

# Higher Education Evaluation System Based on PLS\_SEM Model, Fuzzy Comprehensive Evaluation Model

Huizhi Zhang<sup>1,\*</sup>, Xin Liu<sup>1</sup>, Ran Xie<sup>1</sup>

<sup>1</sup>Anhui University of Finance and Economics, Bengbu, 233030, China

\*1279027093@qq.com

## Abstract

The thesis first extracted 14 data indicators from three aspects of education input, education output and education process, and established a Multi-Dimensional higher education evaluation system. According to the PLS\_SEM model and using Smart PLS software to normalize the data, the weights of 14 indicators are obtained. Then, a fuzzy comprehensive evaluation model was established to obtain a comprehensive score for the development of higher education, and the development status was divided into 3 categories according to the score level. In addition, we chose China for in-depth analysis. Finally, we found that the application of this model is very common, but in fact it is very difficult to reform.

## Keywords

PLS\_SEM model; A fuzzy comprehensive evaluation model; Higher education system.

## 1. Introduction

With the development and progress of science and technology in the world and the continuous improvement of human culture, people gradually realize that the higher education system plays a vital role in the development of the national economy and even the development of the whole society. With the advent of the era of universal higher education, strengthening the quality assurance of higher education has become the core of the development policy of higher education in all countries (regions). The establishment of a universal teaching quality model as an important part of higher education quality assurance is the focus of global higher education reform [1]. Western developed countries, such as the United States, are leading the world in higher education. There are abundant researches on the evaluation of the development quality of higher education. Through continuous exploration, higher education quality evaluation standards and testing systems have been established one after another. In the United States, the "National Student Participation Survey" method of education quality evaluation is proposed to evaluate the teaching quality standards of universities [2].

The criterion is students' participation in learning. Britain in order to promote college students' learning quality, build a university undergraduate teaching quality evaluation system "distinguished teaching framework" (TEF). The system of the undergraduate teaching quality in colleges and universities can be divided into three levels [3]. On this basis, they build the quality of teaching and learning environment and study the results of the three dimensions of performance evaluation system for the undergraduate teaching quality and pay attention to students' learning results and quality combining [4-8]. In Australia, the establishment of the Agency for Quality and Standards in Education is mainly aimed at ensuring that students receive high-quality education [4]. OECD Indicators [9] is A relatively representative international evaluation of higher Education. This index focuses on the analysis of the relationship between education and social economy, emphasizing the quality of education. "World Education Report" [10-12] published by UNESCO is a relatively authoritative and

universal Education indicator system, including the proportion of Education expenditure in GDP, student-teacher ratio, gross enrollment rate, literacy rate and other indicators.

## 2. Basic Assumptions

In order to simplify the problem and solve the problem better, we put forward reasonable assumptions. Each of our assumptions is true and consistent with the basic facts: (i) assume that only to study the current situation of higher education, excluding population factors, primary education and other effects on higher education; (ii) assume that the obtained data can outline the overall situation of the country and provide scientific and accurate data; (iii) assume that the model does not consider the influence of personal factors such as educational preferences on higher education; (iv) assume that the model does not consider the impact of wars and plagues on higher education; (v) assume that the future economy of the country where the model is applied will grow steadily in order to ensure uninterrupted implementation of reform policies;

## 3. How to Evaluate the Development of a Country's Higher Education System and Quantify Its Comprehensive Results.

### 3.1. Analysis Approach

According to the research of Johnstone(1981) and J. H. Sun (2010) , we divide the higher education system into three areas, namely, educational input, process and output, in order to establish a model to study the current development status of the international higher education system [13-14]. We select 14 indicators and divided them into four first-level indicators. Specific indicators are classified as follows:

**Table 1.** Health degree index system of higher education system.

Field	First level indicators	Secondary indicators	Symbolic representation
Education input	Educational resources	Teacher student ratio in Higher Education(%)	$a_1, a_2, a_3$
		The proportion of government expenditure on Higher Education in GDP(%)	
Education process	Domestic education	The proportion of employee compensation in total expenditure of public higher education institutions(%)	
		Proportion of college students in Higher Education(%)	
	Cross-border education	Proportion of undergraduate students in Higher Education(%)	$b_1, b_2, b_3, b_4$
		Proportion of master students in Higher Education (%)	
Education output	Educational output	Proportion of doctoral students in Higher Education(%)	$c_1, c_2, c_3, c_4$
		Net current ratio of international students(%)	
		Ratio of international students(%)	
Education output	Educational output	Ratio of overseas students(%)	$d_1, d_2, d_3$
		Foreign enrollment rate by Region(%)	
		Expected length of Higher Education (Years)	
Education output	Educational output	Graduation rate of the first stage of Higher Education (Bachelor and master) (%)	$d_1, d_2, d_3$
		Proportion of people aged 25 and above who have received at least undergraduate education (%)	

Data source: EPS Global Statistics.

