Statistical Evaluation of Green Economic Development Based on Rough Set Calculation Method

Minjing Peng¹,a, Jiajia Fan²,b

¹School of Economics & Management, Wuyi University WYU SEM Jiangmen, China
²School of Economics & Management, Wuyi University WYU SEM Jiangmen, China
a peng.mj@wyu.edu.cn
b Corresponding author: 69580413@qq.com

Abstract. In the context of big data, decisions and evaluations in business, economics and other fields increasingly rely on data analysis rather than experience and intuition. Data grows exponentially, but not all of it is useful. Rough set is a data mining method to deal with data classification, which can effectively remove redundant information. Green economy is an important proposition of the current world development. The evaluation of green economy should not only focus on economic data, but also environmental and social data. How to filter the data and establish a green economic index evaluation model is the problem this paper wants to solve. Based on the data of green economy and rough set algorithm, this paper evaluates the development level of green economy. Taking the data of Jiangmen city as an example, the evaluation model is tested and analyzed.

1. Introduction
The Fifth Plenary Session of the 18th CPC Central Committee pointed out that green is a necessary condition for sustainable development and an important embodiment of people's pursuit of a better life. With the development and progress of society, human beings gradually realize the importance of the environment. How to combine the environment and economy to seek better development and achieve a win-win situation is worth our in-depth thinking and research[1]. Green economy is an irreversible trend and an important proposition in line with sustainable development. However, there is no specific comprehensive evaluation index system for green economy in China.

Rough set theory is mainly used in machine learning and data mining, prediction and control, pattern recognition and classification, etc[2]. In recent years, rough set method has been gradually applied to multi index comprehensive evaluation, and has been used in many fields such as society, science and technology, finance. When dealing with massive fuzzy and uncertain incomplete data, the rough set method can be used to deal with problems efficiently. Without any prior information, rough set can effectively propose redundant information, compare the dependence and importance of different attributes, and derive classification rules[3]. Therefore, it is necessary to apply rough set to green economy evaluation.
2. Construction of Evaluation Index System

2.1. Construction Principles
The construction of green economy evaluation system follows five principles: scientific, systematic, hierarchical, dynamic and comparable. Taking Jiangmen City as an example, this paper collects the relevant index data of 11 years from 2009 to 2019. The data comes from the Statistical Yearbook of Jiangmen City of Jiangmen Bureau of statistics. Through the construction of a comprehensive evaluation model based on rough set calculation method, this paper evaluates the development of green economy in Jiangmen City.

2.2. Framework System
According to the above summary and combing of the meaning of green economy, and in accordance with the idea of analytic hierarchy process, and using the existing research results[4-6] for reference, the green economy evaluation index system is divided into three layers: the first layer is the target layer, that is, the green economy development index. The second layer is the system layer, including three subsystems: economy, environment and society. The third layer is the index layer, which has nine related indexes. Specifically, in the economic subsystem, including the total GDP of the city index, the GDP per capita index and the percentage of tertiary industry in total GDP index. In the environmental subsystem, including the per capita park green area index, the forest coverage index and the greenery coverage of urban area index. In the social subsystem, including the population growth rate index, the unemployment rate index and the per capita disposable income of residents index. Indicator model is shown in Table 1.

| Target Layer | System Layer | Index Layer | Index Number | Index Attribute |
|--------------|--------------|-------------|--------------|----------------|
| Green Economy Development | Economic System | Total GDP of the city (ten thousand Yuan) | A₁ | Positive |
| | | GDP per capita (yuan) | A₂ | Positive |
| | | Percentage of tertiary industry in total GDP (%) | A₃ | Positive |
| | Environmental System | Per capita park green area (square meters) | B₁ | Positive |
| | | Forest coverage (%) | B₂ | Positive |
| | | Greenery coverage of urban area (%) | B₃ | Positive |
| | Social System | Population growth rate (%) | C₁ | Negative |
| | | Unemployment rate (%) | C₂ | Negative |
| | | Per capita disposable income of residents (yuan) | C₃ | Positive |

Table 1  green economy development level evaluation indicator model

3. Rough Set Calculation Method

3.1. Standardize Data

3.1.1. Standardization of positive indicators
Positive indicators refer to the indicators that the greater the value, the better the level of green economy development.

\[ x_{ij}^* = \frac{x_{ij} - \min_{1 \leq j \leq 11} x_{ij}}{\max_{1 \leq j \leq 11} x_{ij} - \min_{1 \leq j \leq 11} x_{ij}} \]  \hspace{1cm} (1)

Where \( x_{ij}^* \) is the standardized data of the \( i \) index in the \( j \) year, and \( x_{ij} \) is the original value.

