Open Journal of Statistics, 2020, 10, 431-452
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Spatial Statistical Analysis and Comprehensive Evaluation of High-Tech Industry Development

Luyao Wang, Binhui Wang*

Department of Statistics, Jinan University, Guangzhou, China
Email: wly@stu2018.jnu.edu.cn, *rstat@126.com

How to cite this paper: Wang, L.Y. and Wang, B.H. (2020) Spatial Statistical Analysis and Comprehensive Evaluation of High-Tech Industry Development. *Open Journal of Statistics*, **10**, 431-452.
https://doi.org/10.4236/ojs.2020.103028

Received: April 29, 2020
Accepted: May 31, 2020
Published: June 3, 2020

Abstract

After 30 years of economic development, the high-tech industry has played an important role in China's national economy. The development of high-level technological industry plays a leading role in guiding the transformation of China's economy from "investment-driven" to "technology-driven". The high-tech industry represents the future industrial development direction and plays a positive role in promoting the transformation of traditional industries. The rapid development of high-tech industry is the key to social progress. In this paper, the traditional analytical model of statistics is combined with principal component analysis and spatial analysis, and R language is used to express the analytical results intuitively on the map. Finally, a comprehensive evaluation is established.

Keywords

Principal Component Analysis, Spatial Statistics, R Language, Comprehensive Evaluation

1. Introduction

1.1. Research Purpose

High technology industries are playing an increasingly important role in the optimization of regional industrial structure. Spatial statistics is an analytical method to analyze statistical information through spatial position. Compared with traditional statistical methods, spatial statistical analysis has its own advantages in understanding the geographical location characteristics and spatial pattern by considering the spatial position and neighborhood relationship according to correlation analysis. This paper combines the principal component analysis method of statistics and spatial analysis, uses R language [1] to make the analysis results intuitively expressed on the spatial map based on the research results.

1.2. Research Status

Spatial statistical analysis was originally proposed by South African geologist Krige, many foreign scholars carried out extensive and in-depth research on it. From 1970s to 1990s, Cliff and Griffith in the analysis [2] not only conducted many meaningful explorations on the verification of spatial autocorrelation, but also establish spatial model. Anselin and Getis in the analysis [3] and [4] conducted research on Local Spatial Statistical Indicators, including ESDA (Exploratory Spatial Data Analysis), LISA (Local Indicators of Spatial Analysis), Moran scatter map, etc. J.R. Friedman in the analysis [5] based on the theory of unbalanced regional economic development, divided the spatial structure of the economic layout into two systems, center and periphery, emphasizing that innovation elements gradually spread to the periphery while strengthening the development capacity and vitality of the "center". Since the 1990s, regional economic research has entered the stage of new spatial economics, in which new economic geography theories such as P.R. Krugman [6], Masahisa Fujita [7], and A. J. Venables [8] have become mainstream. In 2000, Okabe [9] *et al.* discussed spatial interpolation, spatial model and spatial point pattern analysis from the perspective of spatial mosaic, and Lloyd [10] *et al.* systematically summarized the statistical analysis of local models and spatial point patterns. Wu Yuming in the analysis [11] and [12] uses the spatial constant-coefficient spatial autoregressive model, spatial error model, and geographic weighted regression model of spatial statistical analysis to conduct a quantitative analysis of the overall R & D and innovation in China's provinces. Chen Hongchuan in the analysis [13] built a high-tech industry technology innovation capability evaluation index system, used data mining method (K-means clustering) to evaluate the high-tech industry technology innovation capacity. Zheng Shuwang and Xu Zhenlei in the analysis [14] used partial least squares regression model to analyze the influencing factors of the high-tech industries' technological innovation capability in the three northeastern provinces, and established a comprehensive evaluation index system for technical innovation capability. Fan Decheng and Du Mingyue in the analysis [15] used the TOPSIS grey relation projection method to calculate the close degree of grey relation projection, introduced coordination degree model to evaluate regional coordination development level and conducted quadratic weighted calculation. Wu Yanxia and Zhou Chunguang in the analysis [16] used spatial panel Dubin model to test the relationship between financial agglomeration and its spatial overflow and the development of high-tech industries. The results show that the Guangdong-Hong Kong-Macao Greater Bay Area financial agglomeration has the ability to promote the development of high-tech industries.

1.3. Innovation of This Paper

1) In this paper, we conduct comprehensive evaluation of the high tech industry development under principal component analysis (PCA) combined with spa-

tial analysis to find the scientific way to develop high-tech industry.

2) The research results of this paper will be applied in the management and decision-making of high-tech industry in China. Through the research on industrial structure and regional distribution, some feasible countermeasures and suggestions are put forward.

3) This paper introduces the spatial correlation effect analysis for the first time in the domestic industrial statistical analysis research by using the theoretical method of spatial statistical analysis and the integrated statistical model. We systematically and multi-dimensionally conduct empirical test and analysis on the overall research of China's high-tech industry.

2. Construction of Evaluation Index System

2.1. Construction of Evaluation Index System and Data Collection

To conduct comprehensive analysis, it is necessary to construct the evaluation index system. Generally speaking, the principles to be followed in constructing the comprehensive evaluation index system are as follows:

1) Principle of system comprehensiveness.

2) Principle of stable comparability.

3) Principle of scientific simplicity.

4) Principle of flexibility and operability.

In order to fully reflect the situation of national high-tech industry and understand the impact of regional differences on industrial development, it is necessary to conduct classification research according to the similarity of the development situation of high-tech industry in various regions. On the basis of collecting and collating the relevant data released by the country, we selected 10 evaluation indexes in 4 categories, including production and operation conditions as shown in **Table 1**.

2.2. Basic Analysis and Spatial Display of Data

According to the four types of statistical indicators, we establish the astrological map of the development data of each province in 2019.

2.2.1. Comprehensive Statistics

For the 2019 high-tech industry data, we categorize and summarize them according to various indicators, and calculate their minimum, maximum, average, 1/4 quantile, median, and 3/4 quantile, respectively. From the statistical results, there is a huge regional gap in the development of high-tech industries in various provinces of China. For example, the minimum number of high-tech enterprises is 10 and the maximum number is 7583; the profit of high-tech industry development is at least 4.5 billion yuan, and the largest province reached 2969.9 billion yuan as shown in **Table 2** and **Table 3**.

2.2.2. Spatial Statistics of Univariate Data

We then visually displayed the 10 development indicators in 2019 by spatial

Table 1. Index system for comprehensive analysis and evaluation of the development of high-tech industries.

	First-level indicators	Second-level indicators
Index system for comprehensive analysis and evaluation of the development of high-tech industries	A) Production and operation status	A1) Number of high-tech enterprises (x_1)
		A2) Annual average employees (x_2)
		A3) Main business income (x_3)
		A4) Total profit (x_4)
		A5) Export delivery value (x_5)
	B) R & D activities	B1) Number of R & D institutions (x_6)
		B2) Number of R & D institutions (x_7)
		B3) Expenditure on R & D institutions (x_8)
		B4) R & D personnel equivalent to full-time equivalent (x_9)
		B5) R & D internal expenditure (x_{10})
	C) New product development and patents	C1) New product development expenditure (x_{11})
		C2) New product sales (x_{12})
		C3) Technical transformation expenditure (x_{13})
		C4) Number of patent applications (x_{14})
		C5) Number of valid invention patent (x_{15})
	D) Hi-tech industry Fixed asset investment	D1) Number of construction projects (x_{16})
		D2) Completed and put into operation projects (x_{17})
		D3) Project investment amount (x_{18})
		D4) New fixed assets (x_{19})

Table 2. Data summary of high-tech enterprises.

