SPATIAL DIFFERENTIATION OF CHINA'S INDUSTRIAL ENTERPRISE R&D EFFICIENCY

Chengyun Wang, Yue Wang, Nana Li and Tengfei Ma

With 3 figures and 5 tables
Received 26 October 2018 · Accepted 22 August 2019

Summary: Taking China's industrial enterprises above a designated size at provincial level as study objects, this paper explores enterprise R&D efficiency in different regions of China's 31 mainland provinces over time with Data Envelopment Analysis (DEA). It further discusses the factors influencing R&D efficiency. This paper explains R&D efficiency from the three perspectives of comprehensive efficiency, pure technical efficiency and scale efficiency. It is shown that although R&D efficiency displays rising trend, scale efficiency performs better than comprehensive and pure technical efficiency in China. The hot spots and cold spots of R&D efficiency change substantially in spatial distribution over time. Hot spots of comprehensive and pure technical efficiency migrated from Western China to coastal regions, demonstrating strong spatial dependence, whereas scale efficiency did not show a similar spatial distribution. Factors leading to variation in R&D efficiency across regions were studied. The significant factors for comprehensive, pure technical and scale efficiency were found to be different. These findings may be beneficial in planning regional development and effectively encouraging innovation at the regional level.

Zusammenfassung: Die vorliegende Studie untersucht die räumlichen Muster und zeitliche Variabilität regionaler F&E-Effizienz (Forschungs- und Entwicklungseffizienz) größerer chinesischer Industrieunternehmen in den 31 chinesischen Festland-Provinzen mit Hilfe einer Data Envelopment Analysis (DEA). Diskutiert werden zudem die Faktoren, welche die Effizienz von Forschung und Entwicklung beeinflussen. Beleuchtet wird die F&E-Effizienz aus den drei Perspektiven der umfassenden Effizienz, der reinen technischen Effizienz und der Skaleneffizienz. Es wird aufgezeigt, dass die Effizienz von Forschung und Entwicklung zwar einen steigenden Trend aufweist, aber die Skaleneffizienz in China besser abschneidet als die umfassende und rein technische Effizienz. Die räumlichen Muster der Hot Spots und Cold Spots der F&E-Effizienz verändern sich im Laufe der Zeit erheblich. Hot Spots mit umfassender und rein technischer Effizienz wanderten von Westchina in Küstenregionen und zeigten eine starke räumliche Abhängigkeit, während die Skaleneffizienz keine ähnlichen räumlichen Muster aufzeigte. Darüber hinaus werden die Faktoren diskutiert, die zu Unterschieden in der F&E-Effizienz zwischen den Regionen führen. Letztlich sind die wesentlichen Faktoren für eine umfassende, reine technische und skalare Effizienz differenziert zu bewerten. Diese Ergebnisse können bei der Planung der regionalen Entwicklung und der wirksamen Förderung von Innovationen auf regionaler Ebene von Nutzen sein.

Keywords: R&D efficiency, spatial differentiation, influencing factors, China's industrial enterprises

1 Introduction

Technology innovation is a key factor of international competitiveness and promotion of economic growth. As a core driver of technological innovation, research and development (R&D) activity has a direct impact on a nation's level of innovation. Therefore, many countries attempt to strengthen innovation outcomes by increasing expenditure on R&D. However, there is not a simple linear relation between R&D activity and outcomes, and many authors take R&D efficiency as the key indicator of a company or region's capability to innovate. China's spending on R&D has accelerated in recent years, reaching 1.97 trillion yuan (about $286 billion) in 2018, ranking second in the world. This paper focus on discussing the R&D efficiency of China's industrial enterprises.

R&D efficiency refers to the conversion efficiency between input and output of all factors in the R&D process, reflecting the contribution of R&D input to R&D output or the allocation efficiency of R&D resources. Due to different levels of regional development, there are also different levels of R&D performance across these regions. The R&D capabilities of enterprises show significant variation across regions. Lee and Park (2005) classifies 27 countries into four clusters based on their output-specialized R&D efficiency: inventors, merchandisers, academicians and duds, showing for example that Singapore ranks high in total efficiency while Japan demonstrates patent-oriented efficiency. Meanwhile, mainland China, South Korea, and Chinese Taiwan are found to be relatively inefficient in R&D.

https://doi.org/10.3112/erdkunde.2019.03.04 ISSN 0014-0015 http://www.erdkunde.uni-bonn.de

Many theses on R&D efficiency in China have been presented (Zhang 2008; Zhang 2010; Wang and Xiong 2010; Feng et al. 2006; Yan and Feng 2005). Wang (2009) measures both static and dynamic R&D efficiency in mainland China, and finds that production efficiency and technical efficiency are positively associated with regional economic prosperity, but allocation efficiency of resources shows signs of a negative correlation with economic level. In China, scale efficiency causes the relatively low R&D efficiency of large and medium-sized industrial enterprises, and economic level cannot determine the R&D efficiency (Wang et al. 2008). Similarly, Zhang (2008) analyzes R&D production efficiency at provincial level and finds that R&D production efficiency is significantly higher in Eastern China compared to Middle and Western China. Li (2013) finds that R&D efficiency in China displays a long-term upward trend, but the spatial distribution is uneven. R&D efficiency is influenced by R&D capital investment, and is positively related to enterprise scale and FDI, but does not exhibit significant correlation with ownership structure and enterprise performance (Li 2013; Liu and Yang 2012). However, Zhang et al. (2003) argues that ownership is found to be a contributing factor in the cross-sectional variance of both R&D and productive efficiencies, and the state sector has significantly lower R&D and productive efficiency than the private sector. Moreover, financial support from financial agencies has a negative impact on regional R&D efficiency, and the impact of government-sponsored capital on regional R&D efficiency hasn't been shown to be significant (Liu and Yang 2012). Yan and Feng (2005) concluded that R&D efficiency of China's enterprises has a positive relationship with R&D input, but negative correlation with market competition. Conversely, Tang et al. (2009) showed that high R&D input does not lead to high R&D efficiency.

Measurement of R&D efficiency is not trivial, with different models and methods giving different results. Data Envelopment Analysis (DEA) has proven to be the most effective method in measuring efficiency (Chi and Huang 2009). For example, Sharma and Thomas (2008) examined the relative efficiency of R&D across a group of 22 developed and developing countries with DEA. Many existing studies on China's firms also use the DEA model to analyze R&D efficiency (Wu 2005; Chi et al. 2004; Liang et al. 2006; Yu 2006).

