Computer-Aided Design AND APPLICATIONS www.cadanda.com

Research on Economic Development Strategy Assisted by Computational Space Clustering

Zheng Zhang[1]

[1]Northeast University of Petroleum Qinhuangdao, HeBei, QinHuangDao 066004, China, zhangzheng@nepu.edu.cn

Corresponding author: Zheng Zhang, zhangzheng@nepu.edu.cn

Abstract. Economic development is closely related to the region. In order to realize the regional analysis of economic development and explore the law of economic development, based on computer-aided technology, this study combined spatial clustering technology to construct a regional research simulation model and set a number of factors to study the law of regional economic development. At the same time, this paper analyzed the actual development of the region from 2006 to 2018 in China and compared the simulation research data with the actual development of China. In addition, this paper proposed different implementation paths for regional differences, and elaborated on the implementation path from the selection conditions, operational mechanism construction, implementation approaches and strategies. Finally, the paper verified the validity of the model and provided a theoretical reference for subsequent related research.

Keywords: spatial clustering; economic development; regional economy; simulation analysis
DOI: https://doi.org/10.14733/cadaps.2020.S1.137-146

1 INTRODUCTION

There are many factors influencing the regional economic gap, such as capital flows, labor mobility, human capital distribution, the number of regional universities, the degree of urbanization, and institutions and policies. Among all the factors affecting the regional economic gap, capital is an indispensable element in the development of the economy, both at the micro and macro level. In the process of continuous economic globalization, the scope of regional economic operations has become blurred. The organization, flow, agglomeration and management of production factors in the process of urban and regional economic operations are no longer limited to a single spatial extent but occur in a number of different geographical areas. Driven by global governance, cities have a closer relationship with cities and regions in different geographic areas. Under the background of the country's construction of an open economy and the promotion of the Belt and Road Initiative, the spatial scale structure of China's regional economy will undergo major adjustments, and the scale structure of economic operations needs to be put on the agenda [1].

Abdulai first proposed the theory of regional economic growth. The theory divides the economic growth of all regions into five stages: the self-sufficient economic stage, the rural economic interest stage, the rural industrial transformation stage, the industrial economic stage, and the tertiary industry output stage, and believes that all developments have a standard development order. The French economist PeHoux proposed the growth pole theory of regional economic growth. The theory holds that the growth of industrial agglomeration is the center of economic activity, affecting the center and margin of the region. Moreover, the industrial gatherings form economies of scale, which in turn drives the economic development of the entire region [2]. On the basis of Perroux, Fleischmann K believes that the growth pole has a negative effect on the economic growth of the surrounding areas, which is considered to be the echo effect. As the economic benefits diminish from the center of the region to the periphery, all production factors will be concentrated in the regional center, and the economic differences within the region will expand [3].

Fleischmann K [4] proposed the core-edge development theory and believed that the development of the regional core will spread to the periphery of the region, thus promoting the systematic development of the relevant space. Mcfarlane J developed a point-to-axis development theory from the growth pole theory in the The theory takes the growth pole as a point and the traffic between the points as the axis, and various resource elements are gathered at two points, thus generating new growth points [5].

Yu-Ying W further developed new growth theories. The theory holds that knowledge is a new growth force that promotes regional economic growth. This rationality breaks through the assumption that the economic factors have not changed in the past and has found new growth points for economic growth [6]. When the Mirolubova T V studied the economic gap between developed and developing countries, it found that developed countries have the conditions for convergence and convergence of development, while developing countries do not have this trait, and the international economic gap between poor and rich countries will become larger and larger and will not reverse each other's economic status [7].

In the study of the formation of economic differences, Kozlova O A first proposed the Kim curve and showed that the different levels of different factors in different periods have caused different levels of development of the regional economy. Since the elements of input from the edge to the center in a region are different, the differences in economic development are generated, and the economic differences between regions are similar [8]. Yoon believes that the central and marginal regions of the region will have two effects in economic development. Developed regions will attract resources from underdeveloped regions and developed regions will also form economic transmissions for backward regions to form economic transmission, that is, polarization effect and trickle-down effect. In general, because the polarization effect is greater than the trickle-down effect, economic differences are produced [9].