High-tech enterprises	Annual average employees	Main business income	Total profit
Min: 10	Min: 1244	Min: 11	Min: 4.5
1st Qu.: 154	1st Qu.: 55,590	1st Qu.: 475	1st Qu.: 45.3
Median: 533	Median: 222,311	Median: 2395	Median: 210.8
Mean: 1033	Mean: 439,297	Mean: 5026	Mean: 365.4
3rd Qu: 1064	3rd Qu: 394,540	3rd Qu: 5102	3rd Qu: 331.7
Max: 7583	Max: 3,894,182	Max: 37,778	Max: 2969.9

Table 3. Data summary of high-tech enterprises.

Technical transformation funds	Number of patent applications	Number of effective invention patents	Number of construction projects
Min: 1	Min: 1	Min: 20	Min: 17
1st Qu.: 6113	1st Qu.: 160	1st Qu.: 422	1st Qu.: 176
Median: 28,541	Median: 1567	Median: 1787	Median: 566
Mean: 133,164	Mean: 4248	Mean: 8331	Mean: 798
3rd Qu: 133,715	3rd Qu: 4252	3rd Qu: 5348	3rd Qu: 926
Max: 1,001,652	Max: 51,427	Max: 152,519	Max: 4804

map, and divided them into four categories according to statistical classification: high value area, comparably high value area, comparably low value area and low value area.

The classification of four regions was based on the traditional quartile method.

1) Spatial maps of high-tech enterprises

It can be seen from Figure 1 that Guangdong, Jiangsu, Shandong, Zhejiang, Shanghai, Hunan, Sichuan and Henan are the high-value areas for the number of high-tech enterprises. Hubei, Jiangxi, Fujian and Liaoning are the comparably high-value areas. Xinjiang, Xizang, Qinghai and Ningxia are the comparably low-value areas for the number of high-tech enterprises. From the perspective of regional distribution, the coastal areas are basically high value areas and the western inland are low value areas.

2) Spatial map of annual average employees

As can be seen from Figure 2, Guangdong, Jiangsu and Shanghai are high-value areas for employees of high-tech industries, Shandong, Henan, Sichuan, Hubei, Hunan, Jiangxi and Liaoning are the comparably high-value areas. Xinjiang, Xizang, Qinghai and Ningxia are the low-value areas. In terms of regional distribution, except for Sichuan, the high-value areas are basically distributed along the eastern coastal areas, and gradually transit to the western inland into low-value areas.

3) Spatial map of main business income

It can be seen from Figure 3 map displayed, main business income of Guangdong, Jiangsu is of high value area, Zhejiang, Shanghai, Shandong, Sichuan, Beijing, Fujian, Hubei, Hunan and Henan are of comparably high value area. Anhui, Hebei, Shanxi, Guangxi are in relatively low-value areas. Xinjiang,

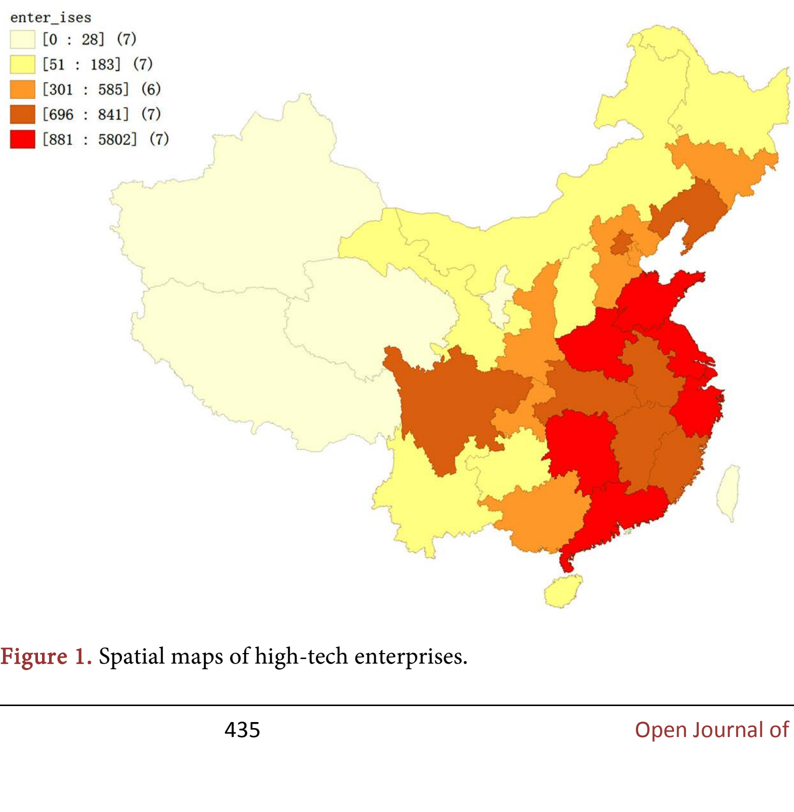


Figure 1. Spatial maps of high-tech enterprises.

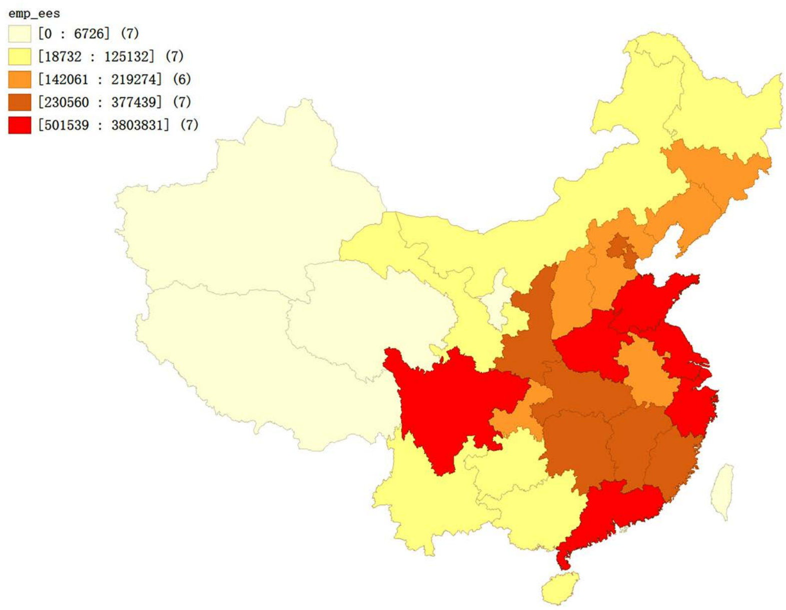


Figure 2. Spatial maps of annual average employees.

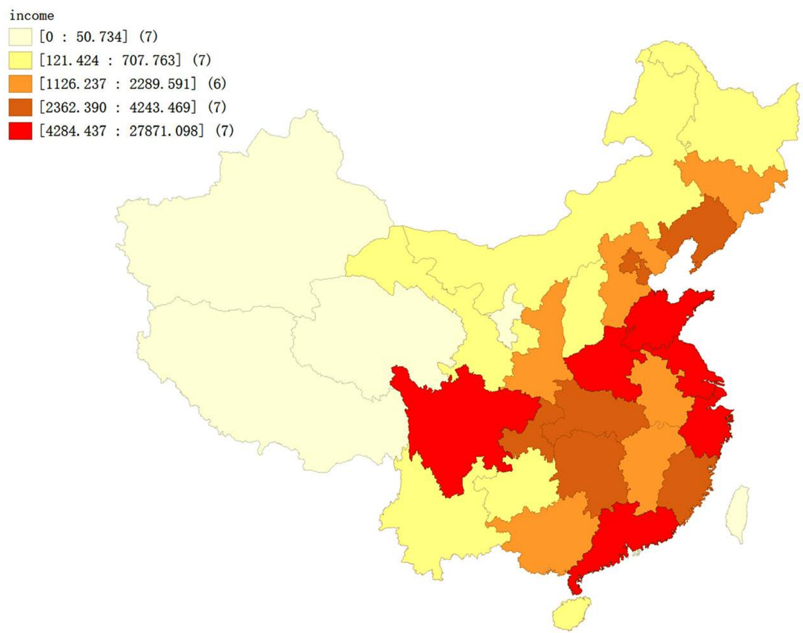


Figure 3. Spatial map of main business income.