From the existing studies, we can see that it is rather difficult to measure R&D efficiency effectively from a single point. It is necessary to examine R&D efficiency from different perspectives. Although there are many studies that have analyzed R&D efficiency in China at a regional level and industrial level, they tend to be static and do not reveal changes over time. Taking China's industrial enterprises above designated size as study object, this paper aims to explore their comprehensive efficiency, pure technical efficiency and scale efficiency separately with the DEA model, and explain their spatial evolution through a hot and cold spot analysis. It addresses the following questions: what is the R&D efficiency performance of China's industrial enterprises under the DEA model? What distribution changes occur over time? What factors influence the performance of R&D efficiency?

The paper is organized as follows. Section 2 introduces the data source. Section 3 explains the DEA model and makes empirical analysis on R&D efficiency by region. Section 4 discusses the spatial evolution of R&D efficiency hot and cold spots. Section 5 explores the correlation between R&D efficiency and selected influencing factors. Section 6 presents concluding remarks.

2 Data sources

This paper takes China's industrial enterprises above a designated size[1)] on provincial level as study object. All data on industrial enterprises in this paper is from the *China Statistical Yearbook (CSY)* and the National Bureau of Statistics website (http://www.stats.gov.cn/). With consideration to the availability and fullness of the data, selecting statistical data across 31 provinces, autonomous regions and municipalities of mainland China in 2008, 2009, 2011 and 2014, this paper analyses R&D efficiency performance and spatial changes. Interpolation is used to fill in missing data.

3 Empirical analysis of R&D efficiency

3.1 Construction of the evaluation index

In this study, an evaluation index of R&D efficiency for industrial enterprises above a designated size is built with respect to inputs and outputs. Referring to the existing studies (Zhao and Zeng 2009; Hu 2009; Yue 2008; Cao et al. 2016), this paper selects number of patents issued, revenue of new products, and the ratio of new product revenue to revenue from principal business as the measurement indicators of output R&D efficiency of enterprises, and uses R&D expend-

[1)] Since 2011, the annual main business revenue of the industrial enterprises above a designated size has been adjusted from 5 million RMB and above to 20 million RMB and above.

iture, total R&D personnel, and number of new product programs to evaluate R&D input. Usually R&D output lags input, and the lag is regarded to occur after 1 year at the latest through cointegration analysis (Shen 2009). Thus R&D output data in year *t* is associated with R&D input in year *t-1*.

3.2 DEA model

The DEA model was first put forward by Charnes, Cooper and Rhodes in 1978, and is a linear programming method for comparing the relative productivity or efficiency of multiple service units (Yang et al. 2013). DEA has become the most commonly used methodology for measuring the R&D efficiency of nations, regions and firms. Based on and with reference to the previous method of Fang and Guan (2011), this paper measures R&D efficiency of China's industrial enterprises above a designated size. When evaluating the R&D efficiency of a group of *m* enterprises, we assume that there are *K* input indicators and *L* output indicators. Supposing that there are *M* enterprises to be evaluated and each enterprise has *K* inputs and *L* outputs. Let x_{mk}, (x_{mk}>0) be the matrix of input variables, represent the *k*th input (*k*=1,2,…, K) of the enterprises in the *m*th (*m*=1, 2, 3, …, M) province or region; y_{ml} (y_{ml}>0) the matrix of output variables, represent the *l*th output in the enterprises of *m*th province or region. For the enterprises in the *m*th province or region, θ ($0<\theta \leq 1$) represents the comprehensive index of elements and resource input-output efficiency, called the comprehensive efficiency index; ε denotes non-infinitesimal; λ_m ($\lambda_m \geq 0$) expresses weight variables, indicating the returns to scale of R&D activities; s^- ($s^- \geq 0$) is the slack variable and relates to the further reduction in input to reach the efficient frontier; s^+($s^+ \geq 0$) is the surplus variable and represents the increases in output to gain the best efficiency. The DEA model for the *m*th enterprise can be defined with the following formula (Charnes et al. 1978):

$$\begin{cases} \min\left(\theta-\varepsilon\left(\sum_{k=1}^{K} s^- + \sum_{l=1} s^+\right)\right) \\ \text{s.t.} \sum_{m=1}^{M} x_{mk}\lambda_m + s^- = \theta x_k^m \quad k=1,2,...,K \\ \sum_{m=1}^{M} y_{ml}\lambda_m - s^+ = y_l^m \quad l=1,2,...,L \\ \lambda_m \geq 0 \quad m=1,2,...,M \end{cases}$$

The above formula provides for constant returns to scale, i.e. a CRS model. By adding the constraint $\sum^{M}_{m=1}\lambda_m = 1$, the above formula becomes a variable return to scale (VRS) model. The VRS model expresses the pure technical efficiency under VRS circumstances. According to the VRS model, comprehensive technical efficiency can be decomposed into pure technical and scale efficiency, defined as $\theta_m = \theta_{TE} \times \theta_{SE}$. Where, θ_m represents the overall measurement of technical efficiency for the *m*th unit; θ_{TE} ($0< \theta_{TE} \leq 1$, $\theta_{TE} \geq \theta_m$) refers to the pure technical efficiency for the corresponding industrial enterprise above a designated size; θ_{SE} ($0< \theta_{SE} \leq 1$, $\theta_{SE} \geq \theta_m$) measures the scale efficiency. The closer θ_{TE} or θ_{SE} is to 1, the better the *m*th region performs in pure technical efficiency or scale efficiency. When θ_{TE} or θ_{SE} equals 1, it indicates that the efficiency in the *m*th region is optimal.

In this paper, comprehensive efficiency refers to the allocation of R&D resources and the utilization efficiency of those resources. Pure technical efficiency measures the production efficiency brought by technological progress, and scale efficiency is used to evaluate the gap between the current scale and the optimal scale of R&D resources.

3.3 Performance of R&D efficiency by regions

According to the DEA model, R&D efficiency by region is evaluated and the result is shown in Tab. 1. The comprehensive R&D efficiency of China's industrial enterprises above a designated size rises from a low base, with values of 0.530 in 2008, 0.505 in 2009, 0.573 in 2011, and 0.704 in 2014. In 2008, comprehensive efficiency in Jilin, Hunan, Hainan and Tibet is 1.000, the optimal level. Comprehensive efficiency is less than 0.5000 in 19 provinces, accounting for 61.3% of all selected regions; only 5 provinces (16.1% of all selected regions) exceed 0.8000. In 2009, the four regions Guangdong, Hunan, Hainan and Tibet reach optimal efficiency. In 2011, the comprehensive efficiency of only two regions (Jilin and Tibet) is at the optimal level. In 2014, there are 5 provinces (16.1% of all selected regions) with comprehensive efficiency below 0.5000, while the comprehensive efficiency in 11 provinces (35.5% of all selected regions) exceeds 0.8000. Although the average comprehensive efficiency increases monotonically, it only accounts for less than 75% of the optimal efficiency in 2014.

