

Research on Regional Economic development of Guangxi based on multiple statistic analysis

Xin Jian ^{a,1,*}, Weizhang Lai ^{a,2}

^aDepartment of mathematics and statistics, Guangxi Normal University, China

¹jianxinaka@163.com ^{*}; ²2817320567@qq.com

^{*} corresponding author

ARTICLE INFO

Keywords

Economic development

PCA

Cluster analysis

ABSTRACT

Economic development is a prerequisite for social stability and development. However, the regional economic development in various regions of our country is unbalanced, especially in the western region. As one of the minority autonomous regions in western of my country, Guangxi Province is undoubtedly of great practical significance to study its regional economic development. In this context, this paper is based on 8 economic indicators and related economic data of 14 cities in Guangxi Province in 2018, and is based on the theory of dimensionality reduction methods and cluster analysis, using principal component analysis (PCA) and hierarchical clustering methods to assess the comprehensive level of Guangxi's economic development. The result shows that most cities' economic development in Guangxi Province is great. However, there are signature differences in economic development between urban areas.

This is an open access article under the [CC-BY-SA](#) license.



1. Introduction

In recent years, the economic development of regions in China has been a topic of wide concern. Meanwhile, the research of it is a major strategy of the country and a significant way to understand the local states of economic and promote economic development. With the instant development of market economy, the economic power gap between regions is also increasing, and most countries are encountering with the problem of unbalanced regional economic development. At the same time, as early as 2017, the 19th National Congress of the Communist Party of China pointed out: China's economy has entered a period of transformation and upgrading, that is, a period of transformation of development mode, optimization of economic structure, and transformation of growth momentum. China's economic development is no longer a singular pursuit of speed, but a steady pursuit of high-quality development. Therefore, how to analyze and handle the relationship of regional economic development correctly is important.

Over the years, a large number of national and international professors and scholars have done many kinds of research on regional economic development, and they generally consent to regional economic development are affected by abounding factors.

During the already-known researches about regional economic development, Myrdal[1] found that market integration will cause the overall economic growth of a country or a region to show a trend of polarization, thereby expanding the difference in regional economic development. Used the data from 1860-1963 which concludes 35 countries, discovering that financial development positively affects economic growth [2]. Pereira and applied the vector autoregressive method to analyze the impact of public investment on the long-term output of the first twelve countries in the euro area during 1980-2003, the result demonstrates that productive public investment has a significant positive impact on the economic growth of eight of the twelve euro area countries [3].

Discovered that the biggest reason for the imbalance of Chinese regional economic development is that the government implements different development strategies and policies in various regions [4]. Average GDP to decompose the their index. Finding the difference of economic represents a range of U curve between Inland areas and coastal areas in our country [5]. Meanwhile, this difference decreased continuously before 1990s, but increased continuously after 1990s. Had ever divided our country into four parts: eastern and western, coastal cities and inland areas [6].

The result showed that the accumulation of our country human capital could promote the economic development. used a total of 43 initial indicators from four aspects: basic indicators, people's livelihood indicators, economic indicators, and public indicators, applying factor analysis and cluster analysis methods to comprehensively analyze the development status of the regional economy in Gansu Province, and get the conclusion that the development level among the various (prefectures) and cities in Gansu Province which has shown an obvious imbalance [7],[8]. used the literature research method and summary analysis method to analyze the status quo of Heilongjiang Province's construction of the "digital Silk Road" and proposed implementing a complete support system, cultivating new digital services and digital talents [8]. The plan promoted the transformation and development of the regional economy in Heilongjiang Province. Mainly analyzes and evaluates the comprehensive level of the economic development of Beijing-Tianjin-Hebei based on principal component analysis and factor analysis, and further recognizes the current regional differences in economic development and pays attention to the development results brought about by differences [9]. Thereby improving the economic quality of the entire urban agglomeration.