Tibet, Qinghai, Ningxia are in low-value areas. In terms of regional distribution, except for Sichuan, the high-value areas are basically distributed along the eastern coastal areas, and gradually transit to the western inland into low-value areas.

4) Spatial map of total profit

It can be seen from **Figure 4** below, high technology industry of Guangdong, Jiangsu have high value of total profits. Zhejiang, Shanghai, Shandong, Sichuan, Beijing, Hubei, Hunan, Henan, Fujian and Anhui are in comparably

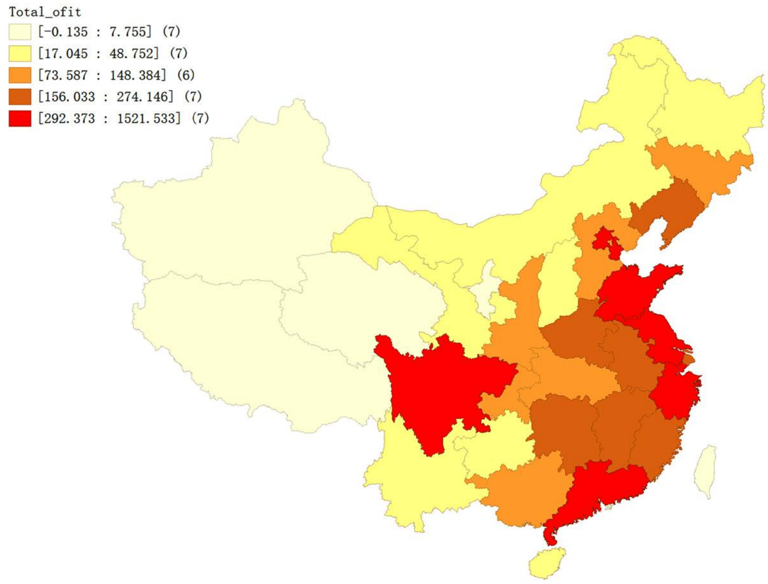


Figure 4. Spatial map of total profit.

high-value areas. Hebei, Shanxi, Guangxi, Guizhou are in relatively low-value areas. Xinjiang, Tibet, Qinghai, Ningxia have the lowest total profit of high-tech industry. In terms of regional distribution, except for Sichuan, the high-value areas are basically distributed along the eastern coastal areas, and gradually transit to the western inland into low-value areas.

5) Spatial map of export delivery value

It can be seen from **Figure 5** map below, Guangdong, Jiangsu are of high value area of the high-tech industry export delivery value. Shanghai, Shandong, Sichuan, Henan, Fujian, Hubei, Hunan, Zhejiang, Shanxi, Liaoning, Beijing are of comparably high value area. Heilongjiang, Jilin, Hebei, Jiangxi, Guangxi, Guizhou and are of relatively low-value, Xinjiang, Tibet, Qinghai, Ningxia, Yunnan are low in export delivery value. From the perspective of regional distribution, the high value region is basically distributed in the eastern part of China, and gradually transfers to the low value region in the western inland region.

6) Spatial maps of R & D institutions

The spatial map (**Figure 6**) shows that Jiangsu is number one for high-tech industry research and development institutions of high value area. Guangdong, Fujian, Zhejiang, Shanghai, Shandong, Beijing are of relatively high value area for high-tech industry R & D institutions. Hebei, Jilin, Heilongjiang, Guangxi, Guizhou are relatively low for the number of R & D institutions. Xinjiang, Tibet, Qinghai, Ningxia, Inner Mongolia are low-value areas. From the perspective of regional distribution, the high-value areas are basically distributed in the eastern part of China, and gradually transit to the low-value areas in the western inland.

7) Spatial map of new product development costs

It can be seen from spatial map (**Figure 7**) above that Guangdong ranks number one in new products for high technology industry development costs. Zhejiang,

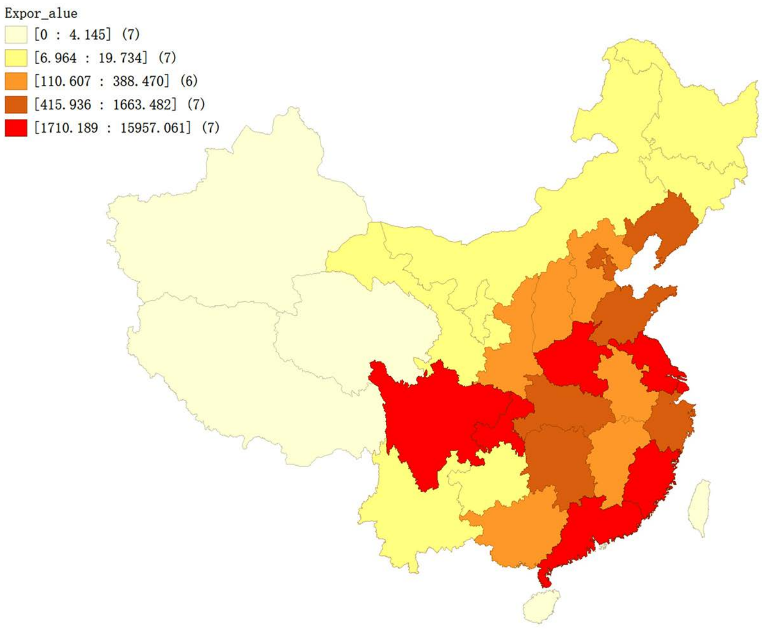


Figure 5. Spatial map of export delivery value.

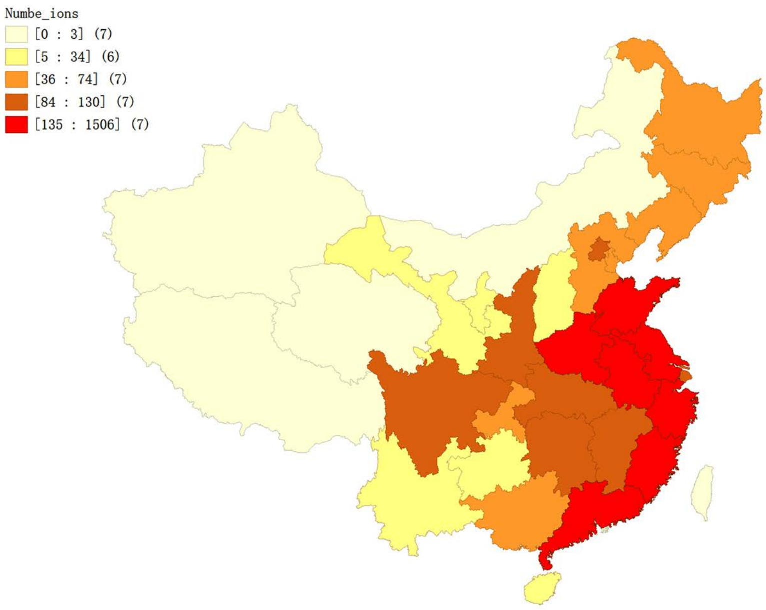


Figure 6. Spatial maps of R & D institutions.

Jiangsu, Shandong, Beijing, Shanghai, Hubei, Sichuan, Hunan, Anhui, Shanxi, Liaoning, Heilongjiang, Tianjin, Fujian are of comparably high value area for new product development costs. Jilin, Henan, Guizhou, Guangxi, Jiangxi are relatively low, while Xinjiang, Tibet, Qinghai, Ningxia, Inner Mongolia, Gansu and Yunnan provinces are low. From the perspective of regional distribution, the high-value areas are basically distributed in the eastern part of China, and gradually transit to the low-value areas in the western inland.

