Quality Evaluation of College Students Based on Moran Index Analysis

Liang Ye [1, a], Shun Ding [2, b] and Jin Lu [3, c, *]

[1]*Office of Admissions and Career Services, Shanghai University, Shanghai, China*
[2]*School of Computer Engineering and Science, Shanghai University, Shanghai, China*
[3]*Office of Admissions and Career Services, Shanghai University, Shanghai, China*
[a]*liangye@shu.edu.cn*
[b]*dingsh@shu.edu.cn*
[c]*lujin221@t.shu.edu.cn*

Abstract

Current studies on student quality are mainly based on college entrance examination results or rankings, without considering the differences in exam content and difficulty, which makes it difficult to compare matriculate quality horizontally and longitudinally. To solve the above problems, this paper proposes a student quality evaluation model based on the provincial ranking of college entrance examination and the number of online students in provincial control. The model firstly proposes a formula to calculate students' comprehensive quality, and then constructs a spatial weight matrix to introduce global and local Moran indexes to comprehensively analyze the spatial autocorrelation of students' quality. Compared to the mainstream student evaluation system, the Moran index can analysis the quality of students in adjacent provinces which gathered distribution on hidden links. It also can reflect the concentration trend of group students' quality compared by horizontal and vertical comparison. The method can provide college enrollment institution some targeted for decision support.

Keywords: *Moran index; evaluation model; matriculate quality;*

1 INTRODUCTION

Enrollment is the beginning of talent training in colleges and universities. The quality of students directly affects the quality of teaching and talent training in colleges and universities [3]. In 2019, 10.31 million students took the national college entrance examination, 5.74 percent more than in 2018, while 4.31 million students took regular undergraduate courses in 2019, 2.16 percent more than in 2018. China's higher education is developing at an unprecedented speed, the competition among universities is becoming increasingly fierce, and the recruitment of college students has attracted more attention than before.

The previous studies on the quality of students mainly focus on three levels. One is the study on the quality evaluation of students. Because colleges and universities in China mainly select new students through the college entrance examination, and there are horizontal differences between regions and vertical differences between years, it is difficult to reach a consensus on how to evaluate the quality of students. Therefore, the study on the quality of students is still a hot spot in the current higher education research. In China, there are mainly the following studies on student quality evaluation: The first type mainly measures the quality of students through the scores of college entrance examination [5][10], such as standardizing the scores of college entrance examination according to the highest scores in the province or standardizing the admission line difference of each university in a province to evaluate the quality of students [1] [11]. In practice, the evaluation system of WuShuLian and Soft science mainly analyzes the source of undergraduate students by taking a normalized average of the admission score of arts and science. The second category is to use ranking data to measure the quality of students [6]; the third category is to measure the overall quality of college students according to the enrollment index, such as the ratio of enrollment to enrollment number [2] and enrollment rate [17], adjustment rate [16]. The above studies have their own advantages in the evaluation of student quality, but they mainly focus on the horizontal comparison of student quality. On the

B. Fox et al. (Eds.): IC-ICAIE 2022, AHCS 9, pp. 1022-1029, 2023.
https://doi.org/10.2991/978-94-6463-040-4_154

basis of the changes of test difficulty in different years, it is difficult to evaluate student quality scientifically and accurately through longitudinal comparison. The second is the study on the influencing factors of student quality. Based on previous studies, there are four main factors affecting the quality of students: Comprehensive conditions of colleges and universities [7], efforts of recruitment publicity [4], measures of talent cultivation and employment prospects [8]. This paper analyzes the main factors affecting the quality of students and puts forward the targeted intervention measures, so as to lay a solid foundation for improving the quality of students. The third is the evaluation of the impact of education policy on the quality of students. Education policy evaluation is an inevitable requirement and an important starting point for promoting the modernization of education governance capacity [12]. The rationality of the policy itself and the implementation effect of the policy need to be verified. For example, Reference (Wen, 2018) evaluated the impact of the "985" Project on the quality of college students by using two-way fixed effect model and event research estimation and other causal inference methods [13].