Meanwhile analyzed the impact of population mobility on economic growth in 14 prefecture-level cities in Liaoning Province and the mechanism from both the theoretical and empirical aspects by comparing the population mobility and opening-up prefecture-level cities [10]. The result showed that economic opening Regional population agglomeration effects and appropriate population policies could attract population inflows. At the same time, population agglomeration has a positive effect on economic growth, and population outflow has a negative economic effect on 14 prefecture-level cities in Liaoning Province. Guided by the new development concept, using the panel data of my country's provinces and cities from 2013 to 2019 to measure and compare the level of high-quality economic development, it is concluded that the level of high-quality economic development in my country is showing an upward trend year by year. , "Coordination" and "green" are relatively weak links. The gap in the level of high-quality economic development between regions is obvious, and it is showing an expanding trend [11]. High-quality development in economically developed areas is generally high, while the level of high-quality development in economically underdeveloped areas is lower. Based on the intermediary role of the industrial structure and the adjustment role of the demographic

dividend, empirically tested the mechanism of energy consumption structure upgrading, industrial structure adjustment and regional economic high-quality development [12]. The study showed that China's overall high-quality development level is on the rise. However, there are regional imbalances and regional high-value clustering problems. The spatial spillover effects of high-quality economic development in various regions are obvious; the upgrading of energy consumption structure can significantly promote the regional economy through the direct path and the intermediary effect of industrial structure adjustment (indirect path) is high-quality development; the demographic dividend plays a positive role in the first half of the intermediary path of industrial structure adjustment, and it can significantly enhance the positive impact of the upgrading of energy consumption structure on industrial structure adjustment.

Selected data from China's prefecture-level cities from 2009 to 2018, and examined the impact of fiscal expenditure, opening to the outside world, and combining the two on the high-quality development of the regional economy [13]. The results of the study show that the fiscal expenditure, opening to the outside world, and the combined effects of the full sample data and the large city sample significantly positively affect the high-quality development of the regional economy, and the fiscal expenditure of the medium-sized and small city samples significantly positively affects the region. The level of high-quality economic development, and the interaction terms of opening up and fiscal expenditure and opening up to the outside world have no significant impact on the high-quality development level of the regional economy, and opening up negatively affects the high-quality development level of the regional economy.

Selected the representative node years (2001, 2005, 2010 and 2015) of GDP growth rates in 31 provinces in China (excluding Hong Kong, Macau and Taiwan), based on the Markov chain model and natural breakpoints law to analyze the state of my country's regional relay growth since the new century [14]. The results show that the relay growth of my country's regional economy presents a state of relaying from the coast to the inland, from the north to the south. During the study period, the medium-to-high-level convergence clubs formed by the economic growth rate and development level respectively showed a trend of shifting from the north-south vertical belt-like distribution in the central and western regions to the grid-like distribution in the central and southern regions; the current structural adjustment stage, the eastern coastal and northeastern regions rely on transformation and reforms to pick up the economic growth rate.

The central and western regions, mainly Chongqing, Guizhou, Anhui, Jiangxi and other provinces and cities, have gradually become the main body of regional relay growth, and the ethnic and border areas have been in a slow growth stage; The evolution of the relay growth pattern is the result of the combined effects of multiple factors, and the regional resource and energy endowments have laid the foundation for relay growth. The distribution of economic development gradually adjusts the rate of regional relay growth, and the transformation of regional development mechanisms and spatial functions promotes the cyclical link of relay growth in different regions. The three interact under the guidance of the national regional development strategy to determine China's future regional relay growth trend jointly. Took Sichuan Province which in China as an example, using cluster analysis and principal component analysis to conduct empirical research on regional economic differences [15]. Through research, it is found that Sichuan Province has large regional economic differences. The 12 cities (prefectures) have negative comprehensive economic development scores, and their development level is poor. Six cities (prefectures) have the lowest level of economic development and comprehensive

economic development. Only 3 cities (states) have a good level of comprehensive economic development.