8) Spatial map of new product sales

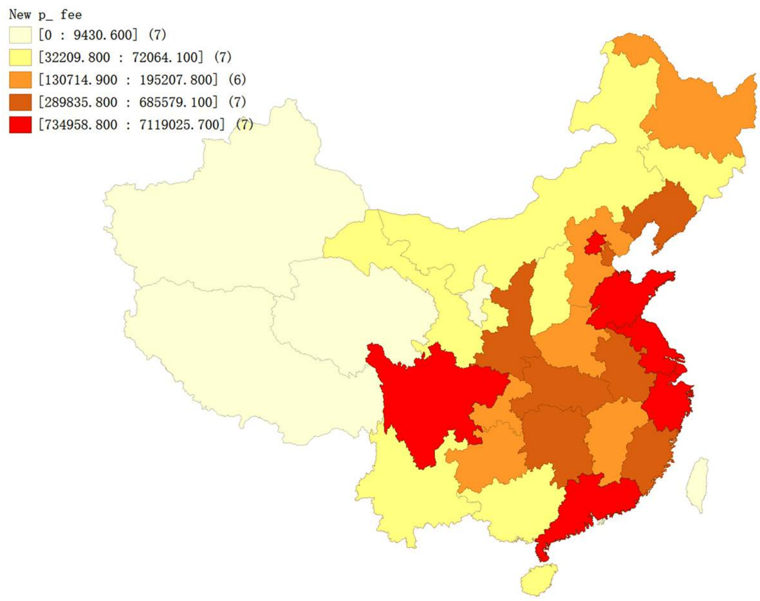


Figure 7. Spatial map of new product development costs.

As the map (**Figure 8**) below shows, new product sales in Guangdong are of high value area. Sales of new products in Fujian, Zhejiang, Jiangsu, Shandong, Beijing, Shanghai, Tianjin, Hunan, Jiangxi, Liaoning and Sichuan are of relatively high value area. Heilongjiang, Jilin, Hebei, Henan, Guizhou, Guangxi are relatively low. Xinjiang, Tibet, Qinghai, Ningxia, Inner Mongolia, Gansu and Yunnan provinces are low. From the perspective of regional distribution, the high-value areas are basically distributed in the eastern part of China, and gradually transit to the low-value areas in the western inland.

9) Spatial map of the number of patent applications

It can be seen from spatial map (**Figure 9**), Guangdong has the highest value of application for patent. Fujian, Zhejiang, Jiangsu, Shandong, Sichuan, Shanghai, Beijing, Hubei, Hunan, Anhui, Henan, Liaoning are of relatively high value area. Heilongjiang, Jilin, Hebei, Guizhou, Chongqing, Shanxi are relatively low. Xinjiang, Tibet, Qinghai, Ningxia, Inner Mongolia, Gansu are low. From the perspective of regional distribution, except that Sichuan is a high-value region, the high-value region is basically distributed in the eastern region of China, and gradually transfers to the western inland region as a low-value region.

10) Spatial map of valid invention patents

It can be seen from **Figure 10** that Guangdong is the best in regards of high value of invention patent. Zhejiang, Jiangsu, Shandong, Sichuan, Beijing, Shanghai, Tianjin, Hunan, Hubei, Anhui, Fujian, Liaoning are of relatively high value area. Jilin, Heilongjiang, Henan, Guizhou, Guangxi, Yunnan are relatively low. Xinjiang, Tibet, Qinghai, Ningxia, Inner Mongolia, Gansu are low in valid invention patents. From the perspective of regional distribution, the high-value areas are basically distributed in the eastern part of China, and gradually transit to the western inland areas. Sichuan province in the central part belongs to the high-value area, showing a prominent spatial display.

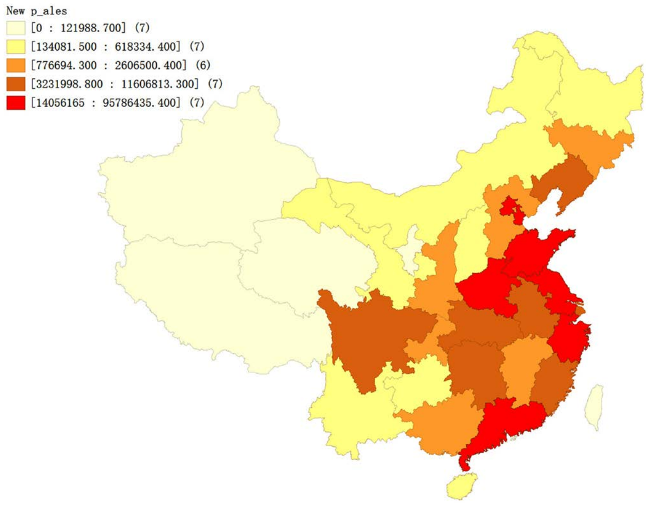


Figure 8. Spatial map of new product sales.

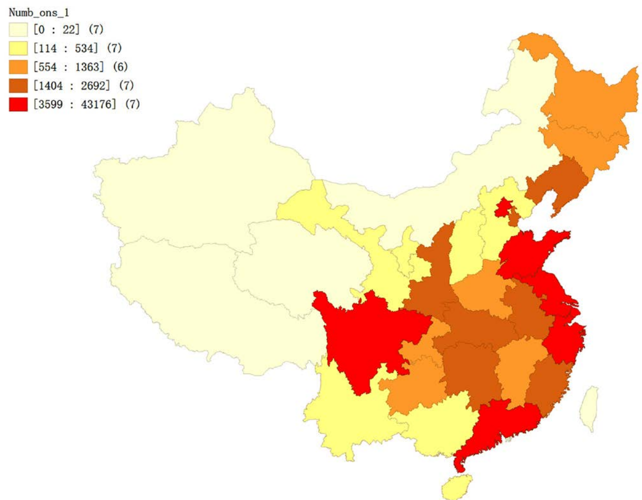


Figure 9. Spatial map of the number of patent applications.

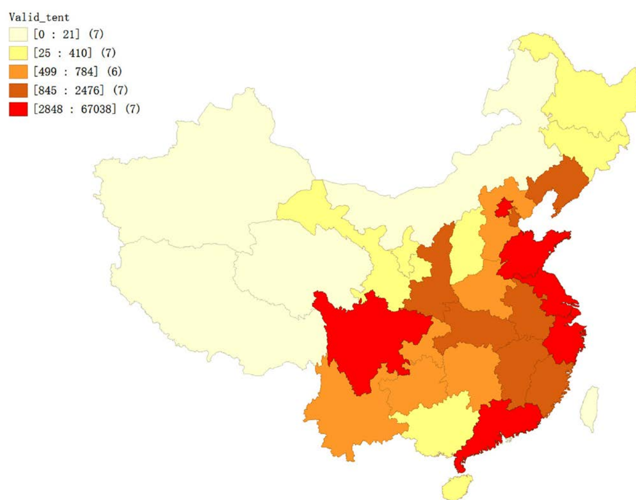


Figure 10. Spatial map of valid invention patents.

3. Spatial Principal Component Analysis Method and Application

3.1. Principal Component Analysis Method

Principal component analysis (PCA) is a powerful tool for dimensionality reduction of variables, and its basic idea is to try to recombine correlated indicators into a new set of independent comprehensive indicators to replace the original ones. The mathematical solution is to take the original p indices as a linear combination of the new indices. The first linear combination, namely the first composite index, is denoted as y_1. In order to make this linear combination unique, it is required that the variance of y_1 is the largest among all linear combinations. If the first principal component is not enough to represent all the information of the original p indexes, then consider selecting the second principal component y_2, and require the existing information of y_1 not to appear in y_2, that is, Cov (y_1, y_2) = 0. Figure 11 shows schematic diagram of principal component analysis.

As shown in Figure 11, the indexes (scatter points) with x_1-x_2 as the axis have a large projection on x_1 and x_2, so both indexes contain the necessary information of the data. However, if we find out there is a strong correlation between x_1 and x_2, and the data could be distributed along y_1, then we can carry out simple transformation.

$$\begin{cases} y_1 = \cos\theta x_1 + \sin\theta x_2 \\ y_2 = -\sin\theta x_1 + \cos\theta x_2 \end{cases} \tag{1}$$

After this transformation, the information of the original data is mainly contained on the y_1 axis under the new coordinate axis, while the information projection on the y_2 axis is very small, which can be ignored. If only y_1 is selected for study, we can successfully achieve the dimensionality reduction processing.

