International Journal of Database Theory and Application
Vol.9, No.11 (2016), pp.353-362
http://dx.doi.org/10.14257/ijdta.2016.9.11.31

Big Data Cloud Computing Platform Application in Performance Evaluation Index of Regional Innovation Capability

Wu Hongxia[1], Zhao Dongmei[2] and Liu Ying[3]

[1] *College of Management, North China University of Science and Technology, Tangshan 063000, China*

[2] *Finance Department, North China University of Science and Technology, Tangshan 063000, China*

[3] *College of Foreign Languages, North China University of Science and Technology, Tangshan, 063000, China*

Abstract

In this paper, the authors analyze the big data cloud computing platform application in performance evaluation index of regional innovation capability. Relevant domestic researches of regional innovation capability performance measures are expounded. In view of this, an evaluation index system of regional innovation capability is created, including the number of R&D personnel, input indicators of R&D funds and the output indexes of the patent application and grant. The regional innovation capability was evaluated by choosing relevant data from 2012 to 2014 in various cities in Hebei Province by DEA method. The results show that the scale efficiency of regional innovation capability in Hebei Province is increasing and the regional innovation capability has the same tendency with the efficiency. The increasing input of innovation resources and improving the efficiency of innovation resources are the inevitable choice.

Keywords: *DEA method; Cloud computing platform; Regional innovation capability; Performance evaluation*

1. Introduction

The world economic development shows the regionalization as characteristics of the economic globalization is deepening. Regional economy is becoming the focus of the global economic activity [1]. The regional innovation capability is not balanced in our country because of the geographical conditions and economic resources. Before China's reform and opening up, the innovation ability of southern coastal areas lags behind that of northern region. The economy of southern coastal areas began to develop rapidly, while the north heavy industry base declined [2]. The innovation capability of southern regions is far stronger than that of the northern regions. Under the background of "Volkswagen entrepreneurial innovation", local governments attach great importance to improving regional innovation ability. In order to enhance the innovational capability, local governments must make use of the region innovation resource utilization level and improve the regional innovation performance. The evaluation index system of regional innovation capability is constructed from the angle of input and output, and evaluated by methods of Data Envelopment Analysis (DEA). Some measures are put forward to enhance the capability of regional innovation according to the evaluation results.

The performance evaluation of regional innovation capability is a very complex system. Domestically, the researches about the performance evaluation of regional innovation ability are mainly concentrated on three aspects: (1) the study of performance evaluation connotation of regional innovation ability. Guan Jiancheng, Liu Shunzhong (2003) pointed out that the regional innovation system means the innovation of human

ISSN: 2005-4270 IJDTA

resources and conversion of financial resources into the output of regional innovation system [4], which means that the regional innovation performance is the ratio of input to output. Tang Houxing, Liang Wei (2005) considered that the innovation performance of regional innovation system is the innovation efficiency and effect. (2) studies on the corresponding evaluation index system of regional innovation ability. Zhou Hongwen (2014) built the evaluation system from R&D staff full-time equivalent [2], with R&D spending as a share of GDP and 14 other factors; Yi Fan, Shan Jieying (2011) built the entropy weight method evaluation system from four primary indexes, manpower, funds input, knowledge output and economic output and so on four primary indexes and the activities of science and technology people and so on seven secondary indexes. (3) studies on the performance evaluation method of regional innovation ability. Tang Houxing (2006) built the performance evaluation model based on Data Envelopment Analysis (DEA) method. The comprehensive performance of decision making units, technical efficiency and scale efficiency were analyzed. The "based on the multiple attribute decision making of maximizing deviations method", "typical example determination method" and "goal setting method" for effective decision making units and inefficient decision making units are analyzed on the basis of DEA analysis. Zhou Hongwen (2014)evaluated the resource allocation efficiency and regional innovation performance in China by the chain link network DEA model method and grey relation entropy method. Bai Junhong (2009) evaluated the innovation efficiency of regional innovation system by the method of nonparametric analysis DEA and revealed the efficiency of science and technology input and output and found out the cause of regional innovation efficiency in all provinces.

