

# A Study of Regional Technological Innovation on the Basis of Grey Clustering Theory

Guo Ju<sup>1,2</sup>, Chen WeiXu<sup>3</sup>

<sup>1</sup>Department of Economics & Management, Fuzhou University, Fujian, 310001, China

<sup>2</sup> Fujian Polytechnic of Information Technology, Fujian, 310001, China

<sup>3</sup>Department of Marxism, Fuzhou University, Fujian, 310001, China

**Keywords:** grey clustering; technological innovation; appraise system.

**Abstract.**Regional technological innovation presents both the general development of regional economy and the sustainable development of the region. Establishing an appropriate appraise system can help us better observe and study the innovation system, thus find new regular patterns of innovative development and quicken the pace of human's technological innovation. The superiority of the grey clustering theory lies in its abilities of overcoming the shortcomings of probability statistics and obtaining useful information from limited and disordered data. Therefore, the use of grey clustering theory in the study of the appraise system of regional technological innovation contributes to solving problems which are intra-class similar and inter-class different. Meanwhile, the calculation of the theory is relatively easy, and the information from which we obtained is more integrated so that conclusions are more in line with the objective law.

## Introduction

Due to the historical reasons, China's economic development is uneven and the level of which differs greatly from each region. The disparities also reflect technological innovative capabilities of different regions. We must objectively evaluate the technological innovative capabilities of different regions; fully understand both the advantages and disadvantages of the technological innovative capabilities in different fields of each region. Then gradually reduce the imbalance in economy while drawing upon each other's strong points. The present study shows that there still exists divergence in many problems on the study of regional innovative evaluation system, which causes great controversy. For instance, in the selection of the index (How to select appropriate index); in quantifying the qualitative index (How to scale quantitative indicators); in the comparison between the importance of indexes (How to assign weights); how to classify (Scale the similarities of a certain category or different categories), etc. The posing and the solution of these problems all contribute to make policies better confirm to the reality, as well as better evaluate technological innovative capabilities of different regions. Good economic and social benefits are obtained through the activities of technological innovation.

## Related Literature Review

The evaluation of technological innovation has witnessed great diversity in the methods used by academics. The most frequent used are clustering methodology, multiple factor analysis method and PCA (principal component analysis), etc. For example, Chen Yunwei(2009) has evaluated the technological innovations of eight economic regions of China[1]; Wu Xianying (2003) analyzed and researched the regional technological innovative capability of 30 provinces in China by using multiple factor analysis method, and reached a conclusion: regional disparity in technological innovative capabilities exists and that of some regions still needs to be improved. His suggestion was that the technological innovative activities should be undertaken on the basis of the region's own situation. Governments should take an active part in them, make appropriate plans for their own regions and boost organic developments in technological innovative system. Shao Yunfei (2003) drew a similar conclusion of Wu Xianying by using clustering methodology, and gained

valuable conclusion by the use of PCA [2][3]. But because of the lack of information, single index can not well reflect regional technological innovative capabilities. Tang Yanzhao (2001) used the grey comprehensive evaluation method, Liu Guoxin and Luo Xianfeng successively used BP artificial neural network, optimality of every level and entropy evaluation method to analyze the technological innovative capabilities of every province, but their conclusions are different, which affects the general understanding[4]. So far, the method of evaluation is still under researched. Due to the complexity of the evaluation system and huge quantities of the related contents, there still exist problems sort of weight matching of ration and qualitative index, and determine of magnitude, etc.

## The Theoretical Basis

Grey clustering theory is a theory proposed by Professor Deng JuLong in 1982, the grey clustering assessment is an important branch of the grey system theory, it is mainly used for the assessment of circumstance under lacking of information (or uncertain information index system) and a small sample of events.

We suppose there are  $n$  observation object, each observed object has  $m$  features of the data, using the matrix expression for:

$$A = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & \mathcal{E}_{2m} \\ \cdots & \cdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

We describe a grey number for the range of different numerical preference by using the definite weighted functions.