Used panel data from 2015 to 2019 in Anhui Province, China to conduct empirical analysis [16]. The results show that regional economic differences in Anhui Province are affected by factors such as system, industry, labor, capital, and nature; policy support Intensity, industrial structure, labor quality, human capital input, and location advantage are the most important and representative five factors; each factor has a positive effect on the difference in regional economic development level and development speed in Anhui Province. Policy support and industrial structure are the two most powerful and significant factors, and the other factors are slightly weaker. Selected panel data from 31 provinces (regions, cities) in China from 2010 to 2018 based on regional differences [17]. They constructed a spatial Dubin model for empirical analysis. The study found that the economic development of various regions in China has a high degree of spatial correlation; regional technological innovation has different effects on economic development. Increasing the scale of technological innovation in the eastern region can synergistically improve economic development and produce positive spatial spillovers to neighboring regions.

The effect of technological innovation in the central and western regions can promote the economic development of the region, but it has rms neighboring regions, and the positive effect is greater than the negative effect; the consumption transition of all regions has the same influence on economic development, showing a promoting effect and positive spatial spillover effects. Selected the economic impact reports issued by 10 world-class universities, based on the theoretical basis of the four-fold classification of economic impact, and used content analysis to analyze the path of first-class universities to serve the regional economic development [18]. The research results show that the specific path of service includes five dimensions: scientific research and innovation, input and output, indirect influence, catalytic influence, and induced influence. These five dimensions each play a unique role in influencing the regional economy and promote each other. At the same time, they also suggest that first-class universities further enhance their service awareness, promote institutional innovation, establish service paths, assess their influence and contribution to regional economic and social development, and regularly release economic impact reports to the society.

The distance coordination degree model to analyze the relationship between higher education and regional economy quantitatively, and conducts an empirical analysis of the temporal and spatial stability of the coordination degree between China's higher education and the regional economy from 2009 to 2018 based on spatial statistics [19]. The results show that: from the perspective of time, the degree of coordination between higher education and regional economy in various provinces is increasing year by year, and has successively experienced four stages of low imbalance, low coordination, moderate coordination and high coordination; from the perspective of space dimension, higher education promotes regional economic development, but also depends on the regional economic level. The three indicators of higher education input, output, and regional economic coordination are stable and agglomerating.

Guangxi is one of the minority autonomous regions in my country, so studying its regional economic development undoubtedly has important practical significance. Looking back at the previous papers, it shows that scholars tend to choose more data and a larger range of research when studying regional economic development, which may lead

to a decrease in the accuracy of the results. On the contrary, they have little research on the data of a certain region, especially Guangxi Province. These shortcomings provide a certain space for further research.

2. Data and Index

This paper discusses the regional economic development of Guangxi based on multivariate statistical analysis, hoping to understand the regional economic development of Guangxi Province and provide policy recommendations for the government.

The data for this paper comes from the Guangxi Provincial Statistical Annual book 2019, including 8 indexes for 14 cities in Guangxi Province in 2018 : GDP(X_1), the GDP of per capital(X_2), Industrial added value(X_3), general public budget revenue(X_4), the balance of various RMB deposits in financial institutions at the end of the year(X_5), the total retail sales of social consumer commodities(X_6), per available income of urban households(X_7), per available income of rural households(X_8). The specific data is listed by Table 1 as follows.