### 3.2. Determination of Index Factors for PLS\_SEM Models

#### 3.2.1. Model Introduction and Establishment.

PLS\_SEM method is a new method proposed by Wold(1982) on the basis of PLS regression, which can be used for multi-index aggregation problems . This paper uses PLS\_SEM model to measure the health degree of higher education system. PLS\_SEM model is composed of measurement model and structural model[15-16]. The measurement model is usually called the external model and the structural model is called the internal model. The specific model is as follows:

External model: assume that the explicit variables of each group are:  $X_j = (x_{j1}, x_{j2}, \dots, x_{jp_j})$ ,  $j = 1, 2, \dots, J$ . Suppose that each group of variables corresponds to a latent variable  $\xi_j$ , ( $j = 1, 2, \dots, J$ ), where the mean value of  $\xi_j$  is 0 and the variance is 1. The external model, that is, the relationship between  $X_j$  and  $\xi_j$ , is shown by formula (1):

$$x_{jh} = \lambda_{jh}\xi_j + \varepsilon_{jh}, \tag{1}$$

Where,  $\varepsilon_{jh}$  is the random error term, the mean value is 0. It is not related to  $\xi_j$  and the above formula should satisfy the condition of formula (2):

$$E(x_{jh} | \xi_j) = \lambda_{jh}\xi_j. \tag{2}$$

There is a linear combination relationship between  $\xi_j$  and  $X_j$ :

$$\xi_j = \sum_{h=1}^{p_j} w x_{jh} + \delta_j, \tag{3}$$

Where  $\delta_j$  is also a random error term, which needs to meet the requirements of the above formula. But,

$$E(\xi_j | x_1, x_2, \dots, x_j) = \sum_{h=1}^{p_j} w x_{jh}. \tag{4}$$

It shows that the mean value of residual  $\delta_j$  is 0, and it is not related to the explicit variable  $x_{jh}$ . Internal model: the internal model represents the causal relationship between latent variables, as shown in formula (5):

$$\xi_j = \sum_{i \neq j} \beta_{ji}\xi_i + \zeta_j \tag{5}$$

Among them,  $\zeta_j$  is a random error term and the mean value of residual is 0.  $\zeta_j$  is not related to  $\xi_j$ . Formula (5) shows that there is a causal relationship between latent variables.

### 3.2.2. Model Estimation

PLS\_SEM model estimates latent variable by iteration, and two methods can be used to estimate latent variable  $X_j$  corresponding to explicit variable group  $\xi_j$ : one is external estimation. The relationship between explicit variables and latent variables is studied and the latent variables are calculated. The second is internal estimation, which calculates the relationship between latent variables.

External estimation: assuming that the estimator is  $Y_j$ . The latent variable  $\xi_j$  can be estimated by the linear combination of the explicit variable  $x_{jh} (j=1,2,B, p_j)$ . In the model setting, it is assumed that the latent variables are standardized:

$$Y_j = \left( \sum_{h=1}^{p_j} \omega_{jh} x_{jh} \right)^* = (X_j w_j)^* \quad (6)$$

Where  $w_j$  is the weight vector and \* represents the standardization of the estimator.