3.1.2. Standardization of negative indicators
Negative indicators refer to the indicators that the smaller the value, the better the level of green economy development.
\[ x_{ij}^* = \frac{\max x - x_{ij}}{\max x - \min x} \]

Where \( x_{ij}^* \) is the standardized data of the \( i \) index in the \( j \) year, and \( x_{ij} \) is the original value.

### 3.2. Discretize Data

After standardizing the index data, it is necessary to discretize the standardized data. The K-means clustering in SPSS was used to discretize the data into three levels. The higher the level, the better the development, and the attribute information table was obtained. Where \( X_n \) represents the data in the \( n \) year after processing, \( X_1 \) is 2019, \( X_2 \) is 2018, and so on. Take economic information system as an example, as shown in Table 2.

| \( X \) | \( A_1 \) | \( A_2 \) | \( A_3 \) |
|-------|-------|-------|-------|
| \( X_1 \) | 3 | 3 | 3 |
| \( X_2 \) | 3 | 3 | 3 |
| \( X_3 \) | 3 | 3 | 3 |
| \( X_4 \) | 2 | 2 | 3 |
| \( X_5 \) | 2 | 2 | 3 |
| \( X_6 \) | 2 | 2 | 2 |
| \( X_7 \) | 2 | 2 | 2 |
| \( X_8 \) | 1 | 1 | 2 |
| \( X_9 \) | 1 | 1 | 1 |
| \( X_{10} \) | 1 | 1 | 1 |
| \( X_{11} \) | 1 | 1 | 1 |

### 3.3. Indicator Attribute Reduction

According to the above steps, the index evaluation information system is constructed, \( S = (U,A,V,F) \). Where \( U \) represents the domain, and in this paper represents the index data of each year, \( U = \{x_1, x_2, x_3 \ldots , x_{11}\} \); All evaluation indicators are regarded as attribute set \( A \) in the information system, \( A = \{a_1, a_2, \ldots , a_{|A|}\} \); \( V \) is the range of each evaluation index; \( F \) is the information function of \( U \rightarrow V \), for \( \forall a \in A \), \( \forall x \in U \), \( f(x,a) \in V_a \) specifies the score value of \( x \) under the Evaluation Index \( a \) in each year.

The indicators selected in this paper can evaluate the development level of green economy in Jiangmen City to a certain extent, but there may still be redundant indicators in the indicators, which makes the evaluation not targeted and accurate. Therefore, it is necessary to reduce the attributes of the indicators through rough sets to make them more targeted. Here are the steps:

1) Calculate the lower approximation \( POS_{\text{ind}(A)}(U) \) of the set \( U \) of \( \text{ind}(A) \) under the equivalence relation:

2) Calculate the lower approximation \( POS_{\text{ind}(A - \{a_i\})}(U) \) under the equivalence relation \( \text{ind}(A - \{a_i\}) \) after eliminating an attribute \( a_i (a_i \in A) \) in attribute \( A \).

3) If \( POS_{\text{ind}(A)}(U) = POS_{\text{ind}(A - \{a_i\})}(U) \), \( a_i \) is an unnecessary attribute in attribute \( A \) that needs to be reduced. Otherwise, it is a necessary attribute, which forms the core \( core_{\text{ind}(A)}(U) \) of attribute set \( A \). A new information system \( S = (U,A',V,F) \) after reduction is established based on the obtained core, and \( A' \) is the attribute set formed by the necessary attributes in \( A \).

