Construction of the Evaluation Index System of "Five Development Concepts" Based on R-type Clustering and Variation Coefficient in Hunan Province

Juan Xu*
Hunan Modern Logistics Vocational and Technical College, Changsha, China

*Corresponding author e-mail: xujuan0822@163.com

Abstract: On the basis of deeply understanding and grasping the connotation of the five major concepts of development, we should closely follow the important spirit of implementing General Secretary Xi Jinping's series of speeches. Drawing on domestic and foreign representative research results on the development of evaluation indicators systems and innovative, coordinated, green, open, and shared evaluation indicators, we conducted preliminary screening of evaluation indicators for sea elections based on the principle of observability. Then, the data of six central provinces were selected by combining R-cluster analysis and variational coefficient analysis to select 39 evaluation indicators, and the effectiveness evaluation index system of the "five major development concepts" suitable for the six central provinces was constructed. Then, according to the regional characteristics and characteristics of Hunan province, two indexes were added from the perspective of rational analysis, so as to construct a quantitative evaluation index system of the "five major development concepts" in Hunan province with 41 indexes.

1. Methods for the selection of evaluation indicators

1.1 Evaluation indicators for sea elections and initial screening

Combing domestic and foreign representative on the development of evaluation indicators research results, audition evaluation indicators. Then the method of observable screening index is used for the initial screening.

Because qualitative indicators are difficult to obtain in data collection, no matter how comprehensive and perfect the evaluation indicator system is designed, if the corresponding data are not collected, statistical evaluation will be difficult to carry out, and the establishment of a system of evaluation indicators is meaningless. Therefore, in order to ensure the availability of indicator data, the selection of evaluation indicators by sea is initially screened on the basis of the principle of observability, and those indicators that are not available except for data or that can not be obtained indirectly by calculation are deleted. In order to ensure that the selected evaluation indicators can be observed, obtained, and statistical analysis can be conducted. Through the observable screening, the indicators belong to the quantifiable quantitative evaluation indicator system, and can be further screened by statistical methods to achieve the simplicity and comprehensiveness of the evaluation indicator system.

1.2 Standardization of indicator data

When using multiple indicators for comprehensive evaluation, because of the different dimensions
and orders of magnitude of the evaluation indicators, if the original data indicators are directly used for analysis, the numerical indicators will have a great impact on the evaluation results. The effect of small numerical indicators on the evaluation results will be reduced. Therefore, in order to eliminate the impact of the dimension on evaluation analysis, the original data need to be standardized to harmonize the standards of the evaluation indicators.

Common data standardization methods include the maximum-minimum standardization method and the means-variance standardization method. This paper adopts the means-standardization method. The specific approach is as follows:

\[
\begin{align*}
    z_{ij} &= \frac{x_{ij} - \bar{x}_j}{s_j} \\
    \bar{x}_j &= \frac{1}{n} \sum_{i=1}^{n} x_{ij} \\
    s_j &= \left( \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2 \right)^{\frac{1}{2}}
\end{align*}
\]  

(1)

(2)

(3)

1.3 Method of screening and evaluating index for R-type cluster analysis

Cluster analysis is a multivariate statistical method to study the classification of multiple samples or multiple indicators. The classification of samples is called Q-type cluster analysis, and the classification of indicators is R-type cluster analysis. [1]

The basic principle of cluster analysis is the process of classifying different classes or clusters according to the degree of similarity. The classification of samples or indicators is actually a study of the relationship between samples or indicators. Commonly used research methods include similarity coefficients and distance methods. This study classifies indicators and chooses to use similar coefficient methods to classify indicators. The so-called similarity coefficient method is the indicator with more similar properties. The greater the similarity coefficient, the closer the absolute value of the similarity coefficient is to 1. Such indicators are grouped into one category; For indicators that are not similar in nature, the closer their similarity coefficient is to 0, they are classified into different categories according to the size of the similarity coefficient.

1.3.1. Steps for R-type cluster analysis

1) The original data is standardized using the means-standardization method. The specific treatment method can be found in the standardization of indicator data.
2) Calculate the matrix R of the correlation coefficient of the standardized data.
3) Classify the indicators according to the correlation coefficient, classify the indicators with a larger correlation coefficient (i.e., a greater degree of similarity) into one category, and classify other indicators with a smaller degree of similarity into one category until all the indicators are aggregated. Classify all indicators into different categories.
4) Determination of Rationality of Cluster Number

How to determine the number of categories? In other words, how many kinds of cores are there for P indicators? In this paper we use the square sum of spread in the system clustering method to cluster the index.

we cluster the index by the method of square sum of spread in the system clustering method.
5) Test on the Rationality of Cluster Number

In order to determine the rationality of the cluster number L, the non-parametric Kendall-W test is performed on each category of indicators after clustering. In order to avoid the subjective arbitrariness of the clustering analysis classification number L to timing.