Pure technical efficiency is higher than comprehensive efficiency and shows a continued upward trend. Pure technical efficiency was 0.657 in 2008, 0.581 in 2009, 0.764 in 2011 and 0.799 in 2014, i.e. ex-

Tab. 1: R&D efficiency of China's industrial enterprises over time

Year	2008			2009			2011			2014		
Region	CE	PTE	SE	CE	PTE	SE	CE	PTE	SE	CE	PTE	SE
Beijing	0.560	0.823	0.681	0.685	0.762	0.899	0.565	1.000	0.565	0.885	1.000	0.885
Tianjin	0.469	0.657	0.715	0.643	0.719	0.894	0.775	0.997	0.778	0.794	0.888	0.894
Hebei	0.323	0.323	1.000	0.288	0.295	0.976	0.478	0.557	0.858	0.575	0.576	0.999
Shanxi	0.355	0.355	1.000	0.276	0.426	0.648	0.556	0.644	0.863	0.563	0.666	0.846
Inner Mongolia	0.421	0.421	1.000	0.233	0.303	0.768	0.548	0.565	0.971	0.594	0.607	0.980
Liaoning	0.445	0.506	0.879	0.352	0.369	0.955	0.558	0.703	0.794	0.580	0.593	0.979
Jilin	1.000	1.000	1.000	0.677	0.809	0.836	1.000	1.000	1.000	0.964	0.964	1.000
Heilongjiang	0.348	0.348	1.000	0.294	0.300	0.981	0.187	0.293	0.639	0.250	0.316	0.792
Shanghai	0.775	1.000	0.775	0.743	0.805	0.923	0.781	1.000	0.781	0.952	1.000	0.952
Jiangsu	0.436	1.000	0.436	0.448	0.842	0.532	0.610	1.000	0.610	0.716	1.000	0.716
Zhejiang	0.238	1.000	0.238	0.456	0.555	0.822	0.554	1.000	0.554	0.908	1.000	0.908
Anhui	0.515	0.567	0.909	0.496	0.577	0.860	0.565	0.879	0.643	0.834	1.000	0.834
Fujian	0.406	0.406	1.000	0.359	0.384	0.934	0.608	0.735	0.827	0.478	0.507	0.942
Jiangxi	0.209	0.209	1.000	0.199	0.212	0.939	0.547	0.547	0.999	0.738	0.738	1.000
Shandong	0.452	1.000	0.452	0.331	0.392	0.843	0.683	1.000	0.683	0.657	0.974	0.675
Henan	0.390	0.408	0.955	0.329	0.341	0.964	0.414	0.492	0.841	0.713	0.713	1.000
Hubei	0.478	0.503	0.952	0.442	0.465	0.951	0.539	0.673	0.801	0.721	0.734	0.983
Hunan	1.000	1.000	1.000	1.000	1.000	1.000	0.793	1.000	0.793	1.000	1.000	1.000
Guangdong	0.679	1.000	0.679	1.000	1.000	1.000	0.697	1.000	0.697	0.883	1.000	0.883
Guangxi	0.452	0.489	0.924	0.326	0.381	0.857	0.509	0.544	0.935	0.772	0.779	0.991
Hainan	1.000	1.000	1.000	1.000	1.000	1.000	0.711	1.000	0.711	0.822	1.000	0.822
Chongqing	0.508	0.951	0.534	0.445	0.666	0.668	0.730	1.000	0.730	1.000	1.000	1.000
Sichuan	0.433	0.538	0.805	0.482	0.599	0.804	0.518	0.784	0.660	0.645	0.792	0.815
Guizhou	0.464	0.542	0.856	0.505	0.582	0.869	0.370	0.614	0.602	0.483	0.767	0.630
Yunnan	0.572	0.637	0.897	0.698	0.750	0.931	0.414	0.657	0.630	0.653	0.822	0.794
Tibet	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Shanxi	0.332	0.410	0.808	0.343	0.421	0.816	0.334	0.646	0.517	0.305	0.382	0.799
Gansu	0.448	0.469	0.955	0.397	0.435	0.913	0.617	0.751	0.821	0.511	0.621	0.822
Qinghai	0.816	0.816	1.000	0.513	0.664	0.771	0.243	0.542	0.448	0.417	0.526	0.793
Ningxia	0.263	0.320	0.820	0.268	0.504	0.531	0.391	0.534	0.732	0.586	0.968	0.605
Xinjiang	0.640	0.676	0.947	0.418	0.448	0.935	0.470	0.537	0.877	0.816	0.849	0.962
Average	0.530	0.657	0.846	0.505	0.581	0.865	0.573	0.764	0.754	0.704	0.799	0.881

Note: CE means comprehensive efficiency; PTE means pure technical efficiency; SE means scale efficiency.
Data source: CSY 2008, 2009, 2011 and 2014.

ceeded 0.6500 except in 2009. The number of regions with the optimal pure technical efficiency was 9 in 2008 and 10 in 2014. Despite remaining steady, the spatial distribution at provincial level changed considerably. The number of regions whose pure technical efficiency exceeds 0.6000 was 15 (48.4% of provinces) in 2008 and increased to 25 in 2014.

Scale efficiency was observed to be higher than comprehensive efficiency and pure technical efficiency in any given period. It shows an upward trend in general, and its value was 0.846 in 2008, 0.865 in 2009, 0.754 in 2011 and 0.881 in 2014 respectively, remaining above 75% of the optimal level. The number of regions whose scale efficiency reached the optimal level was 11 in 2008 and declined significantly to 6 in 2014, accompanied by a significant shift in regional distribution. However average scale efficiency is much closer to the optimal level as compared to comprehensive efficiency and pure technical efficiency.

4 Spatial evolution of hot spots for R&D efficiency

Getis_OrdG_i^* is used to identify the spatial distribution of high-value clusters (hot spots) and low-value clusters (cold spots) (Anselin 1995). In order to explore the distribution of R&D efficiency and identify which regions contribute greatly to global autocorrelation, this part uses ArcGis 10.0 software to get the value of Getis_OrdG_i^* over each period, and classifies G_i^* into four groups from high value to low value according to Jenks Natural Breaks Classification Method: hot spots, sub-hot spots, sub-cold spots and cold spots (Wang et al. 2017; Liu et al. 2017; Zou et al. 2015). The spatial distribution of hot spots and cold spots is shown in Fig. 1, Fig. 2 and Fig. 3.