Fritsch M believes that regional economic differences are "U" type. In the early stage of economic development, regional economic differences will expand and maintain differences in stability for a certain period of time. However, after a period of economic development, especially after entering the stage of mature growth, regional economic differences will gradually shrink [10].

Leigh N G proposed the theory of regional economic development cycle by studying the development process of industrial zones. A regional economic development follows a fixed cycle as well as an organism, and different stages have different problems. Therefore, different regions have differences due to different phases of their respective cycles [11]. Eda U believed that regional economic differences are due to differences in industrial structure. The economic development of the region presents a gradient, and the transformation of industrial structural advantages has formed a gradient of high and low. Therefore, this theory is also called gradient transfer theory [12]. Gaspar proposed a cyclical accumulation causal theory. The theory holds that regional economic development is the result of economic agglomeration, and resource accumulation is the premise of agglomeration, and a variety of factors generate cyclical accumulation of benefits, which leads to regional economic growth [13]. Tsekeris T proposed the bell-shaped theory to explain the changing process of regional economic growth differences. The theory holds that, in general, the process of formation and variation of regional differences is bell-shaped [14]. Tokunaga S introduced geographic location as an important variable in regional economic analysis, which enriched the study of regional economic differences. With the development of concepts and theories such as economic globalization, new geo-economics, and horizon theory, and the further deepening and innovation of quantitative research, the study of regional differences has been deepened and comprehensive, and the academic community has made the study of regional economic differences more objective [15].

2 RESEARCH METHOD

2.1 Data preprocessing

The main task of data preprocessing is to standardize the raw data directly obtained from the problem on the basis of noise removal. This mainly includes data attribute selection, data cleaning, and data centralization and standardization. Attribute selection refers to properly selecting the appropriate fields from the data set, which can effectively reduce the amount of calculation and improve the clustering results. The main task of data cleaning is to kick out null and noise values and correct data errors. Centralization is the adjustment of the observed values of the variables to the same base point, and the average of this dimensional variable is usually subtracted from each data. Data discretization preprocesses the attribute values of the data. It divides the attribute into a finite number of parts and then uses the label of this part instead of the original attribute value. In this paper, data needs to be normalized. Data normalization is the process of transforming the scope of data into a relatively small, deterministic range. Commonly used data normalization methods are minimum and maximum normalization, z-score normalization, and decimal normalization. The minimum and maximum normalization is the method of transforming data into a certain interval with the minimum maximum value of the data. The following formula is an example of the minimum and maximum specification, which maps the data to the interval $[0,1]$ [16].

$$x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

In the formula, $x_{\min}$ is the minimum value of the indicator attribute and $x_{\max}$ is the maximum value of the indicator attribute.

2.2 Cluster analysis

Cluster analysis is a modeling process for constructing a clustering model based on the data to be classified. The complete cluster analysis mainly includes four stages: data preprocessing, feature calculation and extraction, cluster pattern discovery and result interpretation. Clustering is the process of dividing data into groups. According to the similarity between the natural distribution nature of data and the degree of existence between data variables, the most similar data are clustered according to certain criteria. Cluster analysis not only can reflect the similarity between data objects according to the attribute changes of objects, but also can find the dissimilarity of data objects. Cluster analysis is a technique for studying the logical or physical relationship between data. It divides the data set into several categories composed of similarly similar data points through certain rules. The results of cluster analysis can not only reveal the internal connections and differences between data, but also provide important basis for further data analysis and knowledge discovery, such as association rules between data, classification patterns and change trends of data [17].

Cluster analysis has applications on many practical issues. In business, clustering can help market analysts discover different customer segments from the customer base library and use the buying model to characterize different customer segments. Cluster analysis is an effective tool for market segments, and it can also be used to study consumer behavior, find new potential markets, select experimental markets, and act as a pre-processing for multivariate analysis. Biologically, clustering can be used to derive the classification of plants and animals, classify genes, and gain an understanding of the inherent structure of the population. As a data mining function, cluster analysis can be used as an independent tool to obtain data distribution, observe the characteristics of each cluster, and focus on further analysis of certain clusters [18].

