International Journal of Multidisciplinary Research and Publications
ISSN (Online): 2581-6187

Research on the Path of New Quality Productivity Based on Entropy K-Means Clustering-- Take the Yangtze River Economic Belt as an Example

Yichen Zhou
School of Management Science and Engineering, Anhui University of Finance and Economics
Email address: 1754694223@qq.com

***Abstract*—*With the introduction of the important concept of "new quality productivity", many scholars have carried out research from different angles. By manually collecting the data of China Statistical Yearbook, this paper uses entropy method, K-Means algorithm and random forest to conduct comprehensive evaluation and cluster analysis of 11 provinces and cities along the Yangtze River Economic Belt based on the development level of new quality productivity. The results show that the clustering results of the development level of new quality productivity in 11 provinces and cities along the Yangtze River Economic Belt are as follows: Shanghai, Anhui, Jiangxi, Hubei, Hunan, Chongqing and Sichuan are grouped into one category; Jiangsu and Zhejiang provinces are grouped into one group; Yunnan and Guizhou provinces are grouped into one group. Among them, scientific and technological productivity and economic productivity are more important. This result accords with the expectation, and has certain practical significance and potential research value*.

***Keywords*—** *New quality productivity, Entropy method, K-Means clustering, Random forest, Lifting path*.

I. INTRODUCTION

In September 2023, during his inspection and research in Heilongjiang Province, General Secretary Xi Jinping first proposed the important concept of "new quality of productivity." General Secretary Xi Jinping further underlined the necessity of actively cultivating future industries, accelerating the formation of new productivity qualities, and enhancing the development of new momentum at the symposium on promoting the comprehensive revitalization of Northeast China in the new era. These emerging industries include new energy, new materials, advanced manufacturing, and electronic information. As one of the key driving forces for contemporary economic development, the new quality of productivity is of great significance for promoting economic growth, promoting industrial upgrading, enhancing innovation capabilities, and achieving sustainable development:In terms of promoting economic growth, the continuous emergence and application of new qualities of productivity can effectively enhance production efficiency, accelerate economic growth rates, and drive sustained and healthy economic development. Regarding promoting industrial upgrading, the introduction and utilization of new qualities of productivity can facilitate the transformation and upgrading of traditional industries, prompting enterprises to increase technological investment and management innovation, improve industrial structure, and enhance overall economic competitiveness. In terms of enhancing innovation capabilities, the research and application of new qualities of productivity require ongoing technological and management innovation, further enhancing the innovation capabilities of enterprises and nations, and driving technological progress and industrial revolution. Concerning achieving sustainable development, the development of new qualities of productivity facilitates the efficient utilization of resources and environmental protection, promotes sustainable economic development, and fosters the coordinated development of the economy, society, and environment.

Currently, research on new qualities of productivity is in a vigorous stage of development. Scholars are delving deeply into the concept and connotations of new qualities of productivity, gradually refining the relevant theoretical framework. In terms of empirical research, a plethora of academic achievements are dedicated to exploring the impact of new qualities of productivity on economic growth, industrial upgrading, innovation capabilities, and employment levels, providing a solid practical foundation for this field. Concurrently, scholars have begun to focus on the interaction between new qualities of productivity and institutional environments, policy systems, strengthening research on the development of new qualities of productivity and policy support. However, there is relatively less focus on the development of new qualities of productivity itself and related optimization paths. Currently, many scholars have conducted research on new qualities of productivity from different perspectives. Wen Feng'an et al. [1] believe that the formation of new qualities of productivity is the source of driving modernization in other fields, a crucial force for promoting high-quality economic and social development, a key strategy for addressing changes in the primary contradictions of society in the new era, and a critical link for digital integration and promoting future and emerging industries. Yao Shujie, et al. [2] argue that clarifying the interaction between the digital economy and new qualities of productivity is an important theoretical and practical innovation for developing new qualities of productivity and achieving Chinese-style modernization. Additionally, technological innovation, institutional optimization, and factor synergy are important pathways to promote the development of new qualities of productivity through leveraging the digital economy. Zhang Xiaheng, et al. [3] suggest that efforts should be made to