The case of multidimensional variables is similar to that of two-dimensional

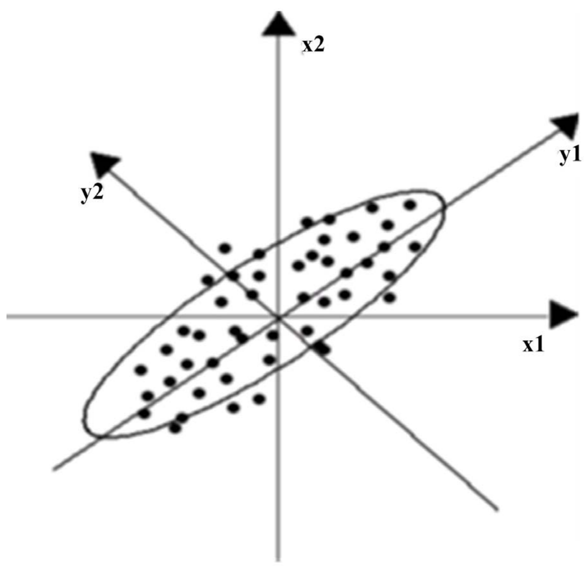


Figure 11. Schematic diagram of principal component analysis.

variables, but we can only imagine the abstract space, such as a multidimensional ellipsoid with high fluctuation. In order to reduce the dimension, the principal component analysis is basically completed by first finding out each principal axis of the ellipsoid. Then calculate the new axis standard which can represent most of the principal axis information as a new variable. Similar to the two-dimensional case, the spindles of the higher dimensional ellipsoid are also required to be perpendicular to each other, and these new spindles that are orthogonal to each other are linear combinations of the original spindles, which are the principal components. The fewer principal components can be selected for study, the better the dimensionality reduction effect will be. Most researchers believe that the criterion for selecting principal components is that the sum of the principal axes represented by the new principal components should account for most of the sum of the original principal axes. Some scholars suggest that the total length of the spindle selected accounts for more than 80% of the total length of all spindles. This principle can be used as a basic treatment principle.

Principal component analysis process:

1) Find the eigenvalues and eigenvectors of the correlation matrix;

2) Calculate variance contribution rate and cumulative variance contribution rate: the contribution rate of each principal component represents the percentage of the total information of the original data;

3) Determine the principal components: let comp.1, comp.2 ... comp.p be p principal components, in which the total amount of data information contained by the first m principal components (*i.e.* their cumulative variance contribution rate) is not less than 80%, the first m principal components should be used to reflect the original evaluation object.

4) Use the linear combination of the original indexes to calculate the scores of each principal component: take the eigenvector as the weight, represent each principal component as the linear combination of the original indexes, and the meaning of the principal component is determined by the comprehensive meaning of the indexes with greater weight in each linear combination.

$$\text{Comp.}j = u_{j1}x_1 + u_{j2}x_2 + \cdots + u_{jp}x_p \quad j = 1, 2, \cdots, m \tag{2}$$

The formula is called principal component score function, which is used to calculate the principal component score of each sample. If $m = 2$, then the p variables of each sample can be substituted into the above equation to calculate the principal component scores of each sample, then they can be used as scatter plots of principal component scores on the plane so that the samples can be classified.

5) Take the variance (characteristic value) of each principal component as weight, obtained comprehensive score.

$$\text{Comp} = \frac{\lambda_1 \text{Comp.1} + \lambda_2 \text{Comp.2} + \cdots + \lambda_m \text{Comp.}m}{\lambda_1 + \lambda_2 + \cdots + \lambda_m} = \sum_{j=1}^{m} W_j \text{Comp.}j \tag{3}$$

where W_j is the weight of the principal component, the ranking can be obtained by using the total score.

The principal component analysis function is used to analyze data, where X is the data box, m is the number of factors, whose default is 2. Plot is the main component graph. When these parameters are TRUE, the function program will automatically calculate and output the result. When the parameter value is FALSE, it is not calculated.

3.2. Example for Principal Component Analysis

Next, due to the fact that there are many years, we will take the data of year 2019 as an example for principal component analysis.

1) find the principal components of the correlation matrix

2) determine the principal component

From Table 4, according to the principle that the contribution rate of cumulative variance is greater than 80% and the variance is greater than 1, two principal components are selected, and the contribution rate of cumulative variance is 97.72%. In this case, $m = 2$.

3) principal component coefficient

Obtained from Table 5

$$y_1 = 0.251x_1 + 0.25x_2 + 0.251x_3 + 0.25x_6 + 0.25x_7 + 0.247x_9 + 0.249x_{12}$$

$$y_2 = 0.266x_{13} + 0.241x_{14} + 0.346x_{15} + 0.312x_{16} + 0.324x_{17} + 0.382x_{18} + 0.415x_{19}$$

Choosing only two principal components, the main two components should be production management, research and development investment. Therefore, the first principal component (y_1) can be regarded as technology research and production. Under the condition that other conditions remain unchanged, whenever x_1 means that the number of high-tech enterprises increases by 1 unit, the corresponding y_1 increases by 0.251 units; whenever x_2 means that average annual employee increases 1 billion, the corresponding y_1 increases by 0.25 units; whenever x_3 means that the main business income of high-tech enterprises increases by 1 billion, the corresponding y_1 increases by 0.251 units. The second principal component (y_2) is the fixed capital investment. Under the condition that other conditions remain unchanged, whenever x_{14} means that the number of patent applications increases by 1 unit, the corresponding y_2 increases by 0.241 units; whenever x_{18} means that the project investment increases by 1 unit, the cor-responding y_2 increases by 0.346 units; whenever x_{19} means that the new fixed assets increases by 1 unit, the cor-responding y_2 increases by 0.415 units.

4) Principal component score

From Table 6, we can see the top four provinces with the highest score in the first principal component C1 are Jiangsu, Shandong, Zhejiang and Guangdong.

Table 4. Contribution of principal component variance of high-tech industry data.

	Variance	Contribution (%)	Cumulative Contribution (%)
Comp.1	17.973	0.8316	0.8316
Comp.2	2.766	0.1456	0.9772

Table 5. Principal component coefficients of high-tech industry data.

	Comp.1	Comp.2		Comp.1	Comp.2
Number of high-tech enterprises	0.251	0.000	New product sales	0.249	0.000
Average annual employee	0.250	0.000	Technical renovation fund	0.179	0.266
Main business income	0.251	0.000	Number of patent applications	0.232	0.241
Total profit	0.240	−0.161	Valid patent for invention	0.210	0.346
Export delivery value	0.238	0.000	Number of construction projects	0.210	0.312
Number of R & D institutions	0.251	0.000	Completed project	0.207	0.324
R & D staff	0.250	0.000	Project investment	0.190	0.382
R & D institution funding	0.232	0.252	New fixed assets	0.174	0.415
R & D equivalent	0.247	0.126	New product development fee	0.236	0.228
R & D internal expenditure	0.238	0.211	-	-	-

Table 6. Principal component scores of high-tech industry.