The H0 of the non-parameter Kendall-W test is: There is no significant difference between different indicators of the same category. If significant level of Sig &gt; 0.05, then accept the original assumption; That is, there is no significant difference between the same type of indicators, and the number of clusters is reasonable; If significant level of Sig &lt; 0.05, then the original assumption H0 is rejected and regrouped.
2. Construction of the Evaluation Index System of the "Five Great Principles of Development" in Hunan Province

2.1 Audience evaluation indicators and initial screening

2.1.1. Evaluation indicators for sea elections. Combing domestic and foreign representative research results on development evaluation indicators, the foreign representative is the United Nations Development Programme (UNDP) founded by "Life expectancy, educational standards and quality of life" The Human Development Index (HDI) of the three basic variables interprets the imbalance between economic growth and social development[2]: The UN Committee on Sustainable Development established a sustainable development indicator system with social, environmental, economic, and institutional dimensions, and established its core indicator system in 2001[3]: United Nations Statistics Office (UNSTAT) proposed the framework for the System of Indicators for Environmentally-economically Sustainable Development (FISD) (FISD), the UN proposed in September 2000 with eight goals "Millennium Development Goals"[4] And related 18 targets and 48 indicators; Highly integrated sustainable development evaluation indicator system proposed by SCOPE and UNEP[5] Commission U.S. government And the British government[6] The establishment of sustainable development evaluation indicators system.

More representative of the country is that the Research Group on Sustainable Development of the Chinese Academy of Sciences proposed the China Sustainable Development Index System in the 1999 "China Sustainable Development Strategy Report". In 2011, the National Bureau of Statistics announced the Comprehensive Development Index (CDI)[8], which includes five dimensions of economic development, improvement of people's livelihood, social development, ecological construction, and scientific and technological innovation. After it was revised and improved into the Regional Development and People's Livelihood Index (DLI); In 2013, the National Bureau of Statistics issued the "System of Statistical Monitoring Indicators for the All-round Construction of a well-off Society" In addition, provinces and cities have also established relevant development indicator systems that are suitable for their own development characteristics, such as the Shenzhen "Benefit Shenzhen" index. Yang Xinhong Take Shenzhen City as an example to build a statistical evaluation index system of the "five major development concepts" and monitor and evaluate the effectiveness of the implementation of the new development concepts. Wangxiaoliang The evaluation index system for the level of regional open economic development has been established in terms of open foundation, open scale, open structure and open efficiency.

2.1.2. Preliminary screening of evaluation indicators. Looking at relevant documents and combing through them, on the basis of deeply understanding and grasping the connotation and significance of the five major development concepts, from the perspective of Hunan Province, Drawing on domestic and foreign representative research results on development evaluation indicators system and innovation, coordination, green, open, and shared evaluation indicators From multiple directions and multiple perspectives, 57 indicators for the evaluation of the effectiveness of the "five major development concepts" were selected.

According to the principle of observability, the evaluation indicators of the sea selection were initially screened, and the indicators that could not be obtained were deleted to ensure the quantitative nature of the post-R-type clustering and variability coefficient analysis. The evaluation indicator system after the initial screening is shown in the table below.
Table 1 The preliminary screening evaluation index system for the effectiveness of the "Five Development Concepts"