Yin Fan, Shan Jieying (2011) defined the regional innovation performance evaluation index system from two aspects of regional innovation efficiency and effect and constructed the composite measure model on the basis of entropy weight method and coordination degree method. The comprehensive evaluation of the regional innovation performance has been carried out by using the composite measure model in Hebei province. Researches of the existing literature show that the selection standard of the domestic scholars in the performance evaluation of regional innovation capability index is different. Different scholars build the evaluation index system from different angles. Most scholars establish the index system from two aspects as innovation input and output. The input indicators include people, goods, content resources, while the output indicators include patents, books, paper and other indicators. Most scholars use the evaluation method such as two stage of DEA, principal component analysis, factor analysis, entropy value and grey correlation analysis methods and so on. On the basis of previous researches the comprehensive evaluation of regional innovation performance in Hebei province is studied by the data envelopment analysis (DEA) method, which guarantees the scientific nature and effectiveness to the largest extent.

2. Big Data Cloud Platform

Big data technology, networking, cloud computing is a technology revolution of the IT industry's subversive. The era of big data on people's lifestyle, business models have had a significant impact. The big data has brought new opportunities to the information technology industry, especially the influence on data mining technology obvious. Data mining technology has entered a new stage of development, to improve the accuracy of big data, big data mining algorithm on the data error rate must be within acceptable limits, it is necessary for the traditional data mining algorithm. To improve the data mining efficiency and accuracy for target and clustering for large data mining algorithm accuracy and efficiency as the research focus on the traditional clustering algorithm was improved and the use of cloud computing platform will be changed The improved clustering algorithm has a good theoretical and practical value, and can be widely used in large data sets.

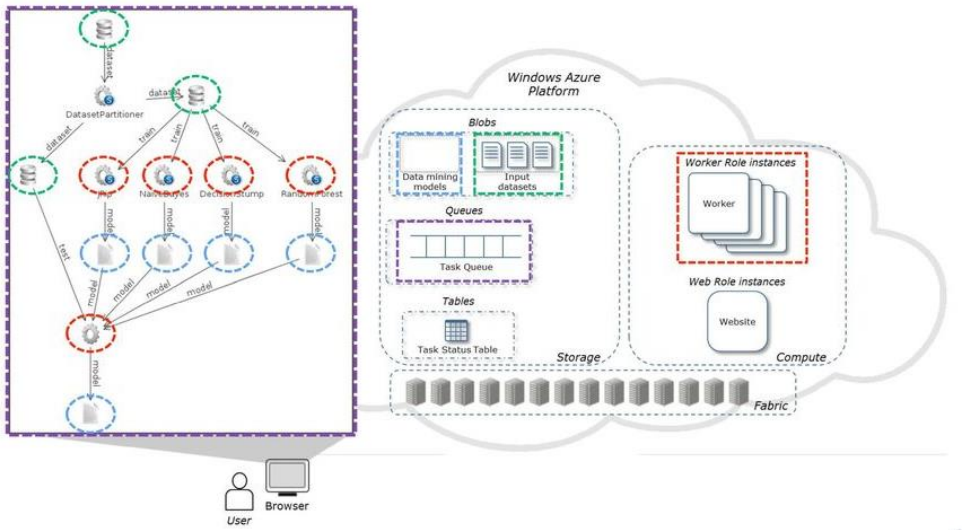

Figure 1. Big data mining platform

Cloud computing as an Internet service delivery and usage model based on, is through the network to the on-demand, scalable way to get the service needed. Its core idea is to use a large number of computing resources connected to the network unified management and scheduling, constitute a pool of resources to provide on-demand services for users of cloud. The calculation to provide the data storage center of the most reliable and safe, the user can no longer worry about data loss viruses and other problems. The application model of cloud computing are: software as a service (SaaS), platform as a service (PaaS), the basic implementation of the service (IaaS). Hadoop an open source framework, can write and run the distributed application to large-scale data processing (platform). Its main characteristic is that users cannot understand the underlying details of the distributed situation, development and distribution the program takes full advantage of the power of cluster computing and storage

1) HDFS: distributed file system, distributed storage can be achieved;

2) MapReduce: distributed programming framework, distributed computing can be achieved