Table 1. Main economic and social statistics of cities in Guangxi Province in 2018

State	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Nanning	4506.56	75172	631.81	326.10	9279.67	1813.20	35276	13654
Liuzhou	3128.35	110059	1288.01	170.26	3171.61	1083.00	34849	13451
Guilin	2105.56	46288	164.78	96.41	2161.07	605.41	34649	14626
Wuzhou	991.4	61095	182.77	55.26	663.91	247.49	31209	12238
Beihai	1300.8	126532	474.37	63.92	820.59	171.27	33954	13998
Fangchenggang	701.23	94510	250.37	35.78	489.51	80.70	34325	14617
Qinzhou	1356.27	47375	100.81	41.84	640.50	234.64	33488	12816
Guigang	1257.53	32463	151.89	35.72	659.35	299.23	30506	13786
Yulin	1679.77	45645	106.93	48.95	815.13	289.65	33960	14984
Baise	1257.78	68889	86.02	19.56	396.41	93.42	30611	11086
Hezhou	700.11	32677	85.11	22.30	467.67	105.06	30864	11548
Hechi	878.1	29859	41.61	15.44	442.33	143.49	27468	9177
Laibing	654.15	33146	73.57	15.23	338.18	87.17	32910	11752
Congzuo	760.46	64421	92.13	9.20	211.40	33.92	30916	12000

3. Method

3.1. The trend and forecast of the regional gross product of Guangxi Province since 2000

The full name of the ARIMA model is the differential autoregressive moving average model, also denoted as ARIMA(p,d,q), and it is one of the commonly used statistical models in time series forecasting. Compared with other time series models, the ARIMA model is very simple. It only needs endogenous variables and does not need to rely on other exogenous variables. It is very suitable for forecasting economic data. Therefore, this paper uses the ARIMA model to conduct a simple empirical analysis firstly.

The data of Guangxi Province's GDP from 2000 to 2018 is originated from the official website of Guangxi Provincial Statistics Bureau. After acquiring the data, we applied R statistical software to build up the time series model to get the ARIMA(0,3,2). At the same time, we try to predict the future 10 years data and portray the time series picture as Fig.1 shows. Fig.1 displays the regional GDP in Guangxi Province constantly increased since 2000, and the overall economic situation is wonderful. It will continue to grow in the next ten years.

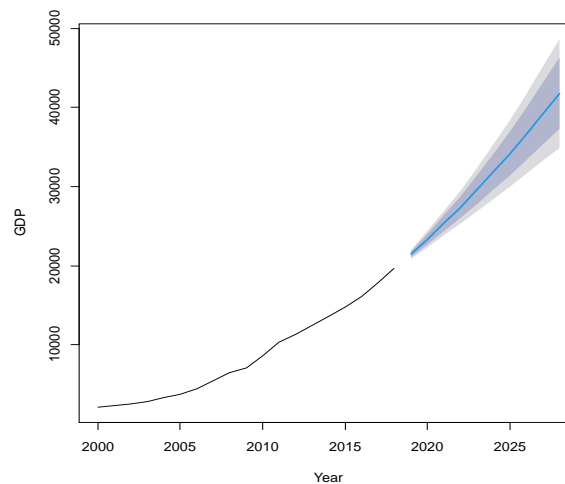


Fig. 1. Example of a figure caption. *(figure caption)*

3.2. Principal component analysis(PCA)

As one of the multiple statistical methods, principal component analysis usually takes the original P indicators as a linear combination to form a new comprehensive index. In short, it is to learn from the dimensionality reduction method to reduce many indicators to a small number of comprehensive indicators, the correlation coefficient between the new comprehensive indicators is 0, and most of the original indicator information is retained.

Therefore, in order to simplify the variable information and better study the regional economic development of Guangxi Province, the paper performs principal component analysis on 8 variables that measure the comprehensive economic level. Firstly, the correlation matrix shows that the 8 indicators are highly correlated and the variables have a strong linear relationship, which is suitable for extracting principal components.

Applying R statistical software to perform PCA shows that it is available to select one or two principal components. The result lists in Fig.2.