As shown in Fig.1, the spatial distribution of hot and cold spots displays significant changes from 2008 to 2014. In 2008, the hot spots are in Tibet, Qinghai, Xinjiang, Guangxi and Hainan, and the sub-hot spots are mainly concentrated in their adjacent areas such as Guangdong, Chongqing and Hunan. In 2009, Xinjiang and Tibet changed into sub-hot spots. In 2011, the hot spots had transferred from Western China to Eastern regions and are concentrated in Zhejiang, Shanghai, Jiangsu and other coastal regions. In 2014, the hot spots are concentrated in the middle and lower reaches of the Yangtze River, and the sub-hot spots are mainly in coastal regions. Meanwhile, the cold spots and sub-cold spots are mainly concentrated in Middle China in 2008 and 2009, and then transferred towards Middle and Western China in 2011 and 2014. The sub-cold spots are observed to be distributed next to cold spots.

The hot and cold spots of pure technical efficiency also change significantly over time (Fig. 2). The hot and sub-hot spots are mainly concentrated in Eastern coastal regions, while cold and sub-cold spots are mostly distributed in Middle and Western regions. Between 2008 and 2014, the numbers of hot spots increased and the distribution also transferred from being scattered across the Eastern region and then being concentrated in Eastern coastal regions and the middle and lower reaches of the Yangtze River. The sub-hot spots moved towards Middle and Eastern regions. The cold spots transferred from the middle regions to being concentrated in the areas connecting Middle and Western regions and sub-cold spots were distributed adjacent to cold spots.

The hot and cold spots of scale efficiency also present distinct spatial differentiation, and their distribution has changed significantly (Fig. 3). The hot and sub-hot spots are mainly distributed in Eastern coastal regions, while the cold and sub-cold spots are mainly distributed in the Middle and Western regions. Between 2008 and 2014, the number of hot spots decreased, and the distribution transferred from Western mainland regions to Guangdong and Liaoning province in the Eastern regions and Hunan, Jiangxi and Hubei province in the middle regions. It is noticeable that the scale efficiency in 2008 was higher in middle regions than coastal regions. In other years, the cold and sub-cold spots were distributed alternately and dispersedly.

5 Influencing factors of R&D efficiency

5.1 DEA model and variation selections

According to the DEA method, setting the efficiency value as the dependent variable, we perform empirical analysis on the influencing factors of R&D efficiency through constrained-regression (Tobit regression) analysis (Xie et al. 2008; Zhang 2008) using stata14 software. The limited dependent variable model is described as follows:

$$Y_i^* = \alpha X_i + \varepsilon$$

$$\begin{cases} Y_i = Y_i^* & Y_i^* > 0 \\ Y_i = 0 & Y_i^* \leq 0 \end{cases}$$

where Y_i^* denotes the potential dependent variable, Y_i is the actual dependent variable, X_i represents the vector of explanatory variables,

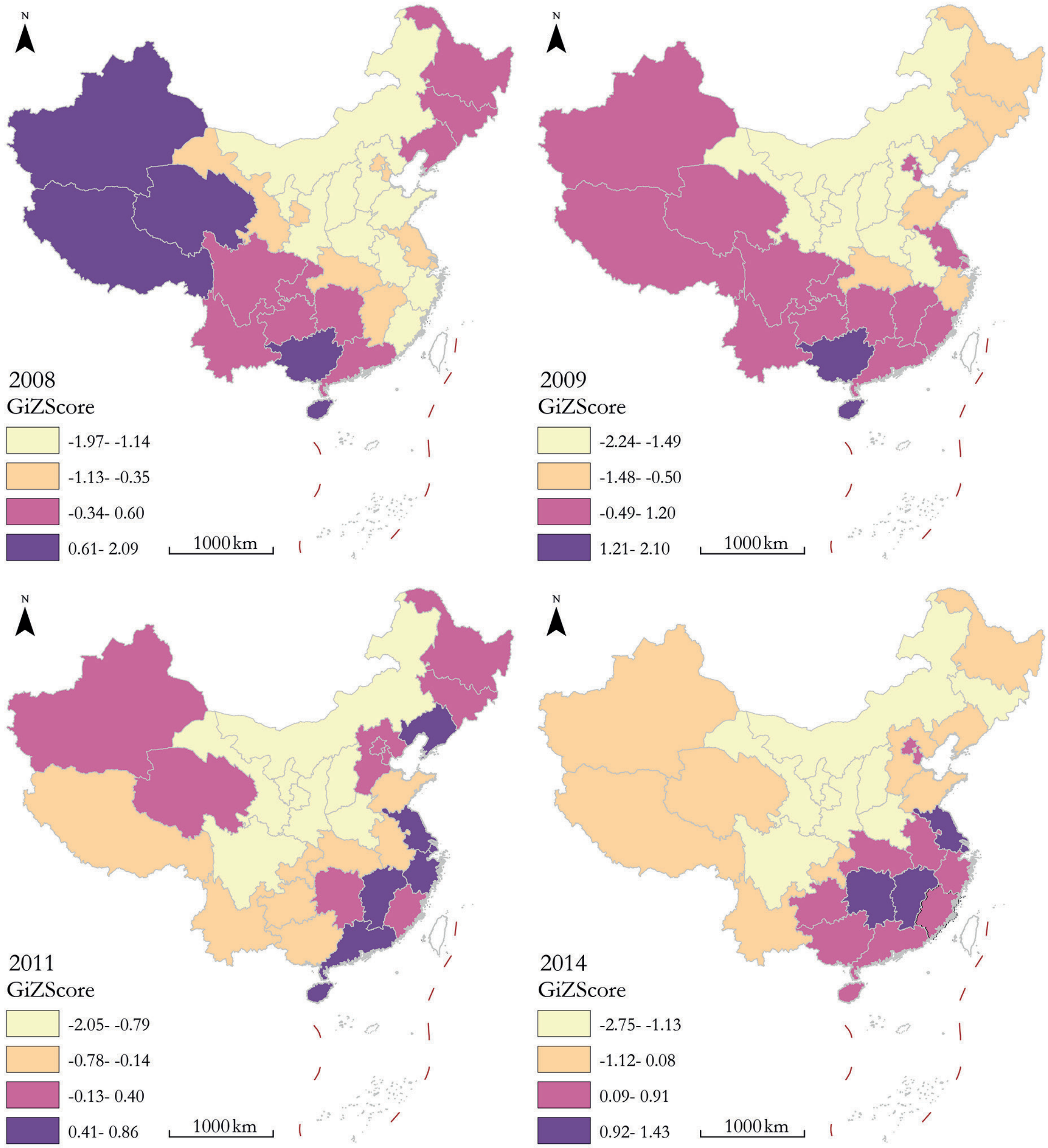

Fig. 1: Hot and cold spots' distribution of the comprehensive efficiency. Data source: CSY 2008, 2009, 2011 and 2014.