Most clustering algorithms can be divided into partitioning clustering method, hierarchical clustering method, density-based clustering method and opportunity grid clustering method. In this paper, the K-means clustering method in partitioning clustering is applied by comprehensive analysis, and the method of dividing clustering is introduced below.

The partitioning clustering method is a prototype-based clustering method. The basic idea is to first select several objects from the data set as the prototype of the cluster, and then assign the other objects to the most similar ones represented by the prototype, that is, the nearest class. The main advantage of this method is that it is less complex and relatively scalable and efficient for processing big data. The disadvantage is that this type of algorithm requires that the number k of clusters to be generated be given in advance, and the selection of the initial points of k numbers has a great influence on the clustering result. In addition, this type of algorithm can only find non-recessed spherical clusters and is sensitive to noise data.

The specific flow of the K-Means algorithm.

The input is the sample set $D=\{x_1,x_2,\cdots,x_m\}$, the clustered cluster tree k, the maximum number of iterations

The output is cluster division $C=\{c_1,c_2,\cdots,c_k\}$.

(1) A sample of k numbers is randomly selected from data set D as the centroid vector of the initial k numbers: $\{u_1,u_2,\cdots,u_k\}$

(2) $n=1,2,\cdots,N$

(a) Cluster division C is initialized to $Ct=\varnothing, t=1,2,\cdots,\mathrm{k}$

(b) For $i=1,2,\cdots,\mathrm{m}$, the distance d_{ij} between the sample x_i and each centroid vector $u_j\left(j=1,2,\cdots,k\right)$ is calculated. (Formula 2).

(c) For $j=1,2,\cdots,k$, all sample points in c_j are recalculated with new centroids (see Formula 4).

(d) If the centroid vectors of all k numbers have not changed, the algorithm proceeds to step (3).

$$d_{ij}=\left\|x_i-u_j\right\|_2^2 \tag{2}$$

$$c_{\lambda i}=c_{\lambda i}\cup\{x_i\} \tag{3}$$

$$u_j=\frac{1}{\left|c_j\right|}\sum_{x\in c_j}x \tag{4}$$

(3) Cluster division $C=\{c_1,c_2,\cdots,c_k\}$ is output.

2.3 Spatial overlay analysis

Overlay analysis is a very important spatial analysis feature in GIS. It refers to the process of generating new data through a series of set operations on two data under the unified spatial reference system. The data mentioned here can be the data set corresponding to the layer, or it can be a feature object. The overlay analysis of multi-layer data not only creates new spatial relationships, but also generates new attribute relationship relationships, and can discover features such as mutual differences, connections, and changes between multiple layers of dataThis topic will use superposition analysis to study the differences and similarities between economic developments between regions. It is more intuitive to discover the similarities and differences between regions through superposition.

There are many applications for spatial superposition analysis. For example, He Yu and Liu Gang used the Jianshi County of Hubei Province as an example to study the relationship between altitude, slope and aspect and agricultural production conditions in the area and established a land suitability evaluation model. With the support of MapGIS, the application of digital elevation model (DEM) and spatial superposition analysis technology is used to realize the application of the model in land suitability evaluation. Practice has shown that the results obtained by the digital quantitative evaluation method are consistent with the actual situation, and it is characterized by high efficiency, accuracy and visualization, and can support government decision-making.

Aiming at the efficiency problem of spatial data superposition analysis of massive land use vector data, Qi Fengying et al proposed a spatial superposition analysis method based on Spark for land use vector data. This method implements index filtering and superposition calculation through elastic distributed data sets and makes a new attempt to solve the bottleneck problem of spatial superposition analysis. The experimental results show that compared with the superposition analysis method based on Oracle data management, this method significantly improves the efficiency of spatial superposition analysis and is more suitable for superposition analysis of massive land use vector data.