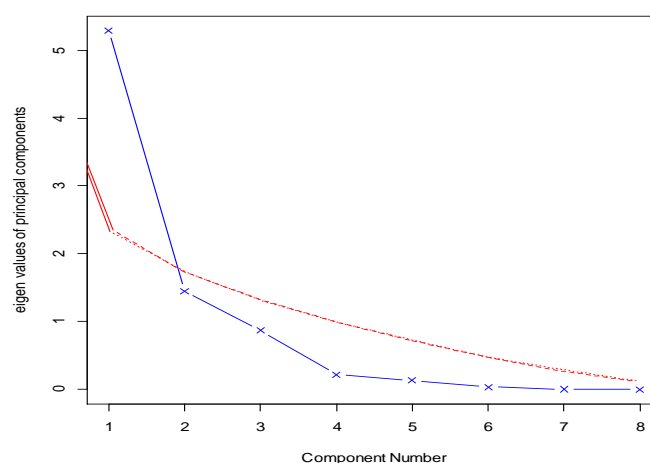


Fig. 2. Parallel Analysis Scree Plot

If we only choose one principal component, we would find that the degree of explanation of this principal component for all variables is only 66%. However, two principal components would achieve 84%. Therefore, this paper select two principal

components to analysis. By calculating component matrix (Table 2), the special vectors of each principal component and the proportion of information extracted from each variable can be verified. So that this paper chooses GDP(X1), Industrial added value(X3), general public budget revenue(X4), balance of various RMB deposits in financial institutions at the end of the year(X5), the total retail sales of social consumer commodities(X6) as the first principal component, which represent the total economic development of various urban areas in Guangxi Province. Meanwhile, this paper selects the GDP of per capital(X2), per available income of urban households(X7), per available income of rural households(X8) as the second principal component, which represent the production level of labor in various urban areas of Guangxi Province.

Table 2. Special Vectors

	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
X ₁	-0.4123	-0.2193	-0.0470	0.0891	0.1225	0.8377	-0.1388	0.1893
X ₂	-0.2331	0.4853	0.5891	-0.5512	0.1536	0.1207	0.0504	-0.1365
X ₃	-0.3471	0.1292	0.5145	0.6967	-0.0378	-0.2388	-0.1890	0.1362
X ₄	-0.4161	-0.2232	-0.0267	-0.1779	0.0333	-0.2720	0.6119	0.5432
X ₅	-0.3897	-0.3231	-0.1011	-0.3728	0.0102	-0.3624	-0.6801	0.0256
X ₆	-0.4081	-0.2798	-0.0467	0.1066	0.0671	-0.0653	0.3212	-0.7936
X ₇	-0.3290	0.4221	-0.3163	-0.0240	-0.7802	0.0557	0.0055	-0.0326
X ₈	-0.2353	0.5376	-0.5224	0.1429	0.5878	-0.1183	-0.0486	0.0211

As the Table2 shows, the main component calculation formula is as follows.

$$F_1 = -0.4123X_1 - 0.2331X_2 - 0.3471X_3 - 0.4161X_4 - 0.3897X_5 - 0.4081X_6 - 0.3290X_7 - 0.5353X_8$$

$$F_2 = -0.2193X_1 + 0.4853X_2 + 0.1292X_3 - 0.2232X_4 - 0.3231X_5 - 0.2798X_6 + 0.4221X_7 + 0.5367X_8$$

Selecting the first two factors as the principal component in this analysis, but the cumulative contribution rate of the two is not equal to one. Therefore, this paper need to apply the adjusted cumulative contribution rate as a weighting factor to calculate as follows.

$$F = 0.53 / (0.53 + 0.31) * F_1 + 0.31 / (0.53 + 0.31) * F_2$$

From this, the comprehensive ranking of Guangxi Province is calculated as shown in Table3.