α is the coefficient of explanatory variables, and $\varepsilon_i \sim N(0,\sigma^2)$, $i = 1,2,3 \ldots$

The value of efficiency is regarded as a bounded variable and influencing factor is taken as explanatory variable, and then Tobit regression is made on the above two variables. The equation is described as the follow:

$$E_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta n X_n + \varepsilon$$

where E_i means explained variable, it represents the efficiency of the ith enterprise; β_0 is constant, β_1, $\beta_2, \beta_3, \beta_4, \beta_5$ are estimation parameters, $X_1 \ldots X_n$ means different influencing factor respectively, ε denotes random interference.

Various factors have an impact on R&D output, and these factors can also influence each other. Some are factors that enterprises can change and control consciously, such as improving a management system, changing strategy, and paying more attention to R&D activities; others are factors that enterprises cannot control, such as political and economic policy, laws and regulations. In this paper, R&D efficiency is mainly measured by quantitative analysis.

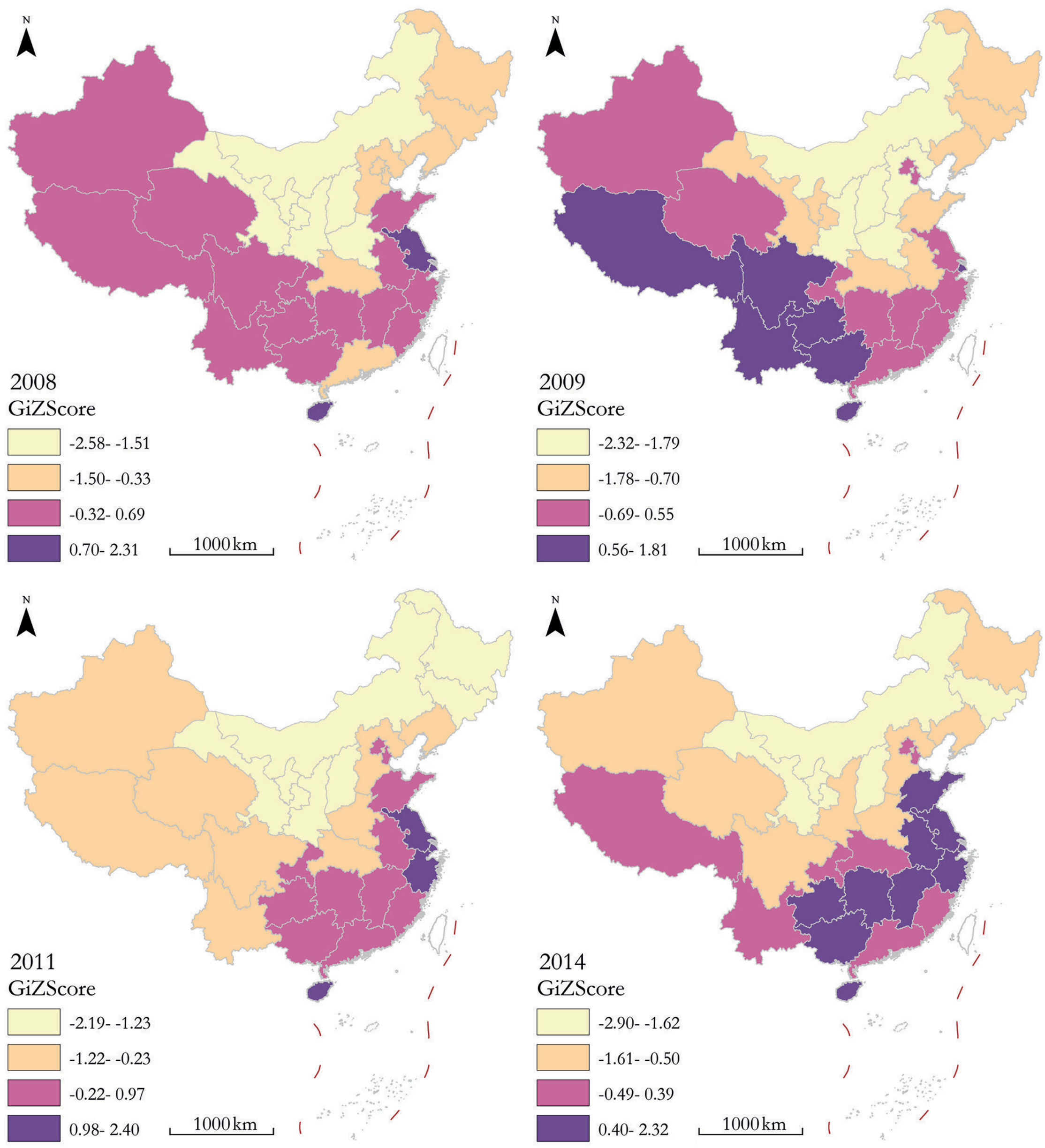

Fig. 2: Hot and cold spots' distribution of the pure technical efficiency. Data source: CSY 2008, 2009, 2011 and 2014.

Considering the availability of relevant data, this part selects seven factors and performs analysis on these factors of R&D efficiency based on the previous study results (Shi et al. 2009; Deng 2010; Liu 2008; Wei and Shen 2008). The selected factors are: R&D intensity, level of opening-up, degree of nationalization, enterprise scale, market concentration, economic level and education level (Tab. 2).

R&D intensity is defined as the ratio of R&D expenditure to main business revenue, and reflects how much investment enterprises put into R&D activities. The level of opening-up aims to measure the exposure of industry to foreign trade and investment under open market economics, as China experiences a gradual shift away from a planned economy towards open competition. In measuring the level of opening-up, some authors use foreign direct investment to reflect the degree of overseas investment involved in regional development. In this paper, the level of opening-up is defined as the ratio of total reported revenue from foreign and Hong Kong, Macao and Taiwan enter-