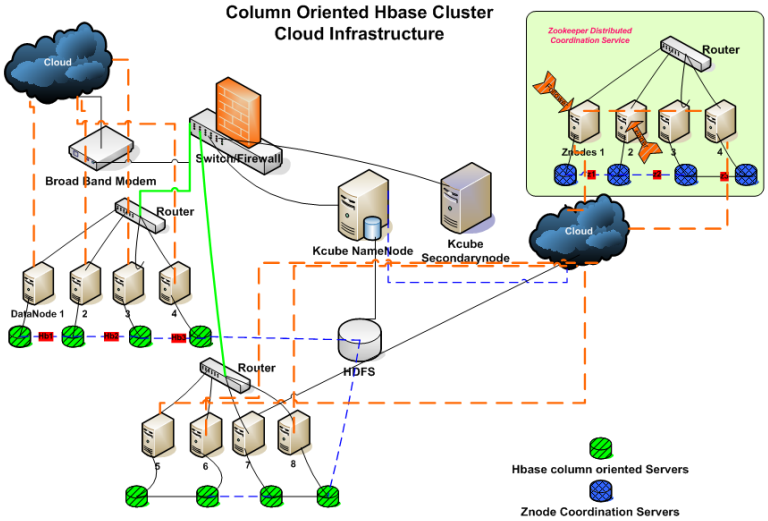

Figure 2. Hadoop cloud platform

Big data technology, because of the rapid development of the Internet, cloud computing, mobile and networking in recent years and became a hot topic of concern. Big data is a large scale in the acquisition, storage, management and analysis greatly beyond the traditional scope of database software tools for data collection, data size, fast data the mass transfer, with various data types and low value density four characteristics. From a technical perspective, the relationship between big data and cloud computing is inseparable. Big data must adopt distributed architecture in the use process, which is characterized by distributed mass data mining. But it must rely on cloud computing distributed processing, distributed database and cloud storage, virtualization technology.

Clustering analysis is the core technique of data mining, clustering analysis based on the Like attracts, the naive idea, according to the clustering or classification of things characteristics. From the implicit level, the result of cluster analysis will be a collection of data objects; we call the objects in the cluster. The cluster is similar to each other however, other clusters of objects is different. In many applications, the data can be an object in the cluster as a whole. The earliest clustering mining technology in the field of statistics and artificial intelligence has been widely studied. In the field of data mining, often face database containing a massive data, therefore, to continue Improve the clustering method for large scale database to meet the new challenges.

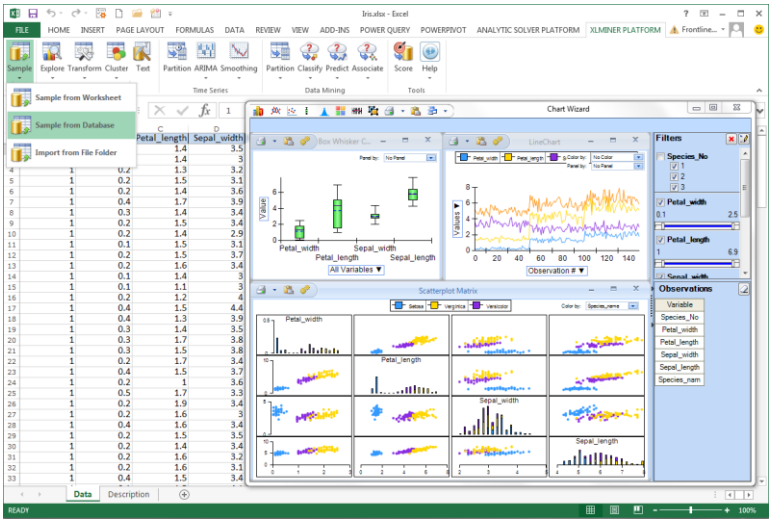

Figure 3. Data mining platform

3. Performance Evaluation Model of Regional Innovation Capability

3.1. Construction of regional innovation performance evaluation model

Data Envelopment Analysis (DEA) is a kind of evaluation index multiple input and output decision making units the relative effectiveness of multi-objective decision using method. The DEA method needn't carry on the parameter estimation, avoiding the subjective factors and accordingly reducing errors. The DEA model include four kinds of models such as CCR, BCC, C2GS2 and C2WH. The most basic model is CCR, which isn't measured by scale reward to assume. The decision making units (DMU) is whether achieve effectively technology and scale at the same time. The BBC model is measured by the size of variable remuneration efficiency. The C2GS2 model is the study of the technical effectiveness of production department. The paper will adopt the CCR model to analyze the DMU decision-making because the input indexes and output indexes are not many.