The GIS studied by Deng Xudong and Wang Jing in 2015 as a well-established means and platform for geospatial data processing and analysis has been valued in many fields of social economy. Moreover, site selection analysis, as one of its applications, is increasingly used. By reading the literature, the steps of using GIS location in retail stores, real estate, warehouses and shopping malls are presented, and the method of combining GIS with MCDA is studied.

3 SYSTEM CONSTRUCTION

The economic subsystem is the core part of the economic development system, and the main indicators involved are shown in figure 1. The number of employees in the three industries is the link between the population subsystem and the economic subsystem, which is determined by the total population and affects the output value of the three industries. The output value of the three industries also determines the real GDP and social fixed assets investment. The GDP interacts with the technology investment and energy consumption respectively, and is the bridge for the economic subsystem to communicate the energy subsystem. The population also affects per capita GDP and per capita fixed asset investment, and these two variables ultimately determine the size of real GDP.

The population subsystem is an important factor in the economic development system. The key variable is the population size, which is determined by the birth rate and mortality rate of the floating population and is the decisive ratio variable of the number of employees in the three industries. At the same time, the impact of population quality on population size and carbon emissions was taken into account. The quality of the population is determined by the investment in science and technology. It is mainly represented by the number of people in higher education. Generally speaking, the higher the quality of the population, the stronger the awareness of low-carbon energy conservation. The system mainly includes the following variables: population size, number of employees in three industries, floating population, birth rate, population mortality, natural population growth rate, per capita energy consumption, labor, GDP, technology investment, and population quality, and the functional relationship among them can be shown in Figure 2.

Compared with provincial units, cities can measure regional economic systems in more detail. Moreover, compared with county, township and other units, the availability of continuity data at the city level is higher. Therefore, this paper selects cities at the prefecture level and above as research objects. Input indicators include capital investment and labor input, capital investment is urban fixed capital stock, labor input is the number of employees, and output indicator is GDP.

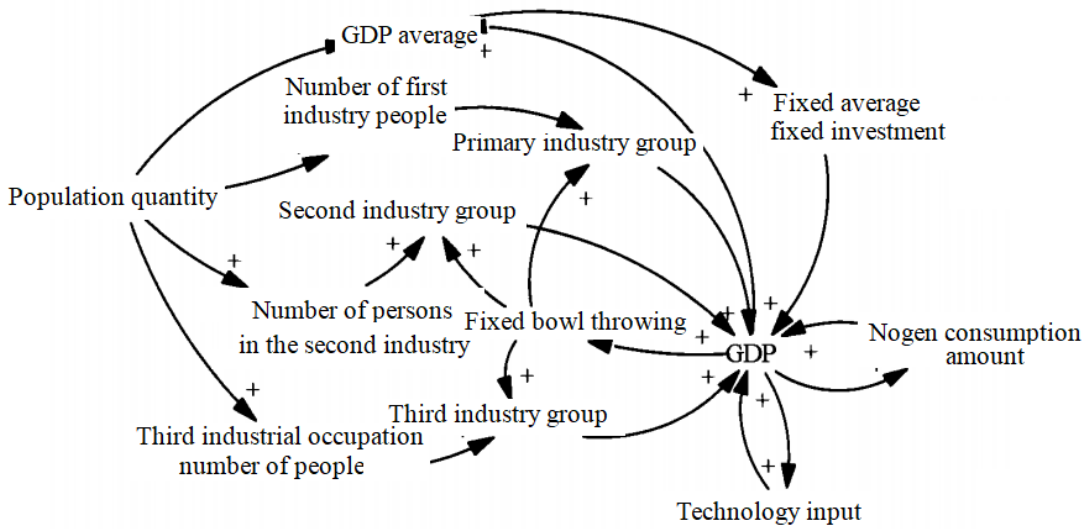

Figure 1: Causal relationship diagram of the economic subsystem.