At present, although there have been a lot of research results on the quality of college students, because of the planned enrollment of colleges and universities in China, there are differences in the quality of students in different provinces, there are few studies on the spatial distribution characteristics and laws of the quality of students. Tobler's First Law of geography states that geographical objects or properties are spatially related to each other. Generally speaking, the closer the distance, the greater the correlation between ground objects; The farther the distance, the greater the dissimilarity between ground objects [9]. Spatial autocorrelation analysis has been widely studied in the field of economy [14]. The economies of different regions interact with each other, but education and economic development are inseparable. Therefore, there is a certain relationship between the quality of students in different regions in theory. Reference [15] analyzed the spatial pattern of the source area and studied the spatial distribution of the number of students by using spatial autocorrelation analysis, which has certain guiding significance to study the regional influence of colleges and universities and formulate targeted policies, but the quantitative analysis of the spatial distribution of student quality needs further research.

On the basis of the current research, this paper puts forward an evaluation model of student quality, aiming to compare student quality horizontally and longitudinally. Taking the enrollment data of UNIVERSITY A as an example, empirical analysis is conducted to study the advancement of the quality evaluation model of the student source. Then the spatial autocorrelation analysis is used to study the spatial distribution characteristics and rules of the quality of college students, so as to promote the rational allocation of enrollment resources and provide decision-making basis for enrollment, which is of great significance to improve the quality of college students.

2 EVALUATION MODELS AND DATA SOURCES

2.1 Data Sources

This study takes the enrollment data of UNIVERSITY A in 2018-2019 as the research object, obtains the data of students' college entrance examination results from the recruitment and employment department of the university, and obtains the corresponding data of college entrance examination ranking from the one-section table of each province published by China online Education. Coordinates and boundary data of Chinese and provincial administrative regions were collected from AmAP.

2.2 Student Quality Evaluation Model

The student quality evaluation model proposed in this paper is mainly composed of two parts: the first part calculates the individual student quality of each province, and the second part accumulates the individual student quality of each province.

Individual student quality q represents the student quality of a certain student. Based on the number of students online in the corresponding batch in the province where the student lives, the student's student quality is measured according to its position in the batch. The calculation method is as follows:

$$q_i = \left(1 - \frac{r_i}{M}\right) \times 100 \tag{1}$$

Among them, q_i represents the individual student quality score of student i, r_i represents the ranking of student i in his province, and M represents the number of online students in his province.

Q represents the comprehensive student quality of a group, is the average of all individual student quality of the group, and can reflect the central trend of student quality of the group. Its calculation method is as follows:

$$\mathrm{Q} = \frac{\sum q_i}{N} \tag{2}$$

Among them, N represents the number of students in this group.

This model has three main characteristics:(1) Taking ranking as the evaluation index of student quality is better than college entrance examination scores. Because the number of people in each score segment is different, the ranking may differ greatly by one point, and the score measurement does not take into account the possibility that there may be a large number of competitors. (2) Taking the number of students on provincial control line as the baseline, the proportion of ranking can be calculated to realize the horizontal and vertical comparison of student quality. (3) The individual student quality of the population was weighted and averaged as the comprehensive student quality of the population, which reflected the centralized trend of the student quality of the population and made it possible to compare the student quality of the population horizontally and longitudinally.

After calculating the comprehensive quality of students, this paper constructs the spatial weight matrix, and calculates the global Moran index and local Moran index respectively, comprehensively carries on the spatial autocorrelation analysis to the quality of students. In addition, the spatial relationship of student quality is visualized by constructing Moran index scatter diagram and LISA agglomeration graph, to more accurately dig out the hidden relationship of student distribution and reflect the spatial distribution law of student quality.