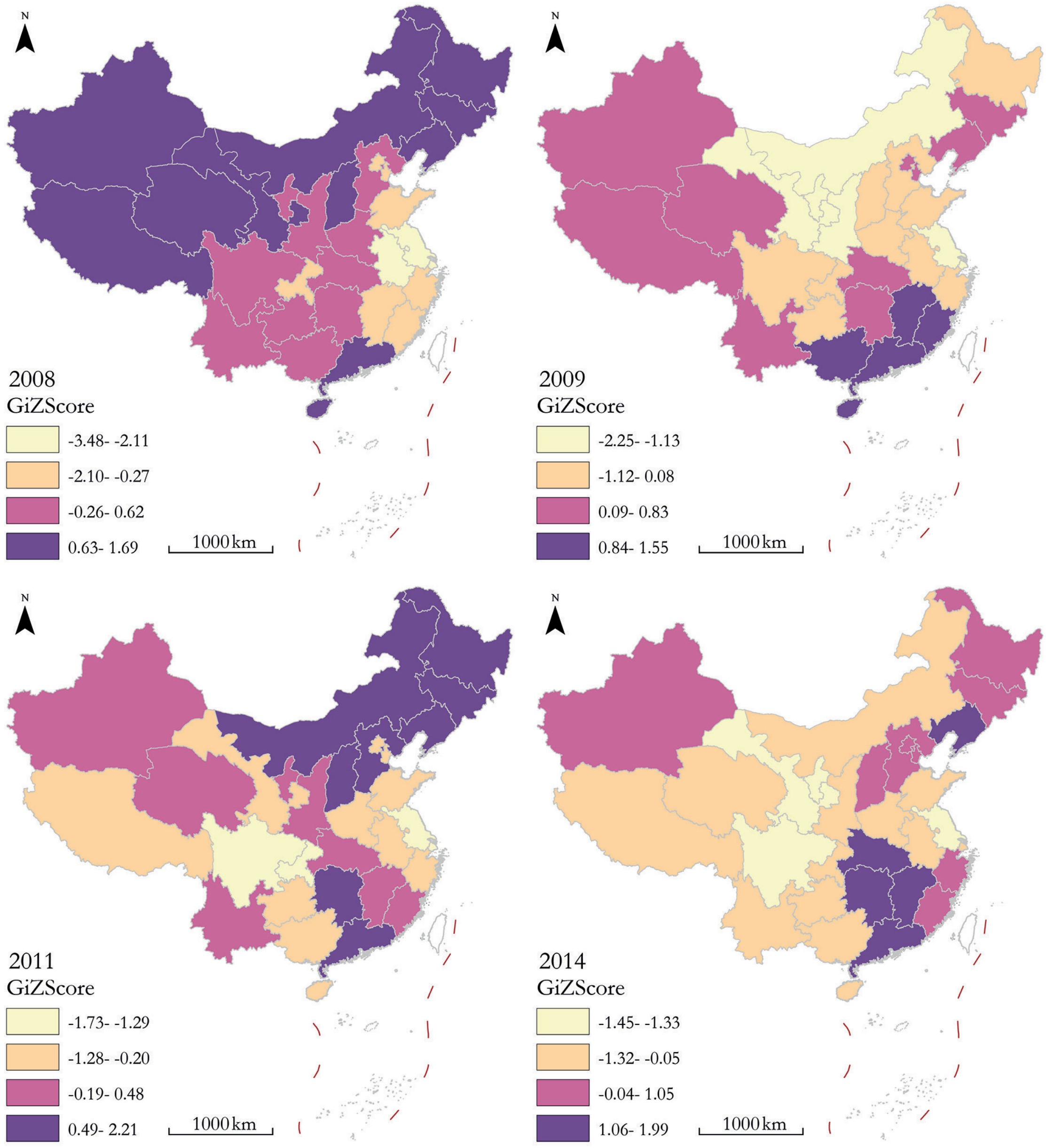

Fig. 3: Hot and cold spots' distribution of the scale efficiency. Data source: CSY 2008, 2009, 2011 and 2014.

prises to the total reported revenue for the region. The degree of nationalization indicates the balance between public and private market participation. Referring to the study of Liu (2008), degree of nationalization is described as the ratio of total reported revenue from state-owned enterprises to the total reported revenue for the region. When measuring the impact on R&D efficiency brought about by enterprise scale, added value (Ye 2006), sales revenue (Zhang 2008), and market share (Wei and Shen 2008) are selected as the indicators to measure enterprise scale in previous studies. In this paper, we use average total assets to measure enterprise scale, and define it as the ratio of total assets of enterprise in the region to the number of enterprises. Market concentration is usually used to measure the degree of market competition, and it is expressed by the number of enterprises above designated size in this paper. Economic level is indicated by GDP per capita and education level is presented by ratio of educational expenditure to fiscal expenditure.

Tab.2: Explanations of indexes

Measurement index	Explanations	Symbol
R&D intensity	Ratio of R&D expenditure to main business revenue	YF
Level of opening-up	Ratio of revenue from foreign and Hong Kong, Macao and Taiwan enterprises to the total enterprise revenue	DW
Degree of nationalization	Ratio of revenue from state-owned enterprises to the total enterprise revenue	GY
Enterprise scale	Ratio of enterprises' total assets to the numbers of enterprises	QY
Market concentration	The numbers of enterprises above designated size	SC
Economic level	GDP per capita (in units of 10,000 Chinese RMB per capita)	GDP
Education level	Ratio of educational expenditure to fiscal expenditure	JY

Tab. 3: Regression results of influencing factors for comprehensive efficiency

	Prob > chi2=0.0000			LR chi2(7) =51.33		
	Coef.	**Std. Err.**	**t**	**P>|t|**	**[95% Conf. Interval]**	
DW	.4455802	.1474076	3.02	0.003	.154992	.7361685
GY	-.4343389	.1531528	2.84	0.005	.7362529	.1324249
JY	-.504858	.5722601	-0.88	0.379	-1.632968	.6232524
GDP	.0093	.0203762	0.46	0.649	-.030868	.0494681
QY	.0317386	.0116456	2.73	0.007	.0087813	.054696
SC	-.0421912	.0160334	-2.63	0.009	-.0737983	-.0105842
YF	19.30961	9.846258	1.96	0.051	-.1005582	38.71979
_cons	.5883176	.0838091	7.02	0.000	.4231027	.7535326

Data source: CSY 2008, 2009, 2011 and 2014.

Tab. 4: Regression results of influencing factors for pure technical efficiency

	Prob > chi2=0.0000			LR chi2(7) =100.25		
	Coef.	**Std. Err.**	**t**	**P>|t|**	**[95% Conf. Interval]**	
DW	.7847694	.1858474	4.22	0.000	.4184039	1.151135
GY	-.2877947	.1750181	-1.64	0.102	-.6328122	.0572228
JY	-.0504151	.694467	-0.07	0.942	-1.419435	1.318605
GDP	-.0179665	.0236861	-0.76	0.449	-.0646594	.0287265
QY	.0669401	.0138014	4.85	0.000	.039733	.0941471
SC	.0712279	.0219517	3.24	0.001	.0279539	.1145019
YF	26.09198	11.24508	2.32	0.021	3.92428	48.25968
_cons	.3827852	.0997442	3.84	0.000	.1861571	.5794134

Data source: CSY 2008, 2009, 2011 and 2014.