Yichen Zhou, "Research on the Path of New Quality Productivity Based on Entropy K-Means Clustering-- Take the Yangtze River Economic Belt as an Example," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 6, Issue 12, pp. 144-147, 2024.

strengthen the development of data factor markets, promote the marketization of data factors, accelerate the construction of data rules and regulations, and expedite the cultivation of high-quality data talent, thereby fully unleashing the enabling effect of data factors to facilitate the generation and development of new qualities of productivity. It can be seen that most of the current research on new qualities of productivity is still at the stage of policy conception, lacking robust quantitative analysis. Therefore, this paper aims to quantitatively study the development level and enhancement paths of new qualities of productivity in the eleven provinces and cities in the Yangtze River Economic Belt using the entropy method and the K-Means clustering algorithm in machine learning.

II. Index System Construction

A. Construction of Evaluation Index System for the Development Level of New Quality Productivity

New quality productivity refers to a new combination of production factors and modes centered around innovation, driven by new technologies such as digitization, intelligence, and networking. It enables the fulfillment of increasingly diverse and personalized demands in a more efficient, flexible, and intelligent manner, thereby promoting economic structural transformation and upgrading, as well as enhancing production efficiency and the quality of economic development. The focus of new quality productivity development lies in technological innovation, which involves the extensive application of new technologies including artificial intelligence, big data, and the Internet of Things to improve production efficiency and product quality. The challenges in the development of new quality productivity include industrial transformation and upgrading, as well as structural adjustment. Overcoming entrenched interests and path dependencies associated with industrial transformation and upgrading, and promoting traditional industries to transition towards environmentally friendly directions are crucial areas of research. Therefore, when constructing the evaluation index system for the development level of new quality productivity in the Yangtze River Economic Belt, it should comprehensively cover aspects such as technological productivity, green productivity, economic productivity, and digital productivity. The development of new quality productivity is reflected in research and development efforts and achievements in technological productivity, pollutant emissions in waste gas and waste water for green productivity, regional gross domestic product and significant enterprise revenue for economic productivity, and the development and income of information technology-related services for digital productivity. In addition to comprehensiveness, when constructing the index system for the development level of new quality productivity, attention should also be paid to the specific research area—the Yangtze River Economic Belt. It is necessary to adhere to the development characteristics of the Yangtze River Economic Belt, focusing on ecological and resource consumption considerations while enhancing the development level of new quality productivity, in line with the principle of sustainable development. At the same time, attention should be paid to the development differences between the upper, middle, and lower reaches of the Yangtze River Economic Belt, planning for more optimal development paths.

B. Construction of Evaluation Index System for the Development Level of New Quality Productivity

The selection of evaluation indicators for the development level of new quality productivity needs to comprehensively cover the development characteristics of various provinces and cities in the Yangtze River Economic Belt, including aspects such as technological productivity, green productivity, economic productivity, and digital productivity. Based on this, this article selects 9 indicators: Full-time Equivalent (FTE) Research and Development (R&D) personnel (hours), effective invention patents (pieces), nitrogen emissions in wastewater per unit of regional gross domestic product (tons/10,000 yuan), sulfur dioxide emissions in exhaust gas per unit of regional gross domestic product (tons/10,000 yuan), regional gross domestic product (100 million yuan), operating income of industrial enterprises above a certain scale (100 million yuan), total telecommunications business volume (100 million yuan), number of mobile phone base stations (10,000), and information technology service revenue (10,000 yuan). FTE R&D personnel and effective invention patents reflect technological productivity; nitrogen emissions in wastewater per unit of regional gross domestic product and sulfur dioxide emissions in exhaust gas per unit of regional gross domestic product reflect green productivity; regional gross domestic product and operating income of industrial enterprises above a certain scale reflect economic productivity; total telecommunications business volume, number of mobile phone base stations, and information technology service revenue reflect digital productivity.

C. Variable Definition

The comprehensive evaluation indicator system constructed in this paper consists of 9 indicators, defined as follows: Full-time Equivalent (FTE) Research and Development (R&D) personnel (x1), effective invention patents (x2), nitrogen emissions in wastewater per unit of regional gross domestic product (x3), sulfur dioxide emissions in exhaust gas per unit of regional gross domestic product (x4), regional gross domestic product (x5), operating income of industrial enterprises above a certain scale (x6), total telecommunications business volume (x7), number of mobile phone base stations (x8), and information technology service revenue (x9).