3 INSTANCE ANALYSIS

3.1 Comparative Analysis of Regional Student Quality in Different Years

Based on the enrollment data of A University in each province from 2018 to 2019, the advancement of the student source quality model proposed in this paper is analyzed. There is a model to measure the quality of students among the many university evaluation indexes of WuShuLian and Soft Science. This model only calculates the scores of college entrance examination, takes few factors into consideration, and does not have the ability of longitudinal evaluation. The model in this paper takes the provincial ranking as the starting point and carries out the class normalization calculation according to the ranking of admission line. In addition, it is assumed that the rankings of the three loci of provincial admission line, median score and highest score of the university represent the quality of students admitted by the target universities at high, middle and low levels respectively, then, the difference between the student quality measured by the above indicators and the student quality calculated by the model in 2018-2019 was compared, and the change trend of student quality represented by each indicator was reflected by the difference, so as to analyze the rationality of the student quality model in this paper. If the difference value is positive, it means that the quality of students measured by this index is getting better; if it is negative, it means that the quality of students measured by this index is getting worse; if it is 0, it means there is no change.

As shown in Fig. 1, in terms of the enrollment data of A University from 2018 to 2019, the changing trend of student quality calculated by the model in this paper is consistent with the changing trend of admission line ranking, median ranking and highest score ranking in most provinces, but there is still a slight difference. For example, the ranking of the admission line in Shandong and Henan decreases, while the ranking of the median score and the highest score increases, and the student quality score calculated by the model in this paper improves, indicating that the high score group admitted in 2019 is conducive to improving the student quality of these two provinces. In Jiangsu, Qinghai and Guizhou, the ranking of the admission line and the highest score decreased, while the ranking of the median score increased, and the quality score of the students calculated by the model improved, indicating that the number of high-stratified students enrolled in 2019 increased, further revealing the quality of the students in these provinces. In Hunan, Sichuan and Jilin, the median ranking fell, and the line with the highest ranking increase, explain skewness distribution degree graduate matriculate quality is bigger, the median ranking representative is not strong, but in fact these provinces matriculate quality is rising, and the improvement of student quality calculated by the model in this paper can reflect this upward trend. In Guangxi, Fujian and Jiangxi, the ranking of the highest score decreased, while the ranking of the other two loci increased, indicating that the quality of low-stratified students was improved and the overall quality of students was further improved, which can be reflected in the model proposed in this paper. In addition, in Qinghai, Hunan and Jiangxi provinces, the change range of admission line, median score and highest score rank is small, while the change range of student quality calculated by the model in this paper is large. This is due to an increase in the number of people who went online in 2019, Therefore, in the case of similar admission rankings in 2018 and 2019, the actual increase of the quality of the university's students is greater than the change in the rankings, and according to the analysis of the questionnaire after admission, the real performance of the students after admission is closer to the model in this paper, so the quality of the students calculated by the model in this paper is more advanced and reasonable.

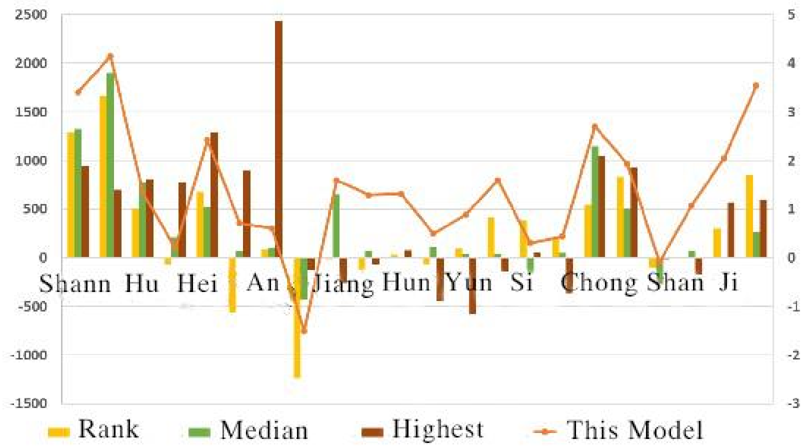

Figure 1: A comparison between the change of admission ranking in each province from 2018 to 2019 and the change of student quality calculated by the model in this paper