| Primary indicator | Secondary indicator | Three-level indicator | Indicator type | Observability |
|-------------------|---------------------|-----------------------|----------------|---------------|
| Innovation and development | Innovation investment | R&D expenditure intensity (%), Average expenditure on R&D expenditure of industrial enterprises above designated size (10,000 yuan/piece), Full-time equivalent of R&D personnel of industrial enterprises above designated size (person year), The proportion of science and technology expenditure to public budget expenditure | X11 Positive | |
| | Innovation output | Number of domestic patents granted (pieces), Number of new product development projects of industrial enterprises above designated size (items), Number of R&D projects of industrial enterprises above designated size (items) | X15 Positive, X16 Positive, X17 Positive | |
| | Innovation | Full labor productivity (yuan/person), High-tech industry development funds accounted for the proportion of sales revenue (%), The proportion of the total output value of high-tech industries to the total industrial output value (%) | X18 Positive, X19 Negative, X11 0 Positive Not available | |
| | Urban and rural regional coordinatio n | Ratio of per capita disposable income of urban and rural residents, Ratio of students in urban and rural junior high schools, Ratio of health technicians per 1,000 people in urban and rural areas, The ratio of the number of beds per 1,000 medical and health institutions in urban and rural areas | X21 Negative, X22 Negative, X23 Negative, X24 Negative | |
| | Material civilization and spiritual civilization coordinatio n | The proportion of rural cable audio users (%), Cultural sports and media spending (100 million yuan), Education expenditure (100 million yuan), Per capita possession of public library books (book) | X26 Positive, X27 Positive, X28 Positive, X29 Positive | |
| | Coordinate d development of economy and society | The proportion of defense expenditure to public finance expenditure (%), Social security and employment expenditure as a percentage of fiscal expenditure (%), Public security expenditure as a percentage of fiscal expenditure (%) | X21 0 Positive, X21 1 Positive, X21 2 Positive | |
| | ECO developme nt | 10,000 yuan of gross domestic product electricity consumption (kWh / yuan), The number of days in the provincial capital city reached and better than level 2 (days), Green coverage rate in built-up areas (%) | X31 Negative, X32 Positive, X33 Positive | |
| | Resources and Environment | Number of public transportation vehicles per 10,000 people (standard platform), Per capita water resources (cubic meters per person), Per capita cultivated area (hectare / person) | X34 Positive, X35 Positive, X36 Positive | |
| | Low carbon cycle development | Harmless treatment rate of domestic garbage (%), Sewage treatment rate (%) | X37 Positive, X38 Positive | |
| | Ecological Protection | Ecological construction and protection investment (ten thousand yuan), Industrial pollution control investment (ten thousand yuan), Total afforestation area (hectare) | X31 0 Positive, X31 1 Positive, X31 2 Positive | |
| | Agriculture, forestry and forestry expenditure (100 million yuan) | X31 3 Positive | |
2.2 Selection and standardization of sample data

This study is based on the 2016 China Statistical Yearbook, 2016 China Urban Statistics Yearbook, 2016 China Science and Technology Statistics Yearbook, 2016 China Foreign Trade Statistics Yearbook. According to the index system after the initial screening, 53 third-level evaluation indicators were selected in 6 provinces (Shanxi, Anhui, Jiangxi, Henan, Hubei, and Hunan) (see Table 2). In order to simplify the evaluation index system, R-type cluster analysis and variable coefficient analysis are further selected.

Table 2 After the initial results of the “Five Development Concepts”, the original data of the evaluation index system of the six provinces in the central region

| province | R&D expenditure intensity (%) = R&D expenditure as a percentage of GDP (%) | Technical market turnover (ten thousand yuan) | Ratio of per capita disposable income of urban and rural residents | Public security expenditure as a percentage of fiscal expenditure (%) | Per capita road area (m2) |
|----------|-------------------------------------------------|---------------------------------------------|-------------------------------------------------------------|-------------------------------------------------|--------------------------|
| Shanxi   | 1.04                                            | 512007.1358                                 | 2.732                                                       | 5.077                                           | 14.81                    |
| Anhui    | 1.96                                            | 1904669.26                                   | 2.489                                                       | 3.742                                           | 19.86                    |
| Jiangxi  | 1.04                                            | 148.372                                      | 2.379                                                       | 4.814                                           | 17.61                    |
| Henan    | 1.18                                            | 161.115                                      | 2.357                                                       | 4.814                                           | 14.01                    |
| Hubei    | 1.9                                             | 248.14                                       | 2.284                                                       | 4.740                                           | 12.77                    |
| Hunan    | 1.43                                            | 251.962                                      | 2.623                                                       |                                                  |                          |

Source: 2016 China Statistical Yearbook, 2016 China City Statistical Yearbook, 2016 China Science and Technology Statistical Yearbook, 2016 China Foreign Trade Statistics Yearbook
2.3 Standardization of sample data

In order to eliminate the impact of dimensions on evaluation and analysis, the original indicators are standardized. Since the larger the value of positive indicators, the better the development effectiveness, and the smaller the negative indicators, the better the development effectiveness. Therefore, according to the type of indicator (positive and negative), negative signs are taken on the basis of standardization. The specific standardized data are shown in the table below.