Definition 1: The ascending on left and descending on right continuous function called typical definite weighted functions  $f_j^k(\bullet)$ . If  $\lambda_j^k = 1/2[f_j^k(2) + f_j^k(3)]$ , then  $\lambda_j^k$  is called a critical value of the subclass  $k$  of  $j$  index.

Definition 2: If the definite weighted functions  $f_j^k(\bullet)$  has not the first and the second turning point  $x_j^k(1), x_j^k(2)$ , then it is called the lower measure definite weighted functions  $f_j^k[-, -, x_j^k(3), x_j^k(4)]$ . If  $\lambda_j^k = x_j^k(3)$ , then  $\lambda_j^k$  is called a critical value of the subclass  $k$  of  $j$  index.

Definition 3: If the second turning point  $x_j^k(2)$  and the third turning point  $x_j^k(3)$  of the definite weighted functions  $f_j^k(\bullet)$  Coincidence, then it is called the moderate measure definite weighted functions  $f_j^k[x_j^k(1), x_j^k(2), -, x_j^k(4)]$ . If  $\lambda_j^k = x_j^k(2)$ , then  $\lambda_j^k$  is called a critical value of the subclass  $k$  of  $j$  index.

Definition 4: If the definite weighted functions  $f_j^k(\bullet)$  has not the third and the fourth turning point  $x_j^k(3), x_j^k(4)$ , then it is called the upper measure definite weighted functions  $f_j^k[x_j^k(1), x_j^k(2), -, -]$ . If  $\lambda_j^k = x_j^k(2)$ , then  $\lambda_j^k$  is called a critical value of the subclass  $k$  of  $j$  index.

Definition 5: If  $\lambda_j^k$  is called a critical value of the subclass  $k$  of  $j$  index, then the  $\eta_j^k = \lambda_j^k / \sum_{j=1}^m \lambda_j^k$  is the weight of the subclass  $k$ .

Definition 6: If the  $\eta_j^k$  is called the weight of the subclass  $k$ , then the  $\sigma_i^k = \sum_{j=1}^m f_j^k(x_{ij}) \eta_j^k$  is the object  $i$  belongs to the  $k$  grey clustering.

Define 7: We suppose  $x_{i,j}$  ( $i=1,2,\dots,n; j=1,2,\dots,m$ ) is the samples value on the object  $i$  of the index  $j$ , then  $f_j^k(\bullet)$  ( $j=1,2,\dots,m; k=1,2,\dots,s$ ) is called the definite weighted functions on the subclass  $k$  of  $j$  index. If the weight  $\eta_j^k(\bullet)$  ( $j=1,2,\dots,m; k=1,2,\dots,s$ ) of the subclass  $k$  of  $j$  index has nothing to do with  $k$ . To any  $k_1, k_2 \in (j=1,2,\dots,m)$ , there is  $\eta_j^{k_1} = \eta_j^{k_2}$ , it is called  $\eta_j$ .

Define 8: We suppose  $\eta_j$  is the weight of the subclass  $k$  of  $j$  index, then  $\sigma_j$  is the grey fixed weight clustering coefficient, to clustering coefficient matrix:

$$\sum(\sigma_i^j) = \begin{bmatrix} \sigma_1^1 & \sigma_1^2 & \dots & \sigma_1^s \\ \sigma_2^1 & \sigma_2^2 & \dots & \sigma_2^s \\ \dots & \dots & \dots & \dots \\ \sigma_n^1 & \sigma_n^1 & \dots & \sigma_n^s \end{bmatrix}$$

Define9: We suppose  $\sigma_i^{k*} = \max_{1 \leq k \leq s} \{\sigma_i^k\}$ , then Object  $i$  belongs to the gray clustering  $k^*$ .

### Realization of evaluation index based on grey system theory

We consider 31 regional data of China are too large, so the first six areas are analyzed, and then we can extended to all provinces. We formulate the following formto to analyze the technology innovation ability, the quantitative part of the data come from the National Bureau of Statistics (2011), the qualitative part come from ten experts in related fields we can visit.