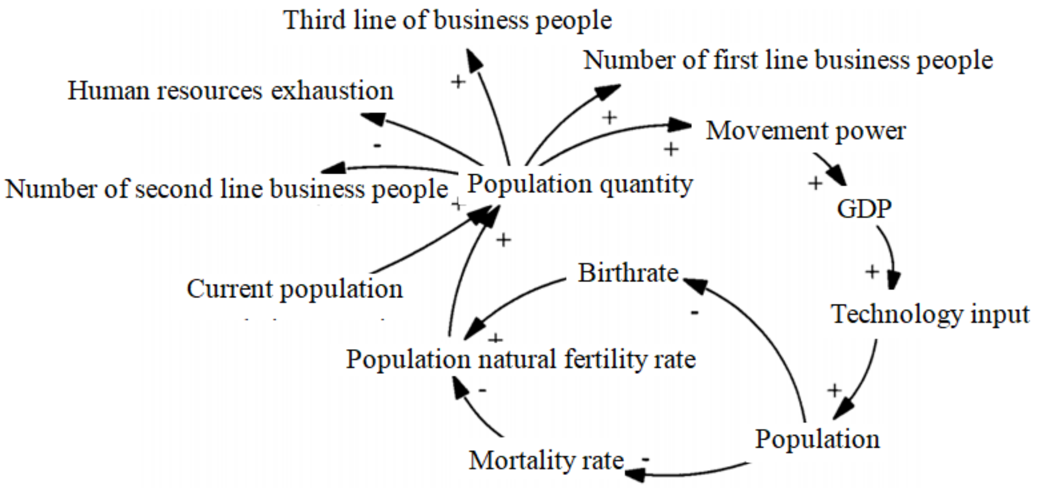

Figure 2: Causal diagram of the population subsystem.

Before 2003, the administrative divisions of prefecture-level and above cities in China changed a lot, and the adjustment of relevant indicators was difficult, and it was not the focus of this paper. Therefore, this paper will set the research period to 2006-2018. The calculation of the fixed capital stock needs to use the average of the previous three years, and all price-related indicators are uniformly adjusted to the 2001 constant price. The GDP deflator is about the weighted average of the consumer price index and the fixed-asset price index, and the weight is the proportion of consumption and fixed-asset investment in GDP (Shen Lisheng and Wang Huogen, 2008). This paper takes the investment and consumption as the weight of GDP as the weight and calculates the GDP deflator according to the fixed asset investment price index and the consumer price index of the province where the city is located. The GDP was also adjusted to the constant price in 2001. The data comes from the China Statistical Yearbook and the China City Statistical Yearbook.

Since the number of subsystems is too large, it cannot be displayed completely. Therefore, this article uses ArcGIS10.3 to visualize the hierarchical diagram of the connection between two points. The rules are as follows: According to the national average (0.665), 40470 pairs of subsystems were divided into strong synergy group (>0.665) and weak coordination group (<0.665), and weak synergy group was not displayed. The subsystems below the mean (0.740) in the strong synergy group are divided into three levels of synergy, and then the interval is divided: first level synergy 0.870-1.000; second level synergy 0.740-0.870; third level synergy 0.665-0.740 (not displayed). On this basis, the sum of the outward coordination of each city is obtained, and it is judged whether it is a synergy center. The maximum and minimum values are 0.838 and 0.359, respectively, and the upper and lower limits are limited to 0.840 and 0.350. Cities below the national average (0.665) are not considered collaborative centers. Moreover, the cities below the mean (0.680) in the collaborative center city are divided into three-level centers, and then divided into equal intervals: the first-level center is 0.760-0.840; the second-level center is 0.680-0.760; the third-level center is 0.665-0.680; the non-center is 0.350-0.665 (not displayed).

