80.34                | 73.27                | 67.71         | 70.63      |
To systematically cluster this matrix, SPSS software was used to divide 27 provinces into 5 clusters. Taken province as an example, and six indicators and their values can be used as variables, thus the cluster can be performed. The following clustering conclusions can be obtained:

Table 2. Classification of cluster groups by province

| Case&Province | 5 Clusters | Case&Province | 5 Clusters |
|---------------|------------|---------------|------------|
| 1:Beijing     | 1          | 14:Jiangxi    | 3          |
| 2:Tianjin     | 1          | 15:Shandong   | 2          |
| 3:Hebei       | 2          | 16:Henan      | 2          |
| 4:Shanxi      | 2          | 17:Hubei      | 4          |
| 5:Inner Mongo | 2          | 18:Hunan      | 4          |
| 6:Liaoning    | 2          | 19:Chongqing  | 4          |
| 7:Jilin       | 3          | 20:Sichuan    | 3          |
| 8:Heilongjian | 3          | 21:Xizang     | 5          |
| 9:Shanghai    | 1          | 22:Shaanxi    | 2          |
| 10:Jiangsu    | 4          | 23:Gansu      | 4          |
| 11:Zhejiang   | 4          | 24:Qinghai    | 5          |
| 12:Anhui      | 2          | 25:Ningxia    | 2          |
| 13:Fujian     | 4          | 26:Xinjiang   | 3          |

Different from other clustering methods, the hierarchical clustering method supports the display of pedigree diagrams to more accurately characterize the clustering process of each case.

![Figure 1. Cluster pedigree map of each province](image-url)
4. Conclusion
According to the clustering results, Beijing, Tianjin and Shanghai, as large cities with abundant resources and sound protection policies, can be clustered into one category. Industrial provinces such as Hebei, Shandong, and Shanxi with more energy-intensive industries and relatively serious environmental pollution and ecological damage can be merged into one group. Jilin, Heilongjiang, Jiangxi and Sichuan are in the same category as provinces whose economic development is slightly slower than the former but relatively good in environmental protection. Jiangsu, Zhejiang, Hubei and other provinces that have shown a trend of high-quality economic development can be classified into one category. Qinghai and Tibet, as provinces whose economic development is still relatively lagging across the country but the ecological environment has not been destroyed, can be classified into one category.

Acknowledgments
Support by Science and Technology Project of State Grid Corporation of China (SGNY0000NGJS1900127): Research on Strategic Planning Technology and Evaluation Method of Green Power Development Based on Thought of Ecological Civilization

References
[1] Dimitri Pinel. Clustering methods assessment for investment in zero emission neighborhoods’ energy system[J]. International Journal of Electrical Power and Energy Systems, 2020, 121.
[2] S. Madadi, B. Mohammadi-Ivatloo, S. Tohidi. Probabilistic Small Signal Stability Evaluation of Power Systems with High Penetration of Wind Farms[J]. Computers and Electrical Engineering, 2020, 85.
[3] Jimmy Bordarie. A highlight of the relationship between themata and social representations using a hierarchical clustering method[J]. Review of Psychology, 2016.
[4] Lin-Chih Chen. Building a web-snippet clustering system based on a mixed clustering method[J]. Online Information Review, 2011, 35(4).
[5] Moez Davodi, Ehsan Reihani, Mehdi Davodi. Application of hierarchical clustering method for separation of power system[J]. Int. J. of Power and Energy Conversion, 2010, 2(3).
[6] Hisao Miyano. Hierarchical Clustering Method Using Probabilistic Similarity Measure[J]. The Behaviormetric Society, 1995, 22(2).
[7] Communique on the Annual Evaluation Results of Ecological Civilization Construction in 2016. http://www.gov.cn/xinwen/2017-12/26/content_5250387.htm