3.2 Based on the comparative analysis of regional student quality of a university from 2018 to 2019

As shown in Fig. 2, in terms of the distribution of student quality in each province, the distribution of student quality in Beijing, Jilin, Chongqing and Qinghai is relatively scattered, while that in Hebei, Shandong and Henan is relatively concentrated. Longitudinally, Shaanxi, Liaoning and Heilongjiang provinces showed significant differences in the distribution of student quality in 2018 and 2019. The lowest and highest scores of student quality increased in several other provinces except Inner Mongolia, so the difference was positive.

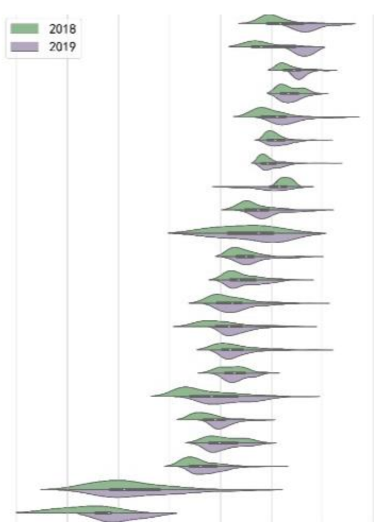

Figure 2: Quality distribution of students in A University by province in 2018-2019

As shown in Fig. 3, from the perspective of the quality of enrollment sources in each province, the quality of students enrolled in Shaanxi, Liaoning and Hebei is higher, while the quality of students enrolled in Beijing, Jilin and Jiangxi is lower. In 2019, the university saw an improvement in the quality of students in most provinces, notably Shaanxi, Liaoning, Heilongjiang and Chongqing, while the quality of students in Inner Mongolia and Shanxi declined to varying degrees. Therefore, colleges and universities in Inner Mongolia, Shanxi, Beijing and Jilin should be especially strengthened.

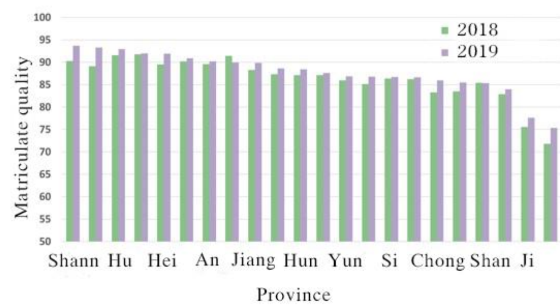

Figure 3: Enrollment quality of A University in each province from 2018 to 2019

3.3 Global Spatial Autocorrelation Analysis

3.3.1 Construct spatial weight matrix

Before the spatial autocorrelation analysis of student quality is carried out, the spatial weight matrix is constructed to define the space between the regions where students are located, to describe the correlation degree between the regions of students and provide a basis for the subsequent global and local spatial autocorrelation analysis. The definition expression of spatial weight matrix W is as follows:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \quad (3)$$

where, w_{ij} represents the spatial distance between region i and j, and w_{ij} satisfies:

$$w_{ij} = \begin{cases} 1, & region\ i\ is\ adjacent\ to\ j \\ 0, & region\ i\ is\ not\ adjacent\ to\ j \end{cases} \quad (4)$$

when $i = j$, $w_{ij} = 0$, we know that the matrix W is a symmetric matrix, and $W^T = W$.

3.3.2 Global Moran index calculation method

The global Moran index is mainly used to measure the average correlation degree and difference degree of each region in space as a whole. The calculation formula is constructed as follows:

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S^2(\sum_i \sum_j w_{ij})} \quad (5)$$

where n is the number of samples and S^2 is the sample variance. The value range of the global Moran index is generally between -1 and 1. If the value is greater than 0, it indicates that the overall positive correlation exists, and there is correlation phenomenon between high values and high values, and low values and low values. If the value is less than 0, the overall negative correlation exists, and there is a difference between the high value and the low value. If the value is close to 0, it indicates that the spatial distribution is random and there is no spatial correlation.