D. Data sources

The original data of this article is sourced from the China Statistical Yearbook, which manually collected nine relevant indicator data for evaluating the development level of new quality productivity in eleven provinces and cities along the Yangtze River Economic Belt, including Shanghai, Jiangsu, Zhejiang, Anhui, Jiangxi, Hubei, Hunan, Chongqing, Sichuan, Yunnan, and Guizhou from 2018 to 2022.

Yichen Zhou, "Research on the Path of New Quality Productivity Based on Entropy K-Means Clustering-- Take the Yangtze River Economic Belt as an Example," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 6, Issue 12, pp. 144-147, 2024.

III. Method Introduction

A. Basic Theory of Entropy Method

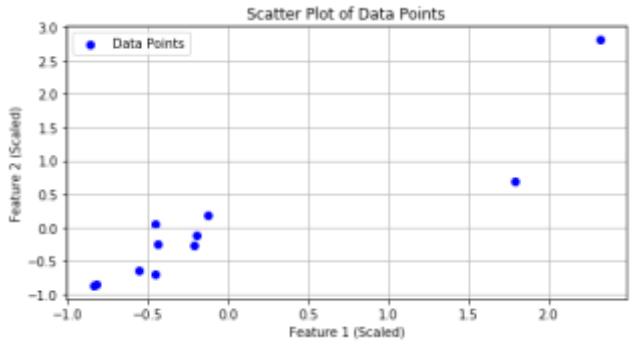
The entropy method is a multicriteria decision-making approach used to evaluate the overall merits of different solutions. Its basic principle is as follows: First, identify decision criteria and quantify the weights of each criterion. Then, assess each solution on each criterion to obtain an evaluation matrix. Next, calculate the entropy value for each solution on each criterion, and subsequently compute the comprehensive entropy value for each solution. Finally, compare the comprehensive entropy values of the solutions to determine the optimal one. Additionally, the linear weighting method can be employed to calculate the overall score based on the calculated weights.

B. Basic Theory of K-Means Algorithm

The K-Means algorithm is a commonly used unsupervised learning algorithm for partitioning a dataset into K distinct groups or clusters. Its basic principle is as follows: Initialization: Choose K random data points as initial cluster centroids. Assignment of data points: For each data point, calculate its distance to the K cluster centroids and assign it to the cluster corresponding to the nearest centroid. Update of cluster centroids: For each cluster, calculate the average of all data points and use this average as the new cluster centroid. Iterative repetition: Repeat steps 2 and 3 until the cluster centroids no longer change or until a predetermined number of iterations is reached. Convergence: The algorithm converges to a set of cluster centroids, dividing the dataset into K clusters.

IV. Analysis Process

Firstly, present a scatter plot of the dataset as a whole:



A. K-means clustering

In K-Means clustering analysis, different classification results are obtained by adjusting the k value:

When k=3:

From the results, it can be seen that Shanghai, Anhui, Jiangxi, Hubei, Hunan, Chongqing, and Sichuan are relatively close in terms of the development level of new quality productivity, and are grouped together; Jiangsu and Zhejiang are grouped together; Yunnan and Guizhou are grouped together.

K-Means Clustering Results:

	province	kmeans_cluster
0	Shanghai	0
1	Jiangsu	1
2	Zhejiang	1
3	Anhui	0
4	Jiangxi	0
5	Hubei	0
6	Hunan	0
7	Chongqing	0
8	Sichuan	0
9	Yunnan	2
10	Guizhou	2

When k=4:

K-Means Clustering Results:

	province	kmeans_cluster
0	Shanghai	2
1	Jiangsu	0
2	Zhejiang	0
3	Anhui	3
4	Jiangxi	3
5	Hubei	3
6	Hunan	3
7	Chongqing	3
8	Sichuan	3
9	Yunnan	1
10	Guizhou	1

Compared with the results when k=3 mentioned above, the difference is that when k=4, Shanghai is classified separately.