Table 3 Standardization data of the “Five Development Concepts” effectiveness evaluation index system

| Province | R&D expenditure intensity (%) = R&D expenditure as a percentage of GDP (%) | Technical market turnover (ten thousand yuan) | Ratio of per capita disposable income of urban and rural residents | Public security expenditure as a percentage of fiscal expenditure (%) | Per capita road area (m²) |
|----------|--------------------------------------------------------------------------------|---------------------------------------------|---------------------------------------------------------------|---------------------------------------------------------------|-------------------------|
| Shanxi   | -0.9238                                                                         | -0.53952                                    | -1.48003                                                     | 1.16723                                                      | -0.42142                |
| Anhui    | 1.28378                                                                         | -0.56075                                    | 0.70226                                                      | -0.23818                                                     | 0.73119                 |
| Jiangxi  | -0.9238                                                                         | -0.49246                                    | 0.57179                                                      | -0.23704                                                     | 0.66278                 |
| Henan    | -0.5879                                                                         | -0.56075                                    | 0.70226                                                      | -0.23818                                                     | 0.66278                 |
| Hubei    | 1.13981                                                                         | 2.00581                                     | 1.12418                                                      | 0.59708                                                      | 0.66278                 |
| Hunan    | 0.012                                                                          | -0.3538                                     | -0.84902                                                     | 0.43649                                                      | -1.21134                |

2.4 Screening and Evaluation Indices Based on R-type Cluster and Variable Coefficient Analysis

2.4.1. Clustering analysis is conducted on the five levels of innovation, coordination, green, open, and shared indicators, and the K-W non-parameter test is performed on the cluster structure. For classes that can not pass the test, further classification is carried out. The results of cluster analysis of evaluation indicators at various levels are as follows.

2.4.2. Calculate coefficient of variation. The variation coefficient of each index is calculated, and the indexes in the same category are screened and preserved according to the variation coefficient. The specific deletion principle is to retain the index with the largest variation coefficient for the same type of indicators, and delete other indicators to ensure that the selected indicators have a large impact on the evaluation results and avoid duplication of information on the same type of indicators.

Table 4 “Five Development Concepts” effectiveness evaluation index system based on R-type clustering and coefficient of variation analysis

| Innovation investment | Three-level indicator                                                                 | Clusterin g category | KW test sig value | Adjusted category | Coefficient of variation | Wheth er to keep |
|-----------------------|----------------------------------------------------------------------------------------|----------------------|-------------------|--------------------|--------------------------|------------------|
|                       | R&D expenditure intensity (%)                                                          | 1                    | 0.2924            | delete             |
|                       | Average expenditure on R&D expenditure of industrial enterprises above designated size (10,000 yuan/piece) | 3                    | 0.2539            | Reserve d          |
|                       | Full-time equivalent of R&D personnel of industrial enterprises above designated size (person year) | 2                    | 0.5183            | Reserve d          |
|                       | The proportion of science and technology expenditure to public budget expenditure (%)   | 1                    | 0.4311            | Reserve d          |
Table 4, cont.

| Innovation output | Number of domestic patents granted (pieces) | 1 | 0.4852 | delete |
|-------------------|------------------------------------------|---|--------|--------|
|                   | Number of new product development projects of industrial enterprises above designated size (items) | 1 | 0.6436 | delete |
|                   | Number of R&D projects of industrial enterprises above designated size (items) | 1 | 0.5610 | Reserve |
|                   | The added value of high-tech industries accounts for the proportion of industrial increase (%) | | | Increase |
| Innovation        | Full labor productivity (yuan/person) | 2 | 0.14 | 3 | 0.5610 | Reserve |
|                   | High-tech industry new product development funds accounted for the proportion of sales revenue (%) | 1 | | 1 | 0.4852 | delete |
|                   | Technical market turnover (ten thousand yuan) | 1 | | 2 | 0.6436 | Reserve |
| Urban and rural regional coordination | Ratio of per capita disposable income of urban and rural residents | 1 | | | 0.0694 | delete |
|                   | Ratio of health technicians per 1,000 people in urban and rural areas | 1 | | | 0.1397 | Reserve |
|                   | The ratio of the number of beds per 1,000 medical and health institutions in urban and rural areas | 1 | | | 0.1267 | delete |
|                   | Basic pension insurance coverage rate for urban and rural residents (%) | 2 | | | 0.0677 | Reserve |
|                   | Ratio of students in urban and rural junior high schools | 3 | | | 0.2780 | Reserve |
| Material civilization and spiritual civilization coordination | The proportion of rural cable audio users (%) | 1 | | 1 | 0.2975 | Reserve |
|                   | Per capita possession of public library books (book) | 1 | | 2 | 0.2418 | Reserve |
|                   | Cultural sports and media spending (100 million yuan) | 2 | | 3 | 0.1934 | Reserve |
|                   | Education expenditure (100 million yuan) | 2 | | 4 | 0.2450 | Reserve |
| Economic and social coordination | The proportion of defense expenditure to public finance expenditure (%) | 1 | | | 0.3878 | Reserve |
|                   | Social security and employment expenditure as a percentage of fiscal expenditure (%) | 2 | | | 0.0955 | Reserve |
|                   | Public security expenditure as a percentage of fiscal expenditure (%) | 3 | | | 0.1017 | Reserve |
| Resources and Environment | The number of days in the provincial capital city reached and better than level 2 (days) | 1 | | | 0.2631 | delete |
|                   | Green coverage rate in built-up areas (%) | 1 | | | 0.0611 | delete |
|                   | Per capita water resources (cubic meters per person) | 1 | | | 0.8607 | Reserve |
|                   | Number of public transportation vehicles per 10,000 people (standard platform) | 2 | | | 0.1844 | Reserve |
|                   | 10,000 yuan of gross domestic product electricity consumption (kWh / yuan) | 3 | | | 0.4039 | Reserve |
|                   | Per capita cultivated area (thousand hectares per 10,000 people) | 3 | | | 0.0328 | delete |
| Low carbon cycle development | Harmless treatment rate of domestic garbage (%) | 1 | | | 0.0328 | delete |
|                   | Sewage treatment rate (%) | 1 | | | 0.0503 | Reserve |
|                   | Sulfur dioxide and nitrogen oxide emissions per unit of GDP (ton/100 million yuan) | 2 | | | 0.6701 | Reserve |
After R-type cluster analysis and variation coefficient analysis, a system of "five major development concepts" effectiveness evaluation indicators with 39 indicators was established. Specifically, 5 indicators for innovative development, 10 indicators for coordinated development, 8 indicators for green development, 7 indicators for open development and 8 indicators for shared development are shown in the table above.