Internal estimation: according to the structural model, latent variables can also be estimated by other latent variables associated with them. Assuming that the internal estimator obtained is  $Z_j$ , the following formula can be obtained:

$$Z_j = \left( \sum_{i:\beta_{ji}} e_{ji} Y_i \right)^* \quad (7)$$

Among them,  $\beta_{ji}$  is the coefficient in equation (5) and  $e_{ji}$  is the internal weight. The calculation method of  $e_{ji}$  is as follows:

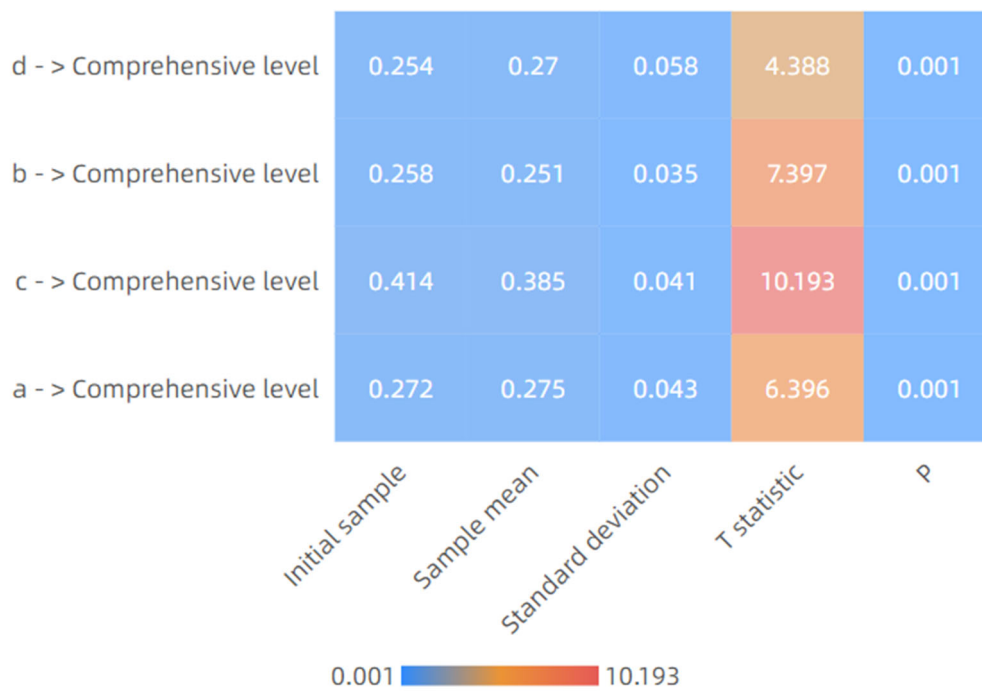
$$e_{ji} = \text{sign}(r(Y_j, Y_i)) = \begin{cases} 1, & r(Y_j, Y_i) > 0, \\ 0, & r(Y_j, Y_i) = 0, \\ -1, & r(Y_j, Y_i) < 0, \end{cases} \quad (8)$$

$r(Y_j, Y_i)$  is the correlation coefficient between  $Y_j$  and  $Y_i$ .  $\text{sign}$  is a sign function.

### 3.2.3. Model Results and Analysis

Using Smartpls software to build a model, assuming that education resources are represented by  $a$ , domestic education by  $b$ , overseas education by  $c$ , and education output by  $d$ .

We first normalize the data and then test the bootstrap method. The bootstrap method is used to calculate the bootstrap distribution and estimate and infer the population parameters.



**Figure 1.** Bootstrap method test results

From the above table, we can see that the data pass the bootstrap test. Then we test the latent variables. The test results are as follows:

**Table 2.** The latent variable Bootstrap test

	Standard deviation	T statistic	P
c -> a	0.116	4.735	0.001
b -> a	0.115	5.566	0.001
b -> c	0.102	6.581	0.001
d -> a	0.066	11.939	0.001
d -> c	0.055	16.090	0.001
d -> b	0.063	13.341	0.001
Comprehensive level -> a	0.130	3.808	0.001
Comprehensive level -> c	0.108	5.690	0.001
Comprehensive level -> b	0.118	4.445	0.001
Comprehensive level -> d	0.060	12.908	0.001

As can be seen from the above table, the data passed the test, so we can establish PLS\_SEM model. The 14 index data that passed the test were substituted into the model and the PLS program was used for iteration. The results are as follows:

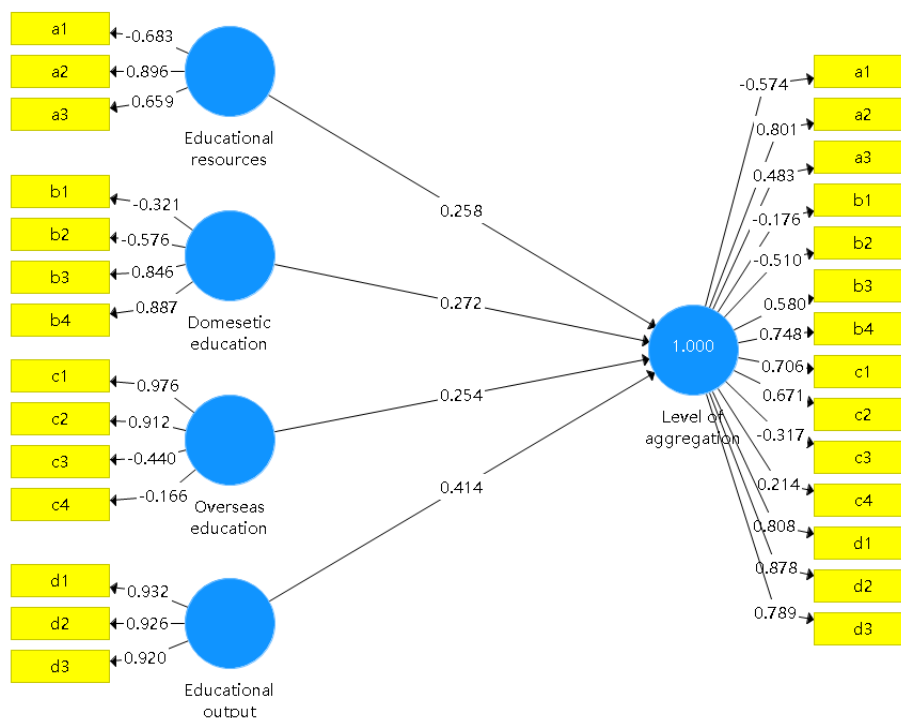


Figure 2. Higher education system comprehensive level evaluation model

As can be seen from the above figure, in the index of education resources, the proportion of government higher education expenditure in GDP has the greatest impact on education resources, followed by the proportion of higher education teacher/student ratio and the proportion of staff salary in the total expenditure of public higher education institutions. In domestic education, master's and doctoral education of higher education plays a decisive role in the quality of higher education, with path coefficients of 0.846 and 0.887 respectively. Therefore, increasing investment in master's and doctoral education can effectively promote the development of higher education. In overseas education, the net flow rate of international students and the rate of inbound students occupy an important position and their path coefficients are 0.976 and 0.912. The high rate of international student turnover and the rate of inbound students indicates that the country's domestic higher education is in good condition, which is an important feature of a country's higher education system. Finally, in the index of educational output, the expected years of higher education, the graduation rate of the first stage of higher education and the proportion of the population aged 25 and above who have received at least a bachelor's degree are all important criteria to judge the quality of local education, which reflects the health degree of the higher education system from the side. The higher the number of years of education, the higher the graduation rate and the higher the proportion of higher education in the population, which indicates that the quality of education is good and the higher education system is relatively healthy.

By using Smartpls software, we get the path coefficient of each index and the structural equation model as follows:

$$S = 0.258a + 0.272b + 0.254c + 0.414d . \tag{9}$$

It can be seen that the main factors influencing the higher education system are educational output, followed by domestic education, educational resources and overseas education. An important measure of a good, healthy higher education system is the output of education. All indicators will influence and promote each other. [17] Therefore, to establish a healthy higher

education system, we should not only pay attention to the quality of education, but also give systematic support in various aspects, constantly increase education investment and improve the level of domestic higher education [18].

### 3.3. Fuzzy Comprehensive Evaluation Model

#### 3.3.1. Model Introduction and Establishment

The fuzzy comprehensive evaluation method can evaluate the problems that are difficult to quantify and fuzzy[19-20]. In order to quantify the data of each country and establish the scoring system of higher education system, we establish a fuzzy comprehensive evaluation model. First of all, we must unify the attribute of evaluation index. In this paper, in addition to the ratio of overseas students and the ratio of foreign enrollment in various regions are cost-effective indicators, other indicators are benefit indicators. Then we establish the relative superior membership matrix:

$$R = \begin{pmatrix} r_{11} & r_{12} & B & r_{1n} \\ r_{21} & r_{22} & B & r_{2n} \\ C & C & E & C \\ r_{m1} & r_{m2} & B & r_{mn} \end{pmatrix}, \quad (10)$$