3.3.3 Global spatial autocorrelation analysis of student quality

The global Moran's I index of student quality of 22 provinces in 2018-2019 calculated according to the above formula is shown in Table I. In 2018, the global Moran's I index of student quality was -0.179, and the Z value was -0.929, with the absolute value less than 1.96 (significance level was 0.05). The results showed that the global Moran's I index of student quality in 2018 did not pass the significance test, there was no spatial autocorrelation, and the global distribution was discrete. In 2019, the global Moran's I index of student quality was -0.241, and the absolute value of its Z value was 2.340 greater than 1.96 (significant level was 0.05), indicating that the distribution of student quality in 2019 showed a significant global spatial negative correlation, presenting a clustering distribution.

Table 1: Student Quality Global Moran's Index.

Year	Moran's I	Z	Spatial autocorrelation
2018	-0.179	-0.929	Yes
2019	-0.241	-2.340	No

3.3.4 Local Spatial Autocorrelation Analysis

According to the above results of spatial global autocorrelation analysis, it can be concluded that the quality of students in 2018 is globally spatially irrelevant. This section will further discuss whether local spatial autocorrelation exists. In addition, the specific location and clustering mode of the spatial agglomeration of student quality in 2019 will be analyzed. In this section, local Moran's I is used for local spatial autocorrelation analysis, aiming to further explore the agglomeration distribution of student quality through visualization of Moran's I scatter diagram and LISA agglomeration diagram

1) Local Moran index calculation method

The construction of local Moran index is mainly used to help determine the types and locations of spatial aggregation, so as to make up for the deficiency of global Moran index. The calculation formula is constructed as follows:

$$I = \frac{(x_i - \bar{x}) \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})}{S^2} \quad (6)$$

The meaning of each symbol is the same as the formula of global Moran index.

2) Local spatial autocorrelation analysis of student quality

As shown in Fig. 4, student quality in 2018 is mainly distributed in the first, second and fourth quadrants of Moran's I scatter plot, indicating that the student quality is spatially concentrated in high-high, low-high and high-low types. When the significance level is 0.05, it can be seen from the LISA agglomeration diagram (Fig. 6) and agglomeration results (Table II) of the local spatial relationship of student quality in 2018, the high-value cluster centers are Jiangsu, Shandong and Henan, which are the three provinces with the most difficult college entrance examination in China, with high quality of students. The low-value outliers are Beijing and Shanxi, and there are no low-value cluster centers and high-value outliers. In conclusion, although there is no global spatial correlation of student quality in 2018, there is local spatial autocorrelation.