5.2 Results

In the Tobit regression shown in Tab. 3, the model p-value is 0.0000 and Chi squared is 51.33, which implies that the model is effective in modelling the influence of the proposed factors of comprehensive efficiency. There are four significant factors at the $p<0.05$ level, namely the level of opening-up, degree of nationalization, enterprise scale and market concentration. Education, economic level and R&D intensity are not significant ($p>0.05$). Among the significant factors, the coefficient of opening-up and enterprise scale is positive, and the coefficient of nationalization and market concentration is negative.

As shown in Tab. 4, the model p-value is 0.0000 and Chi squared is 100.25, which implies that the model is effective in modelling the influence of the proposed factors of pure technical efficiency. There are four significant factors at the $p<0.05$ level, namely the level of opening-up, enterprise scale, market

Tab. 5: Regression results of influencing factors for scale efficiency

	Prob > chi2=0.0000			LR chi2(7) =47.69		
	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
DW	.049888	.1012339	0.49	0.623	-.1496769	.2494529
GY	-.2229899	.1039048	-2.15	0.033	-.4278199	-.0181598
JY	-.7218028	.3851661	-1.87	0.062	-1.48109	.0374847
GDP	.027769	.0138007	2.01	0.045	.0005634	.0549746
QY	-.0155245	.0078144	-1.99	0.048	-.0309291	-.0001198
SC	-.0760427	.0107996	-7.04	0.000	-.0973323	-.0547532
YF	-.886289	6.811679	-0.13	0.897	-14.31432	12.54174
_cons	1.086862	.0573089	18.96	0.000	.9738876	1.199836

Data source: CSY 2008, 2009, 2011 and 2014.

concentration and R&D intensity. There are three insignificant factors: education, economic level and nationalization. The coefficient of all significant factors is positive.

As presented in Tab. 5, the model p-value is 0.0000 and Chi squared is 47.69, which implies that the model is effective in modelling the influence of the proposed factors of scale efficiency. There are four significant factors at the $p<0.05$ level, namely nationalization, economic level, enterprise scale and market concentration. Three factors are not significant: level of opening-up, education level and R&D intensity. Among the significant factors, only the coefficient of economic development is positive, while the coefficients of nationalization, enterprise scale and market concentration are negative.

It is evident that a higher level of opening-up has a positive effect on comprehensive R&D efficiency and pure technical efficiency. This may be explained by more opportunities for enterprises to communicate with the outside world and acquire advanced technology and management experience. Promotion of opening-up may assist firms in better allocating limited R&D resources towards the most beneficial streams of research and development. On the contrary, degree of nationalization has a significant negative effect on comprehensive and scale efficiency. This may be explained by considering company culture and decision making in Chinese state-owned enterprises, which are known for their highly hierarchical management structures, potentially leading to administrative interference and inflexibility in allocation of limited R&D resources. Restructuring of state-owned enterprises is one measure suggested in raising R&D performance.

Enterprise scale plays a positive role in promoting comprehensive efficiency and pure technical efficiency, but has a negative effect on scale efficiency. This may be explained as smaller firms having insufficient resources to produce effective R&D outcomes. However, in large firms with large R&D budgets, the focus of R&D efforts may be diluted across multiple fields, also resulting in decreased efficiency. Market concentration has a positive impact on pure technical efficiency, but a negative effect on comprehensive efficiency and scale efficiency. High Market concentration helps to promote communication among enterprises and then enhance technology advancement. However, extremely high market concentration intensifies competition and enterprises are forced to enlarge and increase R&D inputs, which is against the efficient use of R&D resources and causes the decrease of R&D efficiency.

It is found here that economic level does not have a significant correlation with R&D efficiency as also found by Wang et al. (2008). As commented earlier, there exists different opinions on which factors have the greatest influence on R&D efficiency. Our analysis shows clearly that comprehensive efficiency, pure technical efficiency and scale efficiency are influenced by different factors, which means that both internal and external factors should be considered in increasing R&D efficiency.

6 Conclusions and discussions

Using the DEA model, this paper measures the R&D efficiency of China's industrial enterprises above a designated size at provincial level, and explores spatial distribution and influencing factors. In contrast to existing studies, this paper explains R&D efficiency more objectively from the three perspectives of comprehensive efficiency, pure technical efficiency and scale efficiency. It is found that comprehensive efficiency, pure technical efficiency and scale efficiency show improvement over time, and scale efficiency is more significant than comprehensive efficiency and pure technical efficiency. Although comprehensive efficiency is low in general, it is rising rapidly. R&D efficiency displays significant spatial distribution at the provincial level. The hot and cold spots show distinct spatial differences and change significantly over time. The hot spots of comprehensive efficiency have shifted from Western China to coastal regions, in line with current regional economic development in China. Pure technical efficiency hot and cold spots express spatial dependence, but scale efficiency does not.

In contrast to previous studies, this paper also evaluates the factors influencing three distinct components of overall R&D efficiency. The level of opening-up, degree of nationalization, enterprise scale and market concentration have significant impact on comprehensive efficiency. The level of opening-up, enterprise scale, market concentration and R&D intensity correlate strongly with pure technical efficiency, while scale efficiency is largely determined by the degree of nationalization, economic level, enterprise scale and market concentration. It is seen that the factors which play a positive role in improving R&D efficiency in different regions and enterprises are different. Therefore, it is not possible to pinpoint one factor as the sole factor with the biggest role in increasing R&D efficiency. Analysis of R&D efficiency from various perspectives can reveal which factor is the most significant in different regions and enterprises. This new knowledge may guide enterprise decision making and regional and national government policy decisions, with the aim of improving the regional innovation environment and enterprise innovation capability.

Finally, the limitations of this study also should be discussed. Because of the limitations of available data, this paper only uses statistical data which has good stability and enables reliable conclusions. This leads to analysis and results at the macro level, which can explain general industrial and regional differences, but cannot resolve individual differences between enterprises and zones smaller than the provincial regions. Therefore, it is necessary to conduct more surveys on enterprises and regional policies in future research. In addition, this paper selects seven indices as the factors of R&D efficiency by referring to previous studies. In fact, enterprise R&D is affected by many factors, including measurable and unmeasurable factors. In future research, more factors should be called into question in order to give a more detailed and objective analysis.

Acknowledgement

This study was funded by the National Natural Science Foundation of China (41571110).

References

Anselin, L. (1995): Local indicators of spatial association - *LISA*. In: Geographical Analysis 27 (2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Cao, X.; Zeng, G; Zou, L. and Liu, G. (2016): Impact of R&D input on regional economic growth in case of the Yangtze River economic zone in China. In: Resources and Environment in the Yangtze Basin 25 (2), 208–218.