Among them:

$$u_{ij} = \begin{cases} a_{ij} / \max \{a_{ij}\} & a_{ij} \text{ is the benefit index,} \\ \min \{a_{ij}\} / a_{ij} & a_{ij} \text{ is the cost index,} \end{cases} \quad i \leq j \leq n, i = 1, 2, B, m$$

Comprehensive evaluation method is based on a number of indicators to evaluate. If the weight of each evaluated object in an index is large, the value of the index can clearly distinguish each evaluated object. On the contrary, if the value difference of each evaluated object on an index is small, and the ability of the index to distinguish each evaluated object is weak, the weight of the index is small. The coefficient of variation formula of each index is as follows:

$$v_i = \frac{s_i}{\bar{x}_i}, \quad (11)$$

Among them,  $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n a_{ij}$  is the average value of index  $i$ .  $s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (a_{ij} - \bar{x})^2$  is the variance of index  $i$ . By normalizing  $v_i$ , we can get the weight of each index as follows:

$$w_i = \frac{v_i}{\sum_{i=1}^m v_i}, \quad (12)$$

According to the relative superior membership matrix and the weight of coefficient of variation method, we can get the comprehensive evaluation model as follows:

$$F_j = \sum_{i=1}^{10} w_i u_{ij} (j = 1, 2, B, 10). \quad (13)$$

### 3.3.2. Results and Analysis

According to the index classification of PLS\_SEM model, we use the same division method in establishing the fuzzy comprehensive evaluation model. Using MATLAB software to process the data and establish the model, the fuzzy comprehensive evaluation results are obtained and the specific weights are as follows:

**Table 3.** Fuzzy comprehensive evaluation of each indicator weight table.

First level indicators	Weight	Secondary indicators	Weight
Educational resources	0.3585	Teacher student ratio in Higher Education(%)	0.2679
		The proportion of government expenditure on Higher Education in GDP(%)	0.2440
		The proportion of employee compensation in total expenditure of public higher education institutions(%)	0.4881
Domestic education	0.1112	Proportion of college students in Higher Education(%)	0.3469
		Proportion of undergraduate students in Higher Education(%)	0.3502
		Proportion of master students in Higher Education (%)	0.2673
		Proportion of doctoral students in Higher Education(%)	0.0356
Cross-border education	0.2215	Net current ratio of international students(%)	0.1448
		Ratio of international students(%)	0.1383
		Ratio of overseas students(%)	0.4216
		Foreign enrollment rate by Region(%)	0.2953
Educational output	0.3088	Expected length of Higher Education (Years)	0.2893
		Graduation rate of the first stage of Higher Education (Bachelor and master) (%)	0.3559
		Proportion of people aged 25 and above who have received at least undergraduate education (%)	0.3548

From the weight of each index in the above table, we can get the secondary index formula as follows:

$$\begin{aligned}
 f_1 &= 0.2679a_1 + 0.244a_2 + 0.4881a_3 \\
 f_2 &= 0.3469b_1 + 0.3502b_2 + 0.2673b_3 + 0.0356b_4 \\
 f_3 &= 0.1448c_1 + 0.1383c_2 + 0.4216c_3 + 0.2953c_4 \\
 f_4 &= 0.2893d_1 + 0.3559d_2 + 0.3548d_3
 \end{aligned} \quad (14)$$

Then, the first level index formula can be obtained as follows:

$$F = 0.3585f_1 + 0.1112f_2 + 0.2215f_3 + 0.3088f_4. \quad (15)$$

Based on the above model, we established the health evaluation model of higher education system, which can be used to evaluate the health status of higher education system in various countries, so as to put forward suggestions for the construction and development of national higher education system in the future.