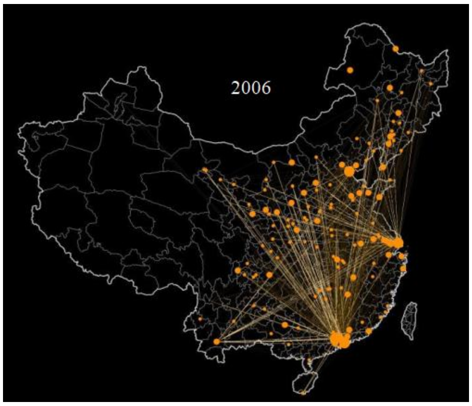

Figure 3: Coordinated development pattern of regional economy in 2006.

Figure 4: Coordinated development pattern of regional economy in 2009.

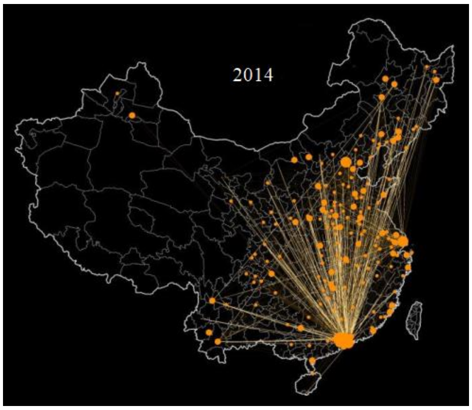

Figure 5: Coordinated development pattern of regional economy in 2014.

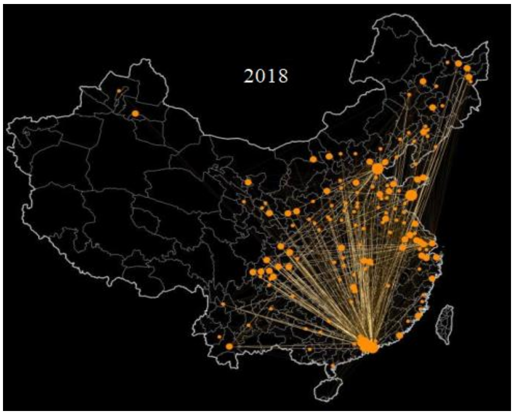

Figure 6: Coordinated development pattern of regional economy in 2018.

It can be seen from Figure 3-6 that during 2006-2018, China's overall cooperative radiation network centered on the core cities of the Pearl River Delta, the Yangtze River Delta and the Beijing-Tianjin-Hebei region. The network density shows a "strong-weak-strong-weak" volatility change, which is consistent with the major impact of China's macroeconomics and the time inflection point of fundamental changes. The regional economy scatter diagram is shown in Figure 7.

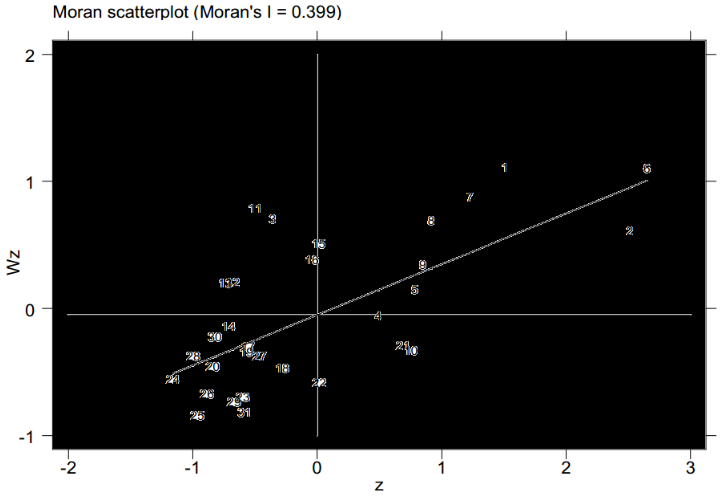

Figure 7: Regional economic scatter plot.