4) Taking the economic system as an example, set as \( S_1 = (U,A,V,F) \), the index reduction process is as follows:

\[ U = \{x_1, x_2, \ldots , x_{11}\}, \quad A = \{A_1,A_2,A_3\} \]
\[ U/\text{ind}(A) = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\}, \]
\[ \text{POS}_{\text{ind}(A)}(U) = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\}. \]

Remove an attribute \( A_i \) (\( A_i \in A \)) from attribute \( A \), then:
\[ U/\text{ind}(A - \{A_i\}) = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\}, \]
\[ \text{POS}_{\text{ind}(A - \{A_i\})}(U) = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\} \neq U \]
\[ U/\text{ind}(A - \{A_2\}) = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\}, \]
\[ \text{POS}_{\text{ind}(A - \{A_2\})}(U) = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\} \]
\[ \text{POS}_{\text{ind}(A - \{A_3\})}(U) = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\}. \]

In the same way, the information system of environmental and social indicators is calculated. Finally, \( A_2, B_2, B_3, C_1, C_2, C_3 \) are the necessary evaluation indicators.

3.4. Building a New Information System

It can be seen from the above that \( A_2, B_2, B_3, C_1, C_2, C_3 \) are the necessary attributes in the evaluation index, and other unnecessary attributes can be reduced. The simplified index system can be obtained, as shown in Table 3.

| System Layer  | Index Layer                        |
|---------------|-----------------------------------|
| Economic System | GDP per capita (yuan)             |
| Environmental System | Forest coverage rate (%)       |
| Social System   | Greenery coverage of urban area (%)|
|                | Population growth rate (%)       |
|                | Unemployment rate (%)             |
|                | Per capita disposable income of residents (yuan) |

According to formulas (3) and (4), the significance of each index to the corresponding attribute set is calculated. \( R \subseteq A \), \( |U| \) represents the number of sets, and approximate quality \( \gamma_{\text{ind}(R)}(U) \) represents the degree of dependence, the value range is \([0,1]\), the closer to 1 the higher the degree of dependence.

\[ \gamma_{\text{ind}(R)}(U) = \frac{|\text{pos}_{\text{ind}(R)}(U)|}{|U|} \quad (3) \]
\[ \text{sig}(a_j, A) = \gamma_{\text{ind}(A)}(A) - \gamma_{\text{ind}(A - \{a_j\})}(A) \quad (4) \]

The significance of each evaluation index is normalized to get the weight of each evaluation index, where \( w_j \) represents the weight of the \( j \) evaluation index, \( j = 1, 2, 3, \ldots, |A| \).

\[ w_j = \frac{\text{sig}(a_j, A)}{\sum_{j=1}^{|A|} \text{sig}(a_j, A)} \quad (5) \]

Taking the economic system as an example. Firstly, calculate the significance of necessary index \( A_2 \).

\[ \gamma_{\text{ind}(A - \{A_2\})}(U) = \frac{|\text{pos}_{\text{ind}(A - \{A_2\})}(U)|}{|U|} = \frac{3}{11} \]
\[ \text{sig}(A_2, A) = 1 - \gamma_{\text{ind}(A - \{A_2\})}(U) = \frac{8}{11} \]

Then calculate the weight according to the equation (5).
Other necessary evaluation indexes are calculated in the same way. The weights of the calculated index system are shown in Table 4.

Table 4  Weight of Green Economy Development Level Index System

| Evaluation Index | $A_2$ | $B_2$ | $B_3$ | $C_1$ | $C_2$ | $C_3$ |
|------------------|------|------|------|------|------|------|
| Weight           | 1.00 | 0.17 | 0.83 | 0.29 | 0.29 | 0.41 |

4. Comprehensive Evaluation

4.1. Calculate Model

Combined with the weight of the indicators, the evaluation results and comprehensive evaluation values of the development level of green economy in Jiangmen City from 2009 to 2019 are calculated respectively. Equation (6) was used to calculate the comprehensive evaluation score, where $V_{ij}$ is the score of year $i$ under the evaluation index $j$.

$$T = \sum_{i=1}^{[n']} V_{ij}w_j$$

The scores of green economic development indicators of Jiangmen City in each year are shown in Table 5.