### 3.4. Developments in National Education Systems

Since the comprehensive evaluation model is a benefit model, the higher the comprehensive evaluation score is, the better the health status of the higher education system of the country is. Combining 25 countries' comprehensive evaluation scores, we have sequenced the different health status ranges. Among them, a comprehensive score above 26.3 is "excellent", a subdivision score between 23.2 and 26.3 is "good", and a score below 23.2 (that is, below the world average) is "poor".and the results were as follows:

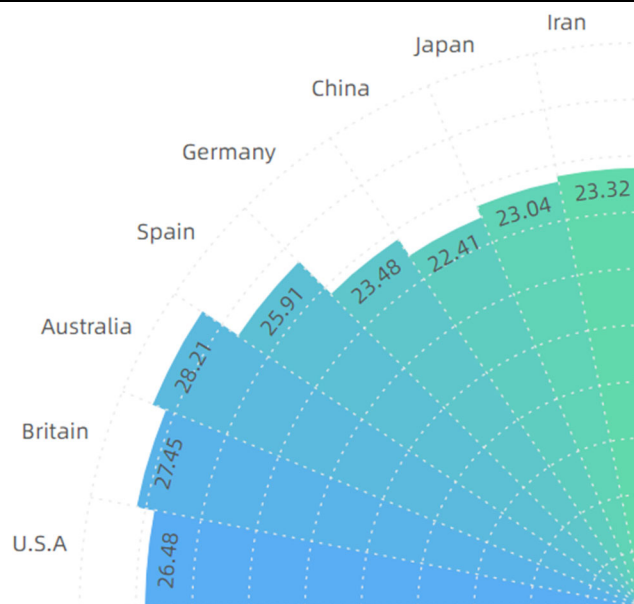
**Table 4.** Rating level

Comprehensive score	<23.2	[23.2,26.3]	26.3<
Higher education grades	Poor	Good	Excellent

Representative countries such as the United States, the United Kingdom, Australia, Spain, Germany, China, Japan and Iran were selected to calculate their scores using the fuzzy comprehensive evaluation model, and the results were as follows:

**Table 5.** Comprehensive evaluation scores for countries

Country	U.S.A	Britain	Australia	Spain	Germany	China	Japan	Iran
Score	26.48	27.45	28.21	25.91	23.48	22.41	23.04	23.32
Health condition	Excellent	Excellent	Excellent	Good	Good	Poor	Poor	Good



**Figure 3.** Comprehensive evaluation scores for countries

As can be seen from the above table, the overall rating of each country is ranked in descending order: Australia, the United Kingdom, the United States, Spain, Germany, Iran, Japan and China. Australia and the UK have relatively perfect higher education system construction and excellent health status, while China and Japan have relatively poor higher education system, which has a large room for improvement. Therefore, we choose China as the research object, combine the model, according to China's national conditions and the situation of higher education, give suggestions for the improvement of China's higher education system and put forward feasible and reasonable vision.

## 4. Conclusions

The path coefficients between the 14 indicators selected and the comprehensive level of higher education system were obtained through PLS\_SEM model and it was found that educational output had the most important impact on the health and sustainable development of higher education system. The Bootstrap test shows that all the 14 indicators of the selection have passed the test, that is, the 14 indicators have a significant impact on the health and sustainable development of the higher education system.

Then the fuzzy comprehensive evaluation model is established. The weight of each index is determined by the coefficient of variation method and the fuzzy comprehensive evaluation model is obtained to further calculate the comprehensive evaluation score of each country[21]. The results show that education resources and educational output have a great impact on the health of higher education system. We then selected representative countries such as the United States, the United Kingdom, Australia, Spain, Germany, China, Japan and Iran to rank their comprehensive evaluation scores. The results show that Australia, the United Kingdom and the United States rank the first three, and China ranks the last, with great room for development [22]. In the future, China should increase the investment of higher education funds, promote the improvement of higher education quality, train the new generation of talents and form a healthy and sustainable higher education system.

## 5. Data Availability

The data in this paper come from Question F of the 2021 American College Students Mathematical Modeling Competition and the statistics of the World Bank.