4 ANALYSIS AND DISCUSSION

The first weakening of collaborative network density occurred in 2009. Comparing the synergistic spatial pattern of 2006 and 2009, it can be seen that the color of the connecting lines radiating from the core cities of the Pearl River Delta and the Yangtze River Delta region is generally shallower, which indicates that the number of dark-colored connecting lines of the first-level coordination is significantly reduced, and the number of connecting lines radiating outward from the Beijing-Tianjin-Hebei region also shows a sharp change trend, and the number of connecting lines radiating outward from the Beijing-Tianjin-Hebei region has also shown a sharp change trend, and the combined effect has led to the first weakening of the synergistic network density. Combined with the international situation and China's macroeconomic analysis, it is known that due to the impact of the subprime mortgage crisis, a large number of financial institutions such as banks, trusts, insurance, and fund management went bankrupt, and both large-scale listed companies and small and medium-sized enterprises suffered severe setbacks. The impact of the financial market has directly weakened and even narrowed the exchange of elements between the regions with funds as the main factor. Moreover, the break of the capital chain also brought about layoffs and even bankruptcies, and further weakened the labor-based exchange of elements, which led to a sharp drop in the intensity of cross-regional factors. In addition, inter-regional synergies as a visual representation are also significantly weakened. The conclusion that the subprime mortgage crisis has impacted China's economic situation is consistent with the conclusions of some experts and scholars.

The second weakening of collaborative network density occurred in 2015. After the subprime mortgage crisis, China's economy gradually recovered. Since 2013, the synergy network density has rebounded significantly and reached its highest value in 2014. Comparing the synergistic spatial pattern between 2001 and 2014, the density of the connecting lines radiating outward from the Pearl River Delta and the Yangtze River Delta region is enhanced, the color is deepened, and the status of the synergy center in the Beijing-Tianjin-Hebei region has also rebounded significantly, and the overall synergistic network density has increased significantly compared to 2010, and is significantly stronger than in 2006, and reached the highest value in nearly ten years. However, this also entered the second weakening of the synergy network density. China is caught in the economic "new normal" of the economic growth slowdown period, the structural adjustment pain period and the "three-phase superposition" of the previous policy digestive period. At this stage, the crisis of supply and demand is blocked, the crisis of overcapacity is revealed, and the contradiction between factor supply efficiency and low resource allocation efficiency is gradually highlighted, which directly leads to low efficiency of inter-regional resource allocation and weakened synergy network density. Although the "new normal" of China's economy was first proposed in 2014, the contradiction between the low efficiency of factor supply and allocation has already appeared. The macro data show that the time impact of the major impacts and fundamental changes in China's regional economy is consistent with the time inflection point of the change in the density of collaborative networks, which not only fully explains the reasons for the synergistic changes, but also proves the reliability of the efficiency-added model.

5 CONCLUSION

This paper combined computer-aided systems to study the path of coordinated development of regional economic growth in China. Moreover, this paper empirically tested the economic growth effect of regional economic synergy development from the two dimensions of inward synergy and outward coordination and explored whether synergy can tap the new potential of regional economic growth in China. Thirdly, this paper empirically tested the synergistic driving effect of regional economy from the aspects of single-drive effect and multi-drive effect and explored the internal root causes of regional differences in synergistic growth effect from the front-end of coordinated development of regional economy. Finally, according to the regional differences, this paper proposes different implementation paths, elaborates on the selection conditions, operation mechanism construction, implementation approaches and strategies, and proposes targeted policy recommendations for the problems revealed by the empirical test to enhance the coordinated growth momentum of China's regional economy.

6 ACKNOWLEDGEMENTS

Youth Fund Project of Heilongjiang Provincial Undergraduate Colleges and Universities, Name: Study on evaluation and optimization of investment and business environment in Heilongjiang province, NO:2018QNQ-13.