B. Random Forest Analysis

This article calculates the comprehensive evaluation scores of the development level of new quality productivity for eleven provinces and cities in the Yangtze River Economic Belt from 2018 to 2022 using the entropy method. Using the selected 9 indicators as x and the comprehensive score as y, the random forest method was introduced to study the importance of each indicator on the development level of new quality productivity. The results are as follows:

	Feature	Importance
0	x1	0.296862
1	x2	0.075752
2	x3	0.012140
3	x4	0.005334
4	x5	0.099361
5	x6	0.333443
6	x7	0.001980
7	x8	0.116540
8	x9	0.058588

Yichen Zhou, “Research on the Path of New Quality Productivity Based on Entropy K-Means Clustering-- Take the Yangtze River Economic Belt as an Example,” *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 6, Issue 12, pp. 144-147, 2024.

V. CONCLUSION

According to the comprehensive score calculated by the entropy analysis method, it can be seen that from 2018 to 2022, the comprehensive score of the development level of new quality productivity in various provinces and cities in the Yangtze River Economic Belt has been increasing, indicating that the new quality productivity of each province and city in the Yangtze River Economic Belt has gradually improved. And the overall comprehensive score shows an increasingly fast growth trend, reflecting the strong development strength of new quality productivity in the Yangtze River Economic Belt and its potential for sustained growth.

According to the K-Means clustering analysis results, the clustering results of the development level of new productive forces in the eleven provinces and cities of the Yangtze River Economic Belt are: Shanghai, Anhui, Jiangxi, Hubei, Hunan, Chongqing, and Sichuan, seven provinces and cities are clustered into one category; Jiangsu and Zhejiang provinces are grouped together; Yunnan and Guizhou provinces and cities are grouped together. This result is consistent with the predicted results before the experiment, and also conforms to the regional distribution of eleven provinces and cities in the Yangtze River Economic Belt. Shanghai, Anhui, Jiangxi, Hubei, Hunan, Chongqing, and Sichuan are seven provinces and cities located in the middle reaches of the Yangtze River Economic Belt; The lower sections of the Yangtze River Economic Belt are home to the provinces of Jiangsu and Zhejiang, while the upper reaches are occupied by the cities and provinces of Yunnan and Guizhou. Based on comprehensive scores, the lower reaches of the Yangtze River Economic Belt have the highest level of development in new quality productivity. Therefore, the development of new productive forces cannot be separated from economic growth, progress in technological productivity, and improvement in digital productivity. In the future development, provinces and cities in the Yangtze River Economic Belt should learn from each other's development strategies based on the clustering results. Regions located in the middle and upper reaches should learn from the development experience of downstream regions and jointly contribute to the improvement of the new quality productivity level in the Yangtze River Economic Belt.

According to the results of random forest analysis, among the 9 selected evaluation indicators, the two indicators representing scientific and technological productivity and the two indicators representing economic productivity have the highest importance. Therefore, provinces and cities in the Yangtze River Economic Belt should focus their development on science and technology and economic productivity, and increase investment in scientific and technological production.

ACKNOWLEDGEMENT

This work is supported by Anhui University of Finance and Economic Undergraduate Scientific Research Innovation Fund Project under Grant No. XSKY24150.

REFERENCES

[1] Wen Feng'an, Xiao Huajian Exploring the Promotion of High Quality Economic and Social Development by New Productive Forces [J/OL] Journal of Yantai University (Philosoph and Social Sciences Edition), 1-11 [2024-04-22]

[2] Yao Shujie, Wang Jiefei The Theoretical Logic and Implementation Path of Digital Economy Promoting the Development of New Quality Productivity [J/OL] Journal of Yantai University (Philosophy and Social Sciences Edition), 1-12 [2024-04-22]

[3] Zhang Xiaheng, Liu Caixia Research on the intrinsic mechanism and path of data elements promoting the realization of new quality productivity [J/OL] Industrial Economic Review, 1-21 [2024-04-22]

Yichen Zhou, “Research on the Path of New Quality Productivity Based on Entropy K-Means Clustering-- Take the Yangtze River Economic Belt as an Example,” *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 6, Issue 12, pp. 144-147, 2024.