Table 5  Evaluation Model of Green Economy Development of Jiangmen City from 2009 to 2019

| Years | Indexes                                      | 2019 | 2018 | 2017 | 2016 | 2015 | 2014 | 2013 | 2012 | 2011 | 2010 | 2009 |
|-------|---------------------------------------------|------|------|------|------|------|------|------|------|------|------|------|
| GDP per capita (yuan)         | 3.00 | 3.00 | 3.00 | 2.00 | 2.00 | 2.00 | 2.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Forest coverage rate (%)      | 0.34 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.34 | 0.34 | 0.17 | 0.17 | 0.17 |
| Greenery coverage of urban area (%) | 2.49 | 1.66 | 2.49 | 1.66 | 1.66 | 0.83 | 0.83 | 1.66 | 1.66 | 0.83 | 0.83 |
| Population growth rate (%)    | 0.87 | 0.29 | 0.29 | 0.29 | 0.29 | 0.58 | 0.58 | 0.58 | 0.58 | 0.58 | 0.58 |
| Unemployment rate (%)         | 0.58 | 0.29 | 0.29 | 0.29 | 0.29 | 0.58 | 0.58 | 0.87 | 0.58 | 0.58 | 0.58 |
| Per capita disposable income of residents (yuan) | 1.23 | 1.23 | 0.82 | 0.82 | 0.41 | 0.41 | 1.23 | 0.82 | 0.82 | 0.41 | 0.41 |
| Comprehensive evaluation score | 8.51 | 6.98 | 7.40 | 5.57 | 5.16 | 4.91 | 5.56 | 5.27 | 4.81 | 3.28 | 3.28 |

4.2. Discussion

As can be seen from the Table 5, only the per capita GDP index showed an upward trend in recent years, and other indicators showed certain fluctuations during the observation years. In terms of the environment, the forest coverage rate decreased for the first time in 2019, and the greenery coverage of urban area rate also fluctuated. As to protect the environment, we need to adhere to the call of "clear water and green mountains are as good as gold and silver mountains". In terms of social systems, the natural population growth rate has been on the rise from 2015 to 2018, reflecting the public's positive response to the two-child policy introduced in 2015. But the unemployment rate between 2015 and 2018 is at its highest level in recent years, and population growth may be putting more pressure on jobs. For
example, working women of child-bearing age are more likely to be discriminated against and face a higher risk of unemployment. The natural population growth rate dropped significantly in 2019, and the unemployment situation also eased. China opened the three-child policy in 2021, while encouraging childbirth, the employment environment should be relatively fair, and the government should improve relevant policies and welfare.

In general, the comprehensive score of Jiangmen's green economic development level showed ups and downs in 2014 and 2018. The reason is that the per capita disposable income of residents decreased in 2014 and the greenery coverage of urban area decreased in 2018. The development of green economy includes three aspects: economy, environment and society. Fluctuations in any aspect will affect the overall development of green economy.

5. Conclusions

Environmental protection and economic development are two important themes in today's world. As the core of these two themes, green economy has important theoretical and practical significance to study. Based on rough set calculation method, incomplete, inconsistent and imprecise information can be effectively analyzed and processed, and hidden knowledge can be found without changing the classification ability of information system, thus revealing the potential law of information.

This paper constructs a comprehensive evaluation model of green economy development based on rough set calculation method, which can provide important theoretical basis for the evaluation of green economy development in China. Green economy should not blindly pursue the growth of economic data, but also pay attention to the balanced development of environment and society. However, the indicators used in this paper are typical, so the amount of index data is small, and the practicability of the model needs to be verified with more indicators in the future.

Acknowledgment

This research was supported by The Natural Science Foundation of Guangdong Province (Grant No. 2015A030313642), and Science and Technology Planning Project of Guangdong Province (Grant No. 2014B040404072).

References

[1] Jia Xueying. (2020). Research on Evaluation of low carbon economy development level of China coal enterprises based on Rough Set (Masters thesis, Xian University of Science and Technology)
[2] Jian Zhenfu, & Xu Jiayu. (2016). Big data analysis and data mining. Tsinghua University Press.
[3] PAWLAK.Z. Rough sets and intelligent data analysis[J]. Information Sciences, 2002, 147(1):1-12.
[4] Zheng Hongxia, Wang Yi & Huang Baorong. (2013). Review of green development evaluation index system. Industrial technology economy (02), 142-152
[5] Zhu Hailing. (2017). Construction of green economy evaluation index system. Statistics and decision (05), 27-30
[6] Liu Qian & Guo Xiaomei. (2015). Construction of green economy development evaluation index system. Enterprise Guide (17), 1-2