Area	Comp.1	Comp.2	Area	Comp.1	Comp.2
Beijing	−0.56922	1.20239	Hubei	−0.31982	−0.25273
Tianjin	−1.06464	0.28377	Hunan	−0.15529	−1.09958
Hebei	−1.08827	−0.39530	Guangdong	16.07246	5.42759
Shanxi	−1.99377	0.39603	Guangxi	−1.47885	−0.30745
Inner Mongolia	−2.08348	0.31796	Hainan	−2.31786	0.68101
Liaoning	−1.66353	0.59022	Chongqing	−0.77357	−0.37851
Jilin	−1.44458	−0.30802	Sichuan	−0.01793	−0.28281
Heilongjiang	−1.95841	0.47696	Guizhou	−1.83069	0.40188
Shanghai	−0.32989	1.17059	Yunnan	−2.16332	0.54259
Jiangsu	12.21032	−5.59727	Tibet	−2.38676	0.69376
Zhejiang	2.07358	−0.32189	Shaanxi	−0.90553	0.08546
Anhui	0.13197	−1.19680	Gansu	−2.17284	0.47988
Fujian	0.32475	−0.69193	Qinghai	−2.32522	0.62475
Jiangxi	−0.56062	−0.83220	Ningxia	−2.29162	0.64178
Shandong	2.71445	−1.34866	Xinjiang	−2.31486	0.63766
Henan	0.68303	−1.64114	-	-	-

The research input and output level of these regions is far higher than that of other regions. Jiangsu and Shandong are significantly ahead of other regions. The level of research input and production output is relatively low in Tibet, Qinghai and Ningxia. The second principal component indicates that Jiangsu, Shandong, Zhejiang and Sichuan spend more on fixed assets. It can also be seen from the scatter map that Guangdong and Jiangsu are in the leading position in the development of high-tech industry, followed by Shandong, Zhejiang, Shanghai, Beijing and Sichuan, other regions have relatively low scores, and it is difficult to make intuitive distinctions due to the impact of image size. For composite scores, Jiangsu, Guangdong and Zhejiang have highest score. Tibet, Gansu, Ningxia and Xinjiang score in the end of the country. Therefore, Jiangsu, Guangdong, Zhejiang and Shandong are the four provinces in the national front row of high technology industry development while Tibet, Gansu, Ningxia and Xinjiang fall behind in high technology industry development.

3.3. Space Comprehensive Evaluation Based on Principal Components

Finally, the comprehensive score was estimated by the weighting method, and the proportion of each principal component's variance contribution to the total variance contribution of the two principal components was used as the weight to make the weighted summary, and the comprehensive score and ranking of the provinces were obtained.

In terms of comprehensive scores from Table 7, Guangdong and Jiangsu scored the highest, Shandong, Zhejiang and Henan ranked 3rd, 4th and 5th, and Ningxia, Qinghai, Xinjiang and Tibet ranked last in the country. In terms of development of high-tech industry, it is true that Guangdong, Jiangsu, Shandong, Zhejiang and other coastal areas are in the leading position of development, while the vast western region lags behind, the central region is relatively backward, but Sichuan is a relatively high point in the central and western regions.

From Figure 12, uses the results of spatial statistical function analysis, we

Table 7. Score of principal components of high-tech industry in 2019.

Area	Score	Rank	Area	Score	Rank
Beijing	−0.32732	12	Hubei	−0.31066	11
Tianjin	−0.88053	16	Hunan	−0.28423	10
Hebei	−0.99365	17	Guangdong	14.61901	1
Shanxi	−1.66747	23	Guangxi	−1.31891	9
Inner Mongolia	−1.75559	24	Hainan	−1.90839	28
Liaoning	−1.35580	20	Chongqing	−0.71963	14
Jilin	−1.28939	18	Sichuan	−0.05410	8
Heilongjiang	−1.62588	22	Guizhou	−1.52585	21
Shanghai	−0.12501	9	Yunnan	−1.79386	25
Jiangsu	9.77886	2	Tibet	−1.96615	31
Zhejiang	1.74650	4	Shaanxi	−0.77022	15
Anhui	−0.04946	7	Gansu	−1.81064	26
Fujian	0.18593	6	Qinghai	−1.92243	30
Jiangxi	−0.59770	13	Ningxia	−1.89110	27
Shandong	2.15967	3	Xinjiang	−1.91172	29
Henan	0.36569	5	-	-	-

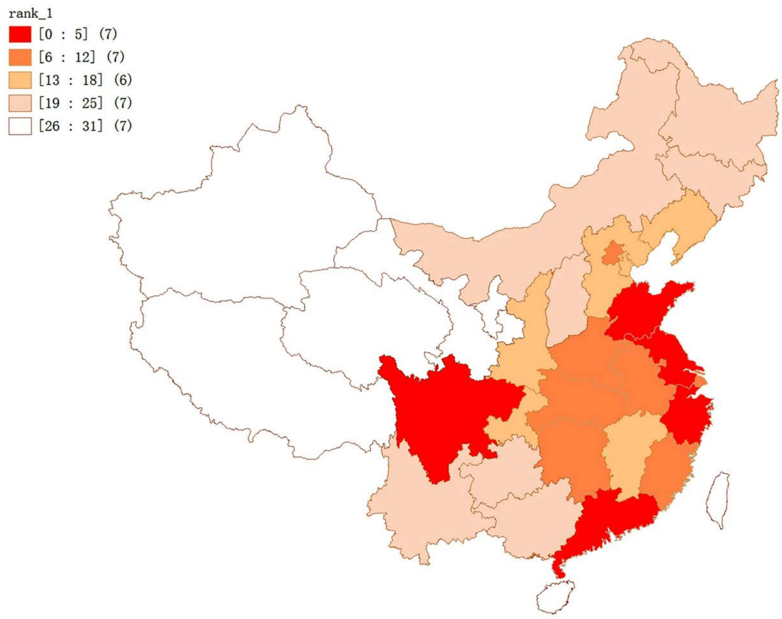


Figure 12. Principal component scores of high-tech industry for different provinces in 2019.

got this picture, which is very intuitive to reflect the various provinces in 2019 of high technology industry in China. Guangdong, Jiangsu and Zhejiang are the best, south China, east China and Sichuan is of high value, the whole picture presents the characteristics of the southeast of China is more advanced while the northwest part of China falls behind.

3.4. Comprehensive Evaluation of Principal Component Space from 2016 to 2019

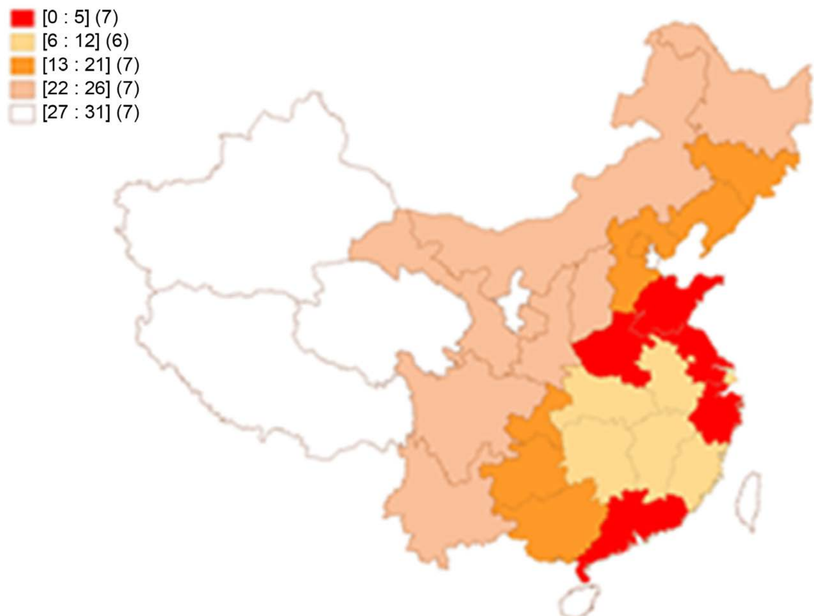
From Table 8 based on the principal component analysis of the first-level indicator, it can be seen that the scores of Guangdong and Jiangsu have been stable in the first and second positions, and the scores of these two provinces are relatively close. The two provinces have their own advantages in the development of high-tech industries. Shandong and Zhejiang have been in the third and fourth place, relatively stable; Hebei province has made progress; however, the scores of Beijing, Shanghai, Tianjin and Sichuan have been declining. Hubei province has the most obvious decline, falling from the 8th place to the 30th place. Ningxia, Xinjiang, Qinghai are at the bottom of the list. Tibet made little progress. From 2016 to 2019 Guangdong, Jiangsu, Shandong, Zhejiang scored the highest on the first principal component C1, which is the research level of input and output. Jiangsu significantly ranks ahead of other areas. The level of research input and production output is relatively low in Tibet, Qinghai and Ningxia. The second principal component indicates that Jiangsu, Shandong, Zhejiang and Sichuan spend more on fixed assets. It can also be seen from the scatter map that Guangdong and Jiangsu are in the leading positions in the development of high-tech industry. Shandong, Zhejiang, Shanghai, Beijing and Sichuan