. Table1:The indexes of different areas of regional technological innovation of basic data (1)

	Regional innovation input				Regional innovation output			
	PA	RDIN	RDP	RDI	PG	TMCA	SIE	HTNP
Jiangxi	348381	8998944	287447	31933	199814	333.43	27946	137552098
Sandong	128614	7431254	180832	25193	75496	140.02	13491	109198091
Sichuan	66316	1044666	36839	6712	42218	111.24	12119	20640382
Shanxi	16786	895891	32476	2348	7196	30.61	2709	8907109
Heilongjiang	30610	838042	39661	4343	20268	100.45	11141	5719207
Ningxia	1985	118879	3967	853	844	2.91	184	1531043

Note: PA:Patent apply, RDIN:R&D input, RDP: R&D personnel FTE, RDIT: R&D, PG:Patent grant, TMCA: Technology market clinch a deal amount, SIE: SCI\ISTP\EI, HTNP: High technology industry new product output, ES: Education spending, TGD: Transparency of the government decision-making, MSD: Market and safety degree, SDA: Sustainable development ability.

Table2:The indexes of different areas of regional technological innovation of basic data (2)

	Regional innovation environment			
	ES	TGD	MSD	SDA
Jiangxi	15882132	9	9	9
Sandong	13727939	9	8	8
Sichuan	10244130	8	7	8
Shanxi	5494903	7	7	7
Heilongjiang	4838173	7	7	8
Ningxia	1313862	7	6	7

We will according to the various index of the normal distribution of regional technology innovation ability is divided into four levels, and we determine the index of the standard value:

Table3:The index standard limit value of the regional technology innovation system(1).

	Regional innovation input				Regional innovation output			
	PA ≥	RDIN ≥	RDP ≥	RDI ≥	PG ≥	TMCA ≥	SIE ≥	HTNP ≥
Rank I	70000	2000000	45000	10000	40000	150	14000	35000000
Rank II	35000	1000000	35000	6000	20000	60	10000	20000000
Rank III	15000	800000	27000	4000	15000	30	5000	8000000
Rank IV	0	0	0	0	0	0	0	0

Table4:The index standard limit value of the regional technology innovation system(2).

	Regional innovation environment			
	ES ≥	TGD ≥	MSD ≥	SDA ≥
Rank I	7000000	9	9	9
Rank II	6000000	8	8	8
Rank III	5000000	7	7	7
Rank IV	0	0	0	0

We will standard limit value as a threshold of grey clustering definite weighted functions, and make to analyze the definite weighted functions by the target "patent applications", according to:

1. Bigger the elements of grey clustering are better, the upper measure definite weighted functions is adopted.
2. Smaller the elements of grey clustering are better, the lower measure definite weighted functions is adopted.
3. If grey clustering elements around an area, the moderate measure definite weighted functions is adopted.
4. In this case all indicators is the bigger the better, so grade I adopts the lower measure definite weighted functions, grade II adopts the moderate measure definite weighted functions, grade III adopts the moderate measure definite weighted functions, grade IV adopts the upper measure definite weighted functions.

$$f_1^I(x) = \begin{cases} 1 & x < 15000 \\ \frac{35000 - x}{35000 - 15000} & 15000 \leq x < 35000 \\ 0 & x \geq 35000 \end{cases}$$

$$f_1^{II}(x) = \begin{cases} 1 & x < 15000 \\ \frac{x - 15000}{35000 - 15000} & 15000 \leq x < 35000 \\ 0 & 35000 \leq x < 70000 \end{cases}$$

$$f_1^{III}(x) = \begin{cases} 1 & x < 35000 \\ \frac{x - 35000 - x}{70000 - 35000} & 35000 \leq x < 70000 \\ 0 & 70000 \leq x < 100000 \end{cases}$$

$$f_1^{IV}(x) = \begin{cases} 1 & x < 70000 \\ \frac{x - 70000}{100000 - 70000} & 70000 \leq x < 100000 \\ 0 & x \geq 100000 \end{cases}$$