7 ORCID

Zheng Zhang, https://orcid.org/0000-0001-6237-1198

REFERENCES

[1] Liu, S.; et al: A three-scale input-output analysis of water use in a regional economy: Hebei province in China, Journal of Cleaner Production, 156, 2017, 962-974, https://doi.org/10.1016/j.jclepro.2017.04.083
[2] Abdulai, A-G.: The political economy of regional inequality in Ghana: Do political settlements matter?, The European Journal of Development Research, 29(1), 2017, 213-229, https://doi.org/10.1057/ejdr.2016.11
[3] Fleischmann, K.; Daniel, R.; Welters, R.: Developing a regional economy through creative industries: innovation capacity in a regional Australian city, Creative Industries Journal, 2017, 1-20, https://doi.org/10.1080/17510694.2017.1282305
[4] Fahmi, F. Z.; Koster, S.: Creative industries and regional productivity growth in the developing economy: evidence from Indonesia, Growth and Change, 2017, https://doi.org/10.1111/grow.12212

[5] Mcfarlane, J.; Blackwell, B.; Mounter, S.: Good gardening for a perennial economy: What's the optimal growth path for a regional economy?, The Journal of Developing Areas, 52(1), 2018, 29-44, https://doi.org/10.1353/jda.2018.0002
[6] Yu-Ying, W.: Coordinated development of Yunnan's frontier culture and economy under the regional economic integration pattern, Ecological Economy, 2017(02), 67-73, https://doi.org/CNKI:SUN:STJY.0.2017-02-008
[7] Mirolubova, T. V.; et al: The problem of reproduction of basic assets in industry as the Barrier for neoindustrial transformation of regional economy, SHS Web of Conferences, 2017, 35, https://doi.org/10.1051/shsconf/20173501094
[8] Kozlova, O. A.; Nifantova, R. V.; Makarova, M. N.: Methods of the assessment of economic losses caused by the mortality of the population employed in regional economy, Economy of Region, 1(2), 2017, 511-523, https://doi.org/10.17059/2017-2-16
[9] Yoon, D.: The regional-innovation cluster policy for R, &, D efficiency and the creative economy: with focus on Daedeok Innopolis, Journal of Science and Technology Policy Management, 8(2), 2017, JSTPM-09-2016-0025, https://doi.org/10.1108/JSTPM-09-2016-0025
[10] Fritsch, M.; Wyrwich, M.: The effect of entrepreneurship on economic development—an empirical analysis using regional entrepreneurship culture, Journal of Economic Geography, 2016, lbv049, https://doi.org/10.1093/jeg/lbv049
[11] Leigh, N. G.; Kraft, B. R.: Emerging robotic regions in the United States: insights for regional economic evolution, Regional Studies, 2017, 1-13, https://doi.org/10.1080/00343404.2016.1269158
[12] Eda, U.; Carlo, L.; Boris, P.: Examining lag effects between industrial land development and regional economic changes: The Netherlands experience, PLOS ONE, 12(9), 2017, e0183285, https://doi.org/10.1371/journal.pone.0183285
[13] Gaspar, J. M.; Castro, S. B. S. D.; Correia-Da-Silva, J.: Agglomeration patterns in a multi-regional economy without income effects, Economic Theory, 2017, https://doi.org/10.1007/s00199-017-1065-9
[14] Tsekeris, T.; Papaioannou, S.: Regional determinants of technical efficiency: evidence from the Greek economy, Regional Studies, 2017, 1-12, https://doi.org/10.1080/00343404.2017.1390312
[15] Tokunaga, S.; Okiyama, M.: Impacts of industry clusters with innovation on the regional economy in Japanese depopulating society after the Great East Japan Earthquake, Asia-Pacific Journal of Regional Science, 2017, https://doi.org/10.1007/s41685-017-0041-5
[16] Husgafvel, R.; et al: Forest sector circular economy development in Finland: A regional study on sustainability driven competitive advantage and an assessment of the potential for cascading recovered solid wood, Journal of Cleaner Production, 2017, S0959652617331475, https://doi.org/10.1016/j.jclepro.2017.12.176
[17] Santos, A. M. P.; et al: Assessment of port economic impacts on regional economy with a case study on the Port of Lisbon, Maritime Policy & Management, 2018, 1-15, https://doi.org/10.1080/03088839.2018.1471536
[18] Dong, B.; et al: Impacts of exchange rate volatility and international oil price shock on China's regional economy: A dynamic CGE analysis, Energy Economics, 2017, S0140988317303171, https://doi.org/10.1016/j.eneco.2017.09.014