Each decision unit DMUj has a corresponding efficiency evaluation index:

$$h_j = \frac{u^T y_i}{v^T x_j} = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{mn} v_i x_{ij}}, j = 1,2,\ldots \quad (1)$$

If the efficiency index of j0 decision unit is taken as the goal and the efficiency index of all the decision units is taken as the constraint, the following CCR model is constructed.

$$\begin{cases} \max \quad h_{j_o} = \frac{\sum_{r=1}^{s} u_r y_{rj_o}}{\sum_{i=1}^{m} v_i x_{ij_o}} \\ s.t. \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, j = 1,2,.. \\ u \ge 0, v \ge 0 \end{cases} \quad (2)$$

Above planning model is a fractional programming and we can change it by Charnes-Cooper.

$$t = \frac{1}{v^T x_0}, w = tv, \mu = \quad (3)$$

Above model can be transformed into a linear programming model P as follows:

$$(P)\begin{cases} \max \ h_{j_o} = \mu^T y_0 \\ s.t. \ w^T x_j - \mu^T y_j \ge 0, j = 1,2,\ldots \\ w^T x_0 = 1, w \ge 0, \mu \ge 0 \end{cases} \quad (4)$$

The effectiveness of decision unit j0 is defined by using the optimal solution of linear programming. It can be seen from the model that the effectiveness of decision unit j0 is relative to all other decision making units.

The CCR model can be used to express the P planning, and the linear programming is an important effective theory of duality. It is easier to analyze the theory and economy from the dual model.

The duality of planning P is the planning D/:

$$(\quad D/)\begin{cases} \min\theta \\ s.t. \ \sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{i0}, i = 1,2,\ldots \\ \sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{r0}, r = 1,2,\ldots,p \\ \lambda_j \ge 0, j = 1,2,\ldots,n \\ \theta \text{ is Unconstrained} \end{cases} \quad (5)$$

The CCR model can be changed into the following equality constraints by changing the Charnes-Cooper and using the duality theory.

$$(D)\begin{cases} \min\theta \\ s.t. \sum_{j=1}^{n} \lambda_j x_j + s^+ = \epsilon \\ \sum_{j=1}^{n} \lambda_j y_j - s^- = \theta y \\ \lambda_j \ge 0, j = 1,2,\ldots,n \\ s^+ \ge 0, s^- \ge 0 \end{cases} \quad (6)$$

The stands for the DMU RMS(input relative to the output of effective utilization degree); stands for the input collection; stands for the output collection; stands for the combination ratio of the DMUi; $s^+$and stand for slack variable.

Assuming the DEA model there exists an optimal solution and the optimal solution is θ^*, s^{*+},, then the following results will appear:

1）If $\theta^* = 1$ and $s^{*+} = 0$ and $s^{*-} =$, then j0 is DEA efficiency and at the same time the technical and scale are efficiency.

2） if θ^* = and at least an input or output greater than zero, then j0 is weak DEA efficient and the economic activity of decision making unit is not best for the technical and scale efficiency at the same time.

3） if θ^* < then j0 is noneffective state and the efficiency of technical and scale are poor.

We can also use the of CCR model to determine the scale returns of DMU:

If exists to make $\sum \lambda_j^*$ =, then the DMU is the scale returns variable.

1） If doesn't exists to make $\sum \lambda_j^*$ = and $\sum \lambda_j^*$ <, then the DMU is the increasing scale returns.

2） If doesn't exists to make $\sum \lambda_j^*$ = and $\sum \lambda_j^*$ >, then the DMU is the diminishing scale returns.