According to the above algorithm in the data of jiangsu province:  $f_1^I = 1, f_1^{II} = 0, f_1^{III} = 0, f_1^{IV} = 0$ . We can get the definite weighted functions of different indicators of different regions by repeating to the above algorithm. To reference GuoJu (2012) description of the regional technology innovation index weights, we can define the clustering weights:

Table5: The grey clustering weights(1)

rank	Regional innovation input				Regional innovation output			
	PA	RDIN	RDP	TDI	PG	TMCA	SIE	HTNP
I II III IV	0.36	0.16	0.06	0.06	0.02	0.05	0.02	0.02

Table6: The grey clustering weights(2)

rank	Regional innovation environment			
	ES	TGD	MSD	SDA
I II III IV	0.04	0.04	0.04	0.06

Base on  $\sigma_i^k = \sum_{j=1}^m f_j^k(x_{i,j}) \eta_j$  to construct the grey clustering matrix:

$$\sigma_c = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2k} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nk} \end{bmatrix} = \begin{bmatrix} 1^* & 0 & 0 & 0 \\ 0.7177^* & 0.0911 & 0.1887 & 0 \\ 0.0411 & 0.4097 & 0.5686^* & 0.04 \\ 0 & 0 & 0.1722 & 0.8277^* \\ 0 & 0.0564 & 0.6544^* & 0.4086 \\ 0 & 0 & 0 & 1^* \end{bmatrix}$$

In the matrix, if  $\sigma_{ik}^* = \max_{1 \leq k \leq 4} \{\sigma_{ik}\}$ , then the clustering objects  $i$  belongs to the grey clustering  $k^*$ . The line that the number marked\* (i.e., a rank) is the grey clustering that the clustering object belong to. In the matrix  $\sigma_c$  we can Find out all the various maximum clustering coefficient value (the numer with \*), their clustering is the object of clustering.

From the table5 and table6 we know that Shandong province and Jianguo province belongs to the first clustering, Sichuan province and Heilongjiang province belongs to the third clustering, Shanxi Province and Ningxia belongs to the fourth clustering.

**Summary**

We can use this method to evaluated regional technical innovation capability in 31 provinces, municipalities and autonomous regions in 2011, Overall, by comparing the different regions of the country innovation capability, we can compare different regions index to find the gaps. We can analyze the reasons for the difference by analyzing the standard limit in order to provides the basis for the development of regional innovation policies. By comparing this method with other methods, we can find this article accurate indicators (Only depict regional technical innovation capability by 12 index value, which cover the investment in technological innovation, production and environment), most indicators are generally recognized scholars, which enhances the conclusions convincing. Algorithm design is simple but relatively rigorous in theory and fact-based, evaluation of regional technological innovation capability. Model calculations can be obtained MATLAB or R software programming, saving time and reducing the statistical error, the result of authenticity has been verified. By comparison to other methods, the results similarity great, and are associated with the country's economic development, which are in line with the current development law macroeconomic environment, contribute to the implementation of government policies. Gray clustering method can effectively solve the problem of uncertainty exists in the evaluation process of innovation obtain more rigorous results through limited information. On the whole, research on regional innovation issues, gray clustering method helps to analyze the innovation process, and the effect is satisfactory.

## References

- [1] Chen Yunwei, Yang Zhiping, Fang Shu, Hu Zhengyin, Martin Meyer, Sujit Bhattacharya. A patent based evaluation of technological innovation capability in eight economic regions in PR China.[J]. World Patent Information,2009,(31): 104-110
- [2] Shao Yun-fei, Tang Xiao-wo, Chen Guang. Cluster Empirical Analyses on the Capacity of Region Technical Innovation of China [J]. China Soft Science , 2003(5):113-118
- [3] Shao Yun-fei, Tang Xiao-wo. Principal-Components Empirical Study on the Capacity of Region Technical Innovation of China[J]. Journal of Industrial Engineering and Engineering Management, 2005,19(3):71-76
- [4] Liu Guo-xin, Feng De-xiong, Yao Han-jun, Luo Xian-feng. A Synthetic Evaluation of Regional Innovation & Enterprise Capability[J]. Journal of Wuhan Automotive Polytechnic University, 2003,25(1):84 - 88