Charnes, A.; Cooper, W. W. and Rhodes, E. (1978): Measuring the efficiency of decision making units. In: European Journal of Operation Research 2, 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Chi, R. and Huang, Y. (2009): Analysis on inefficiency determinants of R&D activities in Zhejiang province. In: Science & Technology Progress and Policy 26 (17), 41–44.

Chi, R.; Yu, X. and Li, Z. (2004): An analysis of differences in innovation efficiency between the eastern and the western regions in China and its causes. In: China Soft Science (8), 127–131.

CSY (2008-2014): China statistical yearbook. http://www.stats.gov.cn/

Deng, L. (2010): The effect of FDI spillover and efficiency of independent technological innovation: the perspective of industrial characteristics of high-tech industries. In: Collected Essays on Finance and Economics 149 (1), 12–18.

Fang, C. and Guan, X. (2011): Comprehensive measurement and spatial distinction of input-output efficiency of urban agglomerations in China. In: Acta Geographica Sinica 66 (8), 1011–1022.

Feng, G.; Liu, J. and Xu, Z. (2006): Study on the factors on Chinese Industries R&D efficiency. In: China Industrial Economy (11), 46–51.

Hu, B. (2009): An empirical analysis of the efficiency of R&D activities in Chinese provinces and cities based on improved DEA model. In: Journal of Industrial Technological Economics 6, 80–83.

Lee, H. Y; and Park, Y. T. (2005): An international comparison of R&D efficiency: DEA approach. In: Asian Journal of Technology Innovation 13 (2), 62–75. https://doi.org/10.1080/19761597.2005.9668614

Li, X. (2013): An analysis of provincial large-middle industrial enterprise R&D efficiency's space difference and influence factors. Zhejiang Gongshang University.

Liang, L.; Zhang, H. and Yuan, Y. (2006): Study on R&D efficiency of enterprises based on DEA. In: Science Research Management 27 (6), 68–73, 143.

Liu, H.; Ma, L. and Li, G. (2017): Pattern evolution and its contributory factor of cold spots and hot spots of economic development in Beijing-Tianjin-Hebei region. In: Geographical Research 36 (1), 97–108.

Liu, L. (2008): Analysis of the efficiency changing of the S&T resources allocation and the influential elements in China. In: Science of Science and Management of S&T 7, 13–19.

Liu, Y. and Yang, X. (2012): Regional R&D efficiency in Chinese high-tech industries: empirical analysis based on provincial panel data. In: Chinese Journal of Management Science 20 (11), 653–658.

Sharma, S. and Thomas, V. J. (2008): Inter-country R&D efficiency analysis: an application of data envelopment analysis. In: Scientometrics 76 (3), 483–501. https://doi.org/10.1007/s11192-007-1896-4

Shen, W. (2009): Comparative analysis of R&D efficiency in Jiangsu and Zhejiang provinces based on DEA method. In: Northern Economy and Trade 2, 34–36.

Shi, X.; Zhao, S. and Wu, F. (2009): Analysis of regional innovation efficiency and spatial discrepancy in China. In: The Journal of Quantitative & Technical Economics 3, 55–61.

Tang, Q.; Lu, B. and Yuan, Y. (2009): Industrial firm's resources input and innovation efficiency. In: The Journal of Quantitative & Technical Economics 2, 3–17.

Wang, F.; Bai, Y.; Zhou, L.; Zhang, Y.; Qiao, F. and Ji, X. (2017): Spatio-temporal pattern and influencing factors of scientific and technological innovation capability of universities in China. In: Economic Geography 37 (12), 49–56.

Wang, L. (2009): Empirical study on the factors and regional difference of Chinese industrial enterprises R&D efficiency. Shenyang Ligong University.

Wang, L. and Xiong, Y. (2010): Provincial differences of R&D efficiency: the empirical analysis based on the data of China's large-medium industries. In: Science & Technology Progress Policy 27 (23), 32–36.

Wang, M.; Gu, L. and Yuan, G. (2008): The performance evaluation of regional R&D input of industry enterprise in China. In: Journal of Hebei University (Philosophy and Social Science) 33 (4), 38–41.

Wei, C. and Shen, M. (2008): Industry performance, technology efficiency and its affecting factors: an empirical analysis based on the 2004 Zhejiang economy census. In: The Journal of Quantitative & Technical Economics 7, 18–30.

Wu, H. (2005): A relative efficiency appraisal on regions of R&D. In: Science and Technology Management Research 25 (1), 54–56.

Xie, W.; Hu, W. and Xia, S. (2008): Study on R&D efficiency and its determinants of Chinese high-tech industry. In: Science of Science and Management of S&T 3, 144–149.

Yan, B. and Feng, G. (2005): Study on China industrial R&D efficiency based on stochastic frontier production and cost function estimation. In: Modern Economic Science 27 (6), 14–18.

Yang, G.; Liu, W. and Zheng, H. (2013): Review of data envelopment analysis. In: Journal of Systems Engineering 28 (6), 840–860.

Ye, Z. (2006): An analysis of technical efficiency and its influencing effects on hi-tech industry in China. Hefei University of Technology.

Yu, Y. (2006): Study on the evaluation of China's regional R&D efficiency. In: Economic Forum (24), 21–24.

Yue, S. (2008): R&D efficiency differences and affecting factors of Chinese regions: empirical analysis based on the provincial data. In: Science Research Management 5, 173–179.

Zhang, A.; Zhang, Y. and Zhao, R. (2003): A study of the R&D efficiency and productivity of Chinese firms. In: Journal of Comparative Economics 31 (2003), 444–464. https://doi.org/10.1016/S0147-5967(03)00055-6

Zhang, H. (2008). R&D production efficiencies and determinants in Chinese manufacturing. In: Studies in Science of Science 26 (5), 970–978.

Zhang, H. (2010): On the provincial total factor R&D efficiency and its determinants in China:1999-2007. In: China Economic Quarterly 9 (3), 1029–1050.

Zhao, J. and Zeng, G. (2009): Spatial Measurement of innovation: data and indicator. In: Economic Geography 29(08), 1250–1255.

Zou, L.; Zeng, G. and Cao, X. (2015): The spatial variation and spatial-temporal evolution of R&D input in the Yangtze River Delta city group: based on ESDA analysis. In: Economic Geography 35 (03), 73–79.

Authors

Prof. Dr. Chengyun Wang
Shanghai Normal University
Department of Geography
No.100 Guilin Road
Shanghai
China
chengyun@shnu.edu.cn

Yue Wang
yuew711@163.com

Dr. Nana Li
nanali06@163.com

Tengfei Ma
mtf_0518@163.com