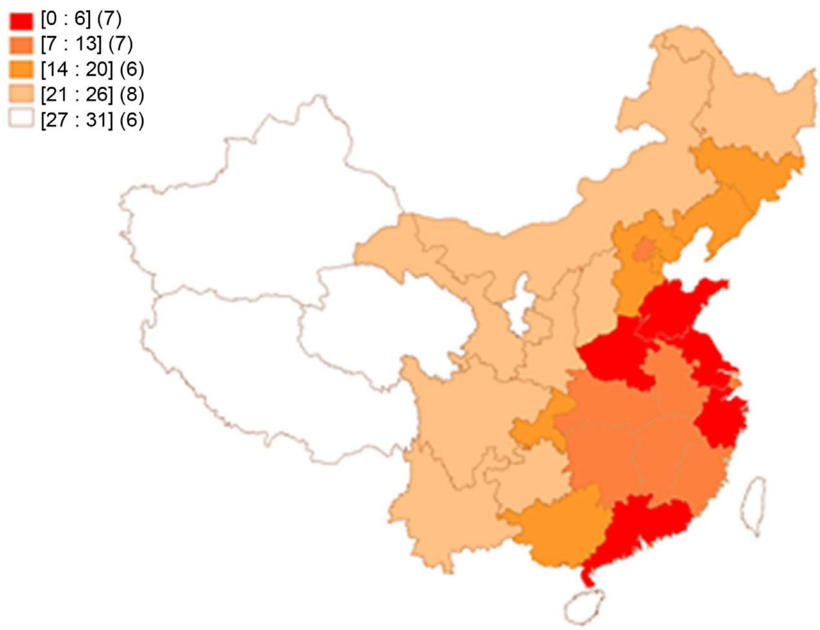
Table 8. Results of principal component scores of first-level indicators in provinces from 2016 to 2019.

	2016		2017		2018		2019	
Area	Score	Rank	Area	Score	Rank	Area	Score	Rank
Beijing	−0.8708	14	−0.9109	10	−0.82417	11	−0.32732	12
Tianjin	−0.9408	17	−0.6755	14	−0.79291	14	−0.88053	16
Hebei	−0.8068	16	−0.8551	17	−0.92349	17	−0.99365	17
Shanxi	−1.8386	23	−1.7824	22	−1.85246	22	−1.66747	23
Inner Mongolia	−1.9051	24	−1.9123	24	−1.95255	24	−1.75559	24
Liaoning	−1.5769	20	−0.6838	15	−0.55329	15	−1.35580	20
Jilin	−1.1869	18	−1.21222	19	−1.186144	19	−1.28939	18
Heilongjiang	−1.14681	22	1.07014	21	−1.0958	21	−1.62588	22
Shanghai	0.44844	11	0.73822	7	0.68685	7	−0.12501	9
Jiangsu	6.52389	2	6.94451	2	6.72942	2	9.77886	2
Zhejiang	2.0813	4	2.22688	4	2.35884	4	1.74650	4
Anhui	0.4433	7	−0.12663	12	0.00789	12	−0.04946	7
Fujian	0.4599	6	−0.1417	8	−0.2144	8	0.18593	6
Jiangxi	−0.2567	12	−0.2441	13	−0.2871	13	−0.59770	13
Shandong	2.8926	3	2.7656	3	2.86395	3	2.15967	3
Henan	0.7569	5	0.4996	6	0.77045	6	0.36569	5
Hubei	−0.1423	11	−0.0331	9	0.10690	9	−0.31066	11
Hunan	0.1273	9	−0.09853	11	0.047849	10	−0.28423	10
Guangdong	14.0240	1	12.37934	1	13.49664	1	14.61901	1
Guangxi	−1.2282	19	−1.34103	20	−1.273958	20	−1.31891	9
Hainan	−2.2003	29	−2.10835	27	−2.202171	27	−1.90839	28
Chongqing	−0.5461	13	−1.16314	18	−0.950924	18	−0.71963	14
Sichuan	0.1497	24	0.39331	24	0.289262	5	−0.05410	8
Guizhou	−1.6921	21	−1.90070	23	−1.889755	23	−1.52585	21
Yunnan	−2.0290	25	−1.96673	25	−2.051084	25	−1.79386	25
Tibet	−2.2705	31	−2.22929	31	−2.301347	31	−1.96615	31
Gansu	−2.0277	26	−1.94005	26	−1.939874	26	−1.81064	26
Qinghai	−2.1986	30	2.20968	30	−2.8465	30	−1.92243	30
Ningxia	−2.1708	27	−2.16429	28	−2.043669	28	−1.89110	27
Xinjiang	−2.1933	28	−2.1759	29	−2.03732	29	−1.91172	29

follow. Other regions have relatively low scores, and it is difficult to make intuitive distinctions due to the impact of image size. For composite scores, Guangdong, Jiangsu and Zhejiang have high score. Tibet, Gansu, Ningxia, Xinjiang score at the end of the country. Thus, Jiangsu, Guangdong, Zhejiang, Shandong are in the national front row of high technology industry development. With the high level of research and development investment, patent number, fixed assets, high technology industry development is relatively high. In the economically backward western regions, the number of new product patents is reduced due to the low investment in high-tech research and development. In addition, the fixed asset investment in northwest regions is small, that is why the development of high-tech industries in these regions falls behind.

Through space comparison of annual statistics (**Figure 13**), we can clearly see that China's high-tech industries have continuously to be concentrating in the coastal and southern part of the country. Guangdong province in the Pearl River delta coastal areas has remained number one from 2016 to 2019. Jiangsu province in the Yangtze River delta coastal area remains steadily in the second place