### 2.5 Determination of Reasonable Evaluation System after Screening

Calculate the trace of the 39 covariance matrices after screening and the trace of the 57 covariance matrices before screening, and bring it into $I_n = trS_r / trS_p$. That is $I_n = 89.31\%$, the
evaluation index system after screening with R-type cluster analysis and variable coefficient analysis reflects 89.31% of the original information. It can be considered that the evaluation index system constructed in this study is relatively reasonable.

2.6 Rational analysis of the addition of indicators and construction of the "five major development concepts" evaluation index system in Hunan Province

Based on the evaluation index system finally established through R-type clustering and variation coefficient analysis, combined with Hunan's own regional characteristics and Hunan's characteristics, from the perspective of rational analysis, the proportion of the total output value of high-tech industries to the total industrial output value(%) is increased. This indicator, In addition, the number of indicators of residents' satisfaction with the Party and the government that were deleted because they could not be measured (the issue of malfeasance and corruption reported by the Central Commission for Discipline Inspection) was increased, and the index values were obtained by checking the data published by the CPC Central Commission for Discipline Inspection. In this way, the evaluation index system of the "five major development concepts" in Hunan Province with 41 indicators is established as shown in table 4. The two additional indicators appear in bold form.

3. Conclusion

Based on the analysis of R-type clustering and variable coefficients, the evaluation index system of "five major development concepts" in Hunan Province was constructed. Drawing on domestic and foreign representative research results on the development of evaluation indicators systems and innovative, coordinated, green, open, and shared evaluation indicators, we conducted preliminary screening of evaluation indicators for sea elections based on the principle of observability. Then, the data of six central provinces were selected by combining R-cluster analysis and variational coefficient analysis to select 39 evaluation indicators, and the effectiveness evaluation index system of the "five major development concepts" suitable for the six central provinces was constructed. Then, according to the regional characteristics and characteristics of Hunan province, two indexes were added from the perspective of rational analysis, so as to construct a quantitative evaluation index system of the "five major development concepts" in Hunan province with 41 indexes.

Acknowledgments

This work was financially supported by the Hunan Philosophy and Social Science Fund Project (Project number: 17YBA319)

References

[1] Huyonghong, He Sihui. Comprehensive evaluation method[M]. Beijing: Science Press, 2000
[2] United Nations Development Programme(UNDP). Human Development Reports[R]. New York: Oxford University Press, 2002: 141-162
[3] United Nations Commission on Sustainable Development. Indicators of Sustainable Development Framework & Methodologies [M]. New York, 2001
[4] The United Nations development goals in one thous. 2000
[5] European Commission. Euro Stat. Towards Environmental Pressure Indicators for the EU[M]. EU, 1999.
[6] Cao Feng-zhong. American sustainable development indicators[J]. Environmental Science Trends, 1996, 2: 5-8. (in Chinese)
[7] Institute of Sustainable Development, Chinese Academy of Sciences. 1999 China Sustainable Development Strategy Report[M]. Beijing: Science Press, 1999.
[8] Statistical Office of the People's Republic of China, Comprehensive Index Development Index Preparation Programme, 2011