3.2. Regional innovation capability evaluation index system

The performance evaluation system of regional innovation ability has a dynamic character and the evaluation index system is affected by many aspects, therefore, the choice of indicators should not only meet the definition of regional innovation performance, but also to ensure the scientificity, rationality and objectivity of index data, so as to ensure that the evaluation result has practical significance. The paper follows the principle being comprehensive, hierarchical and operational when the performance evaluation index system of regional innovation is constructed. On the basis of the principle of regional innovation and relevant statistical data in Hebei province, the input and output indexes are selected. The input indexes include R&D personnel number and R&D expenditure. R&D personnel number indicator reflects the human input of scientific and technological activities in different areas. R&D expenditure reflects the capital investment of science and technology activities in different areas. The output indicators include patent filings and patent grant. The patent filings reflect the overall situation of the patent technology application. The patent grant reflects the number of authorized patent invention area and utility model patent and appearance design patent.

Table 1. The performance evaluation index system of regional innovation

Variable	entry	index
input and output indexes	A1	R&D personnel number
	A2	R&D expenditure
	B1	patent filings
	B2	patent grant

4. Empirical Analysis

4.1. The regional innovation efficiency analysis

This paper takes 11 cities in Hebei Province as the research sample, and the sample selection is based on the scientific and technological data of 2012-2014.Data is from the original data of the Hebei Province Statistical Yearbook (2013, 2014 and 2015) and the Hebei Province Science and Technology Statistical Yearbook (2013, 2014 and 2015). Using DEAP2.1 analysis software, the original data of the regional innovation input and output of cities in Hebei Province in 2012-2014 years is respectively imported into the software to calculate the city's DEA efficiency values. The results are in the following tables.

Table 2. Original data in Hebei Province in 2012

City	R&D personnel number(A1)	R&D expenditure(A2)	patent filings(B1)	patent grant(B2)
ShijiazhuangCity	4.08	63.8	4962	3447
ChengdeCity	1.64	6.4	387	282
ZhangjiakouCity	1.05	6.1	459	259
QinhuangdaoCity	2.21	15	2149	1199
TangshanCity	5.38	58.2	2976	1798
LangfangCity	1.61	8.4	2086	1473
BaodingCity	3.54	41.2	4034	2714
CangzhouCity	1.87	7.7	1611	1049
HengshuiCity	3.65	4.1	1186	956
XingtaiCity	2	9.7	1273	976
HandanCity	4.3	24.5	2118	1162

Table 3. Original data in Hebei Province in 2013

City	R&D personnel number	R&D expenditure	patent filings	patent grant
ShijiazhuangCity	6.7	72.9	5996	3799
ChengdeCity	2.6	6.8	415	299
ZhangjiakouCity	2.7	8.2	689	420
QinhuangdaoCity	2.7	12.2	2257	1387
TangshanCity	5.2	67.6	3419	2398
LangfangCity	3.1	10.9	2250	1594
BaodingCity	5.5	50.9	4837	3267
CangzhouCity	4.3	10.9	1780	1312
HengshuiCity	2.7	6.5	1408	1084
XingtaiCity	3.8	11.2	1967	1146
HandanCity	4.7	20.4	2601	1480

Table 4. Original data in Hebei Province in 2014

City	R&D personnel number	R&D expenditure	patent filings	patent grant
ShijiazhuangCity	7.1	83.4	6373	4433
ChengdeCity	2.8	8	474	300
ZhangjiakouCity	2.8	8.8	1058	525
QinhuangdaoCity	2.9	13.4	2298	1310
TangshanCity	5.6	72.4	3794	2636
LangfangCity	3.1	14.3	2965	2156
BaodingCity	5.7	52.3	4723	3385
CangzhouCity	4.5	12.6	2307	1545
HengshuiCity	2.9	8.3	1392	1061
XingtaiCity	4.2	12.4	1975	1283
HandanCity	4.9	27.2	2641	1498

Table 5. DEA efficiency values of regional innovation in various regions in 2012-2014