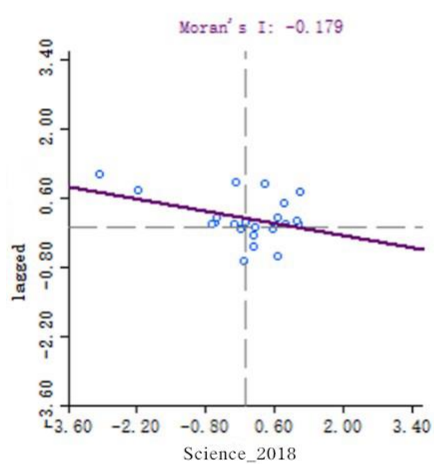

Figure 4: Moran's I scatter chart of student quality in 2018

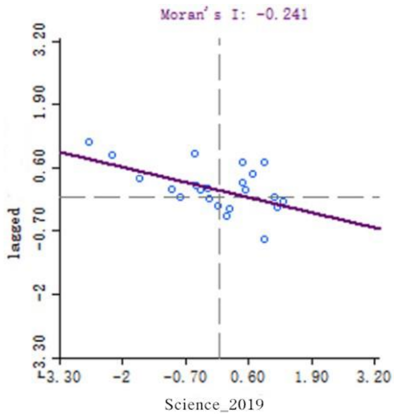

Figure 5: Moran's I scatter chart of student quality in 2019

As shown in Fig. 5, the student quality in 2019 is mainly distributed in the first, second and fourth quadrants of Moran's I scatter plot, and the student quality is mainly spatially clustered in high-high, low-high and high-low patterns. LISA cluster diagram (Fig. 7) and cluster results (Table II) of local spatial relationship of student quality in 2019 show that the high-value cluster centers are the same as those in 2018, mainly distributed in East China, without low-value cluster centers. The low-value outliers are Beijing, Jilin and Shanxi, while the high-value outliers are Heilongjiang.

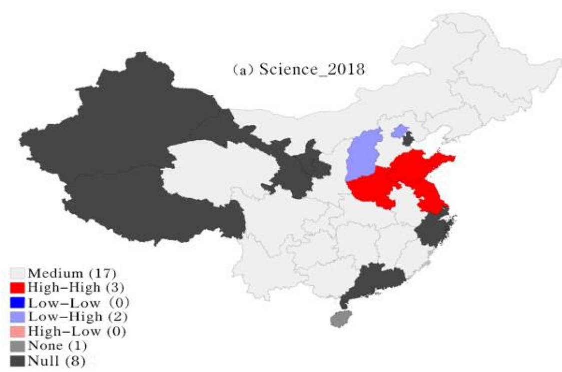

Figure 6: LISA cluster map of student quality in 2018

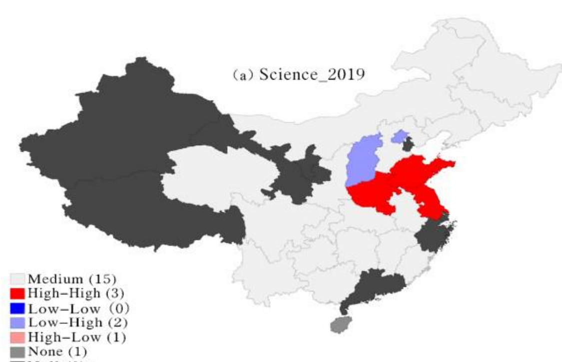

Figure 7: LISA cluster map of student quality in 2019

Table 2: Student Quality Lisa Cluster Results.

Cluster type	2018	2019
High-high (high-value cluster center)	Henan, Jiangsu, Shandong	Henan, Jiangsu, Shandong
Low-low (Low-value cluster center)	none	none
Low-high (low value outlier)	Beijing, Shanxi	Beijing, Jilin, Shanxi
High-low (high value outlier)	none	Heilongjiang

From the above analysis, it can be seen that in 2019, there were two newly added low-value outliers and one newly added high-value outliers, and their spatial autocorrelation was stronger. It can be seen from the above description that the application of the student quality evaluation model in this paper on moran index can obtain the distribution characteristics of student quality, find low aggregation of high and low student quality, and make a more scientific comparison of longitudinal student quality. On this basis, the analysis of recruitment work will be more conducive to scientific allocation of recruitment resources.

4 CONCLUSIONS

In view of the problem that existing researches cannot compare the horizontal differences between regions and longitudinal differences between years in exam content and difficulty, this paper proposes a student quality evaluation model, and analyzes the advanced nature of the model by taking the actual enrollment data of A University in 2018-2019 as the research object. Using spatial autocorrelation analysis, the spatial distribution characteristics and rules of student quality are found. Through the research, the following conclusions can be drawn:

The individual ranking was normalized according to the number of online students in provincial control, and the central trend of student quality was used to represent the overall student quality, considering both the change of student base and the distribution of the overall student quality. Therefore, the evaluation results generated by the student quality evaluation model proposed in this paper can more scientifically calculate the longitudinal results and aggregation distribution of the student quality comparison, which has a certain practical significance for college enrollment and talent training.