Table 3. The comprehensive ranking of Guangxi Province

State	F ₁	F ₂	Total F	Rank
Nanning	3.3055	-0.0730	2.0587	1
Liuzhou	1.1530	1.2831	1.2010	2
Guilin	0.2446	0.4345	0.3147	3
Beihai	-0.6260	1.5881	0.1911	4
Yulin	-0.3501	0.5800	-0.0069	5
Fangchenggang	-0.9161	1.3763	-0.0701	6
Qinzhou	-0.2931	-0.0100	-0.1886	7
Wuzhou	-0.2494	-0.3145	-0.2734	8
Guigang	-0.2273	-0.4182	-0.2978	9
Baise	-0.3415	-0.6081	-0.4399	10
Laibin	-0.4819	-0.4805	-0.4814	11
Chongzuo	-0.6126	-0.3117	-0.5016	12
Hezhou	-0.3444	-0.8923	-0.5466	13
Hechi	-0.0093	-2.1538	-0.8007	14

As the result shows, we could find that the cities with positive overall scores are Nanning, Liuzhou, Guilin and Beihai, among which Nanning has the highest overall economic development strength. Nevertheless, the overall scores of other cities are

negative, which allocates that the economic development of these cities is not significant, among which the overall economic development of Hechi is worst.

3.3. Cluster analysis

Firstly, standardizing the obtained data. Then, according to cluster classification standards, dividing the urban areas of Guangxi Province into categories, we could find that the best result is when the number of clusters is 4. Finally, using hierarchical clustering method to draw cluster vertical tree diagram(Fig.3).

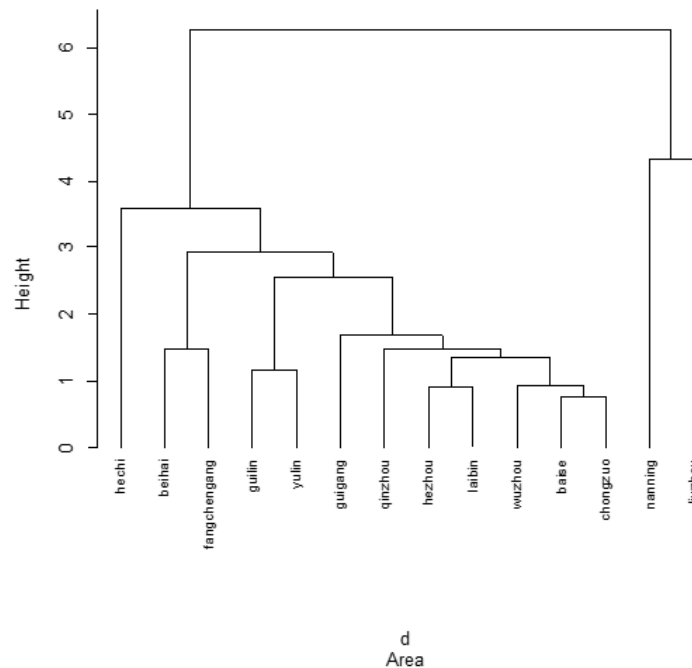


Fig. 3. cluster vertical tree diagram

Therefore, this paper classifies Beihai, Fangchenggang, Guilin, and Yulin into one category, and Hechi City alone. Meanwhile, Guigang, Qinzhou, Hezhou, Laibin, Wuzhou, Baise, Chongzuo are grouped into one category. Finally, Nanning and Liuzhou are also classified into one category.

4. Conclusion

According to the data collected in this paper and the multiple statistic analysis on the regional economic development in Guangxi Province through PCA and cluster methods, we can comfortably come to the following conclusions. On the one hand, it can be seen from the results of the principal component comprehensive score that the overall economic development level of Guangxi Province is relatively unbalanced, and the comprehensive scores of various cities vary greatly. Among them, Nanning and Liuzhou have relatively high comprehensive scores, so their comprehensive economic strength is relatively strong; On the other hand, according to the cluster analysis, we can also find that cities with higher comprehensive scores, such as Nanning and Liuzhou, are classified into one category, and cities with lower comprehensive scores, such as Hechi City, are classified into the same category. Consistent with the results of principal component analysis. Based on the above conclusions, this paper believes that coordinating the regional economic development of Guangxi Province is one of the important contents of the future government work. In the future, the government should make full use of the advantages of various regions, strengthen regional cooperation, encourage regions with higher economic levels to better

promote the development of regions with lower economic levels, strive to narrow the gap in regional economic development, and promote the overall development of Guangxi Province.