	2012			2013			2014			3-year average of municipalities		
City	TE	PTE	SE	TE	PTE	SE	TE	PTE	SE	ATE	APTE	ASE
ShijiazhuangCity	0.939	1.000	0.939	1.000	1.000	1.000	0.938	1.000	0.938	0.959	1.000	0.959
ChengdeCity	0.244	0.902	0.270	0.285	1.000	0.285	0.286	1.000	0.286	0.272	0.967	0.280
ZhangjiakouCity	0.337	1.000	0.3370	0.400	0.968	0.413	0.580	1.000	0.580	0.439	0.989	0.443
QinhuangdaoCity	0.751	0.753	0.997	1.000	1.000	1.000	0.828	1.000	0.828	0.860	0.918	0.942
TangshanCity	0.427	0.441	0.967	0.776	0.809	0.960	0.708	0.727	0.974	0.637	0.659	0.967
LangfangCity	0.995	1.000	1.000	0.975	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BaodingCity	0.860	1.000	0.840	1.000	1.000	1.000	0.866	0.967	0.858	0.945	0.996	0.919
CangzhouCity	0.832	0.949	0.877	0.760	0.777	0.978	0.883	0.936	0.944	0.825	0.887	0.933
HengshuiCity	1.000	1.000	1.000	1.000	1.000	1.000	0.848	1.000	0.848	0.949	1.000	0.949
XingtaiCity	0.570	0.751	0.759	0.834	0.841	0.991	0.768	0.849	0.905	0.724	0.814	0.885
HandanCity	0.380	0.381	0.998	0.678	0.800	0.847	0.564	0.613	0.920	0.541	0.598	0.922
Annual average	0.669	0.834	0.820	0.794	0.927	0.861	0.752	0.919	0.827	0.738	0.893	0.836

Note: TE indicated the comprehensive efficiency, PTE indicated the pure technical efficiency, SE indicated scale efficiency, ATE indicated the average comprehensive efficiency, APTE indicated the average pure technical efficiency, ASE indicated the average scale efficiency

As is indicated in Table 5, we can see the comprehensive efficiency of regional innovation in recent 3 years in the cities of Hebei province is relatively low, the efficiency of Langfang is 1, in front of DEA value effectively, and the comprehensive efficiency of Zhangjiakou City, Tangshan City, Xingtai city and Handan city is lower than the average value, the DEA value is invalid, the comprehensive efficiency valueof the city of Chengde a value of 0.286, is the lowest among all cities. From the analysis of pure technical efficiency and scale efficiency of nearly 3 years in Shijiazhuang, Langfang and Hengshui City, the pure technical efficiency andscale efficiency value are high, indicating that these regional innovation resources allocation relative to other regions reached the optimal investment, innovation resources to improving the best combination, also achieved maximum output effect. From the view of the annual mean value of this line, Hebei Province this year is relatively low, the lowest value reaching 0.669 (2012), the highest is 0.794 (2013), the value of pure technical efficiency is slightly greater than the value of the scale efficiency.

Thus regional innovation efficiency in Hebei province is lower, but the pure technical efficiency value is slightly larger than that of the scale efficiency. Therefore, the

advancement of the regional innovation capacity must strengthen technical innovation and improve the efficiency of management. As is known much scale efficiency is close to the production frontier and have no chance to expand its scale. Thus, only pure technical efficiency can improve overall innovation in Hebei province.

4.2. Scale return analysis

The scale returns are divided into three types: scale returns constant, scale returns increasing and scale returns decreasing. The three stages are the technological innovation activities. Table 6 shows the scale returns of every city in Hebei Province in the past 3 years, which DRS, IRS, - , represent the scale of diminishing returns, increasing and constant respectively.

Table 6. 2012-2014 annual return of cities in Hebei Province

Region	2012	2013	2014
ShiJiazhuang City	drs	-	drs
Chengde City	irs	Irs	irs
Zhang Jiakou City	irs	Irs	irs
QinHuangdao City	drs	-	irs
Tangshan City	drs	Irs	drs
Langfang City	-	-	-
Baoding City	drs	-	drs
Cangzhou City	irs	Drs	irs
Hengshui City	-	-	irs
Xingtai City	irs	Drs	irs
Handan City	drs	Drs	irs

According to the calculation results in table 6, in 2012, only constant return in Lang fang and Heng shui city scale, scale of Cheng de, Zhang Jiakou, Cang zhou and Xing tai belong to increasing, Shi Jiazhuang, Qin Huangdao and other five cities belong to the diminishing, overall performance for the scale trend of diminishing returns. In 2013the cities with the constant returns four, those with diminishing returns are 3 and those with the increasing return has the same number, thus overall scale reward in Hebei province remains stable; Significant changes have taken place in 2014, of which seven cities exhibit the diminishing returns to scale. Among them, Langfang keeps the constant position in the scale. Generally speaking, this belongs to the increasing tendency of return of scale.