## 6. Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

- [1] J. H. Sun. Empirical Research on Evaluation of Regional Education Development in China . Nanjing University of Aeronautics and Astronautics, 2010. Anderson, L.W., Ryan, D. W.&Shapiro, B. J. The IEA Classroom Environment Study. Oxford: Pergamon, 1989:3~15.
- [2] Wold H. Soft modeling: The basic design and some extensions. Reskog K. G. J. Wold H. Systems under Indirect Observation. Amsterdam: North-Holland: 1982, 1-54.
- [3] R. Tippelt, B. Schmidt and A. V. Hippel, "Higher Education Evaluation in Germany", Research in Comparative and International Education, vol.5, no.1, pp.98, 2010.
- [4] Wold H. Partial Least squares. Kots S, Johnson N. L. Encyclopedia of Statistical Science [M]. New York: John Wiley&Sons, 1985(6): 581-591.
- [5] H. W. Wang, Linghui Fu. Application of PLS Path Model in Establishing Comprehensive Evaluation Index. Systems Engineering Theory & Practice, 2004(10):80-85.
- [6] Y. Liang, Fitting Index of Partial Least Squares Method and Its Application in Satisfaction Research. Mathematical Statistics and Management, 2005(2):40-44.
- [7] H. R. Kells, "Higher education evaluation systems for Latin America: An analysis of recent experiences and the formulation of a generalized model", Higher Education Policy, vol.9, no.3, pp.239-253, 1996.
- [8] B. Quinn, "Reformation or Transformation? Policy Reform in Ireland's Higher Education System", Higher Education Policy, vol.33, no.6, pp. 159-177, 2020.

- [9] B. R. Xu, "Research on the Evaluation System of Higher Education Service Quality——Based on the SERVPERF Model", *Journal of Nanjing Institute of Technology (Social Science Edition)*, vol.17, no.2, pp. 65-72, 2017.
- [10] H. Liu, Y. Yu, W. S. Jiang, "Blockchain technology, differences in educational resources and high quality economic development: An Empirical Analysis Based on the allocation of higher education resources in China", *Journal of Shaanxi Normal University (Philosophy and Social Sciences Edition)*, vol.49, no.1, pp.145-158, 2020.
- [11] Z. Q. Wang, Y. Wang, X.Y. Sun, "Study on the Measurement of the Competitiveness of Higher Education and Related Factors in Countries Along the Belt and Road", *Journal of Southwest University (Social Science Edition)*, vol.47, no.1, pp.112-123+227,2021.
- [12] Y. Liu, N. Li, "Research and Judgement on the Evaluation Index System of Higher Education Internationalization\_Based on the Comparison of 9 Evaluation Index Systems", *Heilongjiang Higher Education Research*, vol.38, no.8, pp.77-83, 2020.
- [13] L. Ge, Z. Y. Liu, "Research on the evaluation index system of University Entrepreneurship Education Ability Based on CIPP", *Journal of Northeast University (Social Science Edition)*, vol.16, no.4, pp.377-382, 2014.
- [14] H. T. Zhang, X. R. Li, T. Kong, "The application of BP neural network in missing data estimation", *Computer Engineering and Design*, no.14, pp.3457-3459, 2007.
- [15] S. L. Yin, J. Liu, L. Teng, "A Sequential Cipher Algorithm Based on Feedback Discrete Hopfield Neural Network and Logistic Chaotic Sequence", *International Journal of Network Security*, vol.22, no.5, pp. 869-873, 2020.
- [16] H. Xu, F. Ding, M. Gan, et al, "Two-stage recursive identification algorithms for a class of nonlinear time series models with colored noise", *International Journal of Robust and Nonlinear Control*, vol.30, no.17, pp. 7766-7782,2020.
- [17] Z.W. Liu, F.K. Zhu, "A New Extension of Thinning-Based Integer-Valued Autoregressive Models for Count Data", *Entropy*, vol.22, no.1, pp.62, 2020.
- [18] C. X. Xie, F. X Wang, "GDP forecast based on ARIMA-DGM-BP combined model in Kashgar area", *Mathematical Practice and Knowledge*, vol.50, no.15, pp.43-48, 2020.
- [19] M. Q. Zheng, J. M Zhu, "Forecast and analysis of Zhejiang Province GDP based on ARIMA model", *Natural Science Journal of Harbin Normal University*, vol.36, no.3, pp.56-61, 2020.
- [20] J. H. Sun. PLS Path Model Analysis of Statistical Characteristics of Higher Education Development in China. *Mathematical Statistics and Management*, 2010, 29(2): 362-371.
- [21] W. S. Wu, M. Fan, "Research on Regional Innovation Capability Evaluation and Spatial Distribution of Each City in Anhui Province Based on Entropy TOPSIS Method", *Journal of Changchun University of Science and Technology (Social Science Edition)*, vol.32, no.2, pp.82-87, 2019.
- [22] H. T. Zhang, X. R. Li, T. Kong, "The application of BP neural network in missing data estimation", *Computer Engineering and Design*, no.14, pp.3457-3459, 2007.