References

- [1] G. Myrdal, *Economic Theory and Underdeveloped Regions*. Duckworth, 1957.
- [2] Goldsmith R W, *Financial structure and development*. Yale University Press, 1969.
- [3] A. M. Pereira and M. F. Pinho, "Public investment, economic performance and budgetary consolidation: Var evidence for the first 12 Euro countries," *J. Econ. Dev.*, vol. 36, no. 1, 2011.
- [4] P. Aguighier, "Regional disparity since 1978, in Feuch thtwang Sephen Hussa in A thar and pairault," *Transform. China's Econ. eightiesthe urban Sect.*, 1988.
- [5] G. Y. Long, "China changing region in equalities," *J. Comp. Econ. Rev.*, 2000.
- [6] Fleisher, M. Belton, H. Li, and M. Q. Zhao, "Human Capital, Economic Growth, and Inequality in China," *J. Dev. Econ.*, vol. 92, no. 2, 2009.
- [7] H. Shuyuan and M. Chou, "Multivariate statistical analysis of regional economic development in Gansu Province," *Product. Res.*, 2020.
- [8] Z. Jian, W. Meizheng, and H. Jiayi, "Research on Digital Economy Promoting Regional Economic Development—Taking Heilongjiang Province as an Example," *Econ. Res. Guid.*, pp. 44–46, 2021.
- [9] Y. Li, "Research on the Economic Development of Beijing-Tianjin-Hebei Integration Region Based on Multivariate Statistical Analysis," *Mark. Mod.*, 2021.
- [10] C. Benchang and Z. Xuan, "Economic opening, population flow and regional economic unbalanced development—An empirical analysis based on the fluctuation of economic growth rate in Liaoning Province," *J. Shenyang Univ. Technol. (Social Sci. Ed.)*, 2021.
- [11] C. Jingjing and X. Yongxiang, "Measurement and comparison of the high-quality development level of my country's provincial economy based on the new development concept," *Ind. Technol. Econ.*, 2021.
- [12] J. jie Huang, M. zheng Xiong, Y. Chen, and F. yan Zhao., "E-commerce Product Sales Plan Based on LDA and ARIMA," *Int. J. Comput. Eng.*, vol. 5, no. 3, 2020.
- [13] jingbo Zhang, "Fiscal expenditure, opening to the outside world and high-quality development of regional economy," *Econ. Manag. Dig.*, 2021.
- [14] Z. Zhang, "Analysis of the evolution of my country's regional economic relay growth pattern and its influencing factors," *West. Econ. Manag. Forum*, vol. 32, no. 03, 2021.
- [15] Y. Tan, "Analysis of Regional Economic Differences Based on Cluster Analysis Model—Taking Sichuan Province as an Example," *China Mark.*, 2021.
- [16] J. Tingting and Z. Bin, "Research on the Factors Affecting Regional Economic Differences in Anhui Province—An Empirical Analysis Based on City Panel Data.," *J. Chongqing Univ. Sci. Technol. (Social Sci. Ed.)*, 2021.
- [17] Z. Nannan and L. Xiuyu, "Technological Innovation, Consumption Transformation and Regional Economic Development—Analysis Based on the Perspective of Regional Differences," *Res. World*, 2021.
- [18] W. Yan, C. Rui, and J. Bo, "Research on the path of world-class universities serving regional economic development—Based on the content analysis of the economic impact reports of 10 universities," *Comp. Educ. Res.*, vol. 43, no. 05, 2021.
- [19] H. Qian, "Analysis of the coordination degree between higher education and regional economy," *Stat. Decis.*, vol. 37, no. 09, 2021.