Through the global and local spatial autocorrelation analysis of student quality of target universities, it can be concluded that there is no global correlation of student quality in 2018, but there is local spatial autocorrelation. The global positive correlation of student quality in 2019 is consistent with the spatial influence of local Moran's I scatter plot.

Through the local spatial autocorrelation analysis of student quality of target universities, we can identify the high value cluster center, low value cluster center, low value isolated point and high value isolated point of student quality. It is necessary to give full play to the "leading role" of high-value agglomeration centers to improve the quality of students in surrounding areas and strengthen the recruitment publicity of low-value isolated points and low-value cluster centers to attract high-quality students.

It can be seen from the results of calculation of Moran index and the visualized graph results of Moran index by constructing the spatial weight matrix that the application of Moran index in student quality evaluation can accurately analyze the law and distribution characteristics of student quality in spatial level. Compared with the mainstream student evaluation system, the application of Moran index can compare students' quality more scientifically and provide targeted decision support for college enrollment.

REFERENCES

[1] Bei, D. Ye, B. (2015). Quality of undergraduate students in "Project 985": A statistical analysis based on 2011-2012 Data. Tsinghua University Education Research, 36(03): 27-38+75.

[2] Chen, B. Wang, (2015). Analysis and reflection on the enrollment quality of clinical medicine undergraduates in Beijing. China Health Talents, (6): 84-88.

[3] Guo, C. (2017). Analysis on the application of educational big data in educational management. Journal of Teach. (09): 37-39.

[4] Hou, X. Yang, H. Mao, (2014). L. Research on the Influencing factors and Countermeasures of student source quality in College enrollment [J]. Science and Education Guide, (01): 27-28.

[5] Liu, L. Lin, H. LI, H. Zeng, H. (2019). Student Quality evaluation of Water Supply and Drainage Science and Engineering major from 2013 to 2017 under the background of Engineering Education Certification -- A Case study of Guilin University of Technology Education Modernization, (56): 237-242.

[6] Li, n, X. Zhang, L. The Influence of "985 Project" on the quality of College students Peking University Education Review, 2014, 12(04): 157-171+187-188.

[7] Pei, X. Wang, K. Zhang, X. (2017). China Metallurgical Education, (06): 104-105+110.

[8] Shao, F. (2019). Factors Influencing the quality of college students and countermeasures. Beijing Education, (12): 65-67.

[9] Tobler, W R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography, 46: 234-240.

[10] Xiao, Y. Chen, S. (2019). Analysis of Main Factors influencing university Comprehensive Ranking and Its Econometric Model -- Taking University of Electronic Science and Technology of China as an example Value engineering. 38(14): 125-128.

[11] Xiong, J. (2008). Quality Evaluation Model and Empirical Research of College Students. Central South University.

[12] Yan, W. Han, Y. (2020). International Frontiers and Methods of Education Policy Evaluation: An exclusive interview with Professor Yan Wen fan. Journal of Soochow University (Education Science), 8(03): 76-85.

[13] Wen, W. Lian, Z. (2018). Disadvantaged Students under the Preferential Policy of Enrollment: Access Opportunities and Educational Equity: An Empirical Study Based on the enrollment data of a key University [J]. Education Research, Tsinghua University, 39(2): 116-124

[14] Zhang, A. Fan, C. (2019). Spatial autocorrelation and overall promotion of regional integration development in Yangtze River Delta. Modern Economic Discussion, (08): 15-24.

[15] Zhang, M. Bi, G. (2020). Spatial difference and pattern evolution of urban land economic density in The Yangtze River Delta. Chinese Journal of Management, 33(04): 17-26.

[16] Zhang, W. Tang, Y. Guan, X. Tian. (2016). Research on Optimization Scheme of College Enrollment Plan Based on Grey Prediction Model and Cluster Analysis. Journal of North Western Polytechnical University (Social Science Edition), 36(01): 96-100+112.

[17] Zheng, Y. Shi, G. (2016). Journal of Hefei University of Technology (Natural Science), 39(11): 1571-1575.