5. Conclusions

Data Envelopment Analysis (DEA) method is a kind of evaluation method of non-parameter system to solve, the problem of multi-input and output index. Using the DEA analysis method, we comparatively analyzed the performance of regional innovation capability in Hebei Province in recent 3 years. In general, the regional innovation of Hebei province is still in the increasing scale returns. As a conclusion, Hebei province must carry on effective innovation, constantly increasing the input of innovation resources. It is the inevitable choice to improve the efficiency of innovation resources. As

for the input of innovation resources, high-end technical personnel and adequate funding need not only to be provided, but also create a good environment for innovation. In particular, innovation ability of enterprises need to be strengthened through the guiding of the policy tilt for the innovation of the enterprise. Due to the fact that enterprise is the main body of regional innovation, only through the promotion of innovation ability will improve the regional innovation capability.

Acknowledgments

This study was financially supported by Foundation Item: Item from Technology Department of Hebei Province in 2016 “Research on construction and performance evaluation of resource-based enterprise technology innovation system under the regional innovation system in HebeiProvince” (Project number: 16454901D).

References

[1] J. Bai, K. Jiang and J. Li, “Research on the Innovating Efficiency of Regional Innovation Systems of China”, Management Renew, vol. 9, **(2009)**, pp. 1-7.
[2] H. Zhou, L. Song and W. Liu, “A Comparative Study of the Performance Evaluation of Regional Innovation System”, Journal of Industrial Technological Economics, vol. 9, **(2014)**, pp. 22-33.
[3] S. Liu and J. Guan, “The Evaluation on the Innovating Performance of Regional Innovation Systems”, Chinese Journal of Management Science, vol. 01, **(2002)**, pp. 75-78.
[4] J. Guan and S. Liu, “Research Framework and Content of Regional Innovation System”, China Science and Technology Forum, vol. 3, **(2003)**, pp. 24-26.
[5] X. Yuan, W. Liu and W. Sun, “Review of Domestic Research on the Regional Innovation Performance”, Scientific Management Research, vol. 8, **(2011)**, pp. 37-41.
[6] M. Liu and G. Li, “Research on Innovation Performance Evaluation of Regional Innovation System in China”, Modern Management Science, vol. 3, **(2011)**, pp. 72-74.
[7] W. Chen, Z. Feng, H. Jiang and X. Kang, “Evaluation Research on the Innovation Efficiency of Regional Innovation Systems of China - A New View Based on Chain Relational Network DEA Model”, Journal of Intelligence, vol. 12, **(2010)**, pp. 24-29.
[8] H. Zhao and W. Zhao, “Research on Performance of Regional Innovation Ability in Southwestern China”, Techno economics and Management Research, vol. 6, **(2009)**, pp. 37-39.
[9] F. Yin, Y. Shan, C. Su and C. Yan, “Construction of Regional Innovation Performance Evaluation Model in Hebei Province”, Statistics and Decision, vol. 14, **(2011)**, pp. 77-79.
[10] K. Rennings, “Redefining Innovation-Eco-Innovation Research and the Contribution from Ecological Economics”, Ecological Economics, vol. 32, **(2000)**, pp. 319-332.
[11] S. J. M. Harkema and W. Baets, “Customerized Innovation through the Emergence of a Mutually Adaptive and Learning Environment”, Europer Journal of Economic and Systems, vol. 1, **(2001)**, pp. 111-129.
[12] L. Fleming and O. Sorenson, “Technology as a complex adaptive system: evidence from patent data”, Research Policy, vol. 30, **(2001)**, pp. 1019-1039.
[13] R. Rutten and F. Boekema, “Regional social capital: Embeddedness, innovation networks and regional economic development”, Technological Forecasting and Social Change, vol. 74, no. 9, **(2007)**, pp. 1834-1846.