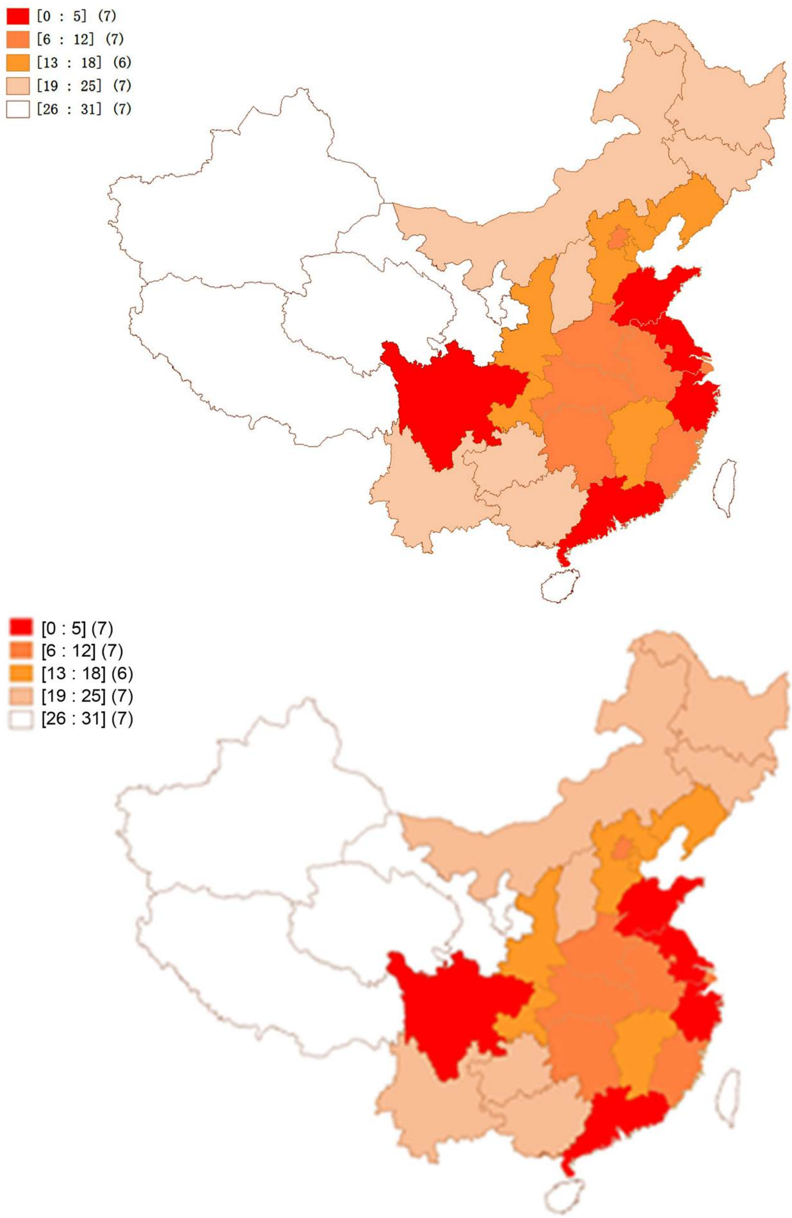


Figure 13. Principal component scores of high-tech industry for different provinces from 2016 to 2019.

from 2016 to 2019. Shandong and Zhejiang provinces have remained third and fourth in recent years. Sichuan province is an important support for the continuous rise of central China. The development of high-tech industries in central regions such as Hunan, Jiangxi is relatively slow. Northeast China, such as Heilongjiang and Jilin, are declining, while western regions, such as Xinjiang, Qinghai, Gansu, are lagging far behind.

4. Conclusion and Suggestions

From the research results of this paper, high technology industry in China has made rapid progress. However, since the disparity of China's vast circumstances is great, we should gradually establish and improve scientific management according to different regional classified guidance. In addition, we should also promote balanced development among different regions. In this paper, we set up a statistical analysis according to the data of high technology industry development of 31 provinces from 2016 to 2019. The indexes we used in this paper include the production and operation and situation of research and development (R & D), new patent, fixed investment. We also adopt the method of principal component analysis to high technology industry in China. Statistical analysis of the traditional model combined with the spatial correlation analysis is used in this paper. Moreover, the integrated use of R language makes the results of the analysis on the spatial maps available to be visualized. From the perspective of regional distribution, the development of high-tech industry in China is extremely unbalanced, which is fast in the east and slow in the west. The development resources are concentrated in the Yangtze River delta and the Pearl River Delta.

Based on the statistical research of the high-tech industry, this paper puts forward the following Suggestions:

First, establish a set of comprehensive indexes for the development of high-tech industries. Make sure that production and business operation activities, research and development, patents, fixed asset is feasible and effective. Forming a comprehensive index is not only more advantageous to simple index, but also is clearer to show a regional industrial development level. Therefore, we suggest establishing a comprehensive index for the development of high-tech industry, so as to facilitate the reference of government when making decisions.

Second, carry out the classification to guide the development according to the region difference. China can implement regional classified development. Since Guangdong and Jiangsu lead the development of high-tech industries, we should take the two provinces as examples for other provinces. Guangdong and Jiangsu should have the courage to innovate and vigorously develop high-tech industries by extending existing advantages. Northeast China should further accelerate the pace of industrial catch-up. In central China, Sichuan province is the center to accelerate the realization of industrial upgrading. Since the western region, represented by Tibet, Qinghai and Xinjiang, is sparsely populated and lacks the industrial foundation. Therefore, it should give up the development of high-tech industries and adopt the development of basic industries.

In addition, the development of high-tech industry in Guangdong province has made great contributions to the country. However, the development of surrounding provinces of Guangdong is in a relatively backward state due to their small correlation with the development of Guangdong province. It is suggested that Guangdong should encourage and promote the transfer of Guangdong's high-tech industry to surrounding areas in industrial policy and other aspects, so as to strengthen the support of Guangdong's development.

Moreover, Hong Kong and Macao bay area financial agglomeration has the ability to promote the development of high technology industry, therefore, this area should develop direct financing to alleviate financial institutions. What is more, we should encourage large technological enterprises to set up production in rural area so as to promote the development of high technology industries.

Lastly, the western region should take the opportunity of "One Belt and One Road" construction to expand its economic openness to the outside world and promote the coordinated development of domestic and foreign markets. The vast western region, Inner Mongolia, Xinjiang, Tibet, Yunnan should not only seize the construction opportunity, but also actively participate in the cooperation of high technology industries both at home and abroad. The undeveloped area in China should promote innovation ability to improve the research level of science and technology for high-tech industry development.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Chen, Y.F. and Wang, B.H. (2004) R Flexible, Powerful Free Statistical Analysis Software. *Statistics and Decision-Making*, No. 9, 49-50.

[2] Cliff, A.D. and Ord, J.K. (1981) Spatial Processes: Models and Applications. Pion Limited, London.

[3] Anselin, L. (1988) Spatial Econometrics: Methods and Models. Kluwer, Dordrecht. https://doi.org/10.1007/978-94-015-7799-1

[4] Getis, A. and Ord, J.K. (1992) The Analysis of Spatial Association by the Use of Distance Statistics. *Geographical Analysis*, **24**, 189-206.

[5] Friedman, J.R. (1966) Regional Development Policy: A Case Study of Venezuela. MIT Press, Cambridge.

[6] Krugman, P.R. (1991) Increasing Returns and Economic Geography. *Journal of Political Economy*, **99**, 483-499. https://doi.org/10.1086/261763

[7] Fujita, M. (1988) A Monopolistic Competition Model of Spatial Agglomeration: A Differentiated Products Approach. *Regional Science and Urban Economics*, **18**, 87-124. https://doi.org/10.1016/0166-0462(88)90007-5

[8] Venables, A.J. (1996) Equilibrium Locations of Vertically Linked Industries. *International Economic Review*, **37**, 341-359. https://doi.org/10.2307/2527327

[9] Okabe, A., Boots, B., Sugihara, K., *et al.* (2000) Spatial Tessellations: Concepts and Applications of Voronoi Diagrams. 2nd Edition, John Wiley & Sons, Chichester.

[10] Lloyd, C.D. (2007) Local Models for Spatial Analysis. CRC Press, Boca Raton.

[11] Wu, Y.M. (2006) Local Spatial Econometric Analysis of University, Enterprise R & D and Capital Region Innovation. *Scientific Research*, No. 3, 398-404.

[12] Wu, Y.M. (2006) Application Research of Spatial Econometric Model in Provincial R & D and Innovation. *Quantitative Economic Technical Economic Research*, No. 5, 101-108.

[13] Chen, H.C. (2010) Empirical Research on Evaluation of Technological Innovation Capability of High-Tech Industry. *Science and Technology Management Research*, **30**, 20-22+49.

[14] Zheng, S.W. and Xu, Z.L. (2016) PLS-Based High-Tech Industrial Technological Innovation Capability and Evaluation of the Three Northeast Provinces. *Science and Technology Management Research*, **36**, 86-93.

[15] Fan, D.C. and Du, M.Y. (2017) Dynamic Comprehensive Evaluation of Technological Innovation Capability of High-Tech Industries Based on TOPSIS Grey Correlation Projection—From the Perspective of Beijing-Tianjin-Hebei Integration. *Operations Research and Management*, **26**, 154-163.

[16] Wu, Y.X. and Zhou, C.G. (2018) Research on the Evaluation of the Technological Innovation Capability of High-Tech Industries in the Western Region—Comprehensive Evaluation Model Based on the Combined Weighting Method of Entropy Weight and Variation Coefficient. *Xinjiang Agricultural Reclamation Economy*, No. 12, 83-88.